

# Imitating Unfamiliar Sequences of Connected Linear Motions

Yigal Agam<sup>1</sup>, Daniel Bullock<sup>2</sup> and Robert Sekuler<sup>1,2</sup>

<sup>1</sup>Volen Center for Complex Systems, Brandeis University

<sup>2</sup>Cognitive and Neural Systems, Boston University

## ABSTRACT

A fundamental challenge in neuroscience is to understand the mechanisms by which multi-component actions are represented and sequenced for production. We addressed this question with a movement imitation task in which subjects viewed the quasi-random, 2D movements of a disc and then used a stylus to reproduce the remembered trajectory. The stimulus disc moved along straight segments, which differed sufficiently from one another that it was possible to trace individual segments' fate in the resulting movement imitation. A biologically-based segmentation algorithm decomposed each imitation into segments whose directions could be compared to those of homologous segments in the model. As the number of linked segments in a stimulus model grew from three to seven, imitation became less accurate, with segments more likely to be deleted, particularly from a model's final stages. When fidelity of imitation was assessed segment by segment, the resulting serial position curves showed a strong primacy effect and a moderate recency effect. Analysis of pairwise transposition errors revealed a striking preponderance of exchanges between adjacent segments, which, along with the serial position effects, supports a competitive queuing model of sequencing. In analogy to results with verbal serial recall, repetition of one directed segment in the model reduced imitation quality. Results with longer stimulus models suggest that the segment-by-segment imitation generator may be supplemented in the final stages of imitation by an error-signal driven overlay that produces a late-course, real-time correction. Results are related to neural mechanisms that are known to support sequential motor behavior and working memory.

## INTRODUCTION

Unlike members of many other species, humans are able to observe and later reproduce the actions of con-specifics, an ability that is critical for learning and perfecting a wide range of motor and social behaviors (e.g., Heyes, 2001; Bandura, Ross, & Ross, 1961; Wolpert, Doya, & Kawato, 2003). Imitation is a complex process, recruiting diverse functions such as visual processing, working memory, motor

control, sequential organization and learning. Even though these functions individually rank among neuroscience's most intensely studied areas, their theoretical integration has been difficult (Schaal, 1999). One source of this difficulty is the paucity of useful, theory-related data on human imitation.

Although many imitated behaviors are complex and contain multiple components that must be performed in proper sequence, studies of imitation have tended to limit themselves to very simple actions, such as the flexion of a single finger, or well-practiced stereotyped movements, such as grasping and moving an object from one location to another pre-defined, fixed location (but see Matarić & Pomplun, 1998). Moreover, such studies typically assess imitation quality on a "pass-fail" scale, whose binary values correspond to extremes of imitation present or absent (e.g. Iacoboni et al., 1999; Makuuchi, Kaminaga, & Sugishita, 2005). The all-or-none character of this approach is poorly suited to assessing subtle changes in performance such as those that come with learning (Schaal, 1999).

"Mirror neurons", reported first in area F5 of the monkey brain (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996) and subsequently postulated in Broca's area (Iacoboni et al., 1999; Buccino et al., 2004), the human brain's putative F5 homologue, may be implicated in action understanding and in movement imitation (Rizzolatti & Craighero, 2004). But, understanding how events are encoded and represented for subsequent imitation must take account of how separable components are organized and sequenced. Competitive queuing (CQ) models provide an attractive computational architecture for sequencing (Bullock & Rhodes, 2003; Bullock, 2004; Grossberg, 1978; Houghton, 1990; Rhodes, Bullock, Verwey, Averbeck, & Page, 2004), and recent neurophysiological results from awake behaving monkeys have strongly supported this approach (Averbeck, Chafee, Crowe, & Georgopoulos, 2002, 2003; Averbeck, Crowe, Chafee, & Georgopoulos, 2003). In these studies of prefrontal area 46 in monkeys, small neural ensembles corresponding to as many as five forthcoming drawing strokes were simultaneously active well before sequential movement initiation. The relative activation levels among these ensembles also specified the relative order of the corresponding strokes in the forthcoming sequences: Higher preparatory activation reliably predicted earlier execution of the corresponding stroke when the sequence was eventually performed. Such a preparatory representation is fundamental in CQ models, but is not predicted by classical or recurrent-state chaining models (Page & Norris, 1998; Henson, Norris, Page, & Baddeley, 1996).

---

Supported by AFOSR grant F49620-03-1-0376, NSF grant SBE-0354378, and NIH grant R01MH068404. Address correspondence to sekuler@brandeis.edu

Even though the studies by Averbeck and colleagues on monkeys are highly suggestive for human imitation, those studies were limited to well-learned performance of just a few highly-distinctive models. To better understand sequencing of component actions during human imitation of novel sequences, we carried out two experiments on human delayed imitation of simple, multi-component models. The models for imitation were abstract movement sequences of graded complexity. Such two-dimensional models afford several advantages for studying sequencing in imitation, such as minimal intrusion from semantic memory and little influence from subjects' previous, domain-specific experience. By using a new model on each trial, we avoided contamination from model-specific learning that results from even a few encounters with the same model (Sekuler, Siddiqui, Goyal, & Rajan, 2003). Moreover, the models and task allow us to examine theoretically important relationships among errors in imitating various components of a model. As will become clear later, this analysis is essential for evaluating CQ accounts of sequencing. So, using a biologically-based, automated algorithm for assessing imitation quality segment by segment, we examined not only the overall performance for models of varying complexity, but also the way in which magnitude and types of errors depended upon components' serial position. One aim was to mine the details of behavioral performance for clues to the representations that subjects used when reproducing the just seen models.

## METHODS

### *Overview*

To study imitation learning, Sekuler et al. (2003) synthesized abstract stimulus models of graded complexity. Their stimuli were 2-D motion patterns, generated by a disc that moved along an irregular, unpredictable trajectory comprising variable numbers of segments. As only the disc's momentary position was visible on the display screen, viewers had to construct a mental representation of the trajectory, using the mind's eye to knit together those momentary positions. This product had to be held in working memory until the entire trajectory was to be reproduced with a stylus on a graphic tablet. Unfortunately, the analytic methods of that study were insensitive to an essential feature of complex sequential events and actions, namely the serial ordering of their components (Lashley, 1951; Zacks, Tversky, & Iyer, 1999; Rhodes et al., 2004). We therefore modified the methods of Sekuler et al. (2003) in order to facilitate the detection and analysis of effects associated with individual components' serial order.

In order to link visuomotor imitation to the wealth of parametric observations and models of serial order behavior in various sequence learning and performance tasks (e.g., Rhodes et al., 2004), we wanted to evaluate how quality of imitation might vary with the serial position that a segment occupied in the model. We also wanted to examine ways in which this relationship was influenced by model complexity,

which was operationalized as the number of segments in a model. Finally, we wanted a way to examine the relationship between the fidelity with which successive components were imitated. Alternative views of serial recall from working memory predict different relationships between the quality of recall of successive items. Classical chaining models, for example, predict a "collapse" of the entire sequence after a single error is made (Henson et al., 1996; Henson, 2001). CQ models, in contrast, assert that the recall process can be robust in the face of an error on a single component, and that the most common type of error would be a temporally local one, that is a transposition between two adjacent segments.

To serve our analytic goals, we devised an algorithm that automatically decomposed subjects' reproductions into component segments, which made it possible to analyze imitation quality on a segment-by-segment level. This algorithm was grounded in current understanding of the way in which human observers normally segment complex actions. Behavioral and brain imaging studies show that observers rely on spatio-temporal discontinuity as the principal evidence that one component of an event has ended and another has begun (Newton, Engquist, & Bois, 1977; Zacks et al., 1999, 2001; Zacks & Tversky, 2001). Therefore, we built our segmentation algorithm on a foundation of spatio-temporal discontinuities in the subjects' performance. The segmentation algorithm is a simplified version of ones used in previous research (for example, Abend, Bizzi, & Morasso, 1982; Lewis et al., 2003).

### *Subjects*

A total of 21 human subjects (ages 18-34), with normal visual acuity and no history of visual or motor disorder, took part in two different experiments. Thirteen participated in the first experiment and eight in the second experiment. Fifteen of the subjects were right-handed. All subjects gave informed consent, and all procedures were approved by the Institutional Review Boards of Brandeis University and Boston University.

### *Stimulus Trajectory Generation and Presentation*

Motion pattern stimuli were generated by the steady movement of a yellow disc (0.57 deg visual angle in diameter) against a black background on a computer screen, which subjects viewed from a distance of 75 cm. The disc moved along a series of linked, straight segments, each 1.15 deg visual angle long (Figure 1A). The display was refreshed at 60 Hz. At the start of a trial, the yellow disc appeared at the center of the screen for 750 ms, and then began moving at a constant speed of 4.57 deg visual angle/s, pausing for 150 ms at the end of each segment. This brief pause was meant to minimize illusory, perceptual smoothing of the disc's path (Brown & Voth, 1937). The disc remained on the display screen for 400 ms after movement was over, then disappeared. A formal description of our stimuli will facilitate understanding of the rationale behind this design. The model trajectory, *S*, traced out by the moving disc can be described as

$$S = \{s_1, \dots, s_n\}$$

where  $s_i$  represents the segment in a model comprising  $n$  segments. Trajectory complexity, defined by the number of segments in a model, varied randomly from trial to trial, with  $3 \leq n \leq 7$  (20 trials of each). Regardless of  $n$ , model segments were all the same length and duration. Stimulus models were constructed iteratively by generating random directions of movement for each candidate segment, and then verifying that its specifications fit desired constraints. The direction of the  $i^{\text{th}}$  segment is given by

$$\theta_i = \theta_{i-1} + \Delta\theta$$

where  $\theta_i$  is the direction of the  $i^{\text{th}}$  segment. In the present experiments, the movement direction,  $\theta$ , of each new segment was restricted to multiples of  $45^\circ$ .  $\Delta\theta$  was a quasi random variable with a minimum value of  $45^\circ$ . The disc was not allowed to retrace the preceding segment or to move in the same direction as the previous segment. These constraints were designed to facilitate subsequent analysis of segment-by-segment errors in reproduction.

Pilot studies (see Sekuler et al., 2003) showed that closed trajectories produced anomalous imitations, probably by allowing subjects to apply and use verbal labels for the models. A comparable result was reported recently for reproductions of felt objects (Henriques, Flanders, & Soechting, 2004). To eliminate verbal encoding of the stimuli, models were also constrained so that segments could not intersect, cross one another, or even approach one another more closely than one half the length of a segment. The generation algorithm also ensured that each stimulus model contained movements in at least one of the four cardinal and in at least one of four diagonal directions. Finally, the change in direction between two consecutive segments,  $\Delta\theta$ , could not be constant along the entire stimulus pattern. These constraints prevented the generation of motion trajectories whose shapes would be easy to label verbally, e.g., regular polygons.

### Reproduction of Disc Trajectory

Subjects reproduced the movement of the disc by drawing with a stylus on a  $15 \times 20$  cm graphic tablet positioned directly in front of a subject's preferred hand. One second after each model presentation, a brief beep sounded and a small blue disc appeared on the screen, at the same point from which the movement of the yellow stimulus disc had begun. The blue disc's location on the screen was yoked to the position of the stylus, moving along with the stylus. Furthermore, the moving disc left a thin blue, visible trace of its movement path, giving subjects a real-time record of their movements (Figure 1B).

The recording of an imitation began when the stylus first touched the graphic tablet, and ceased when the stylus was lifted from the tablet. When the subject's reproduction was

complete, the entire  $n$ -segment trajectory of the stimulus model was displayed in yellow. Subjects could then see the trace of their attempted imitation superimposed on the trajectory of the stimulus model (Figure 1C). Such outcome feedback was meant to provide subjects with trialwise information about the quality of their performance.

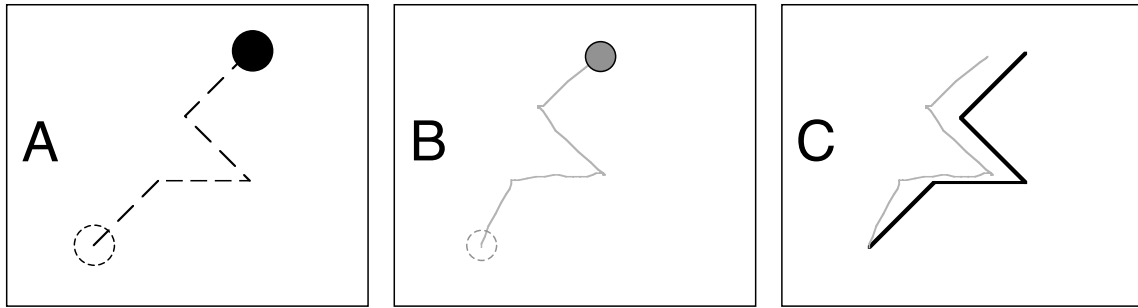
### Recording of Stylus Movement

The stylus' position on the graphic tablet was sampled at 50 Hz. To scale the tablet's drawing area to the larger area of the computer display, the graphic tablet's software driver was configured so that one cm of stylus movement on the tablet produced 1.4 cm of movement on the computer display. After the imitation had ended and the stylus was lifted from the graphic tablet, the captured array of stylus coordinates and associated timebase information were saved to computer disk for later analysis.

### Segment Identification

One of the greatest challenges in studying movement imitation is the development of sound methods for characterizing the quality of an imitation. As Schaal, Ijspeert, and Billard (2003) noted, the quality of any movement imitation can be expressed in a number of alternative metrics. When multi-component models are being imitated, part of the challenge is to identify the breakpoints between components in the imitation. An algorithm we used to segment each imitation into its component parts was based on behavioral and functional imaging studies of human observers' segmentation of complex everyday actions (e.g. Newton et al., 1977; Zacks et al., 2001). Such studies show that perceptual segmentation depends mainly on marking spatio-temporal discontinuities (see also, Abend et al., 1982; Lewis et al., 2003). To identify segment breakpoints in subjects' imitations, our algorithm scanned sampled imitations for points at which the stylus remained motionless for more than 40 ms. From these potential segment breakpoints, the algorithm identified points at which the direction of stylus movement shifted instantaneously by at least  $15^\circ$ . In addition, points where a shift of  $30^\circ$  or more occurred (without regard to time) were also selected as potential breakpoints. When this selection process left several candidate points clustered together less than 40% of model segment length apart, we designated the cluster's midpoint as the breakpoint. The segmentation algorithm yielded a set of coordinates corresponding to the reproduction's presumed segment breakpoints. We discarded trials that yielded fewer than two, or more than ten segments (about 4%), as they likely resulted from algorithm errors or technical glitches during data acquisition.

In scoring each trial's imitation for its similarity to the corresponding model, we were not interested in subjects' drawing ability *per se*. Therefore, the algorithm minimized errors that resulted exclusively from motor artifacts, such as high-frequency "wiggles", by breaking the reproduced path into discrete segments and treating them as straight lines (see Figure 2). This approach is consistent with



*Figure 1.* Example of a 4-segment stimulus model presentation and imitation attempt. A: The stimulus disc moved along a series of four straight, connected segments. The dashed line represents the disc's path, with the dashed circle signifying the trajectory's starting point. Here the disc is shown at the end of its trajectory. Shortly after completing the trajectory, the stimulus disc disappeared. Although its complete path is shown here, subjects saw only the instantaneous position of the disc. B: When subjects reproduced the trajectory, a smaller disc moved on the screen according to the movements of the stylus on the tablet. The path that the stylus traveled was displayed to the subject as a thin line. The disc in this example is shown at the end of the subject's reproduction of the model. C: Upon lifting the stylus from the tablet, subjects were shown their own reproduction (thin line), as well as the original model (thick line), so information about the quality of the imitation was available to them.

the natural assumption that subjects parsed each trajectory into a set of discrete, directed line segments. Once the imitation had been broken down into individual segments, we used similarity in segment orientation as a measure of accuracy. For this calculation, the absolute orientation of each stimulus segment was compared to the orientation of the corresponding segment in the reproduced path, and the absolute value of the difference between them was regarded as the reproduction error for that segment

As would be the case for any segmentation algorithm's performance with a noisy spatio-temporal input stream (e.g., Calic & Izquierdo, 2002), we recognize that our algorithm was not foolproof. In an attempt to examine the validity of the algorithm, we randomly selected 50 imitations of 6- and 7-segment models, and asked four human judges to view them and break each one into its component segments. The judges, none of whom had been subjects in the actual experiments, watched as the reproductions were shown in real time; they saw the path forming on the screen at the same speed as it had during the actual experiment. The judges could replay the path, and the path remained on the screen once it had appeared. With temporal and spatial data available to them, the judges were asked to "read the mind" of the subject who originally generated the path, marking any and all points that seemed to represent the beginning or end of a segment.

On almost a third (16/50) of the trials, the four independent judges failed to agree on how many segments were in the models. On 30 of the 34 trials where all the judges agreed on the number of segments, the algorithm generated the same number of segments as the judges did. Analogous differences in segmentation were found when subjects observed and attempted to segment videos of complex everyday tasks (Zacks et al., 1999; Speer, Swallow, & Zacks, 2003). Lack of unanimity among human judges exemplifies the challenge presented by

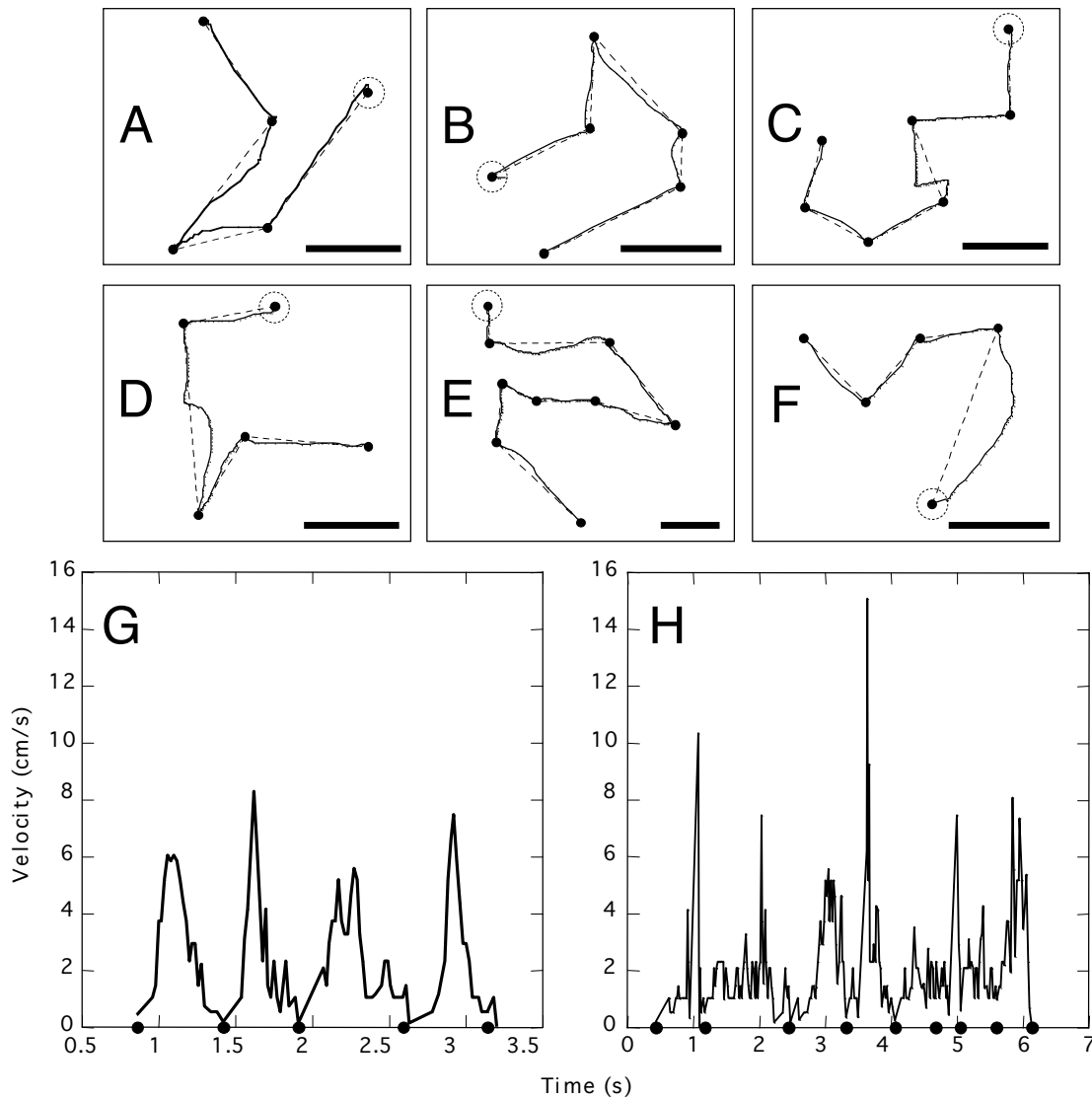
inherently noisy data, such as those generated by subjects' unconstrained movements, and implies that there may not be an "ideal segmenter" against which we could benchmark our algorithm. Figure 2 shows examples of the segmentation algorithm's output, including cases where the algorithm seemed to be at risk of error. An additional, parametric analysis of the segmentation algorithm is found in the supplementary material for this article.

## RESULTS

### *Number of Reproduced Segments*

The analysis of subjects' performance began by identifying imitations in which the number of segments matched the number of segments in the corresponding stimulus model. Figure 3A shows the proportion of segment-matching imitations for models of varying length. The proportion of matching imitations clearly decreases as the number of segments in the model increases, a result confirmed by a repeated measures ANOVA ( $p < .001$ ).

Two kinds of errors characterized imitations in which the number of segments did not match that of the corresponding model: the insertion of an extra segment, i.e., reproduction of  $n + 1$  segments for models with only  $n$  segments, or the omission of one segment, i.e., reproduction of  $n - 1$  segments for models with  $n$  segments. We were interested in how often these two types of errors occurred for different model lengths, so we sorted the identified trials according to the actual number of segments in the corresponding model stimuli. In cases where an extra segment had been inserted into the reproduction, the proportion of errors appeared to be unrelated to stimulus length (Figure 3B,  $p > .20$ , ANOVA). However, when a segment had been omitted, the proportion of omission errors clearly increased with model length (Figure 3C,  $p < .001$ , ANOVA).



**Figure 2.** A-F: Segmentation algorithm examples. The thin, solid lines represent imitation attempts by various subjects. The filled circles represent the segment endpoints found by the algorithm, based on the imitation's spatial and temporal characteristics. The detected endpoints were used to form straight segments, shown here as dashed lines, with the dashed circles representing the models' starting points. Good reproductions usually resulted in successful automatic detection of endpoints (A, B). Unclear or ambiguous attempts (C-F), however, tend to cause difficulty in correctly interpreting the intended number of segments. Note that model segments were of constant length; here, models and their reproductions have been scaled so as to occupy roughly a constant area within the figure. The horizontal scale bars in each panel represent the actual size of a segment in the stimulus model. G-H: Movement speed profiles for two representative trials. Panel G shows the speed at which the stylus moved during the imitation described in panel A. The filled dots along the horizontal axis indicate the times at which the algorithm identified segment breakpoints. Panel H shows the speed profile and identified segment breakpoints for the imitation presented in panel E. The jaggedness of the speed profiles reflects the relatively coarse, 50 Hz sampling of stylus position.

To characterize these two types of error in more detail, we set out to determine *where* in the reproduction sequence a segment was most likely to have been omitted or inserted. Since there were very few omissions in the shorter models, we concentrated on 6- and 7-segment models. To identify the serial position at which an extra segment had been inserted, we used an iterative procedure, removing one segment from

the reproduction, each time from a different serial position within the reproduction, and calculating the mean orientation error for the remaining segments, as if the correct number of segments had been reproduced. The segment whose removal produced the lowest mean error was taken to be the segment that had been added. We were especially interested in testing a hypothesis suggested by the work of

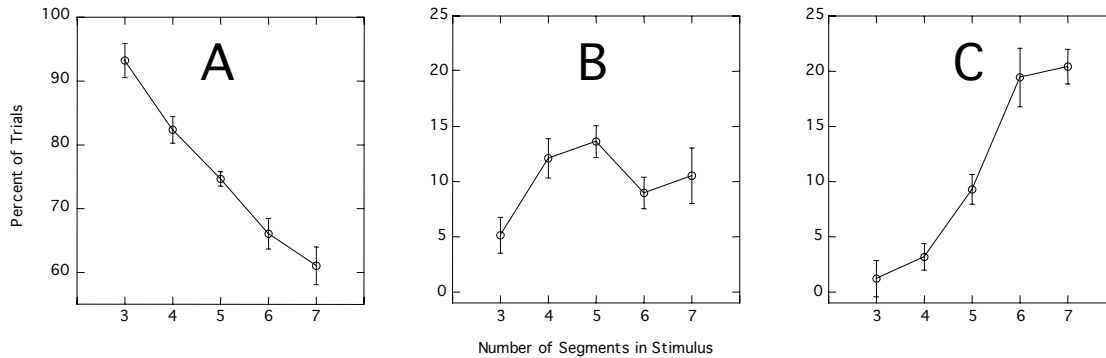


Figure 3. A: Errors in the number of segments in imitations. The plot shows the percent of trials on which the correct number of segments had been reproduced, for different model lengths. B-C: Segment omission, but not insertion, occurs more frequently for longer models. Plots show the percentage, out of the total number of trials, where a false insertion (B) or omission (C) of one segment occurred, for each model length. Error bars are within-subject SEM (Loftus & Masson, 1994).

Fujii and Graybiel (2003), namely that extra segments might result from a failure of neural end-state markers. Those researchers proposed that such markers were as integral to sequential action codes as to other sequential biological codes, including those in DNA. If segment insertion did arise from failures of end-state markers, extra segments should be most common at or near the end of an imitation sequence. Contrary to this hypothesis, extra segments were not restricted to the end of imitations, but occurred at about the same frequency for all serial positions (Figure 4A-B,  $p > .70$  (A),  $p > .90$  (B), ANOVA).

To identify the likely serial position of a segment that was missing from the reproduction, we iteratively removed one segment from the stimulus model, from a different serial position each time, and compared the reproduction to each of these modified stimulus models. The serial position at which a deletion produced the lowest mean orientation error was taken to be the serial position from which the subject's imitation had most likely omitted a segment. This analysis showed that a segment was more likely to be omitted in reproducing a model's end than in reproducing a model's beginning. Figure 4C-D shows the percentage of segment omissions at each serial position, out of all the trials on which an omission occurred ( $p < .035$  (C),  $p < .001$  (D), ANOVA).

### Quality of Segment Reproduction

For a more detailed segment-by-segment analysis we used only trials on which the number of recovered segments had been correct, that is, matched the number of segments in the model. To characterize the dependence of imitation error on model length, we calculated the mean absolute orientation error over each model for which the correct number of segments had been reproduced. The mean orientation error for all segments per trial was taken as the overall error score for that imitation. Figure 5A shows that this mean error increases with the number of segments in the stimulus model (ANOVA,  $p < .001$ ). This is consistent with results reported by Sekuler et al. (2003).<sup>1</sup>

We next examined how serial position affected the accuracy with which individual segments in the model were reproduced. Figure 5B shows the orientation error for segments in each serial position for models of different lengths. Note that there is a strong *primacy* effect: for all model lengths, the error in reproduction is smallest for the earliest serial positions. In addition, all model lengths show at least a modest *recency* effect: the error associated with the final serial position is less than that associated with the next-to-last serial position. In other words, the final segments' errors violate the general upward trend observed across non-final serial positions. Also noteworthy is the difference in error levels at any serial position across models of different lengths. In other words, at any serial position performance degrades with increasing model length. This systematic increase in error at any particular serial position, for example the third, implies a striking retroactive interference, with subsequent events degrading mnemonic storage of preceding information.

### Analysis of Transposition Errors

According to some Competitive Queuing (CQ) models, sequencing errors increase with sequence length because the activity level differences that are available to represent

<sup>1</sup>The literature on orientation anisotropies predicts that our subjects would have represented verticals and horizontals more accurately than other orientations. We found no systematic difference in errors for horizontal vs. vertical segments, but we saw a modest difference between obliques and these cardinal directions, as expected from past literature on the "oblique effect" (e.g., McMahon & MacLeod, 2003). Mean errors were about 2 degrees larger for oblique segments than for cardinal (horizontal or vertical) segments. This effect is largely irrelevant to explanation of all the main error patterns, which are functions of serial position, sequence length, and within-model repetitions. Because segment directions were randomized and uniformly distributed across model lengths and serial positions, the oblique effect would have been averaged out in all but one respect. It would have added about 1 deg of total error, on average, for each additional segment in an imitation.

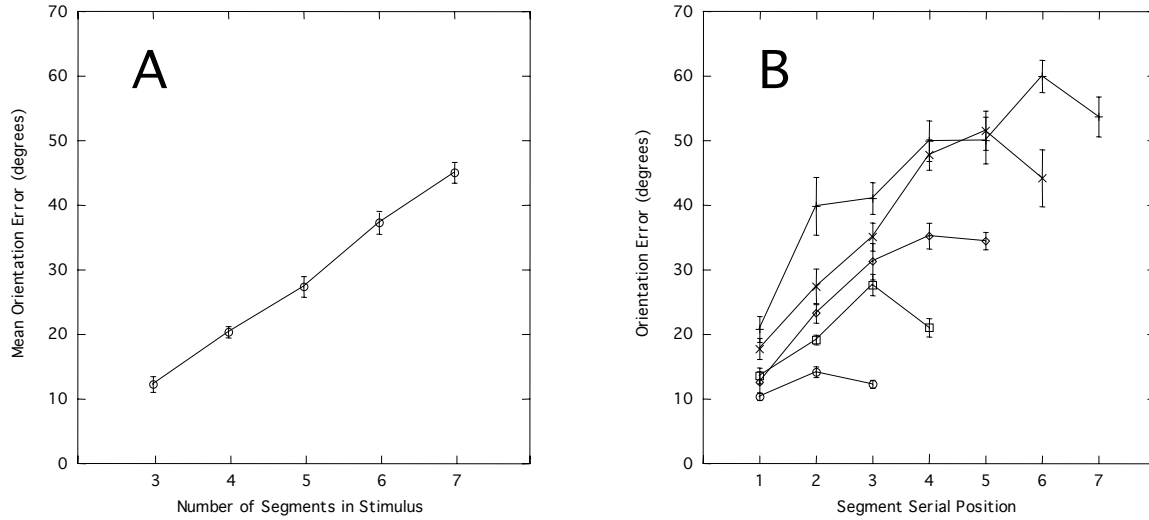


Figure 5. A: The mean orientation error over all trajectory segments as a function of model length. B: Orientation errors broken down by serial position. Each curve corresponds to stimuli of a different length:  $\circ$  = 3 segments;  $\square$  = 4 segments;  $\diamond$  = 5 segments;  $\times$  = 6 segments;  $+$  = 7 segments. Error bars are within-subject SEM, for each curve independently.

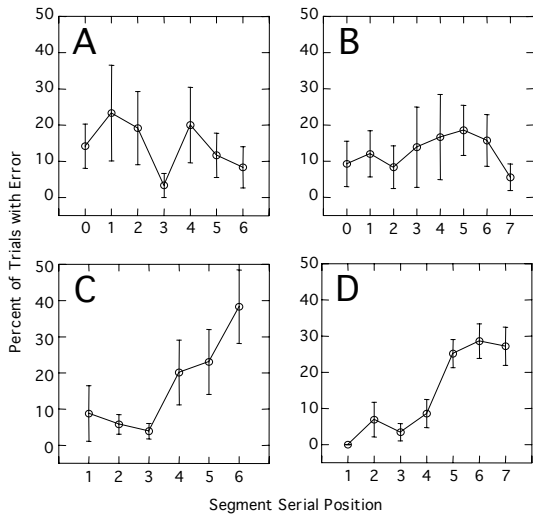


Figure 4. Omission and insertion errors. Panels A and B show the position after which insertion is most likely to have happened, 0 being before the first segment. Panels C and D show the position where a segment is most likely to have been omitted. All error bars are within-subject SEM.

relative priority are necessarily smaller in the representations of longer sequences (see Rhodes et al., 2004, for review). This limit arises because only a small number of distinct analog values can be reliably stored and discriminated within the finite activity range available for neural representation. As the differences between the activities representing individual segments decrease (with longer sequences and

with serial position within a sequence), these representations, and their processing, become more vulnerable to corruption by noise. The resultant mis-processing of relative priority shows up in behavior as transposition errors: Items are recalled, but in the wrong serial positions. In some CQ variants, activity decay (which has a longer time to operate for later items in longer lists) and noise combine to produce a similar result. Because noise is much less likely to erroneously promote an item's relative priority by two (or more) positions than by just one position, all CQ models predict that the preponderance of transposition errors will reflect an exchange of items that were adjacent in the model.

We first considered the full set of opportunities for transpositions between adjacent and non-adjacent serial positions. For every pair of serial positions, we could swap the corresponding two segments in the model in order to "reverse" the candidate transposition involving that pair. We could then recalculate the absolute orientation errors for the two swapped segments and compare those errors with the errors of the unswapped segments. If a transposition error had occurred, then reversing the transposition should, in general, reduce the error at both serial positions. However, if segments at the two positions differed by 0 degrees, their transposition would not inflate error, and so a reversal could not reduce the error. Thus, such cases should be excluded from an analysis of the percentage of cases for which a swap reduces error. In our paradigm, adjacent segments could differ by 45, 90 or 135 degrees, whereas non-adjacent segments could differ by 0, 45, 90, 135 or 180 degrees. To accommodate the 0 degree "exception" just noted, and to remove any potential bias due to the 180 degree differences found only across non-adjacent positions, we used only segment pairs which differed by 45, 90 or 135

degrees. The five matrices in Table 1 show the percentage of opportunities on which segment swaps reduced the error at both serial positions for those segment pairs (As shown in the on-line supplemental materials, nearly identical distributions of these three angular differences were found in the adjacent vs. non-adjacent opportunities for exchange errors). Each matrix shows results for models of a particular length. The position of the first segment in a pair is given by the row number, and the position of the item it could have exchanged with is given by the column number. All values on the main diagonal are zero because a self-exchange leads to no reduction in error; the lower half of each matrix is omitted because it would be redundant with the upper half. The effect of swapping immediately adjacent segments in an imitation is represented by the highlighted values in the cells lying just above the main diagonal. Without exception, for each matrix the largest percentage in the matrix lies on this diagonal ( $p < .015$ ) and is at least twice the size of any of the off-diagonal percentages (which are drawn from non-adjacent cases). Out of 175 comparisons, there are only four minor exceptions (found in the matrices for 5 and 6 segment models) to the rule that percentages for adjacent swaps are larger than any percentages in the remainder of the matrix. This strong ordering supports the prediction from CQ models that the preponderance of transposition errors should arise from mis-orderings of adjacent items. Additional analysis of transposition errors appears in the supplementary material for this article.

### *Absolute vs. Relative Orientation Errors*

To this point, our principal error metric (i.e., the deviation between corresponding segments in the model and the imitation) has presupposed that subjects encode and represent segments' absolute orientations and rely on that representation when reproducing a model. However, it is possible instead that subjects coded each segment in terms of the angular change between it and the prior segment. If noise were introduced as each successive segment were encoded, relative (and therefore absolute) orientation errors would increase with serial position and with the total number of segments in an imitation, so results qualitatively similar to those already seen for absolute orientation errors might arise.

It is difficult to analyze relative orientation errors separately from absolute orientation errors, because the two are highly correlated. However, one would expect that if performers were encoding and controlling a given variable, then its error would be less than that of a correlated, but not explicitly controlled, variable, because performers do not care about the latter, as such. In that case, a transposition analysis that searches for error-inflating exchanges of an actually controlled variable should produce many more cases of error reduction than the same analysis using a correlated, but not explicitly coded, variable. To test whether these two expected rank orderings would favor relative or absolute encoding, we conducted additional error and transposition analyses, but this time under the assumption of relative orientation coding.

For each pair of successive segments in the stimulus, the *change* in orientation of each segment in the imitation was compared to the *change* in orientation of corresponding segments in the model. Finally, the error in imitation was defined as the absolute value of the difference between the two measures of change. Figure 6 shows the serial position curves for errors in relative orientation. Because the analysis compares the difference in orientation between pairs of adjacent segments, only  $n - 1$  data points are available for models of  $n$  segments. With a few exceptions Figure 6 looks much like Figure 5B, as expected given the correlation between absolute and relative errors. However, relative errors are greater than absolute errors by 10 to 20 degrees (chance level is 90 degrees for both). By the logic that a controlled variable should have less error than an uncontrolled correlated variable, this result favors the hypothesis of absolute orientation coding. Moreover, if we compare both sets of serial position curves to the curves commonly found in the immediate serial recall literature, curves defined with the absolute metric better replicate the typical shape, including the recency effect.

A second expectation was that the incidence of error-reducing swaps should be higher when the search for such swaps is based on the actually controlled variable than when purported reversals of a correlated but uncontrolled variable are sought. To examine this prediction, we conducted a transposition analysis identical to that for Table 1, except that relative error was used instead of absolute error. This search for all swaps that could reduce relative error at both swapped positions resulted in much smaller percentages than those presented in Table 1. Thus, when compared to absolute error, plots of relative error produced more error (Figure 6), yet reversals of putative exchanges causing relative errors resulted in many fewer cases of error reduction in a transposition analysis. The two results are mutually consistent, and both favor the hypothesis that what performers cared more about, i.e. were preferentially encoding and controlling, were absolute orientations rather than relative orientations.

Finally, the two encoding hypotheses predict different outcomes for the segment immediately following a transposed pair. Indeed, this predicted difference is an important reason why subjects who are susceptible to transposition errors should prefer absolute orientation representations. Absolute orientation storage allows for recovery from a transposition error, whereas with relative encoding, the next segment is prone to very large error, (if the error criterion being used is absolute error) because its change adds to the incorrect prior orientation. For every pair of adjacent segments for which swapping reduced absolute error at both positions in 6- and 7-segment models, we therefore calculated the difference between the error in the segment immediately following the swapped pair and the average error the subject made at that position (taken from figure 5B). If performance were dominated by encoding of relative orientation, then we would expect the error in those cases to be larger than its respective average (the average error for that position, for a particular subject). Contrary

Table 1  
Results of the transposition analysis.

3 Segments		4 Segments			5 Segments				6 Segments					7 Segments										
0	<b>3.8</b>	1.0	0	<b>4.4</b>	1.1	1.2	0	<b>3.5</b>	2.8	2.0	1.4	0	<b>11.1</b>	5.9	2.8	7.4	3.4	0	<b>13.1</b>	3.6	5.3	1.9	4.1	2.5
	0	<b>2.6</b>		0	<b>5.1</b>	3.0		0	<b>4.8</b>	2.1	2.6		0	<b>9.4</b>	10.5	6.2	4.6		0	<b>18.7</b>	8.7	7.3	7.4	3.0
		0			0	<b>8.2</b>			0	<b>13.1</b>	4.7			0	<b>17.3</b>	6.9	10.2			0	<b>20.9</b>	6.0	5.6	7.4
						0				0	<b>17.0</b>				0	<b>24.2</b>	12.1				0	<b>23.6</b>	9.8	12.5
											0					0	<b>16.8</b>					0	<b>21.4</b>	7.7
																	0						0	<b>32.3</b>
																								0

The value in each row and column represents the percentage of trials for which swapping segments in the corresponding serial positions reduced the error in both positions. Highlighted numbers represent percentage values for adjacent pairs of segments.

to this expectation, the errors were 12.7 degrees and 12.4 degrees smaller than average in 6 and 7 segment imitations, respectively. In summary, analyses consistently favor the hypothesis that imitations are based on a representation of the absolute orientations, rather than the relative orientations, of segments in a model.

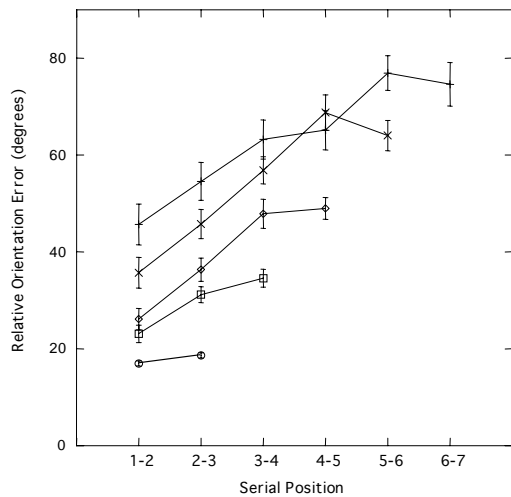


Figure 6. Serial position curves based on errors in relative changes of orientation between segments. The numbers on the x-axis indicate the serial positions of the segments whose relative orientation is being compared. Curves show results from models with different number of segments. The symbols used to represent each model length are identical to those used in Figure 5B. Error bars are within-subject SEM, for each curve independently.

*The Effect of Repetition of Segment Orientation*

Research on the serial recall of verbal items has shown that memory capacity declines (recall errors increase) as items become less discriminable from one another (see Page & Norris, 1998). In the extreme, any pair of segments in our model stimuli would be minimally discriminable from one another when they were identical except for their

serial position, that is, maximum similarity would occur when the same item was repeated at different list positions. Note that the protocol we have used so far to generate models produced an association between model length and repetition of a segment’s orientation: As models grew in length, so too did the probability that a segment orientation presented early in the model might be repeated later in the model. To illustrate this point, Monte Carlo simulations showed that for 4-segment models, the probability that the last segment’s orientation would repeat the orientation of a prior segment was 0.33; for 5-segment models, the corresponding probability was 0.45. Because the probability of a repetition had not been controlled, and because such repetition might influence quality of imitation, we set out to measure the effect of segment repetition directly, controlling the occurrence or non-occurrence of segment repetitions.

To examine the possible effect of repetition, we conducted a second series of sessions with a new group of eight subjects, none of whom had served in the main experiment. These new subjects viewed and reproduced models of four and five segments. On half the trials no segment direction was repeated; on the remaining trials, the final segment in the model repeated the direction of one preceding segment.

Figure 7 shows reproduction errors for stimuli with and without repetitions, for models of four (A) and five (B) segments. Errors are shown for each serial position. With the exception of one set of data points, the difference between performance with and without a repetition of a segment orientation was small. However, in 8 of 9 comparisons, the mean error with a segment repetition was larger than with no repetition ( $p < .05$ , sign test), suggesting that repetition is a weak but significant factor in the quality of reproduction.

*Higher-level Factors: Errors in Segment Location and Length*

There is one likely influence on imitation that has not so far been explicitly mentioned: the heterogeneity of models comprising the *same* number of segments. Complexity in serial recall of sequences is usually defined by the number of elements in the sequence (Rhodes et al., 2004), but it is reasonable to hypothesize other, related influences on

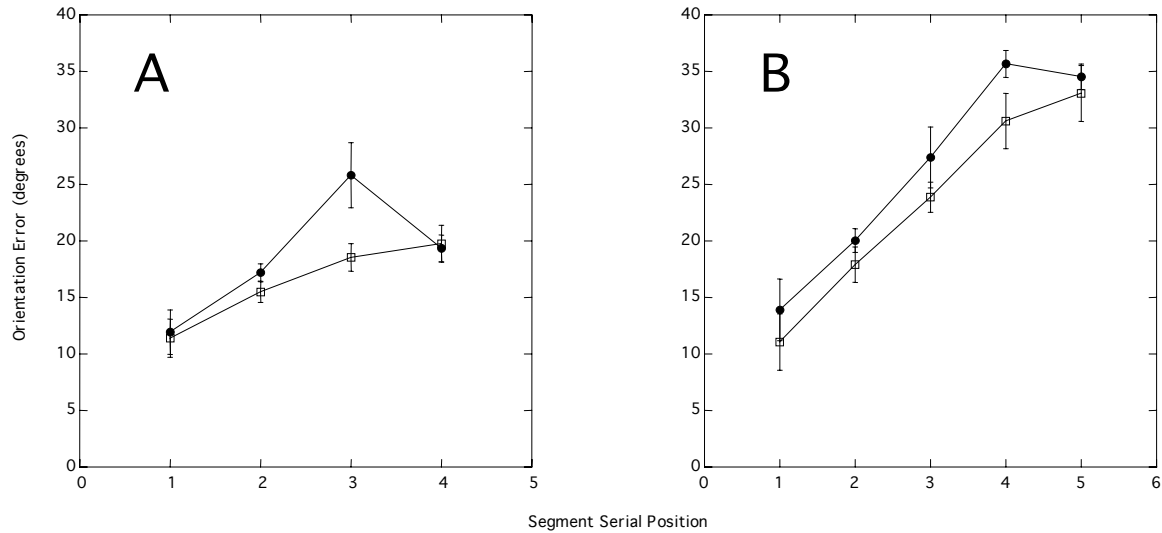


Figure 7. The effect of repetition of segment orientation on the quality of imitation. A: Errors with 4-segment models. B: Errors with 5-segment models. Filled symbols represent performance with segment repetition in models; open symbols represent performance with no repetition. Error bars are within-subject SEM.

movement imitation. In particular, all  $n$ -segment models may not be equally easy to encode and then imitate. We just showed that imitation quality is influenced by the structural relationship among components in a model, that is, whether an element has been repeated (see Figure 7). Also, various factors might cause a moving object's path to be misperceived (Brown & Voth, 1937; Tripathy & Barrett, 2003; Sekuler et al., 2003), which would reduce an ensuing imitation's fidelity to the model's physical characteristics. This source of error would arise from the particular sequence of movements in a trajectory, especially the pattern of directional change from one segment to another. A final source of heterogeneity in our task resides in the perceived, overall trajectory generated by the linked movement segments. As the entire trajectory of the model's movement was never visible on the screen, any representation of that trajectory had to be generated by the subject, probably using an integrative mechanism akin to working memory's visuo-spatial sketchpad (Baddeley, 2003).

Although accounts of movement imitation built on segment by segment production leave out such gestalt effects, we believe that a model's overall shape or trajectory could affect imitation quality. Figure 8A illustrates this point, showing serial position curves for different model lengths, but using a 2D spatial error criterion instead of the discrepancy in segment orientation between model and imitation, as in Figure 5B. To generate this measure of spatial error, the x-y coordinates of the midpoint of every stimulus segment were compared to the corresponding coordinates of the reproduced segment's midpoint for imitations with the correct number of segments. The absolute Euclidean distance was taken as the error score for that segment. Figure 8A shows that for all model lengths, but especially for longer

ones, serial position curves flatten toward the end of an imitation. It is possible that this flattening was driven not by the process responsible for segment-by-segment execution of the initial portion of the imitation, but by a real-time response to a recognition that the ongoing imitation was deviating significantly from the remembered model. This hypothesis assumes that in addition to information about individual segments, subjects have some, perhaps imperfect, memory of the trajectory's overall shape or endpoint. If an imitation's cumulative position error grows with successive segments, memory of the model's overall shape or endpoint allows subjects to recognize in real-time their reproduction's growing deviation from the remembered model, and to make a late-course correction.

There are, however, interesting alternative explanations for the flattening of the longer curves shown in Figure 8A. One possibility is that as imitation of any model proceeds, the length of successive segments is reproduced with increasing fidelity. To evaluate this hypothesis we examined the accuracy in the length of reproduced segments at different serial positions. Figure 8B shows the result, plotting the length of reproduced segments at each serial position, expressed as percent of segment length in the model. The curves represent models of different lengths. A slight decrease in error can, indeed, be seen from the first to the penultimate position, but the sharp rise in error in the last segment rules out accuracy in segment length as the sole underlying factor in the flattening of the position error curves (Figure 8A). Subjects may have attempted to match the distance from the start of the first segment to the end of the last one and adjusted the length of the last segment to compensate for accumulated errors. However, this striking increase in the final segment's length is consistent with predictions of a self-normalizing CQ

model (Bullock, 2004). Because the representation of the final segment in imitation has no competition from a succeeding element, the level of activation for that final segment would be higher than for preceding elements if the working memory were self-normalizing. The result could be larger segment amplitudes for the final movements, especially if the planning field represents movement vectors (as we have assumed), rather than movement endpoints. On the other hand, the longer segments at the end of an imitation might be an artifact of subjects' having to lift the stylus at the end of the final segment, or could result from an accentuation at the end of a drawing, in a manner similar to the flourish that often accompanies the ending of someone's handwritten signature. Finally, note in Figure 8B that the number of segments comprising a model has an influence on the magnitude of this error, with a tendency for reproduced segments in short models ( $n=3$ ) to be larger than reproduced segments in longer models ( $n=7$ ). Although this tendency is consistent with the operation of a re-normalizing process, as postulated in some CQ models (Bullock, 2004), a firm conclusion on this point must await results from experiments specifically designed to address this prediction.

## DISCUSSION

This study introduced a new approach to imitation. By decomposing reproduced models into their individual segments, we were able to examine the fine structure of the imitative process. Studies by Zacks and colleagues showed that humans analyze various seen events into discrete segments with distinct spatio-temporal boundaries, and that this process dominates perception of complex actions and the planning of actions (Zacks et al., 1999, 2001; Speer et al., 2003). It seemed wise, then, to approach imitation as a sequence of components, rather than as a single, integrated unit.

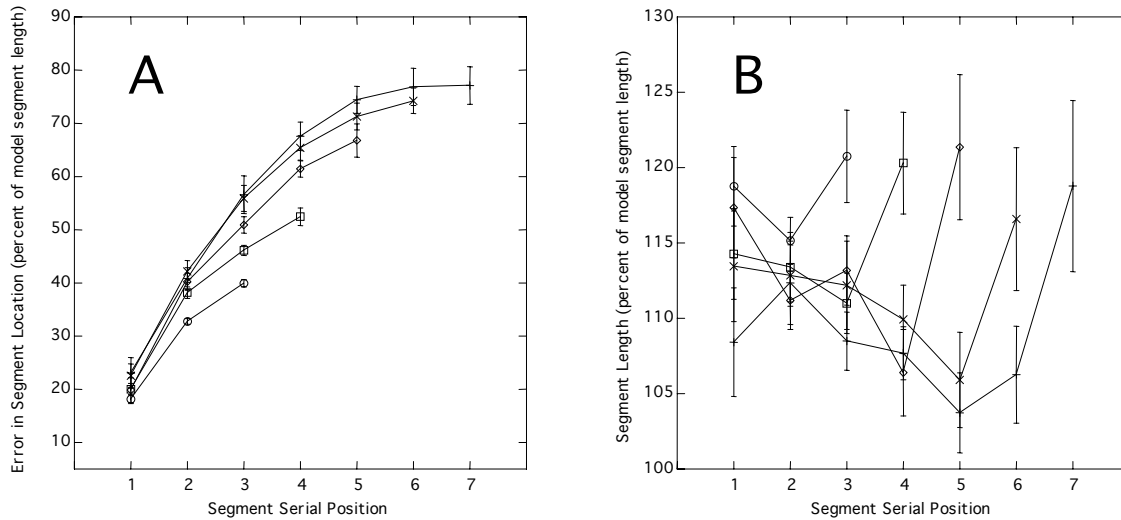
Our results can be compared to those from previous studies of visuo-spatial working memory. Because there was a delay in our task between seeing the trajectory and reproducing it, the representation of information guiding the imitation must have been stored in working memory. Our procedure, however, differs in several ways from those in prior studies, arguably making it more appropriate for generalization to imitation in everyday life. First, the stimulus here is generated by a single, continuously moving object, rather than a series of different objects shown at some single discrete location. Second, serial recall of imitation components was assessed by means of free imitation of the stimulus model, as opposed to the constrained, discrete measures used in most other studies. Finally, error is regarded as a continuous variable, rather than a binary one, such as correct or incorrect (see Farrand, Parmentier, & Jones, 2001; Rhodes et al., 2004). This is arguably more applicable to real-life tasks, which often involve dynamic objects, and are repeated by an observer whose goal is a reproduction that is roughly isomorphic with the model, but not necessarily identical to it (Schaal et al., 2003).

Henriques et al. (2004) followed a related line of

inquiry, asking subjects to feel geometrical shapes and then reproduce them using a robot manipulandum. Their data indicate that errors in reproduction arose principally from distortions of individual segments; integration of segments into a complete shape introduced little error of its own. Despite the difference in modality, this study shares some similarities with ours: Several connected, directional segments are seen or felt, stored in working memory, and then reproduced. Henriques et al. (2004), on the other hand, used only sequences with fewer segments than the span of normal working memory (four segments for closed shapes, three for open shapes), whereas our stimulus models ranged from subspan (three segments) to superspan (seven segments).

### *Insertion and Omission of Segments*

Subjects' ability to reproduce the correct number of segments declined with stimulus length, from nearly total success for 3-segment models to only 60% success for 7-segment models. An interesting finding is the difference between the two most common types of error in the number of reproduced segments: the omission of a segment from the imitation or the insertion of an extra segment. Segment omissions clearly tended to occur more frequently as stimuli grow longer, and to occur more towards the end of the model than at the beginning. This can be thought of as another manifestation of the primacy effect. In contrast, with the other type of error, segment insertion, neither length nor serial position seem to have an effect. This result is inconsistent with the hypothesis that extra segments arise from a failure of a neural end-state encoder, such as that recently proposed to exist in monkey frontal cortex (Fujii and Graybiel, 2003). An important question to be asked is to what extent this type of error might result from incorrect output by the segmentation algorithm. On probability grounds, if the extra segments were generated by algorithm failure, the proportion of trials with such errors would grow with model length, which is not the case (see Figure 3B). To be certain that the extra segments were not a result of the algorithm mistaking small shifts in movement direction for entirely new segments, we examined the distribution of change in direction between a segment that our analysis indicated was wrongly inserted and the segments that preceded and followed it. If algorithm failures were at fault, one should expect such direction changes to be relatively small. However, the size of changes in direction showed a broad, approximately uniform distribution (data not shown), indicating that insertions were not the result of false alarms by the segmentation algorithm, that is a spurious response to a near-threshold, unintended direction change in an imitation. Nevertheless, we cannot currently explain the origin of these inserted segments, which deserve further exploration in future studies. They may be evidence of one or more constructive processes that affect memory for movement sequences. Such processes can be seen as the motor domain equivalents of reconstructive processes that have been demonstrated in other domains of memory (for



*Figure 8.* A: Serial position curves based on a spatial location metric. The y-axis shows the 2D discrepancy between the coordinates of each segment's midpoint and the midpoint of the corresponding segment in the imitation. The units are percent of the length of a segment in the stimulus model; a value of 100, the entire length of a segment, would be an error of 1.15 deg visual angle. B: The length of reproduced segments. The units are percent of the segment length in the stimulus model. In both A and B, curves show results from models with a different number of segments, represented by symbols identical to Figure 5B. Error bars are within-subject SEM, for each curve independently.

example, Bartlett, 1932; Bower, Black, & Turner, 1979; Miller & Gazzaniga, 1998; Roediger, Watson, McDermott, & Gallo, 2001).

### *Accuracy in Reproducing Individual Segments*

Our procedure enabled us to evaluate the quality of imitation on a segment-by-segment level. Consistent with results by Sekuler et al. (2003), we found that the mean error in reproducing a segment's orientation increased with stimulus model length. In addition, our serial position results show a strong primacy effect, consistent with most, if not all, studies of ordered serial recall, verbal (Crowder, 1970; Drewnowski & Murdock, 1980; Drewnowski, 1980) and non-verbal (Farrand et al., 2001; Smyth & Scholey, 1996; Blake, Cepeda, & Hiris, 1997). In addition, also in agreement with results from other serial recall paradigms, there appears to be a strong retroactive effect — errors in a same position are larger if additional, succeeding items must also be remembered. Our results also show a small recency effect, which is expressed mostly in a flattening of the curve, rather than an actual decline in error between the final positions. One should use caution, however, when interpreting this result. Page and Norris (1998) pointed out that in immediate serial recall of short lists, in which error data is dominated by immediate exchanges, i.e. transpositions of adjacent list items (see Henson et al., 1996), the first and last list items are privileged because each faces only one chance to exchange positions with an immediate neighbor, whereas all non-terminal list elements have two such chances. Although our analysis of transposition errors

in imitation is not entirely conclusive, it does seem to support such an account of transpositions.

Our repetition experiment raised the interesting possibility that the stark differences in performance between trials with different model lengths can be accounted for, at least in part, by the higher probability of repeating a segment orientation in longer models. Higher probability of repetition is also more likely toward the end of models, which might help explain the serial position curves. Our results do, indeed, show a modest but consistent effect of repetition. This finding replicates results of prior studies in which lowering the discriminability between items on a list reduced the accuracy of immediate serial recall. Yet, one cannot draw a definitive answer from these data. First, the effect is quite small. Second, note that even when no repetitions are present, performance on 5-segment models is worse than that on 4-segment models (Figure 7). Serial position effects are also apparent in the non-repeating case. This suggests that another process is taking place, along with the interfering effect of repetition. The data from the repetition trials show an interesting pattern, where the error in the next-to-last position (the segment prior to the repeating segment) tends to be the highest. We believe this effect deserves to be explored more deeply in a separate study.

### *Underlying Neural Mechanisms*

What neural mechanisms likely account for the strong effects of model length and segment serial position on imitation quality? Lashley (1951) postulated simultaneous representation of a series of movements prior to their

execution. Lashley also asserted that some errors, notably exchange (or transposition) errors, seem to be caused by interactions between parallel representations. Later theorists (Grossberg, 1978; Houghton, 1990) specified parallel neural models in which relative activity level implicitly specified the relative serial positions of the corresponding movements in a forthcoming serial action. This implicit representation is read out by a competitive queuing mechanism, which iteratively searches for the highest remaining activation, performs the action corresponding to that activation, deletes the representation of that action, searches again, etc. In such models, maintaining the relative activation level differences needed for a reliable representation of relative order becomes more difficult for longer sequences because of the limited 'bandwidth' afforded by the small range of neural activity levels. Representations of items later in a sequence have similarly small activations, which can be corrupted (e.g. reordered) by noise, or simply dropped from the working memory because of decay. Note that this account does not currently include any obvious route by which additional, erroneous components might be inserted into the produced sequence.

Measurements consistent with this coding principle and with CQ were reported recently in awake behaving monkeys (Averbeck et al., 2002, 2003, 2003). Neural activity in the dorsolateral prefrontal cortex (area 46) was recorded while monkeys were cued to reproduce visually displayed geometric figures using a prescribed, highly-practiced, stroke sequence. Although not observed during an imitation task *per se*, these neural responses in area 46, long associated with working memory (Goldman-Rakic, 1990), suggest that this area provides a working memory substrate for imitation of novel sequences. Averbeck and colleagues found evidence for parallel processing of serial movements. Prior to sequence initiation, there was a primacy gradient: the forthcoming sequence was represented by parallel activity in distinct neuronal populations whose relative activation levels anticipate the relative performance order of the represented strokes. During sequence production, the representations were serially deleted, in order, from the initially most active representation to the initially least active, consistent with CQ. Furthermore, by examining error trials, Averbeck et al. showed that copying errors may have been caused by disordering of the relative strength of activity in two populations representing neighboring segments, which results in the monkey "jumping ahead of itself". This is similar to the pairwise exchanges we observed between segments at adjacent serial positions.

The general trend in Table 1 suggests that transpositions involving nearest-neighbor segments are the prevalent type of error, as has been seen in other domains of serial recall from memory (Smyth & Scholey, 1996; Henson et al., 1996). Recently, Farrell and Lewandowsky (2004) noted that several models, in addition to what we have called CQ models, have been formulated or reformulated to explain why transpositions involving nearest-neighbors might be the most prevalent type. Thus, CQ models are not unique in their ability to make this prediction.

However, all CQ models – defined by a primacy gradient across plans, choice of the maximum, and serial deletion of executed plans – make a further prediction that is (so far) unique, namely that the latency of a transposed response should be a monotonically decreasing function of transposition displacement, measured in serial positions. Here, anticipation errors are considered to have a negative transposition displacement and deferral errors a positive transposition displacement. With a task involving speeded verbal sequence recall mediated by working memory, Farrell and Lewandowsky (2004) produced data that confirmed this unique CQ prediction, whereas the same data dis-confirmed the (non-monotonic) predictions of all alternative models. This important result with verbal materials suggests that future tests of models for sequence imitation in non-verbal tasks may benefit from a combined analysis of error and latency. Above, we noted that normalized CQ models may make further predictions (such as enhanced lengths of final segments) that distinguish them from other CQ, as well as non-CQ, models.

Two other areas, aside from the prefrontal cortex, that have been shown to play a crucial role in sequential movement are the supplementary motor area (SMA) and the pre-SMA, just anterior to the SMA. Anatomically, the pre-SMA mediates between prefrontal cortex and the SMA. In monkeys, neurons in SMA and pre-SMA may encode the serial relations among movements in a sequence (Tanji & Shima, 1994; Tanji, 2001). Some SMA neurons are active primarily during execution of one of the movements, again showing a primacy gradient, i.e. more neurons corresponding to the earliest serial positions (Clower & Alexander, 1998). In humans, transcranial magnetic stimulation (TMS) over the SMA interferes with the organization of subsequent movements in a sequence, whereas stimulation over the primary motor cortex affects only the movements that would otherwise have been performed during the stimulation (Gerloff, Corwell, Chen, Hallett, & Cohen, 1997). It is reasonable, therefore, to assume that activity in both SMA and pre-SMA strongly influences the quality of performance in our task.

## Conclusion

By analyzing movement imitation at the level of individual components, we have shown how the quality of imitation depends on the complexity of the modeled movement, and, furthermore, how it varies as the imitation is executed. By examining error patterns across serial positions, we found support for competitive queuing, a parallel processing model of serial behavior. A full understanding of imitation must also take into account other spatial and temporal aspects of individual segments, which we have not considered fully, and higher-order properties of the model, such as the relationship between its individual components and the overall shape that they constitute together.

## References

- Abend, W., Bizzi, E., & Morasso, P. (1982). Human arm trajectory formation. *Brain*, *105*, 331-348.
- Averbeck, B. B., Chafee, M. V., Crowe, D. A., & Georgopoulos, A. P. (2002). Parallel processing of serial movements in prefrontal cortex. *Proceedings of the National Academy of Sciences, USA*, *99*(20), 13172-13177.
- Averbeck, B. B., Chafee, M. V., Crowe, D. A., & Georgopoulos, A. P. (2003). Neural activity in prefrontal cortex during copying geometrical shapes. I. single cells encode shape, sequence, and metric parameters. *Experimental Brain Research*, *150*(2), 127-141.
- Averbeck, B. B., Crowe, D. A., Chafee, M. V., & Georgopoulos, A. P. (2003). Neural activity in prefrontal cortex during copying geometrical shapes. II. decoding shape segments from neural ensembles. *Experimental Brain Research*, *150*(2), 142-153.
- Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), 829-839.
- Bandura, A., Ross, D., & Ross, S. A. (1961). Transmission of aggression through imitation of aggressive models. *Journal of Abnormal and Social Psychology*, *63*, 575-582.
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, England: Cambridge University Press.
- Blake, R., Cepeda, N. J., & Hiris, E. (1997). Memory for visual motion. *Journal of Experimental Psychology: Human Perception & Performance*, *23*(2), 353-359.
- Bower, G. H., Black, J., & Turner, T. (1979). Scripts in memory. *Cognitive Psychology*, *11*, 177-220.
- Brown, J. F., & Voth, A. C. (1937). The path of seen movement as a function of the vector-field. *American Journal of Psychology*, *49A*, 543-563.
- Buccino, G., Vogt, S., Ritzl, A., Fink, G. R., Zilles, K., Freund, H. J., & Rizzolatti, G. (2004). Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. *Neuron*, *42*(2), 323-334.
- Bullock, D. (2004). Adaptive neural models of queuing and timing in fluent action. *Trends in Cognitive Sciences*, *8*(9), 426-433.
- Bullock, D., & Rhodes, B. (2003). Competitive queuing for serial planning and performance. In M. Arbib (Ed.), *Handbook of brain theory and neural networks* (p. 241-244). Cambridge, MA: MIT Press.
- Calic, J., & Izquierdo, E. (2002). Temporal segmentation of MPEG video streams. *EURASIP Journal on Applied Signal Processing*, *2002*(6), 561-565.
- Clower, W. T., & Alexander, G. E. (1998). Movement sequence-related activity reflecting numerical order of components in supplementary and presupplementary motor areas. *Journal of Neurophysiology*, *80*, 1562-1566.
- Crowder, R. G. (1970). The role of one's own voice in immediate memory. *Cognitive Psychology*, *1*(2), 157-178.
- Drewnowski, A. (1980). Attributes and priorities in short-term recall: A new model of memory span. *Journal of Experimental Psychology: General*, *109*(2), 208-250.
- Drewnowski, A., & Murdock, B. B. (1980). The role of auditory features in memory span for words. *Journal of Experimental Psychology: Human Learning*, *6*(3), 319-332.
- Farrand, P., Parmentier, F. B., & Jones, D. M. (2001). Temporal-spatial memory: retrieval of spatial information does not reduce recency. *Acta Psychologica*, *106*(3), 285-301.
- Farrell, S., & Lewandowsky, S. (2004). Modelling transposition latencies: Constraints for theories of serial order memory. *Journal of Memory & Language*, *51*, 115-135.
- Fujii, N., & Graybiel, A. M. (2003). Representation of action sequence boundaries by macaque prefrontal cortical neurons. *Science*, *301*, 1246-1249.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*, 593-609.
- Gerloff, C., Corwell, B., Chen, R., Hallett, M., & Cohen, L. G. (1997). Stimulation over the human supplementary motor area interferes with the organization of future elements in complex motor sequences. *Brain*, *120*, 1587-1602.
- Goldman-Rakic, P. S. (1990). The prefrontal contribution to working memory and conscious experience. In J. C. Eccles & O. Creutzfeldt (Eds.), *Principles of design and operation of the brain* (p. 389-407). Berlin: Springer-Verlag.
- Grossberg, S. (1978). A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In R. Rosen & F. Snell (Eds.), *Progress in theoretical biology* (Vol. 5, p. 233-374). New York: Academic Press.
- Henriques, D. Y. P., Flanders, M., & Soechting, J. F. (2004). Haptic synthesis of shapes and sequences. *Journal of Neurophysiology*, *91*, 1808-1821.
- Henson, R. N. A. (2001). Serial order in short-term memory. *The Psychologist*, *14*(2), 70-73.
- Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology*, *49A*, 80-115.
- Heyes, C. (2001). Causes and consequences of imitation. *Trends in Cognitive Sciences*, *5*(6), 253-261.
- 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* (p. 287-319). New York: Academic Press.
- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, *286*, 2526-2528.
- Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (p. 112-136). New York: Wiley.
- Lewis, S. M., Jerde, T. A., Tzagarakis, C., , Georgopoulos, M. A., Tsekos, N., Amirkian, B., Kim, S. G., Ugurbil, K., & Georgopoulos, A. P. (2003). Cerebellar activation during copying geometrical shapes. *Journal of Neurophysiology*, *90*(6), 3874-3887.
- Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*(4), 476-490.
- Makuuchi, M., Kaminaga, T., & Sugishita, M. (2005). Brain activation during ideomotor praxis: imitation and movements executed by verbal command. *Journal of Neurology, Neurosurgery and Psychiatry*, *76*, 25-33.
- Matarić, M. J., & Pomplun, M. (1998). Fixation behavior in observation and imitation of human movement. *Cognitive Brain Research*, *7*(2), 191-202.
- McMahon, M. J., & MacLeod, D. I. A. (2003). The origin of the oblique effect examined with pattern adaptation and masking. *Journal of Vision*, *3*(3), 230-239.

- Miller, M. B., & Gazzaniga, M. S. (1998). Creating false memories for visual scenes. *Neuropsychologia*, 36(6), 513-520.
- Newton, D., Engquist, G., & Bois, J. (1977). The objective basis of behavior units. *Journal of Personality & Social Psychology*, 37, 847-862.
- Page, M. P. A., & Norris, D. G. (1998). The primacy model: a new model of immediate serial recall. *Psychological Review*, 105(4), 761-781.
- Rhodes, B. J., Bullock, D., Verwey, W. B., Averbeck, B. B., & Page, M. P. A. (2004). Learning and production of movement sequences: Behavioral, neurophysiological, and modeling perspectives. *Human Movement Science*, 23, 699-746.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169-192.
- Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: a multiple regression analysis. *Psychonomic Bulletin & Review*, 8(3), 385-407.
- Schaal, S. (1999). Is imitation learning the route to humanoid robots? *Trends in Cognitive Sciences*, 3(6), 233-242.
- Schaal, S., Ijspeert, A., & Billard, A. (2003). Computational approaches to motor learning by imitation. *Philosophical Transactions of the Royal Society, London, Series B*, 358(1431), 537-547.
- Sekuler, R., Siddiqui, A., Goyal, N., & Rajan, R. (2003). Reproduction of seen actions: Stimulus-selective learning. *Perception*, 32(7), 839-854.
- Smyth, M. M., & Scholey, K. A. (1996). Serial order in spatial immediate memory. *Quarterly Journal of Experimental Psychology*, 49(1), 159-177.
- Speer, N. K., Swallow, K. M., & Zacks, J. M. (2003). Activation of human motion processing areas during event perception. *Cognitive, Affective & Behavioral Neuroscience*, 3(4), 335-345.
- Tanji, J. (2001). Sequential organization of multiple movements: involvement of cortical motor areas. *Annual Review of Neuroscience*, 24, 631-651.
- Tanji, J., & Shima, K. (1994). Role for supplementary motor area cells in planning several movements ahead. *Nature*, 371, 413-416.
- Tripathy, S. P., & Barrett, B. T. (2003). Gross misperceptions in the perceived trajectories of moving dots. *Perception*, 32(11), 1403-1408.
- Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society, London, Series B*, 358, 593-602.
- Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., Buckner, R. L., & Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience*, 4, 651-655.
- Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, 127(1), 3-21.
- Zacks, J. M., Tversky, B., & Iyer, G. (1999). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General*, 130, 29-58.