Animal Learning & Behavior
1998, 26 (1), 85-102

Encoding of spatial information in images of an outdoor scene by pigeons and humans

MARCIA L. SPETCH, DEBBIE M. KELLY, and DAVID P. LECHELT
University of Alberta, Edmonton, Alberta, Canada

Pigeons and adult humans searched for a 2-cm^2 unmarked goal in digitized images of an outdoor scene presented on a touch-screen monitor. In Experiment 1, the scene contained three landmarks near the goal and a visually rich background. Six training images presented the scene from different viewing directions and distances. Subsequent unreinforced tests in which landmark or background cues were removed or shifted revealed that pigeons' search was controlled by both proximal landmarks and background cues, whereas humans relied only on the proximal landmarks. Pigeons' search accuracy dropped substantially when they were presented with novel views of the same scene, whereas humans showed perfect transfer to novel views. In Experiment 2, pigeons with previous outdoor experience and humans were trained with 28 views of an outdoor scene. Both pigeons and humans transferred well to novel views of the scene. This positive transfer suggests that, under some conditions, pigeons, like humans, may encode the three-dimensional spatial information in images of a scene.

Many organisms have been shown to use visual landmarks to remember and locate a goal area (see Collett, 1992, and Gallistel, 1990, for reviews). Use of visual landmarks is demonstrated experimentally by manipulating the visual cues between the opportunity to encode the spatial information and subsequent search tests in which the subject attempts to locate the goal. Two types of evidence indicate that a subject uses a particular landmark: (1) the subject shifts its search location in response to landmark shifts, and (2) accurate search is disrupted by the absence of the landmark. In a series of experiments using a laboratory task in which food is hidden on the floor of a spatial arena, Cheng and his colleagues (e.g., Cheng, 1988, 1989, 1994; Cheng & Sherry, 1992) have clearly demonstrated that pigeons use visual landmarks to locate a hidden goal, and they have identified several principles of pigeons' landmark-based search.

Pigeons' landmark-based search has also been investigated in a touch-screen task in which pigeons search for an unmarked goal on the surface of a monitor (Spetch, Cheng, & Mondloch, 1992; Spetch & Mondloch, 1993; Spetch & Wilkie, 1994). Visual stimuli displayed on the monitor served as landmarks. This two-dimensional (2-D) search task has yielded results that are remarkably similar in certain ways to those found in real-space tasks. In particular, near landmarks are weighted more heavily than are far landmarks (Cheng, 1989; Spetch & Wilkie, 1994), and landmarks near an edge of the search space exert more control in the dimension parallel to the edge than in the dimension perpendicular to the edge (Cheng & Sherry, 1992; Spetch et al., 1992). Recent work (Spetch, Cheng, & MacDonald, 1996; Spetch et al., 1997) has revealed that, on both the touch screen and the laboratory floor, pigeons can use the configuration of an array of landmarks to locate a goal. In both environments, however, they respond to expansions of the array by maintaining the absolute training distance from individual landmarks in the array, rather than adjusting distance so as to maintain the appropriate relative position with respect to the entire array of landmarks.

Although the results obtained to date suggest that the touch-screen task calls forth many of the spatial processes used in more naturalistic search tasks, the visual information provided in studies using the touch-screen task has clearly differed from that in most real-world spatial search situations. In particular, whereas landmarks found in natural settings are typically located within a rich visual context consisting of many distal as well as local stimuli, the landmarks presented in the touch-screen studies have been presented against visually uniform backgrounds. In most studies, the landmarks have been colored 2-D geometric stimuli presented against a uniform gray background (Cheng & Spetch, 1995; Spetch, 1995; Spetch et al., 1992; Spetch & Mondloch, 1993). One study (Spetch & Wilkie, 1994) presented visually richer stimuli in the form of digitized images of an outdoor scene consisting of three objects on a grassy field. However, the grassy field was reasonably uniform, except for a gradient in the texture of the grass from the top (background) to the bottom (foreground) of the image. Most importantly,

This research was supported by a Natural Sciences and Engineering Research Council of Canada research grant held by the first author. Some of the data from Experiment 1 were presented at the Conference on Comparative Cognition, Melbourne, FL, in March 1993, and at the meeting of the Psychonomic Society, Los Angeles, CA, in November 1995. The results of Experiment 2 were presented at the International Conference on Comparative Cognition, Melbourne, FL, in March 1997. The authors thank R. Kelly and A. Soldat for assistance with the research and D. Treit for comments on an earlier version of the manuscript. Correspondence should be addressed to M. L. Spetch, Department of Psychology, University of Alberta, Edmonton, AB, Canada T6G 2E9 (e-mail: mspetch@psych. ualberta.ca).

—Accepted by previous editor, Robert A. Rescorla

there were no objects in the background of the scene that would provide a spatial context for the landmarks.

Distal background cues in a scene could play an important role in spatial search. For example, distal cues may be used to define the general region in which to search, with local landmarks serving to pinpoint the goal within that region. Distal cues may also provide a context that determines the meaning of local landmarks (e.g., the clump of trees beside the river may signify something different than a similar looking clump of trees in an open field). Finally, distal cues may be used as additional redundant sources of information to locate a goal. Indeed, laboratory studies of spatial memory have suggested that both global and local cues may be encoded to remember a goal location (Brodbeck, 1994; Spetch & Edwards, 1988).

Another way in which the visual information during real-world search is very different from that available in the touch-screen studies is that a goal area can often be approached from different directions in the real world. This means that the visual information used in searching for the goal changes because it is viewed from different perspectives. The touch-screen studies published to date, however, have provided only a single perspective of the landmark array.

The present research used the touch-screen task to investigate pigeons' spatial search in images of outdoor scenes that included landmarks near the goal and a visually rich set of background cues. In Experiment 1, subjects were trained with six images that provided some variation in direction and distance from which the scene was viewed. This meant that the three-dimensional (3-D) spatial relationships in the scene depicted by the images were invariant, but the 2-D spatial relationships among the visual features varied somewhat from image to image. A variety of tests were conducted following training to determine what information in the scene was used when searching for the goal and whether accurate searching would transfer to novel views of the same scene. For comparison, adult humans were trained with the same set of training images and then were given some of the same tests presented to pigeons.

In Experiment 2, racing pigeons with prior outdoor experience were trained with 28 different views of another complex scene and then were tested for transfer to novel views. Again, humans were trained and tested with the same stimuli for comparison.

EXPERIMENT 1

Method

Subjects

Pigeons. Four adult Silver King pigeons served in Experiment 1. One pigeon was experimentally naive, and 1 had previous experience in standard operant chambers but had never served in a touch-screen task. The other 2 pigeons were experienced at performing in the touch-screen spatial search task, but with visual stimuli different from those used here. The birds were housed in large individual cages under a 12-h light:dark cycle (lights on at 6:00 a.m.). All birds were maintained at approximately 85% of their free-feeding weights by Kay Tee pigeon pellets obtained after experimental sessions and mixed grain during experimental sessions. Water and grit were available ad lib in the home cages.

Humans. The human subjects were 8 female undergraduate students, ranging from 19 to 47 years of age. They participated in the experiment for credit for their introductory psychology course. Four participants were assigned to the landmark removal test series, and 4 were assigned to the transfer test series.

Apparatus, Search Space, and Images

Pigeons. The experiment was conducted in a large custom-built chamber, 44 cm high, 32 cm deep, and 74 cm wide (inside dimensions). A Zenith 1492 color monitor with attached infrared touch frame (Carroll Touch, 1492 Smart Frame) was placed against an opening centered in the back wall of the chamber. This opening was 10 cm from the raised grid floor of the chamber, and it provided access to the entire surface of the monitor. Spacers were used to recess the touch frame by approximately 3 cm from the opening and to separate the frame from the monitor by approximately 1.6 cm. Two Gerbrands pigeon grain feeders were mounted on the back wall, one on each side of the monitor. The feeder openings began 8.5 cm from the sides of the monitor opening and 17 cm from the floor. Lamps located within each feeder illuminated feeder presentations, and photocells in each hopper measured head entries so as to limit eating times. Microcomputers, located in an adjacent room, controlled experimental contingencies and recorded peck coordinates. The touch frame was programmed to detect individual responses (i.e., detection of a beam break, then a return to unbroken beams before another response would be recorded).

The search space was a rectangular area, approximately 26 × 20 cm, on the surface of the color monitor. The images displayed during training provided six different views of an outdoor scene that contained three landmarks (a chair, a pot of flowers, and a pile of logs) near the goal and several visually distinct background cues (house, flowers, trees, etc.). These six views are shown in Figure 1. The rectangular object that can be seen near the three landmarks was a blue plastic box that served as a marker for the goal area. This goal marker was present in the images only during preliminary training. Thereafter, it was edited out of the images so that it was absent from the scene during the later stages of training and during all testing. Note that the six different views used in training not only provided variation in the vantage point of the scene but also provided variation in the absolute location on the screen of the goal.

The images were digitized and constructed using procedures similar to those described in Spetch and Wilkie (1994) and Wilkie, Mak, and Saksida (1994). The landmarks and goal marker were laid out, and the scene was videotaped from various viewpoints. The videotape was then played into a computer using a frame grabber system (Creative Labs Video Blaster), and still frames were saved in a GIF format (Compuserve, Inc.). Six of the saved images were selected for use in training, and an additional set of six images was selected for use in transfer tests. Removal of the goal marker from the images and manipulation of the visual information in the scenes for the various test phases were accomplished by editing the images with Photofinish software (Zsoft).

For each view, the location of the goal needed to be matched to the touch-screen coordinates. This matching was done by displaying the scene with the goal marker present, touching in the center of the goal marker, and reading off the coordinates recorded by the touch frame. This was repeated several times for each image to ensure reliability.

Humans. The experiment took place in a small private room that contained a touch-screen-equipped computer. The computer monitor (Zenith 1490) and touch screen (Carroll Touch 1490 Smart Frame) provided the same search space and stimuli as those used for the pigeons in Experiment 1. The subjects sat in a chair in front of the monitor and searched by touching the screen with the eraser end of a pencil.

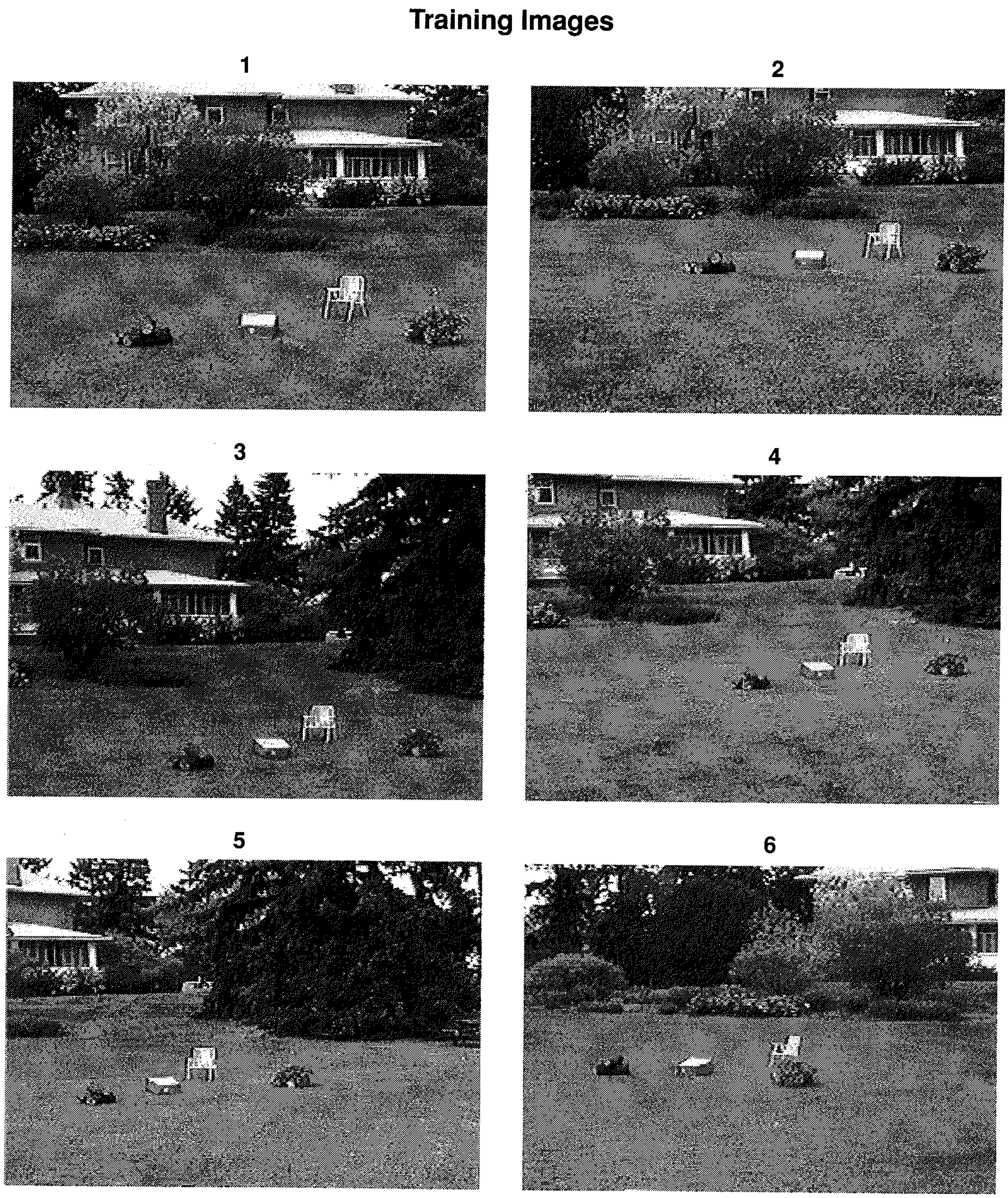

Figure 1. Images showing the six training views used for the subjects in Experiment 1. The goal marker (the box located between the chair and logs) is shown in these images for illustrative purposes. This marker was edited out of the images after preliminary training and was not present during any baseline or test trials.

Procedure for Pigeons

Preliminary training. The 2 pigeons with no previous experience in the touch-screen task received one or two sessions of training to eat from the raised illuminated hopper. All birds then received several sessions of training with a modified autoshaping procedure to establish reliable pecking at the monitor. Initially, a 2-cm yellow square was intermittently presented in various screen locations against a dark gray background, with 60-sec intertrial intervals (ITIs). The yellow square remained on until the bird pecked at the square or until 8 sec elapsed, and then food was presented. Once reliable pecking was established to the yellow square, the background was changed to green and the square was changed to blue. Once a bird pecked reliably at the blue square, the six images shown in Figure 1 were introduced. Pecks in a 2-cm area centered at the blue box that served as the goal marker were reinforced immediately; otherwise, the display terminated with food after 8 sec. For the final session of preliminary training, the ITI was decreased to 5 sec, and the image stayed on until a peck in the goal area was recorded.

Search training. During initial search training sessions, the goal marker was gradually eliminated so that the pigeons needed to rely on visual cues in the scene to find the goal. The marker was removed by editing the images to cover the marker with patches of grass using a "clone" tool. Images were created in which the grass covered about 50%, 75%, and, finally, 100% of the marker (the control image in Figure 2 shows an example of an image with the marker 100% covered). This phase of search training continued until the bird completed (by pecking in the goal area) at least 80 trials in a 60-min session with the goal marker absent. During the next few sessions, the number of pecks in the goal required to terminate the trial and obtain food was increased to two and then to three. Next, a consecutive peck requirement was added such that the last two pecks recorded had to be consecutive pecks in the goal area. Thus, the response requirement during the final stage of training and for all subsequent baseline sessions was a minimum of three pecks in the goal with the final two pecks being consecutive pecks in the goal. Each increment in the response requirement was implemented only if the bird completed 80 trials in a session. During the last set of search training sessions, the percentage of trials on which reinforcement was available was decreased from 100% to 50%. On unreinforced trials, completion of the response requirement terminated the display, but no food was presented. Each bird remained on the 50% reinforcement condition for a minimum of 5 sessions and advanced to the test phase only after it completed at least 80 trials on each of the last two sessions. Search training required between 50 and 70 sessions.

Testing. All birds received several series of test sessions. Two or more baseline sessions, identical to the last phase of training, were inserted between each test series. Baseline sessions were also interspersed among the test sessions (typically, one baseline session after every 3rd or 4th test session) within each test series. Each test session consisted of a mixture of reinforced and unreinforced baseline trials, control trials that were visually identical to baseline trials but were procedurally the same as test trials, and test trials in which visual information in the image was altered in some way. On control and test trials, the image remained on for 8 sec following the second peck recorded anywhere on the screen. In all test sessions, at least 50% of the trials were reinforced baseline trials. Each type of scheduled test trial occurred once in each block of 15 or 16 trials, with the order varying randomly between blocks.

For one test series, the images were edited to remove one or two of the local landmarks. These tests provided an assessment of accuracy with each landmark removed from the scene and with each landmark being the only local landmark present in the scene. Removal of the landmark was accomplished by cloning grass on top of the landmark. Test and control trials were conducted with each of the six training views.

For a second test series, the images were edited to remove all three local landmarks (see "No Lms" image in Figure 2), to remove all of the background information (see "No Bk" image in Figure 2), or to remove both the landmarks and the background (see "None" image in Figure 2). These removals were also accomplished by copying grass over the landmarks or background. Test and control trials were again conducted with each of the six training views.

For a third test series, the set of three landmarks was shifted horizontally in some tests and vertically in others. Only Views 2, 4, 5, and 6 were used for these tests. Vertical shifts were always down. Horizontal shifts were to the left in Views 2 and 4 and to the right in Views 5 and 6. All vertical shifts were by 2 cm. The horizontal shifts were by 2 cm in Views 1, 4, and 5, but by error were only 1 cm in View 6. The image labeled "Lm Shift" in Figure 2 shows an example of a leftward shift of the landmarks in View 4. Shifts were accomplished by cloning the area containing the landmarks to the desired location and then cloning grass over the top of whatever remained of the original landmarks.

For a fourth test series, the backgrounds were swapped for three pairs of training views. This was accomplished by cloning the background of one image (the entire portion of the image above the landmarks) onto the other image. An example of View 4 containing the background of View 1 is shown as the image labeled "Bk Swap" in Figure 2.

The final test series assessed whether the birds would transfer the learned spatial relationships from the set of six training views to a new set of images depicting the same scene but from different distances, perspectives, or orientations. The six images used for the transfer tests are shown in Figure 3. Each test session included 50% reinforced trials with the trained views, 25% unreinforced control trials in which each of the trained views occurred equally often, and 25% unreinforced test trials in which each of the transfer test images was presented equally often. Each bird received four transfer test sessions, each separated by a baseline session.

In all series, each individual test image (i.e., each test type in each view) was presented at least 10 times.

Procedure for Humans

The images used for the human subjects were the same as those used for the pigeons. At the beginning of each session, the subjects were provided with the following information: Their task would be to remember and locate a goal area, which initially would be marked but later would be unmarked. A point would be available on randomly selected trials. They could obtain this point only by touching in the goal area. The required number of touches in the goal would vary across trials, and sometimes consecutive touches in the goal would be required. Whenever they earned a point, their cumulative points would be displayed. When points were not available, the trial would end after a certain number of touches, whether or not they were correct. The session would end after they obtained a certain number of points or after 45 min, whichever came first.

The experimenter then ran a demonstration program that presented three training trials. The experimenter demonstrated the importance of holding the pencil straight and removing it between touches. The first two demonstration trials presented images with the marker present, and the last demonstration trial presented an image with the marker absent. Two touches in the goal were required during each demonstration trial. Each trial ended as a "reinforced" trial in which a point was earned (i.e., the screen was blanked and cumulative points were displayed). Any procedural questions were answered with "I'm sorry but I cannot provide any further information until after you have completed the experiment." After confirming that the subject wished to participate, the experimenter started the program and left the room.

The subject's program began with 6 trials that presented each of the six training images with the goal marker present. The next 6 trials presented images showing the same views but with the marker absent. This was followed by 6 more trials with the marker present. The marker was absent in all remaining trials of the session. The response requirement for all baseline trials included a variable number of minimum touches (one to five) and a variable number of consecutive touches in the goal (ranging from no consecutive touch requirement to three consecutive touches required). Beginning with Trial 13, only 50% of the trials were reinforced (i.e., earned points). On unreinforced baseline trials, the display ended with a blank screen upon completion of the response requirement.

Beginning with Trial 31, control and test trials were interspersed among the baseline trials. During control and test trials, the second touch recorded anywhere on the screen initiated an interval that varied randomly from 2 to 4 sec, after which the first response anywhere on the screen terminated the trial. Each of the six training views was presented equally often on control trials and served equally as the basis for test images. The subjects assigned to the removal test series were tested with images in which all three local landmarks

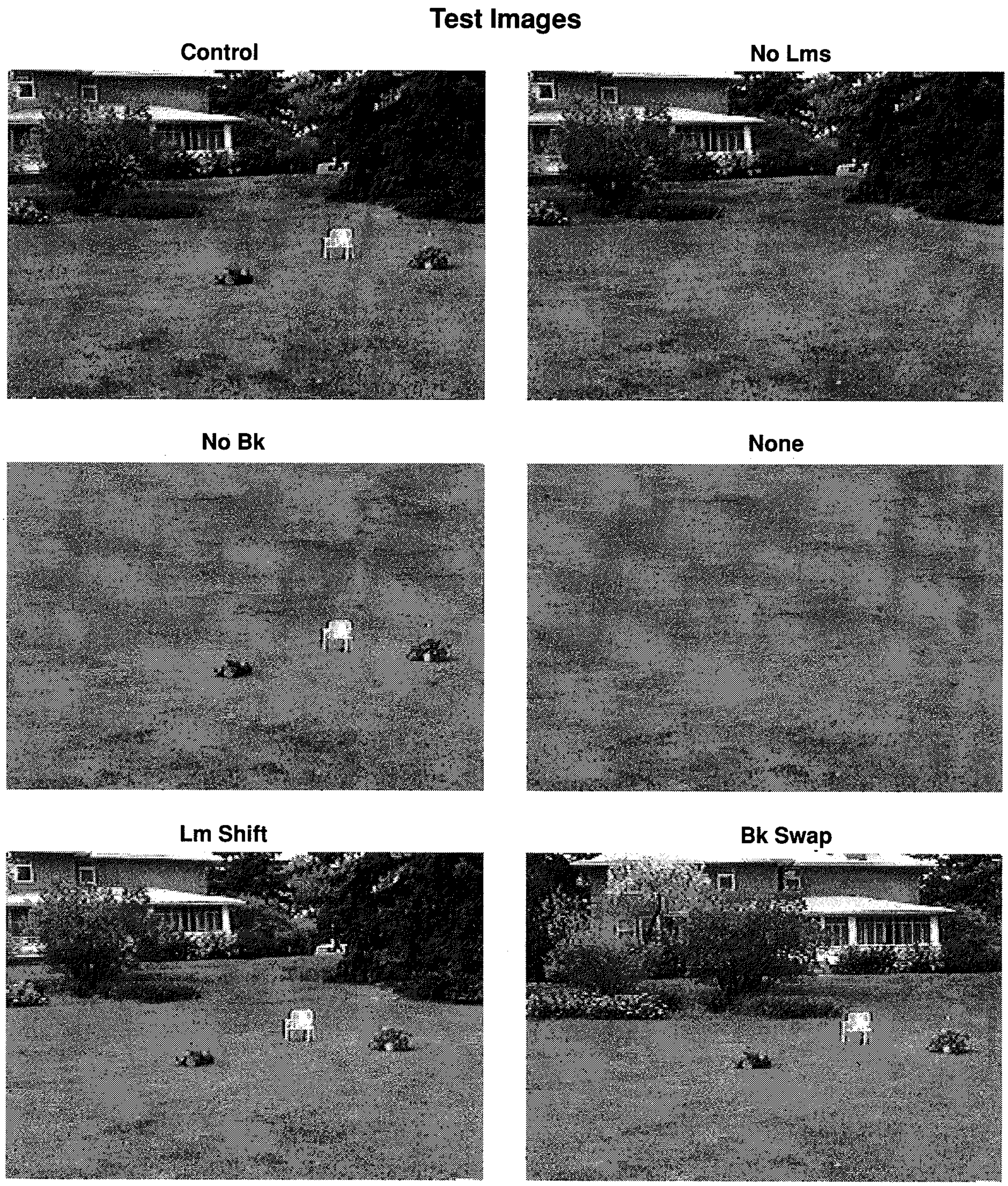

Figure 2. An example of control and test images used in Experiment 1. Lm = landmark; Bk = background. The test images shown are based on Training View 4. Tests based on other training views were also presented.

were removed, images in which all of the background information was removed, and images in which both the landmarks and the background were removed. The subjects who participated in a transfer test series were tested with each of the six transfer images shown in Figure 3. All subjects were debriefed following the session.

Data Recording and Analysis

All data presented are from unreinforced control and test trials. Each response was recorded in terms of touch-screen coordinates that ranged from 0 to 640 in the horizontal dimension and 0 to 480 in the vertical dimension. Analysis of individual trial performance