

Robert G. Cook · Robert Shaw · Aaron P. Blaisdell

Dynamic object perception by pigeons: discrimination of action in video presentations

Received: 30 August 2000 / Accepted after revision: 23 March 2001 / Published online: 28 August 2001

© Springer-Verlag 2001

Abstract. Two experiments examined the discrimination by pigeons of relative motion using computer-generated video stimuli. Using a go/no-go procedure, pigeons were tested with video stimuli in which the camera's perspective went either "around" or "through" an approaching object in a semi-realistic context. Experiment 1 found that pigeons could learn this discrimination and transfer it to videos composed from novel objects. Experiment 2 found that the order of the video's frames was critical to the discrimination of the videos. We hypothesize that the pigeons perceived a three-dimensional representation of the objects and the camera's relative motion and used this as the primary basis for discrimination. It is proposed that the pigeons might be able to form generalized natural categories for the different kinds of motions portrayed in the videos.

Key words Object perception · Motion perception · Visual cognition · Discrimination learning · Pigeons

Introduction

Mobile diurnal animals, such as birds and mammals, live in a highly dynamic world of visual events. These animals navigate and move through a complex landscape of moving and stationary objects. Add the dimension of flight and one might expect birds to be especially adept at understanding their own motion and the relative motion of surrounding objects. Accurately perceiving and recognizing objects and

their associated behaviors, whether they are predators, conspecifics, or obstacles, is critical to interacting with them appropriately (e.g., approaching, avoiding, courting). Any casual observation of birds certainly suggests these animals might experience an object-filled dynamic world much like our own. The experiments described in this article continue our laboratory's recent exploration of this thorny and difficult psychological issue (Cook 2000; Cook and Katz 1999). In the current research, we were specifically interested in whether pigeons could learn to discriminate the action categories of around and through as portrayed with moving computer-animated objects as presented in a semi-naturalistic video setting.

For too long, the psychological study of animal behavior and discrimination learning has concentrated primarily on examining stationary visual stimuli. Recent advances in video and computer animation technology, however, have slowly begun to change this situation (see a recent review by Lea and Ditttrich 1999). Such video technology now allows us to present to animals extended sequences of images that can more realistically recreate the temporal and spatial properties of the real world. Using this technology, three different approaches have emerged in studying bird behavior with video stimuli.

The first approach uses video presentations of actual behavior as playback stimuli, studying the degree to which these stimuli elicit appropriate behavior from different animals. In birds, such techniques have elicited several types of species-typical behavior, suggesting that these video stimuli contain features similar to those normally encountered in the real world (Adret 1997; Evans and Marler 1991; McQuoid and Galef 1993; Shimizu 1998). Some of these studies have also documented a direct contribution of motion to such behaviors. Shimizu (1998) documented, for example, that male pigeons will court a video of a female pigeon longer when the video is presented in motion than as a still image (although the latter still elicited courtship displays).

The second approach uses videos of real behaviors as the basis for teaching animals visual discriminations based on the video's content. This kind of technique has revealed,

Electronic supplementary material to this paper can be obtained by using the Springer Link server located at <http://dx.doi.org/10.1007/s100710100097>.

R.G. Cook (✉) · R. Shaw · A.P. Blaisdell
Department of Psychology, 530 Boston Ave, Tufts University,
Medford, MA 02155, USA
e-mail: rcook1@emerald.tufts.edu,
Tel.: +1-617-6272548, Fax: +1-617-6274132

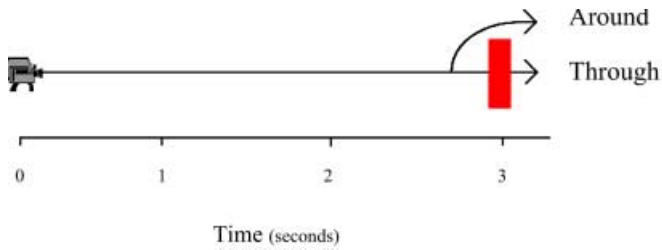


Fig. 1 Overhead diagram of the two motion pathways tested in these experiments

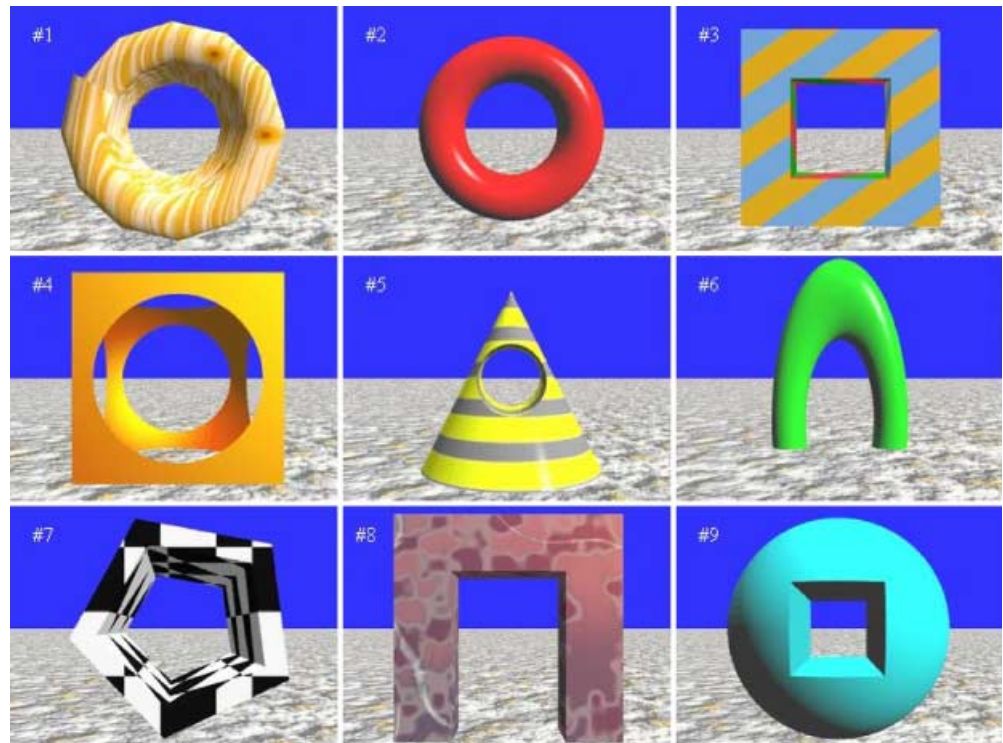
for instance, that pigeons can discriminate static from moving images of the same video (Dittrich and Lea 1993), and that they can categorize the behavioral actions of conspecifics (Dittrich et al. 1998; Jitsumori et al. 1999). For instance, in Dittrich et al. (1998), pigeons learned to discriminate between videos of other pigeons that were either “walking” or “pecking.”

These two approaches have been fruitful in beginning to reveal how stimulus and object motion contribute to stimulus recognition in animal behavior. Despite their success, concerns have been expressed about these approaches. The first concern centers on whether video technology engineered for the human eye and brain accurately portrays the sensory qualities of the real-world stimuli for the species being tested (D’Eath 1998; D’Eath and Dawkins 1996; Jitsumori et al. 1999; Patterson-Kane et al. 1997). Given that pigeon color vision is at least tetrachromatic, “realistic” color videos for the human eye may not appear the same way to pigeons. The second concern stems from the difficulty in controlling the pictorial content of the video stimuli directly taped or captured from living, be-

having animals (a concern often expressed about static pictorial stimuli as well). That is, do these videos only contain the information deemed relevant by the experimenter or are there confounded extraneous features that might influence or control an animal’s reactions? Without the capacity to isolate, manipulate, and identify the many features and properties of such complex stimuli, questions consistently remain about which stimulus properties actually control an animal’s behavior. This difficulty in controlling and analyzing the content of live-action videos and their potential mismatch to an animal’s sensory *umwelt* represent important obstacles to overcome in the future development and application of video technology to the study of animal behavior.

A third approach attempts to minimize these concerns by using video techniques to create and present computer-generated synthetic stimuli as the basis for a visual discrimination (Cook and Katz 1999; Dittrich et al. 1998; Evans et al. 1993; Regolin et al. 2000). Despite the greater control over the video’s content and the less pressing requirement to exactly recreate all of the sensory nuances of a live-action playback, fewer studies of this type have been conducted. For example, Dittrich et al. (1998) used point-light displays (Johansson 1973) to show that movement cues were sufficient for pigeons to discriminate the general actions of “pecking” and “walking.” Regolin et al. (2000) have similarly used point-light displays to show that young chicks can discriminate between a “walking hen” and a scrambled version of this same display. Evans et al. (1993) have shown that the apparent size and speed of an overhead hawk-like computer image influences the degree of alarm calling by chickens. In another example of this approach, we have trained pigeons to discriminate among

Fig. 2 Examples of the nine objects used in creating the motion pathways in experiments 1 and 2. *Identification numbers at upper left correspond to the text, but were not present in the actual videos*



three-dimensional (3D) computer projections of static and dynamic cubes and pyramids (Cook and Katz 1999). We found that the dynamic rotation of these objects within a trial consistently produced better discrimination than did comparable sets of static views. This dynamic object discrimination was also relatively immune to other transformations in object size, rate and direction of rotation, and various combinations of motions around the objects' axes.

The experiments described in this paper also use this third approach. In these studies, we investigated whether the action-event relations produced by the motion of synthetic objects in a naturalistic context could be discriminated by pigeons. Specifically, the pigeons had to discriminate video stimuli that portrayed 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 (e.g., arches, doughnuts) appear to be approached from the camera's perspective (these videos are most appropriately described in terms of the camera's motion because of the frame of reference provided by the motion of the textured ground). Near the end of the video, the camera's point of view either veers off to the left of the display, creating the impression of going "around" the object, or passes "through" the object's central opening. Figure 1 shows the trajectory and timing of these two different motion pathways. These different motions had to be discriminated relative to a variety of dissimilar objects (Fig. 2).

Two sets of experiments were conducted investigating this motion discrimination. Experiment 1 examined the acquisition of this discrimination and its transfer to new video stimuli containing novel objects. Experiment 2 explored the effect of randomizing the order of the individual frames of these video sequences on performance. Collectively, the results are consistent with the hypothesis that the pigeons perceived and discriminated these videos as depicting the 3D approach of an object.

Experiment 1

The first experiment explored the acquisition and transfer of this through/around discrimination. Seven birds were tested in all. Four of these were experienced with other discriminations, including a variation of the dynamic object discrimination used by Cook and Katz (1999). Three were experimentally naïve. The four experienced birds were tested first. Using a go/no-go discrimination, two of these experienced birds were reinforced on a variable interval (VI) schedule for pecking at the around video sequences (Around+) and two for pecking at the through sequences (Through+). The other motion sequence was designated the S- and was presented in extinction. Five different objects were used in presenting each type of motion (see Fig. 2). At first, each video lasted approximately 3 s and was repeated a little over six times in succession during the 20-s period forming each S+ and S- trial. Because these

four birds showed little learning at first, we then made several procedural modifications in an effort to promote learning (timing changes, additional punishment), with a reduction in the number of objects being tested eventually having the most marked effect on acquisition. With this success, we then trained and tested the three naïve birds using these successful conditions. Following acquisition, we transfer-tested the successful pigeons to novel videos containing objects differing in color, shape, and material from those experienced in training. If the pigeons had learned to discriminate these videos based on a generalized representation of the relative motion, then the specific identity of the objects should not matter.

Method

Animals

Four experienced and three naïve male White Carneaux pigeons (*Columba livia*) were tested. The experienced animals had previously been tested in a compound choice discrimination and a dynamic object discrimination task very similar to that reported in Cook and Katz (1999). They were maintained at 85% of their free-feeding weight during testing, with free access to water and grit in their home cages.

Apparatus

Testing was done in a flat-black Plexiglass chamber (38 cm wide×36 cm deep×38 cm high). The stimuli were presented on a color computer monitor (NEC MultiSync C500; McDonough, Ga., USA) visible through a 26 cm×18 cm viewing window in the middle of the front panel of the chamber. The viewing window's bottom edge was 20 cm above the chamber floor. The monitor was protected by a thin piece of glass mounted in this window. Pecks to the monitor were detected by an infrared LED touch screen (resolution of 80×48 locations; EMS Systems, Champaign, Ill., USA) mounted behind a 40-mm-wide Plexiglas ledge that went around the inside edge of the viewing window. A house light was located in the ceiling of the chamber and was illuminated at all times, except during timeouts. A food hopper was located in the middle of the front panel with its access hole flush to the floor.

All experimental events were controlled and recorded with a Pentium II-class computer. Its graphics card operated the monitor at an 800×600 pixel screen resolution with 16-bit color resolution. Computer-controlled relays (Metabyte, Taunton, Mass., USA) operated the hopper and house light. All event programming was done in VisualBasic (Microsoft, Redmond, Wash., USA) using the FXTTools video control component (Pegasus Software, Tampa, Fla., USA). The videos were created and rendered in the AVI video format (384×288 pixels; Cinepak Codec compression) using the Bryce 4 animation package (MetaCreations Corporation, Carpinteria, Calif., USA).

Discrimination training with video stimuli

Two motion sequences or pathways were used in discrimination training. The video stimuli for each motion pathway consisted of 60 individual frames presented in succession. Each frame was programmed to last 50 ms, but the actual measured time was 53 ms, and as a result each video segment lasted approximately 3200 ms. For both motion sequences frames 1–50 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 ten frames. For the *Around* condition the camera's perspective went around to the left of the object over these frames. As a result, the critical ten frames distinguishing the two conditions lasted a little over 500 ms for each clip.

Nine different objects were used in all to illustrate each type of motion sequence over both experiments (see Fig. 2). These consisted of a wood-finished spiral tunnel (object 1), a red doughnut (object 2), a blue-and-gold striped hollow square (object 3), a gold cube with a spherical area removed from the center (object 4), a yellow-and-gray striped cone (object 5), a green arch (object 6), a black-and-white checkered five-sided tunnel (object 7), a purple marbled square arch (object 8), and a hollowed-out blue sphere (object 9). The colors and textures were selected from the available materials in the animation software.

The motion towards these objects was rendered with a semi-naturalistic 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. Its dappled look produced a flow field that created the appearance of the camera's motion over this landscape. The selected sky type was "Caribbean Clear" and contained no movement cues (e.g., clouds). The "sun" was placed at an azimuth of 134° with an altitude of 30° providing shading that contributed to the 3D appearance of the objects, however, shadows were not permitted. Some examples of these videos can be found at the journal's website for electronic supplementary materials. A color version of Fig. 2 is also there.

Procedure

Pre-training

All pigeons were initially trained to peck at the training stimuli prior to the beginning of go/no-go discrimination training. The experienced birds needed no training to peck the displays, while the naïve birds were trained to first peck the warning signal and then the video displays. Each pre-training session consisted of 80 trials, with 40 *Around* and *Through* trials. Each object appeared equally often in each condition. The experienced birds were pre-trained with five objects (1, 2, 3, 4, 5), while the naïve birds were trained with one object (1). Birds were reinforced for

pecking all displays on an increasingly lean VI schedule. This took three to five sessions for the experienced birds and nine sessions for the naïve birds. For the naïve birds, two additional sessions were then conducted in which they were exposed and reinforced for pecking at four additional objects (3, 4, 6, 7) to reduce any neophobia during subsequent transfer testing. Discrimination training began once peck rates to the displays had stabilized for each bird.

Discrimination testing – experienced birds

The four experienced birds were divided into two groups. The *Around+* group was reinforced for pecking at the *Around* video condition and punished for pecks at the *Through* condition. The *Through+* group was reinforced for pecking at the *Through* condition and punished for pecks at the *Around* condition.

Each session consisted of 80 trials (40 *Around* and 40 *Through*). A trial began with the presentation of a 2.5-cm white warning signal in the center of the display. A single peck to this signal caused it to be replaced with one of the video stimuli. Each trial was programmed to last 20 s, so each 3.2-s video was repeated about 6.25 times within a trial. Pecks at the video stimulus on S+ trials were reinforced on a VI-10 schedule. Pecks at the video stimulus on S– trials were simply not reinforced over the first four sessions, but were then punished with a variable dark timeout (1 s for each peck to the video) in subsequent sessions. A small portion of S+ trials (25%) were always conducted as non-reinforced probe trials in order to measure peck rate without the contamination of food delivery. These probe trials simply ended after the 20-s video presentation. Trials were separated by a 3-s inter-trial-interval (ITI).

Because the four experienced birds showed little discrimination over the initial sessions, several successive changes were made in the procedure. Timeouts were implemented starting with session 5. Next the duration of the video segments were reduced beginning with session 15, with the first 30 frames, depicting the lengthy approach to the object, eliminated. This reduced each video segment to approximately 1600 ms and caused them to be repeated approximately 12.5 times each trial. Beginning with session 20, the computer monitor was moved 4 cm back from the front panel in order to reduce the visual angle of the display. Beginning with session 26, four of the five objects were removed from daily testing with each session consisting exclusively of discrimination testing using only object 1.

Discrimination testing – naïve birds

Based on the final results of training with experienced birds, all three naïve birds were tested from the beginning in only the *Around+* condition using object 1. Each session consisted of 80 trials using the timeout procedure and with the monitor moved back 4 cm from the front panel. Only the last 30 frames of each video were shown.

Transfer testing – Around+ birds

Immediately after learning the discrimination as depicted in Figs. 3 and 4, the two experienced and three naïve birds in the Around+ condition were started on a series of transfer tests using new objects. Transfer tests were conducted in blocks of two sessions, with a new object tested in each block. For each transfer test session, six transfer trials (3 Around and 3 Through) with a new object were randomly mixed into a daily session's trials. These test trials were conducted as non-reinforced probe trials. Following two sessions of testing, the transfer object was added to the daily training regime and trials with this new object were differentially reinforced. Once this introduced object supported a discrimination ratio $DR = \frac{S+ \text{ probe peck rate}}{S+ \text{ probe peck rate} + S- \text{ peck rate}} \times 100$ of 65 or higher

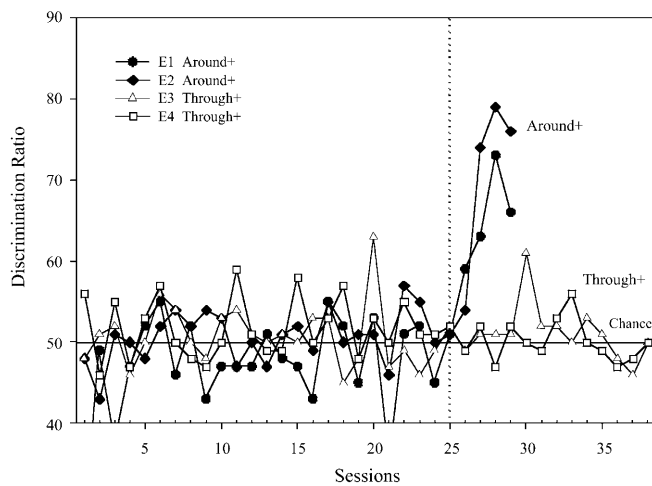


Fig. 3 Mean discrimination ratios over the first 38 sessions for the four experienced birds (*E1*, *E2*, *E3*, *E4*) in experiment 1. The horizontal dotted reference line denotes chance performance. The vertical dotted line denotes the reduction of the total number of objects from five to one

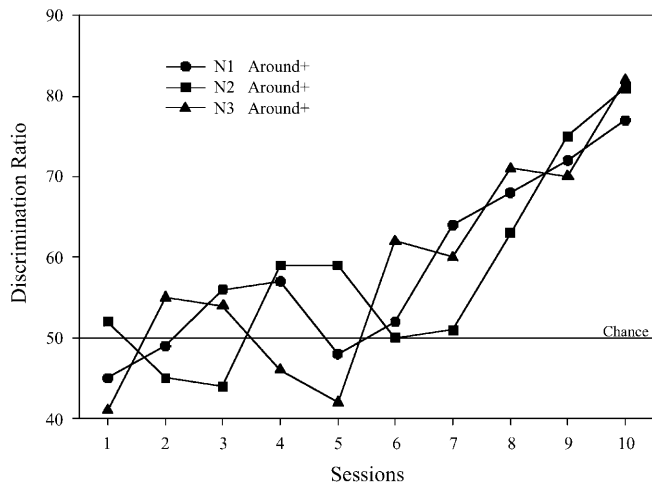


Fig. 4 Mean discrimination ratios over the first 10 sessions for the three naïve birds (*N1*, *N2*, *N3*) in experiment 1. The dotted reference line denotes chance performance

for at least four sessions, another transfer test was conducted using yet another new object. Following two sessions of testing, this object was then added to the training regime, and so on. Using this incremental procedure, four birds (2 experienced and 2 naïve) completed six transfer tests and one naïve bird completed four tests. The experienced birds were tested in order with the objects 2, 3, 4, 5, 6, 7, while the naïve birds were tested with a different ordering of the objects 2, 4, 3, 6, 7, 8. The number of trials per session was adjusted between 80 and 84 daily trials to accommodate and equate the increasing number of objects over testing, but keeping session length approximately the same.

Results

Discrimination testing – all birds

Figure 3 shows the acquisition results for the four experienced birds as a function of discrimination ratio. The S+ peck rates for these scores were derived from only the non-reinforced probe trials. Over the first 25 sessions, none of the four showed any evidence of learning. Beginning with the reduction in the number of objects (the vertical dotted line in Fig. 3), the two Around+ birds rapidly learned to discriminate between the two motion sequences. This was confirmed by a significant main effect of sessions in a repeated measure analysis of variance (ANOVA) of the first five post-reduction sessions, $F_{(4,4)}=9.29$, $P<0.05$. The two Through+ birds showed little change in performance with this reduction over the next 15 sessions. Because they were showing no evidence of learning, we discontinued their testing in order to concentrate on understanding the nature of discriminative control in the Around+ birds.

Figure 4 shows the acquisition results for the three naïve birds as a function of DR over their first ten discrimination sessions. All three readily learned to discriminate the Around+ condition with one object, significantly improving their discrimination over these sessions, $F_{(9,18)}=9.01$, $P<0.001$.

Transfer testing – around+ birds

Figure 5 shows mean DR over the six transfer tests for baseline (shaded bars) for all birds combined and the transfer conditions for the experienced (solid bars) and naïve (stippled bars) birds. Combining all the transfer data, the birds performed significantly above chance (mean=50) with the transfer videos (mean DR=61.7; $t_{(4)}>5.3$, $P<0.01$, single mean t-test), but were significantly below their baseline performance (mean DR=76.7; $t_{(4)}=4.9$, $P<0.01$, paired t-test). Despite this positive transfer, there was some variability in the DR scores across the different transfer tests, as can be seen in Fig. 5. This was due in part to the reduced peck rates observed on the transfer trials. Further, the naïve birds collectively failed their first and fifth transfer tests of the six. We suspect that the failure in the fifth

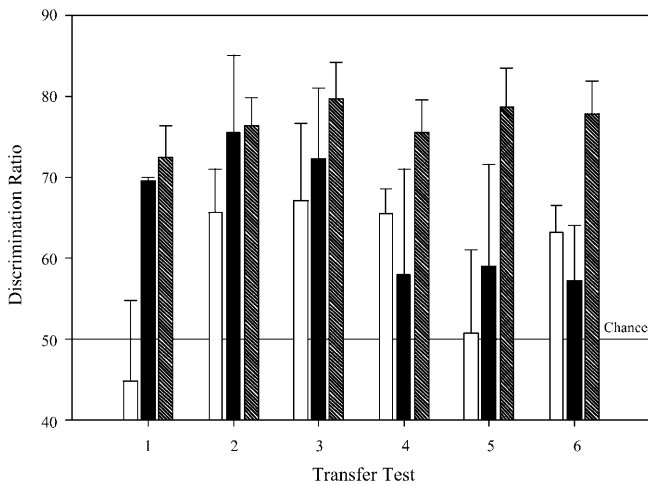


Fig. 5 Mean discrimination ratios for six transfer tests with novel objects. The *white bars* show transfer performance for the naïve birds and the *black bars* show transfer performance for the experienced birds. The *striped bars* show baseline performance for all birds combined. The *reference line* denotes chance performance

test was due in part to the material used to compose the test object (7). The experienced birds also did poorly with this object (their sixth test), and its checkered pattern may have made it difficult for the birds to discriminate the objects' boundaries and surfaces. Finally, it should be noted that the experienced birds generally scored better on the earlier transfer tests than the later ones. Recall that the experienced birds had been differentially reinforced for the objects (2, 3, 4, 5) used in these tests, although they had not expressed any discrimination at that time. This difference suggests some possible carryover (e.g., reduced neophobia or latent discrimination learning) from the earlier experience.

Discussion

This first experiment found that birds could learn to discriminate the motion sequences of around and through across a wide variety of objects used to instantiate these actions. It also revealed that this discrimination generally transferred to novel objects, although at a reduced level of discrimination. It was also found that learning may have depended upon which motion designated the S+: with five birds easily learning to discriminate the Around+ condition when tested with a single object, while the two Through+ birds failed to learn the discrimination after extensive training.

The key question concerns what properties of these stimuli controlled this generalized discrimination. Did the birds experience these video stimuli as depicting looming 3D objects or did they somehow use a simpler 2D interpretation of these images to guide the discrimination? To avoid repetition regarding this central question, we reserve our discussion of this core issue to the end of the paper. Instead, we focus here on two issues specific to Experiment 1's results.

The first issue is why the discrimination initially seemed harder with multiple objects, but could be learned following training with one object. Similar easy-to-hard effects are relatively common in animal discrimination learning. One possible reason may have been due to the wide variety of shapes, colors, and materials used to make the objects. These salient surface features may have drawn attention away from the critical motion properties until they were rendered irrelevant by training with only a single object. A caveat to keep in mind, however, is that this effect was detected with the experienced birds (the naïve birds not being tested for this effect). Recall that these birds had been tested in a dynamic object discrimination similar to that of Cook and Katz (1999). In that particular discrimination, motion facilitated the discrimination of the objects' shape, but was incidental to the discrimination per se. This prior experience may have biased the birds to attend to the object's properties in preference to its motion, resulting in the initially poor discrimination behavior observed. It will be important to see if we can get this one-versus-many effect with naïve birds in a between-groups design.

The second issue is the failure of the two Through+ birds to learn the discrimination. When compared to the Around+ condition, this looks much like a feature+/feature- effect (Jenkins and Sainsbury 1970). Jenkins and Sainsbury (1970) found that pigeons more readily learn discriminations when the distinctive feature is present in the S+ stimulus than vice versa (see Dittrich and Lea 1993 for a F+/F- effect in the discrimination of motion versus static videos). Given this analysis, it suggests the around path may contain a distinctive featural difference from the through path. The most obvious candidate is the change of direction present in the around sequence that is not present in the uniformly expanding optic flow of the through sequence. Another more speculative reason, with an ecological twist, is that it might be more reinforcing or natural to successfully avoid an approaching object than pass through it in a manner similar to a collision. Despite the intriguing nature of the latter hypothesis, birds often fly through the branches of trees and other gaps making this explanation seem unlikely to us. Again, the caveat that this effect was tested and found only with our experienced birds should be heeded. Caution is recommended until a stronger experimental replication with naïve birds is conducted.

Experiment 2

The next experiment asked whether this putative motion discrimination was critically tied to the sequential changes in the camera's perspective relative to the objects. Towards this end, we manipulated whether the frames of the videos were presented in their normal coherent sequence, most consistent with the experience of an approaching object, or in a randomized sequence that broke up this motion and its natural continuity.

This manipulation helps to address two critical issues. The first is to judge the degree to which the relative mo-

tion of the camera's point of view and objects across the frames is needed for the discrimination. If the actions of around and through were the basis for the discrimination, then randomizing the frames should directly hinder the birds' capacity to extract this information. The second issue is whether the birds experience the videos three-dimensionally. If, for example, the discrimination were done exclusively on 2D cues, such as the spatial position of the objects relative to the border during the divergent later frames, then randomizing the order of the frames should not be particularly disruptive. Presumably the key frames with these 2D cues are just as available in either type of presentation, whereas the 3D cues are not. If the objects are seen as being approached in depth, then this frame randomization should be especially disruptive and its effects long lasting.

This experiment consisted of two separate tests of this manipulation. In both tests, two types of presentation were tested with videos constructed with novel objects. One type, labeled coherent, presented the motion relative to the objects just as in experiment 1. The second type, labeled randomized, presented the same video, but randomized the order of the video's individual frames. For test 1, this random ordering was different for each randomized presentation. In test 2, only one fixed random ordering was used on these trials, making it more comparable to the single coherent sequence tested in the coherent condition.

Method

Animals and apparatus

The pigeons, minus the two Through+ birds, and chamber were the same as in experiment 1.

Procedure

Test 1

At the beginning of this test, all five Around+ birds were being tested daily with different videos composed from seven objects, except for one naïve bird who was being tested with only five objects. The first part of test 1 consisted of a transfer test with a novel object (experienced birds – object 8, naïve birds – object 5) for two sessions. This novel object was tested on 12 (6 Around/6 Through) non-reinforced probe trials in each session. Half of the six trials for each motion path were conducted using a coherent ordering of the video frames and half used a randomized ordering of the same frames. This randomized sequence was different for each presentation of this condition. These 12 test trials were randomly intermixed into the 80 training trials testing the familiar objects. For all trials the duration of each video segment was approximately 1600 ms and repeated a fixed 16 times per trial.

After the transfer test, trials testing these coherent and randomized sequences were continued, but were now dif-

ferentially reinforced for the motion they contained. Eight (4S+/4S-) coherent and eight randomized trials were tested in each session, and randomly intermixed with the ongoing training trials. Two of the S+ trials for each type were conducted as probe trials to measure peck rate in the absence of reinforcement. Testing of coherent and randomized sequences using differential reinforcement continued for 18 sessions. The novel object was not added to the daily training regime as done in experiment 1.

Test 2

Test 2 was identical to test 1 except for the following differences. First, all birds were tested with the same new novel object (object 9). Second, one naïve bird was not tested for reasons unrelated to the experiment. Third, only one fixed random sequence was used for the randomized presentations. After two sessions of transfer testing, testing of coherent and randomized sequences using differential reinforcement continued for ten sessions.

Results

Test 1

Figure 6 shows mean DRs in the baseline, coherent, and randomized presentation conditions in two-sessions blocks for test 1. Overall the coherent condition supported better discrimination of the same videos than did the randomized condition. This was confirmed by comparing the DR for the coherent and randomized conditions across the reinforced sessions (repeated measures ANOVA, Blocks of 2 Sessions×Presentation Condition). This revealed both a main effect for both Condition, $F_{(1,4)}=18.5$, $P<0.025$ and Sessions, $F_{(8,32)}=3.65$, $P<0.005$. To examine in more detail the effects of experience on each condition, additional one-

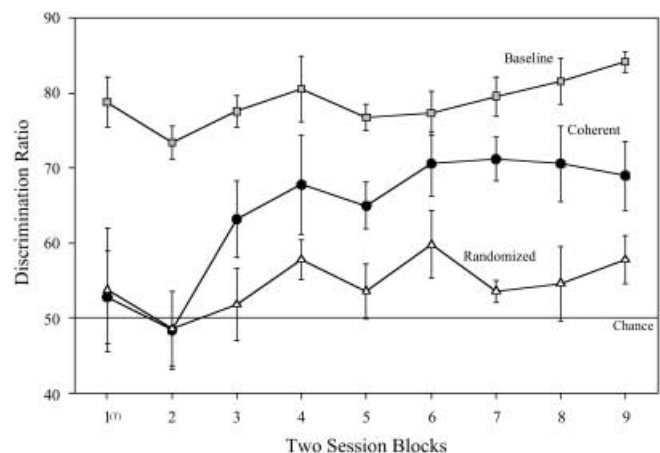


Fig. 6 Mean discrimination ratios in two-session blocks for the baseline, coherent and randomized (multiple sequences) presentation conditions in the test 1 of experiment 2. The *small T* identifies non-reinforced transfer sessions. The *reference line* denotes chance performance

way ANOVAs examining DR across sessions for each presentation condition were conducted. These revealed a significant change across sessions for the coherent condition, $F_{(8,32)}=4.1$, $P<0.005$, but no change in performance across sessions for the randomized condition, $F_{(8,32)}=1.5$, $P>0.05$. To accommodate this difference, performance from the last third of the test (blocks 6–9) was used to judge whether these discriminations exceeded chance. T-tests using these sessions revealed that coherent performance (mean DR=70.4) was significantly above chance, $t_{(4)}=7.8$, $P<0.01$, as was randomized performance (mean DR=56.5), $t_{(4)}=3.0$, $P<0.05$. Finally, separate comparisons with baseline performance (repeated measures ANOVA, Blocks of two Sessions×Baseline vs. Presentation Condition) revealed that performance in both presentation conditions was significantly below baseline, both $F_{S(1,4)}>128.2$, $P_s<0.001$. Initial transfer performance was poor in both presentation conditions. T-tests examining DR from just the first two probe sessions revealed that neither coherent (mean=52.8) nor randomized (mean=53.8) performance was significantly above chance.

Test 2

Figure 7 shows discrimination ratios for the baseline, coherent, and randomized presentation conditions for test 2. The results basically replicate those for test 1, with the coherent condition supporting superior discrimination to the randomized condition. This was confirmed by comparing DR for these two conditions across the reinforced session (Blocks of 2-Sessions×Presentation Condition) which revealed a main effect for Condition, $F_{(1,3)}=73.6$, $P<0.005$, but no effect of Sessions. Further examination with the separate one-way ANOVAs as described previously revealed no significant effect of session for either condition examined separately, $F_{S(4,12)}<1$. Performance over the last

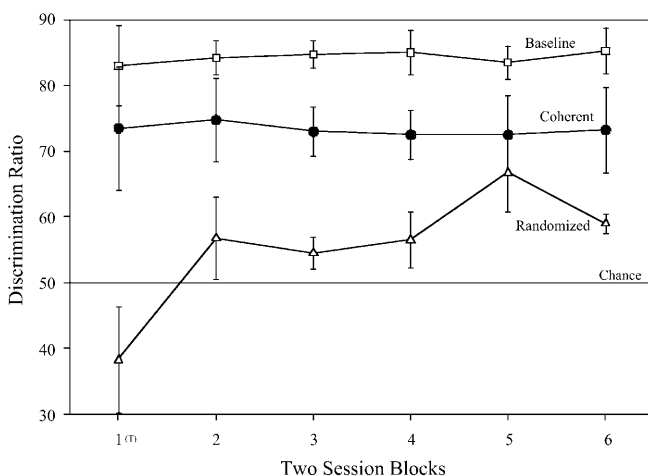


Fig. 7 Mean discrimination ratios in two-session blocks for the baseline, coherent and randomized (fixed sequence) presentation conditions in the test 2 of experiment 2. The *small T* identifies non-reinforced transfer sessions. The *reference line* denotes chance performance

half of the reinforced testing (Blocks 4–6) was used to judge whether discrimination in either presentation condition exceeded chance. T-tests revealed that again both coherent (mean DR=72.7, $t_{(3)}=6.2$, $P<0.01$) and randomized performance (mean DR=60.7, $t_{(3)}=3.9$, $P<0.05$) was significantly above chance. Separate comparisons with baseline performance (repeated measures ANOVA, 2-session Blocks×Baseline vs. Presentation Condition) revealed that performance in both presentation conditions was significantly below baseline, both $F_{S(1,3)}\geq 14.9$, $P_s<0.05$. Transfer performance in the first two sessions was also different for the two presentation conditions, with coherent condition (mean DR=73.5) supporting significantly better discrimination than did the randomized condition (mean DR=38.2).

Discussion

The major result from experiment 2 was that coherently ordered presentation of the motion discrimination supported better discrimination than did the randomized presentation of the same video. This was true whether the randomized presentation was differentially reinforced or not (tests 1 and 2), and whether only a single randomized sequence was always tested (test 2). This coherent/randomized difference indicates that the temporal sequencing of the individual frames was critical to the present discrimination.

Why did the randomization of the frames so disrupt the discrimination? At least three possibilities can be considered. First, this drop in performance may represent a form of generalization decrement or neophobia to the randomized appearance. Such experience-based explanations, however, are not consistent with the general failure to see much improvement with continued training of the randomized condition, although it may play a role in the poor transfer performance in test 1. Instead the randomization of the frames seems to have prevented the birds from acquiring critical sequential information from the video. Two variations of this idea can be entertained.

Part of our original motivation behind the randomization manipulation was to test the role of 2D cues in the videos. One set of such cues might be the asymmetric difference in the position of the object at the end of the video. For instance, in the around condition the objects move off and fill the right side of the display, while in the through condition they symmetrically fill and move 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 match the coherent conditions because the same 2D cues are present in both. The consistently poorer performance in the randomized condition hints that such cues are not the primary mediators of the discrimination, although the consistent above-chance elevation of the randomized condition following some experience suggests that such cues might be present to a limited degree. A simple variation of this possibility is that the birds were relying on the timing of the frames to help

detect these divergent frames. For example, the birds may have used the start of the video to anticipate when to look for the 2D cues in the latter frames. In this case, then, the frame randomization just breaks up this temporal cueing. We do not think this is the case, however. In experiments not reported here, we have subsequently reintroduced the first 30 frames of the videos, and making the video twice as long. This did not dramatically affect the discrimination, suggesting that video timing per se is not critical to performing the discrimination.

The intriguing third possibility is that the birds perceived the videos much as they were intended – as 3D objects being approached in depth over a textured ground with two different actions portrayed. In this case, the randomization manipulation is effective because it disrupts the perception and integration of the movement and depth cues across frames, essentially removing the appearance of moving through or around something.

General discussion

The results of these experiments are consistent with the hypothesis that pigeons can discriminate the motion paths around and through with a variety of approaching objects in a semi-realistic video context. This discrimination moderately transfers to novel objects and clearly depends on the order of the video's frames. While we continue to be concerned about simpler interpretations (e.g., 2D cues+timing), we currently think the results are most concordant with a working hypothesis that this discrimination is mediated by a 3D perception of the objects and the integration of motion and depth from across the frames of the video. This interpretation is consistent with the conclusions drawn by Cook and Katz (1999) regarding the contributions of motion to object perception in a different setting. It extends their observations by showing that object motion itself can be the exclusive basis for visual discrimination, whereas their experiments demonstrated only an indirect contribution of motion to object perception. In both cases, the experimental results imply that the birds perceive the three-dimensional projections portrayed in 2D moving stimuli. Of course, this conclusion must be accepted provisionally upon further experiments testing directly the 2D and 3D hypotheses. What does seem increasingly clear is that by adding motion to our discriminative stimuli, the growing experimental evidence seems to be tipping the balance in favor of a 3D interpretation for least some types of stimuli (see also Reid and Spetch 1998).

Beginning with the seminal observations of Gibson (1950), human perceptual psychologists have focused on the usefulness of optic flow fields as a means for understanding the integration of motion and action in visual scenes (e.g., Harris and Jenkin 1998; Watanabe 1998). Optic flow is the outward expansion of the points in visual space produced by forward self-motion. For humans in laboratory settings (and theme parks) such flow fields produce powerful impressions of motion, even when seated

and stationary. This type of information would be valuable to any mobile animal, and perhaps especially to flying birds (also noted by Gibson in several of his illustrations). The design of the stimuli in the current experiments certainly created strong flow fields because of the camera's movement over the textured ground and the approach of the rigid central object, and is a quite reasonable candidate as the mediator of this video discrimination.

These behavioral results converge nicely with recent studies examining the neural mechanisms of motion perception by the pigeon's visual system. In this area, good evidence indicates that their visual system is sensitive to optic flow information. For instance, Hongjin and Frost (1999) recently described three types of looming-selective neurons in the nucleus rotundus that could be the basis for performing such image expansion computations. These mechanisms, along with ones for object translation (Wylie and Frost 1999), readily provide a physiological basis for the visual discrimination required by our tests.

While a 3D interpretation as the perceptual basis for our discrimination is consistent with the behavioral and physiological evidence, the birds also need to classify the different object motions as well. We offer the hypothesis that the present discrimination may also have been additionally mediated by the representation of the different motions as separate *action categories* by the birds. For some time, it has been accepted that pigeons can form natural categories for polymorphous pictorial stimuli, such as of trees, cars, cats, flowers, birds, mammals, fish, oak leaves, and humans (Bhatt et al. 1988; Cerella 1979; Cook et al. 1990; Herrnstein and de Villiers 1980; Herrnstein et al. 1976). In these cases, the individual exemplars are so variable that many have argued that an abstract categorical representation of these stimuli is required for their discrimination. What has not been widely considered is the existence of another class of highly variable stimulation in which a categorical-like representation might also be valuable – classifying one's own behaviors, those of other organisms, and of surrounding objects. Outside of fixed or modal action patterns, behavior can be both flexible and variable in form and would seem to create similar difficulties in the coding of its specifics in the same manner as with polymorphous objects. Yet being able to recognize the generalized stalking behavior of a predator, regardless of context, individual, and other specific details, would be of obvious value. The ability to form generalized action or behavioral categories would allow an animal to discriminate and respond to the functions or consequences of behavior regardless of its exact form. This kind of categorization is exactly what is done by ethologists and psychologists faced with coding long sequences of complex behavior in the field or the laboratory. Our suggestion is that animals might also be able to form and use behavioral categories. That is, besides encoding generalized nouns, such as trees and chairs, animals may also be able to classify verbs. The current positional discrimination of around and through may be tapping into such a categorical-like capacity. The recent work by Dittrich et al. (1998) on the discrimination by pigeons of the behaviors of "walking"

and “pecking” in video and point light displays is another example of this same general idea.

One interesting question raised by this idea concerns the level or scale of the behavioral units to be classified. Just as object categories have a preferred basic level of encoding (Rosch et al. 1976), classifications of behavior may also have preferred levels of description. Thus, encoding a class of behavior as “walking” may be easier than at a more detailed motoric description (for a clearly related discussion, see the extensive literature on molar versus molecular descriptions of matching behavior) or at a more general level such as “moving”. One speculation is that how species parse behavior into units may be influenced or constrained by the hierarchical organization of its own behavioral systems (e.g., Timberlake 1983).

There are an enormous number of ways of using videos to explore a wide spectrum of issues in animal behavior. We think our current results show that this research approach is a promising one for revealing how organisms perceive and classify the dynamic actions surrounding them, and that both synthetic and natural video approaches can be valuable tools in the analysis of cognition and behavior.

Acknowledgements This research was supported by a grant from the National Science Foundation to R.G. Cook. This research was conducted following the relevant ethics guidelines for research with animals and was approved by Tufts University’s institutional IACUC. Further information about our research can be found at www.pigeon.psy.tufts.edu.

References

- Adret P (1997) Discrimination of video images by zebra finches (*Taeniopygia guttata*): direct evidence from song performance. *J Comp Psychol* 111:115–125
- Bhatt RS, Wasserman EA, Reynolds WF Jr, Knauss KS (1988) Conceptual behavior in pigeons: categorization of both familiar and novel examples from four classes of natural and artificial stimuli. *J Exp Psychol Anim Behav Proc* 14:219–234
- Cerella J (1979) Visual classes and natural categories in the pigeon. *J Exp Psychol Hum Percept Perform* 5:68–77
- Cook RG (2000) The comparative psychology of visual cognition. *Curr Dir Psychol Sci* 9:83–88
- Cook RG, Katz JS (1999) Dynamic object perception in pigeons. *J Exp Psychol Anim Behav Proc* 25:194–210
- Cook RG, Wright AA, Kendrick DF (1990) Visual categorization in pigeons. In: Commons ML, Herrnstein R, Kosslyn SM, Mumford DB (eds) *Quantitative analyses of behavior: behavioral approaches to pattern recognition and concept formation*. Lawrence Erlbaum, Hillsdale, pp 187–214
- D’Eath RB (1998) Can video images imitate real stimuli in animal behaviour experiments? *Biol Rev Cambridge Philos Soc* 73:267–292
- D’Eath RB, Dawkins MS (1996) Laying hens do not discriminate between video images of conspecifics. *Anim Behav* 52:903–912
- Dittrich WH, Lea SEG (1993) Motion as a natural category for pigeons: Generalization and a feature-positive effect. *J Exp Anal Behav* 59:115–129
- Dittrich WH, Lea SEG, Barrett J, Gurr PR (1998) Categorization of natural movements by pigeons: visual concept discrimination and biological motion. *J Exp Anal Behav* 70:281–299
- Evans CS, Marler P (1991) On the video images as social stimuli in birds: audience effects on alarm calling. *Anim Behav* 47:17–26
- Evans CS, Macedonia JM, Marler P (1993) Effects of apparent size and speed on the response of chickens, *Gallus gallus*, to computer-generated stimulation of aerial predators. *Anim Behav* 46:1–11
- Gibson JJ (1950) *The perception of the visual world*. Houghton Mifflin, Boston
- Harris L, Jenkin M (1998) *Vision and action*. Cambridge University Press, Cambridge
- Herrnstein RJ, Villiers PA de (1980) Fish as a natural category for people and pigeons. In: Bower GH (ed) *The psychology of learning and motivation*. Academic Press, New York
- Herrnstein RJ, Loveland DH, Cable C (1976) Natural concepts in pigeons. *J Exp Psychol Anim Behav Proc* 2:285–302
- Hongjin S, Frost BJ (1999) Computation of different optical variables of looming objects in pigeon nucleus rotundus neurons. *Nat Neurosci* 1:296–303
- Jenkins HM, Sainsbury RS (1970) Discrimination learning with the distinctive feature on positive or negative trials. In: Mostofsky D (ed) *Attention: contemporary theory and analysis*. Appleton-Century-Crofts, New York, pp 239–274
- Jitsumori 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. *Anim Learn Behav* 27:303–315
- Johansson G (1973) Visual perception of biological motion and a model of its analysis. *Percept Psychophys* 14:201–211
- Lea SEG, Dittrich WH (1999) What do birds see in moving video images? *Cah Psychol Cogn* 18:765–803
- McQuoid LM, Galef BG Jr (1993) Social stimuli influencing feeding behaviour of Burmese fowl: a video analysis. *Anim Behav* 46:13–22
- Patterson-Kane E, Nicol CJ, Foster TM, Temple W (1997) Limited perception of video images by domestic hens. *Anim Behav* 53:951–963
- Regolin R, Tommasi L, Vallortigara G (2000) Visual perception of biological motion in newly hatched chicks as revealed by an imprinting procedure. *Anim Cogn* 3:53–60
- Reid S, Spetch ML (1998) Perception of pictorial depth cues by pigeons. *Psychonom Bull Rev* 5:698–704
- Rosch E, Mervis CB, Gray WD, Johnson DM, Boyes-Braem P (1976) Basic objects in natural categories. *Cogn Psychol* 8:382–429
- Sands SF, Lincoln CE, Wright AA (1982) Pictorial similarity judgments and the organization of visual memory in the rhesus monkey. *J Exp Psychol Gen* 3:369–389
- Shimizu T (1998) Conspecific recognition in pigeons (*Columba livia*) using dynamic video images. *Behaviour* 135:43–53
- Timberlake W (1983) The functional organization of appetitive behavior: behavior systems and learning. In: Zeiler MD, Harzem P (eds) *Advances in analysis of behavior*, vol 3. Wiley, Chichester, pp 177–221
- Watanabe T (1998) *High-level motion processing*. MIT Press, Cambridge
- Wylie DRW, Frost BJ (1999) Responses to neurons in the nucleus of the basal optic root to translational and rotation flowfields. *J Neurophysiol* 81:267–276