CONCEPT LEARNING BY MONKEYS WITH VIDEO
PICTURE IMAGES AND A TOUCH SCREEN

RAMESH S. BHATT AND ANTHONY A. WRIGHT

GRADUATE SCHOOL OF BIOMEDICAL SCIENCES,
UNIVERSITY OF TEXAS HEALTH SCIENCE CENTER AT HOUSTON

Two rhesus monkeys were trained in a same/different task to discriminate digitized computer-stored picture stimuli. The pictures were digitized from 35-mm slides and presented in pairs on a computer monitor. The monkeys were required to touch the pictures and then make a choice response to indicate whether the pictures were identical or nonidentical. The response areas and stimuli were located to the sides of the picture stimuli. Responses were defined and monitored by an infrared matrix touch screen. After learning the same/different task, both monkeys showed performance accuracy with novel picture stimuli similar to that with training picture stimuli. This accurate novel-picture transfer indicates that a same/different concept had been learned, a concept similar to the one they had previously demonstrated in a different apparatus with rear-projected slide stimuli and a response lever.

Key words: concepts, concept learning, abstract concepts, computer images, picture stimuli, touch screens, same/different task, monkeys

The study of animal cognitive processes has become increasingly popular (see, e.g., special issues of Learning and Motivation, 1987, and the Journal of the Experimental Analysis of Behavior, 1989). This popularity has been facilitated by the development of new experimental techniques and procedures aimed at improving the study of an animal's cognitive abilities (e.g., concept learning and memory). One fundamental characteristic of many of the new procedures has been the trend to use larger numbers of complex multidimensional stimuli (Wright, in press). Large numbers of training stimuli reduce proactive interference and in turn can enhance learning and the final performance level (e.g., Jitsumori, Wright, & Cook, 1988; Wright, Urcuioli, & Sands, 1986). Large numbers of training stimuli produce better concept learning in humans and animals (e.g., Homa & Chambliss, 1975; Homa, Cross, Cornell, Goldman, & Swartz, 1973; Homa, Sterling, & Treple, 1981; Omohundro, 1981; Wright, Cook, Rivera, Sands, & Delius, 1988). Tests for whether or not the concept has been learned also require a considerable number of stimuli, in addition to the training stimuli, because test stimuli must be novel, they can be used only once (by definition of novelty), and a sizable number of test trials often are required for statistical reasons.

One of the few ways to provide a large number of distinctly different stimuli is to use pictures of objects and scenes (i.e., travel slides). When such pictures were first used in this research application they were displayed with Kodak® carousel projectors (e.g., Bhatt, Wasserman, Reynolds, & Knauss, 1988; Overman & Doty, 1980; Roberts & Kraemer, 1981; Sands & Wright, 1980a, 1980b, 1982; Santiago & Wright, 1984; Wasserman, Kiedinger, & Bhatt, 1988; Wright, Santiago, & Sands, 1984; Wright, Santiago, Sands, Kendrick, & Cook, 1985). Recently developed image-processing systems with their electronic storage and display of picture stimuli, however, afford several advantages over carousel slide projectors. One advantage is their easy generation of sequences of to-be-presented pictures. Instead of the several hours required to produce each stimulus sequence by rearranging slides in carousel trays, image-processing systems allow fast daily randomization and stimulus counterbalancing through computer programs. Another
advantage of these computerized image-processing systems is their fast stimulus presentation rates (e.g., 10 per second), which are not easily accomplished by other means (e.g., 8-mm movie film with frame-by-frame projection; Cook, Wright, & Sands, 1991; Sands, Urcuioli, Wright, & Santiago, 1984; Wright, Cook, et al., 1990). Computer storage and presentation of slide pictures also allow easy manipulation of aspects of the pictures, such as color changes, black/white images of color slides, enlargement or shrinkage of portions of the image, and even spatial rearrangement of the image components. Such manipulations may be necessary to explore the range of stimulus dimensions over which the learned concept, category, or “rule” can apply.

In addition to providing superior storage, manipulation, and presentation, the new image-processing systems have another advantage. Touch screens can be added or built into the video monitor to record the subject’s touches (or pecks in the case of birds) to different parts of the monitor screen on which stimuli appear. These touch screens eliminate the need for levers, push buttons, or pecking keys, which often tend to separate the response from the discriminative stimuli. It has been shown with rats, pigeons, and monkeys that contact with the stimuli enhances learning compared to conditions in which the response manipulation and the stimuli are spatially separated (e.g., Harrison, 1984; Harrison, Downey, Segal, & Howe, 1971; Harrison, Iversen, & Pratt, 1977; Stollnitz, 1965; Wright, Shyan, & Jitsumori, 1990). Thus, the use of touch screens should also create a potential for more rapid learning and a greater asymptotic accuracy.

Here we present results from training and testing 2 monkeys with an image-processing and touch-screen system in a same/different concept learning task.

**METH

**Method**

**Subjects**

The subjects were 2 adult male rhesus monkeys (*Macaca mulatta*), Linus and Max. Both monkeys had extensive experience in same/different and serial probe recognition (list) memory tasks, but had never been exposed to video picture displays or had been required to touch picture stimuli. They were maintained on a 14:10 hr light/dark cycle, and their access to food (Purina® monkey chow) was restricted for about 16 hr prior to and for 2 hr following experimental sessions.

**Apparatus**

The completely enclosed experimental chamber (68.6 cm high, 48.3 cm wide, and 53.7 cm deep) was constructed from sheet aluminum (0.635 cm thick) by the authors. The four walls of the chamber extended to the floor to support the chamber. The chamber’s floor consisted of 14 aluminum bars (1.27 cm diameter) running the width of the chamber, spaced 3.8 cm center to center. A snug-fitting stainless steel pan was positioned 15.2 cm below the floor bars and 10.5 cm above the floor. Ventilation was provided by an exhaust fan (Dayton Fan Model 4C440 with rheostat control) located on top of the chamber.

The chamber was designed to be moderately confining with few distractions in order to direct the monkeys’ attention to the video monitor. The design eliminated potential distractions such as screw heads and holes that might otherwise be visible from inside the chamber, and the inside surface of the chamber was sanded to a uniform matte finish. Furthermore, all chamber illumination was provided by the video monitor. An aluminum shelf attached to the outside of the chamber supported the video monitor and a pellet dispenser. The 33-cm NEC video monitor (Model JC-1401P3A Multisync color monitor with an 800 × 560 resolution) and the attached touch screen fit snugly into a cut-out area (approximately 28.1 cm high by 34.9 cm wide) in the front wall of the chamber; the center of this cut-out was approximately 40.6 cm above the chamber floor bars. A Gerbrands pellet dispenser (Model G5120) dispensed 300-mg Noyes banana pellets through a tube into a food cup placed 11.4 cm below the center of the lower edge of the bezel surrounding the video screen.

An AST (Premium 286) computer controlled the pellet dispenser through a peripheral interface (Metabyte PIO-12), collected and analyzed data, digitized stimuli, and controlled display of the stimuli.

**Stimuli**

The stimuli were derived from 35-mm color slides of a wide variety of distinctively different animals, flowers, fruits, people, and other nat-
ural and human-made objects. The training stimuli consisted of 600 pictures that were selected from a pool of 4,000 slides to which the monkeys had been exposed in previous experiments. In addition, 75 pictures that the monkeys had not previously seen were used to test for concept learning following acquisition of the task.

Digitized images of these slides were created by a special camera, two image-processing cards, a computer, and software programs. The camera (Howtek Photomaster, Model 87RU) allowed any 35-mm slide picture to be digitized and stored in the computer. Slide images were displayed on the video monitor just as they would be seen in the experiment. The camera had knobs to crop, zoom, shrink, or rotate the image, and to adjust contrast, focus, brightness, or color. The color adjustment was a joy stick that allowed relative proportions of the three primaries blue, green, and red to be continuously varied. Once each picture was properly adjusted (often requiring only a minute or two), it was stored on the computer hard disk (70-megabyte capacity) with several key strokes.

Pictures were stored in a 256 by 256 format because this resolution closely matched the pixel density of the portion of the video screen on which they were displayed (higher resolutions were easily obtainable but required more memory storage). With this format, and with the disk space occupied by other programs, 200 pictures were stored on the hard disk along with other necessary programs for this experiment.

Slides were digitized with a Truevision TARGA-16 image-processing card, and displayed with a Truevision VISTA card. The advantages of the VISTA card over the TARGA card are its greater memory (4 megabytes), flexibility (e.g., dividing the memory into different images), and speed (e.g., 10 or more images per second), which can be appreciated in the conduct of list-memory experiments.

The size of the video screen, inside the touch-screen bezel, was 21.6 cm high and 27.9 cm wide, and the picture stimuli were each 8.5 cm high and 13.0 cm wide. One stimulus was displayed above the other. The sample stimulus was displayed in the top half of the screen, and the probe stimulus was displayed in the bottom half of the screen. Both pictures were centered with respect to an imaginary vertical center line.

In addition to the two picture stimuli, there were two choice-response areas on the monitor for the monkeys to touch and make their same/different choice responses. The choice-response area to indicate that the sample and probe were identical (same) was located 1 cm to the right of the division between the two stimuli; this area was indicated by a green ellipse (7.6 cm high and 5.1 cm wide) with a black letter “S” (3.8 cm high) superimposed upon it. The choice-response area to indicate that the sample and probe were nonidentical (different) was located 1 cm to the left of the division between the two stimuli; this area was indicated by a red rectangle (7.6 cm high and 5.1 cm wide) with a white letter “D” (3.8 cm high) superimposed upon it.

Procedure

The monkeys were initially trained to touch the picture images on the monitor screen by reinforcing successive approximations to the desired touch response. The experimenters viewed the monkeys with a high-sensitivity TV camera (Sony® SSC-D5) mounted on a chamber wall and pointed at an angle of about 45° toward the center of the stimulus video screen. A transparent Plexiglas template (0.95 cm thick) was superimposed on the video screen and was held in place with Velcro® strips and double-sided sticky tape. The template had holes (6.4 cm diameter) cut out over the sample and probe locations and holes (5.1 cm diameter) cut out over the same and different response locations. This template guided the monkeys’ fingers to the appropriate locations on the screen and was removed after the monkeys became proficient at touching the appropriate images.

Following pretraining, the monkeys were trained on a simultaneous same/different task. Daily training sessions consisted of 100 trials; an equal number of same and different trials was presented in a different random sequence each day. Each trial was preceded by a 10-s intertrial interval, after which a 1000-Hz tone was presented for 800 ms, and then the sample stimulus for that trial was presented. The monkey was required to touch the sample stimulus a programmed number of times (fixed-ratio observing response requirement); then the probe stimulus was presented. Next, the
ural and human-made objects. The training stimuli consisted of 600 pictures that were selected from a pool of 4,000 slides to which the monkeys had been exposed in previous experiments. In addition, 75 pictures that the monkeys had not previously seen were used to test for concept learning following acquisition of the task.

Digitized images of these slides were created by a special camera, two image-processing cards, a computer, and software programs. The camera (Howtek Photomaster, Model 97RU) allowed any 35-mm slide picture to be digitized and stored in the computer. Slide images were displayed on the video monitor just as they would be seen in the experiment. The camera had knobs to crop, zoom, shrink, or rotate the image, and to adjust contrast, focus, brightness, or color. The color adjustment was a joy stick that allowed relative proportions of the three primaries blue, green, and red to be continuously varied. Once each picture was properly adjusted (often requiring only a minute or two), it was stored on the computer hard disk (70-megabyte capacity) with several key strokes.

Pictures were stored in a 256 by 256 format because this resolution closely matched the pixel density of the portion of the video screen on which they were displayed (higher resolutions were easily obtainable but required more memory storage). With this format, and with the disk space occupied by other programs, 200 pictures were stored on the hard disk along with other necessary programs for this experiment.

Slides were digitized with a Truevision Targa-16 image-processing card, and displayed with a Truevision VISTA card. The advantages of the VISTA card over the TARGA card are its greater memory (4 megabytes), flexibility (e.g., dividing the memory into different images), and speed (e.g., 10 or more images per second), which can be appreciated in the conduct of list-memory experiments.

The size of the video screen, inside the touch-screen bezel, was 21.6 cm high and 27.9 cm wide, and the picture stimuli were each 8.5 cm high and 13.0 cm wide. One stimulus was displayed above the other. The sample stimulus was displayed in the top half of the screen, and the probe stimulus was displayed in the bottom half of the screen. Both pictures were centered with respect to an imaginary vertical center line.

In addition to the two picture stimuli, there were two choice-response areas on the monitor for the monkeys to touch and make their same/different choice responses. The choice-response area to indicate that the sample and probe were identical (same) was located 1 cm to the right of the division between the two stimuli; this area was indicated by a green ellipse (7.6 cm high and 5.1 cm wide) with a black letter “S” (3.8 cm high) superimposed upon it. The choice-response area to indicate that the sample and probe were nonidentical (different) was located 1 cm to the left of the division between the two stimuli; this area was indicated by a red rectangle (7.6 cm high and 5.1 cm wide) with a white letter “D” (3.8 cm high) superimposed upon it.

Procedure

The monkeys were initially trained to touch the picture images on the monitor screen by reinforcing successive approximations to the desired touch response. The experimenters viewed the monkeys with a high-sensitivity TV camera (Sony® SSC-D5) mounted on a chamber wall and pointed at an angle of about 45° toward the center of the stimulus video screen. A transparent Plexiglas template (0.95 cm thick) was superimposed on the video screen and was held in place with Velcro® strips and double-sided sticky tape. The template had holes (6.4 cm diameter) cut out over the sample and probe locations and holes (5.1 cm diameter) cut out over the same and different response locations. This template guided the monkeys’ fingers to the appropriate locations on the screen and was removed after the monkeys became proficient at touching the appropriate images.

Following pretraining, the monkeys were trained on a simultaneous same/different task. Daily training sessions consisted of 100 trials; an equal number of same and different trials was presented in a different random sequence each day. Each trial was preceded by a 10-s intertrial interval, after which a 1000-Hz tone was presented for 800 ms, and then the sample stimulus for that trial was presented. The monkey was required to touch the sample stimulus a programmed number of times (fixed-ratio observing response requirement); then the probe stimulus was presented. Next, the
monkey was required to touch the probe stimulus the same number of times that it touched the sample in order to produce the same/different choice stimuli. The observing response requirement to sample and probe stimuli was gradually increased from one to eight responses from Session 1 to Session 58 of training, and remained at that value for the remainder of the experiment. A single touch to the correct choice image resulted in removal of all images and presentation of a banana pellet accompanied by an 800-ms 500-Hz tone. A touch response to the incorrect choice image resulted in a 10-s blackout and repetition of the trial. Each trial was repeated until a correct response was made. Performance on these correction trials did not contribute to the analyses of daily accuracy, nor did they influence the number of trials (100) presented in a daily session.

The 2 monkeys were each trained for 116 sessions using this procedure. A set of 150 stimuli was used during the first 76 sessions. Each stimulus was presented on only one trial during each session. One hundred pictures were randomly chosen daily and used as sample stimuli on same trials; the remaining 50 pictures were used as probes on different trials. Different sets of 150 pictures were introduced in Sessions 77, 93, and 105; thus, a total of 600 pictures were used during the course of training. These were pictures that the monkeys had seen in prior experiments in a different apparatus. Therefore, to conduct a valid test of the same/different concept, the monkeys were tested with an additional 75 novel pictures that they had never seen before.

The 75 novel pictures used in 50 test trials (25 same and 25 different trials) were distributed over five consecutive transfer test sessions. Each transfer test session consisted of 10 novel-stimuli test trials quasi-randomly intermixed with 90 familiar-stimuli training trials. Correct responses on test trials were reinforced, and apart from the novelty of the sample and probe stimuli, test trials were otherwise identical to training trials.

RESULTS

Figure 1 shows the performance accuracy of the 2 monkeys during acquisition. These accuracy scores are mean percentage correct choices from blocks of four daily sessions and
CONCEPT LEARNING BY MONKEYS

did not include any performance on correction trials. Accuracy rose from near-chance level (50% correct) in the first block to better than 80% correct by the end of training.

There were no apparent disruptions when the different sets of 150 training stimuli (Sets 2, 3, and 4) were introduced in the first session of Blocks 20, 24, and 27. Accuracy upon introducing these new stimulus sets (Sessions 77, 93, and 105 of Blocks 20, 24, and 27, respectively) was 82%, 68%, and 95% for Linus and 73%, 76%, and 74% for Max, respectively. This lack of disruption when the new stimulus sets were introduced indicates that the monkeys generalized their same/different responding to stimuli that they had not previously seen in this particular setting.

The monkeys also generalized their same/different responding to stimuli that they had never seen before, in any setting (Figure 2). Transfer performance to novel stimuli was 88.0% and 74.0% correct for Linus and Max, respectively. At the same time, performance with the training stimuli remained accurate, 88.5% for Linus and 84.4% for Max. Transfer performance did not differ significantly from training trial performance: Linus, t(4) = 0.08, p > .05; Max, t(4) = 2.40, p > .05. This accurate responding with novel stimuli indicates that the monkeys had learned the same/different concept. This same/different concept is an abstract concept because performance transcended the individual stimuli used in training.

DISCUSSION

Digitized video images may be the wave of the future in the testing of animal cognitive processes. The technology is available and the resolution capabilities are good and continue to improve at an ever-decreasing price. With the increase in hard disk storage capacity of personal computers to 100 or more megabytes at a relatively low cost, storage of images for many experiments can be accomplished on hard disks. Another current cost-effective option is the read/write laser disk for applications for which a gigabyte of storage is needed (see Appendix). Monitoring responses is now possible with touch screens built in to the video monitors; these touch screens can identify precisely which portion of the picture was touched (or pecked if pigeons are the subjects, e.g., Pisacreta & Rilling, 1987; Wright et al., 1988).

In the apparatus of the present experiment, the monkeys learned more rapidly (<12,000 trials) the same/different task and concept than they had previously learned (>32,000 trials) a similar task and concept in an apparatus with a response lever and slide projectors (Wright, Santiago, Urcuioli, & Sands, 1984). Although it could be argued that learning in the previous task may account, in part, for the more rapid learning in the present task, it is also possible that the prior training could have interfered with learning in the present task. The monkeys had worked on the previous task for more than 9 years, making responses with a three-position lever and looking at slide pictures projected on screens 61 cm distant. The more rapid learning in the present experiment, we feel, was due to such things as contact with the picture stimuli and daily randomizations of the stimuli rather than to prior experience in the other apparatus.

In any case, the present experiment has shown that monkeys can be efficiently trained to discriminate among stimuli shown on a video monitor, to learn a task requiring separate touch responses to indicate same or different, and to transfer accurately a same/different concept to novel stimuli.

REFERENCES

Bhatt, R. S., Wasserman, E. A., Reynolds, W. F., Jr., & Knauss, K. S. (1988). Conceptual behavior in pigeons: Categorization of both familiar and novel examples from four classes of natural and artificial stim-


Received May 15, 1991
Final acceptance November 7, 1991
APPENDIX

Touch screens. Touch screens are frequently constructed as a bezel from a matrix of infrared light-emitting diodes (LEDs) and photosensors. The one used here (Carroll Touch, Infrared Smart Frame 5002380) was added to the monitor as a bezel. We built an aluminum base and mounting system to support the video monitor and bezel as a single integrated unit. Since this unit was constructed, complete units with integral built-in touch screens have become available. Carroll Touch integrates a touch screen with a Zenith 14-in. flat-screen monitor (#1492 at a current price of $1,260) with a resolution of 640 × 480. Flat screens have the advantage that they eliminate peripheral air gaps if protective glass plates are added to the monitor (e.g., Wright et al., 1988). A monitor with built-in touch screen was just released by IBM (Model 8516) with a resolution of 1,024 × 768. The touch screen is of the pressure-sensitive variety and records strength of the touch as well as location. According to IBM, these monitors can be used with the AT buss via an adapter. For more information on touch screens, contact IBM or Carroll Touch, P.O. Box 1309, Round Rock, Texas 78680; Tel. (512) 244-3500.

Image processing. The Howtek camera made digitizing 35-mm slides easy, but any videocamera with an RGB output should be able to digitize images. It is advisable, however, to obtain a demonstration of the image quality before purchase because image quality can vary considerably. For more information on the Howtek camera contact Howtek, 21 Park Ave., Hudson, New Hampshire 03051; Tel. (603) 882-5200.

The camera images were digitized by a TARGA-16 card in the present experiment because the necessary image-capture software was unavailable for the VISTA card at the time of set-up, but the VISTA card also has capture capability for image digitization. In comparing these two cards for research applications, the TARGA (now called TARGA PLUS, $1,795 to $2,495 for 512K to 2M of memory) is just a frame buffer (see Watson et al., 1986, for their use in visual research), whereas the VISTA ($2,995 to $4,795 for 1 to 4M of memory) is an intelligent processor. Unlike the TARGA, the VISTA can show lists of stimuli at presentation rates of 10 per second, its memory can be expanded to 14 megabytes, and it allows complete flexibility in allocating memory to different images. Either of these cards can be used in any AT style buss computers: 286, 386, or 486 machines. The faster speed and greater transfer rates of the 386 and 486 machines are not necessarily crucial in most applications because all the critical image manipulations are done on the hardware of the VISTA card, which has its own clock (e.g., 28 MHz display clock). For more information on these image-processing cards contact Truevision, 7340 Shadeland Station, Indianapolis, Indiana 46256, Tel. (317) 841-0332.

Digital picture storage. The storage capabilities for 200 pictures were adequate for the present experiment, but greater storage capabilities may be needed in other applications. One inexpensive way to increase image storage, and the one we currently use, is to store images on a streaming tape back-up. Images that are not currently being used are saved and loaded at some later time. A much more flexible storage system is one of the recently available, rewritable laser or CD disks. These laser disk storage systems typically store about 900 megabytes and allow disks to be removed and interchanged. Pictures to be used in each experimental session can be down-loaded to the computer hard disk from the laser disk, and then those pictures to be presented on an individual trial can be loaded into the RAM of the VISTA. For more information on erasable optical storage systems contact Storage Dimensions, 2145 Hamilton Ave., San Jose, California 95125, Tel. (408) 879-0300; MicroNet Technology, Inc., 20 Mason, Irvine, California 92718, Tel. (714) 837-6033; or Tecmar, 6225 Cochran Rd., Solon, Ohio 44139, Tel. (216) 349-0600.