

The Effect of Distinctive Parts on Recognition of Depth-Rotated Objects by Pigeons (*Columba livia*) and Humans

Marcia L. Spetch, Alinda Friedman, and Sheri L. Reid
University of Alberta

To explore whether effects observed in human object recognition represent fundamental properties of visual perception that are general across species, the authors trained pigeons (*Columba livia*) and humans to discriminate between pictures of 3-dimensional objects that differed in shape. Novel pictures of the depth-rotated objects were then tested for recognition. Across conditions, the object pairs contained either 0, 1, 3, or 5 distinctive parts. Pigeons showed viewpoint dependence in all object-part conditions, and their performance declined systematically with degree of rotation from the nearest training view. Humans showed viewpoint invariance for novel rotations between the training views but viewpoint dependence for novel rotations outside the training views. For humans, but not pigeons, viewpoint dependence was weakest in the 1-part condition. The authors discuss the results in terms of structural and multiple-view models of object recognition.

Animals are consistently faced with the problem of recognizing objects from various perspectives in the natural environment, either because of their own movements or because of the movements of some external object. Finding one's way back to the nest, recognizing a familiar conspecific, and avoiding predators would all appear to benefit from quick and accurate object recognition regardless of the position of the object or of the viewer. Nonetheless, processes of object recognition across variations in viewpoint have not received as much experimental attention in animals as in humans.

The research in this article examines recognition of depth-rotated objects in pigeons (*Columba livia*) and humans to address several basic questions. First, from the perspective of avian visual cognition, we examined whether pigeons would show viewpoint dependence (see, e.g., Tarr & Pinker, 1989) in an object discrimination task and whether either its occurrence or degree would change as a function of the number of parts in an object. Second, we examined whether the number of distinctive object parts would affect the accuracy and/or latency of pigeons' discriminative choice behavior. Third, we observed whether pigeons' discrimination performance would decrease systematically as a function of the degree of rotation of an object from its nearest training view.

Fourth, from the perspective of human visual cognition, we examined the effect of number of distinctive parts on humans' object recognition within a simultaneous discrimination task, thereby testing the generality of results recently obtained by Tarr, Bulthoff, Zabinski, and Blanz (1997) with a different procedure. Our simultaneous discrimination task differs from both the sequen-

tial matching task and the naming task used by Tarr et al. in that, for our task, a target and comparison object are physically present at the same time, thereby allowing a direct comparison between objects. However, like the naming task, the simultaneous discrimination task entails a learning process (to determine which object is correct) as well as repeated activation of long-term object representations (to match presented objects to representations of the correct and incorrect objects). Fifth, we examined whether viewpoint invariance in our task would be obtained for a novel view of an object that is between two training views of the object. If so, this would provide further support for the notion that, when trained consistently on two views of an object, humans may either acquire or be able to rapidly generate representation of views between the trained views (Bulthoff & Edelman, 1992).

Finally, from a comparative perspective, this study allows us to determine whether the effects seen in pigeons are similar to those seen in human observers tested in a similar task and with the same stimuli. This question is interesting not only from the perspective of animal cognition but also for advancing understanding of basic object recognition processes. Specifically, this question addresses whether object recognition effects seen in humans represent fundamental properties of visual perception that are general across species or whether the effects reflect specialized cognitive processes that vary depending on phylogenetic or experiential factors.

Whether object recognition is viewpoint dependent or independent has been intensively investigated in the human literature (e.g., see Biederman & Gerhardstein, 1993; Edelman & Bulthoff, 1992; Jolicoeur & Humphrey, 1998; Tarr & Pinker, 1989). Contemporary examples of the viewpoint-independent approach are the recognition-by-components theory (Biederman, 1987) and its more recent version, the geon-structural-description (GSD) theory (Biederman & Gerhardstein, 1993, 1995). Both versions hold that objects are represented in memory as structural descriptions composed of the spatial relations among simple, volumetric, three-dimensional (3-D) parts called geons. The particular geons composing an object and their particular arrangements are equally important for recognition. Thus, provided that (a) an object's image can be decomposed into parts (geons), (b) the arrangement

Marcia L. Spetch, Alinda Friedman, and Sheri L. Reid, Department of Psychology, University of Alberta, Edmonton, Alberta, Canada.

This research was supported by Natural Sciences and Engineering Research Council of Canada research grants. We thank S. Doerksen, K. Steinbring, and S. Nystrom for assistance with the research.

Correspondence concerning this article should be addressed to Marcia L. Spetch, Department of Psychology, University of Alberta, Edmonton, Alberta T6E 2E9, Canada. Electronic mail may be sent to mspetch@ualberta.ca.

of the parts provides a structural description that is distinct from other arrangements, and (c) the object has this same structural description across different views, viewpoint-independent recognition should be attained. The behavioral signature of viewpoint independence is thus a recognition function that does not change with distance from the training view.

In contrast to the viewpoint-independent approach, multiple-view theories (Bulthoff & Edelman, 1992; Humphrey & Khan, 1992; Tarr, 1995; Tarr & Bulthoff, 1995; Tarr & Pinker, 1989) assume that objects are encoded in the specific pose in which they are seen, using a viewpoint-dependent frame of reference. This implies that objects are represented as a collection of familiar views that reflect the precise metric (appearance) properties of these particular viewpoints. Except in cases in which diagnostic features are available, recognition varies systematically with distance to the nearest represented viewpoint. This prediction is made whether recognition of novel views is assumed to depend on (a) some form of generalization from stored views, (b) a transformation of the novel percept to match the frame of reference (viewpoint) of the nearest stored view (Tarr & Pinker, 1989), or (c) the rate of accumulation of activity from neurons selective for the stored object, which should vary systematically as a function of distance between the novel and stored views (see, e.g., Perrett, Oram, & Ashbridge, 1998). In any case, the behavioral signature of the multiple-view approach is that either reaction time increases, accuracy decreases, or both as a function of rotation away from the encoded view(s). More recently, Tarr and Bulthoff (1995; see also Hayward & Tarr, 1997) stated that viewpoint-dependent mechanisms should be expected to dominate recognition when discriminating between visually similar objects, whereas viewpoint-independent mechanisms should predominate when discriminating between visually dissimilar objects.

These two approaches make quite different assumptions about the processes underlying object recognition, but their predictions are not always easy to dissociate. Specifically, both the GSD model and the multiple-view model predict viewpoint-dependent recognition when discriminating between views of the same object that differ in the global shape of their projected image on the retina (or beyond). That is, GSD theory explicitly claims that to achieve viewpoint-independent recognition, the two views of an object must result in the same structural description; with different global images, this is unlikely, though not impossible (see Biederman and Gerhardstein, 1993, Experiment 1). Thus, for the GSD model, viewpoint dependence arises when one of the necessary conditions for viewpoint invariance—in this case, that the different views of the object elicit the same distinctive structural description—has not been met. In the case of the multiple-view model, however, viewpoint dependence is a natural consequence of the basic process of normalization to the nearest stored viewpoint.

On the other hand, both theories predict viewpoint-independent object recognition when discriminating between views of an object that contain a local diagnostic feature. For the GSD model, viewpoint invariance is expected whenever the three fundamental conditions of the basic object recognition process described earlier are met. In the case of a local diagnostic feature, it is likely that the geon that represents that feature is distinctive across all presented views because of the nonaccidental properties that are hypothesized to define geons in the first place (see Biederman, 1987; but see Hayward & Tarr, 1997). For the multiple-view model, the presence of a local diagnostic feature allows for viewpoint-

independent object recognition because the feature, particularly if familiar, precludes the need for the viewpoint-dependent normalization process. For example, Jolicoeur (1985, 1988) has hypothesized that novel percepts are normalized to a familiar orientation to be recognized, whereas, after practice, participants may use diagnostic features to recognize objects without normalization. Thus, the literature on object recognition in humans is extensive in both data and theory, but consensus has not yet been reached on whether object recognition is best conceptualized as viewpoint invariant or viewpoint dependent.

Although object recognition across variations in viewpoint has not been studied as extensively in animals as in humans, it is beginning to enjoy considerable experimental attention (e.g., Cook & Katz, 1999; Logothetis, Pauls, Bulthoff, & Poggio, 1994; Wasserman et al., 1996). Interestingly, some of the theoretical controversies in the human literature have parallels in the literature on object recognition in pigeons (e.g., see Cerella, 1990; Wasserman et al., 1996).

Pigeons are visually sophisticated animals with highly developed abilities to discriminate, categorize, and remember visual details presented in pictures (see Herrnstein, 1979; Vaughan & Greene, 1984). Nevertheless, the literature on pigeons' ability to recognize objects across variation in viewpoint has produced a mixed pattern of results with some studies showing strong generalization across viewpoints and other studies showing viewpoint-specific object recognition.

Perspective processing in the pigeon has been investigated with both pictures of objects rotated in the frontal plane and pictures of objects rotated in depth. Whereas objects rotated in the frontal plane change only in orientation, objects rotated in depth are distorted in shape as well. These distortions may consist of changes in metric properties, such as distance and angle between edges, or changes in structural properties that are due to certain parts of the object being occluded and revealed as it is rotated around an axis. For both types of transformations, the results of studies on orientation invariance in the pigeon have been inconsistent.

Studies investigating animals' recognition of two-dimensional shapes across rotations in the frontal plane have typically used one of two approaches. The first is to train participants on a shape discrimination with the stimuli presented in a particular orientation and then to test for transfer of the learned discriminations when the stimuli are presented in novel orientations. Several studies of this type have found either little or no transfer to novel orientations (see, e.g., Towe, 1954) or transfer to some orientations but not to others (see, e.g., Vetter & Hearst, 1968). For example, Morgan, Fitch, Holman, and Lea (1976) found poor transfer of discriminative responding by pigeons to most novel orientations of the letter A, although some transfer was shown to 180° orientations.

The second approach has been to train pigeons to select an object that matches a sample object and to investigate the effect of rotating the comparison objects to produce orientation disparity. Using this procedure, Hollard and Delius (1982) and Delius and Hollard (1995) found that pigeons showed rotational invariance (measured in terms of both accuracy and reaction time) both when the incorrect comparisons were mirror images of the objects and when the incorrect comparisons were arbitrarily different shapes. Humans, on the other hand, showed longer latencies with increasing angular disparity between the sample and comparison when the

incorrect comparison was a mirror image but not when it was an arbitrarily different shape (Delius & Hollard, 1995).

Recognition studies using depth-rotated stimuli have also produced variable results in the pigeon. Cerella (1977, 1990) has reported that pigeons are unable to recognize objects across rotational transformations. Cerella (1977) reported that, even after extensive training, pigeons were unable to discriminate line drawings of 3-D cubes rotated in depth from random distortions of the cube. Cerella (1990) trained pigeons to discriminate two-dimensional parallelogramlike objects from random distortions of parallelogramlike objects and then conducted transfer tests with the positive stimulus (S+) rotated in depth around the *x*- and *y*-axes. Results indicated that the pigeons perceived the rotated S+ as being equivalent to the negative stimuli (S-). In contrast, other researchers have found that pigeons can generalize to novel depth rotations and that their discriminative performance decreases as an object is rotated further away from the training orientation (Lumsden, 1970, 1977; Lumsden & Pullen, 1970; Reynolds, 1961; Wasserman et al., 1996). For example, Lumsden (1970) trained pigeons to discriminate between real 3-D objects, each presented at a single orientation. The S+ was a rectangular object with a semicircle protruding from it, and the S- objects were two differently shaped objects of approximately the same dimensions (a wedge-shaped object and a flat hourglass-shaped object). Generalization to four novel depth orientations of the S+ was tested; responding to rotations of the S+ was found to decrease systematically as it was rotated further away from the training orientation. Once again, generalization gradients showed secondary peaks at the mirror-image orientation of the S+, a result that has also been found in monkeys (see, e.g., Logothetis et al., 1994) and humans (see, e.g., Tarr & Gauthier, 1998).

Wasserman et al. (1996) conducted an extensive study of pigeons' generalization to novel depth rotations of line drawings of 3-D objects. In one of their experiments, pigeons were trained to discriminate line drawings of four distinct objects (i.e., an airplane, a chair, a lamp, and a flashlight) presented at a single training orientation. Choice accuracy on tests with novel orientations was well above chance, but accuracy decreased as a function of rotation away from the training view. Increasing the number of training views from one to three increased the degree of generalization to novel rotations. It should be noted that the objects in their study had diagnostic features; that is, each object had certain distinctive features or components that were not present in the other objects. Consequently, generalization might have been based on these features alone, rather than on features and relations of the entire object.

Taken together, the results of studies of object recognition in pigeons suggest that pigeons can, but do not always, display recognition of novel objects presented in depth rotations from the training views. Most typically, pigeons' recognition of depth-rotated objects displays viewpoint dependency, with recognition accuracy decreasing as the object is rotated further from its training view.

A design used in a recent study on humans' object recognition by Tarr et al. (1997) offers a particularly informative way to further explore object recognition in pigeons and to assess the extent to which it is similar to or different from that found in humans. The Tarr et al. experiment was designed to test *differential predictions of the view-based versus structural-description approach*. Object recognition across changes in viewpoint was

compared in four conditions. In one condition, the objects were papercliplike objects with no distinctive parts added. Both the view-based and geon theories predict viewpoint-dependent object recognition for these objects. The remaining three conditions had either one, three, or five distinctive parts (geons) added to the objects. In the one-part condition, the 10 objects in the set each contained a different part, and thus, the added part was a locally diagnostic feature. In the three- and five-part conditions, the added parts differed in each object (e.g., a three-part object might have a pyramid, a rectangular solid, and a cone) but were shared across objects so that no single part was diagnostic of an object's identity across the set. According to view-based theories, only the one-part condition should show viewpoint-invariant recognition because only that condition contains a locally diagnostic feature. According to geon theory, all three of these conditions should produce viewpoint-invariant recognition because each unique stimulus should result in the same structural description regardless of viewpoint and because the number of different parts/geons in an object should not affect object recognition (Biederman, 1987). In support of the view-based theory, strong viewpoint dependence was seen in the zero-part, three-part, and five-part conditions. Viewpoint dependence was weaker, although not completely absent, in the one-part condition.

We adopted the Tarr et al. (1997) design but used a simultaneous discrimination procedure that lends itself particularly well to investigations with pigeons. In addition, each of our participants experienced all four object-part conditions, whereas the participants in Tarr et al.'s study experienced only the baseline (zero-part) condition and one test condition (one-, three-, or five-part). Of primary interest was whether pigeons' recognition of depth-rotated objects is similar to humans', in terms of both sensitivity to degree of rotation from the training views and the number of distinctive object parts.

Experiment 1

In this experiment, both pigeons and humans were trained to discriminate between pairs of shaded 3-D objects presented on a computer monitor. Each pair of objects was trained at two depth rotations and then tested at four novel depth rotations. Each participant was trained and then tested in four experimental conditions. As in the Tarr et al. (1997) study, these experimental conditions differed in whether the objects contained either zero, one, three, or five distinctive parts.

Method

Participants. Eight adult pigeons with varied experimental histories served as participants. None had previously served in a discrimination task involving rotated objects. All birds were maintained at approximately 85% of their free-feeding body weights and housed individually under a 12-hr light/dark cycle. Food was obtained during and after experimental sessions. Water and grit were obtained ad libitum in their home cages.

The human participants were 16 undergraduate students (2 men and 14 women) who received extra credit toward their introductory psychology class for participating in the experiment. The participants ranged in age from 17 to 29 years.

Apparatus for pigeons. The stimuli were displayed on Zenith (model no. 1490 or 1492; St. Joseph, MI) flat-screen color monitors (with a viewing area of approximately 26 × 20 cm) with attached touch frames (Carroll Touch model no. 1490 or 1492 Smart Frames; Round Rock, TX).

The touch frames had a resolution of 3 mm and were programmed to detect individual responses (i.e., detection of a beam break and then a return to unbroken beams before another response was recorded). A microcomputer located in an adjacent room controlled stimuli and recorded the peck locations sensed by the touch frame.

Testing was conducted in one of two operant chambers. Chamber 1 was an operant chamber with a 28 × 20-cm opening cut into the back wall for the monitor. The monitor screen rested 9 cm above the chamber floor and was covered with a thin sheet of Plexiglas. Spacers recessed the touch frame by approximately 3 cm and separated the frame from the monitor by approximately 1.5 cm. On each side wall of the chamber, a pigeon feeder was attached 8 cm from the back wall and 9 cm from the floor. Chamber 2 was a large custom-built chamber 44 cm high, 32 cm deep, and 74 cm wide (inside dimensions). The monitor was placed against an opening centered in the back wall, 10 cm above the chamber floor. Spacers recessed the touch frame by approximately 3 cm from the opening and separated the frame from the monitor by approximately 1.6 cm. Two pigeon feeders were mounted on the back wall, one on each side of the monitor. Each feeder opening began 8.5 cm from the side of the monitor opening and 17 cm from the floor. In both chambers, lamps located within each feeder illuminated feeder presentations, and photocells measured head entries into the hoppers to limit feeding duration to 2 s per food presentation.

Apparatus for humans. Stimuli were presented on the Zenith model no. 1490 flat-screen color monitor described above. Participants were seated approximately 45 cm from the monitor in a small private room.

Stimuli. Using Adobe Dimensions, we created four stimulus sets: zero-part, one-part, three-part, and five-part objects (see Figures 1, 2, and 3). The stimuli were created to be similar to those used by Tarr et al. (1997). Each stimulus set consisted of two objects, with one object designated as the S+ and the other object designated as the S-.

Training Orientations

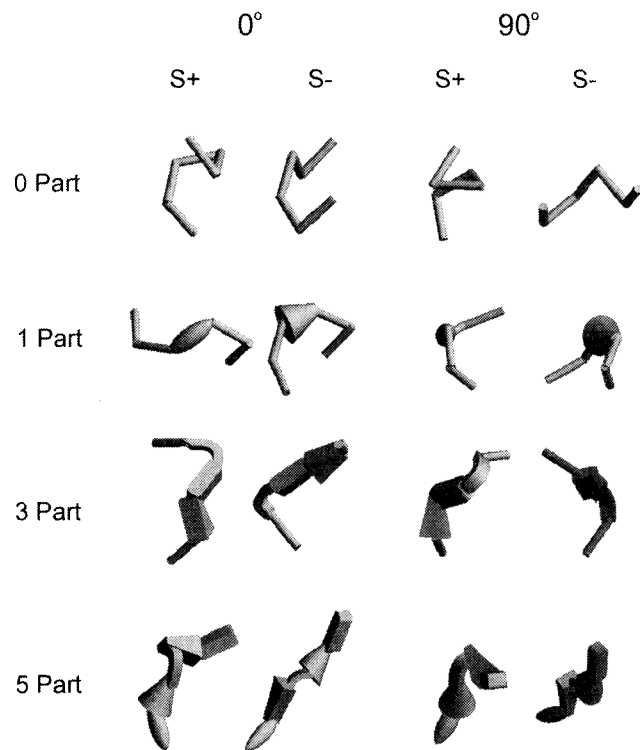


Figure 1. Positive (S+) and negative (S-) objects at each of the two training orientations in each object-part condition.

Novel Orientations

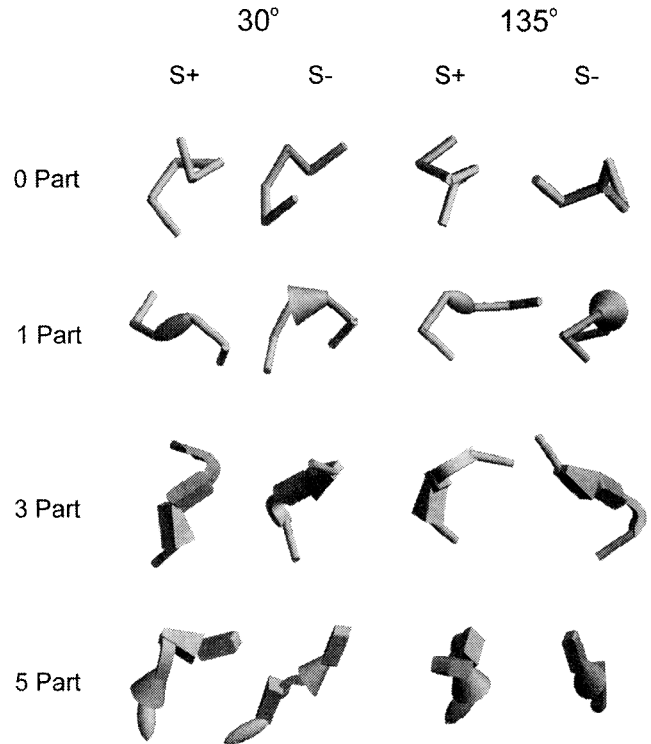


Figure 2. Positive (S+) and negative (S-) objects at two of the novel orientations used in testing for each object-part condition.

Several images were created within each stimulus set. For each image, the S+ object and the S- object were presented side by side. The orientation of the S+ was the same as the orientation of the S- during both training and testing. The initial orientation of each object was oblique with respect to the frontal-parallel (picture) plane and was arbitrarily assigned as the 0° orientation. Across images, the objects were then rotated around the y-axis so that there were two training depth orientations (0° and 90°) and four testing depth orientations (30°, 135°, 160°, and 225°). When objects that are initially oblique with respect to the picture plane are rotated around the vertical axis, they may appear to be tumbling around a different axis. That is, although an object that is vertical with respect to the picture plane appears to spin around the y-axis, oblique objects may appear to tumble rather than spin, giving the impression that they are being rotated around a different axis. Nevertheless, operationally, the objects were rotated around the y-axis. Zero-part objects consisted of five tubelike, nondistinct, 3-D volumes that were linked end-to-end at varying angles. The one-part, three-part, and five-part objects, respectively, had one, three, or five qualitatively distinct volumes (geons) inserted into the middle of the objects. For the one-part objects, a different geon was inserted into each of the two objects. For the three-part and five-part object pairs, the same set of geons was inserted into each object of the pair, but the geons were arranged in differing configurations and with differing angles of attachment (for example, the three-part objects differed in whether the rectangular object was attached to the narrow or the wide end of the triangular object).

Procedure for pigeons. All birds had previous experience in touch-screen tasks but were given one or more sessions with a modified autoshaping procedure to reestablish reliable pecking. On each trial of these peck-shaping sessions, one of the training stimulus pairs was presented either until a peck at the positive object was made or for a maximum of 8 s,

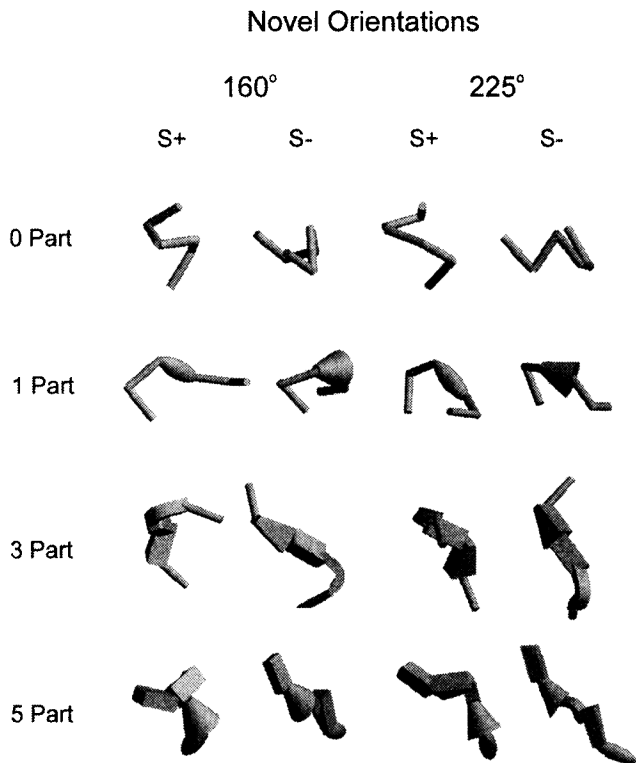


Figure 3. Positive (S+) and negative (S-) objects at the remaining two novel orientations used in testing for each object-part condition.

and then food was presented. Trials were separated by a 40-s intertrial interval (ITI).

Each bird was then trained and tested in each of four conditions that differed in type of objects (zero-part, one-part, three-part, or five-part). The order of exposure to the four conditions was counterbalanced across pigeons.

The task was a simultaneous discrimination between an S+ and an S- object. Each trial began with the presentation of a stimulus pair that remained on until the pigeon pecked at one of the two objects. During training sessions, pecks to the S+ object were always rewarded with food, and pecks to the S- object were unrewarded. The objects were presented at 0° on half of the training trials and at 90° on the remaining training trials. For each orientation, the S+ was equally often on the left and right. To prevent a position bias, incorrect choices were followed by correction trials in which the same stimulus pair was represented until the pigeon made a correct choice. Correction trials were not used in the determination of accuracy. Each session consisted of 120 trials, separated by a 2-s ITI during which the monitor was blank. Training continued until the bird met an accuracy criterion of 80% or greater over a block of five consecutive sessions.

Three test sessions followed completion of each training condition. Each test session consisted of 60 reinforced baseline trials at the training orientations and 60 unreinforced test trials. On test trials, the object pairs were presented equally often at the two training orientations (0° and 90°) and at four novel orientations (30°, 135°, 160°, 225°). For each trial, the two simultaneously presented objects were always displayed in the same orientation, and hence, they were both shown in either a novel or a trained orientation. At each orientation, the S+ was equally often on the left and right.

Sessions were conducted at approximately the same time each day, 5 or 6 days per week. Sessions lasted until all scheduled trials were com-

pleted or for a maximum of 1 hr. The monitor screen was cleaned at the end of each day and between sessions as needed.

Procedure for humans. The participant was seated at the computer and provided with verbal instructions that included the following information:

On each trial, a pair of objects will be presented, one of which is correct and the other of which is wrong. You should use the arrow key to choose the object on the right or left. A point will be awarded for choice of the correct object but not for choice of the incorrect object. A double point will be earned if the correct object is chosen quickly. On most trials, choices will be followed by feedback that indicates whether the choice was correct or wrong and any points earned. After reading the feedback, you should press the enter key to advance to the next image. You should use the feedback to learn which object is correct and then should try to respond as quickly and accurately as possible. Occasionally, the trial will end without feedback, regardless of whether the choice was correct or incorrect.

The experimenter then started the program and left the room. The simultaneous discrimination task was similar to that used for pigeons except that (a) the choice response consisted of pressing the right or left arrow key on the keyboard and (b) reinforcement consisted of points. Participants received 1 point for each correct choice. An extra point was awarded if the participant chose correctly within 0.5 s of onset of the stimulus display. The session consisted of four blocks, each with 20 training trials followed by 20 test trials. The zero-part, one-part, three-part, and five-part objects (identical to those used for pigeons) were trained and then tested in one block each, with the order of exposure to the four object types counterbalanced across participants. During the 20 training trials, the objects were presented equally often at 0° and 90°, and, within each orientation, equally often with the S+ on the left or right. The test block consisted of a mixture of 4 training trials at each of the two training orientations and 12 test trials. During the 12 test trials, the objects were presented for 2 trials each (once with the S+ on the right and once with the S+ on the left) at the training orientations (0° and 90°) and four novel orientations (30°, 135°, 160°, and 225°). On test trials, the trial ended with the message "No feedback trial." Order of presentation of trial types within each block was randomly determined. The accuracy and latency of each response were recorded. All data presented are from test trials.

Results

An alpha level of .05 was used for all statistical tests.

Accuracy. The left panels of Figure 4 show accuracy as a function of orientation for each object-part condition. Results for humans are shown with solid lines, and results for pigeons are shown with dashed lines. Both similarities and differences in the patterns of results for pigeons and humans are apparent. Strong viewpoint dependence is shown by both species for the zero-part and three-part objects. Pigeons also showed strong viewpoint dependence for the one-part and five-part objects. Humans showed some viewpoint dependence for the five-part objects but no viewpoint dependence for the one-part objects. These observations were confirmed by a mixed-design analysis of variance (ANOVA), with species as the between-subjects factor and parts and orientation as within-subjects factors. The ANOVA revealed significant main effects of orientation, $F(5, 110) = 44.25$, and parts, $F(3, 66) = 6.54$, but not of species, $F(1, 22) = 2.36$. In addition, all of the two-way interaction terms were significant (Orientation \times Parts, $F(15, 330) = 2.65$; Orientation \times Species, $F(5, 110) = 2.92$; Parts \times Species, $F(3, 66) = 7.60$), as was the three-way interaction, $F(15, 330) = 2.20$.

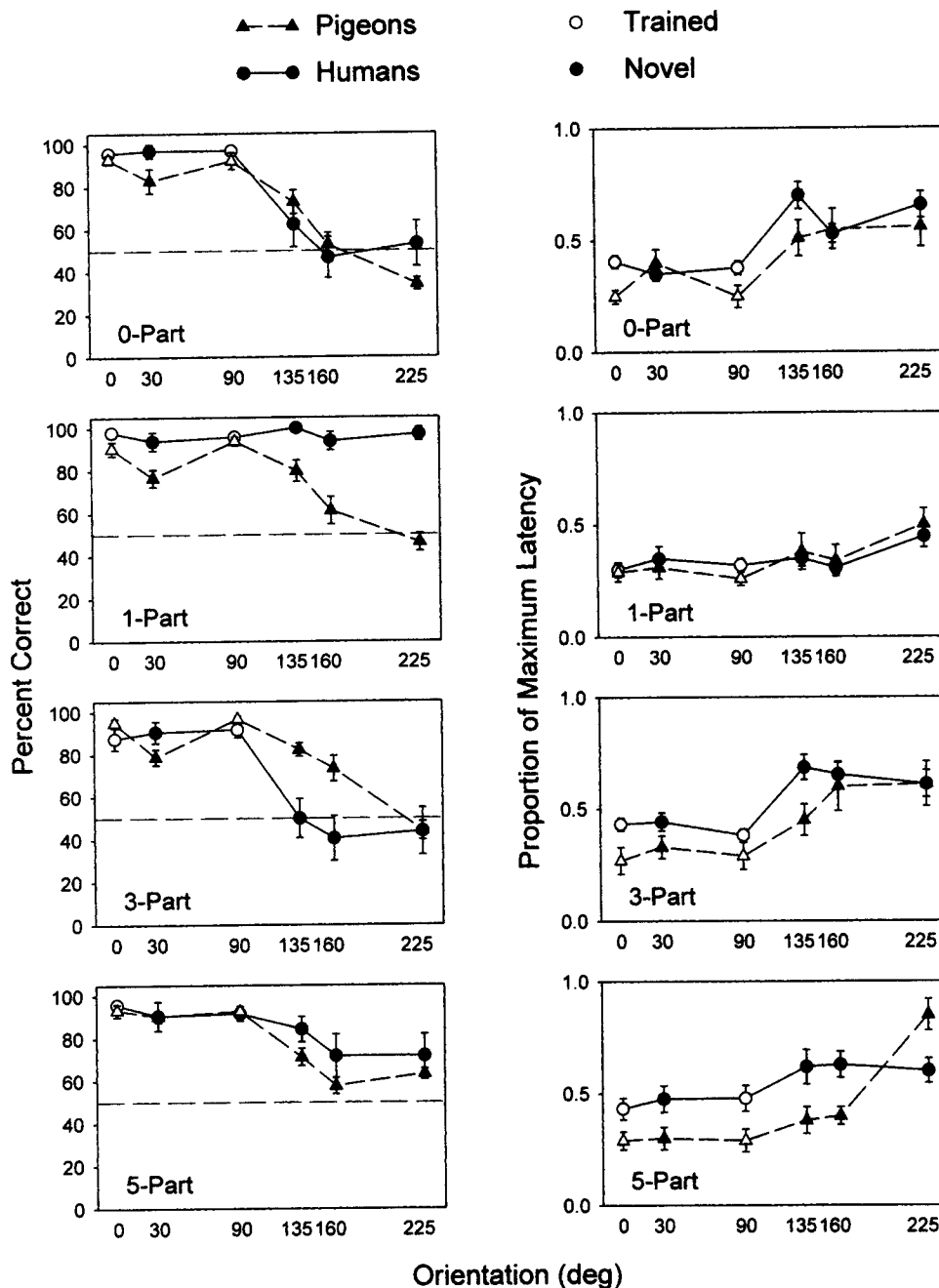


Figure 4. Choice accuracy (left panels) and proportion of maximum latency (right panels) in pigeons and humans on test trials with trained and novel orientations in the four object-part conditions in Experiment 1. Error bars show standard error of the mean (absent error bars indicate that error was smaller than the symbol size).

Similarities and differences in the patterns of results shown by pigeons and humans are particularly apparent when one looks at accuracy as a function of rotation away from the trained values. The left panels of Figure 5 show accuracy scores for both species in each object-part condition as a function of the degree of rotation away from the nearest trained orientation. The data point shown as 0° rotation is the average of the accuracy scores at the two trained (0° and 90°) orientations. The 30° rotation represents the novel orientation that fell between the two trained values; its nearest

training orientation was 0°. The remaining stimuli fell outside the range of training stimuli and so are shown as the degree of rotation from the 90° training orientation.

In general, pigeons showed orderly generalization functions in each object-part condition, with accuracy declining as the stimulus was rotated farther away from the nearest trained value. In contrast, humans showed perfect transfer to the novel rotation that fell between the two trained orientations in all conditions, but their accuracy was at chance to all values outside the training range for

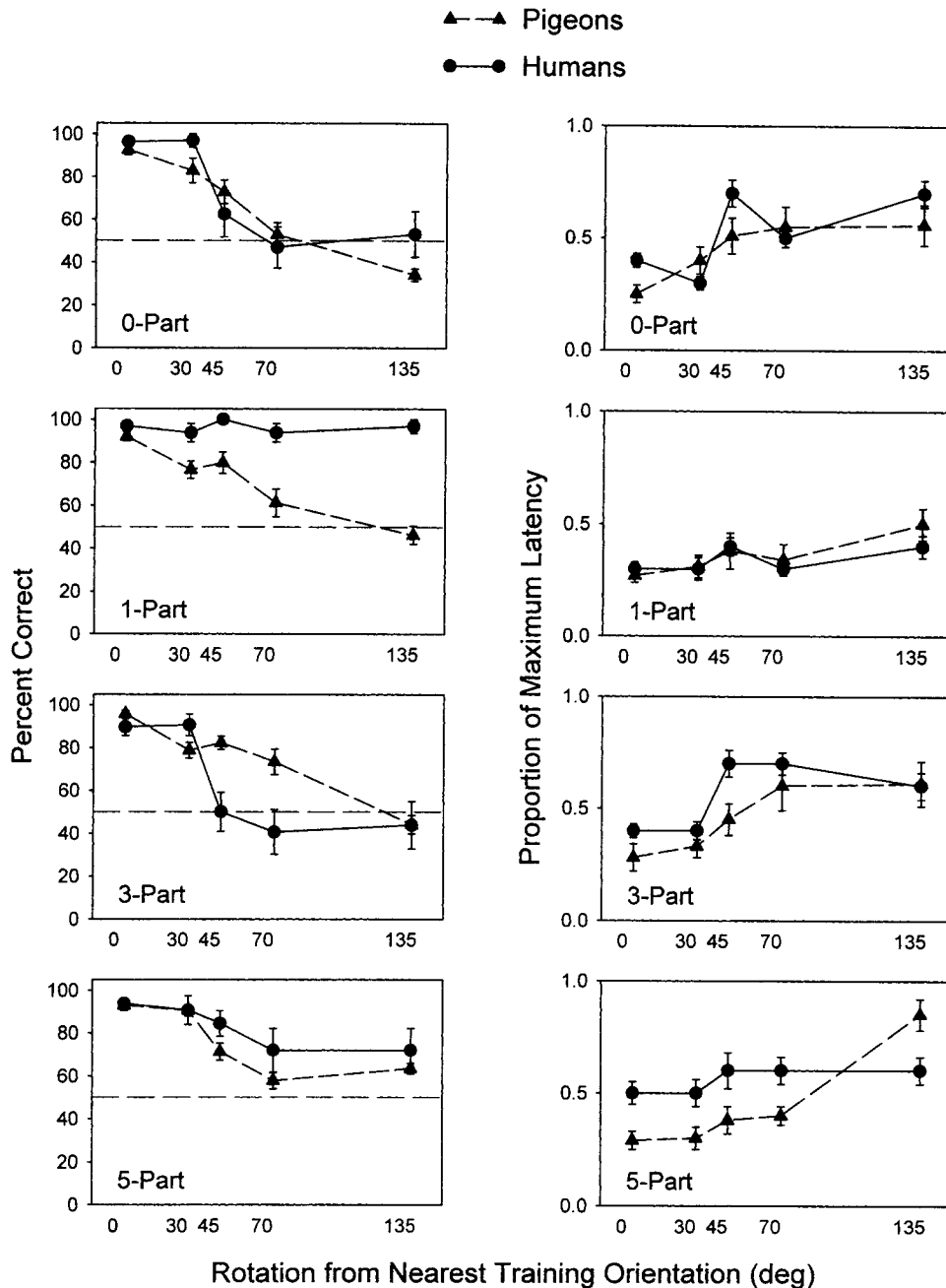


Figure 5. Choice accuracy (left panels) and proportion of maximum latency (right panels) in pigeons and humans as a function of rotation from the nearest training view in the four object-part conditions in Experiment 1. The 0° rotation is the average accuracy on test trials with the two trained views. Error bars show standard error of the mean (absent error bars indicate that error was smaller than the symbol size).

the zero-part and three-part objects. Humans showed a smaller decline in accuracy for the five-part objects and no decline in accuracy for the one-part objects.

A mixed-design ANOVA on these data revealed significant main effects of rotation, $F(4, 88) = 37.89$, and parts, $F(3, 66) = 7.19$, as well as significant interactions between rotation and parts, $F(12, 264) = 2.42$, rotation and species, $F(4, 88) = 2.99$, parts and species, $F(3, 66) = 7.61$, and rotation, parts, and species, $F(12, 264) = 2.10$.

To further analyze these interactions, we conducted ANOVAs on the data for each species separately. Pigeons showed significant main effects of rotation, $F(4, 28) = 86.09$, and parts, $F(3, 21) = 4.76$, and a significant Rotation \times Parts interaction, $F(12, 84) = 3.05$. Subsequent polynomial contrasts revealed a significant linear trend for rotation within each part condition: $F(1, 7) = 226.80, 48.11, 141.06,$ and 104.17 for the zero-part, one-part, three-part, and five-part conditions, respectively. A quadratic trend was significant only for the five-part condition, $F(1, 7) = 20.53$.

Humans also showed significant main effects of rotation, $F(4, 60) = 17.58$, and parts, $F(3, 45) = 14.88$, and a significant Rotation \times Parts interaction, $F(12, 180) = 3.76$. However, only the zero-part and three-part conditions showed significant linear effects, $F(1, 15) = 18.03$ and 23.54 , respectively. A quadratic trend was also significant for only the zero-part and three-part conditions, $F(1, 15) = 7.80$ and 7.59 , respectively.

Latency. Table 1 shows mean choice latencies at each orientation in each object-part condition for pigeons and humans. Both correct and incorrect trials are included in these scores, and hence, the measure represents the overall time taken to choose between the stimuli, whether correctly or incorrectly. These absolute latencies could not be meaningfully compared across species because of differences in the response measures (pecking vs. pressing the keyboard) and instructions (i.e., humans, but not pigeons, were instructed to respond as quickly as possible). Therefore, we scaled the latency data to meaningfully compare how latencies varied as a function of orientation and object parts across species. For each participant, we identified which of the 24 trial types (six orientations in each of four object-part conditions) had the highest average latency. We then transformed the latency scores for each trial type into a proportion of this maximum score and analyzed these proportional latencies analogously to the accuracy data. The proportional latencies are plotted for both humans and pigeons in the right panels of Figure 4. Again, both similarities and differences in the patterns of results between species emerged. The mixed-design ANOVA revealed significant main effects of orientation, $F(5, 110) = 42.17$, and parts, $F(3, 66) = 7.77$, but not of species, $F(1, 22) = 3.12$. There was a significant two-way interaction between orientation and species, $F(5, 110) = 4.44$, and between orientation and parts, $F(15, 330) = 3.75$, as well as a significant three-way interaction between orientation, parts, and species, $F(15, 330) = 2.91$.

The right panels of Figure 5 plot proportional latencies at each novel orientation as a function of degree of rotation from the nearest trained orientation. With the exception of the one-part condition, pigeons showed an orderly increase in latency as a function of distance from the trained orientation, whereas humans showed a more abrupt increase in latency for novel orientations that fell outside the training range, particularly in the zero- and three-part conditions. In the one-part condition, latencies for both species increased only at the farthest rotation.

A mixed-design ANOVA on these data revealed significant main effects of rotation, $F(4, 88) = 36.60$, and parts, $F(3, 66) = 8.20$, as well as significant interactions between rotation and parts, $F(12, 264) = 3.64$, rotation and species, $F(4, 88) = 4.63$, and rotation, parts, and species, $F(12, 264) = 3.15$.

To further analyze these interactions, we conducted ANOVAs on the data for each species separately. Pigeons showed a significant main effect of rotation, $F(4, 28) = 86.09$, and a significant Rotation \times Parts interaction, $F(12, 84) = 3.51$. Subsequent polynomial contrasts revealed a significant linear trend within each part condition: $F(1, 7) = 16.88, 14.72, 16.89$, and 39.57 for the zero-part, one-part, three-part, and five-part conditions, respectively. The quadratic trend was significant only in the five-part condition, $F(1, 7) = 9.48$.

Humans showed significant main effects of rotation, $F(4, 60) = 23.84$, and parts, $F(3, 45) = 9.10$, and a significant Rotation \times Parts interaction, $F(12, 180) = 3.55$. Polynomial contrasts revealed a significant linear trend within each condition: $F(1, 15) = 24.03, 13.01, 8.20$, and 5.09 for the zero-part, one-part, three-part, and five-part conditions, respectively. The quadratic trend was also significant in the one-part, three-part, and five-part conditions, $F(1, 15) = 8.32, 18.27$, and 8.10 , respectively.

Rotation within versus outside trained orientations. In a final set of analyses, we focused on a subset of the data to examine pigeons' and humans' recognition of the novel orientations that fell between versus outside the trained values. Performance on the novel orientation between the trained values is particularly interesting for theoretical reasons because of implications for object recognition processes. One approach holds that participants acquire a viewpoint-specific representation for each object orientation experienced during training. When confronted with a novel orientation, a normalization process, such as mental rotation (Cooper & Shepard, 1973; Shepard & Metzler, 1971) from the nearest stored view (Tarr & Pinker, 1989), is required to align the novel orientation with stored representations (for a recent discussion, see Jolicoeur & Humphrey, 1998). Except in cases for which recognition can be based on features and does not require attention to the spatial arrangement of elements, this normalization process should produce a systematic decrement in recognition speed as a function of rotation from the nearest trained view. Thus, in the three conditions of the present study that require attention to the spatial arrangement of features for object recognition (i.e., the zero-

Table 1
Mean (SEM) Choice Latency (in Seconds) as a Function of Orientation in Each Object-Part Condition for Humans and Pigeons in Experiment 1

Participants and condition	Orientation					
	0 ^a	30	90 ^a	135	160	225
Humans						
0-part	0.79 (0.05)	0.68 (0.05)	0.74 (0.06)	1.41 (0.14)	1.15 (0.15)	1.37 (0.17)
1-part	0.57 (0.04)	0.66 (0.07)	0.61 (0.05)	0.68 (0.07)	0.59 (0.04)	0.88 (0.10)
3-parts	0.92 (0.11)	0.98 (0.18)	0.79 (0.08)	1.62 (0.33)	1.35 (0.14)	1.33 (0.21)
5-parts	0.83 (0.07)	0.94 (0.13)	0.96 (0.13)	1.21 (0.15)	1.29 (0.14)	1.32 (0.22)
Pigeons						
0-part	1.68 (0.15)	2.68 (0.37)	1.65 (0.15)	3.51 (0.63)	4.12 (1.08)	3.76 (0.50)
1-part	1.98 (0.28)	2.14 (0.27)	1.87 (0.33)	2.42 (0.28)	2.18 (0.28)	3.40 (0.37)
3-parts	1.74 (0.21)	2.59 (0.81)	1.87 (0.23)	3.15 (0.55)	4.41 (1.04)	4.30 (0.81)
5-parts	2.08 (0.35)	2.12 (0.36)	1.95 (0.23)	2.85 (0.70)	2.97 (0.59)	6.69 (1.53)

^a Trained orientations.

three-, and five-part conditions), this approach predicts some decrement in recognition speed, even for the novel view that falls between the training orientations.

Other approaches, however, accord special status to novel views that fall between two trained orientations. For example, Bulthoff and Edelman (1992) have argued that recognition of novel views that fall between two trained orientations can be accomplished by an interpolation process. In support of this view, both humans (Bulthoff & Edelman, 1992) and monkeys (Logothetis et al., 1994) have been found to show much smaller error rates for views that fall between training orientations than for views outside the training range. Alternatively, it is possible that, during training with two orientations, participants may mentally rotate the object between the training views. Thus, participants may store not only the training views but, through a mental rotation process in training, also the views that fall between them. According to this notion, participants should show viewpoint invariance on tests in which the object is rotated between the trained orientations but viewpoint dependence on tests in which the object is rotated outside the trained orientations.

To assess these possibilities, we collapsed across the four object-part conditions and examined accuracy and latency for three rotations: 0°, 30°, and 45°. The 0° rotation is the average of the two trained views, the 30° rotation is the 30° test orientation (which was 30° from the 0° training view and fell between the training views), and the 45° rotation is the 135° test orientation (which was 45° from the 90° training view and fell outside the trained range). These data are shown in Figure 6 for both pigeons and humans. As can be seen, pigeons showed a systematic linear decrease in accuracy and an increase in latency as a function of degree of rotation from the nearest trained view. Humans, by contrast, showed complete viewpoint invariance on both measures for the rotation between the trained values but strong viewpoint dependence for the rotation outside the trained values.

For pigeons, an ANOVA on the accuracy scores showed a significant effect of rotation, $F(2, 14) = 16.22$. Subsequent pairwise comparisons (Newman-Keuls) revealed that accuracy was significantly higher at the 0° (trained) rotation than at either of the test rotations, which did not differ significantly from each other. The ANOVA on proportional latency also showed a significant effect of rotation, $F(2, 14) = 15.55$. Newman-Keuls comparisons revealed that latencies were significantly lower at the 0° (trained) rotation than at either of the test rotations. Latencies were also lower for the 30° test rotation than for the 45° rotation. Thus, pigeons clearly showed viewpoint dependence even for the novel rotation that fell between the trained values.

For humans, an ANOVA on the accuracy scores showed a significant effect of rotation, $F(2, 30) = 26.08$. Subsequent pairwise comparisons (Newman-Keuls) revealed no difference in accuracy between the 0° (trained) rotation and the 30° (between) rotation but higher accuracy for both of these rotations than for the 45° (outside) rotation. The ANOVA on proportional latency also showed a significant effect of rotation, $F(2, 30) = 36.89$. Again, Newman-Keuls comparisons revealed no difference in latency between the 0° (trained) rotation and the 30° (between) rotation but lower latencies for both of these rotations than for the 45° (outside) rotation. Thus, in contrast to pigeons, humans clearly showed *viewpoint invariance* for the rotation that fell between the training orientations.

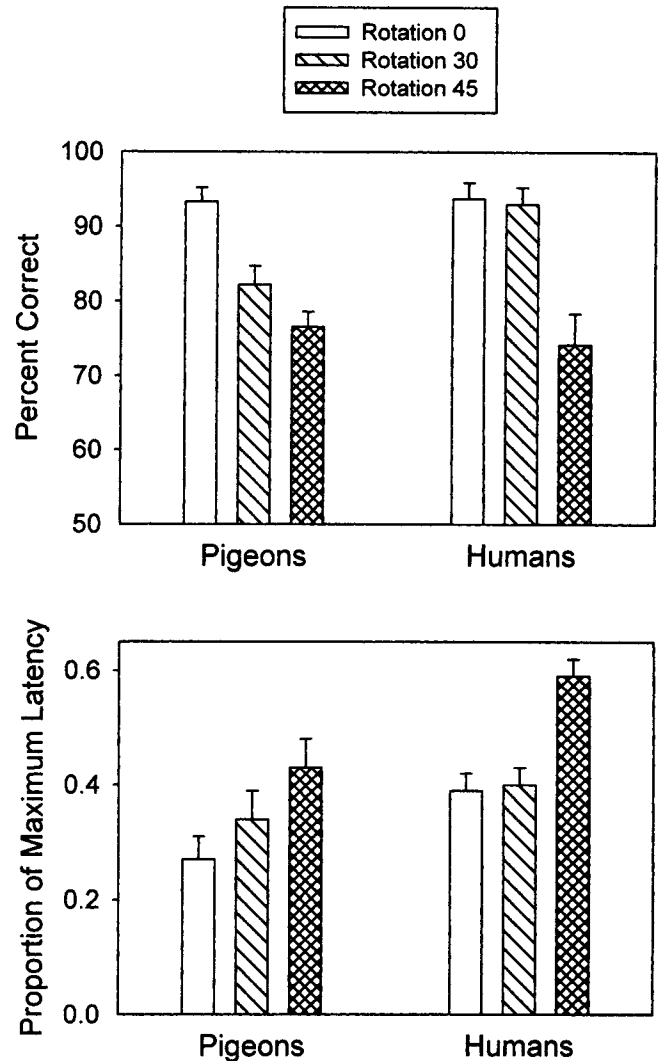


Figure 6. Choice accuracy (top panel) and proportion of maximum latency (bottom panel) in pigeons and humans on test trials with the trained orientations (Rotation 0), with the novel orientation that fell between the trained orientations (Rotation 30), and with the nearest novel orientation that fell outside the training range (Rotation 45). The data are averaged across the four object-part conditions in Experiment 1. Error bars show standard error of the mean.

Discussion

The results for humans were generally consistent with those found by Tarr et al. (1997) in that the degree of viewpoint dependence clearly differed across the four object types. Specifically, in the one-part condition, humans showed no viewpoint dependence in accuracy and only weak viewpoint dependence in latency. By contrast, in both the zero-part and three-part conditions, humans showed strong viewpoint dependence both in accuracy and latency. In the five-part condition, humans also showed clear viewpoint dependence in latency but somewhat weaker viewpoint dependence in accuracy.

Pigeons, on the other hand, showed strong viewpoint dependence on both measures in all four conditions. In general, pigeons' object recognition resembled typical generalization functions, with

latency increasing systematically, and accuracy decreasing systematically, as a function of the degree of rotation away from the nearest training view.

The functions for humans were subtly different from those of pigeons in a particularly interesting way. Specifically, the accuracy and latency functions for humans depended not just on the degree of rotation from the nearest training value but also on whether the novel orientation was inside or outside the training range. Humans showed rotational invariance for the novel orientation within the training range but strong viewpoint dependence for the nearest novel orientation outside the training range. Pigeons showed viewpoint dependence within and outside the training range.

Experiment 2

This experiment was conducted as a control experiment to ensure that the effects of rotation seen in Experiment 1 were not due to inherent differences in the difficulty of recognizing objects at the various orientations. Previous research has found that, even with artificial objects, some views are recognized more quickly and accurately than other views despite equivalent training (Edelman & Bulthoff, 1992). In Experiment 2, we wanted to ensure that the orientations we used for the novel rotations tests were not inherently more difficult to recognize than the orientations we used in training (i.e., that our rotation effects were not confounded by the canonical views phenomenon). Accordingly, participants in this experiment were trained from the outset with all six orientations within each object-part condition.

Method

Participants. Four adult pigeons with varied experimental histories but no prior experience with object discrimination tasks served as participants. The birds were housed and maintained as described in Experiment 1.

The human participants were 8 undergraduate students (3 men and 5 women) who received extra credit toward their introductory psychology class for participating in the experiment. The participants ranged in age from 18 to 27 years.

Apparatus and stimuli. The apparatus and stimuli were identical to those used in Experiment 1.

Procedure. The procedure used for pigeons was identical to that used in Experiment 1 except that all six orientations were presented equally throughout the experiment. Training in each condition continued until overall accuracy averaged across five consecutive sessions equalled or exceeded 80%. As in Experiment 1, we counterbalanced order of exposure to the four conditions. One bird originally assigned to start in the zero-part condition failed to learn the task (overall accuracy remained close to 50% even after 75 sessions) and was replaced with another pigeon with an equivalent experimental history.

The instructions given to humans were identical to those in Experiment 1 except that mention of no-feedback trials was omitted. The procedure was similar except that all orientations were trained equally within each part condition. For each part condition, each orientation was presented twice, once with the S+ on the left and once with it on the right, within each of three sets of 12 trials. Order of presentation within each 12-trial set varied randomly. Thus, each participant received a total of 6 trials with each orientation for each part condition. As in Experiment 1, order of exposure to conditions was counterbalanced across participants.

Results and Discussion

For human participants, the first block of trials was considered practice and was excluded from the data presentation and analysis.

Therefore, the data used were the last 24 training trials in each part condition (4 at each orientation). For pigeons, because acquisition was highly variable across birds and conditions, we presented and analyzed the data from the five sessions that met the accuracy criterion in each part condition. Although this accuracy criterion required that overall accuracy be at or above 80%, it did not require that accuracy or latency be evenly distributed across orientations. It is theoretically possible, for example, to meet the accuracy criterion through very high accuracy with some orientations and poor accuracy with others.

The left panels of Figure 7 show accuracy as a function of orientation for each object-part condition. A mixed-design ANOVA, with species as the between-subjects factor and with parts and orientation as within-subjects factors, revealed no significant main effects and no significant interactions (all p values $> .2$). Table 2 shows latency as a function of orientation for each object-part condition. As in Experiment 1, we scaled these data by converting them to a proportion of the maximum latency for each participant. These proportional latencies are plotted in the right panels of Figure 7. For both species, latencies tended to be highest for the five-part objects, but latency did not vary systematically as a function of orientation. A mixed-design ANOVA, with species as the between-subjects factor and with parts and orientation as within-subjects factors, revealed a significant main effect of species, $F(1, 10) = 5.40$, and of parts, $F(3, 30) = 3.08$. The Species \times Parts interaction approached but did not reach significance, $F(3, 30) = 2.36$. Most importantly, neither the main effect of orientation nor any of the interactions involving orientation approached significance (all p values $> .3$).

The lack of significant orientation differences in this experiment supports our conclusion that the strong effects of orientation seen in the first experiment were a function of the difference between the training and test orientations rather than a function of inherent differences in difficulty among the different views.

Experiment 3

One of the more interesting findings that emerged from Experiment 1 was that humans, but not pigeons, appeared to easily recognize the objects at the novel orientation that fell within the training range but not at orientations outside the training range. Pigeons, on the other hand appeared to show a straightforward generalization function in which the determining factor was distance from the nearest training orientation.

One could argue, however, that distance from the nearest training view could also be the single determining factor underlying the human observers' apparent viewpoint invariance for objects between the training views. Although humans showed no decrement in accuracy when the objects were rotated to a value within the training range and a huge decrement in accuracy when the objects were rotated outside the training range, the within versus between distinction was confounded by a difference in distance to the nearest training value (i.e., a distance of 30° for the within-range test orientation compared with a distance of 45° for the nearest outside-range test orientation). It is possible that the 30° rotation was simply below the threshold for producing view-dependent recognition deficits.

Experiment 3 was designed to replicate the results obtained in Experiment 1 with humans while eliminating the difference in distance from the nearest training orientation for the within-range

Trained with all Views

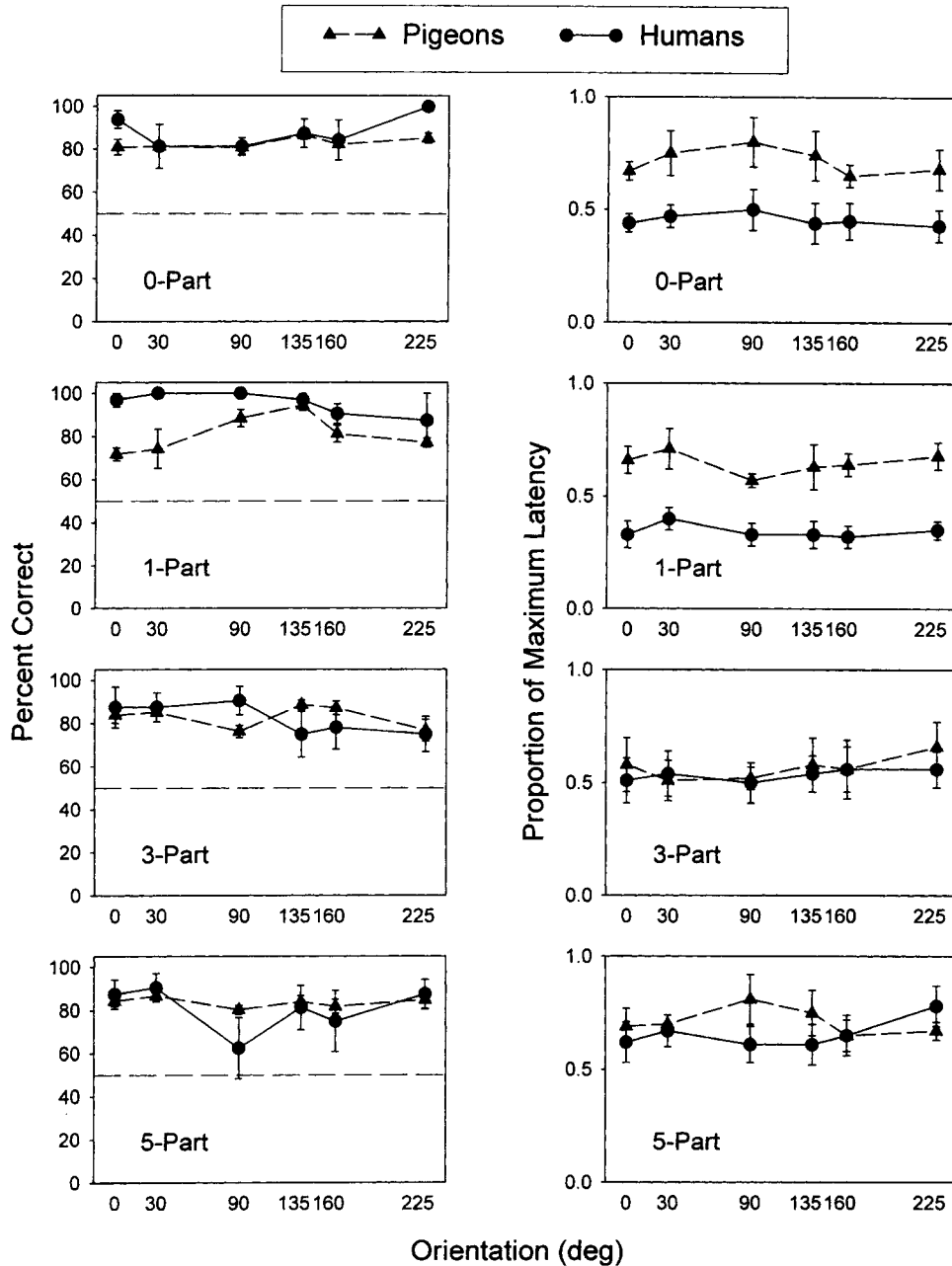


Figure 7. Choice accuracy (left panels) and proportion of maximum latency (right panels) in pigeons and humans as a function of orientation in the four object-part conditions in Experiment 2. Error bars show standard error of the mean (absent error bars indicate that error was smaller than the symbol size).

versus outside-range test orientations. For comparison purposes, we also tested two additional groups of humans, one trained only at the 0° orientation and one trained only at the 90° orientation. Because these conditions provide only a single training view, viewpoint invariance should not occur to any of the test views.

Method

Participants. The participants were 36 undergraduate students (9 men and 27 women) who received extra credit toward their introductory psy-

chology class for participating in the experiment. The participants ranged in age from 17 to 42 years.

Apparatus and stimuli. The apparatus was similar to that used in Experiments 1 and 2 except that a different monitor (NEC M700 [NEC U.S.A., Melville, NY]) with a larger viewing area (approximately 32 × 23 cm) was used. The stimuli were the same as in previous experiments except that the 30° test orientation was replaced with a 45° test orientation and the 160° orientation was not used.

Procedure. Participants were assigned to one of three groups (n = 12): Both, 0, or 90. The instructions given to humans were the same as those

Table 2
Mean (SEM) Choice Latency (in Seconds) as a Function of Orientation in Each Object-Part Condition for Humans and Pigeons in Experiment 2

Participants and condition	Orientation					
	0	30	90	135	160	225
Humans						
0-part	0.92 (0.19)	0.94 (0.17)	0.94 (0.18)	0.81 (0.11)	0.84 (0.11)	0.82 (0.16)
1-part	0.61 (0.13)	0.69 (0.16)	0.84 (0.22)	0.59 (0.10)	0.62 (0.13)	0.71 (0.17)
3-parts	0.95 (0.24)	0.94 (0.22)	1.05 (0.30)	1.01 (0.17)	1.02 (0.15)	1.16 (0.29)
5-parts	1.30 (0.37)	1.29 (0.35)	1.53 (0.43)	1.26 (0.32)	1.35 (0.32)	1.79 (0.52)
Pigeons						
0-part	1.88 (0.39)	2.16 (0.54)	2.20 (0.40)	2.08 (0.46)	1.89 (0.50)	1.87 (0.36)
1-part	1.76 (0.17)	1.85 (0.10)	1.56 (0.24)	1.65 (0.19)	1.70 (0.17)	1.81 (0.18)
3-parts	1.45 (0.22)	1.30 (0.11)	1.38 (0.14)	1.50 (0.29)	1.49 (0.38)	1.73 (0.18)
5-parts	2.03 (0.57)	1.96 (0.38)	2.34 (0.65)	2.20 (0.69)	1.91 (0.52)	1.90 (0.39)

described for Experiment 1. The procedure was the same as that used in Experiment 1 except for the orientations presented during training and testing.

Group Both was trained with both the 0° and the 90° orientations, as in Experiment 1, whereas Group 0 was trained with only the 0° orientation and Group 90 was trained with only the 90° orientation. For all groups, the orientations used during testing were 0°, 45°, 90°, 135°, and 225°. As in Experiment 1, each participant was trained and then tested in each of the four geon-part conditions (zero-, one-, three-, and five-parts), with order of exposure to conditions counterbalanced across participants. Each condition began with 16 training trials with the training orientation(s) followed by a test block that consisted of a mixture of 12 training trials and 10 test trials (2 test trials at each of the five test orientations).

Results and Discussion

The left columns of Figures 8, 9, and 10 show accuracy as a function of orientation in each object-part condition for each group. Latencies are shown in the right-hand column of each figure. These data were analyzed with two-way ANOVAs for each group and each measure.

For Group Both, accuracy scores revealed significant main effects of parts, $F(3, 33) = 4.92$, and orientation, $F(4, 44) = 23.23$, and a significant interaction, $F(12, 132) = 2.22$. Similarly, latency scores showed significant effects of parts, $F(3, 33) = 3.42$, and orientation, $F(4, 44) = 11.80$, and a significant interaction, $F(12, 132) = 2.08$. For Group 0, accuracy scores showed significant main effects of parts, $F(3, 33) = 4.52$, and orientation, $F(4, 44) = 23.09$, but the interaction was not significant. Latency scores showed a significant effect only of orientation $F(4, 44) = 9.32$. For Group 90, accuracy scores showed a significant effect of orientation, $F(4, 44) = 26.36$, and a significant Parts \times Orientation interaction, $F(12, 132) = 3.41$, but the main effect of parts was not significant. Latency scores showed a significant effect only of orientation, $F(2, 44) = 15.06$. Thus, consistent with results reported by Tarr et al. (1997), all groups showed a strong viewpoint dependence, and this viewpoint dependence was typically weakest in the one-part condition. This is visualized most clearly in Figure 11, in which we have replotted the data in terms of degree of rotation from the nearest training view(s), using only the rotations common to all groups and averaging data points when necessary.

The main focus of this experiment, however, concerned the distinction between the novel rotation within the training range and the novel rotation outside the training range. To make this com-

parison, we collapsed across part conditions and examined a subset of the orientations (0°, 45°, 90°, and 135°) for each group (see Figure 12). For all groups, one-way ANOVAs returned a significant effect of orientation for both measures. Subsequent post hoc comparisons (Newman-Keuls) were used to further assess the pattern of differences. For Group Both, accuracy was lower, and latencies were longer, at the 135° orientation than at any of the other orientations, which did not differ from each other on either measure. Thus, this result replicates the previous finding of viewpoint invariance within the training range (i.e., recognition at the novel 45° orientation was comparable to that at the 0° and 90° training orientations) and viewpoint dependence outside the training range (135°).

As expected, the two groups trained with only a single orientation showed a different pattern than the group trained with both orientations. Group 0 showed higher accuracy and shorter latencies at the training (0°) orientation than at any of the novel orientations (45°, 90°, and 135°). Accuracy at the 45° orientation was higher than at either 90° or 135°, but latency did not differ across the three novel orientations. Group 90 also showed significantly higher accuracy and shorter latency at the training orientation (90°) than at any of the three novel orientations (0°, 45°, or 135°). Accuracy was higher at the 45° than at the 135° orientation. Latencies were longer at the 0° orientation than at either the 45° or 135° orientations.

These results clearly support our previous finding with humans of rotational invariance within a training range but not outside that range. A rotation of 45° from the nearest training view produced a significant decrease in accuracy and an increase in latency in all cases except when the rotation fell between two trained views (i.e., the 45° orientation for Group Both).

General Discussion

The present results can be summarized as follows. Following training at two orientations, pigeons showed viewpoint-dependent object discrimination both in accuracy and choice latency, regardless of whether the objects were composed of zero, one, three, or five distinct parts. Moreover, their recognition varied as a simple linear function of degree of rotation from the nearest training view, and they showed viewpoint dependence even when the objects were rotated within the training range. Humans also showed view-

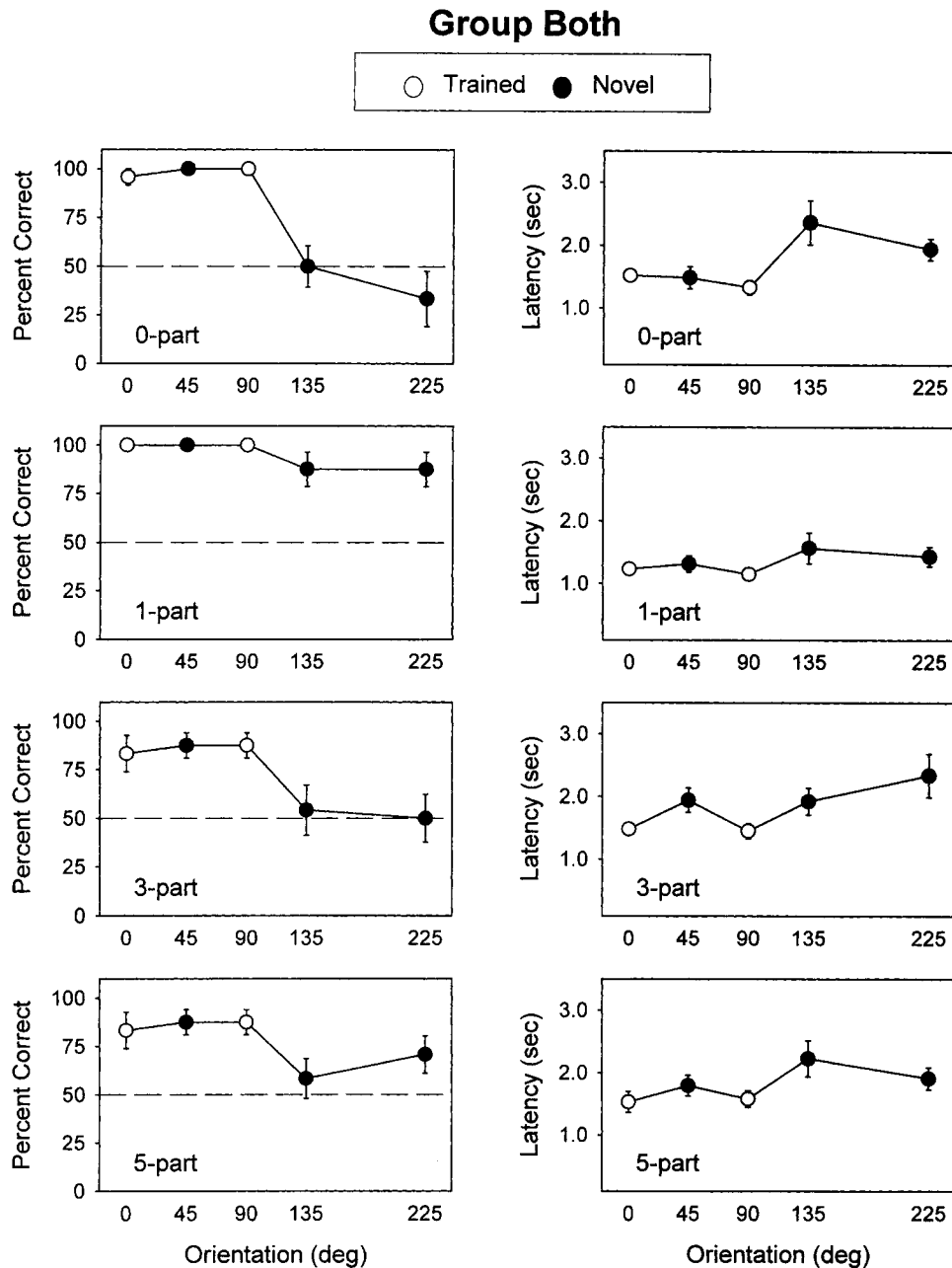


Figure 8. Choice accuracy (left panels) and proportion of maximum latency (right panels) for humans in Group Both as a function of orientation in the four object-part conditions in Experiment 3. Error bars show standard error of the mean (absent error bars indicate that error was smaller than the symbol size).

point dependence on both measures, but the degree of viewpoint dependence varied with the composition of the stimuli. Specifically, viewpoint dependence was weaker in the one-part condition than in the zero-part or multipart conditions. Moreover, in contrast to the pigeons, humans trained with two orientations showed viewpoint invariance for the novel rotation that was between the trained orientations.

These results are interesting from several perspectives. First, from the perspective of avian visual cognition, the present results advance knowledge of object recognition processes in pigeons.

Specifically, this is the first systematic investigation of the effect of number of distinct object parts on recognition across novel viewpoints. The finding that pigeons showed viewpoint dependence in all object-part conditions suggests that viewpoint dependence is a general feature of pigeons' object recognition processes, a conclusion that is consistent with most previous studies in the literature.

The finding that pigeons showed viewpoint dependence even when the target object contained a distinctive geon that could serve as a diagnostic feature is also consistent with the results of another

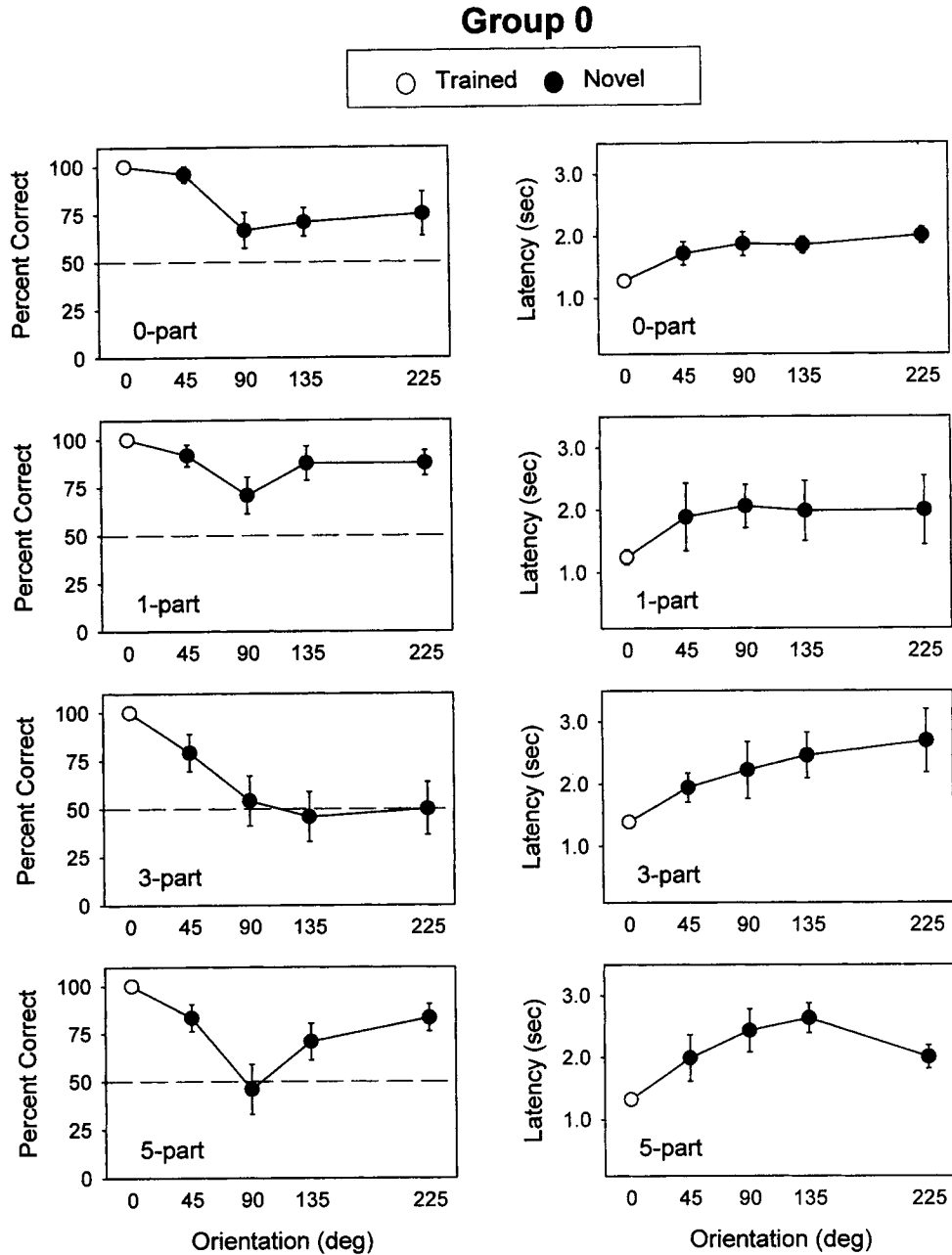


Figure 9. Choice accuracy (left panels) and proportion of maximum latency (right panels) for humans in Group 0 as a function of orientation in the four object-part conditions in Experiment 3. Error bars show standard error of the mean (absent error bars indicate that error was smaller than the symbol size).

recent study from our laboratory (Spetch, Kelly, & Reid, 1999). Two groups of pigeons were trained to discriminate one target object from three distractor objects. For one group, the target object and the distractor objects were composed of differently shaped parts, whereas for the other group, the target and distractors were composed of the same parts but were configured differently, so as to form globally different shapes. Both groups were trained at six orientations and then tested at several novel orientations. Although the distinct-parts group showed higher overall accuracy levels, indicating that they were able to benefit from the distinctiveness of the parts in their object discriminations, they showed

decrements in accuracy when tested with novel rotations that were equivalent to those shown by the same-parts group. Thus, the presence of the distinctive parts facilitated pigeons' object discrimination but did not facilitate their recognition of the objects at novel rotations.

Although pigeons' object recognition showed strong viewpoint dependence, it should be noted that their recognition accuracy remained well above chance levels at the nearest rotations both in the present study and in Spetch et al. (1999). Moreover, the orderly nature of the functions seen in the present study, with accuracy decreasing and latency increasing as a linear function of degree of

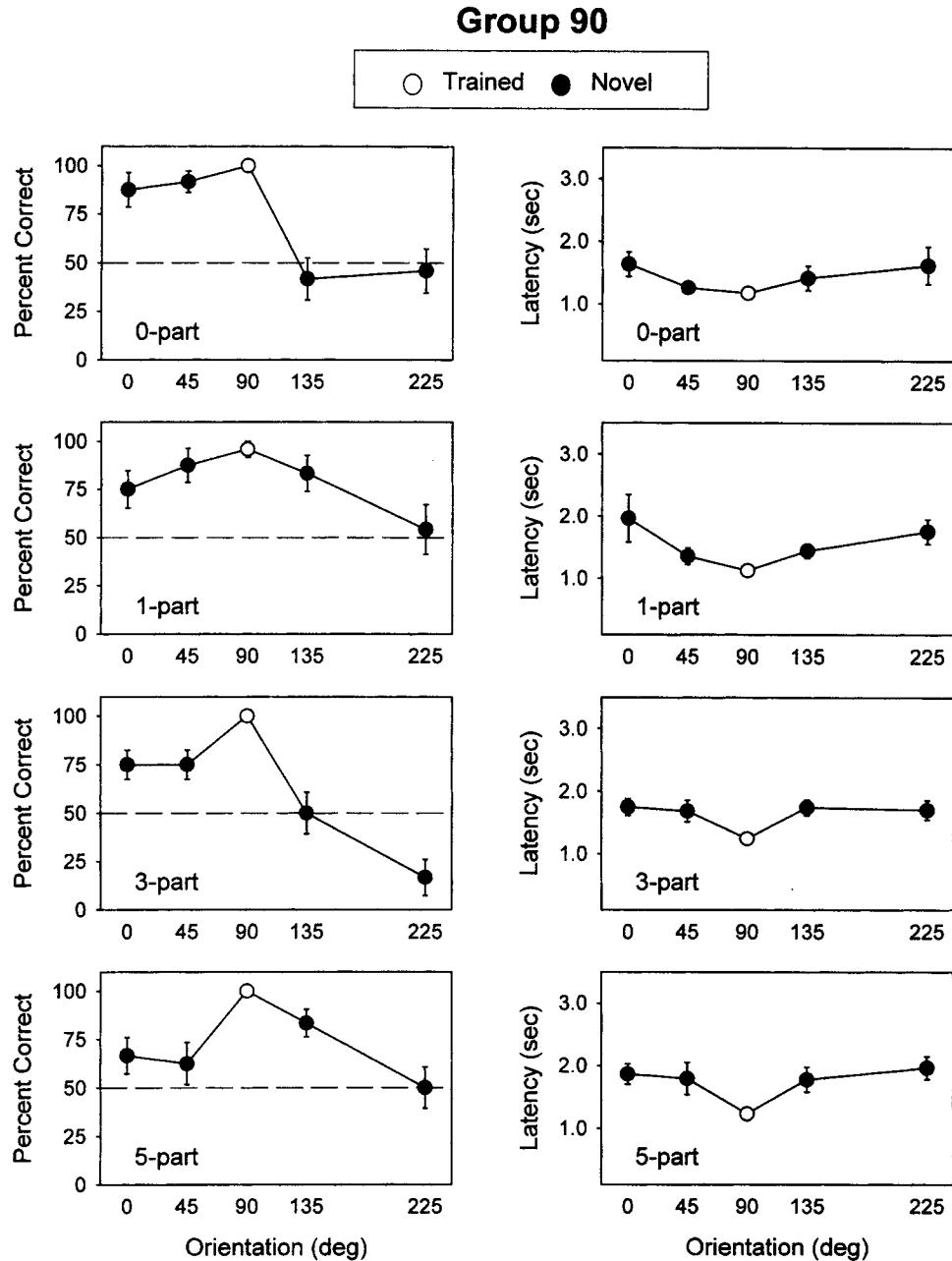


Figure 10. Choice accuracy (left panels) and proportion of maximum latency (right panels) for humans in Group 90 as a function of orientation in the four object-part conditions in Experiment 3. Error bars show standard error of the mean (absent error bars indicate that error was smaller than the symbol size).

rotation from the nearest training view, indicates that to some extent, the pigeons recognized the similarity of the rotated objects to those seen during training. However, this detection of similarity was not necessarily based on a recognition of the identity of the objects and instead could reflect similarity of local features or abstract patterns.

From the perspective of the human object perception literature, the results for humans are interesting in two ways. First, the results are quite consistent with those obtained by Tarr et al. (1997) despite procedural differences, specifically, our use of a discrim-

ination task rather than a same/different or naming task and the fact that each of our participants was exposed to all object-part conditions. Consistent with Tarr et al.'s results, humans showed strong viewpoint dependence in both the zero-part and the three-part condition and weaker or absent viewpoint dependence (manifest primarily in latencies changes) in the one-part condition. Our results for the five-part objects were somewhat inconsistent across the two experiments and the two measures but, for the most part, showed viewpoint dependence, as found by Tarr et al. A difference in degree of viewpoint dependence between the zero-part and

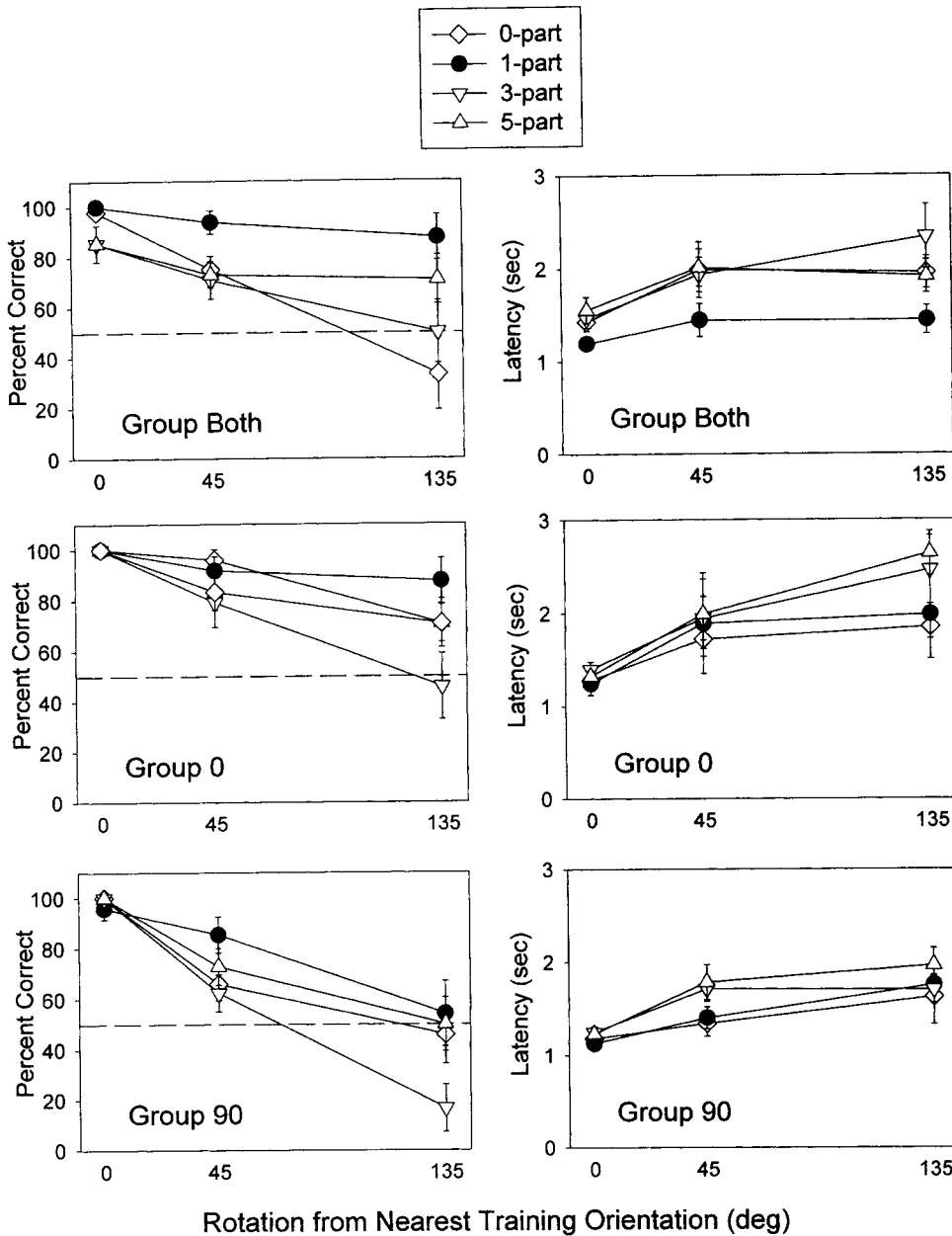


Figure 11. Choice accuracy (left panels) and proportion of maximum latency (right panels) for the three groups of humans in each object-part condition as a function of rotation from the nearest training orientation in Experiment 3. For Group Both, the 0° rotation is the average of the training views, the 45° rotation is the average of the 45° and 135° orientations, and the 135° rotation is the 225° orientation. For Group 0, the rotations correspond directly to the orientations. For Group 90, the 0° rotation represents the 90° training view, the 45° rotation is the average of the 0° and 135° orientations, and the 135° rotation is the 225° orientation. Error bars show standard error of the mean (absent error bars indicate that error was smaller than the symbol size).

one-part conditions is consistent with both view-based and geon theories. However, the strong viewpoint dependence seen in the three-part condition favors view-based theories.

Second, the finding of complete viewpoint invariance in both accuracy and latency for the novel rotation that fell between the two training orientations is interesting because it supports the notion that, when trained consistently on two views of an object, humans may acquire or be able to rapidly generate representation

of views between the trained views. This result is consistent with previous findings and with the view interpolation model of Bulthoff and Edelman (1992).

Finally, from a comparative perspective, the results are interesting because they suggest both similarities and differences between species in object recognition processes. For both species, object recognition was generally viewpoint dependent, with either accuracy decreasing, latency increasing, or both as a function of degree

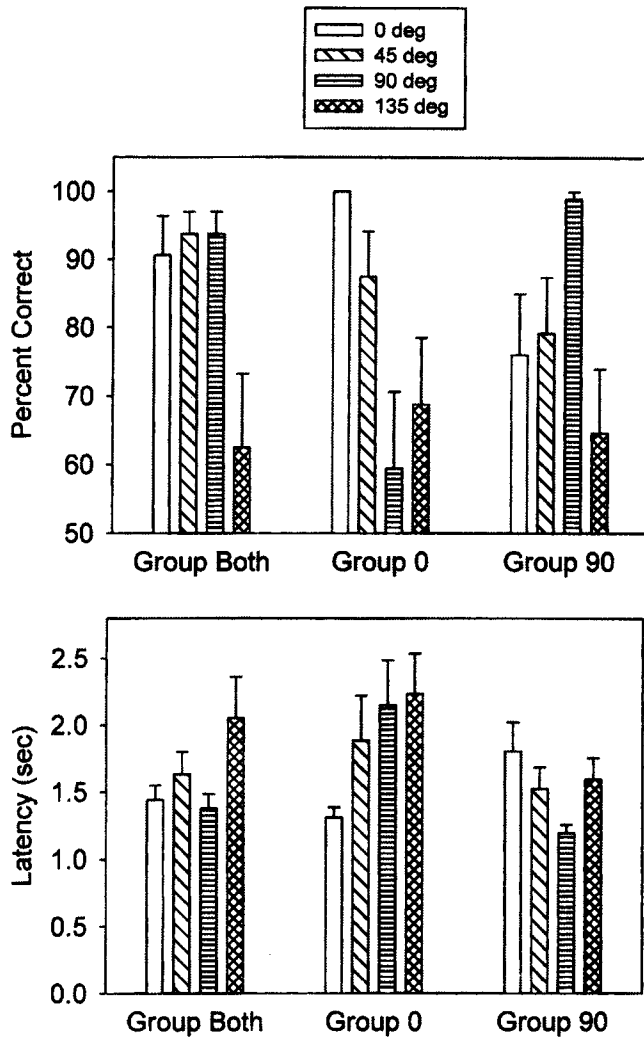


Figure 12. Choice accuracy (top panel) and proportion of maximum latency (bottom panel) in the three groups of humans on test trials at four orientations, collapsed across the four object-part conditions in Experiment 3. Error bars show standard error of the mean.

of rotation from the nearest training view. However, the presence of a single distinctive geon that differed between the positive and negative objects (i.e., the one-part condition) facilitated object recognition more in humans than in pigeons. Thus, humans seemed able to use the distinctive feature to accurately recognize the objects across variation in viewpoint, whereas pigeons showed clear viewpoint dependency in their object recognition even in the one-part condition.

The difference between pigeons and humans in their response to the novel rotation that fell between the training orientations is particularly interesting and may suggest a difference in normalization processes. The invariance shown by humans to the rotation between the training values, coupled with the large decrement in recognition to rotations outside the training range, suggests that during training, humans either formed a representation of or could rapidly and accurately infer how the objects looked in all views between the trained ones. They appeared to show "categorical" object perception (see Livingston, Andrews, & Harnad, 1998), in

which the object was perceived as the same object within the rotational boundary formed by the training views. By contrast, the results shown by pigeons suggest that they learned how the objects looked at the two training views but not at the views in between. For pigeons, recognition of the novel views appeared to reflect generalization from the nearest training view.

The "graded" rather than categorical perception observed in pigeons, compared with humans, suggests that object recognition in pigeons may be based entirely on generalized resemblance detection and that pigeons may lack a special purpose mechanism that allows for local invariance in object recognition. This difference between pigeons and humans is particularly interesting because Logothetis et al. (1994) found that monkeys and humans showed the same pattern of results on tests with novel views between and outside the training range. They suggested that object recognition in both species may be accomplished by a nonlinear interpolation between stored representations. If their interpretation is correct, an interesting possibility to explore in future research is that such interpolation processes may be absent or reduced in avian species. Additional research using a variety of tasks and other avian species is needed to explore this exciting possibility.

In summary, the present results identify some interesting similarities and differences between pigeons and humans in object recognition processes. We suggest that cross-species comparisons may be a fruitful complement to the large literature on object recognition in people. Object recognition is not a uniquely human problem but is also fundamental to a wide range of adaptive behaviors in animals. Cross-species comparisons may reveal which processes are general across diverse object recognition systems and which processes reflect mechanisms that are specialized through phylogeny or experience.

References

- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115-147.
- Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 1162-1182.
- Biederman, I., & Gerhardstein, P. C. (1995). Viewpoint-dependent mechanisms in visual object recognition: Reply to Tarr and Bulthoff (1995). *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1506-1514.
- Bulthoff, H. H., & Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences*, *89*, 60-64.
- Cerella, J. (1977). Absence of perspective processing in the pigeon. *Pattern Recognition*, *9*, 65-68.
- Cerella, J. (1990). Pigeon pattern perception: Limits on perspective invariance. *Perception*, *19*, 141-159.
- Cook, R. G., & Katz, J. S. (1999). Dynamic object perception by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *25*, 194-210.
- Cooper, L. A., & Shepard, R. N. (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual information processing* (pp. 75-176). New York: Academic Press.
- Delius, J. D., & Hollard, V. D. (1995). Orientation invariant pattern recognition by pigeons (*Columba livia*) and humans (*Homo sapiens*). *Journal of Comparative Psychology*, *109*, 278-290.
- Edelman, S., & Bulthoff, H. H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision Research*, *32*, 2385-2400.

- Hayward, W. G., & Tarr, M. J. (1997). Testing conditions for viewpoint invariance in object recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 1511–1521.
- Herrnstein, R. J. (1979). Acquisition, generalization, and discrimination reversal of a natural concept. *Journal of Experimental Psychology: Animal Behavior Processes*, *5*, 116–129.
- Hollard, V. D., & Delius, J. D. (1982, November 19). Rotational invariance in visual pattern recognition by pigeons and humans. *Science*, *218*, 804–806.
- Humphrey, G. K., & Khan, S. C. (1992). Recognizing novel views of three-dimensional objects. *Canadian Journal of Psychology*, *46*, 170–190.
- Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory and Cognition*, *13*, 289–303.
- Jolicoeur, P. (1988). Mental rotation and the identification of disoriented objects. *Canadian Journal of Psychology*, *42*, 461–478.
- Jolicoeur, P., & Humphrey, G. K. (1998). Perception of rotated two-dimensional and three-dimensional objects and visual shapes. In V. Walsh & J. Kulikowski (Eds.), *Perceptual constancy: Why things look as they do* (pp. 69–123). Cambridge, England: Cambridge University Press.
- Livingston, K. R., Andrews, J. K., & Harnad, S. (1998). Categorical perception effects induced by category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 732–753.
- Logothetis, N. K., Pauls, J., Bulthoff, H. H., & Poggio, T. (1994). View-dependent object recognition by monkeys. *Current Biology*, *4*, 401–414.
- Lumsden, E. A. (1970). Implication of the equivalence of mirror-image stimuli for object constancy. *Psychonomic Science*, *19*, 55–56.
- Lumsden, E. A. (1977). Generalization of an operant response to photographs and drawings/silhouettes of a three-dimensional object at various orientations. *Bulletin of the Psychonomic Society*, *10*, 405–407.
- Lumsden, E. A., & Pullen, M. R. (1970). Object orientation as a dimension of stimulus generalization. *Psychonomic Science*, *18*, 149–150.
- Morgan, M. J., Fitch, M. D., Holman, J. G., & Lea, S. E. G. (1976). Pigeons learn the concept of 'A.' *Perception*, *5*, 57–66.
- Perrett, D. I., Oram, M. W., & Ashbridge, E. (1998). Evidence accumulation in cell populations responsive to faces: An account of generalisation of recognition without mental transformations. *Cognition*, *67*, 111–145.
- Reynolds, G. S. (1961). Contrast, generalization, and the process of discrimination. *Journal of the Experimental Analysis of Behavior*, *4*, 289–294.
- Shepard, R. N., & Metzler, J. (1971, February 19). Mental rotation of three-dimensional objects. *Science*, *171*, 701–703.
- Spetch, M. L., Kelly, D. M., & Reid, S. (1999). Recognition of objects and spatial relations in pictures across changes in viewpoint. *Current Psychology of Cognition*, *19*, 729–764.
- Tarr, M. J. (1995). Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin and Review*, *2*, 55–82.
- Tarr, M. J., & Bulthoff, H. H. (1995). Is human object recognition better described by geon structural descriptions or by multiple views? Comment on Biederman and Gerhardstein (1993). *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1494–1505.
- Tarr, M. J., Bulthoff, H. H., Zabinski, M., & Blanz, V. (1997). To what extent do unique parts influence recognition across changes in viewpoint? *Psychological Science*, *8*, 282–289.
- Tarr, M. J., & Gauthier, I. (1998). Do viewpoint-dependent mechanisms generalize across members of a class? *Cognition*, *67*, 73–110.
- Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Science*, *21*, 233–282.
- Towe, A. L. (1954). A study of figural equivalence in the pigeon. *Journal of Comparative and Physiological Psychology*, *47*, 283–287.
- Vaughan, W., Jr., & Greene, S. L. (1984). Pigeon visual memory capacity. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 256–271.
- Vetter, G. H., & Hearst, E. (1968). Generalization and discrimination of shape orientation in the pigeon. *Journal of the Experimental Analysis of Behavior*, *11*, 753–765.
- 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.

Received September 8, 1998

Revision received January 5, 1999

Accepted February 22, 2000 ■