An automated apparatus for presenting depth-rotated three-dimensional objects in human and animal object recognition research

ALINDA FRIEDMAN, MARCIA L. SPETCH, and ISAAC LANK
University of Alberta, Edmonton, Alberta, Canada

For practical reasons, research on the recognition of objects from different viewpoints has relied almost exclusively on the use of two-dimensional representations of three-dimensional objects. We describe an apparatus that enables the presentation of three-dimensional objects in a discrimination learning paradigm. Three chambers positioned on a movable table allow each of two objects to be presented on either the left or the right side; a viewing window exposes only two of the objects at a time. The objects can be arbitrarily designated as either an S+ or an S−. In addition, they can be placed precisely in any arbitrary start position and rotated in depth in 100 steps of 3.6° each. We have successfully used this apparatus to investigate recognition of depth-rotated objects by both pigeons and humans. By varying the stimuli, number of stimulus chambers, and software programs, the apparatus can be used for other types of tasks and to investigate other types of processes.

The ability to recognize objects in the environment quickly despite changes in viewpoint is one of the most fundamental and impressive tasks that the visual system must accomplish. Object constancy—the interpretation of varying two-dimensional retinal information as representing a constant external stimulus—is critical for the survival of humans and animals alike. Yet much remains to be understood about the visual and cognitive mechanisms that allow organisms to accomplish this task. Not surprisingly, a great deal of research and theory has been devoted to understanding these processes in humans and, more recently, in animals.

For practical reasons, research on the recognition of rotated objects has relied almost exclusively on the use of two-dimensional representations of three-dimensional (3-D) objects, in the form of drawings, photographs, or digitized images. Such pictorial displays allow rapid and automated presentations of the objects from exact preselected viewpoints. However, pictorial representations of objects—even photographs—do not provide all of the visual information that is available and that may influence object recognition in the real world. Thus, theories of object recognition based exclusively on research using pictorial stimuli may not provide a complete description of object recognition processes. Furthermore, recognizing objects from pictures may involve interpretive or other processes in addition to those used in real-world object recognition, and such processes may vary with evolution, development, or experience (e.g., Deregowski, 1989, 2000; Fagot, 2000; Watanabe, 2000). Therefore, using pictorial stimuli to investigate object recognition may be tapping into processes that underlie the interpretation of pictures, instead of or in addition to processes that underlie object recognition. This may be an especially important consideration in comparative research on differences between species or across development. For example, when differences between humans and animals in the recognition of depth-rotated objects are revealed (e.g., Spetch & Friedman, 2003; Spetch, Friedman, & Reid, 2001), it is important to determine whether these differences reflect mechanisms of picture perception, object recognition, or both.

We designed an apparatus and a method for automated presentations of precise depth rotations of real 3-D objects. The apparatus can be used in research with humans and can be attached to a custom operant chamber for research with animals. Our apparatus complements the automated apparatus designed by Delius (1992) but serves a different function. The apparatus used by Delius was designed to present multiple real objects to pigeons for studies of categorization, whereas our apparatus allows depth rotations of individual objects and is designed for studies of viewpoint effects in object recognition.

Procedure, Design, and Programming Considerations

Although our object rotation device can be used with various experimental procedures, we designed it for use

with a discrimination task in which subjects chose between two simultaneously displayed objects. For each subject, one object was arbitrarily designated at the positive/ reinforced stimulus (S+), and the other was designated as the negative/unreinforced object (S−). Humans chose between the objects by using a keyboard or a response box, whereas pigeons chose by pecking at a Plexiglas panel in front of the objects and their pecks were recorded with a touch frame (Carroll Touch 1490 smart frame). During training, humans received auditory feedback about the accuracy of their choices, whereas pigeons received food for correct choices and no food for incorrect choices.

One design issue with a simultaneous discrimination task concerns the ability to randomly alternate the left–right location of the S+ and S− objects. Our apparatus allowed us to do this by having three chambers to contain the objects, located on a sliding table (see Figure 1). The viewing screen allowed the subjects to see only two chambers at a time (see Figure 2). Two identical objects were placed in the two outside chambers, and a different object was placed in the center chamber. Either object could be designated as S+ or S−, which is another important design consideration, because idiosyncrasies of shape may make one object more or less difficult to recognize than the other at particular positions in depth. By sliding the table back and forth across trials, we could present the left and center object or the right and center object, thus allowing us to alternate the S+ object randomly between the left and the right halves of the viewing screen.

A critical consideration in being able to randomize the side of S+ presentation was the need to prevent the subjects from detecting, via visual or auditory cues, movements of the sliding table that could be used to signal the correct response. If, for example, the table moved only when the S+ location shifted from one trial to the next and if the subjects could detect this movement, they might be able to anticipate which side would contain the S+ object. More generally, all of the issues to be considered below concern eliminating the ways that subjects might learn the discrimination on the basis of unintended cues. Although the "Clever Hans" effect is normally attributed to unintended cuing from a live experimenter, it is certainly possible for unintended cuing to occur in an automated apparatus, and it is thus important to control for or eliminate the sources of such cuing.

To eliminate visual detection of the sliding table movement, we tested the subjects in a completely darkened room. The lights within the object chambers were automatically turned off during all object movement; timing for each trial coincided with turning on the lights within the chambers. The time spent in darkness between trials was insufficient to allow dark adaptation (e.g., <10 sec).

To prevent the use of auditory cues to anticipate table movement, the table was programmed to move on all trials, whether or not the S+ was to stay on the same side or shift sides. For trials in which the S+ was programmed to stay on the same side as in the previous trial, the table moved halfway, paused, and then moved back to the orig-
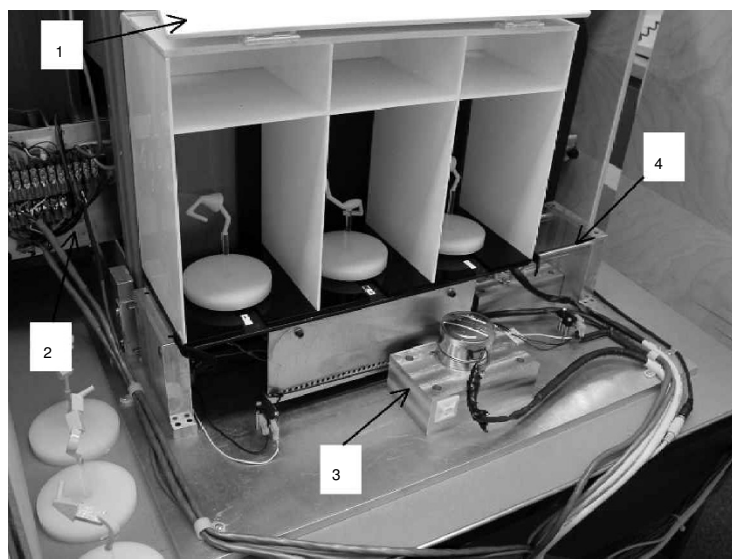

Figure 1. Back of the apparatus, showing the movable table and object chambers. (1) Partial view of the opaque Plexiglas door that covers the back of the three chambers once the objects are in place. (2) Connection terminal for all controllable pieces of the apparatus—for example, feeder hoppers, lights, touchscreen, moving table, sliding cover, and rotating bases. (3) Hurst electric motor to control table movement. (4) Precision-ground 0.5-in. metal rods with four linear bearings. These fit the underside of the table, which slides back and forth on them.
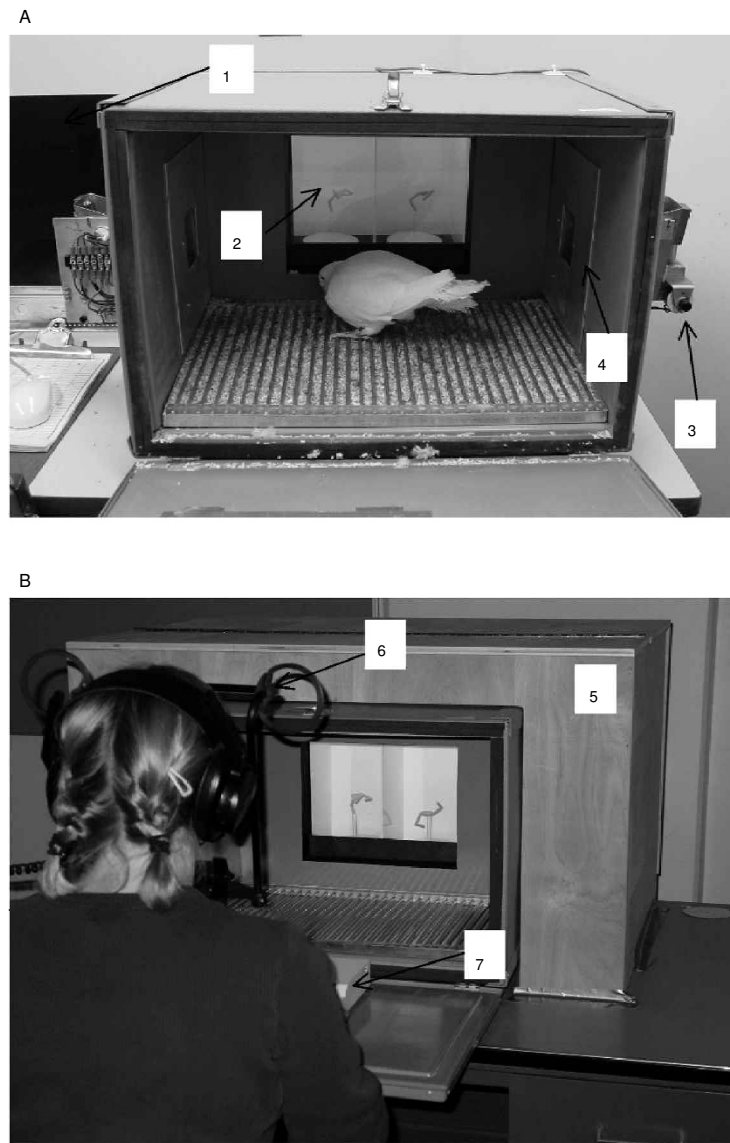
Figure 2. Panel A shows the front of apparatus, showing the view of objects provided to pigeon subjects. Panel B shows the apparatus as it appears to human subjects. (1) Partial view of the sliding door when it is in the open position. (2) Computer touch screen in front of the two displayed objects. (3) One of the two food hoppers. (4) One of two hopper openings with infrared beam. (5) Hinged wooden cover that hides the apparatus from the human subjects. (6) Chinrest; the circular rings attached to the chinrest pivot to touch the subject's temples and keep the head in place. (7) Response box resting on the door to the pigeon box.

inal side. For trials in which the S+ was programmed to move to the other side from that in the previous trial, the table moved halfway, paused, and then continued to move. The pause was included on shift trials to mimic the time and sound of the motor pausing to change directions on stay trials. Pilot observations with humans instructed to try to respond on the basis of auditory cues indicated that there was no detectable difference in the sound of the table's movement as a function of the direction of the movement (i.e., either continuing in the same direction or shifting after the pause). In addition, out of more than 60 subjects tested, none indicated on a subsequent strategy questionnaire that he or she had used auditory cues as a basis for responding. However, if this issue remains a concern, providing a masking noise over the headphones (for the humans) or in the operant chamber (for the pigeons)

should be used to eliminate the noise of the table's movement from being used as a cue.

A second design issue concerned the means by which we could rotate the objects between trials automatically without the subjects determining whether or by how much they had been rotated. We designed the apparatus (described below) so as to allow each object to be rotated rapidly and automatically in 3.6° steps, which allowed us to select from 100 different depth rotations. Turning out the lights within the object chambers prevented the subjects from seeing the objects while they were rotating. In addition, to preclude the use of auditory cues to signal larger or smaller rotations, the objects were rotated on every trial, first by the shortest rotation required (i.e., clockwise or counterclockwise) to move to the orientation selected for the current trial and then by an additional 360°. The latter movement prevented the subjects from discriminating how much the objects had been rotated on any given trial, as well as preventing them from discerning when no rotation occurred between trials.

A third set of design issues concerned preventing the subjects from responding on the basis of (1) visual differences among the three object chambers themselves, (2) small differences between the two versions of the same object, or similarly, (3) small imperfections or idiosyncrasies on a single object (e.g., a scratch), particularly when it was in the middle chamber and, thus, unique. Each of these potential confounds needs to be dealt with by using counterbalancing in the main experiment and/or by running control experiments.

With respect to the first issue, for instance, it is virtually impossible to create three identical chambers. However, it is possible to determine whether a subject responded on the basis of cues distinguishing the chambers by conducting a control experiment after the discrimination has been learned, in which the same object is placed in all three chambers.

With respect to the second and third issues, even with the automated means that we used to create the stimulus objects (see below), there were likely to have been at least some minor differences between the two "identical" objects that were placed in the outer chambers. But because the objects in the outer chambers were assigned to be the S+ for some of the subjects, it is unlikely that the discrimination they learned could be based on idiosyncrasies belonging to a particular object (e.g., a scratch) and more likely that their responses were based on the object's shape. Even so, whichever object was in the middle chamber might have had uniquely identifying marks that could have been used by some subjects to cue its identity. Thus, to eliminate the possibility that subjects are responding to unintended cues on specific exemplars, one should create two or three exemplars of each of the two shapes to be discriminated and counterbalance not only how they are assigned to be the S+ or the S−, but also which of the exemplars appear in the center and outer chambers, on a block-to-block, session-to-session, or subject-to-subject basis.

Stimulus Creation

The 3-D models for the stimuli were created with the Rhino 1.1 computer-assisted design (CAD) program (Robert McNeel & Associates, 2001). They were based on the one- and three-geon stimuli used by Spetch et al. (2001). There were two one-geon stimuli and two three-geon stimuli; for each pair, one of the objects was arbitrarily designated *A*, and the other was designated *B* (see Figure 3). For each A and B pair, we produced one example of the A object and two examples of the B object. During discrimination trials, the A object was placed in the middle chamber of the apparatus, and the B objects were placed in the chambers to the left and the right of it (see Figure 1). As was previously noted, either the A or the B object could be designated as S+ or S−. Also, as was previously noted (although we did not do this), several exemplars of each A and B shape should be produced and counterbalanced in terms of their assignment to the middle and outside chambers and, when assigned to the middle chamber, in terms of which exemplar appeared during a given session or block.

The 3-D coordinates generated by the CAD software were input to a Stratasys Genisys Xs 3-D printer, which makes 3-D objects to scale by first determining an optimal pose for the object (in terms of what surface will form the most stable base) and then laying down threads of polyester in solid layers, from the base up. Because the polyester threads were visible on the surfaces of the resulting objects, we painted the objects with Krylon Living Color 46410 Canary enamel paint, to render a smooth surface. The objects were then glued onto an acrylic rod, and the rod was glued onto a 94-mm-diameter circular base.

The undersides of the circular bases used to support the objects were machined 8 mm into pentagonal shapes that matched metal shapes that were set into the three turntables of the apparatus (see Figure 4). In addition, a vertical line (invisible to the subjects) was drawn on the underside of each base, and another line was drawn on the Plexiglas panels that the bases rested on. These lines ensured that the bases were exactly aligned relative to each other. These measures also permitted the stimulus objects to be mounted in the same positions and aligned precisely each time they were placed in the apparatus. In addition, both the undersides of the object bases and the tops of the turntables in the chambers had Velcro pieces to prevent the shapes from moving.

Apparatus Overview

We will first describe the apparatus as it was configured for use with pigeons, and then we will describe the modifications required for use with humans.

The apparatus had a large box (Grason-Stadler Animal Chest Model E3125AA, 54 × 36 × 36 cm) designed to contain a pigeon subject (see Figure 2, panel A). Behind the box was a small movable table with three objects on it (Figure 1). Between the movable table and the box was a sliding door cover and a computer touch screen mounted on a piece of Plexiglas (Figure 2). When the sliding door
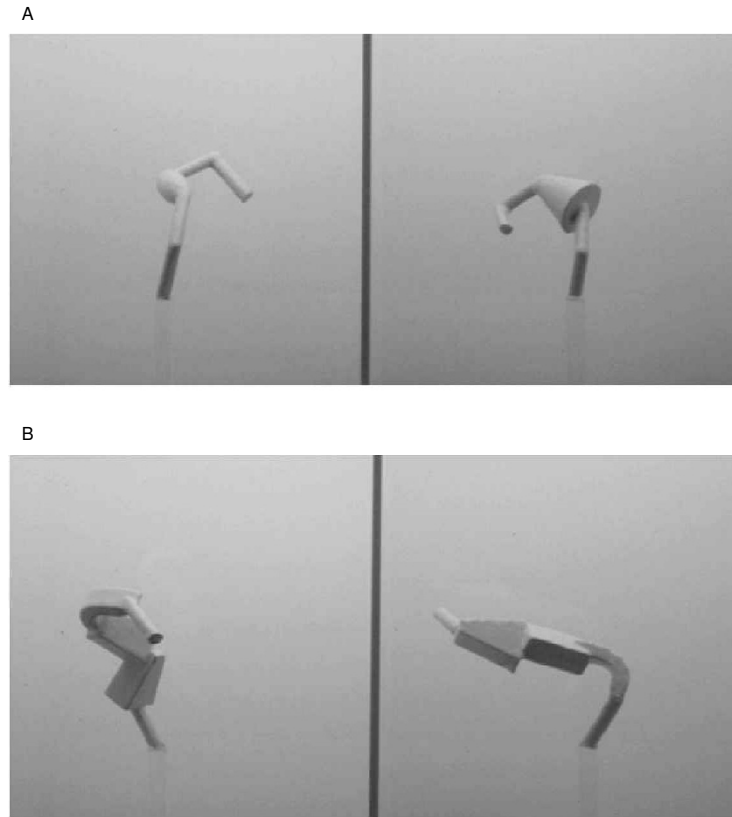
Figure 3. The top panel shows a pair of one-geon objects, and the bottom shows a pair of three-geon objects. Both photographs were taken from the front of the apparatus through the touch screen, as the objects would be seen by both human and pigeon subjects.

was opened, two of the three objects on the movable table could be seen through the Plexiglas viewing area. The middle object was always present but could be on the left or the right, depending on which position the movable table was in. The sliding door was in the closed position only when human subjects entered the room; it prevented them from seeing the starting position of the objects.

There was also a hinged Plexiglas cover attached to the top of the back of the three-chamber box (see Figure 2). It was closed once the objects were in place, to provide a backdrop for them.

The platforms on which the three objects were mounted were attached to individual stepper motors mounted under the table's surface. These motors and their controller cards were salvaged from 5 ½-in. floppy disc drives and, hence, were quite accurate. Because the motors had 100 steps, they allowed the objects to be positioned every 3.6°. The exact alignment and firm attachment of the objects with the bases of the turntables, together with the precision of the stepper motors, ensured that the objects would rotate positions accurately and without discernable variability in the final poses achieved.

The controller card from the floppy drive was activated by 5-V TTL signals from an 82C55 I/O card. For the floppy drive to function, the drive select line (pin 12) must be pulled low. The motor can be set to move clockwise by pulling the direction line (pin 18) high or counterclockwise by pulling the pin low. The motor was pulsed for 3 msec on pin 20 for each 3.6° step. Since the motor can be activated only if it senses a floppy disc in the drive, the presence of a disc is indicated to the controller card by pulling the motor on line (pin 16) low.

There were solenoid-type bird feeders on both sides of the pigeon box that were used to reinforce correct responses from the pigeons (see Figure 2, panel A). These feeders had an infrared beam of light (IR) across the opening of the feeder, which sent a 5-V signal back to the I/O card to indicate when a bird was feeding.

A few simple modifications to the apparatus were made for use with human subjects. First, we opened the door to the operant chamber and placed a chinrest in front of the door so that the subjects would view the objects from a fixed position in front of the chamber. Second, rather than using the touch screen, we provided a response box so that the subjects chose the left or the right object by pressing the left- or rightmost button on the box. Third, the food hoppers were not used, and instead we provided headphones through which the subjects were given feedback.
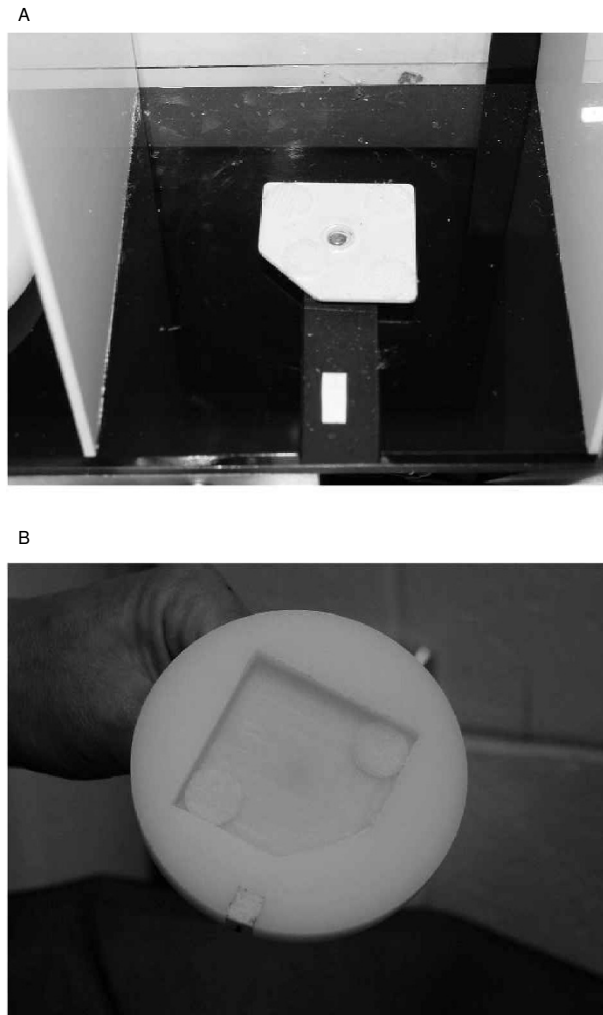
A

B

Figure 4. Panel A shows one of the pentagonal metal bases in the chamber that holds the objects, and panel B shows the Plexiglas bottom of one object's base, machined to the same shape as the metal bases in the object chambers. Both the metal base and the Plexiglas have Velcro pieces at two of their corners. The metal base is attached to a stepping motor.

Fourth, to prevent the subjects from seeing all three chambers when they entered the experimental room, a wooden box was built that covered the entire apparatus except for the front (Figure 2, panel B). The top of the box was hinged so that the entire back of the apparatus could be exposed when the stimuli were positioned.

Specifications

The movable table and sliding cover door slid back and forth on linear bearings. They were powered with reversible Hurst electric motors that were wired to solid-state relays (Crouzet GA8-4B02).

The solenoid type bird feeders were also wired to solid-state relays. These relays were controlled with an 82C55 I/O card installed in a 486 IBM-compatible computer.

Microswitches sensed when the moving table or sliding cover reached its limit and switched off the electric motor driving it by sending back a 5-V signal to the 82C55 I/O card.

The I/O card directly controlled the stepper motor operation by providing the 5-V signal that was required to control the direction of movement and the steps of the rotation.

The 82C55 I/O card controlled other relays that turned on and off lights located inside the feeder hoppers, the pigeon box (houselight), and each object chamber. All the lights were 24-V DC and were attached to the I/O card via solid-state relays.

To summarize, the 82C55 I/O card controlled the lights in each of the three object chambers, the houselight, the hopper lights, the hopper presentations, and the IRs within each hopper; it was used to move the table left and right

and to sense when the table was at the far left or the far right; it closed and opened the cover in front of the chambers and sensed when the cover was at the far left and the far right; and it rotated the stepper motors and controlled the direction of rotation of the motors.

The touch screen that recorded the pigeons' responses was a Carroll Touch (P/N 8100-9162-01) that communicated directly with the computer via the serial port.

The response box used for human subjects was a Psychology Software Tools serial response box (Model 200A).

Programs to operate the apparatus and collect the data were created using E-Prime software (Psychology Software Tools, 2001).

Summary

We have successfully used this apparatus to investigate recognition of depth-rotated objects by both pigeons and humans, and a series of comparative experiments is near completion. By varying the stimuli, the software programs, or the characteristics of the moving table (e.g., number of platforms), the apparatus can be used with other types of tasks (e.g., successive discriminations or classification tasks) or to investigate other kinds of processes (e.g., recognition of spatial layouts).

REFERENCES

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

DEREGOWSKI, J. B. (1989). Real space and represented space: Cross-cultural perspectives. *Behavioral & Brain Sciences*, **12**, 51-119.

DEREGOWSKI, J. B. (2000). Pictorial perception: Individual and group differences within the human species. In J. Fagot (Ed.), *Picture perception in animals* (pp. 397-429). Philadelphia: Taylor & Francis.

FAGOT, J. (ED.) (2000). *Picture perception in animals*. Philadelphia: Taylor & Francis.

PSYCHOLOGY SOFTWARE TOOLS (2001). *E-prime* (Version 1.0) [Computer software]. Pittsburgh: Author.

ROBERT MCNEEL & ASSOCIATES (2001). *Rhinocerous* (Version 1.1) [Computer software]. Seattle: Author.

SPETCH, M. L., & FRIEDMAN, A. (2003). Recognizing rotated views of objects: Interpolation versus generalization by humans and pigeons. *Psychonomic Bulletin & Review*, **10**, 135-140.

SPETCH, M. L., FRIEDMAN, A., & REID, S. (2001). The effect of distinctive parts on recognition of depth-rotated objects by pigeons and humans. *Journal of Experimental Psychology: General*, **130**, 218-225.

WATANABE, S. (2000). How do pigeons see pictures? Recognition of the real world from its 2-D representation. In J. Fagot (Ed.), *Picture perception in animals* (pp. 71-90). Philadelphia: Taylor & Francis.