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

Mark J. Hurlstone

University of Western Australia

Graham J. Hitch

University of York

Word count: 19,092 (approximate count due to use of LaTeX)

Author Note

Mark Hurlstone: School of Psychology, University of Western Australia and Graham Hitch: Department of Psychology, University of York. This paper is based on part of the first author's doctoral dissertation completed at the University of York, England, which was supported by a research studentship from the Economic and Social Research Council of the United Kingdom. The first author is now based at the University of Western Australia. Correspondence concerning this article should be addressed to Mark Hurlstone, School of Psychology, University of Western Australia, Crawley, W.A. 6009, Australia. Email: mark.hurlstone@uwa.edu.au. URL: http://mark-hurlstone.github.io

Abstract

A central goal of research on short-term memory (STM) over the past two decades has been to identify the mechanisms that underpin the representation of serial order, and to establish whether these mechanisms are the same across different modalities and domains (e.g., verbal, visual, spatial). A fruitful approach to addressing this question has involved comparing the transposition error latency predictions of models built from different candidate mechanisms for representing serial order. Experiments involving the output-timed serial recall of sequences of verbal (Farrell & Lewandowsky, 2004) and spatial (Hurlstone & Hitch, 2015) items have revealed an error latency profile consistent with the prediction of a competitive queuing mechanism within which serial order is represented via a primacy gradient of activations over items, associations between items and position markers, with suppression of items following recall. In this paper, we extend this chronometric analysis of recall errors to the serial recall of sequences of visual, non-spatial, items and find across three experiments an error latency profile broadly consistent with the prediction of the same representational mechanism. The findings suggest that common mechanisms and principles contribute to the representation of serial order across the verbal, visual, and spatial STM domains. The implications of these findings for theories of short-term and working memory are considered.

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

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

***

A critical feature of short-term memory (STM) is its capacity to encode temporal relations between events (Marshuetz, 2005)—it is often important to remember not only the specific events that we experienced, but also the order in which we experienced them. This is true of imitative behaviors (Agam, Bullock, & Sekuler, 2005; Agam, Galperin, Gold, & Sekuler, 2007) as well as linguistic behaviors, such as vocabulary acquisition (Baddeley, Gathercole, & Papagno, 1998; Page & Norris, 2009) where the ordering of sub-elements is important.

The study of how people recall serial order information from STM is one of the oldest topics in experimental psychology (cf. Ebbinghaus, 1886). A popular account of how people accomplish this serial recall task is based on a working memory model (Baddeley & Hitch, 1974) comprising a phonological loop—dedicated to the retention of verbal sequences—and a visuospatial sketchpad—dedicated to the retention of visuospatial sequences. The latter system is hypothesized to contain two separate components: a "visual cache"—dedicated to the retention of visual sequences—and an "inner scribe"—dedicated to the retention of spatial sequences (Logie, 1995). Although this model has been hugely influential and offers a qualitative account of the effects of a number of key variables on serial recall performance, a widely acknowledged criticism is that it fails to offer a mechanistic account of how people actually accomplish the serial recall task (e.g., Burgess & Hitch, 1992; Hurlstone & Hitch, 2015; Hurlstone, Hitch, & Baddeley, 2014).

This shortcoming highlights the need for more quantitative theoretical accounts of serial recall (Henson & Page, 1999; Page, 2005) and in recent years several computational theories of verbal STM have been advanced that specify explicit mechanisms for representing serial order (Brown, Neath, & Chater, 2007; Brown, Preece, & Hulme, 2000; Botvinick & Plaut, 2006; Burgess & Hitch, 1999; Farrell, 2012; Farrell & Lewandowsky, 2002; Grossberg & Pearson, 2008; Hartley, Hurlstone, & Hitch, 2016; Henson, 1998; Lewandowsky & Farrell, 2008; Page & Norris, 1998). Based on numerous recent diagnostic results (reviewed in Hurlstone et al., 2014; Lewandowsky & Farrell, 2008), several mechanisms and principles of serial order that feature in different models have been identified that must be instantiated in any adequate theoretical account of serial order

in verbal STM. We discuss these shortly, but first we note that in contrast to the theoretical developments in understanding verbal STM for serial order, progress in understanding spatial and visual STM for serial order has been much slower. This is because—perhaps out of experimental convenience (viz. it is harder to test memory for serial order with spatial and visual materials)—the lion's share of research has employed verbal materials (e.g., letters, digits, words) as stimuli. Nevertheless, as we shall see next, a burgeoning body of evidence indicates that the processing of serial order across different STM domains is functionally similar, suggesting that mechanisms and principles of serial order in verbal STM may extend to visual and spatial STM.

Evidence of Functional Similarities

A long line of studies have now revealed that spatial and visual STM exhibit various phenomena of serial order previously thought to be unique properties of verbal STM. One behavioral phenomenon that has received much scrutiny is the serial position curve, which plots recall accuracy or latency by the serial position of items. When plotting recall accuracy, the serial position curve exhibits a bowed form, such that error rates are smallest for the first several items in the sequence (the primacy effect) and the last few items (the recency effect). When plotting recall latency—viz. inter-response times—the serial position curve follows an inverted U shape trend, and additionally exhibits a long initial recall latency for the first item (Anderson, Bothell, Lebiere, & Matessa, 1998; Farrell & Lewandowsky, 2004; Maybery, Parmentier, & Jones, 2002; Parmentier & Maybery, 2008; Thomas, Milner, & Haberlandt, 2003). These characteristic features of serial position curves are not unique to verbal STM. Accuracy serial position curves exhibiting primacy and recency effects have been witnessed in studies of spatial STM in which participants recalled sequences of seen spatial locations (Avons, 2007; Farrand, Parmentier, & Jones, 2001; Guérard & Tremblay, 2008; Jones, Farrand, Stuart, & Morris, 1995; Tremblay, Guérard, Parmentier, Nicholls, & Jones, 2006), and in studies of visual STM in which participants recalled sequences of novel visual patterns (Avons, 1998; Avons & Mason, 1999) or unfamiliar faces (Smyth, Hay, Hitch, & Horton, 2005; Ward, Avons, & Melling, 2005) presented in the same spatial location. Similarly, latency serial position curves resembling those witnessed with verbal stimuli have been observed with spatial materials (Hurlstone & Hitch, 2015; Parmentier, Andrés,

Elford, & Jones, 2006; Parmentier, Elford, & Maybery, 2005).

Another behavioral phenomenon that has been the subject of a great deal of comparisons across domains is the vulnerability of serial recall to transposition errors. Such order errors occur when an item is recalled in the wrong serial position. Transpositions can be classified according to their displacement—the numerical difference between an item's presentation and recall positions. Anticipation errors are transpositions with negative displacement values and occur when an item is recalled before its correct position; conversely, postponement errors are transpositions with positive displacement values and occur when an item is recalled after its correct position. Transposition gradients plot the probability of transpositions according to their displacement value and exhibit three empirical regularities (Farrell & Lewandowsky, 2004; see Figure 3b): (a) the gradients peak at displacement 0 (most responses are correct); (b) the probability of a transposition decreases as the absolute displacement increases—the *locality constraint* (Henson, 1996; Henson, Norris, Page, & Baddeley, 1996); and (c) the error gradients for anticipations and postponements are approximately symmetrical. Like the serial position curves, these three hallmarks of transpositions are also not confined to verbal memoranda—transposition gradients exhibiting these functional characteristics have also been witnessed with spatial (Hurlstone & Hitch, 2015; Jalbert, Saint-Aubin, & Tremblay, 2008; Parmentier et al., 2006; Smyth & Scholey, 1996) and visual (Avons & Mason, 1999; Smyth et al., 2005) memoranda.

Functional similarities across domains are not limited to serial position curves and transposition gradients. Visual and spatial STM exhibit several additional phenomena of serial order in common with verbal STM, including similar distributions of item and order errors (Avons & Mason, 1999; Guérard & Tremblay, 2008) and similar effects of sequence length (Smyth et al., 2005; Smyth & Scholey, 1996), item similarity (Avons & Mason, 1999; Jalbert et al., 2008; Smyth et al., 2005), and Hebb repetition learning (Couture & Tremblay, 2006; Horton, Hay, & Smyth, 2008) amongst other kindred effects (see Hurlstone et al., 2014 for a review). In the next section, we delineate mechanisms and principles of serial order in computational theories of verbal STM that have been proposed to explain serial recall phenomena such as those just reviewed.

Seriating Mechanisms and Principles

Due to a certain amount of co-evolution in their development, there has been some theoretical convergence amongst computational theories of verbal STM and several mechanisms and principles for the representation and control of serial order have been proposed that are widely employed in different models. Existing models represent serial order either: (a) by using a competitive queuing sequence planning and control mechanism, (b) by imposing a primacy gradient of activations over items, (c) by forming associations between items and some representation of their list position—viz. position marking, (d) by incorporating response suppression, and (e) by implementing output interference, or through some union of these mechanisms and principles.

Competitive Queuing

Most models of serial recall employ a mechanism known as competitive queuing (Bullock, 2004; Bullock & Rhodes, 2003; Davelaar, 2007; Glaspool, 2005; Grossberg, 1978; Houghton, 1990) to plan, represent, and recall sequences. A schematic of such a mechanism—realized as a neural network model—can be inspected in Figure 1. The model comprises two layers of localist item nodes—a parallel planning layer and a competitive choice layer. The nodes in the planning layer represent the pool of items from which sequences are generated. 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 obey recurrent-competitive-field dynamics, meaning that each item node excites itself and sends lateral inhibition to competitor nodes in the same layer. This sets up a 'winner-takes-all' response competition over items, 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 planning layer. This process iterates until recall of the sequence is complete.

Primacy Gradient

The main difference between different competitive queuing models concerns the nature of the activation gradient used to represent serial order in the planning layer. In the most parsimonious models (Farrell & Lewandowsky, 2004; Grossberg, 1978; Page & Norris, 1998), a single monotonically decreasing activation gradient—known as a primacy gradient—is established over items during serial order encoding, such that the earlier an item occurred in a presentation sequence, the stronger the activation it is assigned. This gradient is then held static during sequence generation and serial recall is accomplished via an iterative process of selecting the strongest item before suppressing its activation (viz. response suppression; see below)—the suppression of an item after it has been retrieved removes it from the cohort of recall candidates at the subsequent position, allowing the next strongest item to win the output competition.

Position Marking

In more sophisticated competitive queuing models, the activation gradient established over items is not static, but instead varies dynamically over time via the output of a context signal—separate from the item representations in the planning layer—during the course of sequence generation (Brown et al., 2000; Burgess & Hitch, 1999; Hartley et al., 2016; Henson, 1998; Lewandowsky & Farrell, 2008). This introduces a positional component to the representation of serial order because the state of the context signal at any given moment confers information about the current position in the sequence. Accordingly, this dynamic process of representing serial order is known as position marking.

A specific example that serves to highlight this general approach is provided by the seriating mechanism embodied in the model of Burgess and Hitch (1999). In their model, when an item is presented as part of a to-be-remembered sequence its representation is activated in a planning layer and an association is formed—via Hebbian learning—between the item representation and the current state of a time-varying (distributed) positional context signal. The context signal exhibits the property of local-self similarity, meaning that neighboring states (viz. adjacent serial positions) are more similar to one another than states that are separated in time (viz. non-adjacent serial positions). Recall of the sequence is accomplished by reactivating the different

states of the positional context signal in order—which produces a dynamically varying activation gradient over items in the planning layer—and recalling the most activated item at each position.

Some models incorporate an activation gradient with both static (viz. a primacy gradient) and dynamic (viz. position marking) properties—generating a hybrid ordinal-positional representation of serial order. For example, in some models, a primacy gradient is incorporated into the strength of the associations between items and the different states of the positional context signal (Brown et al., 2000; Lewandowsky & Farrell, 2008). In other models, a primacy gradient is established over items but is then modulated by the output of the positional context signal during serial recall (Burgess & Hitch, 1999).

Response Suppression

Response suppression refers to the inhibition or removal of items from memory following recall and is an assumption incorporated in almost all theories of verbal STM (e.g., Brown et al., 2000; Burgess & Hitch, 1999; Farrell & Lewandowsky, 2002; Grossberg & Pearson, 2008; Henson, 1998; Lewandowsky & Farrell, 2008; Page & Norris, 1998). In competitive queuing models, response suppression occurs as a result of the inhibitory feedback signal from the competitive choice layer to the parallel planning layer following the retrieval of an item. In other models (Farrell, 2006; Lewandowsky & Farrell, 2008), response suppression is implemented through the unlearning—viz. Hebbian anti-learning (Anderson, 1995)—of the association between the item just retrieved and its position marker. This has the effect of reducing the strength with which the item competes for recall when memory is probed with subsequent position markers.

In models that rely on a primacy gradient to represent serial order, the incorporation of response suppression is crucial for sequencing, since it serves to prevent perseveration on the same response. It is a less crucial ingredient in models that rely on position marking to represent serial order because the dynamically re-evolving context signal relieves the suppression mechanism of the burden for sequencing. Nevertheless, even models that represent serial order via position marking must incorporate response suppression to minimize the occurrence of erroneous repetitions, which occur infrequently in serial recall (Henson, 1996; Vousden & Brown, 1998).

Output Interference

Output interference refers to the assumption that the act of recalling an item from STM interferes with the representation of items that are yet to be retrieved. It is 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; Lewandowsky, Duncan, & Brown, 2004; Lewandowsky & Farrell, 2008). The key feature of output interference is that recall of early items interferes with later items in the sequence. This interference occurs irrespective of whether serial order is represented via a primacy gradient or position marking (or both) and regardless of whether or not a recalled item is subsequently suppressed.

Model Selection

Functional similarities across domains suggest that at least some of the mechanisms and principles just reviewed, might also be implicated in the representation of serial order in visual and spatial STM. Indeed, in a recent comprehensive review of the serial recall literature, Hurlstone et al. (2014) identified evidence from behavioral, electrophysiological, and modeling studies that supports the contention that all short-term memories (verbal, visual, spatial) utilize the competitive queuing mechanism to plan, represent, and recall sequences. However, whilst Hurlstone et al. (2014) identified direct evidence for the operation of a primacy gradient, position marking, response suppression, and output interference in the verbal STM competitive queuing system—viz. the phonological loop—they noted that the principles that contribute to the representation of serial order in the visual and spatial STM competitive queuing systems—viz. the visuospatial sketchpad—are not yet known because the existing data in these domains—which has focused largely on serial position curves and transposition gradients—can be handled equally well by mechanisms embodying various different combinations of the representational principles.

A specific illustration of this problem is provided in Figure 2, which shows the theoretical predictions of five models of serial order originally studied by Farrell and Lewandowsky (Farrell & Lewandowsky, 2004; Lewandowsky & Farrell, 2008; see also Hurlstone & Hitch, 2015). The models were built from different combinations of a primacy gradient, position marking, response suppression, and output interference—representative of the combinations of these principles

employed in theories of serial recall; see Hurlstone et al. (2014) and Hurlstone & Hitch (2015)—and implemented within a common competitive queuing neural network architecture that permitted the generation of response probability and recall latency predictions (see Appendix A for precise details of the modeling). It is apparent from inspection of this figure that the five models generate qualitatively similar accuracy serial position curves (Figure 2A), transposition gradients (Figure 2B), and latency serial position curves (Figure 2C), which makes identification of the preferred mechanism difficult. These are not the only behavioral phenomena for which different models generate comparable predictions—most phenomena of serial order can be accommodated equally well by different mechanisms for representing serial order (Hurlstone et al., 2014; Lewandowsky & Farrell, 2008). Although several more diagnostic results have been identified and studied in the context of verbal STM that confer direct support for the operation of specific representational principles, with one noteworthy exception that we discuss next (Hurlstone & Hitch, 2015), these have yet to be examined in the visual and spatial domains.

One such phenomenon is known as the *latency–displacement function* (henceforth LDF; Farrell & Lewandowsky, 2004). The LDF is the latency equivalent of the transposition gradient and plots the mean recall latency of transpositions as a function of transposition displacement. Figure 3 shows the LDFs predicted by the five models of serial order. It is clear from the figure that unlike their predicted serial position curves and transposition gradients, the models' predicted LDFs differ considerably from each other. Specifically, when serial order is represented by position marking alone (PM), the LDF exhibits a symmetric V-shaped function, whereas the addition of response suppression (PM + RS) or output interference (PM + OI) reduces the slope for postponements, producing a partially asymmetric V-shaped LDF. In stark contrast to the above models, the combination of a primacy gradient with response suppression dramatically alters the shape of the LDF rendering it monotonically negative (PG + RS), whilst the addition of position marking (PG + PM + RS) flattens the slope of the function for postponements, but without removing the overall negative latency–displacement relationship.

Across three experiments involving the output-timed recall of verbal sequences, the LDFs observed by Farrell and Lewandowsky (2004) were consistently negative, and additionally exhibited a reduction in slope for postponements compared to anticipations, indicating that serial

order in verbal STM is represented via a mechanism combining a primacy gradient with position marking and response suppression. More recently, Hurlstone and Hitch (2015) reported three experiments exploring the dynamics of transpositions in a spatial serial recall task involving memory for sequences of seen spatial locations in which the observed LDFs were also consistently negative, with the functions being flatter for postponements than for anticipations, conferring support for the operation of the same representational mechanism in spatial STM.

Current Study

The current study sought to extend the analysis of transposition latencies to the recall of visual, non-spatial sequences in order to identify: (1) whether a combination of the four representational principles is responsible for coding serial order in visual STM, and (2) whether those principles are the same as those previously identified in verbal and spatial STM. The structure of the remainder of this article is as follows. First, we report three new experiments exploring the dynamics of transpositions in a visual serial recall task involving memory for sequences of unfamiliar faces. To foreshadow, across manipulations of sequence length (Experiments 1 & 2), articulatory suppression (Experiment 2), and temporal grouping (Experiment 3), the observed LDFs exhibited an overall negative trend, and in addition the slopes of the functions for postponements were relatively flat compared to those for anticipations. Next, we report quantitative fits of the models to representative data, which confirm that they are best accommodated by a model embodying a primacy gradient, position marking, and response suppression. Combined with the results of Farrell and Lewandowsky (2004) and Hurlstone and Hitch (2015), these findings suggest that the same mechanism is responsible for representing serial order across the verbal, visual, and spatial STM domains. The implications of these findings for theories of working memory are subsequently discussed.

Before reporting our experiments, we briefly motivate our choice of visual stimuli.

Choice of Stimuli

Previous studies of serial order in visual STM have employed either novel visual matrix patterns (Avons, 1998; Avons & Mason, 1999) or unfamiliar faces (Smyth et al., 2005; Ward et al.,

2005) as memoranda. A limitation of the former class of visual stimuli is that they are complex and artificial, requiring slow presentation rates, which may foster a reliance on supplementary verbal encoding strategies. Faces, by comparison, are also complex visual stimuli, but they benefit from a familiar form, which adults are extremely adept at processing. Smyth et al. (2005) have shown that serial memory phenomena—viz. serial position effects on accuracy; the sequence length effect; the locality constraint on transpositions—can be obtained with sequences of unfamiliar faces presented at fast presentation rates and are not based on verbal encoding strategies.

The latter conclusion is buttressed by the independent effects of visual similarity and articulatory suppression reported by Smyth et al. The visual similarity effect refers to the finding that sequences of visually similar items are recalled less accurately than sequences of visually dissimilar items (Avons & Mason, 1999; Logie Della Sala, Wynn, & Baddeley, 2000; Logie, Saito, Morita, Varma, & Norris, 2015; Saito, Logie, Morita, & Law, 2008), whereas the articulatory suppression effect refers to the finding that serial recall performance for verbal materials (and nonverbal materials that have been subject to verbal encoding) is depressed when participants must repeat a verbal token—or sequence of verbal tokens—out loud concurrent with the presentation of the study sequence (a secondary task which occupies the speech output system, thus blocking verbal encoding).

At first blush, the deleterious effect of articulatory suppression suggests that verbal coding strategies contribute to serial memory for faces. However, the authors failed to observe a reliable interaction between visual similarity and articulatory suppression. If verbal coding strategies do contribute to serial memory for faces then the visual similarity effect should be stronger in magnitude in the presence, than in the absence, of articulatory suppression due to the increased demands placed by suppression on visual encoding processes. That this was not the case suggests that suppression interfered instead either with attentional resources required during the encoding of the faces (cf. Meisser & Klauer, 1999) or with the representation of their serial order (cf. Henson, Hartley, Burgess, Hitch, & Flude, 2003).

Given their sensitivity to serial memory phenomena and resistance to verbal encoding strategies, we employed unfamiliar faces as visual stimuli in the three experiments that follow.

Experiment 1

Experiment 1 examined the LDFs underpinning visual STM for sequences containing different numbers of faces. The rationale for the sequence length manipulation was manifold. First, it permitted an analysis of potential performance related variability in the LDFs, since serial recall performance for sequences of visual items is known to deteriorate with sequence length (Smyth et al., 2005; Ward et al., 2005), as it does for sequences of verbal (Anderson et al., 1998; Crannell & Parish, 1957; Maybery et al., 2002) and spatial items (Smyth, 1996; Smyth & Scholey, 1994, 1996). Second, it permitted an assessment of the LDFs under conditions that should engender changes in response latencies, since chronometric studies of the sequence length effect in verbal serial recall have shown that recall times at each serial position increase approximately linearly with sequence length (Anderson et al., 1998; Maybery et al., 2002). Third, it enabled an examination of the sensitivity of the LDFs to changes in the range of possible displacements that transpositions could span, which naturally increases with sequence length.

Methods

Participants and materials. Twenty-six undergraduate students from the Department of Psychology at the University of York took part in the experiment in exchange for course credit or payment of £10 (approximately $15). All had normal or corrected-to-normal vision.

The stimuli were sequences of four to six faces of the same gender drawn randomly without replacement from a stimulus ensemble of 814 front profile images of unfamiliar faces, subject to the constraint that no face was presented in the study sequence on more than two occasions across the entire experiment. The faces were drawn from various public domain face databases, were edited to remove any background noise and maximize the size of a face, and presented in greyscale on a white background at a standard height of 1.5 inches. Approximately two-thirds of sequences contained only males faces, whilst the remaining third contained only female faces.

Design. The experiment manipulated two within-participant factors: sequence length with three levels (four vs. five vs. six) and serial position (with as many levels as items in the sequence). Participants undertook two approximately 70 minute experimental sessions, which were spaced at least 24 hours apart. Within each session, there were 150 experimental sequences,

which were divided into three blocks of 50 sequences, one block for each sequence length. The ordering of the blocks was counterbalanced across participants. Each block began with two practice sequences and there were enforced 1-minute rest periods after every twenty-five experimental sequences.

Procedure. Participants were tested individually in a quiet room in the presence of the experimenter. They initiated each trial by selecting a "begin trial" icon situated in the central screen position using a mouse-driven pointer. A central fixation cross then appeared for 1500 ms and was replaced by a sequence of faces presented singly for 500 ms each and separated by a 500-ms blank interval. The final face was followed by a 1000-ms delay after which the set of faces reappeared simultaneously at fixed positions within a circular array centered on the middle of the screen (see Figure 4). The allocation of faces to positions around the circular array was determined at random. Participants were required to click on the faces in their presentation order using the mouse-driven pointer. Once an item was selected, it disappeared temporarily for 50 ms to indicate that the response had been registered. Participants were encouraged to guess whenever they were unsure of the correct item for a given position, otherwise they could select a question mark located in the centre of the reconstruction array to omit that item. Once a response had been registered at each output position, the contents of the screen cleared and the reconstruction time for the sequence was displayed in the central screen position for 3000-ms, before the "begin trial" icon for the next trial was displayed.

Results and Discussion

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: (1) accuracy serial position curves, (2) transposition gradients, (3) latency serial position curves, and (4) LDFs. Effect size estimates are provided—for focused comparisons only—using Pearson's r.

Accuracy serial position curves. The accuracy serial position curves can be inspected in Figure 5A. The curves are representative of those witnessed in typical serial recall studies, showing a clear deterioration in performance with increasing sequence length and extended

primacy and restricted recency within sequences. Statistical confirmation of the effect of sequence length was obtained via a one-way Analysis of Variance (ANOVA) performed on the mean proportion of correct responses collapsed across serial positions for the different sequence lengths. This revealed a significant main effect of sequence length, $F(2,50) = 153.52$, $p < .001$, with four-item sequences being recalled better than five-item sequences, $t(25) = 9.13$, $p < .001$, $r = .88$, and with five-item sequences being recalled better in turn than six-item sequences, $t(25) = 9.97$, $p < .001$, $r = .89$.

Statistical confirmation of the effects of serial position was obtained by conducting one-way ANOVAs on the mean proportion of correct responses as a function of serial position at each sequence length. There was a significant main effect of serial position for sequences of four-items, $F(3,75) = 22.82$, $p < .001$, five-items, $F(4,100) = 23.13$, $p < .001$, and six-items, $F(5,125) = 41.13$, $p < .001$, reflecting the apparent primacy and recency effects in the data.

Transposition gradients. Figure 5B shows the transposition gradients, which exhibit the three hallmark characteristics delineated at the outset. Specifically, the gradients peak at displacement 0; the proportion of transpositions decreases as a function of increasing displacement; and the error gradients for anticipations and postponements are approximately symmetrical. Consistent with the accuracy serial position analysis, the frequency of anticipations and postponements increased with sequence length.

Latency serial position curves. The mean recall latencies for correct responses can be inspected in Figure 5C. The latency curves are similar to those witnessed for the output-timed recall of verbal (Farrell & Lewandowsky, 2004; Maybery et al., 2002) and spatial sequences (Hurlstone & Hitch, 2015; Parmentier et al., 2006), showing an elevation in recall times at each serial position with increasing sequence length along with a markedly longer recall latency for the first item than all subsequent items. However, in studies of verbal and spatial serial recall, once recall has been initiated the recall times over subsequent serial positions typically rise to a saddle point mid-sequence before accelerating thereafter, giving rise to an inverted-U shaped latency curve like those generated by the models in Figure 2C. In contrast to those data and model predictions, the latency curves associated with the recall of visual sequences in Figure 5C are monotonically decreasing showing a speed-up in recall times over serial positions.

The effect of sequence length was statistically verified by performing a one-way ANOVA with sequence length as the independent variable and mean recall latencies for correct responses collapsed across serial positions as the dependent variable. There was a significant main effect of sequence length, $F(2,50) = 18.33$, $p < .001$, with shorter recall times for four-item than five-item sequences, $t(25) = -4.47$, $p < .001$, $r = .67$, but the difference in recall times between five-item and six-item sequences fell marginally short of conventional significance levels, $t(25) = -1.78$, $p = .09$, $r = .34$. A comparison of performance within each serial position curve revealed a significant main effect of serial position for sequences of four-items, $F(3,75) = 155.04$, $p < .001$, five-items, $F(4,100) = 104.64$, $p < .001$, and six-items, $F(5,125) = 125.40$, $p < .001$, reflecting the long initial recall latency and speed-up in recall times over serial positions apparent within each serial position curve.

LDFs. Turning to the data most central to the present article, Figure 6A shows the LDFs which plot the mean recall latencies of transpositions as a function of transposition displacement. Note that the effect of output position on the LDFs—viz. the speed-up in recall times over output positions visible in Figure 5C—has been removed by subtracting from each individual recall latency the mean of all responses for that sequence length condition and serial position, for each individual participant (this procedure was also adopted in the generation of the model predictions—see Appendix A). Removal of the effect of output position is necessary because it is correlated with transposition displacement—anticipations predominantly occur at early output positions, whereas postponements predominantly occur at late output positions. This is problematic because recall times are slower at the start of recall, which artificially elevates the recall latencies of anticipations, whereas recall times are faster toward the end of recall, which artificially accelerates the recall latencies of postponements. The negative latencies at some transposition displacements are a consequence of this filtering process. It is apparent from inspection of Figure 6A that the LDFs for the different sequence lengths exhibit an overall negative trend, and additionally the functions are flatter for postponements than for anticipations. The only visible effect of the sequence length manipulation was that the slope of the LDF for anticipations was steeper for four-item sequences than five-item and six-item sequences.

The LDF for each sequence length was analyzed separately using a multilevel regression

model. Multilevel modeling is a way of analyzing hierarchical data in which some variables are nested within other variables. The hierarchy in the current data is that the effect of transposition displacement is nested within individual participants. The advantage of using multilevel modeling is that a regression model can be fit simultaneously to data from all participants, whilst at the same time allowing for individual differences in the model parameters by incorporating them as random effects.

The regression model examined here included fixed effects for transposition direction (anticipations vs. postponements), transposition displacement (with as many levels as the sequence length *sl* being modeled—viz. ranging from -*sl*+1 to 0 for anticipations and 0 to *sl*-1 for postponements), and the interaction between the two variables, with recall latency as the dependent measure of interest. Random effects were included for the intercept, transposition direction, and transposition displacement. To establish whether it was appropriate to model individual differences in the latter parameters, the fit of the model to each sequence length was compared with that of a fixed-effects-only model in which the intercept and slopes were not allowed to vary across individual participants. The model comparisons confirmed that the model incorporating random effects provided a better fit to the data for all sequence lengths: $\chi^2(10) = 99.81$, $p < .001$, for four-item sequences, $\chi^2(10) = 226.73$, $p < .001$, for five-item sequences, and $\chi^2(10) = 78.38$, $p < .001$, for six-item sequences.

The maximum likelihood estimates of the regression parameters—averaged across individual participants—for each analysis are summarized in Table 1. It can be seen by inspection that there were reliable effects of transposition direction and displacement for each sequence length. The negative slope estimates for these parameters indicate, respectively, that anticipations were slower than postponements, and the recall latency of transpositions became faster with increasing (more positive) transposition displacements. There was also a reliable interaction between transposition direction and displacement for all three sequence lengths, confirming that the effect of displacement was different for anticipations and postponements. Each interaction was scrutinized further by performing two additional multilevel regression analyses, one examining the effect of transposition displacement for anticipations and one examining the effect of transposition displacement for postponements. These models were specified the same as the main model, except

they excluded the effect of transposition direction and the interaction term. The analyses revealed that the effect of transposition displacement was significant for anticipations and postponements for all three sequence lengths. However, consistent with the LDFs illustrated graphically in Figure 6A, it can be seen in Table 1 that the regression slope estimates for anticipations were strongly negative, whereas the regression slope estimates for postponements were weakly positive.

As a first demonstration, the results of Experiment 1 show that across different sequence lengths, the relationship between recall latency and transposition displacement is negative but with a reduction in the slope of the LDF for postponements compared to anticipations. However, one potential limitation of Experiment 1 is that a manipulation check was not incorporated to determine if memory for sequences of faces was mediated by a verbal encoding strategy. Participants may, for example, have generated a verbal description of each face in a sequence and then rehearsed those descriptions as a sequence of verbal tokens. This means that we cannot be certain that the LDFs in Figure 6A are not underpinned in part by a verbal component. As noted previously, Smyth et al. (2005) have shown using an experimental protocol akin to our own that a verbal encoding strategy does not contribute to serial memory for faces. This renders it unlikely that such an auxiliary strategy was brought to bear on performance in the present experiment. Nevertheless, in order to examine the robustness of the results of Experiment 1 and to ensure that such a strategy did not impact upon the shape of the LDFs shown in Figure 6A, in the next experiment we sought to replicate the current results under conditions where we could be certain that a verbal recoding strategy could not be deployed by participants.

Experiment 2

The second experiment was identical to Experiment 1 in all respects except that participants were required to engage in articulatory suppression—viz. speak the digits “1”, “2”, “3”, “4” aloud repeatedly—during the encoding of study sequences in order to block the speech output system and prevent the deployment of a verbal recoding strategy. The question of interest is whether the LDFs observed in the previous experiment will hold when the opportunity to engage in verbal recoding is obstructed and performance must necessarily be based on a visual code only. Given the unrealistically large number of trials that would have been required to have

participants complete both a control and an articulatory suppression condition, only the latter condition was used and an estimate of the impact of the articulatory suppression manipulation on performance was obtained by means of a between-experiment comparison with Experiment 1.

Methods

Participants. Twenty-six undergraduate students from the Department of Psychology at the University of York took part in the experiment in exchange for course credits or an honorarium of £10 (approximately $15). None of the participants took part in the previous experiment and all had normal or corrected-to-normal vision.

Materials, design, and procedure. The materials, design, and procedure were the same as for Experiment 1, with one noteworthy exception: at the beginning of each trial, once the participant had selected the "begin trial" icon they were required to repeat the sequence of digits "1", "2", "3", "4" out loud, at the rate of three digits per second until the reconstruction array appeared. The articulation rate was demonstrated to the participant prior to the first practice sequence using a digital metronome. The experimenter remained present at all times to ensure compliance with the suppression protocols. If the participant failed to keep to the rate of three utterances per second, the rate was demonstrated again using the digital metronome and they were instructed to try harder.

Results and Discussion

Accuracy serial position curves. The accuracy serial position curves for this experiment are shown in Figure 7A alongside the corresponding curves for Experiment 1 to aid interpretability. Like the curves observed in the preceding experiment, there is a clear deterioration in performance with longer sequences and extended primacy and restricted recency effects within sequences. Also apparent is that the articulatory suppression manipulation reduced performance slightly for four-item sequences, but had no effect for five-item and six-item sequences.

Statistical confirmation of the effect of articulatory suppression on performance was obtained via a between experiment comparison with Experiment 1. A 2 (suppression:

no-suppression vs. suppression) × 3 (sequence length: four vs. five vs. six) ANOVA performed on the mean proportion of correct responses collapsed across serial position revealed no significant main effect of suppression, $F(1,50) = 0.18$, $p = .68$, $r = .03$, a significant main effect of sequence length, $F(2,100) = 271.02$, $p < .001$, with recall accuracy decreasing with increasing sequence length, and a significant interaction between the two variables, $F(2,100) = 5.13$, $p < .01$. The interaction materialized because for four-item sequences the small detrimental effect of suppression noted previously approached significance, $t(50) = 1.18$, $p = .24$, $r = .16$, whereas the effect of suppression fell considerably short of significance for five-item, $t(50) = .54$, $p = .59$, $r = .08$, and six-item sequences, $t(50) = -.50$, $p = .62$, $r = .07$.

To compare performance within each serial position curve for the present experiment only, separate one-way ANOVAs were conducted on the mean proportion of correct responses as a function of serial position for each sequence length. There was a significant main effect of serial position for sequences of four-items, $F(3,75) = 32.38$, $p < .001$, five-items, $F(4,100) = 32.04$, $p < .001$, and six-items, $F(5,125) = 60.91$, $p < .001$, reflecting the apparent primacy and recency effects in the data.

Transposition gradients. The transposition gradients are shown in Figure 7B. They are similar to those observed in the previous experiment (Figure 5B) and exhibit the expected hallmark characteristics.

Latency serial position curves. Mirroring the latency serial position curves observed in Experiment 1 (Figure 5C), the curves for the current experiment illustrated in Figure 7C exhibit a monotonically negative trend, with a high amplitude peak at the first serial position. A clear effect of sequence length is also visible, with the recall times at each serial position increasing with longer sequences. Statistical confirmation of the effect of sequence length was provided by performing a one-way ANOVA with sequence length as the independent variable and mean recall latencies for correct responses collapsed across serial positions as the dependent variable. There was a significant effect of sequence length, $F(2,50) = 47.64$, $p < .001$, with shorter recall times for four-item than five-item sequences, $t(25) = -8.64$, $p < .001$, $r = .87$, and with shorter recall times in turn for five-item than six-item sequences, $t(25) = -4.02$, $p < .001$, $r = .63$.

A comparison of performance within each serial position curve revealed a significant main

effect of serial position for sequences of four-items, $F(3,75) = 110.15$, $p < .001$, five-items, $F(4,100) = 154.14$, $p < .001$, and six-items, $F(5,125) = 117.54$, $p < .001$, reflecting the long initial recall latency and acceleration in recall times over serial positions apparent within the latency curves.

LDFs. The LDFs shown in Figure 6B parallel those reported in the previous experiment. They once again exhibit an overall negative latency–displacement relationship, with the functions being flatter for postponements than for anticipations. As before, the only visible effect of the sequence length manipulation is that it reduced the slope of the LDF for anticipations. As for Experiment 1, the LDFs were analyzed using a multilevel regression model, with fixed effects for transposition direction, transposition displacement, and the interaction between the two predictors; random effects were again included for the intercept and transposition direction and transposition displacement. Model comparisons confirmed that for each sequence length, a model including random effects provided a better fit to the data than a fixed-effects-only model: $\chi^2(10) = 161.52$, $p < .001$, for four-item sequences, $\chi^2(10) = 134.11$, $p < .001$, for five-item sequences, and $\chi^2(10) = 178.21$, $p < .001$, for six-item sequences.

The overall results of the multilevel regression analyses are shown in Table 2 and provide statistical confirmation of the pattern illustrated graphically in Figure 6B. There were once again reliable effects of transposition direction, with slower recall times for anticipations than postponements, and the effects of transposition displacement were reliable—or nearly so; viz. five-item sequences—with recall times becoming faster with increasing transposition displacements. The interaction between transposition direction and displacement was reliable for all three sequence lengths, indicating that the effect of transposition displacement was different for anticipations and postponements. As before, each interaction was broken down by performing separate multilevel regression analyses, one examining the effect of transposition displacement for anticipations and one examining the effect of transposition displacement for postponements (with the effect of transposition direction and the interaction term from the main model omitted). The effects of transposition displacement for anticipations and postponements were once again reliable—or nearly so; viz. the effect for anticipations for five-item sequences—for all three sequence lengths. However, as in Experiment 1, it can be seen from inspection of Table 2 that the

regression slope estimates were strongly negative for anticipations, whereas they were weakly positive for postponements.

In brief, the current experiment has shown that when the opportunity to verbally recode a sequence of unfamiliar faces is precluded—by blocking the speech output system—the observed LDFs are virtually indistinguishable from those witnessed in Experiment 1. Indeed, but for a small and unreliable negative effect on the accuracy of recall of four-item sequences, serial recall performance was unaffected by the articulatory suppression manipulation. This result suggests that verbal STM codes are unlikely to have contributed to the shape of the LDFs in Experiment 1. We note also that the small negative effect of articulatory suppression on performance for four-item sequences need not reflect the disruption of a verbal recoding strategy. As noted previously, the results of Smyth et al. (2005) suggest that such disruption is more likely to reflect either competition for attentional resources during serial order encoding or interference with the representation of serial order.

Experiment 3

To further examine the generality of the error latency profiles witnessed in Experiments 1 and 2, we next report a third experiment, which examined the impact of a temporal grouping manipulation. This manipulation involves inserting extended temporal pauses after every few items in a study sequence in order to segregate it into sub-groups. Grouping a sequence of verbal items in this manner has been shown to exert a number of systematic effects on performance. First, grouping improves serial recall performance and alters the shape of the accuracy serial position curve—the serial position curve exhibits mini primacy and recency effects within each group (Frankish, 1985, 1989; Hitch, Burgess, Towse, & Culpin, 1996). Second, grouping alters the shape of the latency serial position curve—as well as leaving a long pause before recalling the sequence, participants leave a long pause before recalling each group (Farrell & Lewandowsky, 2004; Maybery et al., 2002; Parmentier & Maybery, 2008). Third, grouping alters the pattern of transposition errors—grouping reduces the frequency of adjacent-neighbor transpositions that straddle a group boundary (e.g., items 3 and 4 exchanging positions in the sequence 123—456; Maybery et al., 2002; Ng & Maybery, 2005; Parmentier & Maybery, 2008), but increases the

frequency of transpositions between groups that preserve their within-group serial position (e.g., items 2 and 5 exchanging positions; Ng & Maybery, 2002, 2005; Ryan, 1969), a class of errors known as *interpositions* (Henson, 1996). With the exception of the increase in the frequency of interpositions, the effects of grouping with verbal stimuli just reviewed have also been documented with spatial materials (Hurlstone & Hitch, 2015; Parmentier et al., 2006).

It is widely accepted that temporal grouping effects are an empirical referent of the operation of position marking. Positional models account for such effects by assuming that order information in grouped sequences is represented using two sets of position markers, one set that encodes the position of groups (Brown et al., 2000; Hartley et al., 2016; Henson, 1998; Lewandowsky & Farrell, 2008) or items (Burgess & Hitch, 1999) in the sequence, and a second set that encodes the position of items within groups. The latter set of markers are crucial for explaining the pattern of interpositions errors observed in grouped verbal serial recall. By contrast, the absence of interpositions in grouped spatial serial recall has been taken to confer support for a subtly different representational scheme whereby position markers encoding the position of groups in the sequence are augmented by position markers encoding the position of items in the sequence as a whole, rather than within groups (Hurlstone & Hitch, 2015).

The rationale for incorporating the grouping manipulation was two-fold. First, grouping effects have not previously been examined for serial recall of visual materials, and to the extent that such effects are observed this will provide an additional test of the role of position marking in visual STM. Of particular interest is whether grouped visual serial recall is characterized by an increase in the probability of interpositions—like grouped verbal serial recall—or whether the probability of interpositions is unaffected by grouping—like grouped spatial serial recall. Assuming that grouping does exert effects on recall, the latter result might indicate that positional information in grouped visual sequences—like grouped spatial sequences—may be represented in a subtly different manner to that of grouped verbal sequences.

Second, given the multifarious effects of grouping on other aspects of recall documented earlier, it is reasonable to ask whether grouping might also exert systematic effects on the shape of the LDF. Indeed, if positional information in grouped sequences is represented via markers coding the position of groups in the sequence complemented by markers coding the position of

items within groups then the positional models predict that grouping should simultaneously increase the probability of interpositions and accelerate their response latencies. However, at variance with this prediction, Farrell and Lewandowsky (2004) found that although grouping a sequence of verbal items increased the probability of interpositions, this was not mirrored by faster recall times for these errors. Similarly, Hurlstone and Hitch (2015) also found that grouping did not exert any systematic effects on the LDF for spatial sequences (although they also failed to observe an effect of grouping on interposition rates). Thus, in both studies grouping exerted discernible effects on recall in terms of accuracy and latency serial position curves (and transposition gradients in the experiments of Farrell and Lewandowsky, 2004), but had only negligible—if any—effect on the LDFs, suggesting that representations other than position marking must underpin verbal and spatial serial recall. In the current experiment, we ask whether the LDF for visual serial recall is similarly unaffected by grouping.

Methods

Participants. Forty-two undergraduate students from the School of Psychology at the University of Western Australia took part in the experiment in exchange for course credits. All had normal or corrected-to-normal vision.

Design and procedure. The experiment manipulated two independent variables: grouping (ungrouped vs. grouped) was a between-participants factor, whereas serial position (1–6) was a within-participants factor. Half the participants received the ungrouped sequences, whereas the remaining half received the grouped sequences. Grouping was manipulated between-participants in order to reduce the likelihood of individuals spontaneously grouping the ostensibly ungrouped sequences, a tendency which can increase in within-participant designs in which some participants are exposed to objectively grouped sequences prior to ungrouped sequences (e.g., Henson, 1996; Farrell & Lelièvre, 2009; Parmentier et al., 2006).

The procedure was the same as for Experiment 1 with the following exceptions: participants completed only a single experimental session; the sequence length was fixed to six-items; and in the grouped condition the temporal interval separating the third and fourth item in the sequence was extended from 500 ms to 1500 ms to segregate the sequences into two

sub-groups of three items. Participants attempted 80 experimental trials, which were preceded by two practice trials. Enforced 30-second rest periods were included after every 20 experimental trials. The experiment lasted approximately 40 minutes.

Results and Discussion

Accuracy serial position curves. The accuracy serial position curves are shown in Figure 8A. Consistent with previous studies of grouped verbal (Frankish, 1985, 1989; Hitch et al., 1996; Maybery et al., 2002; Ryan, 1969) and spatial (Parmentier et al., 2006; Hurlstone & Hitch, 2015) serial recall, it can be seen that grouping modified the shape of the serial curve: The curve for grouped sequences exhibits effects of primacy and recency within groups and the sequence overall, whereas the curve for ungrouped sequences exhibits these effects for the sequence as a whole only. However, whilst grouping enhanced the accuracy of recall, its beneficial effect was highly localized, being restricted only to those items straddling the group boundary—viz. Serial Positions 3 and 4. This is at variance with the effects of grouping observed with verbal and spatial memoranda, which are larger in magnitude and witnessed across most—if not all—serial positions.

That grouping modified the shape of the accuracy serial position curve was statistically verified by a 2 (grouping: ungrouped vs. grouped) × 6 (serial position: 1–6) ANOVA. As suspected, there was no significant main effect of grouping, $F(1,40) = .23$, $p = .63$, $r = .04$, indicating that grouping did not foster an overall improvement in performance. However, the expected main effect of serial position, $F(5,200) = 61.03$, $p < .001$, and interaction between grouping and serial position, $F(5,200) = 5.03$, $p < .001$, were both significant.

Transposition gradients. In contrast to the effects of grouping on the accuracy serial position curve, Figure 8B shows that the transposition gradients were unaffected by the grouping manipulation. If grouping had fostered an increase in interposition errors then the transposition gradient for grouped sequences should exhibit peaks at ±3 displacements. However, the transposition gradients for grouped sequences—like those for ungrouped sequences—decrease monotonically with increasing transposition displacement, indicating that grouping did not foster an increase in these interposition errors. This result is at odds with the pattern observed in studies of grouped verbal serial recall (Ng & Maybery, 2002, 2005; Ryan, 1969), but it is

consistent with the pattern observed in studies of grouped spatial serial recall (Hurlstone & Hitch, 2015; Parmentier et al., 2006) where an increase in the frequency of interpositions in grouped sequences is also notably absent.

Although grouping did not increase the frequency of interpositions, it did reduce the frequency of adjacent-neighbor transpositions straddling the group boundary, consistent with studies of both grouped verbal (Maybery et al., 2002; Ng & Maybery, 2005; Parmentier & Maybery, 2008) and spatial serial recall (Hurlstone & Hitch, 2015). This was confirmed by calculating the frequency of adjacent inter-group errors and interposition errors for each individual participant from the ungrouped and grouped conditions. A 2 (grouping: ungrouped vs. grouped) × 2 (error-type: adjacent inter-group vs. interposition) ANOVA performed on the error frequencies revealed a significant main effect of grouping, $F(1,40) = 4.30$, $p < .05$, $r = .56$, with more errors for ungrouped than grouped sequences, a significant main effect of error-type, $F(1,40) = 99.33$, $p < .001$, $r = 1$, with more interposition than adjacent inter-group errors, and a significant interaction between the two variables, $F(1,40) = 6.49$, $p < .05$. The interaction arose because grouping decreased the frequency of adjacent inter-group errors (ungrouped $M = 23.33$ vs. grouped $M = 13.14$), $t(40) = 4.20$, $p < .001$, $r = .55$, but exerted no effect on the frequency of interposition errors (ungrouped $M = 35.81$ vs. grouped $M = 34.19$), $t(40) = .40$, $p = .69$, $r = .11$.

Latency serial position curves. The latency serial position curves are plotted in Figure 8C. In accordance with previous studies examining the effects of grouping on response timing in verbal (Farrell & Lewandowsky, 2004; Maybery et al., 2002) and spatial serial recall (Hurlstone & Hitch, 2015; Parmentier et al., 2006), it can be seen that grouping modified the shape of the latency curve. Whereas the ungrouped curve peaks at the first serial position and then decreases monotonically, the grouped curve exhibits a second lower amplitude peak at the fourth serial position, indicating that participants left a brief pause before outputting items from the second group.

The effect of grouping on the latency serial position curve was corroborated by a 2 (grouping) × 6 (serial position) ANOVA, which revealed no significant main effect of grouping, $F(1,40) = .24$, $p = .63$, $r = .04$, a significant main effect of serial position, $F(5,200) = 264.39$, $p < .001$, and the expected interaction between the two variables, $F(5,200) = 2.78$, $p < .05$.

LDFs. The LDFs for the present experiment are displayed graphically in Figure 6C and are in accordance with the now familiar pattern reported in the two previous experiments—once more the relationship between recall latency and displacement is a negative one, and in addition the slope of the functions for postponements continues to be flatter than for anticipations. That the overall shape of the LDF was unaffected by the grouping manipulation is noteworthy given that grouping engendered qualitative changes in the accuracy and latency serial position curves. However, the grouping manipulation did exert some subtle effects on the LDF; thus, although we have seen that grouping did not affect the probability of interpositions, Figure 6C shows that it did nevertheless affect their response latencies, although not quite in the manner predicted by positional models. Specifically, although grouping accelerated the recall times of -3 displacements, it decelerated—rather than accelerated—the recall times of +3 displacements. It is unclear whether these discontinuities in the LDF merely coincidentally occur at displacement distances corresponding to interpositions, or whether they reflect the impact of within-group positional codes. We explore this issue later using quantitative model comparisons.

The multilevel regression model used to analyze the LDFs once again incorporated fixed effects for transposition direction, transposition displacement (with six levels—viz. -5 to 0 for anticipations and 0 to +5 for postponements), and the interaction between the two predictors; random effects were included for the intercept and transposition direction and transposition displacement. Model comparisons once again confirmed that a model incorporating random effects provided a better fit to the data than a fixed-effects-only model: $\chi^2(10) = 135.48$, $p < .001$, for ungrouped sequences, and $\chi^2(10) = 103.29$, $p < .001$, for grouped sequences.

The maximum likelihood estimates of the regression parameters—averaged across individual participants—for each analysis are summarized in Table 3. There were once again reliable effects of transposition direction, with slower recall times for anticipations than for postponements; transposition displacement, with recall latencies becoming faster with increasing transposition displacement; and the interaction between the two predictors, indicating that the effect of transposition displacement differed for anticipations and postponements. The interactions were broken down once again by conducting separate multilevel regression analyses, one examining the effect of transposition displacement for anticipations and one examining the

effect of transposition displacement for postponements (with the effect of transposition direction and the interaction term from the main model excluded). The effect of transposition displacement was reliable for anticipations and postponements for both ungrouped and grouped sequences but as in the two preceding experiments, it is apparent from inspection of Table 3 that the regression slope estimates for anticipations were strongly negative, whereas the regression slope estimates for postponements were weakly positive.

Summary of Experiments

The results of the three experiments are unambiguous—across manipulations of sequence length (Experiments 1 & 2), articulatory suppression (Experiment 2), and temporal grouping (Experiment 3), the LDFs underpinning serial recall of sequences of visual items were consistently negative, and additionally exhibited a reduction in slope for postponements compared to anticipations. This empirical pattern is consistent with the theoretical prediction of a mechanism embodying a primacy gradient, position marking, and response suppression—viz. the PG + PM + RS model; it is incompatible with the error latency predictions of the four competing models and mechanisms for representing serial order (Figure 3).

Although we have hitherto restricted our analysis to the aggregate LDFs, inspection of the LDFs for individual participants reveals a similar pattern. Figure 9A and B show, respectively, the distributions of slopes of LDFs for anticipations and postponements for each individual participant for each experiment and condition. It is apparent from inspection of Figure 9A that the majority of slopes of LDFs for anticipations are negative (75%), and a substantial percentage of these slopes are steep slopes falling in the range of -100 ms to -1500 ms (47%). By contrast, it is visible from inspection of Figure 9B that a larger percentage of slopes of LDFs for postponements are positive (78%), and the majority of slopes are shallow slopes concentrated in the region of -100 ms to 100 ms (91%). Figure 9 thus confirms that the steep negative anticipation slopes and relatively flat postponement slopes of the aggregate LDFs are an accurate reflection of the individual participant LDFs from which they are constructed, and therefore represent a robust and general feature of visual STM.

Model Fitting

Although the LDFs observed across the three experiments are most compatible with the error latency prediction of the PG + PM + RS model, one limit of the initial simulations is that they are based on a priori predictions generated from a single set of model parameter values chosen somewhat arbitrarily to produce comparable levels of performance across the models. It remains possible therefore that models other than the PG + PM + RS model might be able to accommodate the observed LDF under different model parameter values. To address this question, we next report further simulations in which the parameters of the models were estimated by fitting them to the recall latency distributions of the different sequence length conditions of Experiment 2.

Another aspect of the empirical data that would benefit from further quantitative modeling is the results obtained with grouped sequences in Experiment 3. A puzzling feature of those data is that whilst grouping did not affect the probability of interpositions, it nevertheless affected their recall latencies—there was a speed up in the recall times of -3 displacements and a slow down in the recall times of +3 displacements. On the one hand, the lack of an effect of grouping on response probabilities is seemingly at variance with the standard approach to modeling grouping effects, whereby position markers representing the position of groups in sequence are combined with position markers representing the position of items within groups. These results might be consistent instead with an alternative approach advanced by Hurlstone and Hitch (2015) to explain grouping effects in spatial STM, whereby position markers representing the position of groups in the sequence are combined with position markers representing the position of items in the sequence as a whole, rather than within groups. However, complicating the matter, the effects of grouping on the recall times of interpositions are consistent with the former model but not the latter. Thus, there is some uncertainty regarding how positional information is represented in grouped visual sequences, and this is therefore a situation in which model comparisons might help to adjudicate between the two competing approaches.

Accordingly, we fitted two versions of the positional models to the grouped data of Experiment 3, one in which position was coded via markers representing the position of groups and items within groups, and one in which position was coded via markers representing the

position of groups and items within sequence. This yielded a total of eight models for comparison—two versions of each of the models incorporating position marking (viz. PM, PM + RS, PM + OI, and PG + PM + RS; see Appendix B for precise details of how the models were extended to grouped sequences). To distinguish between the two sets of models, we use the super-scripts ^{pwg} (denoting position within group) and ^{pws} (denoting position within sequence). For example, the acronym PM^{pwg} + RS, refers to the version of the PM + RS model using position within-group markers, whereas the acronym PM^{pws} + RS refers to the version of the PM + RS model using position within-sequence markers.

To summarize, the aims of the modeling were: (1) to verify that models other than the PG + PM + RS model cannot account for the shape of the LDF for ungrouped sequences, and (2) to shine further light on the nature of the positional representations underlying grouped visual serial recall. For each target data set (viz. the three sequence length conditions of Experiment 2 and the grouped condition of Experiment 3), the models were fit to the recall latency distributions for each output position using a maximum likelihood method for quantiles (cf. Heathcote, Brown, & Mewhort, 2002). A polytope optimization algorithm was used to find the parameters of each model, for each data set, that minimized the discrepancy between the observed and predicted frequencies of recall latencies from different input positions, at each output position, falling into bins defined by categories obtained from quantile-averaged group data. To control for model complexity, the models were evaluated and compared using Akaike information criterion (AIC; Akaike, 1973) and Bayesian information criterion (BIC; Schwarz, 1978) scores, which were converted into AIC and BIC weights (Burnham & Anderson, 2002; Wagenmakers & Farrell, 2004) to facilitate identification of the best model (see Appendix C for precise details of the model fitting and evaluation procedure).

Model Selection

Ungrouped sequences. Starting with the results for ungrouped sequences, the best fitting model parameters and associated goodness-of-fit statistics can be inspected in Tables 4 and 5, respectively. It can be seen from inspection of the latter table that for each sequence length, both the AIC and BIC scores were smallest for the PG + PM + RS model by a considerable

margin. The PG + RS model had the next smallest AIC and BIC scores, followed by the PM + RS model, then the PM + OI model. The PM model consistently had the largest AIC and BIC scores. Although the differences in AIC and BIC are large enough to be considered non-trivial, the AIC and BIC weights shown in Table 5 confirm that the evidence in support of the PG + PM + RS model is decisive.

Of the five models under comparison, the best three fitting models involved different combinations of a primacy gradient, position marking, and response suppression. To provide further support for the necessity of a model based on all three representational principles, we conducted likelihood ratio tests (see Appendix C) comparing the reliability of the differences in goodness-of-fit between: (1) the general model (viz. the PG + PM + RS model) and a restricted version in which position marking was eliminated (viz. the PG + RS model), and (2) the general model and a restricted version in which the primacy gradient was eliminated (viz. the PM + RS model). The first set of comparisons confirmed that a model incorporating all three principles provided a better fit than a model excluding position marking: $\chi^2(2) = 1042$, $p < .001$, for four-item sequences, $\chi^2(2) = 1404$, $p < .001$, for five-item sequences, and $\chi^2(2) = 2354$, $p < .001$, for six-item sequences. Similarly, the second set of comparisons confirmed that a model incorporating all three principles provided a better fit than a model excluding a primacy gradient: $\chi^2(2) = 2724$, $p < .001$, for four-item sequences, $\chi^2(2) = 5606$, $p < .001$, for five-item sequences, and $\chi^2(2) = 8446$, $p < .001$, for six-item sequences.

Grouped sequences. Turning to the results for grouped sequences, Tables 6 and 7 show, respectively, the best fitting model parameters and associated goodness-of-fit quantities for the eight models under comparison. It is apparent from inspection of the latter table that the smallest AIC and BIC scores were obtained by the two versions of the PG + PM + RS model, followed by the two versions of the PM + RS and PM + OI models. The two versions of the PM model had the largest AIC and BIC scores. Of the two variants of each of the four models, the AIC and BIC scores were consistently smaller for the versions incorporating position within-group markers than the versions incorporating position within-sequence markers. The AIC and BIC weights identify the PG + PM^{pwg} + RS model unambiguously as the preferred model of the data.

Simulation Results

Ungrouped sequences. Figure 10 shows the accuracy serial position curves, transposition gradients, and latency serial position curves predicted by the models under their best-fitting parameters (the first, second, and third row of panels show the results for four-, five-, and six-item sequences, respectively). Figures 10A, D, and G show the predicted accuracy serial position curves from which it can be seen that all models generated the effects of primacy, recency, and sequence length observed empirically (Figure 7A).

The transposition gradients predicted by the models in Figures 10B, E, and H mirror the empirical data in Figure 7B in exhibiting a sharp peak at displacement 0, approximately symmetrical error gradients for anticipations and postponements, and a locality constraint on movement errors. However, it can be seen that the PG + RS model under-predicts the probability of postponements at large displacements. This is because the PG + RS model exhibits a strong "fill-in" tendency—due to the gradient-based representation of order—which means that early items will be activated more strongly than later items—the longer an item is missed in recall, the more likely that it will be recalled at the next output position, due to the increasing disparity between its activation and that of its recall competitors.

Figures 10C, F, and I show the predicted latency serial position curves. To accommodate preparatory processes that precede the production of the first response, a 3000-ms constant has been added to the recall time for the first output position for each model to increase graphical correspondence between the model predictions and the data illustrated in Figure 7C.[1] It is apparent from inspection of the predicted latency curves that with the exception of the PM model—which predicts a relatively flat serial position curve for all serial positions but the first, with a recency effect for the final item—the models predict a speeding up of recall over serial

[1]These preparatory processes that precede the production of the sequence—and groups in grouped sequences—have variously been attributed to the priming of a low-level motor output buffer (Sternberg, Monsell, Knoll, Wright, 1978) or memory search through a hierarchical representation of the order of elements in a sequence (Anderson & Matessa, 1997; Farrell, 2012; Farrell & Lelièvre, 2012). We do not model these processes here because they do not assist in discriminating between the models—since the effects of output position on recall times are removed from the model LDFs—and because additional ancillary assumptions are necessary to accommodate them.

positions, but the extent of that speed up is less pronounced than witnessed in the empirical data.[2]

In summary, notwithstanding some minor differences between models, their predictions are qualitatively similar, rendering it difficult to adjudicate between them on the basis of the three conventional recall measures.

Turning to the simulations results of chief interest, Figure 11 shows the LDFs predicted by the models under their best-fitting parameters for sequences of four- (panel A), five- (panel B), and six-items (panel C). The model LDFs for each sequence length do not differ qualitatively from the initial predictions in Figure 3. As before, the PM model predicts V-shaped LDFs, whereas the PM + RS and PM + OI models predict partially asymmetric V-shaped LDFs in which the slope for postponements is shallower than for anticipations. The error latency predictions of these models are clearly at variance with the empirical data shown in Figure 6B. In contrast to these models, the PG + RS model predicts a monotonically negative LDF. The predictions of this model are a much better approximation of the empirical pattern shown in Figure 6B. However, the PG + RS model predicts a negative LDF slope for postponements across all sequence lengths, which is still at variance with the relatively flat postponement slopes observed empirically. Like the PG + RS model, the PG + PM + RS model predicts a negative LDF, however, consistent with the empirical data the slope of the displacement function for postponements is relatively flat.

Grouped sequences. Turning to the predictions for grouped sequences, Figure 12 shows the accuracy serial position curves, transposition gradients, and latency serial position curves for the models pairing position of group with position within-group markers (upper panels), and the models pairing position of group with position within-sequence markers (lower panels). Looking at the accuracy serial position curves, it can be seen that the former models all predict within-group primacy and recency effects (Figure 12A), consistent with the empirical data (Figure 8A), whereas

[2]This departure from the non-monotonic latency curves predicted at the outset arose for different reasons in the four models. In the PM + RS model, it was due to an increase in the distinctiveness of the position markers (θ); in the PM + OI model, it was due to a decrease in the weighting of the position markers (Ω); and in the PG + RS and PG + PM + RS models it was due to an increase in value of the parameter ρ, which rendered the primacy gradient less steep than in the initial simulations.

the latter models struggled to reproduce these mini primacy and recency effects (Figure 12C).

The models combining position of group with position within-sequence markers did, however, predict monotonically decreasing transposition gradients (Figure 12E) in accordance with the data (Figure 8B), whereas the models combining position of group with position within-group markers predicted non-monotonic transposition gradients characterized by localized peaks at ±3 displacements. Thus, the latter models predicted an increase in the frequency of interpositions that was not witnessed empirically (Figure 8B).

The latency serial position curves for the models coupling position of group with position within-group markers can be scrutinized in Figure 12C, whereas the corresponding curves for the models coupling position of group with position within-sequence markers can be interrogated in Figure 12F. As for the predictions for ungrouped sequences, a 3000-ms constant has been added to the recall times for the first output position but, in addition, a 1000-ms constant has also been added to the recall times for the fourth output position to once more increase graphical correspondence between the model predictions and the data in Figure 8C. Notwithstanding these augmentations to the model predictions, it is visible from inspection of Figure 12C and Figure 12F that the latency curves predicted by both sets of models are a good approximation of the empirical data. However, with the possible exception of the two instantiations of the PG + PM + RS model, the models do not predict as marked a speed up in recall across serial positions as witnessed in the data.

In summary, as per the predictions for ungrouped sequences, it is difficult to adjudicate between the four models on the basis of the conventional serial recall measures. However, the models integrating position of group with position within-group markers had the edge over the models integrating position of group with position within-sequence markers. Although the former models predicted an increase in the frequency of interpositions—which is incompatible with the data—only these models were able to reproduce the effects of grouping on the accuracy serial position curve.

Considering now the error latencies for transpositions, the predicted LDFs for the models pairing position of group with position within-group markers can be inspected in Figure 13A, whereas the predicted LDFs for the models pairing position of group with position

within-sequence markers can be inspected in Figure 13B. Irrespective of the nature of the positional representations employed, the overall shape of the LDFs predicted by the four models is similar to their corresponding predictions for ungrouped sequences—the PM^{pwg} and PM^{pws} models predict symmetric V-shaped LDFs; the PM^{pwg} + RS and PM^{pws} + RS models predict partially asymmetric V-shaped LDFs, as do the PM^{pwg} + OI and PM^{pws} + OI models; whereas the PG + PM^{pwg} + RS and PG + PM^{pws} + RS models predict asymmetric LDFs characterized by an overall negative latency–displacement relationship, but with a flatter slope for postponements than for anticipations. The key difference between the two sets of models is that the models combining position of group with position within-group markers predict a speed-up in the recall times for ±3 displacements, whereas the models combining position of group with position within-sequence markers do not predict such discontinuities. Looking at the data in Figure 6C, it is apparent that the PG + PM^{pwg} + RS and PG + PM^{pws} + RS models provide the best account of the empirically observed LDF. However, of the two models, the PG + PM^{pwg} + RS model arguably provides the better account, since it captures the speed-up in the recall time for -3 displacements; it did not, however, capture the slow down in recall times for +3 displacements, predicting instead a speed-up in the recall time for these errors. The PG + PM^{pws} + RS model failed to capture either of these subtleties of the LDF for grouped sequences.

General Discussion

The results of the experiments and quantitative modeling suggest that the serial recall of a sequence of visual items is driven by a competitive queuing mechanism, within which serial order is represented via a primacy gradient of activations over items, associations between items and position markers, and with suppression of items once they have been recalled. Across manipulations of sequence length, articulatory suppression, and temporal grouping the LDFs observed in the three experiments were consistently negative overall but with a reduction in slope for postponements compared to anticipations. This empirical pattern was a robust feature of the aggregate LDFs as well as the individual participant LDFs from which they were derived, thus confirming that the three representational principles are core ingredients in any adequate model of serial order in visual STM. None of the four alternative mechanisms for the representation of

serial order were able to reproduce the shape of the observed LDFs, even when the models were fit directly to the recall latency distributions for grouped sequences and ungrouped sequences of varying lengths. In previous work (Hurlstone & Hitch, 2015), we have also shown that a representational mechanism embodying a primacy gradient, position marking, and response suppression is the only one of the five mechanisms considered that consistently predicts the observed latency–displacement relationship across broad variations of its parameters. The data and modeling are consistent with those reported by Farrell and Lewandowsky (2004) and Hurlstone and Hitch (2015) with sequences of verbal and spatial items, respectively, and give general support for the notion that common mechanisms and principles are implicated in the representation of serial order across the verbal, visual, and spatial STM domains.

The Pivotal Role of a Primacy Gradient and Response Suppression

If there are two representational principles that the empirical LDFs speak to the most, it is a primacy gradient coupled with response suppression. Across the three experiments, the observed LDFs were consistently negative overall, such that the recall times for items reported too soon were markedly slower than for items reported too late. The theoretical LDF analyses clearly show that only the models incorporating a primacy gradient with response suppression—viz. PG + RS; PG + PM + RS—predict this asymmetry in the recall times of anticipations and postponements. Indeed, the same analyses show that the positional models consistently predict a symmetric or partially symmetric V-shaped LDF, thus ruling out a representational mechanism based on position marking alone. Qualified support for the role of these two representational principles was provided by the results of the AIC and BIC model comparisons for ungrouped sequences, which confirmed that the two models incorporating a primacy gradient and response suppression provided a better fit to the data than the models omitting these principles.

Nevertheless, whilst a primacy gradient and response suppression are necessary to explain the finding that anticipations are slower than postponements, it is apparent that a representational mechanism based on a primacy gradient and response suppression alone is not sufficient to fully account for the shape of the observed LDFs. Specifically, it is necessary to augment the basic primacy gradient and response suppression mechanism with a set of position

markers in order to reproduce the observed flattening of the LDF slopes for postponements. This was once again corroborated by the AIC and BIC model comparisons, which provided unambiguous support for the PG + PM + RS model. Converging evidence for the operation of position marking was also provided by the results of the temporal grouping manipulation, which we consider next.

Converging Evidence for Positional Representations

Converging evidence for the operation of position marking in visual STM was provided by the results of the temporal grouping manipulation employed in Experiment 3. Consistent with previous studies of grouped verbal (Farrell & Lewandowsky, 2004; Frankish, 1985, 1989; Hitch et al., 1996; Maybery et al., 2002; Ryan, 1969) and spatial serial recall (Hurlstone & Hitch, 2015; Parmentier et al., 2006), grouping a sequence of visual items induced a number of qualitative changes in performance including effects of primacy and recency within groups; longer recall latencies at the beginning of groups; and a reduction in adjacent-neighbor transpositions straddling group boundaries. However, the effects of grouping were smaller in magnitude than those observed with verbal and spatial stimuli—we did not observe a reliable main effect of the grouping manipulation on performance. One possible reason for this is that there is no output mechanism by which visual non-spatial stimuli can be rehearsed, whereas verbal stimuli can be rehearsed sub-vocally (Baddeley, 1986) and spatial stimuli can be rehearsed via covert shifts of spatial selection attention (Awh & Jonides, 2001) or eye-movements (Postle, Idzikowski, Della Sala, Logies, & Baddeley, 2006). That output processes contribute to the magnitude of grouping effects is supported by the finding that grouping effects with visually—but not auditorily—presented verbal stimuli increase in size as the length of the inter-group pause (and the opportunity for rehearsal) increases (Frankish, 1985), and by the finding that grouping effects are eliminated—or nearly so—under conditions of concurrent articulatory suppression (Hitch et al., 2006).

There is one hallmark feature of grouped verbal serial recall that did not materialize with grouped visual serial recall, however; this is the increase in the frequency of interpositions—transpositions between groups maintaining their within-group serial position. The

same discrepancy was also noted in our earlier work exploring effects of grouping on the serial recall of spatial sequences (Hurlstone & Hitch, 2015; see also Parmentier et al., 2006). We have previously speculated that this discrepancy might be attributable to a subtle difference in the way positional information is represented in grouped verbal and nonverbal sequences. Specifically, we postulated that position in grouped nonverbal sequences might be represented via markers coding the position of groups and the position of items within the sequence as a whole, rather than the standard approach to modeling grouping effects in the verbal domain, whereby markers coding the position of groups in the sequence are combined with markers coding the position of items within groups. However, we did not provide a formal test of this hypothesis, whereas one of the aims of the quantitative modeling reported here was to pit our alternative representational scheme against the standard approach to modeling grouping effects.

The results of the model fitting of grouped sequences call into question our alternative representational scheme—at least as applied to the current data. Neither of the models implementing this approach provided a satisfactory account of the observed effects of grouping. In particular, although the models predicted some within-group serial position effects on accuracy, they were not as marked as witnessed empirically. At first blush, this result was surprising because this representational scheme is capable of generating pronounced within-group serial position effects provided that the position of group markers are sufficiently distinctive. However, upon closer scrutiny it seems that the problem for the models implementing this approach is that the parameter settings that generate these effects simultaneously produce too many within-group transposition errors and too few between-group transposition errors. These parameter settings are associated with poorer fits to the data than parameter settings that produce little or no within-group serial position effects, but which provide a better account of the distribution of transpositions within and between groups. One might wonder whether this representational scheme would fare better if we had allowed the parameter that weights the influence of the position of group and position within-sequence markers (viz. λ) to vary during the fitting (this parameter was fixed so that equal weight was assigned to the position of group and position within-sequence markers). We have explored this possibility in a separate set of simulations and found that it did not improve the ability of the models to accommodate within-group serial

position effects (indeed, varying λ actually resulted in the models generating less scalloping than the limited amount they already produce, with the resultant serial curves bearing a closer resemblance to those expected for ungrouped sequences).

Although the results of the model comparisons cast doubt on our alternative perspective, they do not support the standard approach either. This is because the models combining position of group with position within-group markers predicted an increase in interpositions (i.e., errors at transposition displacements ± 3) that was not observed empirically. Nevertheless, the models also predicted shorter recall times for these errors on the LDF, which is partially consistent with the data. Notably, reduced response latencies were observed for -3 displacements but not for +3 displacements; on the contrary, we actually observed increased response latencies to these errors. The partial mismatch between models and data regarding the effects of grouping on interposition latencies raises the question of whether the effects in the data are merely coincidental or whether they constitute evidence for the operation of within-group positional codes? In our previous work (Hurlstone & Hitch, 2015) and that of Farrell and Lewandowsky (2004), discontinuities in the LDF were also observed under conditions of temporal grouping. Although these discontinuities occasionally manifested at locations corresponding to interpositions, most of the time they did not. Moreover, as we have observed here, the direction of the effects was inconsistent, such that sometimes shorter recall latencies were observed—characterized by local troughs in the LDF—whereas sometimes longer recall latencies were observed—characterized by local peaks (by comparison, the models considered here can only produce troughs—but not peaks—in the LDF). This leads us to a second related point, namely that the discontinuities observed here and in the work just alluded to arose independently of any increase or decrease in the probability of the specific errors. By contrast, in the models any change in the recall time of an error must necessarily be accompanied by a corresponding change in that error's probability. These considerations lead us to suspect that the discontinuities in the LDF are coincidental rather than emblematic of the operation of within-group positional codes. They might, for example, be symptomatic of increased motor response variance induced by the grouping manipulation.

That the effects of grouping reported here cannot be reconciled with either of the two approaches considered raises the question of how best to explain those results? One possibility is

that the data may be explicable in terms of a more complex model in which positional information is coded along three rather than two dimensions—viz. a model incorporating group-position-in-sequence, item-position-within-group, and item-position-within-sequence position markers. This might represent a general approach to coding positional information in STM, and what might vary across the verbal, visual, and spatial domains is the relative weight attached to the three sets of position markers. In an additional set of simulations (not reported for brevity), we applied versions of the four positional models built on this assumption to the grouped condition of Experiment 3. We found that whilst these models fared better against the data than the models pairing position of group with position within-group markers, once the models were penalized for their additional model complexity—viz. via AIC and BIC—their advantage was neutralized. Moreover, at variance with the data, the models continued to predict interpositions.

A more radical but nevertheless plausible interpretation of the results of Experiment 3 is that rather than reflecting the action of a multidimensional representation of serial order based on positional information, they simply reflect the action of a selective encoding strategy adopted by participants. To explain, it may be that the localized recall advantage witnessed for items at group boundaries arose because participants paid extra attention to these items during the encoding of the sequence. The effects of grouping on the latency serial position curve are more difficult to explain on this interpretation, but one possible account is that in this instance the additional pause left before accessing the second group is reflective of a pattern matching heuristic—viz. a temporal gap at input affords a temporal gap at output—rather than the consequence of some underlying multidimensional or hierarchical representation of the sequence. Thus, although the effects of grouping observed with visual stimuli may superficially look like those observed with verbal and spatial stimuli, this may merely be coincidental, and they may in fact reflect fundamentally different processes altogether.

In brief, it is unclear how precisely the coding of positional information in grouped visual sequences differs from grouped verbal sequences. We have formally tested two alternative approaches to positional coding and found both to be lacking. Unfortunately, this is a question to which we are currently unable to provide a clear answer, and further work will be required to test alternative hypotheses.

Implications for Theories of Working Memory

We next consider the broader implications of our results for theories of working memory. To frame our discussion, we begin with a reprisal of the conclusions drawn in our previous work (Hurlstone and Hitch, 2015) where—based on a similar chronometric analysis of order errors in spatial serial recall—we provided evidence for the operation of a primacy gradient, position marking, and response suppression in spatial STM. This work itself, replicated and extended the earlier work of Farrell and Lewandowsky (Farrell & Lewandowsky, 2004; Lewandowsky & Farrell, 2008) who examined the dynamics of transpositions in verbal serial recall and also found empirical support for a representational mechanism embodying these three principles.

Hurlstone and Hitch (2015) asked how the competitive queuing mechanism and the three principles for representing serial order map onto the different components of the working memory model of Baddeley and Hitch (Baddeley, 1986, 2000, 2007; Baddeley & Hitch, 1974). They presented arguments and evidence in favor of the view that the competitive queuing mechanism, along with the primacy gradient and response suppression are implemented in a modality-specific manner within the working memory slave systems; that is, the phonological loop and the spatial component of the visuospatial sketchpad—viz. the "inner scribe" (Logie, 1995)—each possess their own competitive queuing sequence planning and control mechanisms and processes for generating a primacy gradient and implementing response suppression (see Hurlstone and Hitch, 2015 for a discussion of the theoretical basis for these claims). However, in common with other serial recall theorists (Burgess & Hitch, 1999; Henson 1998; Smyth, Hay, Hitch, & Horton, 2005; Tremblay, Guérard, Parmentier, Nicholls, & Jones, 2006), they noted the possibility that positional information in the verbal and nonverbal domains might be encoded via a common domain-general mechanism. For example, Burgess and Hitch (1999) postulated 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 later work (Burgess & Hitch, 2005), the same authors postulated that the locus of their context signal within the working memory framework might be the episodic buffer (Baddeley, 2000)—a component of the working memory model responsible for integrating information from the working memory slave systems and long-term memory—and that the buffer might serve as a common positional coding mechanism for items maintained in the

phonological loop and the visuospatial sketchpad.

The present results build on this analysis in two respects. First, they suggest that the competitive queuing mechanism, primacy gradient, and response suppression are also key properties of the visual component of the visuospatial sketchpad—viz. the "visual cache" (Logie, 1995). Second, by showing—as Hurlstone and Hitch (2015) did with spatial sequences—that temporal grouping exerts effects on the recall of visual sequences that are similar and dissimilar to those documented with verbal sequences, the present results place further constraints on the debate regarding the domain-general or domain-specific nature of positional coding in STM. In particular, the difference in the effects of grouping across domains—viz. the presence of interpositions in the verbal domain, but not the visual and spatial domains—seems to cast doubt on the hypothesis that a shared mechanism—perhaps mediated by the episodic buffer—is responsible for coding positional information in the two working memory sub-systems. Instead, this discrepancy seems to point to the existence of distinct positional coding mechanisms that function in a similar manner, but posses subtly different representational characteristics.

Is the notion of distinct mechanisms for representing positional information tied to different content domains controversial? We believe not. Indeed, developments in this direction are already taking place. As a case in point, recently Hartley et al. (2016) have presented a model of auditory-verbal STM for serial order that relies on a context-driven timing signal to represent the serial position of items. The timing signal is based on a population of oscillators—i.e., frequency sensitive detectors—sensitive to local changes in the amplitude envelope of incoming speech. As such, the mechanism is tailored specifically for the spoken modality. Hartley et al. (2016) show that this mechanism provides an impressive account of a data pattern currently believed to be unique to auditory-verbal STM, whereby the product of the group sizes of different temporal grouping patterns—an empirical yardstick for their degree of regularity or rhythm—is a potent predictor of recall performance on those different grouping patterns.

In summary, notwithstanding the evidence for common representational principles presented here, the representation and generation of serial order in STM is likely governed by a mixture of domain-general and domain-specific processes. There is surely some cross-modal ordering of items and events, but at the output level there are equally surely differences between

verbal, visual, and spatial sequences.

Potential Limitations

Before concluding, we briefly consider one potential shortcoming of our modeling approach. For simplicity, our simulations were conducted using a single layer lateral inhibition network architecture resembling the competitive choice layer in competitive queuing models. However, the competitive queuing mechanism is a two layer output module in which the competitive choice layer is complemented by a parallel planning layer. In a fully implemented version of the competitive queuing mechanism both of these layers are implemented in a dynamic fashion through a combination of recurrent excitation and lateral inhibition (Bullock & Rhodes, 2003; Davelaar, 2007; Grossberg & Pearson, 2008). However, given the different roles of the two layers the dynamics of each must necessarily differ. Since the function of the parallel planning layer is to maintain a plan of the sequence the level of lateral inhibition must be weak, otherwise this layer will only contain a plan for a single item rather than a parallel plan for the sequence as a whole. By contrast, since the function of the competitive choice layer is to select a single item from amongst a set of parallel activated items, the level of lateral inhibition must be strong in order to resolve the response competition between items. The different dynamics of the two layers means that they will interact non-linearly in a manner that is not captured by a single-layer architecture. This might, ostensibly, change the behavior of the models under comparison here. Although we are confident that this would not be the case, an important avenue for future work is to establish whether the LDF predictions of the models do indeed generalize to a fully implemented dynamic version of the competitive queuing mechanism.

Efforts along these lines may also provide a fruitful in-road to explaining aspects of the data in the experiments reported here that are beyond the purview of our simple single layer architecture. We refer specifically to the pauses preceding the production of the first item in ungrouped sequences, and the pauses preceding the production of the first item of each group in grouped sequences, which have also been documented with verbal (e.g., Farrell & Lewandosky, 2004) and spatial (e.g., Hurlstone & Hitch, 2015) sequences. The competitive queuing mechanism potentially provides a process explanation of these sequence preparation effects. On a competitive

queuing account, the pause preceding the production of the first item in ungrouped sequences might arise because it takes time for the initial activations in the parallel planning layer to increase above threshold before they can be projected onto the competitive choice layer. The long pause preceding the production of the first item in the first group in grouped sequences would arise for the very same reason. However, when the group position in sequence context cue is withdrawn from the parallel planning layer after the last item in the group has been output, there should be a dip in the activations of items in this layer causing them to temporarily drop below threshold (assuming that the recurrent excitation and inhibition is accompanied by activation decay—which is typically the case e.g., Davelaar, 2007). Accordingly, when the context cue for the next group is presented to the parallel planning layer it will take time for the activations of items to once again raise above threshold, thereby generating a lengthened pause prior to the production of the first item of the new group being accessed.

Conclusions

To conclude, across three experiments exploring the dynamics of transpositions in a visual serial recall task we observed a consistent pattern whereby anticipation errors are slower than postponements. Additionally, whereas the latencies of anticipations increase with displacement distance, the latencies of postponements are relatively insensitive to degree of displacement. This error latency profile is consistent with that observed in previous studies of serial recall with verbal (Farrell & Lewandowsky, 2004) and spatial (Hurlstone & Hitch, 2015) sequences, and supports the prediction of a competitive queuing mechanism within which serial order is represented via a mechanism embodying a primacy gradient, position marking and response suppression. Taken together, these results suggest that the same mechanism is implicated in the representation of serial order across the verbal, visual, and spatial STM domains. However, the same results also point to differences across domains in the manner in which positional information is represented in grouped sequences. A challenge for future research is to elucidate how precisely the coding of positional information differs across the verbal, visual, and spatial domains, and whether the similarities and differences are best understood in terms of a domain-general mechanism or domain-specific mechanisms specialized for different content domains.

References

Agam, Y., Bullock, D., & Sekuler, R. (2005). Imitating unfamiliar sequences of connected linear motions. *Journal of Neurophysiology, 94,* 2832-2843. doi: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. doi:10.1167/7.5.1

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19,* 716-723. doi:10.1109/TAC.1974.1100705

Anderson, J. A. (1995). *An introduction to neural networks.* Cambridge, MA: MIT Press.

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

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

Avons, S. E. (2007). Spatial span under translation: A study of reference frames. *Memory and Cognition, 35,* 402-417. doi: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. doi:10.1080/027249899391296

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and working memory. *Trends in Cognitive Sciences, 5,* 119-126. doi:10.1016/S1364-6613(00)01593-X

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. doi:10.1016/S1364-6613(00)01538-2

Baddeley, A. D. (2007). *Working memory, thought and action.* Oxford, UK: Oxford University Press. doi: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. doi: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: Academic Press. doi:10.1016/S0079-7421(08)60452-1

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

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

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

Bullock, D. (2004). Adaptive neural models of queuing and timing in fluent action. *Trends in Cognitive Sciences, 8,* 426-433. doi: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*, second edition (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. doi:10.1016/0749-596X(92)90022-P

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

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

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multi-modal inference: A practical information-theoretic approach.* New York, NY: Springer-Verlag. doi:10.1007/b97636

Crannell, C. W., & Parrish, J. N. (1957). A comparison of immediate memory span for digits, letters, and words. *Journal of Psychology, 44,* 319-327. doi:10.1080/00223980.1957.9713089

Couture, M., & Tremblay, S. (2006). Exploring the characteristics of the visuospatial Hebb repetition effect. *Memory & Cognition, 34,* 1720-1729. doi: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. doi:10.1177/105971230607625

Ebbinghaus, H. (1886/1964). *Memory: A contribution to experimental psychology.* New York:

Dover. doi:10.1037/10011-000

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. doi:10.1016/S0001-6918(00)00054-8

Farrell, S. (2006). Mixed-list phonological similarity effects in delayed serial recall. *Journal of Memory and Language, 55,* 587-600. doi:10.1016/j.jml.2006.06.002

Farrell, S. (2012). Temporal clustering and sequencing in short-term memory and episodic memory. *Psychological Review, 119,* 223-271. doi:10.1037/a0027371

Farrell, S., & Lelièvre, A. (2009). End anchoring in short-term order memory. *Journal of Memory and Language, 60,* 209-227. doi: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, & Cognition, 38,* 1659-1674. doi:10.1037/a0028469

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review, 9,* 59-79. doi: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. doi:10.1016/j.jml.2004.03.007

Frankish, C. (1985). Modality-specific grouping effects in short-term memory. *Journal of Memory and Language, 24,* 200-209. doi: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. doi:10.1037//0278-7393.15.3.469

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

Grossberg, S. (1978). A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In R. Rosen & Snell (Eds.), *Progress in theoretical biology* (Vol. 5, pp. 233-374). New York: 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. doi: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. doi:10.1037/0278-7393.34.3.556

Hartley, T., Hurlstone, M. J., & Hitch, G. J. (2016). Effects of rhythm on memory for spoken sequences: A model and tests of its stimulus-driven mechanism. *Cognitive Psychology, 87,* 135-178. doi:10.1016/j.cogpsych.2016.05.001

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review, 9,* 394-401. doi:10.3758/BF03196299

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

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

Henson, R. N., Hartley, T., Burgess, N., Hitch, G., & Flude, B. (2003). Selective interference with verbal short-term memory for serial order information: A new paradigm and tests of a timing signal hypothesis. *Quarterly Journal of Experimental Psychology, 56,* 1307-1334. doi:10.1080/02724980244000747

Henson, R. N., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology, 49A,* 80-115. doi: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, 49A,* 116-139. doi:10.1080/713755609

Horton, N., Hay, D. C., & Smyth, M. M. (2008). Hebb repetition effects in visual memory: The roles of verbal rehearsal and distinctiveness. *Quarterly Journal of Experimental Psychology, 61,* 1769-1777. doi:10.1080/17470210802168674

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: Academic Press.

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. doi:10.1037/a0034221

Hurlstone, M. J., & Hitch, G. J. (2015). How is the serial order of a spatial sequence represented? Insights from transposition latencies. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 41,* 295-324. doi:10.1037/a0038223295

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. doi:10.1080/17470210701634537

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. doi:10.1037//0278-7393.21.4.1008

Lamberts, K. (2004). Mathematical Modelling of Cognition. In K. Lamberts and R. Goldstone (Eds.), *Handbook of Cognition* (pp. 407-421). London: Sage.

Lewandowsky, S., Duncan, M., & Brown, G. D. A. (2004). Time does not cause forgetting in short-term serial recall. *Psychonomic Bulletin & Review, 11,* 771-790.

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

Lewandowsky, S., Farrell, S. (2011). Computational modeling in cognition: Principles and practice. Thousand Oaks, CA: Sage. doi:10.4135/9781483349428

Logie, R. (1995). *Visuo-spatial working memory.* Hove: Lawrence Erlbaum Associates. doi: 10.1002/acp.746

Logie, R. H., Della Sala, S., Wynn, V., & Baddeley, A. D. (2000). Visual similarity effects in immediate verbal serial recall. *Quarterly Journal of Experimental Psychology, 53,* 626-646. doi:10.1080/713755916

Logie, R. H., Saito, S., Morita, A., Varma, S., & Norris, D. (2015). Recalling visual serial order

for verbal sequences. *Memory & Cognition.* doi:10.3758/s13421-015-0580-9

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. doi:10.1016/S0749-596X(02)00014-1

Meiser, T., & Klauer, K. C. (1999). Working memory and changing state hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 1272-1299. doi:10.1037//0278-7393.25.5.1272

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal, 7,* 308-313. doi: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. doi: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. doi:10.1037/0278-7393.31.1.175

Page, M. P. A. (2005). Connectionist models of short-term memory for serial order. In G. Houghton (Ed.), *Connectionist models in cognitive psychology* (pp. 215-240). Hove: Psychology Press.

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

Page, M. P. A., & Norris, D. (2009). A model linking immediate serial recall, the Hebb repetition effect and the learning of phonological word forms. *Philosophical Transactions of The Royal Society B,* 364, 3737-3753. doi:10.1098/rstb.2009.0173

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. doi: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. doi: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. doi:10.1037/a0013258

Postle, I., Idzikowski, C., Della Sala, S., Logie, R. H., & Baddeley, A. D. (2006). The selective disruption of spatial working memory by eye movements. *Quarterly Journal of Experimental Psychology, 59,* 100-120. doi:10.1080/17470210500151410

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86,* 446-461. doi:10.1037/0033-2909.86.3.446

Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology and Aging, 16,* 323-341. doi:10.1037/f0882-7974.16.2323

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

Saito, S., Logie, R. H., Morita, A., & Law, A. (2008). Visual and phonological similarity effects in verbal immediate serial recall: A test with kanji materials. *Journal of Memory and Language, 59,* 1-17. doi:10.1016/j.jml.2008.01.004

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, 40A,* 940-949. doi: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, 58A,* 909-930. doi:10.1080/02724980443000412

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

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

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.

doi:10.1111/1467-9280.01437

Thapar, A., Ratcliff, R., & McKoon, G. (2003). A diffusion model analysis of the effects of aging on letter discrimination. *Psychology and Aging, 18,* 415-429. doi:10.1037/0882-7974.18.3.415

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. doi:10.1037/0278-7393.32.5.1208

Vousden, J. I., & Brown, G. D. A. (1998). To repeat or not to repeat: The time course of response suppression in sequential behaviour. 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: Springer Verlag.

Wagenmakers, E.-J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review, 11,* 192-196. doi: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. doi: 10.1080/09658210344000279

Appendix A: Formal Description of The Generic Network Architecture Used To Model Transposition Latencies

Following Farrell and Lewandowsky (Farrell & Lewandowsky, 2004; Lewandowsky & Farrell, 2008) and our own earlier modeling (Hurlstone & Hitch, 2015), 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 in competitive queuing models. For each of the representational principles being modeled, we specified 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, since the selection mechanism is insensitive to the exact mechanisms generating the initial activations used to drive recall.

A Common Lateral Inhibition Response Selection Network

A schematic of the response selection network employed for the simulations is illustrated in Figure 14. 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 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 amongst a set of parallel activated representations. As noted above, 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 below). 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:

$$\text{net}_j(t) = a_j(t-1)\,\alpha + \beta \sum_{i \neq j} a_i(t-1) + \epsilon(0, \sigma), \tag{A1}$$

Where net_j is the netinput a node receives from within the layer, a_j represents the initial activation of node j determined by its external input, a_i constitutes the activation of all other nodes in the layer, α and β represent the excitatory and inhibitory weight values, respectively, t corresponds to time, and ϵ is zero mean Gaussian noise with standard deviation σ ($\sigma = .04$ for the initial simulations). (Note that negative activations are not allowed to spread in equation A1 otherwise a node with negative activation would excite rather than inhibit its competitors). The first term on the right hand side of equation A1 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 towards a response. In order to bring the predicted recall times of the network within the range of the observed latencies in the experiments they were multiplied by a scaling factor S ($0 < S \leq 200$; where $S = 50$ for the initial simulations—1 iterative cycle = 50 ms).

Implementation of Representational Principles

The representational principles were implemented through different settings of the starting activations at each output position, which were computed 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 at output

position p was strongest, whilst the activations of item nodes from neighboring serial positions decreased as an accelerating function of their distance from the target item:

$$a_j = \Omega\theta^{(|j-p|)}, \tag{A2}$$

Where θ is a parameter controlling the distinctiveness of the position marking activations ($0 < \theta \leq 1$; $\theta = .65$ for the initial simulations) and Ω is a weighting parameter that determines the distance of each item's initial activation from the response threshold ($0 < \Omega \leq 1$; $\Omega = 1$ for the initial simulations). For each output position, the activations generated by θ 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 Ω. 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 15A shows example starting activations for position marking for the fourth output position in a six-item sequence

Primacy Gradient

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

$$a_j = \phi\rho^{(j-1)}, \tag{A3}$$

Where ϕ is the activation of the item node corresponding to the first input position ($0 < \phi \leq 1$; $\phi = .6$ for the initial simulations) and ρ is a parameter controlling the steepness of the primacy gradient ($0 < \rho \leq 1$; $\rho = .85$ 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 towards 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 below) of those nodes corresponding to previously recalled items. Example starting activations for a primacy gradient for the first output position are shown in Figure 15B.

Primacy Gradient + Position Marking

In line with the seriating mechanisms instantiated in several theories of serial recall (Burgess & Hitch, 1999; Brown et al., 2000; 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:

$$a_j = (1 - \omega)\, \phi\, \rho^{(j-1)} + \omega\, \Omega\, \theta^{(|j-p|)}, \tag{A4}$$

Equation A4 integrates equations A2 and A3 above and incorporates an additional weighting parameter ω ($0 < \omega \leq 1$; $\omega = .5$ 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 15C 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-τ, where τ represents the extent of response suppression ($0 < \tau \leq 1$; $\tau = .95$ for the initial simulations). Example starting activations for a primacy gradient complemented by response suppression for the fourth output position are illustrated in Figure 15D.

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 principles being modeled (e.g., position marking) and was

determined by $\delta \times \sigma \times p$, where δ is a parameter controlling the weighting of output interference across output positions ($0 < \delta \leq 1$; $\delta = .5$ for the initial simulations) and σ 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 15E.

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: (1) position marking (PM); (2) position marking and response suppression (PM + RS); (3) position marking and output interference (PM + OI); (4) a primacy gradient and response suppression (PG + RS); and (5) a primacy gradient, position marking, and response suppression (PG + PM + RS). These are the same set of models as those examined by Farrell and Lewandowsky (Farrell & Lewandowsky, 2004; Lewandowsky & Farrell, 2008) and Hurlstone and Hitch (2015), and they are representative of the range of mechanisms instantiated in contemporary theories of serial recall (see Table 1 of Hurlstone & Hitch, 2015 and Table 2 of Hurlstone et al., 2014). Predictions were generated for each model using 50,000 simulation trials of six-item sequences.

Appendix B: Extension to Grouped Sequences

In this section, we describe how the models implementing position marking can be extended to account for the recall of grouped sequences. As noted in the main text, positional models of serial recall account for grouping effects by assuming that grouped sequences recruit two sets of position markers—one set that encodes the position of groups in the sequence, and a second set that encodes the position of items within groups. This multidimensional representation has been shown to be sufficient to account for the effects of grouping documented with verbal materials. In particular, the use of a set of position markers to represent the within-group position of items is crucial to explaining the between group interposition errors that are a hallmark feature of grouped verbal serial recall.

However, the failure to observe an increase in the frequency of interpositions in grouped visual (current Experiment 3) and spatial (Hurlstone & Hitch, 2015) serial recall raises the possibility that positional information might be represented differently in the visual and spatial domains. Specifically, Hurlstone and Hitch (2015) speculated that in the visual-spatial domain, position markers encoding the position of groups in sequence might be augmented by position markers encoding the position of items in the sequence overall, as opposed to within each group. This should in principle produce the usual effects of grouping on response probabilities, but without fostering an increase in the frequency of interpositions.

To provide a formal test of this hypothesis, we contrasted two approaches to extending the positional models to grouped sequences, one in which position makers encoding the position of groups in sequence were combined with position markers encoding the position of items within groups—viz. the standard approach to modeling grouping effects—and a second in which position markers encoding the position of groups in sequence were combined with position markers encoding the position of items in the sequence overall—viz. the revised approach proposed by Hurlstone and Hitch (2015).

In the following, we outline the equations used to calculate the starting activations for position marking for grouped sequences under the two different approaches to implementing grouped positional representations. When applying the PM, PM + RS, PM + OI, and PG + PM + RS models to grouped sequences, these equations were used in substitute for equation A2.

Position of Group + Position Within-Group

In the implementation of position marking combining information about the position of groups and the position of items within groups, starting activations were chosen that directly reflected the confusability of group positions in the sequence and item positions within groups:

$$a_j = (1 - \lambda)\Omega\theta^{(|g-l|)} + \lambda\Omega\theta^{(|i-r|)}, \tag{B1}$$

Where j indexes an item's input position, g indexes its group's input position, l represents the input position of the group of the target item to-be-recalled at the current output position, i indexes the within-group input position of item j, and r represents the within-group input position of the target item to-be-recalled at the current output position. To illustrate, suppose we wish to calculate the activation of the second item at the sixth output position in a six-item sequence grouped into threes. In this example, $g = 1$ (since item 2 belongs to the first group), $l = 2$ (since the target item 6 belongs to the second group), $i = 2$ (since item 2 appears in the second position in the first group), whilst $r = 3$ (since item 6 appears in the third position in the second group).

As in equation A2, the parameter θ governs the distinctiveness of the position markers ($0 < \theta \leq 1$), whilst Ω is a scaling parameter ($0 < \Omega \leq 1$). The first term in equation B1 generates gradients of activations representing the confusability of the positions of groups in the sequence, whereas the second term generates gradients of activations representing the confusability of the positions of items within groups. The parameter λ weights the amount of attention allocated to the two positional dimensions ($0 < \lambda \leq 1$). When $\lambda = .5$, attention is directed equally to the two dimensions; when $\lambda < .5$, more attention is allocated to the group-position-in-sequence dimension of order; when $\lambda > .5$, more attention is allocated to the item-position-in-group representation of order. Example starting activations based on equation B1 for all output positions in a six-item sequence grouped into threes are shown in Figure 16.

Position of Group + Position Within-Sequence

In the implementation of position marking combining information about the positions of groups and the position of items within sequence, starting activations were chosen that directly reflected the confusability of group and item positions within sequence:

$$a_j = (1 - \lambda)\Omega\theta^{(|g-l|)} + \lambda\Omega\theta^{(|j-p|)}, \tag{B2}$$

Where p represents the output position, and g, l, and j are as before. To illustrate, using the earlier example of calculating the activation of the second item at the sixth output position in a six-item sequence grouped into threes, $g = 1$ (since item 2 belongs to the first group), $l = 2$ (since the target item 6 belongs to the second group), $j = 2$ (since this is the item whose activation is being calculated), and $p = 6$ (since this is the current output position). The first term in equation B2 is the same as in equation B1 and generates gradients of activations representing the confusability of the positions of groups in the sequence. The second term generates gradients of activations representing the confusability of the position of items in the sequence, and is identical to equation A2 used to generate starting activations for position marking for ungrouped sequences. As in equation B1 the parameter λ weights the amount of attention allocated to the two positional dimensions. Figure 17 shows example starting activations based on equation B2.

Incorporating a Primacy Gradient

When position marking was combined with a primacy gradient—viz. the PG + PM + RS model—equations B1 and B2 were augmented as follows:

$$a_j = (1 - \omega)(\text{Eq. B1} \,|\, \text{Eq. B2}) + \omega\phi\rho^{(j-1)}, \tag{B3}$$

Where Eq. B1 | Eq. B2 implements equation B1 or B2, $\phi\rho^{(j-1)}$ corresponds to equation A3 used to compute the primacy gradient over the input position of items, and ω is the same weighting parameter used in equation A4 to determine the attentional weight assigned to the primacy gradient and position markers.

Appendix C: Description of Parameter Fitting Procedure

For each of the four target data sets, the models were fit to the response time distributions for each output position—with the effects of output position subtracted, as per the experiments and initial qualitative model predictions—using a maximum likelihood method for quantiles. For each output position, the response times associated with items recalled from different input positions on the input sequence were sorted in ascending order by each participant and the reaction times for the .5 quantiles that divided the data into two bins were calculated for each individual. These quantile reaction times were then averaged over participants to obtain group response time distributions—a procedure known as Vincent averaging (Ratcliff, 1979). The to-be-fitted data for each output position therefore took the form of a $2 \times sl$ (where sl represents the sequence length) matrix of bins, where the rows represent categories defined by quantile-averaged statistics (one category corresponding to scores at or below the .5 quantile average; the second corresponding to scores above the .5 quantile average) and the columns represent input positions.[3] The values in each bin represent the number of recall latencies for a given category and input position. We opted to fit quantile-averaged group data—rather than individual participant data—to compensate for the fact that participants contributed few or no response times in bins corresponding to large output-input position displacements. However, we note that quantile-averaging preserves information about the shape of the individual response time distributions (Ratcliff, 1979) and yields parameter estimates that are comparable to the average parameter estimates obtained by fitting individual participants (Ratcliff, Thaper, & McKoon, 2001; Thaper, Ratcliff, & McKoon, 2003).

The models were used to obtain predicted frequencies in each bin for each output position.

[3]We have also applied the model fitting and evaluation procedure described here to a larger number of bins using categories defined either by the .25, .5, and .75 quantile averages, or the .1, .3, .5, .7, and .9 quantile averages (both of which are common choices in the evidence accumulation model fitting literature). The relative fits of the models—examined using the goodness-of-fit criteria specified below—using these three- and five-quantile summaries was comparable to our single-quantile summary, except that the discrepancy between the observed and predicted accuracy serial position curves was too large in the former instances. Accordingly, we report the fits of the models using the single-quantile summary, which provided the best approximation of the accuracy and latency data of the three quantile summary approaches.

The discrepancy between observed and predicted frequencies was evaluated using a maximum likelihood criterion. Specifically, for a given model and set of parameter values, the likelihood of the observed response frequencies for a given output position p was determined using the multinomial log-likelihood function:

$$\ln L_{pos}(p) = \sum_{j} \sum_{i} N_{ij} \ln(\pi_{ij}), \tag{C1}$$

Where N_{ij} is the frequency of observations in the bin for the ith category at the jth input position, π_{ij} is the corresponding probability in the bin predicted by the model, and ln is the natural logarithm. The number of bins varied according to the sequence length of the data set being fitted. For four-item sequences, there were 8 bins per output position (32 in total); for five-item sequences there were 10 bins per output position (50 in total); for six-item sequences there were 12 bins per output position (72 in total). A total log-likelihood was then calculated by summing the individual log-likelihoods for each output position:

$$\ln L = \sum_{p=1}^{P} \ln L_{pos}(p), \tag{C2}$$

Where $\ln L$ is the joint multinomial log-likelihood, which was converted to a negative value. Model parameters were varied systematically using the SIMPLEX function minimization algorithm (Nelder & Mead, 1965) until the smallest possible value of this objective function was obtained. Each parameter vector explored by the minimization algorithm involved 5,000 model simulation trials of the sequence length and grouping condition being simulated.

The parameters that were free to vary for the PM model were the weighting (Ω) and distinctiveness (θ) of the position markers. These free parameters were augmented by the amount of response suppression (τ) in the PM + RS model and the amount of output interference (δ) in the PM + OI model. The PG + RS model took as its free parameters the starting point (ϕ) and steepness (ρ) of the primacy gradient, and the degree of response suppression (τ). Finally, the free parameters for the PG + PM + RS model were the weighting (Ω) and distinctiveness (θ) of the position markers, the starting point (ϕ) and steepness (ρ) of the primacy gradient, and the degree of response suppression (τ). The weighting of the position markers and primacy gradient

(ω) was frozen to a value of .5.[4] For the simulations of grouped sequences, the weighting of the two sets of position markers (λ) in the two versions of the PM, PM + RS, PM + OI, and PG + PM + RS models was also set to a fixed value of .5.[5] In addition to the above-mentioned 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 six for the PG + PM + RS model.

The models were initially fit to the data according to the procedure described above, which yielded for each data set 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. 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 and Bayesian information criterion scores (AIC, Akaike, 1973; BIC, Schwartz, 1978, respectively). The AIC was calculated as:

$$\mathrm{AIC}_i = -2 \ln L_i + 2 V_i, \tag{C3}$$

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

[4]We opted to freeze ω because there is some redundancy in this parameter, given the other parameters varied in the fitting. Specifically, the free parameters ϕ and Ω will determine the relative weight given to the primacy gradient and position markers, thus removing the need to also systematically vary ω. An alternative approach—adopted previously (Hurlstone & Hitch, 2015)—would be to fix the latter parameters to values of 1—rendering them inactive—and incorporate ω as a free parameter to govern the relative influence of the primacy gradient and position markers. Varying all three parameters simultaneously improves the quantitative fit of the PG + PM + RS model, but we have not pursued this option here since it is not necessary to bring out the qualitative effects in the data.

[5]We fixed the setting of λ because the main focus of our model comparisons for grouped sequences was to establish whether the data is better understood in terms of a model combining position-of-group with position-within-group markers, or a model combining position-of-group with position-within-sequence markers. The problem with allowing λ to vary is that it opens up the possibility that the models may preferentially engage only one of the two sets of position markers, which spoils this comparison (we consider the impact of varying λ in General Discussion). With λ fixed, the magnitude of grouping effects in the simulations is governed by the setting of the free parameter θ.

seen from equation C3, 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:

$$\mathrm{BIC}_i = -2 \ln L_i + V_i \ln(n), \tag{C4}$$

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 n 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. As the BIC was calculated on quantile-averaged group data, the value of n in the penalty term was set equal to the average number of observations per participant in the data set.

To aid interpretation, 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 degree of support for each model on a continuous measure of evidence. The IC weight for model i was calculated by:

$$w\mathrm{IC}_i = \frac{\exp(-0.5\,\Delta\mathrm{IC}_i)}{\sum_{k=1}^{K} \exp(-0.5\,\Delta\mathrm{IC}_k)}, \tag{C5}$$

Where $\Delta\mathrm{IC}_i$ is the difference in IC between model i relative to the best model, and each $\Delta\mathrm{IC}_k$ 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.

We also conducted model comparisons using likelihood ratio tests (Lamberts, 2004), which compared the improvement in fit of a general model (viz. the PG + PM + RS model) with restricted versions of that model in which one of its components was removed (viz. PM + RS and PG + RS models). The likelihood ratio statistic is calculated by:

$$\chi^2 = -2[\ln L(\text{restricted}) - \ln L(\text{general})], \tag{C6}$$

Where $\ln L(\text{restricted})$ is the log-likelihood of the restricted model and $\ln L(\text{general})$ is the log-likelihood of the general model. The likelihood ratio statistic provides a measure of the reliability of the difference in goodness-of-fit between the restricted and general model. If χ^2 exceeds the critical value corresponding to the 95th percentile of the chi-squared distribution—with degrees of freedom determined by the number of free parameters in the restricted model that were removed from the general model—then the null hypothesis that the restricted model is the best model can be rejected.

Table 1

Estimated multilevel regression parameters for Experiment 1.

Sequence length	Parameter	Estimate	*SE*	*t*	*p*
Four-items	Intercept	1312.35	269.56	4.87	.000
	Direction	-1432.22	279.75	-5.12	.000
	Displacement	-337.75	78.49	-4.30	.000
	Direction × Displacement	390.79	84.64	4.72	.000
	Anticipation	-332.59	76.95	-4.32	.000
	Postponement	50.62	15.65	3.23	.001
Five-items	Intercept	925.01	334.89	2.76	.006
	Direction	-1078.35	350.77	-3.07	.002
	Displacement	-187.71	76.33	-2.46	.014
	Direction × Displacement	223.92	79.47	2.82	.005
	Anticipation	-179.59	76.44	-2.35	.021
	Postponement	36.30	11.56	3.14	.002
Six-items	Intercept	885.35	239.10	3.70	.000
	Direction	-1021.65	274.73	-3.72	.000
	Displacement	-143.96	47.82	-3.01	.003
	Direction × Displacement	176.45	57.82	3.05	.003
	Anticipation	-142.17	47.96	-2.96	.000
	Postponement	33.92	13.88	2.44	.016

Table 2

Estimated multilevel regression parameters for Experiment 2.

Sequence length	Parameter	Estimate	*SE*	*t*	*p*
Four-items	Intercept	2000.52	419.26	4.77	.000
	Direction	-2138.46	424.77	-5.03	.000
	Displacement	-345.20	76.94	-4.49	.000
	Direction × Displacement	389.54	80.61	4.83	.000
	Anticipation	-344.63	77.33	-4.46	.000
	Postponement	46.47	10.17	4.57	.000
Five-items	Intercept	1651.32	774.48	2.13	.034
	Direction	-1791.25	793.50	-2.26	.025
	Displacement	-290.91	155.19	-1.87	.062
	Direction × Displacement	332.25	164.10	2.03	.044
	Anticipation	-287.13	157.37	-1.82	.071
	Postponement	46.02	16.19	2.84	.005
Six-items	Intercept	1467.57	501.96	2.92	.004
	Direction	-1585.63	512.84	-3.09	.002
	Displacement	-260.66	100.95	-2.58	.010
	Direction × Displacement	277.31	100.95	-2.58	.010
	Anticipation	-246.60	96.67	-2.55	.012
	Postponement	15.97	7.28	2.19	.030

Table 3

Estimated multilevel regression parameters for Experiment 3.

Grouping	Parameter	Estimate	*SE*	*t*	*p*
Ungrouped	Intercept	1275.73	412.22	3.09	.002
	Direction	-1429.77	425.44	-3.36	.000
	Displacement	-219.59	81.01	-2.71	.007
	Direction × Displacement	246.38	85.24	2.89	.004
	Anticipation	-204.13	79.26	-2.58	.011
	Postponement	26.79	7.39	3.62	.000
Grouped	Intercept	1188.44	331.87	3.58	.000
	Direction	-1331.75	346.89	-3.84	.000
	Displacement	-201.34	63.77	-3.16	.002
	Direction × Displacement	332.25	164.10	2.03	.044
	Anticipation	-186.92	61.97	-3.02	.003
	Postponement	31.55	9.40	3.36	.001

Table 4

Parameter estimates for the fits of the models to the different sequence length conditions of Experiment 2.

Sequence length	Model	Ω	θ	ϕ	ρ	ω^*	τ	δ	S
Four-items	PM	0.48	0.65	-	-	-	-	-	70.92
	PM + RS	0.91	0.80	-	-	-	0.91	-	107.77
	PM + OI	0.57	0.66	-	-	-	-	0.44	71.08
	PG + RS	-	-	0.43	0.87	-	0.97	-	85.48
	PG + PM + RS	0.43	0.56	0.60	0.91	0.5	0.91	-	77.50
Five-items	PM	0.47	0.66	-	-	-	-	-	91.62
	PM + RS	0.82	0.82	-	-	-	1	-	71.31
	PM + OI	0.48	0.64	-	-	-	-	0.39	82.62
	PG + RS	-	-	0.49	0.93	-	0.87	-	77.06
	PG + PM + RS	0.42	0.61	0.55	0.93	0.5	0.96	-	81.23
Six-items	PM	0.40	0.62	-	-	-	-	-	80.67
	PM + RS	0.98	0.84	-	-	-	0.98	-	83.72
	PM + OI	0.44	0.61	-	-	-	-	0.53	82.16
	PG + RS	-	-	0.48	0.95	-	0.92	-	106.03
	PG + PM + RS	0.23	0.45	0.91	0.96	0.5	0.87	-	92.70

Note—Ω = weighting of position markers; θ = distinctiveness of position markers; ϕ = initial value of primacy gradient; ρ = steepness of primacy gradient; ω = weighting of primacy gradient and position markers; τ = degree of response suppression; δ = amount of output interference; S = iteration-to-ms scaling; * = fixed parameter; - = parameter not associated with the model indicated by the row.

Table 5

AIC and BIC weights with associated goodness-of-fit quantities for the fits of the models to the different sequence length conditions of Experiment 2.

Sequence length	Model	k	ln L	AIC	ΔAIC	wAIC	BIC	ΔBIC	wBIC
Four-items	PM	3	-24419	48845	14406	0	48854	14397	0
	PM + RS	4	-18576	37160	2721	0	37171	2714	0
	PM + OI	4	-24292	48593	14154	0	48604	14147	0
	PG + RS	4	-17735	35479	1040	0	35490	1033	0
	PG + PM + RS	**6**	**-17214**	**34439**	**0**	**1**	**34457**	**0**	**1**
Five-items	PM	3	-35356	70718	19677	0	70728	19668	0
	PM + RS	4	-28318	56643	5602	0	56656	5596	0
	PM + OI	4	-35125	70259	19218	0	70272	19212	0
	PG + RS	4	-26217	52443	1402	0	52455	1395	0
	PG + PM + RS	**6**	**-25515**	**51041**	**0**	**1**	**51060**	**0**	**1**
Six-items	PM	3	-46034	92075	25096	0	92085	25086	0
	PM + RS	4	-37707	75421	8442	0	75435	8436	0
	PM + OI	4	-46001	92009	25030	0	92023	25024	0
	PG + RS	4	-34661	69329	2350	0	69343	2344	0
	PG + PM + RS	**6**	**-33484**	**66979**	**0**	**1**	**66999**	**0**	**1**

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; BIC = Bayesian information criterion; ΔBIC = difference in BIC with respect to the best fitting model; wBIC = BIC weight. The bold items indicate the best fitting model for each sequence length.

Table 6

Parameter estimates for the fits of the models to the grouped condition of Experiment 3.

Model	Ω	θ	λ^*	ϕ	ρ	ω^*	τ	δ	S
PM^{pwg}	0.33	0.53	0.5	-	-	-	-	-	97.49
PM^{pwg} + RS	0.20	0.57	0.5	-	-	-	1	-	111.07
PM^{pwg} + OI	0.36	0.48	0.5	-	-	-	-	0.32	121.99
PG + PM^{pwg} + RS	0.04	0.45	0.5	0.69	0.95	0.5	1	-	113.63
PM^{pws}	0.33	0.68	0.5	-	-	-	-	-	100.12
PM^{pws} + RS	0.31	0.67	0.5	-	-	-	0.99	-	100.87
PM^{pws} + OI	0.36	0.65	0.5	-	-	-	-	0.52	97.99
PG + PM^{pws} + RS	0.03	0.61	0.5	0.75	0.96	0.5	0.75	-	121.57

Note—Ω = weighting of position markers; θ = distinctiveness of position markers; λ = weighting of the two sets of position markers; ϕ = initial value of primacy gradient; ρ = steepness of primacy gradient; ω = weighting of primacy gradient and position markers; τ = degree of response suppression; δ = amount of output interference; S = iteration-to-ms scaling; * = fixed parameter; - = parameter not associated with the model indicated by the row.

Table 7

AIC and BIC weights with associated goodness-of-fit quantities for the fits of the models to the grouped condition of Experiment 3.

Model	k	$\ln L$	AIC	ΔAIC	wAIC	BIC	ΔBIC	wBIC
PM^{pwg}	3	-29427	58861	16071	0	58870	16061	0
PM^{pwg} + RS	4	-23936	47880	5090	0	47892	5083	0
PM^{pwg} + OI	4	-29440	58889	16099	0	58901	16092	0
PG + PM^{pwg} + RS	**6**	**-21389**	**42790**	**0**	**1**	**42809**	**0**	**1**
PM^{pws}	3	-29576	59157	16367	0	59167	16358	0
PM^{pws} + RS	4	-24022	48051	5261	0	48064	5255	0
PM^{pws} + OI	4	-29570	59147	16357	0	59160	16351	0
PG + PM^{pws} + RS	6	-21654	43321	531	0	43339	530	0

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; BIC = Bayesian information criterion; ΔBIC = difference in BIC with respect to the best fitting model; wBIC = BIC weight. The bold item indicates the best fitting model.

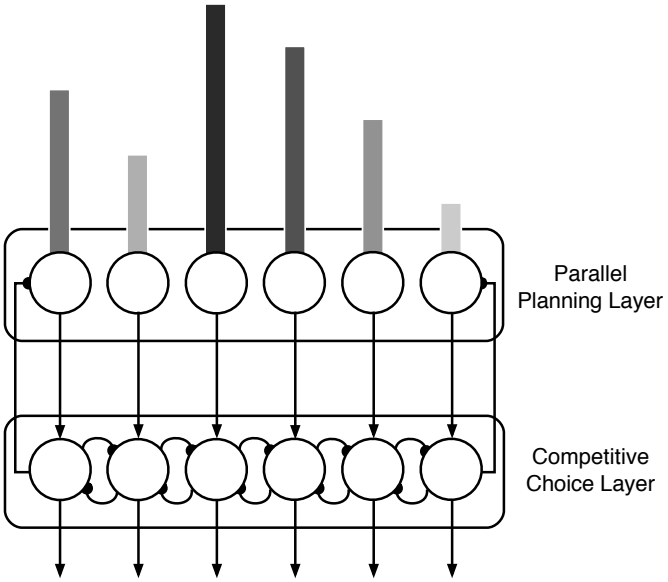

Figure 1. Schematic of a two-layer competitive queuing sequence planning and control mechanism comprising a parallel planning layer (upper field of nodes) and a competitive choice layer (lower field of nodes). Lines terminating with arrows represent excitatory connections, whereas lines terminating with semi-circles 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 main text for further details.

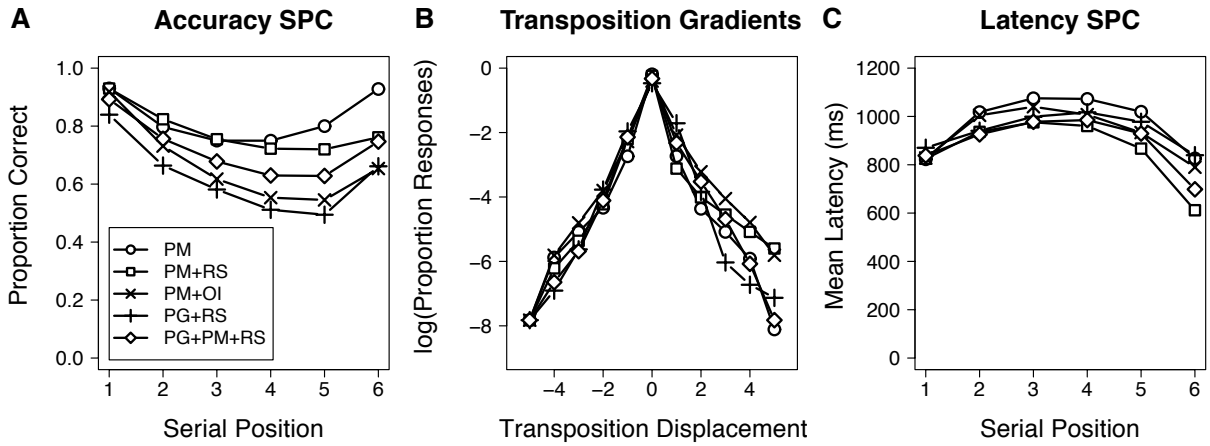

Figure 2. Predicted accuracy serial position curves (A), transposition gradients (B), and latency serial position curves (C) of five models of serial order. PM = position marking; RS = response suppression; OI = output interference; PG = primacy gradient.

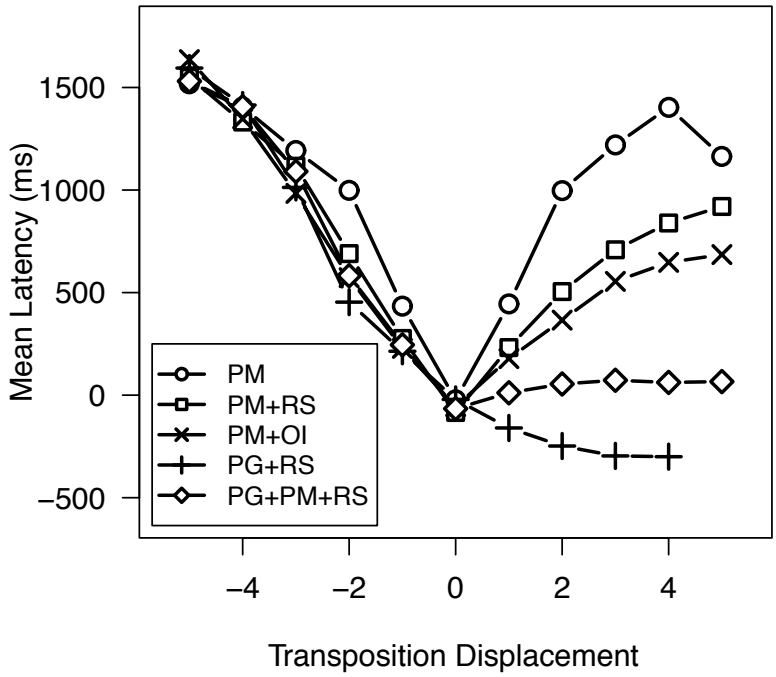

Figure 3. Predicted latency–displacement functions of five models of serial order. PM = position marking; RS = response suppression; OI = output interference; PG = primacy gradient.

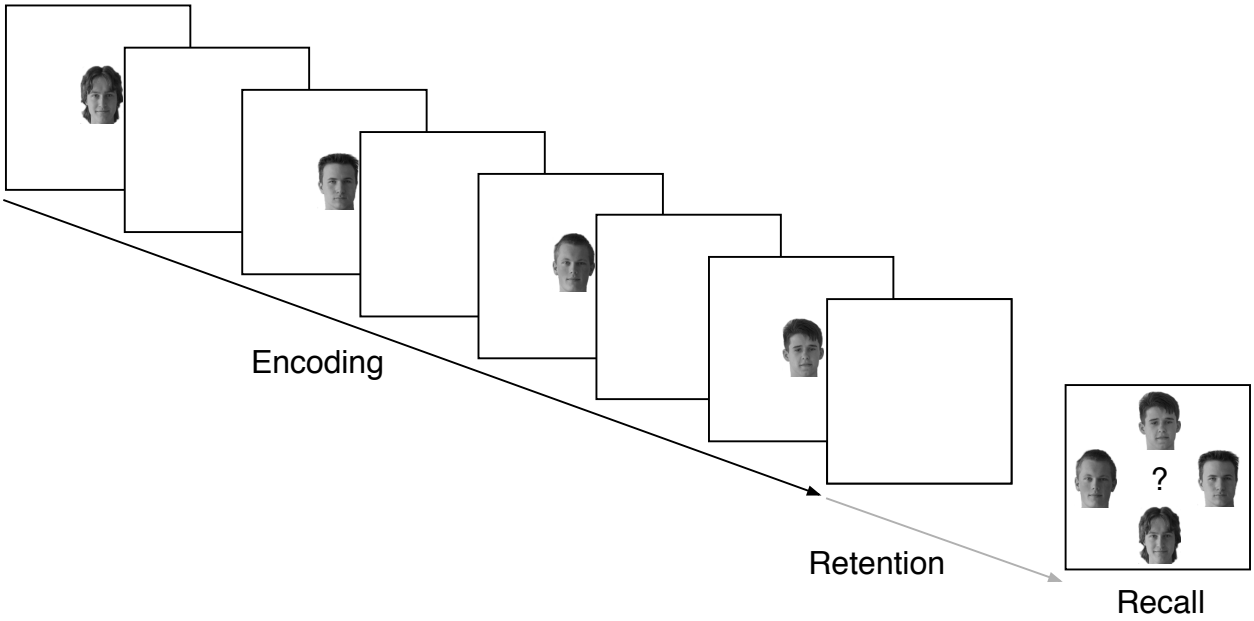

Figure 4. Schematic of the time course of events on each trial.

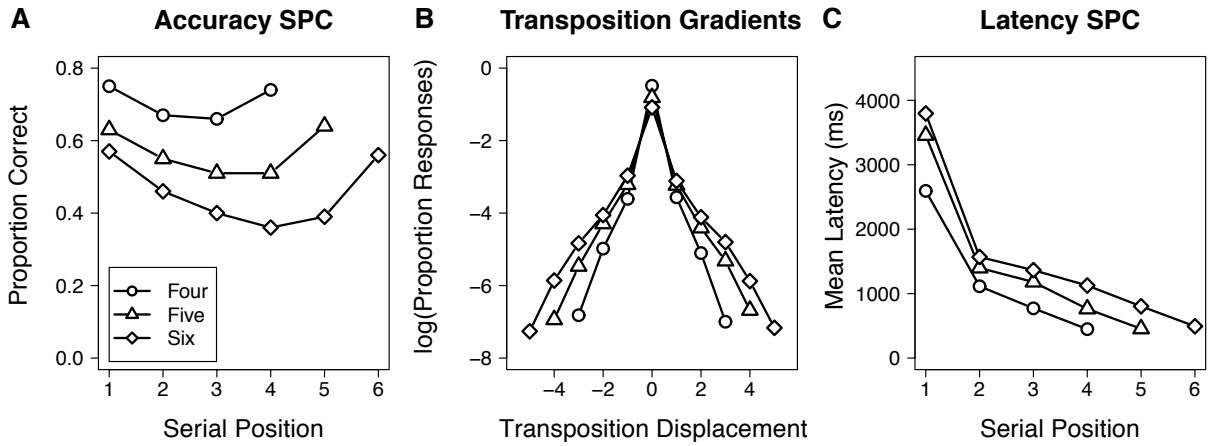

Figure 5. Accuracy serial position curves (A), transposition gradients (B), and latency serial position curves (C) for Experiment 1.

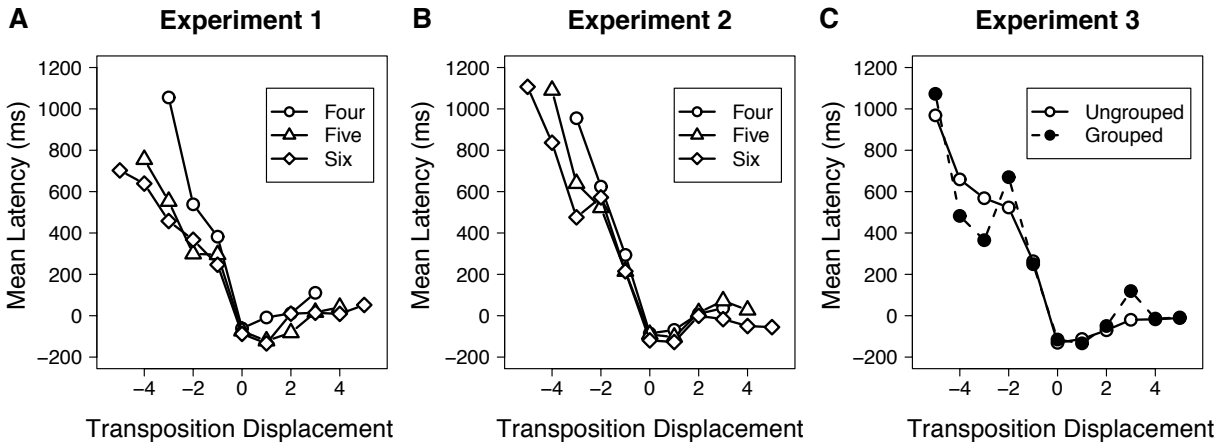

Figure 6. Latency–displacement functions for Experiment 1 (A), Experiment 2 (B), and Experiment 3 (C).

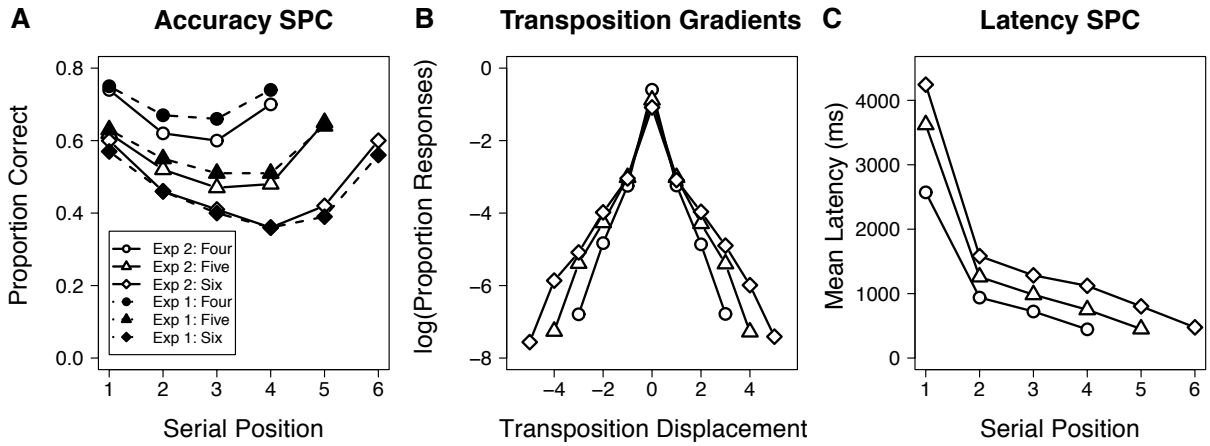

Figure 7. Accuracy serial position curves (A), transposition gradients (B), and latency serial position curves (C) for Experiment 2.

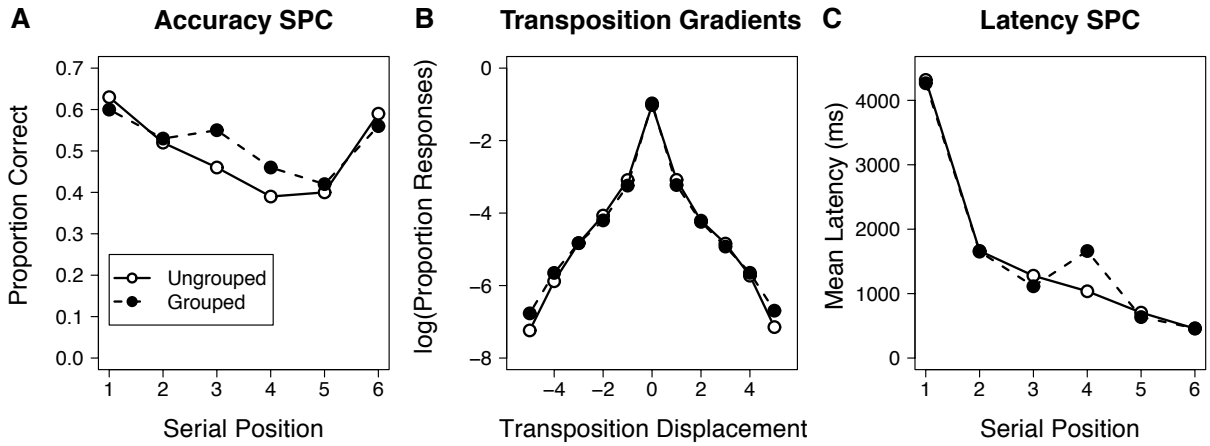

Figure 8. Accuracy serial position curves (A), transposition gradients (B), and latency serial position curves (C) for Experiment 3.

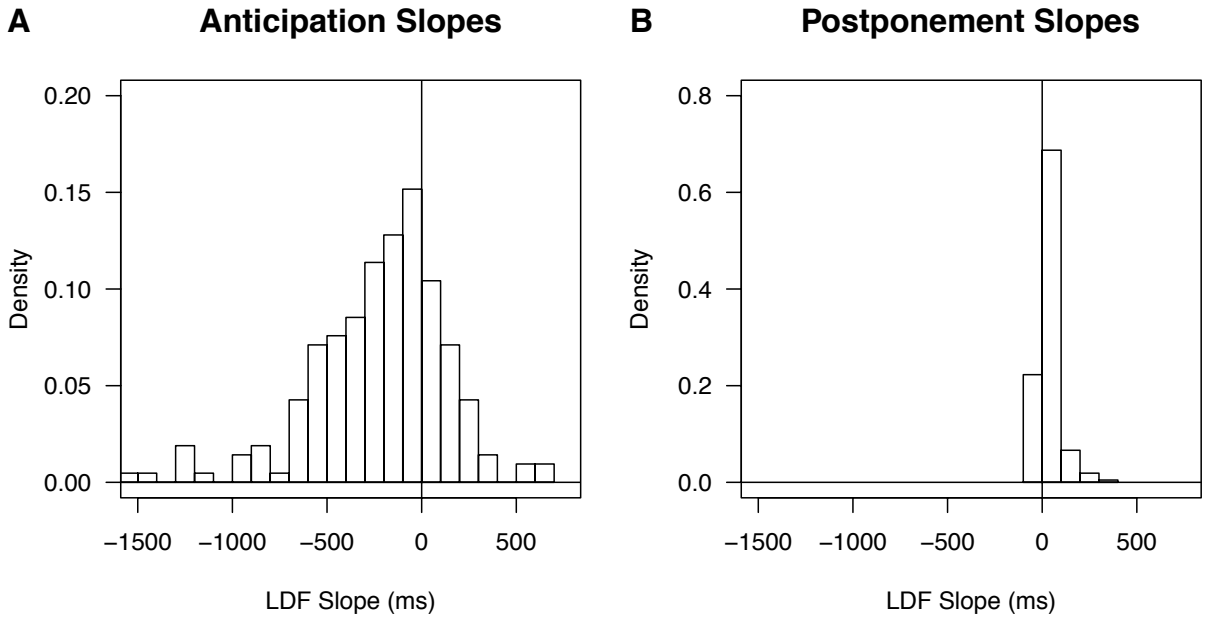

Figure 9. Distributions of slopes of latency–displacement functions (LDF) for anticipations (A) and postponements (B), across individual participants for all three experiments and conditions. The vertical line in each panel corresponds to a slope value of 0.

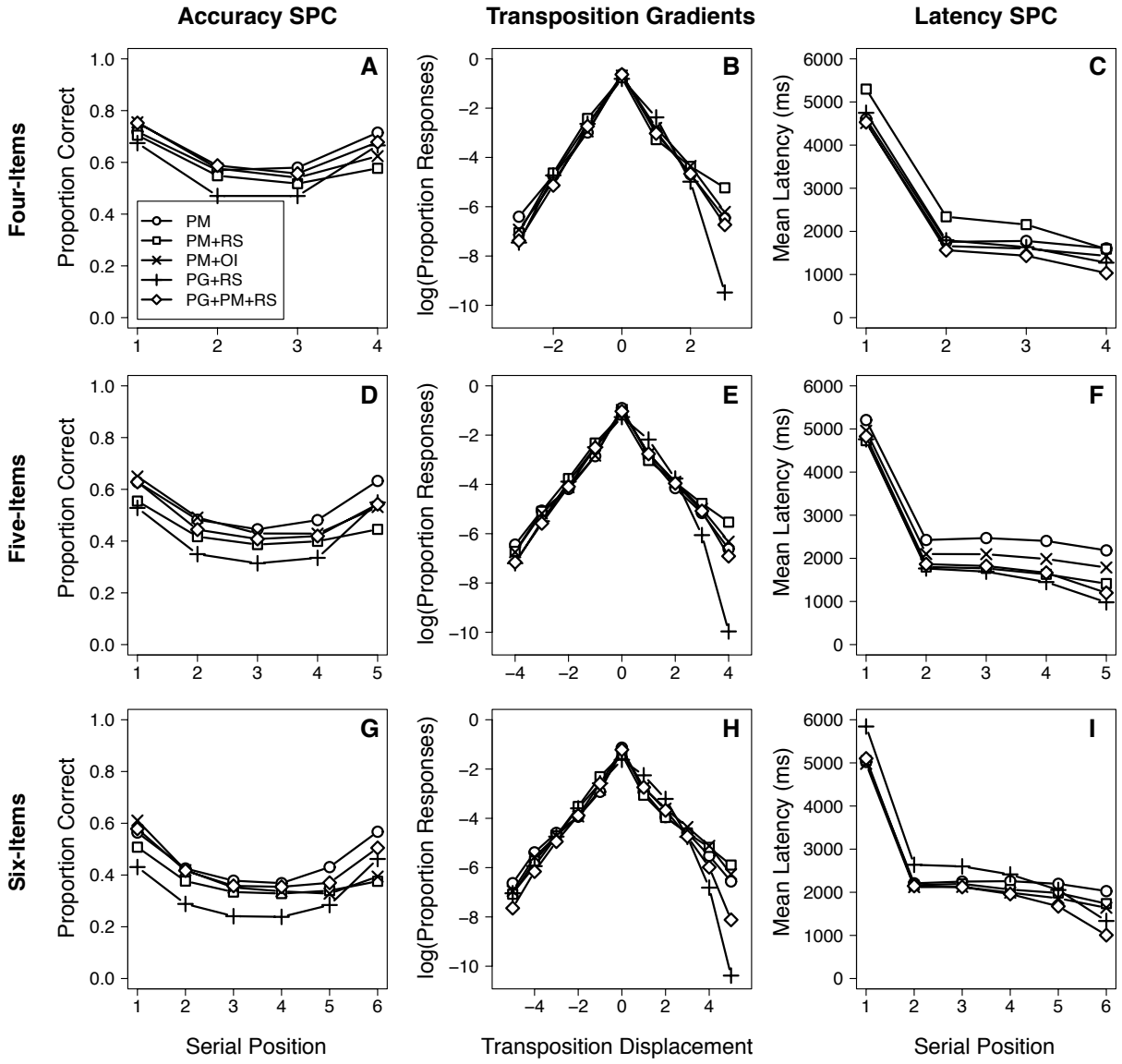

Figure 10. Fits of five models of serial order to the different sequence length conditions of Experiment 2. Panels show accuracy serial position curves (A, D, & G), transposition gradients (B, E, & H), and latency serial position curves (C, F, & I). The upper panels show predictions for four-item sequences; middle panels show predictions for five-item sequences; whilst lower panels show predictions for six-item sequences. PM = position marking; RS = response suppression; OI = output interference; PG = primacy gradient.

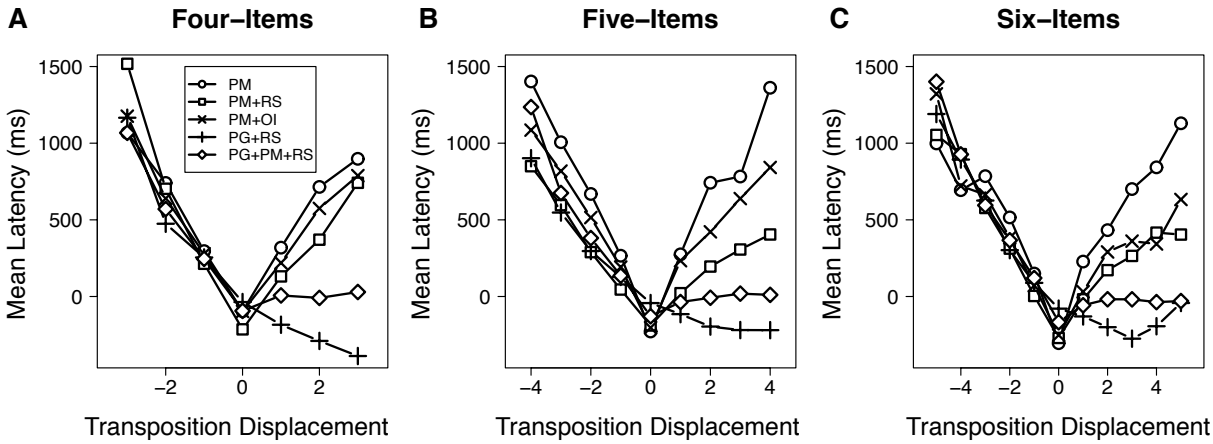

Figure 11. Fits of the five models to the latency–displacement functions of Experiment 2. Panels show predictions for four-item sequences (A), five-item sequences (B), and six-item sequences (C). PM = position marking; RS = response suppression; OI = output interference; PG = primacy gradient.

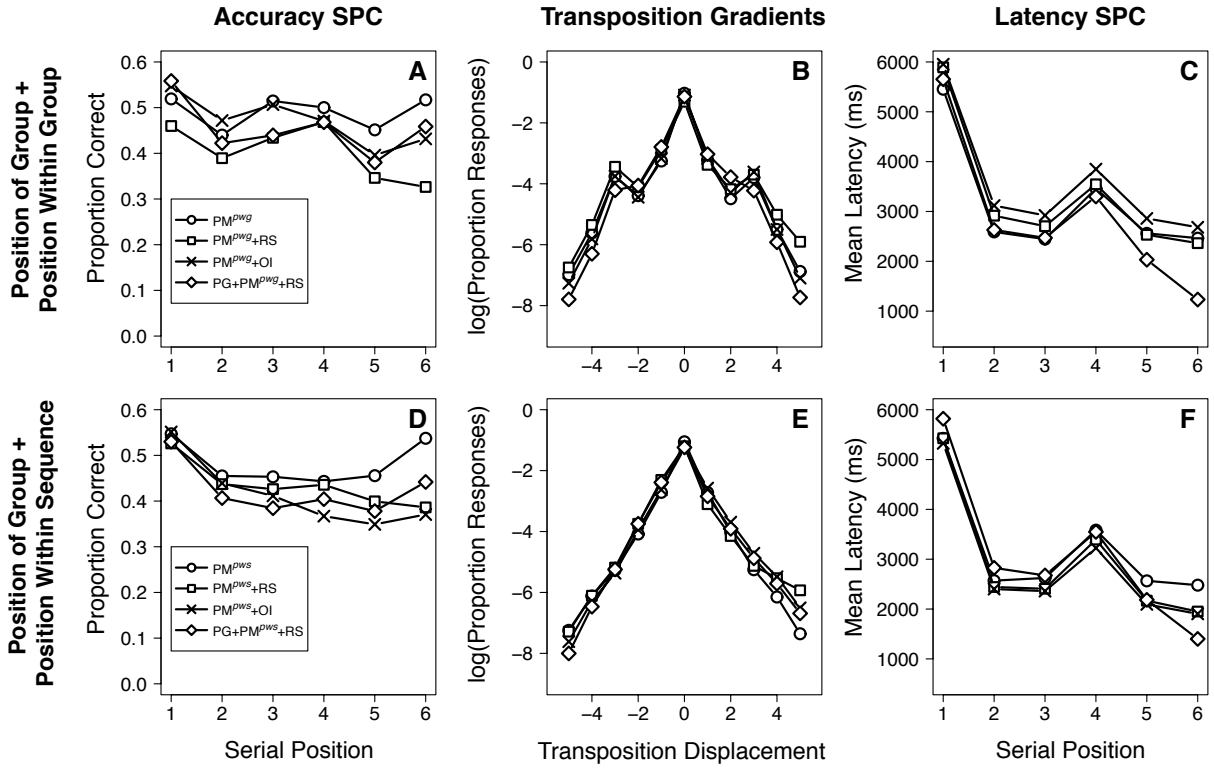

Figure 12. Fits of eight models of serial order to the grouped condition of Experiment 3. Panels show accuracy serial position curves (A & D), transposition gradients (B & E), and latency serial position curves (C & F). The upper panels show the predictions for the models pairing position of group with position within-group markers; the lower panels show the predictions of the model pairing position of group with position within-sequence markers. PM^{pwg} = position marking of groups and position within-groups; PM^{pws} = position marking of groups and position within-sequence; RS = response suppression; OI = output interference; PG = primacy gradient.

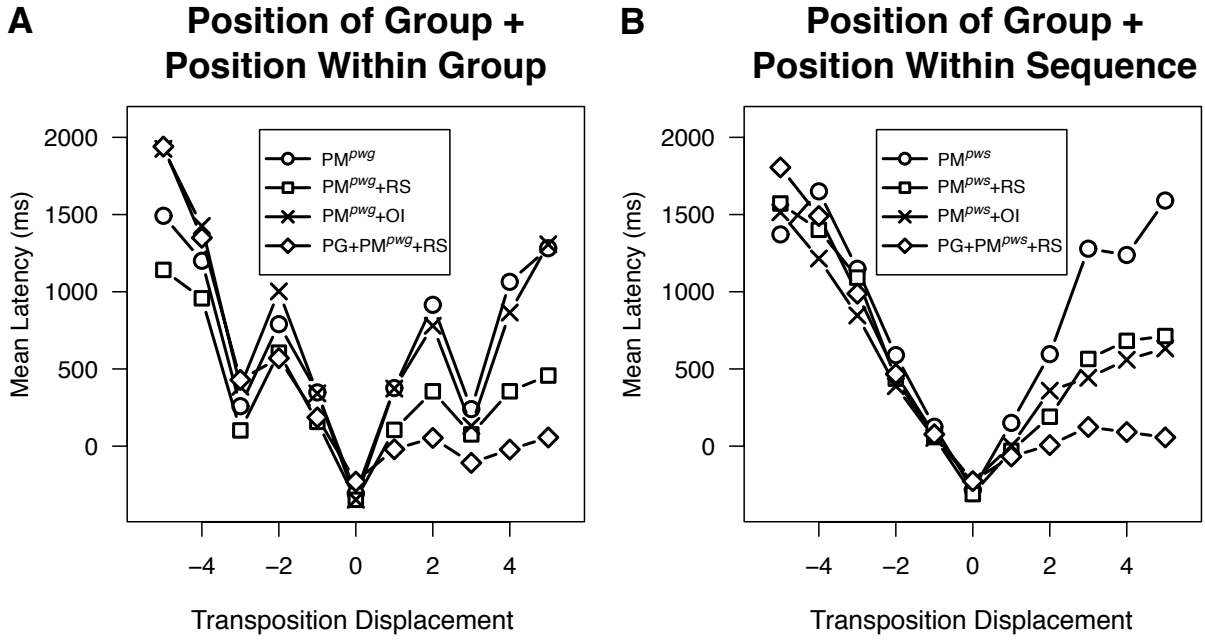

Figure 13. Fits of the eight models to the latency–displacement functions of the grouped condition of Experiment 3. Panels show predictions for the models pairing position of group with position within-group markers (A), and the models pairing position of group with position within-sequence markers (B). PM^{pwg} = position marking of groups and position within-groups; PM^{pws} = position marking of groups and position within-sequence; RS = response suppression; OI = output interference; PG = primacy gradient.

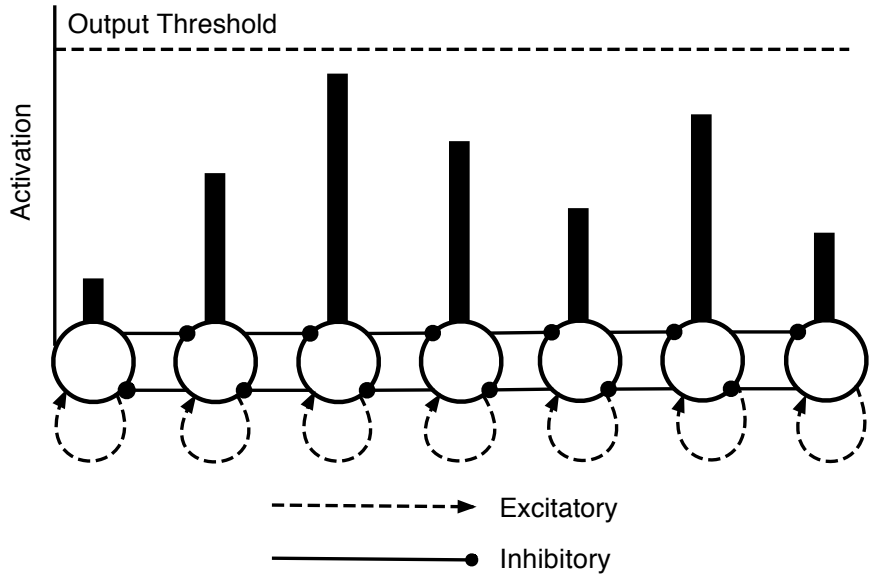

Figure 14. A schematic of the lateral inhibition neural network architecture employed for the simulations. The network is modeled on the competitive choice layer employed in competitive queuing models of serial behavior (Bullock, 2004; Rhodes & Bullock, 2003). Each localist item node possesses a single 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. Note—the number of nodes in the network is dependent on the sequence length being modeled.

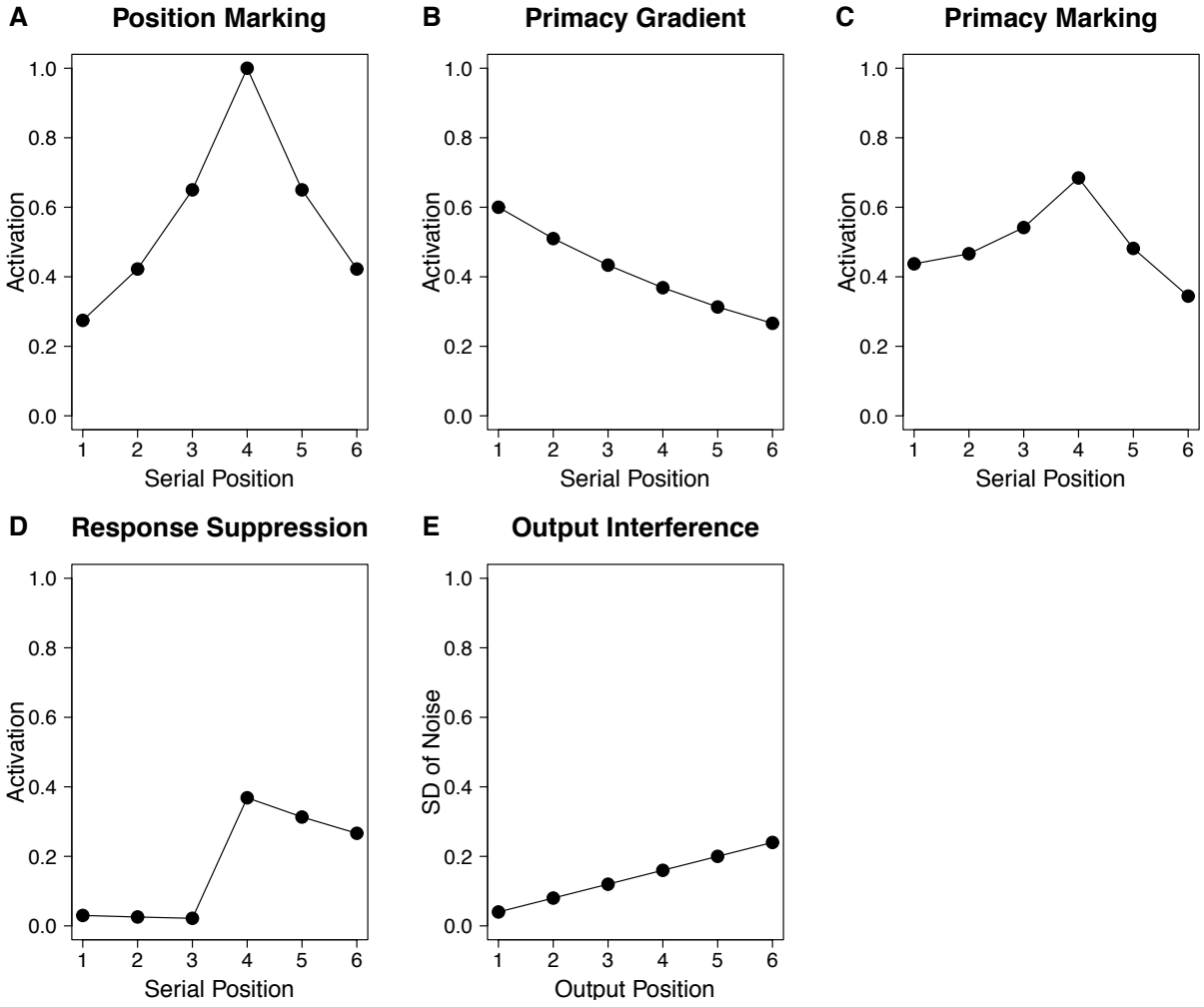

Figure 15. Example starting activations for the four representational principles—and combination of principles—for six-item sequences, based on the parameter settings employed for the initial simulations: (A) position marking (showing activations for the fourth output position), (B) a primacy gradient (showing activations for the first output position), (C) 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).

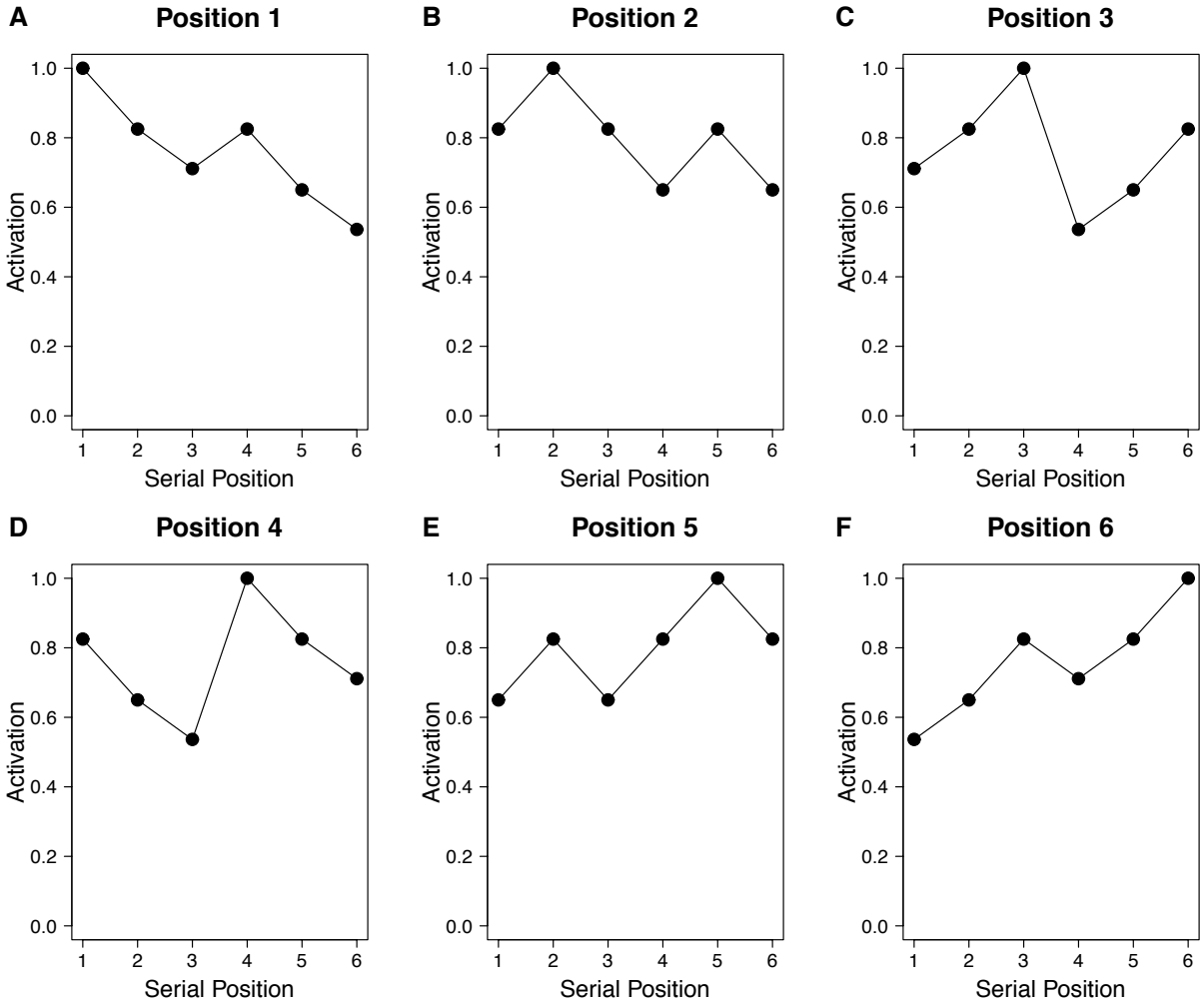

Figure 16. Example starting activations for the position-of-group and position-within-group implementation of position marking for a six-item sequence grouped into threes. Activations were generated using the following parameter values: $\Omega = 1$; $\theta = .65$; $\lambda = .5$.

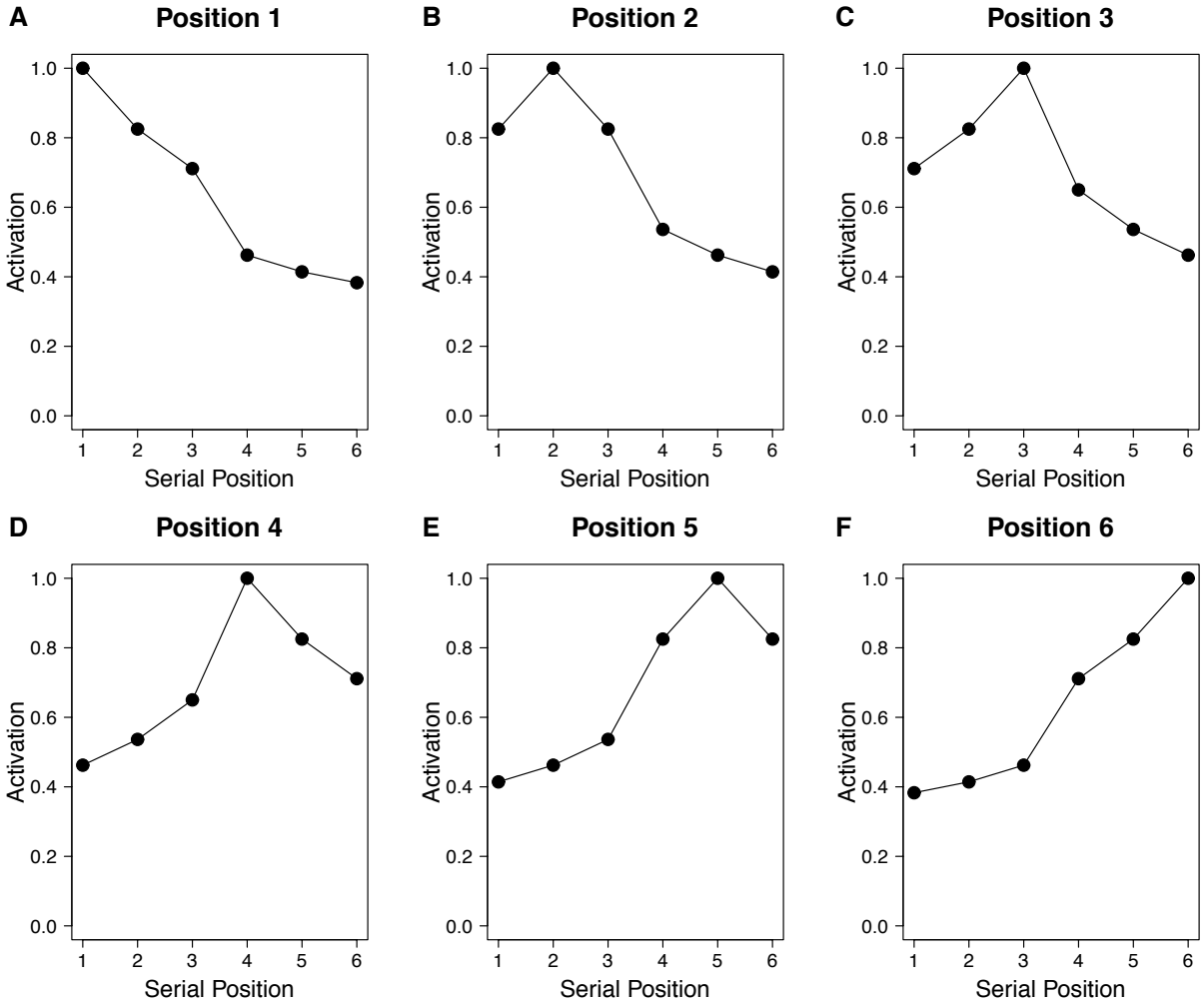

Figure 17. Example starting activations for the position-of-group and position-within-sequence implementation of position marking for a six-item sequence grouped into threes. Activations were generated using the following parameter values: $\Omega = 1$; $\theta = .65$; $\lambda = .5$.