Imitating Unfamiliar Sequences of Connected Linear Motions

Yigal Agam,[1] Daniel Bullock,[2] and Robert Sekuler[1,2]

[1]*Volen Center for Complex Systems, Brandeis University, Waltham; and* [2]*Cognitive and Neural Systems, Boston University, Boston, Massachusetts*

Agam, Yigal, Daniel Bullock, and Robert Sekuler. Imitating unfamiliar sequences of connected linear motions. *J Neurophysiol* 94: 2832–2843, 2005. First published July 13, 2005; doi:10.1152/jn.00366.2005. A fundamental challenge in neuroscience is to understand the mechanisms by which multicomponent actions are represented and sequenced for production. We addressed this challenge with a movement-imitation task in which subjects viewed the quasi-random, two-dimensional 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 with 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 that, 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 conspecifics, an ability that is critical for learning and perfecting a wide range of motor and social behaviors (e.g., Bandura et al. 1961; Heyes 2001; Wolpert et al. 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 predefined, fixed location (but see Mataríc and 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 et al. 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 et al. 1996) and subsequently postulated in Broca's area (Buccino et al. 2004; Iacoboni et al. 1999), the human brain's putative F5 homologue, may be implicated in action understanding and in movement imitation (Rizzolatti and Craighero 2004). However, 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 2004; Bullock and Rhodes 2003; Grossberg 1978; Houghton 1990; Rhodes et al. 2004), and recent neurophysiological results from awake behaving monkeys have strongly supported this approach (Averbeck et al. 2002, 2003a,b). 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 (Henson et al. 1996; Page and Norris 1998).

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, multicomponent models. The models for imitation were abstract movement sequences of graded complexity. Such two-dimensional (2D) 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

Address for reprint requests and other correspondence: R. Sekuler, Volen Center for Complex Systems, Brandeis University, Waltham, MA 02454 (E-mail: sekuler@brandeis.edu).

model (Sekuler et al. 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 on 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 2D motion patterns, generated by a disc that moved along an irregular, unpredictable trajectory constituting variable numbers of segments. Because 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: the serial ordering of their components (Lashley 1951; Rhodes et al. 2004; Zacks et al. 1999). We therefore modified the methods of Sekuler et al. (2003) to facilitate the detection and analysis of effects associated with individual components' serial 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 sought to evaluate how quality of imitation might vary with the serial position that a segment occupied in the model. We also sought 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 sought 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 2001; Henson et al. 1996). 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 spatiotemporal discontinuity as the principal evidence that one component of an event has ended and another has begun (Newtson et al. 1977; Zacks and Tversky 2001; Zacks et al. 1999, 2001). Therefore we built our segmentation algorithm on a foundation of spatiotemporal discontinuities in the subjects' performance. The segmentation algorithm is a simplified version of ones used in previous research (e.g., Abend et al. 1982; Lewis et al. 2003).

Subjects

A total of 21 human subjects (ages 18–34 yr), 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° 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 ° visual angle long (Fig. 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° 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 and 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

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

where s_i represents the segment in a model composed of n segments. Trajectory complexity, defined by the number of segments in a model, varied randomly from trial to trial, with $3 < n < 7$ (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 ith segment is given by

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

where θ_i is the direction of the ith segment. In the present experiments, the movement direction θ of each new segment was restricted to multiples of 45°. $\Delta\theta$ was a quasi-random variable with a minimum value of 45°. 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 et al.
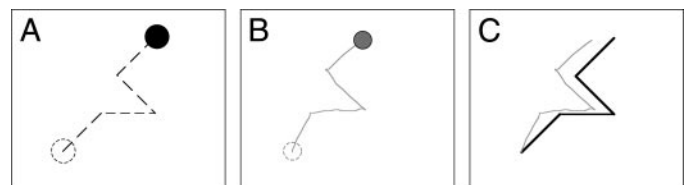

FIG. 1. Example of a 4-segment stimulus model presentation and imitation attempt. A: stimulus disc moved along a series of 4 straight, connected segments. Dashed line represents the disc's path, with the dashed circle signifying the trajectory's starting point. Here the disc is shown at the end of its 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 disc's trajectory, a smaller disc moved on the screen according to the movements of the stylus on the tablet. Path that the stylus traveled was displayed to the subject as a thin line. Disc in this example is shown at the end of the subject's reproduction of the model. C: on 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.

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, such as regular polygons.

Reproduction of disc trajectory

Subjects reproduced the movement of the disc by drawing with a stylus on a 15×20-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 (Fig. 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 (Fig. 1C). Such outcome feedback was meant to provide subjects with trialwise information about the quality of their performance.

Recording of stylus movement

The position of the stylus 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 1 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 time-based 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 et al. (2003) noted, the quality of any movement imitation can be expressed in a number of alternative metrics. When multicomponent 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., Newtson et al. 1977; Zacks et al. 2001). Such studies show that perceptual segmentation depends mainly on marking spatiotemporal 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 >40 ms. From these potential segment breakpoints, the algorithm identified points at which the direction of stylus movement shifted instantaneously by $\geq 15°$. In addition, points where a shift of $\geq 30°$ occurred (without regard to time) were also selected as potential breakpoints. When this selection process left several candidate points clustered together <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%) because 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 Fig. 2). This approach is consistent with 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 with 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 spatiotemporal input stream (e.g., Calic and 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 six- and seven-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 (Speer et al. 2003; Zacks et al. 1999). 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.[1]

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 ($P < 0.001$).

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

---

[1] The Supplementary Material for this article is available online at http://jn.physiology.org/cgi/content/full/00366.2005/DC1.
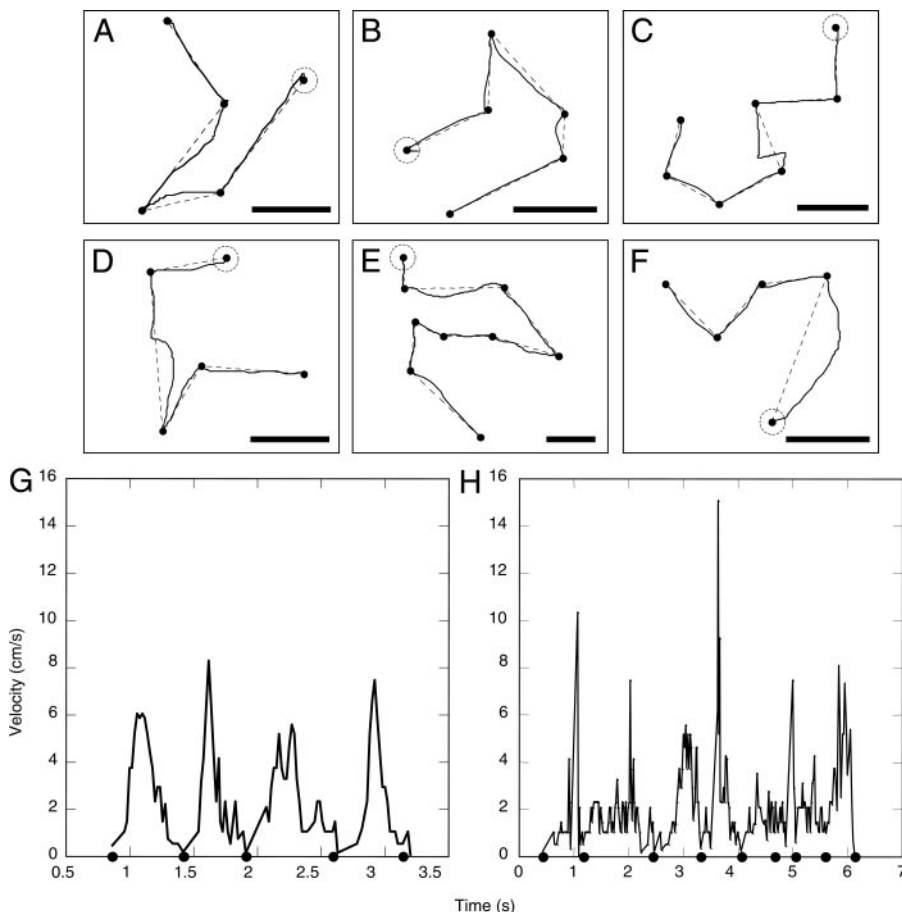
FIG. 2. *A–F*: segmentation algorithm examples. Thin, solid lines represent imitation attempts by various subjects. Filled circles represent the segment endpoints found by the algorithm, based on the imitation's spatial and temporal characteristics. Detected endpoints were used to form straight segments, 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* and *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. Horizontal scale bars in each panel represent the actual size of a segment in the stimulus model. *G* and *H*: movement speed profiles for 2 representative trials. *G*: speed at which the stylus moved during the imitation described in *A*. Filled dots along the horizontal axis indicate the times at which the algorithm identified segment breakpoints. *H*: speed profile and identified segment breakpoints for the imitation presented in *E*. Jaggedness of the speed profiles reflects the relatively coarse, 50-Hz sampling of stylus position.

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 (Fig. 3*B*, $P > 0.20$, ANOVA). However, when a segment had been omitted, the proportion of omission errors clearly increased with model length (Fig. 3*C*, $P < 0.001$, ANOVA).

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. Because there were very few omissions in the shorter models, we concentrated on six- and seven-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 Fujii and Graybiel (2003), that is, 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
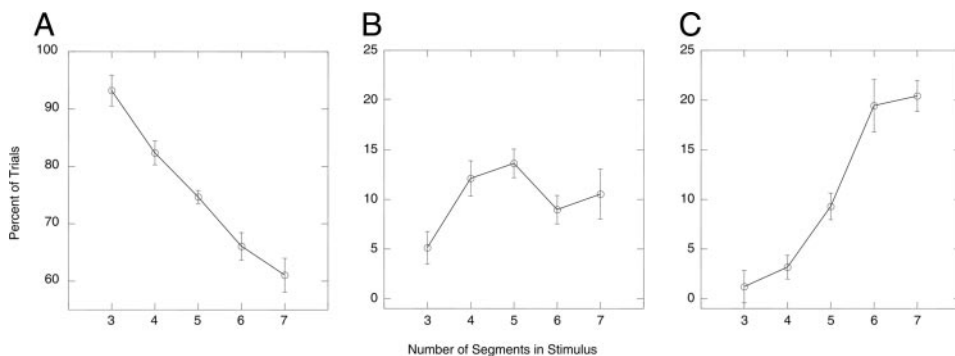

FIG. 3. *A*: errors in the number of segments in imitations. Plot shows the percentage of trials on which the correct number of segments had been reproduced, for different model lengths. *B* and *C*: segment omission, but not insertion, occurs more frequently for longer models. Plots show the percentage, 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 SE (Loftus and Masson 1994).
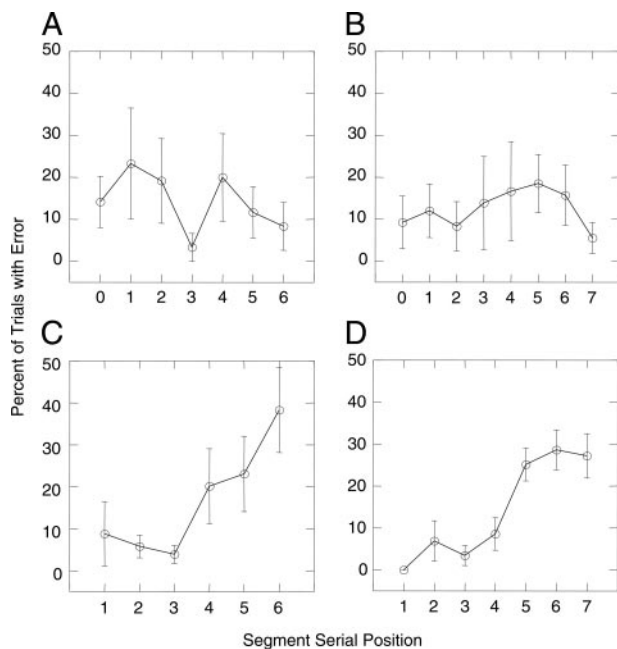
FIG. 4. Omission and insertion errors. *A* and *B*: position after which insertion is most likely to have happened, 0 being before the first segment. *C* and *D*: position where a segment is most likely to have been omitted. All error bars are within-subject SE.

at about the same frequency for all serial positions [Fig. 4, *A* and *B*: $P > 0.70$ (*A*), $P > 0.90$ (*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 4, *C* and *D* shows the percentage of segment omissions at each serial position, out of all the trials on which an omission occurred [$P < 0.035$ (*C*), $P < 0.001$ (*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 dependency 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 5*A* shows that this mean error increases with the number of segments in the stimulus model (ANOVA, $P < 0.001$). This is consistent with results reported by Sekuler et al. (2003).[2]

We next examined how serial position affected the accuracy with which individual segments in the model were reproduced. Figure 5*B* 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 nonfinal 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 relative priority are necessarily smaller in the representations of longer sequences (for review see Rhodes et al. 2004). 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 misprocessing 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 nonadjacent serial positions. For every pair of serial positions, we could swap the corresponding two segments in the model 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°, their transposition would not inflate error, and so a

---

[2] 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 versus 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 and MacLeod 2003). Mean errors were about 2° 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° of total error, on average, for each additional segment in an imitation.
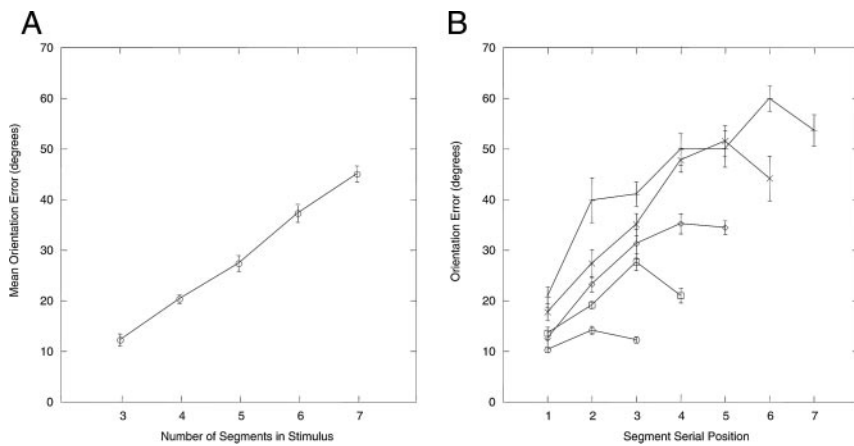
FIG. 5. *A*: mean orientation error over all trajectory segments as a function of model length. *B*: orientation errors broken down by serial position. Each curve corresponds to stimuli of a different length: ○ = 3 segments; □ = 4 segments; and ◇ = 5 segments; ✕ = 6 segments; + = 7 segments. Error bars are within-subject SE, for each curve independently.

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°, whereas nonadjacent segments could differ by 0, 45, 90, 135, or 180°. To accommodate the 0° "exception" just noted, and to remove any potential bias arising from the 180° differences found only across nonadjacent positions, we used only segment pairs that differed by 45, 90, or 135°. 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 online supplemental materials, nearly identical distributions of these three angular difference were found in the adjacent vs. nonadjacent 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 ($P < 0.015$) and is at least twice the size of any of the off-diagonal percentages (which are drawn from nonadjacent cases). Out of 175 comparisons, there are only four minor exceptions (found in the matrices for five- and six-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 misorderings of adjacent

items. Additional analysis of transposition errors appears in the supplementary material for this article.

Absolute versus 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

TABLE 1. *Results of the transposition analysis*

3 Segments			4 Segments				5 Segments					6 Segments						7 Segments						
0	**3.8**	1.0	0	**4.4**	1.1	1.2	0	**3.5**	2.8	2.0	1.4	0	**11.1**	5.9	2.8	7.4	3.4	0	**13.1**	3.6	5.3	1.9	4.1	2.5
	0	**2.6**		0	**5.1**	3.0		0	**4.8**	2.1	2.6		0	**9.4**	10.5	6.2	4.6		0	**18.7**	8.7	7.3	7.4	3.0
		0			0	**8.2**			0	**13.1**	4.7			0	**17.3**	6.9	10.2			0	**20.9**	6.0	5.6	7.4
						0				0	**17.0**				0	**24.2**	12.1				0	**23.6**	9.8	12.5
											0					0	**16.8**					0	**21.4**	7.7
																	0						0	**32.3**
																								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. Numbers in boldface type represent percentage values for adjacent pairs of segments.

compared with 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 Fig. 6 looks much like Fig. 5*B*, as expected, given the correlation between absolute and relative errors. However, relative errors are greater than absolute errors by 10 to 20° (chance level is 90° 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 with absolute error, plots of relative error produced more error (Fig. 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 about, i.e., were preferentially encoding and controlling, were absolute orientations rather than relative orientations.
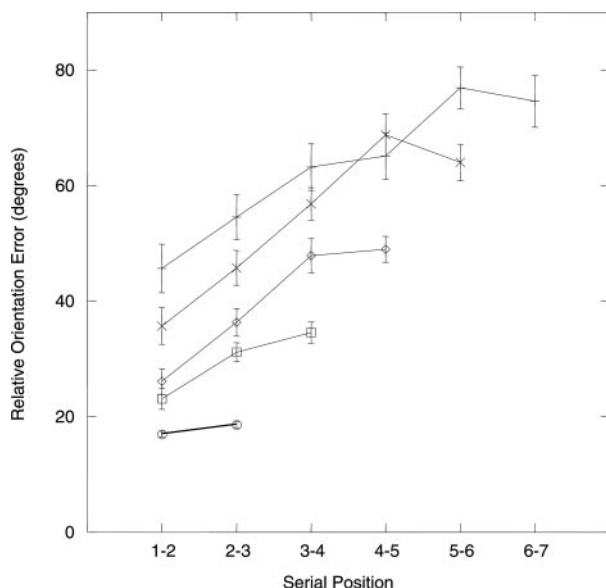

FIG. 6. Serial position curves based on errors in relative changes of orientation between segments. Numbers on the *x*-axis indicate the serial positions of the segments whose relative orientation is being compared. Curves show results from models with different number of segments. Symbols used to represent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

Finally, the two encoding hypotheses predict different outcomes for the segment immediately after 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 six- and seven-segment models, we therefore calculated the difference between the error in the segment immediately after the swapped pair and the average error the subject made at that position (taken from Fig. 5*B*). 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 to this expectation, the errors were 12.7 and 12.4° smaller than average in six- and seven-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.

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 and 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 four-segment models, the probability that the last segment's orientation would repeat the orientation of a prior segment was 0.33; for five-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 nonoccurrence 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 eight of nine comparisons, the
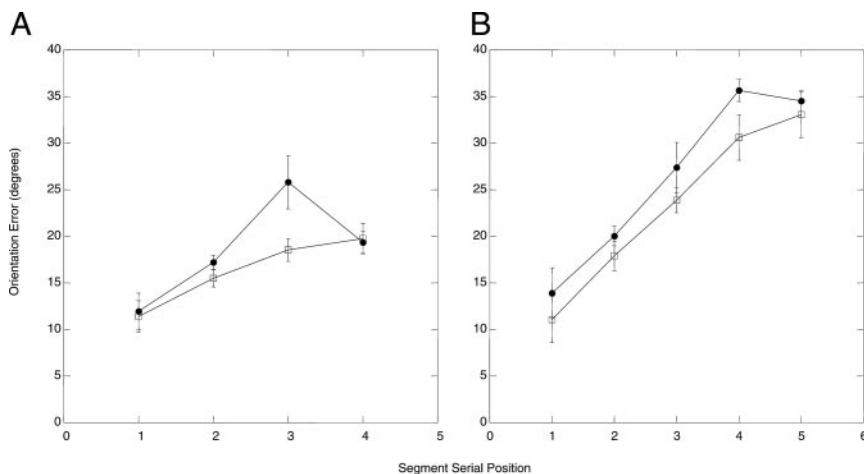
mean error with a segment repetition was larger than that with no repetition ($P < 0.05$, sign test), suggesting that repetition is a weak but significant factor in the quality of reproduction.

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 composed of 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 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 Fig. 7). Also, various factors might cause a moving object's path to be misperceived (Brown and Voth 1937; Sekuler et al. 2003; Tripathy and Barrett 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. Because 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 visuospatial 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 8*A* 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 Fig. 5*B*. To generate this measure of spatial error, the *x–y* coordinates of the midpoint of every stimulus segment were compared with 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 8*A* 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
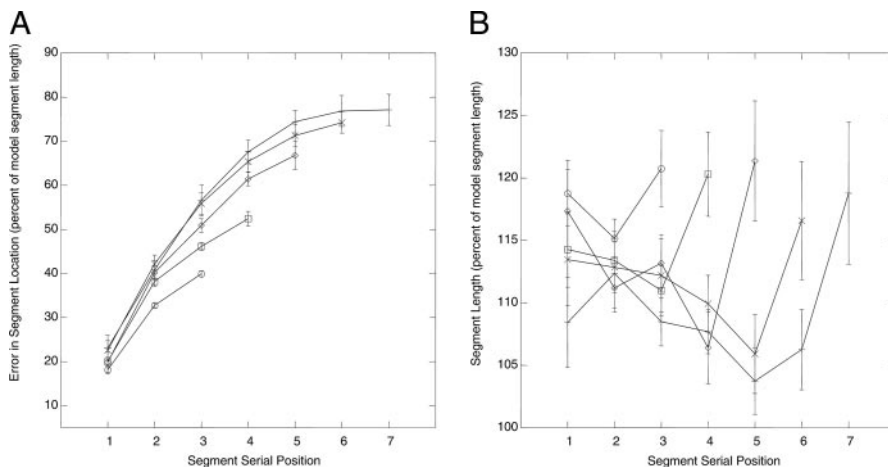
allows subjects to recognize in real time their reproduction's growing deviation from the remembered model, and to make a late-course correction.

There are, however, interesting alternative explanations for the flattening of the longer curves shown in Fig. 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 percentage 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 (Fig. 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 that 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 Fig. 8B that the number of segments constituting 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 renormalizing 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 spatiotemporal boundaries, and that this process dominates perception of complex actions and the planning of actions (Speer et al. 2003; Zacks et al. 1999, 2001). It thus seemed wise to approach imitation as a sequence of components, rather than as a single, integrated whole.

Our results can be compared with those from previous studies of visuospatial 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 et al. 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 three-segment models to only 60% success for seven-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 toward 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 Fig. 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 (e.g., Bartlett 1932; Bower et al. 1979; Miller and Gazzaniga 1998; Roediger et al. 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 1980; Drewnowski and Murdock 1980) and nonverbal (Blake et al. 1997; Farrand et al. 2001; Smyth and Scholey 1996). 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) indicated that in immediate serial recall of short lists, in which error data are 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 nonterminal 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 five-segment models is worse than that on four-segment models (Fig. 7). Serial position effects are also apparent in the nonrepeating 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 before 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 before 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, and so forth. 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, 2003a,b). 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. Before 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 and colleagues 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 (Henson et al. 1996; Smyth and Scholey 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: 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 disconfirmed the (nonmonotonic) predictions of all alternative models. This important result with verbal materials suggests that future tests of models for sequence imitation in nonverbal tasks may benefit from a combined analysis of error and latency. Above, we noted that normalized CQ models might 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 2001; Tanji and Shima 1994). 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 and 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 et al. 1997). It is therefore reasonable to assume that activity in both SMA and pre-SMA strongly influences the quality of performance in our task.

In 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.

GRANTS

REFERENCES

Abend W, Bizzi E, and Morasso P. Human arm trajectory formation. *Brain* 105: 331–348, 1982.

Averbeck BB, Chafee MV, Crowe DA, and Georgopoulos AP. Parallel processing of serial movements in prefrontal cortex. *Proc Natl Acad Sci USA* 99: 13172–13177, 2002.

Averbeck BB, Chafee MV, Crowe DA, and Georgopoulos AP. Neural activity in prefrontal cortex during copying geometrical shapes. I. Single cells encode shape, sequence, and metric parameters. *Exp Brain Res* 150: 127–141, 2003a.

Averbeck BB, Crowe DA, Chafee MV, and Georgopoulos AP. Neural activity in prefrontal cortex during copying geometrical shapes. II. Decoding shape segments from neural ensembles. *Exp Brain Res* 150: 142–153, 2003b.

Baddeley A. Working memory: looking back and looking forward. *Nat Rev Neurosci* 4: 829–839, 2003.

Bandura A, Ross D, and Ross SA. Transmission of aggression through imitation of aggressive models. *J Abnorm Soc Psychol* 63: 575–582, 1961.

Bartlett FC. *Remembering: A Study in Experimental and Social Psychology.* Cambridge, UK: Cambridge Univ. Press, 1932.

Blake R, Cepeda NJ, and Hiris E. Memory for visual motion. *J Exp Psychol Hum Percept Perform* 23: 353–359, 1997.

Bower GH, Black J, and Turner T. Scripts in memory. *Cogn Psychol* 11: 177–220, 1979.

Brown JF and Voth AC. The path of seen movement as a function of the vector-field. *Am J Psychol* 49: 543–563, 1937.

Buccino G, Vogt S, Ritzl A, Fink GR, Zilles K, Freund HJ, and Rizzolatti G. Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. *Neuron* 42: 323–334, 2004.

Bullock D. Adaptive neural models of queuing and timing in fluent action. *Trends Cogn Sci* 8: 426–433, 2004.

Bullock D and Rhodes B. Competitive queuing for serial planning and performance. In: *Handbook of Brain Theory and Neural Networks*, edited by Arbib M. Cambridge, MA: MIT Press, 2003, p. 241–244.

Calic J and Izquierdo E. Temporal segmentation of MPEG video streams. *EURASIP J Appl Signal Process* 6: 561–565, 2002.

Clower WT and Alexander GE. Movement sequence-related activity reflecting numerical order of components in supplementary and presupplementary motor areas. *J Neurophysiol* 80: 1562–1566, 1998.

Crowder RG. The role of one's own voice in immediate memory. *Cogn Psychol* 1: 157–178, 1970.

Drewnowski A. Attributes and priorities in short-term recall: a new model of memory span. *J Exp Psychol Gen* 109: 208–250, 1980.

Drewnowski A and Murdock BB. The role of auditory features in memory span for words. *J Exp Psychol Hum Learn* 6: 319–332, 1980.

Farrand P, Parmentier FB, and Jones DM. Temporal-spatial memory: retrieval of spatial information does not reduce recency. *Acta Psychol* 106: 285–301, 2001.

Farrell S and Lewandowsky S. Modelling transposition latencies: constraints for theories of serial order memory. *J Mem Lang* 51: 115–135, 2004.

Fujii N and Graybiel AM. Representation of action sequence boundaries by macaque prefrontal cortical neurons. *Science* 301: 1246–1249, 2003.

Gallese V, Fadiga L, Fogassi L, and Rizzolatti G. Action recognition in the premotor cortex. *Brain* 119: 593–609, 1996.

Gerloff C, Corwell B, Chen R, Hallett M, and Cohen LG. Stimulation over the human supplementary motor area interferes with the organization of future elements in complex motor sequences. *Brain* 120: 1587–1602, 1997.

Goldman-Rakic PS. The prefrontal contribution to working memory and conscious experience. In: *Principles of Design and Operation of the Brain*, edited by Eccles JC and Creutzfeldt O. Berlin: Springer-Verlag, 1990, p. 389–407.

Grossberg S. A theory of human memory: self-organization and performance of sensory-motor codes, maps, and plans. In: *Progress in Theoretical Biology*, edited by Rosen R and Snell F. New York: Academic Press, 1978, vol. 5, p. 233–374.

Henriques DYP, Flanders M, and Soechting JF. Haptic synthesis of shapes and sequences. *J Neurophysiol* 91: 1808–1821, 2004.

Henson RNA. Serial order in short-term memory. *The Psychologist* 14: 70–73, 2001.

Henson RNA, Norris DG, Page MPA, and Baddeley AD. Unchained memory: error patterns rule out chaining models of immediate serial recall. *Q J Exp Psychol A* 49: 80–115, 1996.

Heyes C. Causes and consequences of imitation. *Trends Cogn Sci* 5: 253–261, 2001.

Houghton G. The problem of serial order: a neural network model of sequence learning and recall. In: *Current Research in Natural Language Generation*, edited by Dale R, Mellish C, and Zock M. New York: Academic Press, 1990, p. 287–319.

Iacoboni M, Woods RP, Brass M, Bekkering H, Mazziotta JC, and Rizzolatti G. Cortical mechanisms of human imitation. *Science* 286: 2526–2528, 1999.

Lashley KS. The problem of serial order in behavior. In: *Cerebral Mechanisms in Behavior*, edited by Jeffress LA. New York: Wiley, 1951, p. 112–136.

Lewis SM, Jerde TA, Tzagarakis C, Georgopoulos MA, Tsekos N, Amirikian B, Kim SG, Ugurbil K, and Georgopoulos AP. Cerebellar activation during copying geometrical shapes. *J Neurophysiol* 90: 3874–3887, 2003.

Loftus GR and Masson MEJ. Using confidence intervals in within-subject designs. *Psychon Bull Rev* 1: 476–490, 1994.

Makuuchi M, Kaminaga T, and Sugishita M. Brain activation during ideomotor praxis: imitation and movements executed by verbal command. *J Neurol Neurosurg Psychiatry* 76: 25–33, 2005.

Mataríc MJ and Pomplun M. Fixation behavior in observation and imitation of human movement. *Cogn Brain Res* 7: 191–202, 1998.

McMahon MJ and MacLeod DIA. The origin of the oblique effect examined with pattern adaptation and masking. *J Vision* 3: 230–239, 2003.

Miller MB and Gazzaniga MS. Creating false memories for visual scenes. *Neuropsychologia* 36: 513–520, 1998.

Newtson D, Engquist G, and Boi J. The objective basis of behavior units. *J Pers Soc Psychol* 37: 847–862, 1977.

Page MPA and Norris DG. The primacy model: a new model of immediate serial recall. *Psychol Rev* 105: 761–781, 1998.

Rhodes BJ, Bullock D, Verwey WB, Averbeck BB, and Page MPA. Learning and production of movement sequences: behavioral, neurophysiological, and modeling perspectives. *Hum Move Sci* 23: 699–746, 2004.

Rizzolatti G and Craighero L. The mirror-neuron system. *Annu Rev Neurosci* 27: 169–192, 2004.

Roediger HL, Watson JM, McDermott KB, and Gallo DA. Factors that determine false recall: a multiple regression analysis. *Psychon Bull Rev* 8: 385–407, 2001.

Schaal S. Is imitation learning the route to humanoid robots? *Trends Cogn Sci* 3: 233–242, 1999.

Schaal S, Ijspeert A, and Billard A. Computational approaches to motor learning by imitation. *Philos Trans R Soc Lond Ser B Biol Sci* 358: 537–547, 2003.

Sekuler R, Siddiqui A, Goyal N, and Rajan R. Reproduction of seen actions: stimulus-selective learning. *Perception* 32: 839–854, 2003.

Smyth MM and Scholey KA. Serial order in spatial immediate memory. *Q J Exp Psychol A* 49: 159–177, 1996.

Speer NK, Swallow KM, and Zacks JM. Activation of human motion processing areas during event perception. *Cogn Affect Behav Neurosci* 3: 335–345, 2003.

Tanji J. Sequential organization of multiple movements: involvement of cortical motor areas. *Annu Rev Neurosci* 24: 631–651, 2001.

Tanji J and Shima K. Role for supplementary motor area cells in planning several movements ahead. *Nature* 371: 413–416, 1994.

Tripathy SP and Barrett BT. Gross misperceptions in the perceived trajectories of moving dots. *Perception* 32: 1403–1408, 2003.

Wolpert DM, Doya K, and Kawato M. A unifying computational framework for motor control and social interaction. *Philos Trans R Soc Lond Ser B Biol Sci* 358: 593–602, 2003.

Zacks JM, Braver TS, Sheridan MA, Donaldson DI, Snyder AZ, Ollinger JM, Buckner RL, and Raichle ME. Human brain activity time-locked to perceptual event boundaries. *Nat Neurosci* 4: 651–655, 2001.

Zacks JM and Tversky B. Event structure in perception and conception. *Psychol Bull* 127: 3–21, 2001.

Zacks JM, Tversky B, and Iyer G. Perceiving, remembering, and communicating structure in events. *J Exp Psychol Gen* 130: 29–58, 1999.