Recognition of static and dynamic images of depth-rotated human faces by pigeons

MASAKO JITSUMORI and HIROSHI MAKINO
Chiba University, Chiba, Japan

In three experiments, we examined pigeons' recognition of video images of human faces. In Experiment 1, pigeons were trained to discriminate between frontal views of human faces in a go/no-go discrimination procedure. They then showed substantial generalization to novel views, even though human faces change radically as viewpoint changes. In Experiment 2, the pigeons tested in Experiment 1 failed to transfer to the faces dynamically rotating in depth. In Experiment 3, the pigeons trained to discriminate the dynamic stimuli showed excellent transfer to the corresponding static views, but responses to the positive faces decreased at novel viewpoints outside the range spanned by the dynamic stimuli. These results suggest that pigeons are insensitive to the three-dimensional properties of video images. Consideration is given to the nature of the task, relating to the identification of three-dimensional objects and to perceptual classifications based on similarity judgments.

The retinal image of an object varies as a function of orientation, distance, lighting condition, background scene, and other aspects at the time of viewing. Invariance operations are required for observers to identify objects despite the variability of retinal stimulation. Observers may recognize the variations between images of the same object, but the differences should not obscure the identity of the object. One of the most sophisticated invariance operations of the visual system is the recognition of objects from novel views. When an object rotates in depth, the retinal image often changes drastically; different parts of the object come into or move out of view for an observer. Similarly, an observer's locomotion changes the retinal image of an object. There is no doubt that orientation invariance is of great ecological relevance for most animals living in the three-dimensional (3-D) visual world. Casual observations of birds suggest that they readily identify conspecifics, predators, and surrounding objects, despite changes in viewpoint.

An earlier study (Cerella, 1977; see also Cerella, 1990a, 1990b) showed that pigeons failed to exhibit viewpoint invariance with two-dimensional (2-D) projections of simple geometric objects. Generalization testing with different types of geometric stimuli (cube and noncube, alphabet letter, irregular quadrilateral, and so forth) consistently revealed that pigeons fail to generalize to stimuli rotated on the projection plane or around the *x*- (horizontal) or *y*- (vertical) axis. More recently, Wasserman et al. (1996) used line drawings of a chair, a flashlight, a desk lamp, and an airplane as discriminative stimuli that contained much richer depth information than did simple geometric stimuli. The pigeons trained in a four-alternative forced-choice procedure successfully transferred the discrimination to the novel views of these objects rotated around the *y*-axis. Similarly, Peissig, Young, Wasserman, and Biederman (1999, 2000) examined pigeons' rotation generalization, using computer-rendered versions of single shapes, such as an arch, a barrel, a brick, and a wedge, as stimuli. Although significant rotation generalization was found in these studies, the generalization was not in practice viewpoint invariant. That is, discrimination accuracy decreased systematically as the objects were rotated farther away from the trained view. The rotation generalization shown by pigeons may be better described as viewpoint-dependent perspective transformations or else as simple stimulus generalization from training view(s), making it unnecessary to posit a more complex 3-D object recognition. Viewpoint dependence in pigeons was recently confirmed by Spetch, Friedman, and Reid (2001). They trained pigeons and humans to discriminate pairs of pictures depicting objects seen from two different viewpoints. Humans showed viewpoint invariance for novel rotations between training views but viewpoint dependence for novel rotations outside the training views. Pigeons' performance always declined systematically with degree of rotation from the nearest training view. Spetch et al. (2001) stated that viewpoint dependence is a general feature of pigeons' object recognition processes and that avian species, such as pigeons, may lack a mechanism that allows for invariance within the rotation boundary formed by training views. It is rather doubtful that the pigeons actually recognized the different training views as representing an

identical object that changes in its 3-D orientation. These experimental findings appear to be inconsistent with casual observations of the superior abilities of birds in recognizing natural objects in the normal environment.

Following the pioneering work by Herrnstein and Loveland (1964), it has been well documented that animals, particularly pigeons, can classify photographs that contain a particular type of natural object (see the reviews in Huber, 2001; Jitsumori & Delius, 2001; and Watanabe, Lea, & Dittrich, 1993). Cerella (1990a) stated that the excellent categorization of photographs of natural objects by pigeons "might seem to entail recognition of a single object across its perspective transformations, as well as recognition of many other objects and their perspective transformations. The successful demonstration of natural object classification by the pigeon may be thought to ensure a positive outcome to the simpler pattern recognition problem" (p. 141). However, as we have noted earlier, generalization testing with different types of geometric stimuli consistently revealed that pigeons fail to generalize to novel views. These findings led Cerella to argue that pigeons learn to recognize and discriminate stimuli by means of 2-D pattern matching of local features. Cerella (1990a) pointed out that "the pigeon does not in fact recover objects in color slide . . . . There is still no evidence for an 'object concept', in the sense of following a target through its three-dimensional rotations" (p. 157).

However, pigeons, like most other animals, are consistently faced with the problem of recognizing objects in the natural environment. A moving object, as well as a pigeon's own locomotion, yields images of the object continuously changing its orientation in depth ("runs," or sequences of views), which may enable it to integrate different views as a unified 3-D object. With regard to computational systems of object recognition, Edelman (1999) has pointed out the importance of the cohesiveness of objects and the smoothness of the physical transformations that objects undergo with time. Also, most natural objects contain much richer information than do simpler geometric shapes, including the diagnostic feature(s) that are more or less distinctive across different views. Thus, one might speculate that, in the natural environment, pigeons learn to selectively attend to diagnostic feature(s) so as to recognize familiar objects in different viewpoints.

In a recent study, Cook and Katz (1999) trained pigeons to discriminate between computer-generated 3-D projections of cubes and pyramids in a go/no-go procedure involving static and dynamically rotating presentations of these stimuli. On static trials, an object was presented randomly at one of 360 angular orientations around the y-axis. On dynamic trials, the object appeared to revolve around its y-axis. Over a number of tests with different types of variations of the objects (rotation around a new axis, new complex paths of motion, changes in surface color, and elimination of surface and contour information), the pigeons showed substantial transfer to test stimuli, with better testing performances to dynamic

stimuli than to static stimuli. The results indicated that there was more discriminative information in dynamic displays than in static ones, and therefore, the authors argued that movements of the objects resulted in a superior 3-D perception of their structure. However, it remains unresolved whether pigeons recognize static views as belonging to an object identical to the one that is also presented dynamically within the same sessions.

In the present study, we examined to what extent pigeons show rotation generalization with pictures of natural objects and whether seeing the objects dynamically rotating in depth promotes viewpoint independence within and outside the range of the dynamic views. More specifically, static and dynamic images of human faces rotated around the y-axis were used as stimuli in the present study. Human faces are highly variable across rotations in depth. Not only the global shape, but also the individual facial features, such as eyes, nose, and mouth, change when viewpoint changes. Employing several measures of physical similarity, Moses (1993, cited in Ullman, 1998; see also Moses, Ullman, & Edelman, 1996) revealed that differences of the same face under different viewing conditions are generally larger than differences between distinct faces under the same viewing conditions.

Nevertheless, Moses et al. (1996) found that human subjects exhibited almost error-free recognition of upright human faces across changes in viewing directions on the basis of an example from a single view. That is, a single view of an object that is unfamiliar for subjects but drawn from a highly familiar stimulus class (i.e., upright human faces) allows human observers to generalize to novel viewpoints. With top–bottom inverted faces, on the other hand, the single-view generalization decreased monotonically (as measured by error rates and response time) as the inverted faces were rotated away from the trained view. The authors argued that the human visual system overcomes variations between upright images of human faces because the visual system relies on the general properties of upright faces to compensate for the effects of viewing conditions. The class-based processing built up for upright faces does not generalize to inverted faces, which are rarely seen in daily life.

In the literature on humans, viewer-centered models (e.g., Edelman & Bülthoff, 1992; Tarr & Pinker, 1989; Ullman, 1989) assume that the structure of a 3-D object is represented by a collection of its 2-D views previously stored in memory and that recognition varies systematically with similarities to the stored views. In contrast, object-centered models (e.g., Corballis, 1988; Marr & Nishihara, 1978; see also Biederman, 1987, for object recognition based on nonaccidental properties) assume that an object in the visual field is compared with 3-D representations of that object that are inferred on the basis of the metrics of the 2-D retinal images and the estimated rotation in depth. The results of Moses et al. (1996), summarized above, are in line with viewer-centered models.

From the perspective of phylogeny, it is unlikely that pigeons form class-general viewpoint-specific (i.e., upright-

only) representation of human faces. Phelps and Roberts (1994, Experiment 3) used pictures of human faces, great ape faces, monkey faces, and outdoor scenes as stimuli for pigeons and found that there were no effects of picture orientation in any of the four types of picture categories. That is, pigeons did not show an inversion effect in memorizing upright and inverted pictures (for a review of the literature on humans, see Valentine, 1988). In another study, Jitsumori and Yoshihara (1997) trained pigeons to classify human faces in terms of facial expressions (happiness vs. anger). The pigeons successfully discriminated between the faces even when the facial features (the eyes and eyebrows and the mouth) were top–bottom inverted. Thus, the pigeons did not show the so-called Thatcher illusion (Thompson, 1980). Given that pigeons, having been trained to discriminate human faces seen from a single viewpoint, compare novel views with the previously seen trained views to determine their responding, they should show more limited rotation generalization than that obtained in previous studies (e.g., Peissig et al., 1999, 2000; Wasserman et al., 1996), when diagnostic features were not available to discriminate between the faces at novel viewpoints.

In Experiment 1, we examined to what extent pigeons show rotation generalization with pictures of human faces. In Experiment 2, the pigeons that had been trained and tested with static views in Experiment 1 were tested for transfer to dynamic views. A question of interest was whether the pigeons would show immediate transfer to the dynamic views. We expected that if pigeons are capable of recognizing static and dynamically presented views of the same face as depicting an identical 3-D object, transfer would occur from the static- to the dynamic-view conditions. Additional questions addressed in Experiment 2 were whether and to what extent dynamic-view training provides a broadening of the range of invariance. In Experiment 3, transfer from the dynamic- to the static-view conditions was examined with experimentally naive pigeons as subjects. They first were trained with the dynamic views and then were tested for transfer to the static views, including those presented in novel orientations outside the range spanned by the dynamic views.

EXPERIMENT 1
Transfer From Static Views to New Static Views

Five pigeons were trained to discriminate frontal (0º) views of two positive and two negative human faces in a go/no-go discrimination procedure. After completion of the training, they were tested for transfer to static images of the faces rotated around the *y*-axis over the range of ±90º.

Method

Subjects

Five experimentally naive homing pigeons of retired racing stock were maintained at 80%–85% of their free-feeding weights throughout the experiment. They were caged individually in an outdoor aviary on the roof of the department building. Water and grit were freely available in the home cages.

Apparatus

The stimuli were displayed via a laser-videodisk player (Sony LVA3500) on the 6-in. screen of a color monitor (Sony PVM6041Q) positioned 2 cm in front of an aluminum panel of the experimental chamber (35 × 35 × 38 cm). The pigeons could view the screen through a transparent rectangular key (screen key) 6 cm high and 6 cm wide. The screen key was positioned 17.5 cm above the floor on the panel. A 1.5 × 4.0 cm food aperture, located on the floor and centered below the screen key, gave 3-sec access to a solenoid-operated food tray containing a mixture of grains. When the food tray was presented, a small light bulb (2 W) immediately below the aperture turned on. A 2-cm-diameter key (start key) was on the rear wall of the chamber. It was centered on the wall 19 cm above the floor. The start key allowed the pigeons to initiate trials and to see the stimulus at different distances while they were approaching the screen key. A houselight (3 W) placed at the center of the ceiling dimly illuminated the chamber. The chamber and the video monitor were in a darkened testing room. A microcomputer (NEC-9821) selected an appropriate frame on a videodisk (Sony LVM3AA0) on each trial, controlled experimental events, and collected responses.

Stimuli

Figure 1 shows examples of the black-and-white reproductions of the stimuli. The stimuli were still video pictures of four undergraduate male students (A and C are Chinese, B and D are Japanese), with no glasses or beards and a neutral expression. They were sitting in a pivot chair looking straight ahead at a small brown square on the wall about 1.6 m from the chair. Their faces were all taken in color with a digital camera (Victor, GR-DVM1) under the same lighting condition. The images at nine depth orientations (−90º, −67º, −45º, −22º, 0º, 22º, 45º, 67º, and 90º rotations around the *y*-axis) were taken by rotating the pivot chair to different orientations with respect to the camera located approximately 1.5 m from the chair. The 0º orientation was a full-face frontal view, with the center point halfway between the eyes. The height and width of the 0º faces were approximately 35 and 29 mm, on average, when they were displayed on the color monitor. The videos showing the faces dynamically rotating around the *y*-axis were prepared, but they were not employed in Experiment 1. Readers may see a compressed version of the digitized stimuli on the homepage (http://cogsci.L. chiba-u.ac.jp/pigeon/face/).

Procedure

Pretraining. The pigeons were first trained to peck on the screen key, using a conventional hand-shaping procedure. As soon as they were pecking consistently, the number of pecks required was gradually increased to 30 (fixed ratio 30), and at least three sessions were given in which 30 reinforcers were provided. During this phase, the video monitor displayed a white unpatterned stimulus. The pigeons were then trained to peck on the start key, and at least three sessions were given on a schedule of continuous reinforcement.

In the final phase of preliminary training, the pigeons were given 60-trial sessions, with an intertrial interval (ITI) of 5 sec, during which the houselight was turned on and a blank frame was displayed on the monitor. At the start of each trial, only the start key was illuminated. A single response on the start key turned it off and presented a white unpatterned stimulus on the video monitor. Thirty pecks on the screen key produced food delivery. This training phase, which was included to ensure that the pigeons moved to the screen key immediately after they pecked on the start key, lasted at least three sessions.

Discrimination training. The pigeons were trained to discriminate the 0º faces on a go/no-go discrimination task similar to that used by Vaughan and Greene (1984). The positive faces were frontal views of C and D and the negative faces were those of A and B for
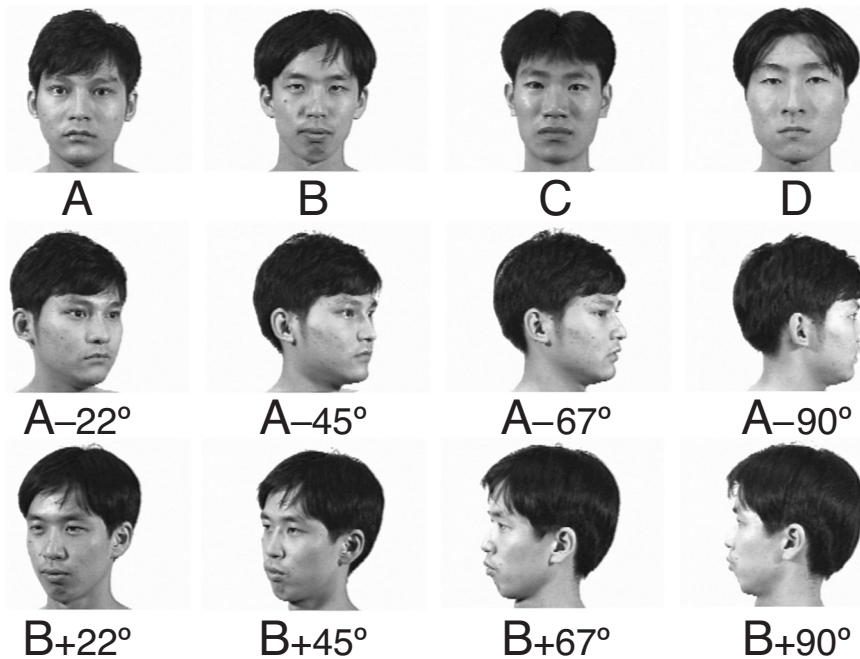
Figure 1. Examples of the black-and-white reproductions of the stimuli. On the top row are the frontal views used for training. The middle and bottom rows are examples of the images from novel viewpoints.

Birds 1, 2, and 3, and vice versa for Birds 4 and 5. The pigeons started trials by pecking once on the start key, as in the final phase of preliminary training. A session consisted of 30 positive and 30 negative trials. Each of the four 0° faces (two positives and two negatives) appeared on 15 trials in pseudorandom order, with the restriction that no more than 3 positive or negative trials could occur in succession. For the first 10 sec after the stimulus was displayed on the monitor, pecks on the screen key were recorded but had no consequence. Following this 10-sec period, a variable interval (VI) of 10 sec started (range, 0–20 sec; step, 2 sec). On positive trials, a peck after the end of a scheduled interval was reinforced. On negative trials, pecks were not reinforced. A negative trial terminated after its scheduled duration had expired and 5 sec had passed without responding. After an ITI (as in pretraining), the next trial started. Training continued until 90% or more of the total responses occurred on positive trials.

Generalization testing. Each pigeon was next tested with the faces at nine depth orientations, including the 0° faces used for training. Thus, there were a total of 36 pictures (4 individuals × 9 orientations). A session consisted of two randomized blocks of 36 trials, with each stimulus being tested twice in a session. Two test sessions were given in extinction (no rewards and no penalties), and the response rate was calculated from the number of responses that occurred during the entire trial time. The average presentation time was 20 sec, varying from 10 to 30 sec. Over the two test sessions, the total presentation time for the two positive faces and the two negative faces from a given viewpoint was 160 sec each. Other procedural details were the same as those during training.

Results and Discussion

The pigeons took a mean of 29.1 days to complete the discrimination training. The mean response rates (pecks per minute) to the positive and negative faces, averaged across the birds in the last two sessions of training, were 161.8 and 17.3.

In generalization testing, the pigeons generally showed more responding to the positive faces than to the negative faces across novel viewpoints, but responses to the positive faces decreased systematically as the faces were rotated farther away from the training orientation. Because the pigeons exhibited different rates of pecking, relative responses, calculated for each pigeon as the proportions of the mean number of pecks to the trained positive views, were averaged across the birds. This yielded 1.0 for the positive faces at 0°. The left panel of Figure 2 shows mean relative responses to the positive and negative faces as a function of degree of rotation. The mean generalization gradient for the positive faces had a peak at the training orientation, whereas the mean generalization gradient for the negative faces was almost flat across the viewpoints. In order to confirm these observations, a repeated measures, two-way analysis of variance (ANOVA), with face (positive vs. negative) and orientation (0° vs. ±22° vs. ±45° vs. ±67° vs. ±90°) as variables, was conducted. In this and all other statistical tests, an alpha level of .01 was used. There were significant main effects of face [$F(1,4) = 2,305.76$] and orientation [$F(4,16) = 5.43$]. In addition, the interaction was significant [$F(4,16) = 5.84$], reflecting the differential control of orientation over responding to the positive and the negative faces. Paired comparisons of contrasts revealed that the pigeons responded significantly more often to the positive faces than to the negative faces at 0° [$F(1,16) = 73.71$], ±22° [$F(1,16) =$
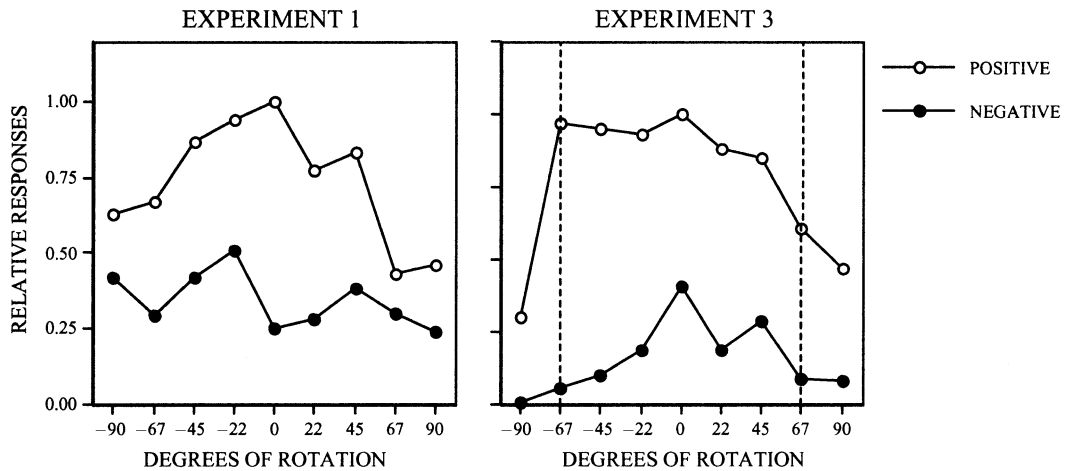
EXPERIMENT 1

EXPERIMENT 3

Figure 2. Left: mean relative responses to the novel views of the positive faces (open circles) and the negative faces (filled circles) during generalization testing in Experiment 1. Right: mean relative responses to the positive faces (open circles) and the negative faces (filled circles) during generalization testing in Experiment 3. The region between the vertical lines represents the range spanned by the dynamic views used for training.

27.31], $\pm45°$ [$F(1,16) = 26.39$], and $\pm67°$ [$F(1,16) = 8.37$]. Although the difference in responding at $\pm90°$ [$F(1,16) = 6.43$] did not meet statistical significance, 3 of the 5 pigeons still responded more often to the positive faces than to the negative faces. The discrimination ratio at $\pm90°$ calculated as the proportion of the total responses in the presence of positive stimuli was .69, .42, .48, .75, and .71 for Birds 1, 2, 3, 4, and 5, respectively.

To explore the effect of disparity, a repeated measures, two-way ANOVA, with disparity (22° vs. 45° vs. 67° vs. 90°) and orientation direction (left vs. right) as variables, was conducted separately on the data with positive and negative faces. A significant effect was found only for disparity of the positive faces [$F(3,12) = 5.88$]. Paired comparisons of contrasts revealed that the rates of responding differed between 45° and 67° [$F(1,12) = 8.64$]. There was no significant difference in responding between 22° and 45° [$F(1,12) = 0.02$] and between 67° and 90° [$F(1,12) = 3.81$]. Thus, the pigeons' responses to the positive faces decreased significantly when the faces were rotated up to $\pm67°$. The effect of disparity was not significant with the negative faces [$F(3,12) = 0.98$]. Generalization gradients of inhibition are often flatter than those of excitation (e.g., Honig & Urcuioli, 1981) and, thus, may be less sensitive to the rotation of negative faces than to that of positive faces.

We expected that, if pigeons simply compare novel stimuli with previously seen trained stimuli to determine whether they should respond, human faces would be far more difficult for pigeons to discriminate over novel viewpoints than artificial objects that differ in shape and geometric components. Nonetheless, the pigeons showed generalized testing performance comparable to that obtained with the highly distinctive artificial objects (see Peissig et al., 1999, 2000; Wasserman et al., 1996). One way to explain this finding is that the pigeons responded

on the basis of 3-D properties. The systematic generalization decrement as a function of distance from the training views is more in line with the viewer-centered models than with the object-centered models in the literature on humans. However, the present findings do not allow one to distinguish among object recognition models. From the perspective of view-independent recognition, Wasserman and his colleagues (Peissig et al., 1999, 2000; Wasserman et al., 1996) argued that the systematic decrement in generalization is possibly due to discrimination between the trained views and other views presented during generalization testing. This notion is consistent with the findings of neurophysiological studies by Logothetis, Pauls, Bülthoff, and Poggio (1995; see also Logothetis et al., 1994). They found cells in the temporal cortex of trained monkey subjects that showed tuning functions to a specific trained view and other cells that responded similarly to different views of an object. In the present experiment, discrimination training with only a single view of each of the four faces and the rotation generalization testing conducted in extinction could have made viewpoint salient for the pigeons that were not explicitly trained to discriminate the stimuli by disregarding changes in viewing orientations.

An alternative explanation of the surprisingly good, although not perfect, transfer to the rotated views is that the pigeons used multiple 2-D features as cues by which to discriminate the stimuli, some of which changed by rotation in depth and some of which did not. It has been well documented that pigeons are capable of using multiple features and additively integrating features to discriminate and categorize highly complex visual stimuli (e.g., Huber & Lenz, 1993; Jitsumori, 1993; von Fersen & Lea, 1990). More recently, Huber, Troje, Loidolt, Aust, and Grass (2000) and Troje, Huber, Loidolt, Aust, and Fieder (1999) reported that pigeons discriminated be-

tween male and female faces on the basis of multiple features and, specifically, that they were sensitive to surface features, such as color, brightness, and texture. If multiple features, including diagnostic feature(s) that are more or less distinctive over some range of rotation, were learned during training, it is not surprising that the pigeons showed substantial transfer to novel views. The present findings thus can be explained in terms of feature learning, without the need to posit more sophisticated processing of 3-D properties.

A key question is whether or not the pigeons recognized different views of a given face as representing an identical 3-D object. With regard to this issue, Delius, Emmerton, Hörster, Jäger, and Ostheim (1999) reviewed a number of recent studies that used pictures of objects and scenes as stimuli for pigeons and chickens. They concluded that the ability of these avian species to recognize equivalence between pictures and real objects is severely limited when a behavioral task requires a relatively precise correspondence. Although there is evidence that pigeons transfer discrimination from objects to pictures (e.g., Cabe, 1976; Delius, 1992; Watanabe, 1993) and that pigeons may have the ability to use 2-D cues to perceive three dimensionality (Cook & Katz, 1999; Reid & Spetch, 1998; Spetch, Kelly, & Lechelt, 1998), little is known about avian recognition of 2-D images of complex natural objects. In Experiments 2 and 3, we explored this issue.

EXPERIMENT 2
Transfer From Static Views to Dynamic Views

In Experiment 2, the pigeons tested in Experiment 1 were tested for transfer to dynamic views continually rotating around the y-axis over the range of $\pm 67°$. If they actually were to perceive different views as belonging to the same object rotated in its 3-D orientation, they would readily recognize the dynamic views and show immediate transfer from the static- to the dynamic-view conditions.

Another issue addressed in this experiment was whether and to what extent dynamic-view training broadens the range of invariance. Seeing dynamic views of a human face continuously changing its orientation in depth may give pigeons the opportunity to integrate the different views as a unified 3-D object. Also, discrimination training under the dynamic-view condition may encourage pigeons to discriminate faces by disregarding changes in viewing orientation. If pigeons can abstract 3-D properties of dynamically rotating faces, dynamic-view training should increase the degree of viewpoint independence, even when the faces are presented at novel orientations ($\pm 90°$) outside the dynamically presented views used for training.

Only one positive and one negative face was dynamically rotated in depth during training. The remaining positive and negative faces were presented at 0°, as in Experiment 1. After completion of the training, the pigeons were given generalization testing, as in Experiment 1. If dynamic-view training with human faces has the effect

that pigeons abstract the 3-D properties that are common to the members of this object class, dynamic-view training may promote rotation independence even with faces that have never been presented dynamically (for the literature on humans, see Vetter, Hurlbert, & Poggio, 1995).

Method

Subjects and Apparatus

The subjects were the 5 pigeons tested in Experiment 1. Housing, maintenance, and apparatus were the same as those in Experiment 1.

Stimuli

Dynamic video images of the four faces (A, B, C, and D) prepared but not used in Experiment 1 were employed. The object person sitting in a pivot chair was requested to rotate the chair at a constant speed from $+67°$ to $-67°$ in 5 sec, with the axis being at the center and his eyes smoothly tracking a number of small brown squares horizontally aligned on the wall about 1.6 m from the chair. Video recordings were repeated until they eventually rotated in the way requested. On training trials for the pigeons, the best 5-sec excerpt of each person was repeatedly played forward and backward so that the face appeared to rotate to the left and the right over a range of $\pm 67°$. The static stimuli used for generalization testing were the same as those used in Experiment 1. There was no notable difference between the images of the static stimuli and the corresponding frozen frames of the dynamic stimuli.

Procedure

Discrimination training. The faces used as dynamic stimuli were determined for each bird. The positive face that elicited more responding than the others and the negative face that elicited less responding than the others during generalization testing in Experiment 1 were presented under the dynamic-view condition. The positive face served as dynamic stimulus was C for Bird 1, C for Bird 2, D for Bird 3, B for Bird 4, and A for Bird 5, respectively. The negative one was B for Bird 1, A for Bird 2, B for Bird 3, C for Bird 4, and D for Bird 5, respectively.

Discrimination training was given immediately after completion of Experiment 1. A session consisted of 30 static trials and 30 dynamic trials. A starting frame on each dynamic trial was selected so that the faces began to rotate to the left or to the right from approximately $-67°$, $-45°$, $-22°$, 0°, 22°, 45°, and 67°, each of which occurred, as nearly as possible, equally often across sessions. The faces were continually rotating to the left and to the right over the range of $\pm 67°$ until each dynamic trial was terminated. In static trials, the 0° views of the remaining one positive and one negative face used in Experiment 1 were presented. Other procedural details were the same as those in Experiment 1. Training continued until the 90% criterion was attained in the static and the dynamic trials.

Generalization testing. After completion of training, the birds received generalization testing as in Experiment 1.

Results and Discussion

Discrimination Training

Discrimination ratios for the static and the dynamic stimuli separately calculated as the proportion of the total responses emitted in the presence of the positive face are shown as a function of training sessions in Figure 3. The discriminative performances with the static stimuli were more or less disrupted at the beginning of the training, probably due to the preceding generalization testing in extinction and/or the introduction of dynamic stimuli. However, the discrimination of the static stimuli recovered quickly in the first few sessions. With the dynamic
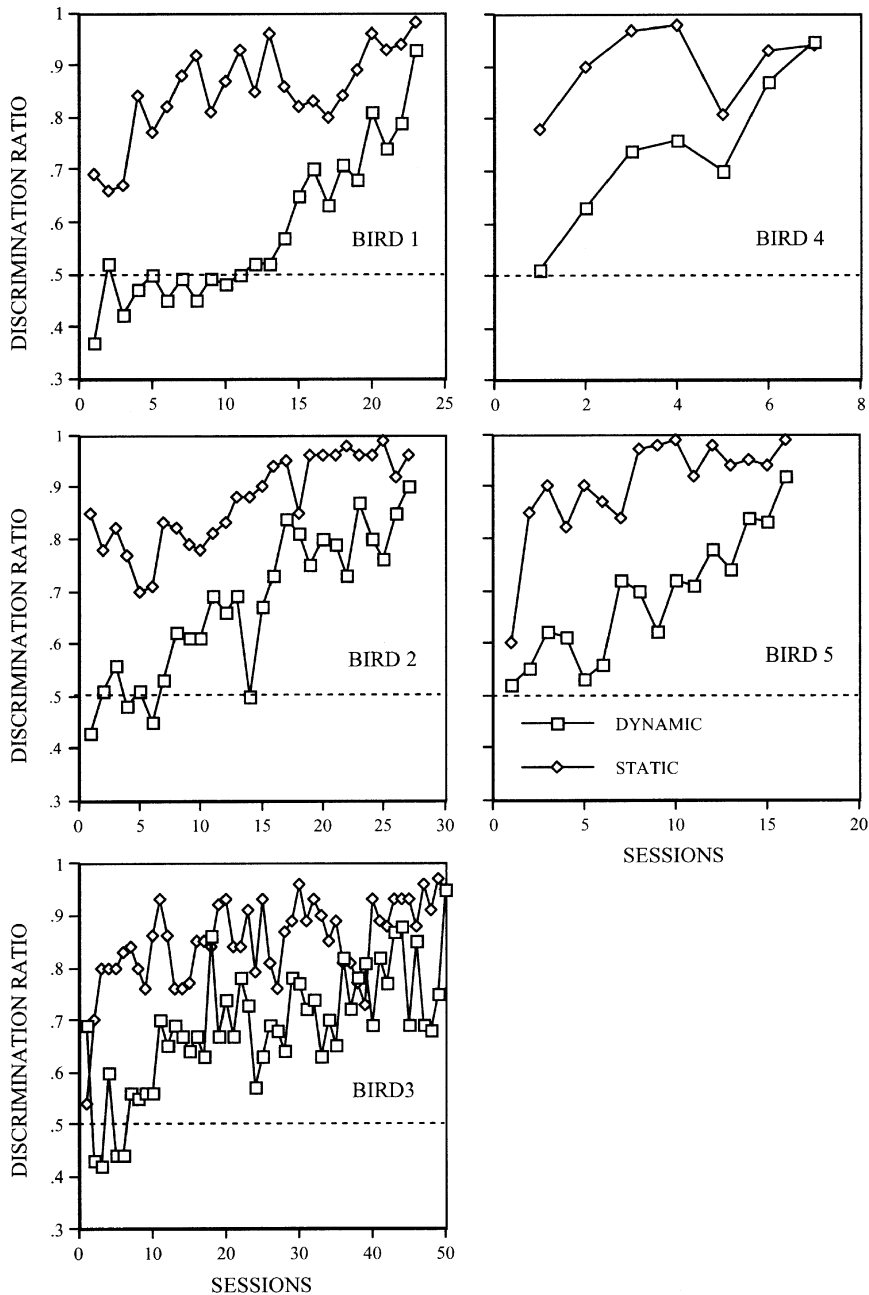
Figure 3. Discrimination ratios for the faces presented in the static-view condition and those for the faces presented in the dynamic-view condition during acquisition in Experiment 2.

stimuli, on the other hand, discrimination was near chance at the beginning of training, despite the finding that the pigeons had shown substantial transfer to novel views of these faces in Experiment 1. The discriminative performances under the dynamic-view condition gradually rose as a function of training sessions. Such a long-lasting deterioration in performance should not be attributable to a novelty effect. Rather, the failure of transfer to the dynamic stimuli suggests that the pigeons treated the static and the dynamic views of the same face differently. Because of the negative finding, we cannot conclude whether or not pigeons recognize static views of a given face as representing a 3-D object identical to that presented under the dynamic-view condition. However, this finding strengthens the notion that the substantial generalization to novel viewpoints shown by the same pigeons in Experiment 1 was based on 2-D processing, rather than on object-centered or viewer-centered 3-D information.

Generalization Testing

Figure 4 shows the mean relative response rates during generalization testing in Experiment 2 (filled circles), in comparison with those in Experiment 1 (open circles). The generalization gradients with the faces used for the dynamic-view training are in the left panels, and those with the remaining faces are in the right panels. Note again that the positive and negative faces used for dynamic-view training were discriminated better than the others during the generalization testing in Experiment 1. The results in Experiments 1 and 2 were remarkably similar, indicating that the dynamic-view training failed to broaden the range of invariance with the static stimuli. A repeated measures, two-way ANOVA, with experiment (Experiment 1 vs. Experiment 2) and orientation (0° vs. ±22° vs. ±45° vs. ±67° vs. ±90°) as variables, was conducted separately on the data shown in each panel of Figure 4. The effect of orientation was significant for the positive faces, regardless of whether they were used [$F(4,16) = 15.06$] or not used [$F(4,16) = 18.91$] for dy-

namic training. The effect of experiment and the interaction were not significant, confirming that generalization over a wider range of orientations than that in Experiment 1 was not obtained in Experiment 2. For the negative faces, a significant effect was found only for orientation of the face that was not used for dynamic training [$F(4,16) = 7.98$].

An unexpected finding is that the pigeons failed to show transfer from the dynamic to the static views even within the range spanned by the dynamic views. This finding is rather surprising in light of the results of Jitsumori, Natori, and Okuyama (1999. Experiments 1 and 2). In this study, two groups of pigeons were trained to discriminate video images of conspecifics based on the individuals or on their actions. Both groups showed rapid acquisition, and the discrimination transferred to still scenes consisted of frozen frames from the dynamic scenes used for training. Note, however, that the pigeons in the present experiment had prior experience discriminating static views (at 0°), whereas the pigeons in Jitsumori
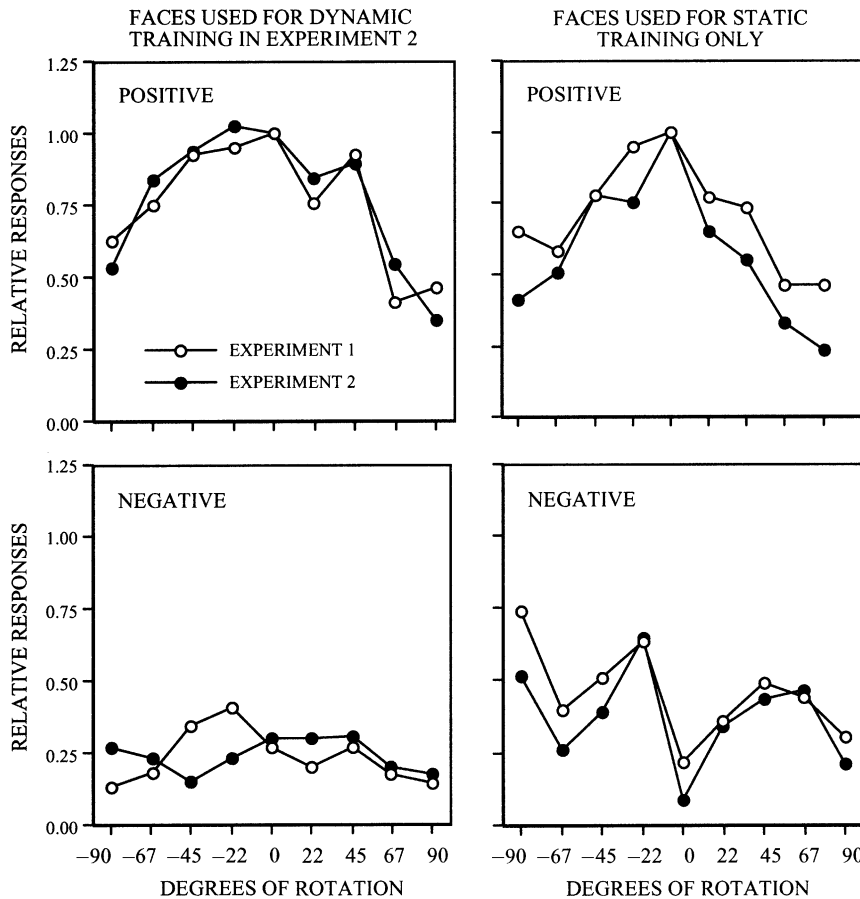

Figure 4. Comparisons of mean generalization gradients in Experiment 1 and Experiment 2. Top left: the positive face used for dynamic-view training in Experiment 2. Top right: the positive face used for static-view training only. Bottom left: the negative face used for dynamic-view training in Experiment 2. Bottom right: the negative face used for static-view training only. The pair of faces used for dynamic-view training was discriminated better than the other pair by each bird during the generalization testing in Experiment 1.

et al. did not. Although video images of conspecifics and human faces may differ in biological significance, visual complexity, and distinctiveness, in the present experiment it is likely that the effects of the extensive training at only one view in Experiment 1 carried over to generalization testing in Experiment 2. One possible explanation of this carryover effect is that the prior experience with the frontal views might have overshadowed other views to acquire control over responding. Or the pigeons might have responded to the dynamic stimuli by relying solely on the frames showing the faces at about 0º.

EXPERIMENT 3
Transfer From Dynamic Views to Static Views

To examine transfer from dynamic views to static views by eliminating any effects of prior experience of seeing pictures of human faces, we trained experimentally naive pigeons to discriminate between dynamically rotating presentations of two positive and two negative faces. The pigeons were then tested for transfer to the static views, including those presented in novel orientations outside the range spanned by the dynamic views.

Method

Subjects, Stimuli, and Apparatus

The subjects were 4 experimentally naive pigeons (Birds 6, 7, 8, and 9). Housing, maintenance, stimuli, and apparatus were the same as those in Experiments 1 and 2.

Procedure

The procedure was the same as that in Experiment 1, except that the dynamic stimuli, instead of the frontal views, were used for training. As in the dynamic trials in Experiment 2, the faces were rotated continually over the range of ±67º until each training trial was terminated. The positive faces were A and B and the negative faces were C and D for Birds 8 and 9, and vice versa for Birds 6 and 7. After completion of training, the birds received testing with the static stimuli as in Experiments 1 and 2.

Results and Discussion

Three of the 4 pigeons took a mean of 48.0 days to complete the discrimination training. The remaining pigeon showed difficulty in discriminating the dynamic stimuli. This bird was dropped after 70 training sessions.

The right panel of Figure 2 shows the results of generalization testing with the static stimuli. Comparisons with the results in Experiment 1 (the left panel) suggest that dynamic-view training improved discrimination of the static stimuli. However, response rates to the positive faces substantially decreased at the novel viewpoints (±90º), and this tendency was pronounced when the faces were rotated to the left (orientations represented by negative values in the figure). A repeated measures, two-way ANOVA was conducted, with face (positive vs. negative) and orientation (0º vs. ±22º vs. ±45º vs. ±67º vs. ±90º) as variables. As in Experiments 1 and 2, reported statistical tests were evaluated using an alpha level of $p \leq .01$. The analysis indicated that there were significant main effects of face [$F(1,2) = 5{,}069.87$] and orien-

tation [$F(4,8) = 16.74$]. In addition, the interaction was significant [$F(4,8) = 15.12$], reflecting the differential control of orientation over responding to the positive and the negative faces. Paired comparisons of contrasts revealed that the pigeons responded significantly more often to the positive faces than to the negative faces at 0º [$F(1,8) = 198.86$], ±22º [$F(1,8) = 287.98$], ±45º [$F(1,8) = 287.98$], ±67º [$F(1,8) = 298.90$], and even at ±90º [$F(1,8) = 66.17$]. To explore the effect of disparity, a repeated measure, two-way ANOVA, with disparity (22º vs. 45º vs. 67º vs. 90º) and orientation direction (left vs. right) as variables, was conducted separately on the data with positive and negative faces. A significant main effect was found only for disparity of the positive faces [$F(3,6) = 14.34$]. Paired comparisons of contrasts revealed that the rates of responding differed between 67º and 90º [$F(1,6) = 19.70$]. There was no significant difference in responding between 22º and 45º [$F(1,6) = 0.01$] and between 45º and 67º [$F(1,6) = 1.32$]. Thus, the pigeons' responses to the positive faces decreased significantly only when the faces were presented in the novel orientations outside the range spanned by the dynamic views. In addition, the interaction was significant [$F(3,6) = 11.79$], reflecting the finding that the responses to the novel views decreased more markedly at −90º than at 90º. With the negative faces, disparity [$F(3,6) = 5.57$], orientation direction [$F(1,2) = 36.44$], and the interaction [$F(3,6) = 3.05$] did not reach statistical significance.

Although there was a large drop in absolute responding to the positive faces at the novel viewpoints (±90º), the pigeons still discriminated between the novel views of the positive and negative faces. Discrimination ratios for the novel views were .88, .90, and .90 for Birds 6, 8, and 9, respectively. However, the improvement of discrimination was due to the fact that there was also a large drop in absolute responding at the novel views of the negative faces. Apparently, the pigeons discriminated between the familiar and the novel views during generalization testing. The novelty of the faces presented at ±90º might have decreased responding to both the positive and the negative faces. One may argue that this finding is consistent with viewpoint-dependent perspective transformation (the viewer-centered models of 3-D object recognition), and the generalization decrement to the novel positive views does not necessarily indicate the inability of pigeons to perceive 3-D objects in video images. However, it is equally likely that the pigeons were responding on the basis of 2-D similarities to the training views; the positive faces at ±90º may be more perceptually similar to the positive training stimuli than are the negative faces at ±90º.

Jitsumori et al. (1999) have pointed out that because pigeons have a higher flicker-fusion threshold than do humans (Emmerton, 1983; Hendricks, 1966; Powell, 1967), pigeons may not perceive the smooth motion of video animation and, thus, may see the animation as we would see a stroboscopic presentation. Although Jitsumori et al. did not test transfer to novel views outside the range of the dynamic stimuli, they argued that the

stroboscopic effects of moving video sequences, rather than rotation invariance recognition, could promote pigeons' transfer to the corresponding static views. The generalization decrement for the positive faces at novel rotations found in the present experiment strengthens the notion suggested by Jitsumori et al. that pigeons do not recognize dynamically presented views as depicting a 3-D object that changes its image as it rotates in depth. This conclusion appears to be inconsistent with that proposed by Cook and Katz (1999) described earlier. But because their training and testing procedures differed from those used in the present study and, also, because their findings were presented in terms of discrimination ratios, direct comparison was not possible.

One anonymous reviewer suggested that the pigeons might not have looked at the video stimuli very long before acting on them, so that the dynamic features would have been lost. As the video stimuli started from approximately −67°, −45°, −22°, 0°, 22°, 45°, and 67° during dynamic training, discrimination based on the starting orientations would explain the excellent transfer to the corresponding static stimuli in Experiment 3. However, if this had been true, the pigeons in Experiment 2 should have shown a transfer to the dynamic views that was comparable to that which occurred to the novel static views during generalization testing in Experiment 1. The lack of transfer to the dynamic views suggests that this limited-view account is not tenable. Also, Jitsumori et al. (1999) reported findings that suggest that pigeons may not respond to dynamic stimuli by relying solely on the starting frames or any other particular frames. In one of their experiments (Experiment 4), experimentally naive pigeons were trained to discriminate between video scenes of two domestic pigeons, each showing a variety of activities in a natural setting. During training, one scene was played in the normal direction, and the other one was played in the reverse direction. On test trials, the scene discrimination was maintained even when each video was played in the opposite direction, relative to that on training trials. And the pigeons transferred the discrimination to a number of frozen frames pseudorandomly selected from the dynamic stimuli used for training. These findings suggest that no particular frame was critical to discrimination of the dynamic displays. Also, it may be worth noting that the pigeons did not use the play direction as a cue by which to discriminate the video scenes. This finding suggests that movement of different kinds (ecologically valid or not valid) is not salient for pigeons to recognize the objects in dynamic displays.

GENERAL DISCUSSION

The present results can be summarized as follows. In Experiment 1, pigeons were trained to discriminate between frontal views of human faces in a go/no-go discrimination procedure. They then generally showed more responding to the positive faces than to the negative ones over novel viewpoints, even though 2-D images of human faces may change radically as viewpoint changes. Although the pigeons appeared to show robust recognition of rotated human faces, responding to the positive faces declined systematically with degree of rotation from the training view. In Experiment 2, the pigeons tested in Experiment 1 failed to show transfer from the static to the dynamic views. They eventually learned to discriminate between the dynamic views, but dynamic-view training did not increase the degree of viewpoint independence, even when the faces were rotated within the range spanned by the dynamic-view training stimuli. Furthermore, in Experiment 3, experimentally naive pigeons were trained to discriminate faces that were rotated in depth. They then showed excellent transfer to the static views, but responses to the positive faces decreased substantially with novel rotation outside the range spanned by the dynamic-view training stimuli.

The present study is the first systematic investigation of the recognition of static and dynamic views of natural objects by pigeons. An interesting finding is that, although human faces are highly variable across rotations in depth, pigeons showed substantial generalization to novel viewpoints. Some processes that may compensate for the effect of viewpoint differences are required to explain the substantial generalization to novel viewpoints shown by the pigeons in Experiment 1. As we have discussed earlier, the most likely explanation, on grounds of parsimony, is that the pigeons responded on the basis of 2-D similarities to the training views. Although the present study was not designed to analyze the features used by pigeons to discriminate between the stimuli, it is likely that the pigeons used multiple features, some of which varied with changes in viewpoint, whereas others did not. It may be that the pigeons recognized the variations between different views of a rotated human face, but 2-D feature(s) that are distinctive over some range of rotation might have yielded substantial generalization to novel views. However, behavioral research that may identify the features actually used by pigeons to discriminate human faces, as well as detailed analyses of the image, would be needed to understand the correlation between pigeons' generalization performance and physical measures of similarity.

Rotation invariance requires identification across different views of the same object in its 3-D rotation. It should be noted, however, that the present task does not necessarily require recognition of unfamiliar views or exact identification of the object. The rotated views may be more similar to the training view of that face than are those of other faces, and the pigeons might have categorized the test stimuli on the basis of broad similarity judgments. This notion is consistent with an exemplar-based approach of classification in the human literature (e.g., Medin & Schaffer, 1978; Nosofsky, 1991) that assumes that classification decisions are based on the similarity of an item to the exemplars of a target category, relative to exemplars of nontarget categories. The relative-similarity rule may well compensate for the effect of viewpoint differences. Classification of different views into the same category, rather than exact identification of

the object, may explain the pigeons' substantial generalization to novel views. Categorization and identification both involve a many-to-one stimulus–response mapping, but object identification requires equivalence across different views, in the sense that they depict an identical 3-D object. An important question for future research concerns the categorization–identification relation underlying the considerable robustness of pigeons' visual processing of natural objects that contain much richer information than do the artificial objects often used as stimuli in this line of research (e.g., Cerella, 1977, 1990a, 1990b; Spetch et al., 2001).

Another important finding of the present study is that the pigeons failed to show transfer from the static- to dynamic-view conditions. This finding, coupled with the failure of transfer to novel orientations outside the range spanned by the dynamic views, suggests that the pigeons did not perceive different views as belonging to the same 3-D object. This conclusion is consistent with the priority of categorization over identification in the present task. However, the present findings by no means indicate that pigeons are not capable of recognizing real objects from various perspectives in the natural environment. We do not really experience a video image presented on a 2-D display as being truly 3-D. Nonetheless, we can "see" real objects in 2-D photos and in videos. The human visual system recognizes 2-D projections of the objects across changes in their 3-D orientation, whereas pigeons may be insensitive to 3-D information in video images.

The present findings converge on the conclusion that pigeons lump different views together on the basis of similarity judgments, rather than determine whether or not particular views belong to the same 3-D object as that previously seen from a different viewpoint. Categorization, rather than exact identification, may be a simple solution by which pigeons deal with unfamiliar views in generalization testing. The priority of categorization over identification shown by the pigeons in the present study is interesting. Edelman (1999) has pointed out that, for computational algorithms as well as human observers, categorization is easier and faster than exact identification. Perspective invariance is of great ecological relevance to animals, including distantly related species, such as pigeons and humans. Additional research using a variety of tasks and stimuli is needed to determine the cross-species generality and differences in underlying mechanisms of object categorization and identification.

REFERENCES

BIEDERMAN, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, **94**, 115-147.

CABE, P. A. (1976). Transfer of discrimination from solid objects to pictures by pigeons: A test of theoretical models of pictorial perception. *Perception & Psychophysics*, **19**, 545-550.

CERELLA, J. (1977). Absence of perspective processing in the pigeon. *Pattern Recognition*, **9**, 65-68.

CERELLA, J. (1990a). Pigeon pattern perception: Limit on perspective invariance. *Perception*, **19**, 141-159.

CERELLA, J. (1990b). Shape constancy in the pigeon: The perspective transformations decomposed. In M. L. Commons, R. J. Herrnstein, S. M. Kosslyn, & D. B. Mumford (Eds.), *Quantitative analyses of behavior* (Vol. 8, pp. 145-163), Hillsdale, NJ: Erlbaum.

COOK, R. G., & KATZ, J. S. (1999). Dynamic object perception by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, **25**, 194-210.

CORBALLIS, M. C. (1988). Recognition of disoriented shapes. *Psychological Review*, **95**, 115-123.

DELIUS, J. D. (1992). Categorical discrimination of objects and pictures by pigeons. *Animal Learning & Behavior*, **20**, 301-311.

DELIUS, J. D., EMMERTON, J., HÖRSTER, W., JÄGER, R., & OSTHEIM, J. (1999). Picture–object recognition in pigeons. *Current Psychology of Cognition*, **18**, 621-656.

EDELMAN, S. (1999). *Representation and recognition in vision.* Cambridge, MA: MIT Press.

EDELMAN, S., & BÜLTHOFF, H. H. (1992). Orientation dependence in the recognition of familiar and novel views of 3D objects. *Vision Research*, **32**, 2385-2400.

EMMERTON, J. (1983). Vision. In M. Abs (Ed.), *Physiology and behavior of the pigeon* (pp. 245-266). London: Academic Press.

HENDRICKS, J. (1966). Flicker threshold as determined by a modified conditioned suppression procedure. *Journal of the Experimental Analysis of Behavior*, **9**, 501-506.

HERRNSTEIN, R. J., & LOVELAND, D. H. (1964). Complex visual concept in the pigeon. *Science*, **146**, 549-551.

HONIG, W. K., & URCUIOLI, P. J. (1981). The legacy of Guttman and Kalish (1956): Twenty-five years of research on stimulus generalization. *Journal of the Experimental Analysis of Behavior*, **36**, 405-445.

HUBER, L. (2001). Visual categorization in pigeons. In R. G. Cook (Ed.), *Avian visual cognition*. Retrieved 2002 from http://www.pigeon. psy.tufts.edu/avc/huber/default.htm.

HUBER, L., & LENZ, R. (1993). A test of the linear feature model of polymorphous concept discrimination with pigeons. *Quarterly Journal of Experimental Psychology*, **46B**, 1-18.

HUBER, L., TROJE, N. F., LOIDOLT, M., AUST, U., & GRASS, D. (2000). Natural categorization through multiple feature learning in pigeons. *Quarterly Journal of Experimental Psychology*, **53B**, 341-357.

JITSUMORI, M. (1993). Category discrimination of artificial polymorphous stimuli based on feature learning. *Journal of Experimental Psychology: Animal Behavior Processes*, **19**, 244-254.

JITSUMORI, M., & DELIUS, J. D. (2001). Object recognition and object categorization in animals. In T. Matsuzawa (Ed.), *Primate origins of human cognition and behavior* (pp. 269-292). Tokyo: Springer-Verlag.

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. *Animal Learning & Behavior*, **27**, 303-315.

JITSUMORI, M., & YOSHIHARA, M. (1997). Categorical discrimination of human facial expressions by pigeons: A test of the linear feature model. *Quarterly Journal of Experimental Psychology*, **50B**, 253-268.

LOGOTHETIS, N. K., PAULS, J., BÜLTHOFF, H. H., & POGGIO, T. (1994). View-dependent object recognition in monkeys. *Current Biology*, **4**, 401-414.

LOGOTHETIS, N. K., PAULS, J., BÜLTHOFF, H. H., & POGGIO, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Current Biology*, **5**, 552-563.

MARR, D., & NISHIHARA, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London: Series B*, **200**, 269-294.

MEDIN, D. L., & SCHAFFER, M. M. (1978). Context theory of classification learning. *Psychological Review*, **85**, 207-238.

MOSES, Y. (1993). *Face recognition: Generalization to novel images*. PhD thesis, Weizmann Institute of Science, Applied Mathematics and Computer Science, Israel.

MOSES, Y., ULLMAN, S., & EDELMAN, S. (1996). Generalization to novel images in upright and inverted faces. *Perception*, **25**, 443-461.

NOSOFSKY, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 3-27.

PEISSIG, J. J., YOUNG, M. E., WASSERMAN, E. A., & BIEDERMAN, I. (1999). The pigeon's perception of depth-rotated shapes. *Current Psychology of Cognition*, **18**, 657-690.

PEISSIG, J. J., YOUNG, M. E., WASSERMAN, E. A., & BIEDERMAN, I. (2000). Seeing things from a different angle: The pigeon's recognition of single geons rotated in depth. *Journal of Experimental Psychology: Animal Behavior Processes*, **26**, 115-132.

PHELPS, M. T., & ROBERTS, W. A. (1994). Memory for pictures of upright and inverted primate faces in humans (*Homo sapiens*), squirrel monkeys (*Saimili sciureus*), and pigeons (*Columba livia*). *Journal of Comparative Psychology*, **108**, 114-125.

POWELL, R. W. (1967). The pulse-to-cycle fraction as a determinant of critical flicker fusion in the pigeon. *Psychological Record*, **17**, 151-160.

REID, S. L., & SPETCH, M. L. (1998). Perception of pictorial depth cues by pigeons. *Psychonomic Bulletin & Review*, **5**, 698-704.

SPETCH, M. L., FRIEDMAN, A., & REID, S. L. (2001). The effect of distinct parts on recognition of depth-rotated objects by pigeons (*Columba livia*) and humans. *Journal of Experimental Psychology: General*, **130**, 238-255.

SPETCH, M. L., KELLY, D. M., & LECHELT, D. P. (1998). Encoding of spatial information in images of an outdoor scene by pigeons and humans. *Animal Learning & Behavior*, **26**, 85-102.

TARR, M. J., & PINKER, S. (1989). When does human object recognition use a viewer-centered reference frame? *Psychological Science*, **1**, 253-256.

THOMPSON, P. (1980). Margaret Thatcher: A new illusion. *Perception*, **9**, 483-484.

TROJE, N. F., HUBER, L., LOIDOLT, M., AUST, U., & FIEDER, M. (1999). Categorical learning in pigeons: The role of texture and shape in complex static stimuli. *Vision Research*, **39**, 353-366.

ULLMAN, S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition*, **32**, 193-254.

ULLMAN, S. (1998). Three-dimensional object recognition based on the combination of views. *Cognition*, **67**, 21-44.

VALENTINE, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, **79**, 471-491.

VAUGHAN, W., & GREENE, S. L. (1984). Pigeon visual memory capacity. *Journal of Experimental Psychology: Animal Behavior Processes*, **10**, 256-271.

VETTER, T., HURLBERT, A., & POGGIO, T. (1995). View-based models of 3D object recognition: Invariance to imaging transformations. *Cerebral Cortex*, **5**, 262-269.

VON FERSEN, L., & LEA, S. E. G. (1990). Category discrimination by pigeons using five polymorphous features. *Journal of the Experimental Analysis of Behavior*, **54**, 69-84.

WASSERMAN, E. A., GAGLIARDI, J. L., COOK, B. R., KIRKPATRICK-STEGER, K., ASTLEY, S. L., & BIEDERMAN, I. (1996). The pigeon's recognition of drawings of depth-rotated stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, **22**, 205-221.

WATANABE, S. (1993). Object–picture equivalence in the pigeon: An analysis with natural concept and pseudo-concept discriminations. *Behavioural Processes*, **30**, 225-232.

WATANABE, S., LEA, S. E. G., & DITTRICH, W. H. (1993). What can we learn from experiments on pigeon concept discrimination? In H. P. Zeigler & H. J. Bischof (Eds.), *Vision, brain, and behavior in birds* (pp. 351-376). Cambridge, MA: MIT Press.