

# The Role of Video Coherence on Object-Based Motion Discriminations by Pigeons

Robert G. Cook and Shelley Roberts  
Tufts University

Two experiments examined the effects of video coherence on the discrimination of relative motion by pigeons using a go/no-go procedure. Pigeons were trained to discriminate video stimuli in which the camera's perspective went either "around" or "through" the interior opening of 2 approaching objects. Experiment 1 used a within-groups design and Experiment 2 used a between-groups design to examine how sequencing these videos in a coherent smooth forward order versus a randomly scrambled order influenced learning. Discrimination learning was significantly faster with the coherent sequences. It is suggested that the pigeon visual system integrates 3-dimensional motion signals across space and time to produce a stable, object-based, perceptual world.

*Keywords:* motion, motion perception, pigeons, comparative perception

One of the more remarkable feats of the human visual system is the apparent stability of the objects and events in the world. Regardless of the source of motion, objects within the visual field cohere and maintain their identity, despite the confusing and constant smear of spatial and temporal ambiguity that crisscrosses the retina, much akin to a small toddler playing with a camcorder. This derived stability is certainly valuable in interpreting the ever-changing landscape of dynamic visual events in the environment. Any highly mobile diurnal animal likely shares this same need. Perceiving and recognizing the actions of organized perceptual "objects," such as predators, conspecifics, or obstacles, seems critical to interacting appropriately with the world, at least at the spatial scale occupied by animals like birds and mammals. With the added dimension of flight, birds may be especially adept at processing motion and using it to build object-based perceptual representations of the world. Because of the past and vast concentration on studying the visual processing of static images in nonhuman animals, how they perceive, process, and interpret the continuous stream of information that truly constitutes reality is a still poorly described and understood phenomenon. The experiments in this article contribute to correcting this oversight by looking at how pigeons process different types of object-based motion.

Recent advances in video and computer animation technology have started to close this empirical and theoretical gap in the analysis of dynamic stimuli, as both psychologists and biologists increasingly use these powerful techniques to examine how animals perceive and react to extended sequences of images that can recreate the temporal and spatial properties of the real world (Cavoto & Cook, 2006; Dittrich & Lea, 2001). Three general approaches have been used. The first approach has used videos of actual behavior as playback stimuli to examine their capacity to

elicit appropriate action patterns (e.g., Adret, 1997; McQuoid & Galef, 1993; Shimizu, 1998). The second approach has used videos of real-world events as the basis for teaching animals visual discriminations based on the video's content (e.g., Dittrich & Lea, 1993; Dittrich, Lea, Barrett, & Gurr, 1998; Jitumori, Natori, & Okuyama, 1999). The third approach has used computer-generated synthetic video stimuli to study how animals visually process dynamic stimuli (e.g., Cook & Katz, 1999; Cook, Shaw, & Blaisdell, 2001; Dittrich & Lea, 1993; Dittrich et al., 1998; Evans, Macedonia, & Marler, 1993). The research in this article used this last approach, as it substantially reduces concerns about whether video technology accurately recreates the true sensory qualities required for playback studies, for example (D'Eath, 1998; D'Eath & Dawkins, 1996; Fleishman, McClintock, D'Eath, Brainards, & Endler, 1998; Jitumori et al., 1999; Patterson-Kane, Nicol, Foster, & Temple, 1997), and concurrently increases the stimulus control for manipulating and controlling the content and presentation of the video information.

Given these advantages, the experiments described below continue our laboratory's exploration of how pigeons process object-based motion (Cook & Katz, 1999; Cook et al., 2001). We were interested in further understanding how pigeons learn to discriminate the action categories of "around" and "through" as depicted by dynamic computer-animated video objects and, in particular, the coherent superiority effect (CSE) discovered by Cook et al. (2001).

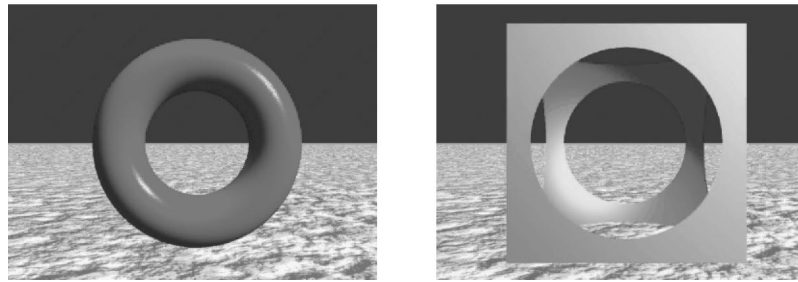
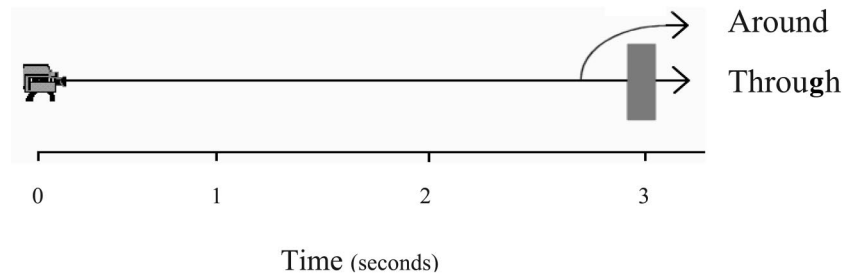
Specifically, Cook et al. (2001) investigated how pigeons discriminate video stimuli that portray the actions of "through" and "around" relative to a number of different objects. Using computer animation software, video stimuli were designed to create a semi-realistic landscape (a textured ground with a clear blue sky and horizon) over which different objects (arches, doughnuts, etc.) appeared to be approaching from the camera's perspective. Near the end of the video, the camera's point of view either veered off to the left of the display, creating the impression of going around the object, or passed through the object's central interior opening. Figure 1 shows the trajectory and timing of these two different motion pathways from that study. They found that pigeons could

---

Robert G. Cook and Shelley Roberts, Department of Psychology, Tufts University.

Shelley Roberts is now at Rogers Wireless, Ottawa, Canada.

Correspondence concerning this article should be addressed to Robert G. Cook, Department of Psychology, Tufts University, Medford, MA 02155. E-mail: robert.cook@tufts.edu



*Figure 1.* Overhead diagram of the two coherent motion pathways tested in these experiments. The lower figures show the two objects used in creating the coherent and randomized video sequences tested in Experiments 1 and 2. A color version of this figure is available on the Web at <http://dx.doi.org/10.1037/0097-7403.33.3.287.sup>

learn this motion discrimination and transfer it to novel objects not previously seen.

To investigate whether this motion discrimination was based on the two- or three-dimensional (2D spatial cues or 3D, respectively) properties of the objects portrayed in the videos, Cook et al. (2001) tested the original videos such that the order of the individual frames was randomly scrambled during their presentation. If the birds were relying on simple 2D cues, these would still be present in the randomized videos, but the 3D object-based properties would be greatly reduced by the chaotic nature of the videos. For example, in the around condition the objects moved off to the right and filled that side of the display, whereas in the through condition they symmetrically filled and moved off to both sides of the screen. If such positional 2D cues were all that controlled the discrimination, then performance in the randomized condition should have matched or come to match that of the coherent baseline conditions because these same spatial cues were present in both arrangements.

If, however, they were seeing the videos as depicting an approaching 3D object, then frame randomization should have disrupted the discrimination by interfering with the coherent processing of these motion and depth signals. That is, the lack of visual coherence across frames would break up the spatial patterns of edge correspondence, common fate, and figure/ground relations associated with smoothly animated objects and disrupt the motion-based 3D cues in the videos. In two different tests, Cook et al. (2001) found that randomization of the videos' frames disrupted transfer and acquisition of their motion discrimination and argued that this was consistent with the hypothesis that the pigeons were interpreting these videos as showing stable 3D objects being approached in depth across an open field.

In this study, we tested several alternative explanations for this CSE. Because Cook et al.'s (2001) frame randomization tests were

done after the pigeons had already had extensive and exclusive training with coherently ordered video stimuli, their randomization effect may plausibly represent a form of generalization decrement related to differential experience. According to this hypothesis, the randomization decrement might have occurred because these scrambled videos looked very different from the extensive experience the birds had already had with coherent videos, and not because they were inherently more difficult to process. A second and closely related hypothesis is that this effect was a sustained form of neophobia related to their differential experience with each type of sequence. A third possibility is that the randomized videos changed the point in time at which the critical discriminative information had previously and reliably appeared in the coherent videos. According to this temporal disruption hypothesis, it was the shift in when the critical frames were expected that reduced performance in the randomized condition. All of these hypotheses, regardless of their exact mechanism, point to the differential experience with coherent and randomized sequences in Cook et al. (2001) as the source of the CSE.

To test these differential experience explanations in the current experiments, we tested naïve pigeons whose experience with coherent and randomized videos was equated over training. In Experiment 1, we tested 5 pigeons in a within-subjects design. Using the same motion discrimination as in Cook et al. (2001), we tested these birds from the beginning with videos that were both coherently and randomly ordered. In Experiment 2, we used a between-groups design to test three new groups of pigeons exclusively trained with either coherent or randomized stimuli. This design prevented any cross-contamination due to experiencing both types of sequences as in the first experiment. In both cases, we were interested in examining the rate of acquisition with each sequence type and the birds' final level of discrimination.

## Experiment 1

The purpose of the first experiment was to examine if the CSE identified by Cook et al. (2001) would occur when exposure to the coherent and randomized video sequences was equated. Five pigeons were trained with both sequence types presented equally often in a within-subjects design in a standard go/no-go visual discrimination.

As mentioned in the introductory paragraphs, another possible difference between the coherent and randomized conditions tested by Cook et al. (2001) is related to the timing of the critical frames. In that experiment, the last 10 frames of each 60-frame sequence video contained the critical information that distinguished the around and through actions. The first 50 frames depicted a common approach to the objects shared by both the around and through conditions. Because the videos were repeated continuously within each 30-s go/no-go presentation, the critical distinguishing frames predictably reappeared in the coherent videos. Thus, those pigeons may have learned to attend to the videos only at certain times. As a result, the frame randomization resulted in moving these critical frames to other points in time, causing the pigeons to be disrupted by this change in the temporal organization of the information. To eliminate this possibility, both the coherent and randomized presentation conditions were tested with two different length videos, making the intervals between the critical frames of the videos unpredictable across trials.

### Method

#### Animals

Five adult pigeons (4 White Carneaux and 1 Silver King) participated in Experiment 1. Each of these pigeons had learned other visual discriminations involving static images unrelated to the dynamic videos tested here. They were maintained at between 80% and 85% of their free-feeding weights. Each was housed separately in a colony room with a 12-hr light/dark cycle with free access to grit and water.

#### Apparatus and Stimulus Generation

Testing was conducted in a flat-black Plexiglas chamber (38 cm wide  $\times$  33 cm deep  $\times$  39 cm high). All stimuli were presented by computer on a color monitor (NEC MultiSync C500, Wooddale, IL) visible through a 29  $\times$  22 cm viewing window in the middle of the front panel of the chamber. A video card controlled the resolution of the monitor at 800  $\times$  600 pixels. The viewing window's bottom edge was 15 cm above the chamber floor. The monitor was moved back 4 cm from the front panel (this distance had been gradually increased during pretraining). Pecks to the monitor screen were detected by an infrared touch screen (originally purchased from Carroll Touch Systems, but now distributed by Elotouch Systems, Fremont, CA) mounted into the front panel. A 28-V house light was located in the ceiling of the box and was illuminated at all times, except when an incorrect response was made. A food hopper was located in the center of the front panel, its access hole flush to the floor. All experimental events were controlled and recorded with a microcomputer using Visual Basic (Microsoft, Redmond, WA) using the FXTools video control component (Pegasus Software, Tampa, FL). The videos were created and rendered in the AVI video format (384  $\times$  288 pixels; Cinepak

Codec compression) using the Bryce 4 animation package (Meta-Creations Corporation, Carpinteria, CA).

*Video sequences.* Two motion pathways were used in discrimination training. These pathways were identical to those used in Cook et al. (2001). The video stimuli for each motion pathway consisted of 60 individual frames. For each pathway type, Frames 1 to 50 were identical and depicted an approaching object from the camera's perspective. Beginning with Frame 51, the two paths from the camera's perspective diverged with respect to the object. In the through condition, the camera's perspective went straight through the central opening of the approaching object over these 10 frames. For the around condition, the camera's perspective went around to the left of the object over these frames (see Figure 1).

Two different objects were used to illustrate each type of motion sequence. These two objects were a red doughnut and a gold cube with a spherical area removed from the center (see Figure 1). The motion toward these objects was rendered with a seminaturalistic context. This consisted of ground and sky that met in the center of the image to create a horizon. The ground was rendered using the "Spotted Clay" material option. This produced a flow field that created the appearance of the camera's motion over this landscape. The sky was the "Caribbean Clear" option and contained no movement cues (e.g., clouds). The sun was placed at an azimuth of 134° with an altitude of 30°, providing surface shading that contributed to the 3D appearance of the objects; however, shadows were not permitted.

*Conditions and timing.* These two motion pathways were presented to the pigeons in two different types of sequences. In the coherent sequence, the order of the frames was sequenced to portray a smooth coherent approach toward the objects. In the randomized sequence, the order of the frames was randomized using Excel to break up the coherence of the objects and their approach. Because there is only one possible coherent sequence by definition, only one randomized sequence was used for testing the randomized condition.

Both the coherent and randomized videos were tested at different lengths. The long duration videos consisted of showing all 60 frames. Each frame lasted 53 ms, resulting in a 3,180-ms video sequence. The short duration videos were constructed using only the last 30 frames of the video, resulting in a 1,590-ms video sequence. The short duration videos removed the majority of the common frames from the around and through videos and increased the frequency of the critical frames containing the differential object movements. For the short duration randomized sequence a different, but fixed, randomized sequence was used to be sure that all 10 critical frames were present.

#### Procedure

The pigeons were first trained to peck all of the stimuli. These pretraining sessions consisted of 96 trials. Each trial began with the presentation of a 2.5-cm white warning signal in the center of the display area. A single peck to this signal caused it to be replaced with a video stimulus. Long duration videos were then presented eight times in immediate succession, and the short durations were presented 16 in succession. The total viewing for each condition was thus a little more than 25 s in duration. During pretraining, pecks at all stimuli were reinforced on a variable interval (VI-10) schedule (gradually increased from the beginning sessions). Once

a bird was responding to all stimuli during pretraining, discrimination testing began. During subsequent discrimination testing, pecks at the S+ videos were reinforced on a variable interval (VI-10) schedule, and pecks to the S- videos were punished with a variable dark timeout (1 s for each peck to the prior S- stimulus). All trials were separated by a 3-s intertrial interval.

The 96 daily trials were counterbalanced across the around versus through pathways, each of the two object types, long and short video duration, and coherent and randomized sequences. Thus, each combination of motion type (around vs. through) object type (Object 1 and Object 2), sequence type (coherent vs. randomized), and duration (short and long) was tested six times within a session. The S+ motion sequence for all 5 pigeons was the around sequence and the S- was the through sequence (see Cook et al., 2001). Eight of the daily S+ trials were presented as nonreinforced probe trials in order to measure S+ peck rates without the contamination of food delivery. These probes tested each combination of object type, sequence type, and duration once each session. On these probe trials, the videos simply ended after the programmed number of loops, with no food delivered at any point in the trial. In scoring the results, only peck rates from these S+ probe trials were used. Experiment 1 consisted of 35 sessions of discrimination training.

### Results

Overall, the coherent sequences were learned faster than the randomized sequences. This was true for 4 of the 5 pigeons, with the fifth bird showing no difference between the two sequence conditions. Figure 2 displays the mean rate of acquisition as a

function of discrimination ratio ( $DR = [S+ \text{ peck rate}/S+ + S- \text{ peck rate}] \times 100$ ; the S+ peck rates were derived from the probe trials) as determined from the last 15 s of each video presentation. This CSE was found for both durations as well. These effects were confirmed in a three-way repeated measures analysis of variance (ANOVA; 5 Session Blocks  $\times$  Duration  $\times$  Sequence Type). This test revealed a significant interaction between sequence type and session,  $F(6, 24) = 6.2$ , reflecting the faster learning of the coherent sequences during acquisition (all statistical tests were evaluated with an alpha level of  $p < .05$ ). There was also a significant main effect of sequence type,  $F(6, 24) = 15.9$ . There was no significant main effect or interactions involving the duration of videos. Further analyses that separately looked at performance with just the coherent or randomized sequences also revealed no effects of duration for either sequence type, although both of these analyses showed significant main effects for sessions,  $F(6, 24) > 6.1$ , as the pigeons did learn to discriminate, albeit at different levels, the videos in both conditions.

Analyses of the last 10 sessions, over which the discrimination appeared to level off, resembled the pattern of effects observed during acquisition. A repeated measures ANOVA (Session  $\times$  Sequence  $\times$  Duration) confirmed that there was no main effect of sessions over this time period,  $F(9, 36) = 0.8$ . There continued to be a significant main effect of sequence type as coherent videos ( $M DR = 78.5$ ) supported better discrimination than did the randomized videos ( $M DR = 68.0$ ),  $F(6, 24) = 10.5$ . Again, there was no significant main effect or interactions involving the duration of the videos over this steady-state period. All four sequence conditions

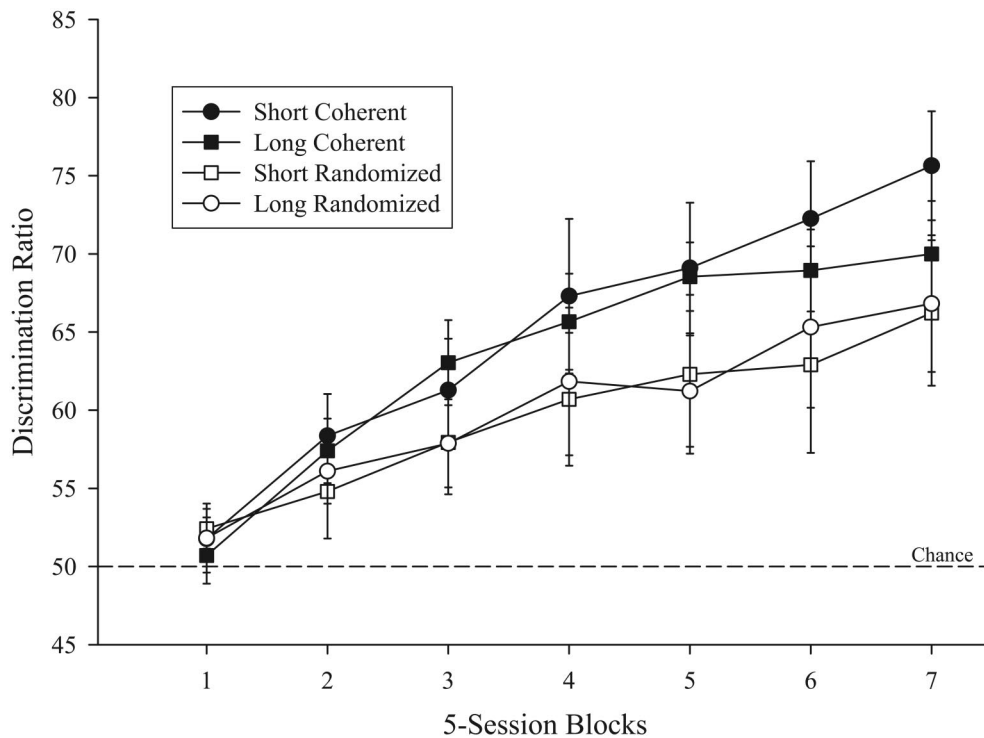
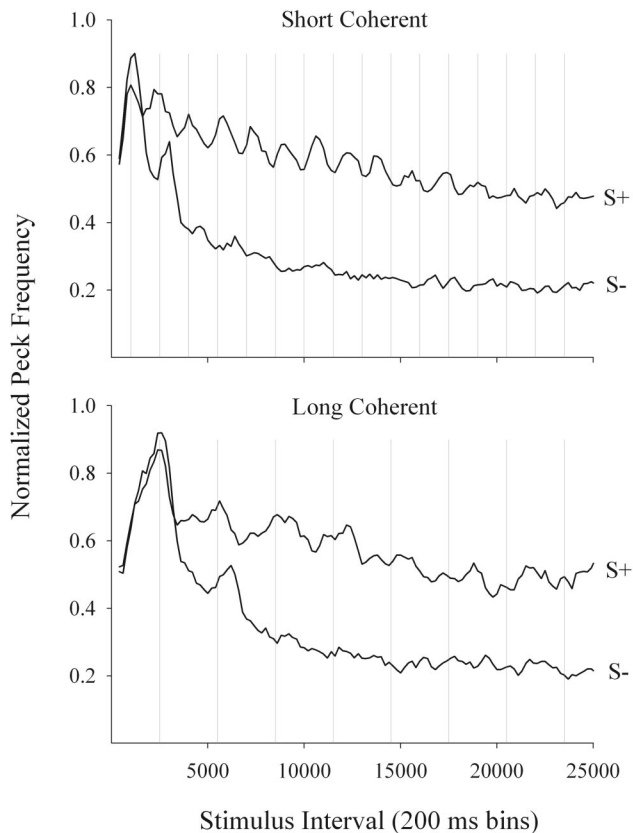


Figure 2. Mean acquisition curves for the coherent and randomized conditions in Experiment 1 as measured by discrimination ratio. Error bars represent the standard error of the mean.

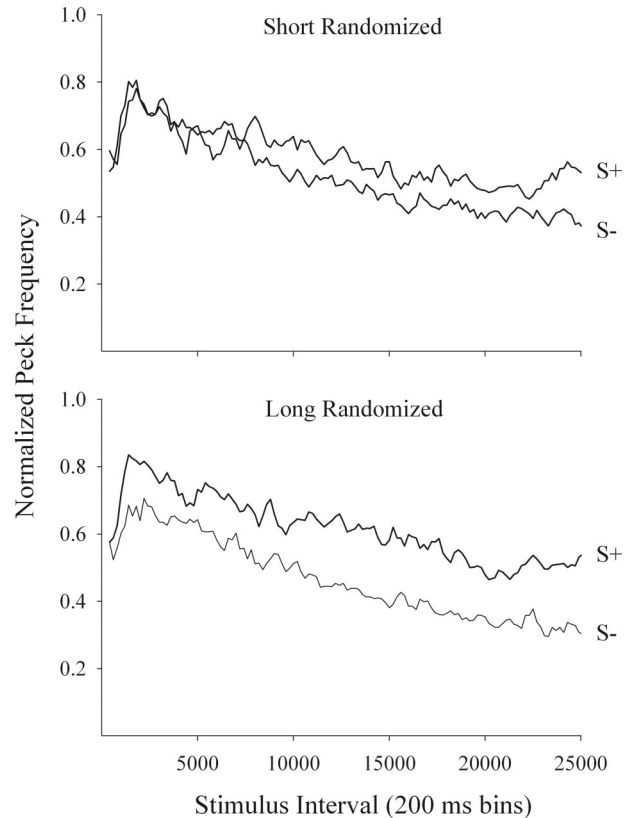
were significantly above the expected value of chance (50) using single mean  $t$  tests: all  $ts(4) > 2.77$ .

We next examined how the different sequence conditions influenced pecking and discrimination within a single trial's presentation. These types of analyses had proven valuable in the past (Cook, Kelly, & Katz, 2003), and we thought they would be revealing here as well. Peck data from the last 10 sessions of training were analyzed. The presentation phase of each trial was divided into 130 bins of 200 ms each, and the number of pecks falling within each of these temporal bins was counted. Figures 3 and 4 show the mean normalized peck frequency as a function of time within a trial for the coherent and randomized video sequences, respectively. Each pigeon's data were normalized relative to the S+ or S- bin with the highest recorded frequency of pecks for that bird. After these normalized frequencies were averaged across pigeons, the resulting curve was smoothed using the method of running means based on the adjacent bins (Tukey, 1977).

The results for the coherent videos, shown in Figure 3, contained several interesting features. First, there was a large and clear



*Figure 3.* Normalized mean peck frequency for S+ and S- trials across the stimulus presentation interval in Experiment 1. Each 200-ms bin shows the relative frequency of pecking occurring at each point in the interval. The top panel shows responding on S+ and S- trials for the short duration coherent sequences. The bottom panel shows responding on S+ and S- trials for the long duration coherent sequences. The light vertical lines in each panel show the temporal location of the critical frames for these videos as they occurred within the presentation interval. See text for details.



*Figure 4.* Normalized mean peck frequency for S+ and S- trials across the stimulus presentation interval in Experiment 1. Each 200-ms bin shows the relative frequency of pecking occurring at each point in the interval. Top: Responding on S+ and S- trials for the short duration randomized sequences. Bottom: Responding on S+ and S- trials for the long duration randomized sequences.

separation between peck rates on S+ and S- trials. A second and more interesting feature was that the pigeons' peck rates varied in a periodic manner tied directly to the video sequence. The top panel of Figure 3 shows S+ and S- peck frequency for the short coherent video presentations with light vertical lines showing the temporal location of the critical frames as they cycled through the presentation interval. As can be seen, changes in the S+ and S- peck rates coincided with the appearance of the critical frames. On S- presentations, the pigeons' peck rates dropped shortly after the appearance of these frames, whereas on S+ presentations the rates generally increased. In both conditions, the ambiguous "approach" frames common to both videos seemed to result in a momentary loss of stimulus control. The majority of the pigeons showed this oscillating pattern. It was exhibited quite strongly by 2 birds, clearly present in the responding of 2 others, but absent in the fifth bird. The long duration videos in the bottom panel of Figure 3 show the same pattern, but it was stretched out because of the 50 ambiguous frames present in these sequences. By comparing the oscillations between the two panels, it is clear that these changes in peck rate were time locked to the videos' contents and were not tied to some inherent cycle related to pecking per se. These oscillations in their peck rates in the coherent conditions suggested

that the pigeons were continually tracking the contents of the entire video and responding appropriately as a function of its momentary properties. The third interesting feature was that despite the multiple repetitions of each video within a presentation, only one cycle of the entire video was necessary for the pigeons to begin discriminating between the two types of motion. In the case of both durations, peck rates between the S+ and S- conditions significantly separated after the pigeon had seen just one set of critical frames. Further repetitions did add information, however, as the S+ and S- peck rates continued to separate.

Examination of the same type of data for the randomized sequences indicated a different pattern of responding on these trials. Consistent with the mean values reported previously, there was separation between the S+ and S- conditions. But this difference was not as large as that seen in the coherent conditions, and moreover this remained true across the entire presentation interval. Furthermore, there was no evidence for the periodicity in peck rates observed with the coherent videos, likely due to the fact that the critical frames were randomly spread out over the entire duration of the video. The absence of any pattern further confirmed that the oscillations in the coherent conditions were under stimulus control and not a byproduct of some inherent rhythm related to pecking behavior.

A series of autocorrelation analyses were conducted looking at the similarity of each curve to itself as it gradually shifted over time. The recurring patterns within each curve were examined by successively displacing the values of each curve by one time bin and computing a Pearson product correlation with the original curve. The results of these analyses are shown in Figure 5. Each line shows the correlation of each of the previous normalized

curves with itself as gradually displaced in successive 200-ms bins. Each line represents the combined calculation for both separately computed positive and negative functions. For the two coherent conditions, the systematically oscillating correlations with time reflect the repetitive sinusoidal structure in these data. Thus, as the coherent curves were temporally displaced, their periodicity produced regular changes in these correlations as each went in and out of phase with itself. For the short condition this period was near 1,500 ms, and for the long condition it was a little more than 3,200 ms as would be expected from the duration of the videos used in these conditions. The randomized condition produced no such pattern in this analysis with no strong oscillating features, as these linear curves generally matched themselves regardless of displacement.

Finally, one last interesting difference was that the randomized sequences were discriminated slightly earlier in their presentation than were the coherent sequences. Examination of the peck rates over the first 2,500 ms revealed a slight separation in S+ and S- pecks in the randomized condition, whereas the pigeons in the coherent conditions were still waiting on the first arrival of the critical frames. Nevertheless, despite the randomized condition's head start, when the critical frames arrived in the coherent condition the pigeons rapidly discriminated these videos and maintained this advantage for the duration of the trial.

### Discussion

The most important finding of Experiment 1 was the faster acquisition and steady-state superiority of the coherent video condition in comparison to the randomized condition. This CSE replicates and extends the findings reported by Cook et al. (2001) in

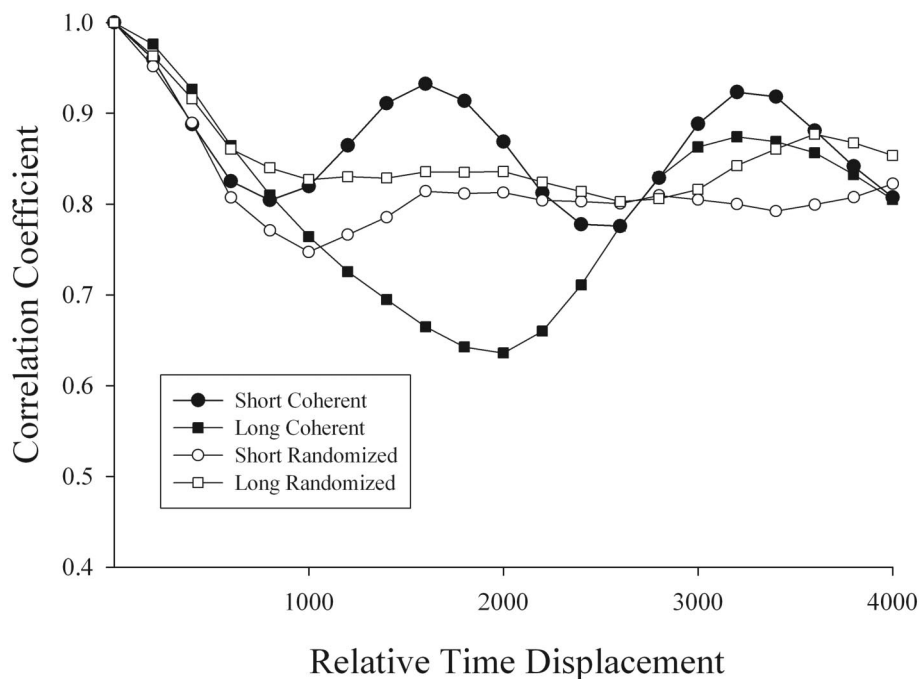


Figure 5. Autocorrelation functions for the four conditions depicted in Figures 3 and 4. Each line shows the correlation of each of the previous curves with itself as one is gradually displaced in successive 200-ms bins. Each line represents the combined calculation for both the positive and negative function for each condition.

two ways. First, it determines that the CSE does not depend on the differential experience with coherent and randomized videos (dismissing, for the moment, the contributions of daily reality). Provided with equivalent experience with both types of sequences, the pigeons still showed the effect. Second, the CSE does not depend directly on the consistent timing of the coherent condition. When tested with two different durations that increased the unpredictability of the critical frames in the coherent videos, the pigeons still showed the effect. These observations eliminate the major alternative accounts of the CSE; the results are thus consistent with the hypotheses that the coherence of the object-based visual information in the video makes a significant contribution to the pigeons' processing of these videos. This interpretation is bolstered by other new observations as well.

First, the examination of the peck rates across a presentation suggest that pigeons are constantly tracking the contents of the coherent videos on a moment-to-moment basis. The cyclic up-and-down pattern to both S+ and S- responding and its time-locked control, depending on the duration of the video, suggest that the pigeons were varying their pecking behavior depending upon what they were seeing. During the ambiguous approach to the object, there was a reduction in stimulus control (increased S- and decreased S+ peck rates), but right after the passage of the critical through and around frames the degree of stimulus control increased (decreased S- and increased S+ peck rates). In addition to demonstrating the constant tracking of the video's content by the pigeons, these data further reveal that significant stimulus control emerged after only one video cycle of the coherent videos. This suggests that the visual organization of these 10 frames was particularly effective. In fact, it is interesting just how good the overall discrimination actually was given that five sixths of the entire presentation interval (> 20 s) provided no information distinguishing between the S+ and S- motions. Thus, despite the spaced, unpredictable, and infrequent nature of the critical frames in the coherent videos, they were still quite effective in promoting a strong degree of stimulus control.

Despite the clear benefits of coherent movement to the discrimination, the randomized videos were capable of supporting the discrimination, albeit at a lower level of performance. This suggests that discriminative cues unrelated to visual coherence were still present. In several personal communications, colleagues have pointed out that some residual "movement" is still apparent in randomized videos. Although possible, we believe a more conservative interpretation is that 2D cues based on the relative position of the objects in the frame are responsible for the discrimination with randomized sequences. Examination of the peck rate data within a trial is consistent with this hypothesis. The more frequent and distributed nature of the randomized critical frames throughout the videos supports an earlier discrimination in these videos. Over the first 2 s or so, there is a separation between S+ and S- peck rates in the randomized condition that cannot occur in the coherent videos because the critical frames have not yet been presented. Despite the continuous and distributed nature of this randomized information, however, the pigeons did not dramatically increase in their discrimination over the course of the presentation, unlike in the coherent condition. These differences suggest that there are additional sources of information in the coherent videos that are not present in or cannot be processed from the randomized sequences.

## Experiment 2

The purpose of the second experiment was to examine if video coherence could produce the same superiority effect as in Experiment 1, but using a between-groups design. Because each pigeon experienced both types of video sequences in Experiment 1, it was impossible to tell how exposure to each type influenced the other. For instance, seeing and discriminating the coherent videos may have helped the pigeons to eventually interpret and learn to discriminate the position of the objects in the randomized sequences. Vice versa, the chaotic organization of the randomized sequences may have interfered with accurate perception and acquisition of the coherent videos. Either way, to prevent such cross-contamination, separate groups of pigeons were tested in each condition.

In addition, we also wanted to add a second type of randomized group. One possible difference in Experiment 1 was that the critical frames of the coherent condition were grouped together in a single block, whereas for the randomized condition these frames were spread out over the entire duration of the video. Thus, this difference in the temporal grouping of the key frames may have contributed to or accounted for the CSE observed in Experiment 1. To examine this issue, we added a second randomized group for which the 50 common frames were randomized with respect to one another, but the 10 critical frames were randomized within a single temporal block similar to in the coherent videos. If temporal grouping per se was the source of the CSE, then this second type of randomized group should be able to perform comparably to the coherent group. If video coherence was the key factor, then the birds seeing this second type of randomized sequence should perform in a manner more similar to the completely randomized group. Thus, in the experiment one group of 5 birds experienced only the two coherent sequences, a second group of 5 birds experienced two randomized sequences with the key frames temporally grouped together, and a third group of 3 birds experienced the two completely randomized sequences.

## Method

### *Animals and Apparatus*

Thirteen naïve adult pigeons participated in Experiment 2 and were kept at approximately 80%–85% of their free-feeding weights. The apparatus was the same as in Experiment 1.

### *Procedure*

The pretraining and discrimination phases of this experiment were conducted identically to those in Experiment 1, except for the use of the between-groups design. The coherent group received only the short and long coherent sequences tested in Experiment 1. The contiguous-randomized group received short and long randomized sequences in which the 10 critical frames were grouped together but randomized within this block. The remaining 20 or 50 frames were also randomized. The completely randomized group received short and long randomized sequences, in which all frames were randomly placed in a sequence similar to that tested in Experiment 1. Again, only one sequence was used with each condition. Each group received 96 total trials per session (48 S+

and 48 S-), with equal numbers of both durations tested. Experiment 2 was conducted for 50 sessions.

*Results*

Figure 6 displays the mean rate of acquisition for the three groups of pigeons in five-session blocks. Once again, DR was computed using peck rates over the last 15 s of each presentation. As in Experiment 1, the coherent sequences supported faster discrimination learning and higher steady-state performance than either type of randomized sequence. These differences were confirmed by a series of mixed design ANOVAs (Groups  $\times$  5-Session Block). Comparison of the coherent group with the completely randomized group revealed significant main effects for both group,  $F(1, 6) = 13.7$ , and session,  $F(9, 54) = 12.2$ . This replicated the CSE established with the within-subjects design in Experiment 1. Comparison of the completely randomized and contiguous ran-

domized groups revealed no significant differences between these groups,  $F(1, 6) = 0.44$ , but did confirm that learning occurred in both groups as there was a significance of sessions,  $F(9, 54) = 75.4$ .

Despite the apparent difference in acquisition in Figure 6, the initial comparison of the coherent and contiguous-randomized groups failed to reveal a significant main effect of group or its interaction across sessions. Examination of the individual data for the contiguous-randomized group revealed the reason for this conflict. This group was a composite of 3 pigeons that learned as slowly as any animal tested with the completely randomized condition, 1 animal that learned at a rate equal that of the poorest pigeon in the coherent group, and 1 bird that learned at a rate near the best performing pigeon in the coherent group. The latter bird's widely divergent performance resulted in considerable variance in our inferential tests. Subsequent analyses in which we removed

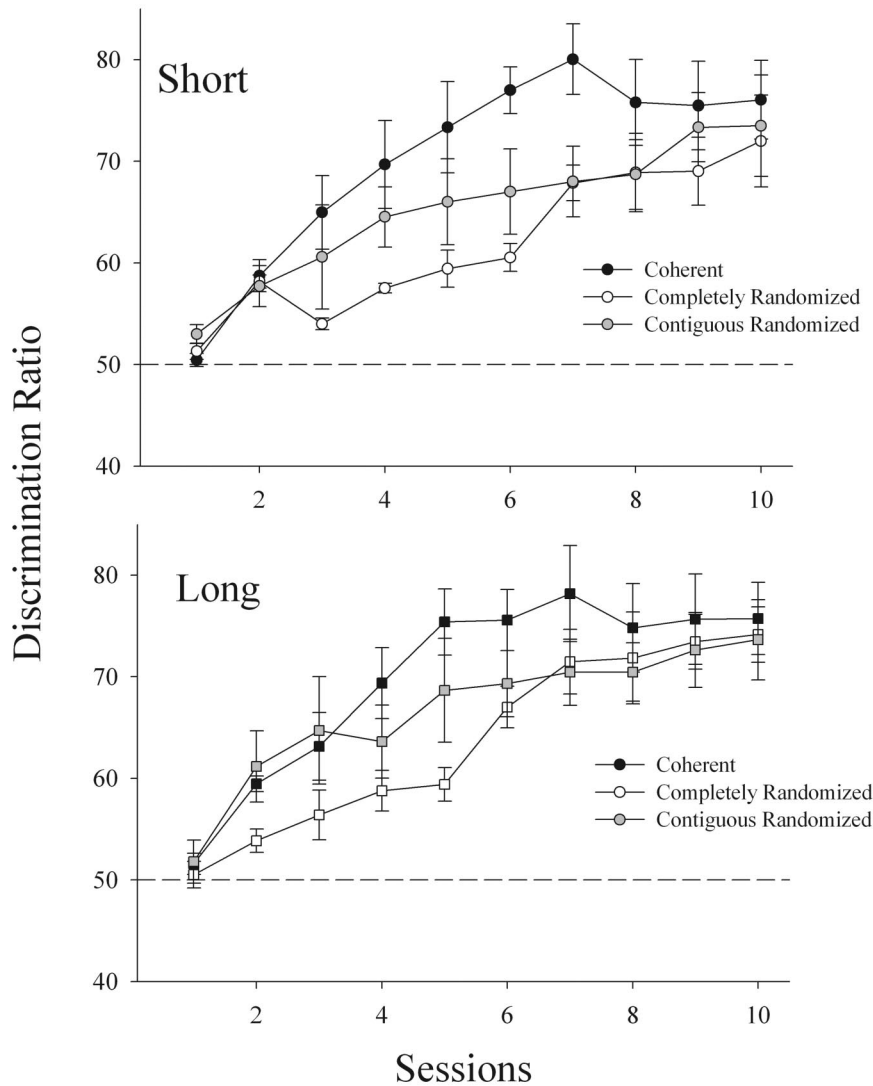


Figure 6. Mean acquisition curves for the coherent, contiguous-, and completely randomized groups in Experiment 2 as measured by discrimination ratio. Error bars represent the standard error of the mean.



this one uncharacteristic bird revealed a significant group difference between the coherent and the contiguous-randomized groups,  $F(1, 7) = 5.5$ . There was also a significant main effect of sessions,  $F(9, 63) = 13.5$ . These results, therefore, suggested that the same CSE generally existed for birds that experienced randomized sequences using this contiguous organization as when the frames were completely randomized, with the standing caveat that 1 bird was not disturbed by this manipulation in the least.

Separate two-way repeated measures ANOVA (Session  $\times$  Duration) of the separate groups further revealed that only the completely randomized group showed an effect of video length across sessions. This was confirmed by the presence of a significant Session  $\times$  Duration interaction,  $F(9, 18) = 3.8$ . This was due to these birds doing better with long randomized sequences than with the shorter duration sequences, especially over the second half of acquisition. No comparable effect of duration was found in the coherent and contiguous-randomized groups, as neither the main effect of duration nor its interaction with session was significant in their analyses.

We again examined how the different sequence conditions influenced pecking behavior across a single presentation. Peck data from the last 25 sessions of training were analyzed using the same procedures as in Experiment 1. Figures 7, 8, and 9 show the mean normalized peck frequency as a function of time for the coherent, completely randomized, and contiguous-randomized video sequences, respectively. The results for the coherent and completely randomized groups replicated those in Experiment 1. The coherent group showed a strong time-locked oscillation in their peck rate correlated with the video's content and critical frames (again depicted by vertical lines). This was observed on both positive and negative trials and was detectable in all 5 pigeons in the coherent group. Like in Experiment 1, the completely randomized group showed no comparable structure in their pattern of responding over a trial. The contiguous-randomized group's results looked like those of the completely randomized group. Despite a temporal pattern of critical frames similar to the coherent group's, no strongly time-locked structure emerged in this group's data. The weak patterning visible on the negative trials was attributable to the 1 bird in this group that did exceptionally well. This bird did exhibit some temporal control on its negative trials, but no comparable organization on its positive trials, unlike the pattern observed with the coherent group.

For both randomized groups, the pattern of discrimination between S+ and S- trials was more gradual in nature, with a growing separation accumulating across time. For the coherent group, the pattern of responding was far more discrete, with a clear separation in peck rate emerging after just one presentation and the birds reaching their asymptotic level of discrimination much earlier, by several seconds, than either randomized group. One difference from the first experiment was that the randomized group did not show the same earlier onset of discrimination. The outcomes of the autocorrelation analyses of these curves are shown in Figure 10. The coherent conditions again produced strong oscillating correlations reflecting their repetitive structure, with peak correlations again near the expected values based on the duration of the videos. The two randomized groups again exhibited little structure in these functions, as each linear curve correlated well with itself regardless of its displacement.

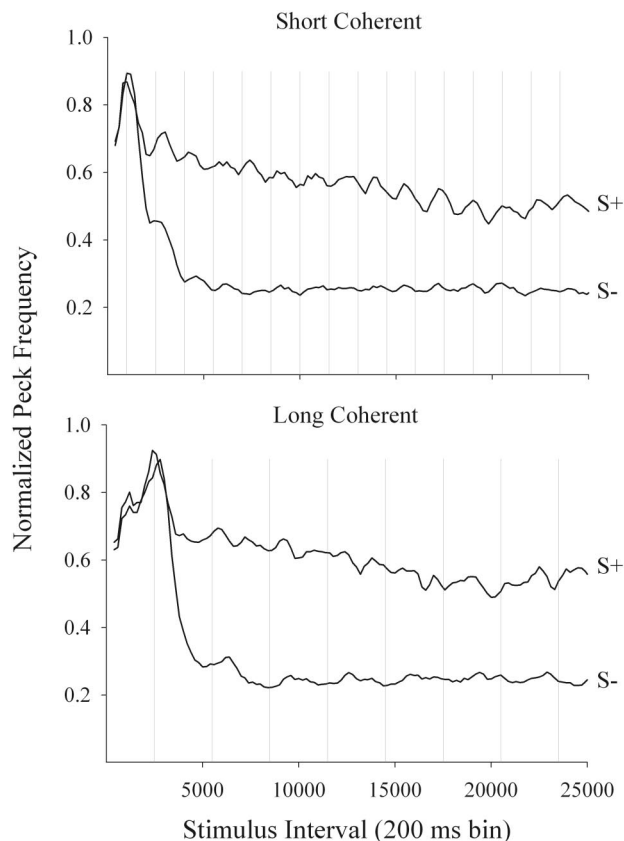
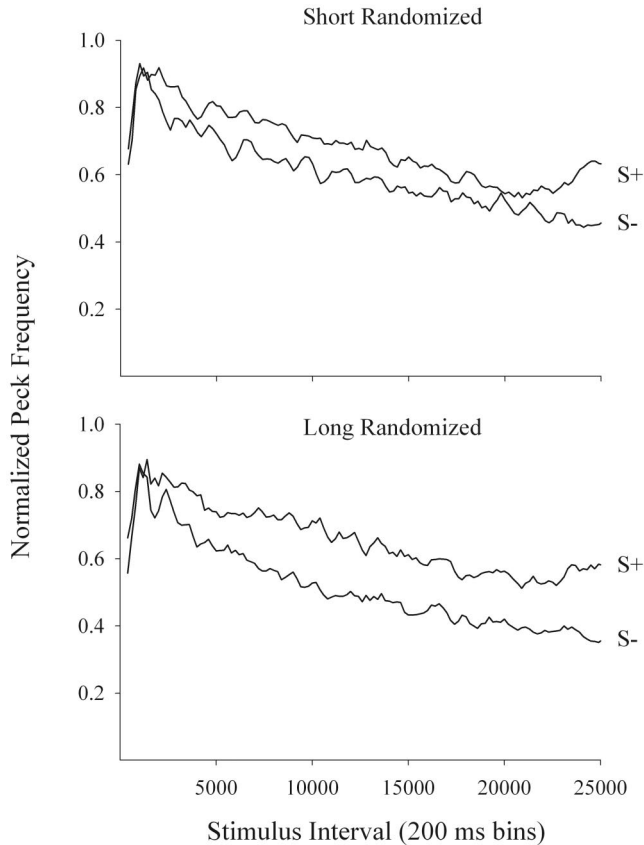


Figure 7. Normalized mean peck frequency for S+ and S- trials across the stimulus presentation interval for the coherent group in Experiment 2. Each 200-ms bin shows the relative frequency of pecking occurring at each point in the interval. The top panel shows responding on S+ and S- trials for the short duration coherent sequences. The bottom panel shows responding on S+ and S- trials for the long duration coherent sequences. The light vertical lines in each panel show the temporal location of the critical frames for these videos as they occurred within the presentation interval.

### Discussion

Experiment 2 replicated and extended the CSE found in Experiment 1 and documents that the CSE is not dependent on the simultaneous experience of both types of video sequences. Furthermore, the learning and steady-state differences between the coherent group and contiguous-randomized group in their overall rate and pattern of pecking behavior indicates that the temporal proximity of the critical frames is not the sole source of the coherent group's advantage, not withstanding 1 pigeon. For the coherent videos, one video cycle was again sufficient to show a clear separation between the S+ and S- peck rates, unlike the gradual pattern exhibited by the randomized groups. When comparing the difference in performance between the random and coherent sequences of Experiments 1 and 2, there appears to be a slightly greater CSE when the pigeons experienced only one of the two sequence types as in Experiment 2. Again, the randomized groups acquired the discrimination. This indicates that the acquisition seen with the randomized condition in Experiment 1 is not being mediated by simultaneous experience with the coherent videos, but can develop solely on its own.



*Figure 8.* Normalized mean peck frequency for S+ and S- trials across the stimulus presentation interval for the completely randomized group in Experiment 2. Each 200-ms bin shows the relative frequency of pecking occurring at each point in the interval. Top: Responding on S+ and S- trials for the short duration completely randomized sequences. Bottom: Responding on S+ and S- trials for the long duration completely randomized sequences.

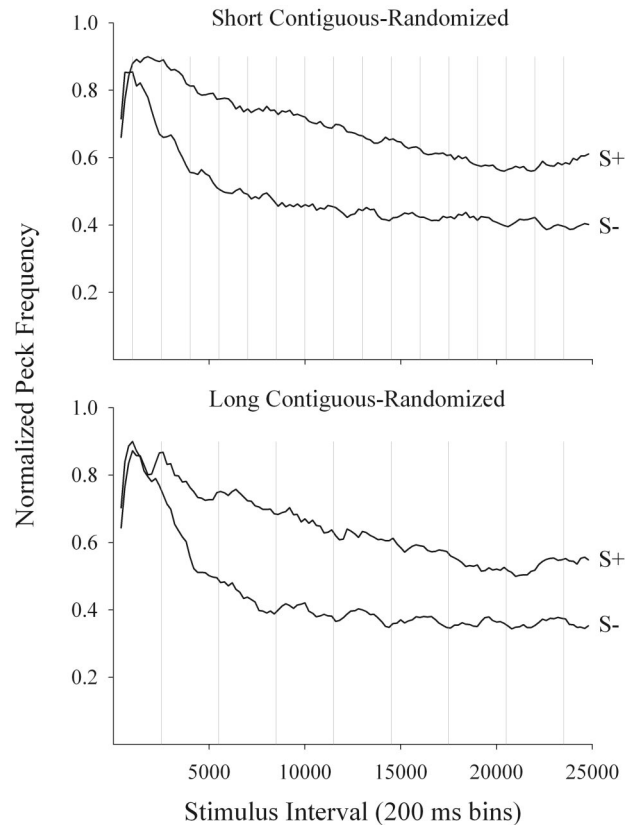
### General Discussion

Both experiments established the strong presence of a CSE. That is, coherently organized motion discriminations are learned faster and to a higher level of discrimination than randomly scrambled versions of the same video information. Improving on the experiments of Cook et al. (2001), the most important and new contribution of the present experiments is that this CSE is present even when experience with the two sequence conditions is equated regardless of the type of experimental design employed. Because of the mixture of multiple video durations in the present studies, these experiments further establish that the CSE does not depend on the predictable timing of the critical frames. The difference between the coherent and contiguous-randomized groups in Experiment 2 also establishes that the temporal grouping of the critical frames is not the explanation for the CSE. Thus, even when experience, temporal predictability and arrangement, and the simultaneous experience of both types of sequences (Experiment 1 vs. Experiment 2) are controlled for, the CSE remains. What is the source of this generalized superiority for coherent presentations?

Consider first performance with the randomized sequences. Unlike in Cook et al. (2001), the randomized videos in the present

experiments did support above-chance discrimination. This above-chance discrimination indicates that the individual frames are capable of supporting discrimination and suggests the likely presence of 2D spatial cues. The most probable source of such 2D cues in the present motion discrimination would be the central and displaced locations of the objects relative to the videos' surrounding frame. Such 2D cues apparently do not need to be coherently organized or last very long for the pigeons to form some basis for using them to discriminate the putative around and through conditions, and, at least in one case, they didn't need to be particularly organized either.

Nevertheless, the superiority of the coherently organized presentations in each experiment requires that some additional factor needs to be included to explain the pigeons' consistently better identification of the critical discriminative information in such video organizations. Because differential experience, temporal predictability, and arrangement are ruled out by the present results, the most obvious and commonsense candidate for this additional factor would be the organized object-based motion powerfully created by these stimuli (at least to the human eye).



*Figure 9.* Normalized mean peck frequency for S+ and S- trials across the stimulus presentation interval for the contiguous-randomized group in Experiment 2. Each 200-ms bin shows the relative frequency of pecking occurring at each point in the interval. The top panel shows responding on S+ and S- trials for the short duration contiguous-randomized sequences. The bottom panel shows responding on S+ and S- trials for the long duration contiguous-randomized sequences. The light vertical lines in each panel show the temporal location of the critical frames for these videos as they occurred within the presentation interval.

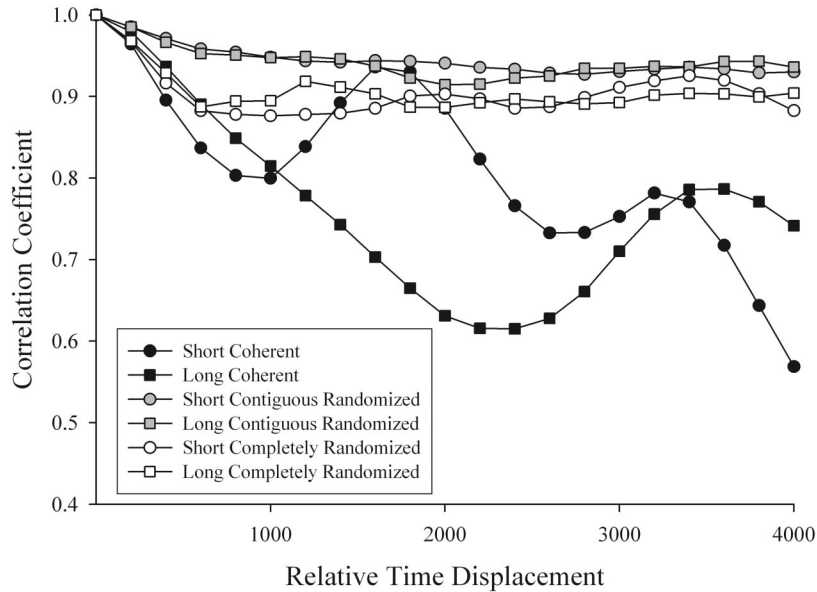


Figure 10. Autocorrelation functions for the four conditions depicted in Figures 7, 8, and 9. Each line shows the correlation of each of these curves with itself as one is gradually displaced in successive 200-ms bins. Each line represents the combined calculation for both the positive and negative function for each condition.

Events always unfold over time. The continuous and constant relative movement of objects (predators, prey, conspecifics, or obstacles) and observers in nature requires the visual system to integrate motion signals from perceptual information that is spaced over time. Because the receptive fields of motion-sensitive cells process only a part of the visual field (Marr, 1982), global integrative processes are needed to disambiguate and group these different motion signals into a coherent percept. The human visual system's capacity to do this effortlessly is in large part responsible for the phenomenological stability of the world. Possessing a similar capacity to integrate motion signals over time and space would seem equally valuable and adaptive for birds, despite the computational demands on their small nervous systems. In the present case, such a capacity to integrate and recognize the relative motion of going around or through objects would provide a seamless explanation of the CSE. That is, when presented with smoothly organized moving sequences of information, the pigeons' motion system integrates this information across frames to produce an extended visual event that can be discriminated over and above the static spatial cues provided by the individual frames. Correspondingly, the randomization of the frames interferes with the operation of this integration process. Consistent with this hypothesis is the fact that the pigeons seem to be continually processing and integrating motion information from these coherent videos, as reflected by the time-locked oscillation in peck rates recorded during these presentations.

The available data on the detection of coherent and apparent motion in pigeons would indicate that the present videos would be fully capable of producing such integrated motion (Bischof, Reid, Wylie, & Spetch, 1999; Siegel, 1970). Bischof et al., for instance, determined the thresholds for integrating spatially distributed dots into coherent directional motion. They found that pigeons were consistently poorer than humans at making such discriminations at

almost all values tested. Nevertheless, when tested with longer dot lifetimes and larger proportions of the dots moving coherently, the pigeons were consistently well above chance. Both of these conditions were met in the current videos. First, a vast proportion of the object information coherently moved small distances from frame to frame. When the moving ground and objects are taken into account, more than 50% of the video changes in a systematic spatial fashion across frames. This proportion is even greater as the objects reach the critical frames later in the videos. Second, the duration over which specific portions of the video could be consistently identified (object edges, bumps on the ground) were within the temporal range (1,590 to 3,180 ms) of the dot lifetimes tested in Bischof et al.'s (83 to 3,320 ms) random dot displays. Thus, these data fit nicely within these established parameters and are consistent with the idea that the extended existence of large identifiable units across time are important conditions for pigeons to see the motion of simulated objects like those tested here.

Now in humans, detecting and establishing the successive correspondence between the edges of an object is a key step in the integration process. Although the CSE was present for both objects tested here, the hollowed-out gold square generally supported a larger CSE than did the red doughnut object. One possible reason for this difference is that that particular object has more visible and interleaved edges that require and promote alignment, while creating more ambiguity as to its structure and location in the randomized condition. Future research directions should examine if pigeons use edge correspondence algorithms or other means to integrate object-based information across frames in order to produce smooth motion and object stability (Dawson, 1991). Furthermore, it would be interesting to reexamine the motion coherence thresholds for stimuli similar to the current ones. Although the pigeons in Bischof et al. (1999) performed consistently more poorly than the humans, their task did require integrating spatially

separated dots into directional motion. This task may not be as biologically relevant as the collision-like object-based discrimination tested here.

Another consideration concerns which part of the pigeon visual system may have mediated the coherent motion of the videos. Given current understanding, it is difficult to know at the moment. All three major divisions of the pigeon visual system (thalamofugal, tectofugal, and accessory optic system) have been found to contain motion-sensitive cells. Because the accessory optic system responds primarily to large-scale signals moving at lower velocities, it has been suggested that it is involved in detecting optic flow for the purposes of determining self-motion rather than object motion, which may be detected by the tectofugal system (Frost, 1985; Frost, Wylie, & Wang, 1990). Because the current videos combined both a considerable degree of optic flow with object-based motion, both factors may have played a role in mediating the entire discrimination. One important use for stimuli similar to the current videos will be to isolate the various contributions and types of information used by these different pathways in supporting motion and object discrimination and their indispensable combination.

While there is strong evidence from the current experiments that video coherence contributed significantly to the discrimination of these videos, a natural question to ask is whether the pigeons actually did see the content of these videos as intended; that is, as approaching 3D objects depicting different types of actions (around and through). No direct evidence from these experiments demands that the pigeons saw these videos as having depth (Cavoto & Cook, 2006; Reid & Spetch, 1998) or as containing objects. The present research does indicate that when presented appropriately, detailed visual information rendered and organized to portray exactly such events greatly benefits a discrimination nominally based on such events. Thus, the present CSE is consistent with the general proposition that pigeons, and birds in general, may experience a stable and object-filled world similar to that of humans.

Overall, the addition of motion to our previous overreliance on static images is an important step forward in the analysis of animal behavior. As always, appropriate analyses and controls are needed to deal with the increased complexity of the additional processes invoked by this type of stimulus. As video presentations are increasingly used in a variety of animal behavior settings, the current randomization procedure seems like a potentially important control condition for comparing with the standard forward and coherent presentation condition ubiquitously tested in such experiments. Interfering with the normal operation of the motion integration system can help to isolate and identify exactly what 2D/3D features and motion cues are being used when an animal responds to dynamic stimulation.

## References

- Adret, P. (1997). Discrimination of video images by zebra finches (*Taeniopygia guttata*): Direct evidence from song performance. *Journal of Comparative Psychology, 111*, 115–125.
- Bischof, W. F., Reid, S. L., Wylie, D. R. W., & Spetch, M. L. (1999). Perception of coherent motion in random dot displays by pigeons and humans. *Perception & Psychophysics, 61*, 1089–1101.
- Cavoto, B. R., & Cook, R. G. (2006). The contribution of monocular depth cues to scene perception by pigeons. *Psychological Science, 17*, 628–634.
- Cook, R. G., & Katz, J. S. (1999). Dynamic object perception by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 25*, 194–210.
- Cook, R. G., Kelly, D. M., & Katz, J. S. (2003). Successive two-item same-different discrimination and concept learning by pigeons. *Behavioural Processes, 62*, 125–144.
- Cook, R. G., Shaw, R., & Blaisdell, A. P. (2001). Dynamic object perception by pigeons: Discrimination of action in video presentations. *Animal Cognition, 4*, 137–146.
- Dawson, M. R. (1991). The how and why of what went where in apparent motion: Modeling solutions to the motion correspondence problem. *Psychological Review, 98*, 569–603.
- D'Eath, R. B. (1998). Can video images imitate real stimuli in animal behaviour experiments? *Biological Reviews of the Cambridge Philosophical Society, 73*, 267–292.
- D'Eath, R. B., & Dawkins, M. S. (1996). Laying hens do not discriminate between video images of conspecifics. *Animal Behaviour, 52*, 903–912.
- Dittrich, W. H., & Lea, S. E. G. (1993). Motion as a natural category for pigeons: Generalization and a feature-positive effect. *Journal of the Experimental Analysis of Behavior, 59*, 115–129.
- Dittrich, W. H., & Lea, S. E. G. (2001). Motion discrimination and recognition. In R. G. Cook (Ed.), *Avian visual cognition*. Retrieved May 21, 2007, from <http://www.pigeon.psy.tufts.edu/avc/dittrich/>
- Dittrich, W. H., Lea, S. E. G., Barrett, J., & Gurr, P. R. (1998). Categorization of natural movements by pigeons: Visual concept discrimination and biological motion. *Journal of the Experimental Analysis of Behavior, 70*, 281–299.
- Evans, C. S., Macedonia, J. M., & Marler, P. (1993). Effects of apparent size and speed on the response of chickens, *Gallus gallus*, to computer-generated simulations of aerial predators. *Animal Behaviour, 46*, 1–11.
- Fleishman, L. J., McClintock, W. J., D'Eath, R. B., Brainards, D. H., & Endler, J. A. (1998). Colour perception and the use of video playback experiments in animal behaviour. *Animal Behaviour, 56*, 1035–1040.
- Frost, B. J. (1985). Neural mechanisms for detecting object motion and figure-ground boundaries contrasted with self-motion detecting systems. In D. Ingle, M. Jeannerod, & D. Lee (Eds.), *Brain mechanisms and spatial vision* (pp. 415–419). Dordrecht, The Netherlands: Nijhoff.
- Frost, B. J., Wylie, D. R. W., & Wang, Y.-C. (1990). The processing of object and self-motion in the tectofugal and accessory optic pathways of birds. *Vision Research, 30*, 1677–1688.
- Jitumori, M., Natori, M., & Okuyama, K. (1999). Recognition of moving video images of conspecifics by pigeons: Effects of individuals, static and dynamic motion cues, and movement. *Animal Learning & Behavior, 27*, 303–315.
- Marr, D. (1982). *Vision*. San Francisco: Freeman.
- McQuoid, L. M., & Galef, B. G., Jr. (1993). Social stimuli influencing feeding behaviour of Burmese fowl: A video analysis. *Animal Behavior, 46*, 13–22.
- Patterson-Kane, E., Nicol, C. J., Foster, T. M., & Temple, W. (1997). Limited perception of video images by domestic hens. *Animal Behavior, 53*, 951–963.
- Reid, S., & Spetch, M. L. (1998). Perception of pictorial depth cues by pigeons. *Psychonomic Bulletin & Review, 5*, 698–704.
- Shimizu, T. (1998). Conspecific recognition in pigeons (*Columba livia*) using dynamic video images. *Behaviour, 135*, 43–53.
- Siegel, R. K. (1970). Apparent motion detection in the pigeon. *Journal of the Experimental Analysis of Behavior, 14*, 93–97.
- Tukey, J. (1977). *Exploratory data analysis*. Boston: Addison-Wesley.

Received September 16, 2003

Revision received February 28, 2006

Accepted March 17, 2006 ■