ELSEVIER

Behavioural Processes 50 (2000) 79–94

BEHAVIOURAL PROCESSES

www.elsevier.com/locate/behavproc

Generalization of ‘same–different’ classification abilities in bottlenosed dolphins

Eduardo Mercado III *, Deirdre A. Killebrew, Adam A. Pack,
Inés V.B. Mácha [1], Louis M. Herman

Kewalo Basin Marine Mammal Laboratory, 1129 Ala Moana Boulevard, Honolulu, HI 96814, USA

Received 27 September 1999; received in revised form 18 April 2000; accepted 5 May 2000

Abstract

Two bottlenosed dolphins taught to classify pairs of three-dimensional objects as either same or different were tested with novel stimulus sets to determine how well their classification abilities would generalize. Both dolphins were immediately able to classify novel pairs of planar objects, differing only in shape, as same or different. When tested on sets of three objects consisting of either all different objects or of two identical objects and one different object, both dolphins proved to be able to classify ‘all different’ sets as different and ‘not all different’ sets as same, at levels significantly above chance. These data suggest that dolphins can use knowledge about similarity-based classification strategies gained from previous training to perform successfully in a variety of novel same–different classification tasks. Visual classificatory abilities of dolphins appear to be comparable to those that have been demonstrated in primates. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Animal; Cetacean; Cognition; Concept formation; Similarity

1. Introduction

Various tasks have been used to demonstrate that animals other than humans are able to develop abstract concepts and generalize rules (for review, see Roitblat and von Fersen, 1992; Roitblat et al., 1993; Thompson, 1995; Vauclair, 1996; Shettleworth, 1998). Tasks that test whether animals can classify stimulus sets as either ‘same’ or ‘different’ have increased in popularity over the last ten years (Oden et al., 1988, 1990; Thompson and Oden, 1993; Cook et al., 1995, 1997; Wasserman et al., 1995; Cook and Wixted, 1997; Young and Wasserman, 1997; Young et al., 1997a,b). In same–different tasks, all elements of a set are typically presented simultaneously. Subjects must perform one response if elements are similar (e.g. identical), and a different response if elements are

---

* Corresponding author. Present Address: Center for Molecular and Behavioral Neuroscience, 197 University Avenue, Rutgers University, Newark, NJ 07102, USA. Tel.: + 1-973-3531080; fax: + 1-973-3531272.

E-mail address: mercado@pavlov.rutgers.edu (E. Mercado III).

[1] Present Address: Institut für Verhaltensbiologie, Haderslebener Strasse 9, 12163 Berlin, Germany.

PII: S0376-6357(00)00100-5

dissimilar. The ability to classify stimulus sets as either same or different is considered to be a basic cognitive skill upon which many other cognitive abilities are based (Delius, 1994).

Most studies of same–different classification abilities have been conducted with either pigeons, chimpanzees, or humans. Few inferences can be safely drawn about general processes underlying judgments of similarity based on the abilities of only three species. By examining same–different classification abilities in a wide variety of species, one can gain broader perspective on the cognitive processes underlying these abilities. Towards this end, two bottlenosed dolphins (*Tursiops truncatus*) were recently trained to classify pairs of three-dimensional objects, presented in air, as either same or different (Herman et al., 1994). After training, these dolphins were tested on their ability to classify pairs of novel three-dimensional objects. One dolphin, Akeakamai, accurately classified novel pairs from the very first trial of testing, whereas a second dolphin, Hiapo, did not. These data demonstrate that a dolphin taught to classify familiar pairs of three-dimensional objects as same or different can accurately classify novel pairs of three-dimensional objects without additional training. The following two experiments were designed to explore whether the strategies the dolphins used to classify pairs of three-dimensional objects would generalize to novel same–different tasks.

In experiment 1, Akeakamai and Hiapo were tested with novel planar objects that differed primarily in terms of shape (i.e. other features such as surface area, material, and color were equated). This greatly reduced the number of visual and tactile cues that were useful for differentiating objects. Akeakamai was tested with planar objects presented underwater, introducing numerous novel cues. Most of these cues were not useful for discriminating ‘same’ pairs from ‘different’ pairs. If dolphins trained to make same–different judgments with pairs of three-dimensional objects can accurately classify novel pairs of planar shapes, then this would provide additional evidence that the dolphins had learned to classify object pairs using an abstract same–different concept. Such generalized abilities would also suggest that (1) shape information is sufficient for the dolphins to visually judge whether a pair of objects are similar or dissimilar, (2) dolphins use classification strategies underwater that are comparable to the ones they use visually in air, and (3) dolphins make same–different judgments based on comparisons of visual percepts (e.g. of shape or homogeneity) and/or on assessments of the overall similarity of stimulus set elements.

Experiment 2 required the dolphins to classify sets consisting of either three different objects, two identical objects and one different object, or three identical objects. For example, the dolphins were presented with the following objects sets (where letters denote objects and letter order reflects the positions of objects): A–B–C, A–A–B, B–A–A, A–B–A, and A–A–A. The dolphins were rewarded for classifying A–B–C sets as different, and all other sets as same. Note that the ‘correct’ response for sets of three objects is usually ambiguous; e.g. A–A–B, B–A–A, and A–B–A sets contain both same and different pairs. Thus, to choose accurately, the dolphins must recognize either that responding ‘same’ is appropriate when at least two of the objects are identical, or that responding ‘different’ is only appropriate when none of the objects are identical. Successful performance in this task would suggest that the dolphins can apply what they know about classifying object pairs as same or different to the classification of multi-object sets with graded levels of ‘sameness’.

2. Experiment 1

2.1. Method

2.1.1. Participants

The subjects in this study were two Atlantic bottlenosed dolphins: Hiapo, a 12-year-old male, and Akeakamai, a 20-year-old female. The dolphins were housed in two sea-water tanks, 15.2 m in diameter and 2 m deep, connected by a channel. Each dolphin was fed approximately one quarter of his/her daily ration of ~9 kg of smelt during each experimental session. Both dolphins had previous experience performing the same–different task (Herman et al., 1994), and a variety of other tasks involving similarity judgments (e.g. visual matching-to-sample and motor mimicry, Herman et al., 1993). The dolphins' training histories prior to this experiment were substantially different.

2.1.2. Materials

Two types of objects were used as visual stimuli: three-dimensional junk objects, and planar wooden objects (see Fig. 1). The three-dimensional objects differed in color, size, shape, and material. For example, stimulus sets included red plastic colanders, blue and orange foam footballs, white plastic bottles, and black and yellow rubber scuba fins. Both dolphins had experience with most of these objects in previous tests of same–different classification. In past tests, ∼80 different objects were used as samples. Twenty-three of these objects were chosen for use in the current experiment. Four planar objects were constructed from plywood such that they differed primarily in shape; i.e. they were painted the same color, and had approximately equal surface areas. The dolphins had no prior experience with these planar objects in any discrimination or classification tasks.

Because the introduction of completely novel objects during transfer tests can often disrupt performance (e.g. because of fear responses, D'Amato et al., 1985; Pack et al., 1991), both dolphins were repeatedly exposed to the planar objects in non-test situations. Objects were individually shown to the dolphins, both in water and in air, and the dolphins were instructed to touch the objects with various body parts. The dolphins were considered to be familiar with a particular planar object when they would touch it with any body part without hesitation.

2.1.3. Procedure

2.1.3.1. Training. Both dolphins initially participated in review sessions to verify that they could still correctly classify three-dimensional object pairs as either same or different. A trial began with a gestural instruction indicating that the dolphin should raise his/her head out of the wa-

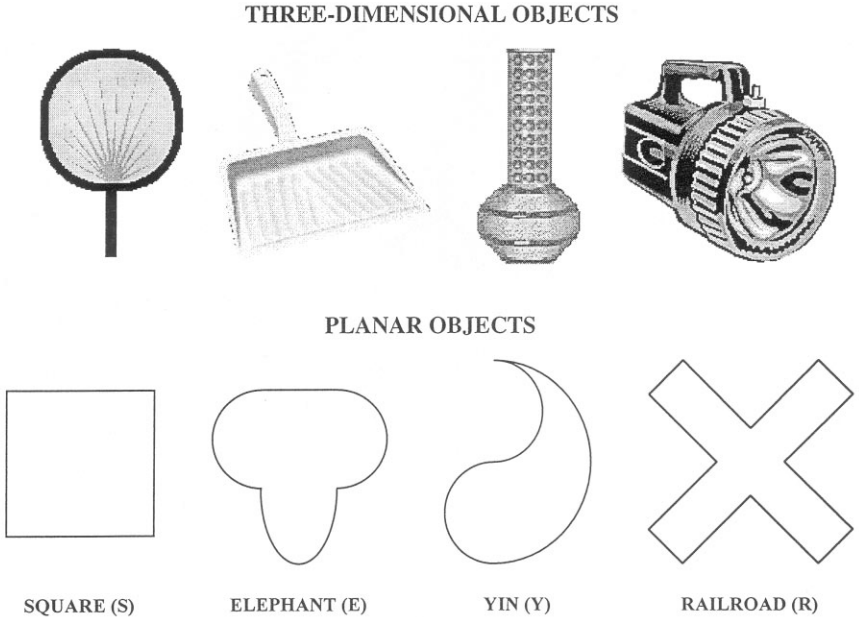

Fig. 1. The top row illustrates the kinds of objects used to train the dolphins to classify pairs of objects as same or different (note that the objects shown are not identical to those used, but are comparable). The bottom row shows the four planar shapes used to test whether the dolphins' same–different classification abilities would generalize to pairs of planar objects.

ter, ventral side facing the trainer. An experimenter (remotely located on a raised platform overlooking the tank) then signaled for the trainer and an assistant to present a pair of objects in front of the dolphin, in air. The dolphin was required to touch each object in the pair, after which the experimenter signaled for response paddles to be presented ~2 m to the right and left of the subject (for Hiapo, the sample objects were removed immediately before the paddles were presented because prior tests showed that this greatly improved his performance level). The dolphins were trained previously to press the right paddle when the objects were the same, and the left paddle when they were different. In the current tests, it was discovered that the subjects would often approach a paddle and remain in front of it, but refuse to press it. To accommodate this idiosyncratic behavior, the dolphins were considered to have chosen a response if either they pressed one of the paddles, or if they remained in front of a paddle continuously for three seconds. The dolphins' responses were judged by a 'blind' observer located near the experimenter. This observer viewed trials through an apparatus that occluded the trainer and sample stimuli, but not the response paddles. If a dolphin's response was correct, this fact was signaled to the trainer, and he or she blew a whistle. The dolphin then returned to the trainer and was rewarded with social reinforcement and a fish. When a response was incorrect, the trainer instructed the dolphin to return to the training station by splashing water, and the dolphin simply remained at the station until the next trial began.

Daily sessions typically consisted of 24 trials each. On half of these trials, a pair of identical objects was presented, and on the remaining trials, a pair of non-identical objects was presented. Trials were pseudo-randomly mixed such that (1) no more than four trials of a particular type (same or different) occurred consecutively, and (2) a particular sample object was not presented in more than two consecutive trials. A session was divided into two blocks of 12 trials each; each block of trials was separated by a minimum two minute interval. There was a minimum 30-s interval between trials.

A number of controls were used to prevent inadvertent cueing of the dolphins. The trainer and her assistant wore opaque goggles to prevent eye gaze cues, and so that they had no knowledge of whether the objects they were presenting were identical or non-identical. Objects were given to the trainer and her assistant by a tankside assistant according to the dictates of a written list. The tankside assistant was positioned such that she could not see the dolphins, and they could not see her. Two other tankside assistants, responsible for inserting the response paddles into the tank, were also goggled, so that no one at tankside could observe or influence the actions of the dolphins during trials. To insure that the participating dolphin was safely positioned during object presentation, and that the actions of other dolphins did not affect performance, the experimenter continuously observed the actions of dolphins during trials. It is unlikely, however, that the experimenter could provide any cues to guide dolphins' responses because he or she was located on a raised platform, several meters behind the dolphins.

After the dolphins demonstrated that they could still correctly perform the same–different discrimination, the presentation of stimuli was gradually shifted from in-air to underwater. Fig. 2 illustrates the differences between in-air and in-water presentation of objects. The objects used to train this transition from air to water were limited to three-dimensional objects that could be safely presented in sea-water without fear of introducing any foreign materials into the tank (e.g. compact plastic and rubber objects rather than bulky foam and cloth objects). This transition also involved some secondary training, peripheral to the discrimination task, to ensure that the dolphins positioned themselves sufficiently far away from the tank wall so that goggled presenters could safely introduce objects into the tank. Additionally, the criteria that the dolphins touch both objects before the response paddles were introduced had to be abandoned because it was extremely difficult for the experimenter to determine whether or not a dolphin was touching objects underwater. Instead, for all trials, the response paddles were introduced ~1 s after the dolphin submerged to inspect the objects.

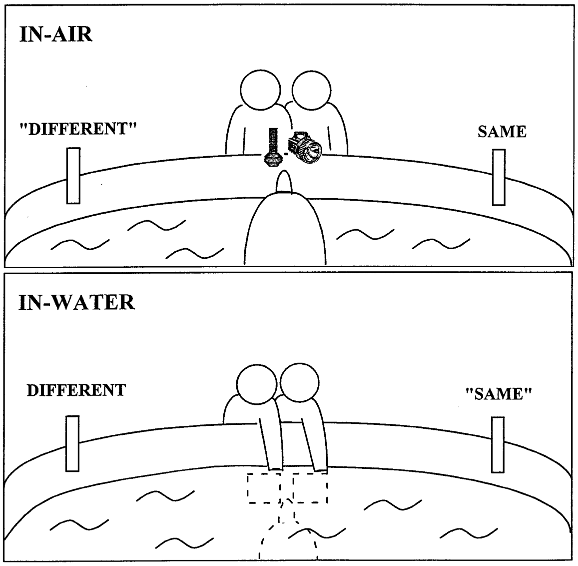

Fig. 2. The top figure shows the positions of the dolphin, trainers, objects, and response paddles for training trials with three-dimensional objects. The bottom figure shows these positions for novel underwater tests with planar objects.

2.1.3.2. Testing. Test sessions with planar objects were conducted once each subject was consistently responding correctly at an above chance level in the underwater classification task (criterion was 75% correct in two consecutive sessions). During test sessions, the dolphins' responses were differentially reinforced using the same reinforcement schedule and criteria for correct responses described above for review and training sessions (i.e. correct responses were rewarded and incorrect responses were not).

Because the pool of potential subjects was greatly constrained, and because there could only be one first exposure to a particular test, extensive efforts were made to develop tasks around the subjects, rather than vice versa. Consequently, test trials were conducted somewhat differently for each dolphin. Tests of Akeakamai were performed more tentatively and conservatively than they were with Hiapo (e.g. she received more extensive training with underwater classification), because it was not clear how she might react to the novel tests. Hiapo's training program was modified in an attempt to take advantage of lessons learned from tests of Akeakamai.

For Akeakamai, eight probe trials were initially inserted into a normal underwater training session (replacing eight 'standard' trials). These probes consisted of the four possible combinations of two of the novel planar objects (E:E, E:S, S:S, S:E, where E and S indicate objects shown in Fig. 1), presented twice. Next, she was tested in one session in which all 24 trials involved comparisons of these two planar objects presented underwater. Akeakamai was then tested on a second pair of novel planar objects (Y and R) in two sessions of 16 trials each. Finally, four full sessions (24 trials) and one half session (12 trials) were conducted in which the four planar objects were intermixed. These sessions included familiar pairings as well as novel cross-pairings of planar objects in 'different' trials (e.g. E:Y and S:R). Occasionally, a second dolphin in Akeakamai's tank would approach her station during testing. When this happened during a trial, the trial was aborted and repeated at a later point in the session.

Tests for Hiapo were conducted subsequently to Akeakamai, and were modified first based on Akeakamai's performance level with E:S pairs versus Y:R pairs, and later in an attempt to improve the accuracy of his classifications. Hiapo was initially tested in one 16-trial session in which all trials involved comparisons of two planar objects (Y and R) presented underwater. Next, he was given four in-air 'review' sessions (24 trials each) involving classification of familiar three-dimensional object pairs. He was then tested on his ability to classify Y–R pairs, presented in air, in 8 probe trials randomly interspersed among 16 standard trials. Subsequent tests with planar objects were conducted with pairs presented in-air. Hiapo's ability to classify Y–R pairs was further tested in three 24-trial sessions. Finally, he was tested on his ability to classify E–S pairs (as well as cross-pairings of planar objects such as R–S and Y–E) in two 24-trial sessions consisting of both familiar and novel (ten of 24 trials) planar pairs.

2.1.3.3. Data analysis. The most conservative measure of generalization in a same–different task is the number of correct responses made the very first time a subject experiences novel pairs of objects (independent of their spatial positions). Tests with novel spatial arrangements of objects can also be used to assess the transfer of classification abilities because spatial organization can affect how the similarity of objects is judged (Cook et al., 1997).

Performance in early sessions can also be used to measure generalization. However, because the dolphins' responses on all test trials in experiment 1 were differentially reinforced, and because each response involved a binary decision (choose same or choose different), the dolphins could potentially have learned how to respond to any given pair of objects after only a single exposure to that pair. For example, the first time the E–E pair was presented to Akeakamai, she might have randomly chosen to press the paddle indicating 'same'. This response would be reinforced, increasing the probability that she would choose same in response to subsequent presentations of the E–E pair. If Akeakamai initially responded to the E–E pair by pressing the paddle indicating 'different', then she would not be reinforced, providing her with sufficient information to choose the correct response the next time an E–E pair was presented. Because correct choices made after the first exposure to any given pair can potentially be explained in terms of stimulus-specific associative learning, choice accuracy in early sessions involving repeated exposures to planar pairs provides a less conservative (but still informative) measure of generalization.

To assess how well the dolphins' classification abilities generalized to pairs of planar objects, we calculated their choice accuracy for (1) completely novel pairs, (2) novel arrangements (i.e. novel pairs and novel arrangements of previously experienced pairs), (3) pairs presented in the first test session, and (4) pairs presented in all test sessions. We calculated the probability that the dolphins' choice accuracy reflected random responses using the cumulative binomial test (probability of choosing the correct response by chance was 0.5).

2.2. Results

Both Akeakamai and Hiapo did well on review trials, and during training sessions in which objects were gradually moved from in-air to underwater presentation. Akeakamai correctly classified 363 of 409 (89%, $P < 0.01$) object pairs as either same or different in training/review underwater sessions, and Hiapo correctly classified 146 of 172 (85%, $P < 0.01$) object pairs. Hiapo's choice accuracy was comparable for same (75 of 86 trials correct, 87%) versus different trials (71 of 86 correct, 83%), whereas Akeakamai was better at classifying different pairs (198 of 205 correct, 97%) than same pairs (165 of 204 correct, 81%).

Akeakamai correctly classified all four novel pairs of identical planar objects (E–E, S–S, Y–Y, R–R) and all six novel pairs of different planar objects (E–S, E–Y, E–R, Y–S, Y–R, S–R) the first time she was exposed to them (100%, $P < 0.01$); she also correctly classified five of the six spatially-reversed pairs of different objects (S–E, Y–E, R–E, etc.) on first exposure. Her overall performance on trials involving planar pairs (130 of 170 trials correct, 76%, $P < 0.01$) was less impressive, but still better than chance. In her first test session, in which test trials were introduced as probes, Akeakamai correctly classified five of eight object pairs. Akeakamai's overall choice accuracy was comparable for same (63 of 85 trials correct, 74%) versus different pairs (67 of 85 trials correct, 79%). It is important to note that during several test sessions, disruption of trials by a second dolphin degraded Akeakamai's performance (e.g. she was 89% correct in uninterrupted sessions, but only 67% correct in sessions where multiple interruptions occurred).

Hiapo did not perform as well as Akeakamai on first trial exposures to novel pairs of planar objects, correctly classifying three of four same pairs and five of six different pairs (80%, $P = 0.05$). Of these, he correctly classified six of seven pairs (86%, $P > 0.05$) that were first exposed to him in air. When responses to novel arrangements of different pairs are added to this in-air total, Hiapo's choice accuracy (nine of ten pairs correct, $P < 0.05$) is better than chance. Like Akeakamai, Hiapo's overall performance level (78 of 104 trials correct, 75%, $P < 0.01$) was lower than on first trial exposures. In his first test session, in which objects were presented underwater, Hiapo only correctly classified eight of 16 object pairs, no better than would be expected by chance. His overall choice accuracy was slightly better for same pairs (42 of 52 trials correct, 81%) than for different pairs (36 of 52 trials correct, 70%).

2.3. Discussion

Akeakamai's performance was errorless when she was presented with completely novel pairs of planar objects underwater. Her ability to classify three-dimensional object pairs presented in air as same or different clearly generalized to pairs of planar objects presented underwater. Neither the decreased number of features differentiating pairs, nor the increased number of irrelevant features introduced in the underwater context, appeared to affect the accuracy of Akeakamai's responses. Although Hiapo's performance was less impressive, his performance with novel arrangements presented in air, and across test sessions, was better than chance. Thus, both dolphins proved to be able to accurately classify pairs of planar objects as same or different based on their prior experience with three-dimensional objects.

Because the dolphins were differentially reinforced during test sessions, they conceivably could have rapidly learned the new classification task rather than generalizing from previous experience. It is because of this possibility that we tested the dolphins with only four planar objects. We were interested in whether the dolphins could generalize skills learned from training with three-dimensional objects, not whether they could generalize based on experience with planar objects. The possibility that the dolphins rapidly learned correct responses to planar objects seems unlikely given that overall performance for both dolphins was worse than performance on first trial exposures. If the dolphins were learning to make correct responses during test trials, then their accuracy should have increased as planar object pairs were repeated, not decreased.

One possible explanation for why the dolphins' choice accuracy decreased after initial exposures is that testing with a small number of highly similar objects may have resulted in high levels of interference across trials. Another possibility is that the planar shapes were more difficult for the dolphins to discriminate than three-dimensional objects, causing the dolphins to become fatigued or disinterested more quickly.

The dolphins' accuracy in first sessions was lower than their overall accuracy. These differences are difficult to interpret because of several confounding variables. First, the novelty of the testing situation may have distracted the dolphins. For example, Hiapo became increasingly fixated on one of the objects (Y) during his first test session. It was for this reason that Hiapo's subsequent tests with planar pairs were conducted with objects presented in air. Second, several trials in Akeakamai's first session were disrupted by a second dolphin that rushed towards her position several times when planar pairs were presented underwater. Finally, the introduction of novel irrelevant cues associated with object pairs may have affected Hiapo's responses. None of these possibilities accounts for why the dolphins were most accurate when they were asked to classify completely novel pairs of planar objects.

Previous studies of underwater matching-to-sample found that dolphins' choice accuracy increased when objects were presented underwater (Harley et al., 1996). Presenting objects underwater introduces numerous novel visual and echoic cues. Most of these cues are not useful for discriminating pairs of identical planar objects from non-identical pairs. We did not determine whether the dolphins used echolocation to inspect pairs of planar objects presented underwater. The role of echoic cues in dolphins' classification of pairs of planar shapes, if any, awaits further study.

The findings of experiment 1 are consistent with prior results showing that dolphins can acquire a generalized ability to visually match objects on the basis of shape features alone (Hunter, 1988; Pack and Herman, 1995). In a previous study, one dolphin (Phoenix) trained to match planar visual shapes proved to be able to match novel planar shapes (Forestell and Herman, 1988; Hunter, 1988). A second dolphin (Elele) trained to visually and echoically match three-dimensional objects differing only in shape, was also immediately able to match novel three-dimensional shapes (Herman et al. 1994; Pack and Herman, 1995). In the current experiment, two dolphins taught to make same–different judgments with pairs of three-dimensional objects differing in color, size, shape, and material, proved to be able to immediately classify novel pairs of planar shapes as same or different.

3. Experiment 2

3.1. Method

3.1.1. Participants

The two dolphins from experiment 1 served as subjects in this experiment.

3.1.2. Materials

Akeakamai was tested with six familiar three-dimensional objects and the four planar objects (E, R, S, Y) used in experiment 1, as well as four novel three-dimensional objects. Hiapo was tested using sixteen familiar three-dimensional objects and four novel three-dimensional objects. Novel objects were individually shown to the dolphins in non-test situations and the dolphins were instructed to touch the objects with various body parts. The dolphins were considered to be habituated to a particular novel object when they would touch it with any body part without hesitation.

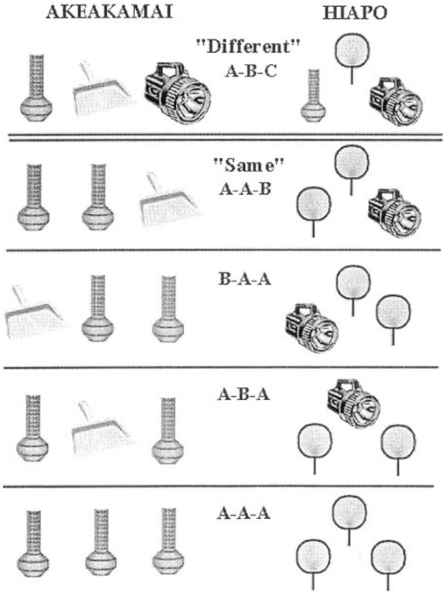

Fig. 3. Differences in the spatial organization of three-object sets presented to Akeakamai and Hiapo. Tables 2 and 3 summarize the dolphins' choice accuracy for each of these configurations.

3.1.3. Procedure

3.1.3.1. Training. The dolphins received no additional training (other than review sessions with pairs of familiar objects) on the classification of objects as either same or different prior to these tests. Hiapo was exposed in four sessions to object pairs in which objects were placed at various angles relative to one another (e.g. two objects might be presented with one directly above the other, or with one diagonally above the other) to familiarize him with the novel object arrangements he would be exposed to during testing. The addition of a third object necessitated the addition of a second goggled assistant to present that object, for a total of three presenters in front of the dolphin. Before testing, both dolphins experienced practice trials involving three presenters; only two objects were presented to the dolphins during these trials.

3.1.3.2. Testing. Testing was conducted in three stages. First, the dolphins were tested on their ability to classify sets of three familiar objects (in air) as either same or different. Objects were presented in different spatial arrangements for each dolphin (see Fig. 3). For Akeakamai, three objects were presented in a horizontal linear arrangement, whereas for Hiapo, three objects were presented in a triangular configuration, with one object on top and two on bottom. These different presentation modes were used to preliminarily assess whether the spatial organization of objects might affect the dolphins' responses in this novel task. The dolphins were reinforced for choosing the 'same' paddle when at least two of the three objects matched (A–A–B, B–A–A, A–B–A, or A–A–A patterns; see Fig. 3), and were reinforced for choosing the 'different' paddle when all objects were different (A–B–C patterns).

Akeakamai was tested using six familiar objects in six sessions of 24 trials each. A session was typically divided into two blocks of 12 trials, separated by a minimum two minute interval, with a minimum 30-s interval between trials. The 24 trials within each session were distributed as follows: A–B–C pattern = 12 trials, A–A–B pattern = five trials, B–A–A pattern = five trials, A–B–A pattern = two trials. Trials involving A–A–A patterns were not included in Akeakamai's initial test sessions to heighten the contrast between 'all-different' and 'not-all-different' sets of objects. Hiapo was tested in 23 sessions of between ten and 30 trials each; the distribution of trial types was comparable to that used with Akeakamai. Correction trials and cued trials were used intermittently within sessions in an attempt to improve Hiapo's performance. For example, the incorrect paddle might not be presented during a trial or might be removed before Hiapo could select it, forcing him to choose the correct paddle.

After initial tests with sets of three familiar three-dimensional objects, both dolphins were tested on their ability to classify sets of three novel three-dimensional objects (in air) as either same or different. The only differences in procedure between these tests and previous tests with three objects were a reduction in the number of trials per session, and a modification of the distribution of 'same' trial types. Akeakamai was tested with novel three-dimensional objects in four sessions. Each session consisted of 20 test trials distributed as follows: A–B–C pattern = ten trials, A–A–B pattern = three trials, B–A–A pattern = three trials, A–B–A pattern = three trials, A–A–A pattern = one trial. Hiapo's tests with novel objects were slightly different from Akeakamai's. He was tested with different novel objects from those used with Akeakamai, and objects were presented to him in a triangular arrangement. Additionally, Hiapo was tested in eight sessions of 24 trials each. These sessions consisted of 10 novel test trials interspersed among 14 trials with familiar three object sets.

Finally, Akeakamai was tested on her ability to classify sets of three planar objects, presented underwater, in four sessions of 18 trials each (the one trial involving A–A–A patterns was eliminated, and one A–B–C trial was removed to balance the number of same versus different trial types). Hiapo was not tested on classification of three planar objects presented underwater.

The sequence of trial types and object arrangements within and across test sessions was randomized using a computerized random number generator.

3.1.3.3. Data analysis. The most conservative measure of generalized classification abilities in this novel same–different task is choice accuracy for completely novel sets (combinations) of three objects. This measure does not take into account the spatial arrangement of objects; e.g. A–A–B, B–A–A, and A–B–A patterns are considered to be equivalent object sets, and A–B–C, B–C–A, C–B–A, A–C–B, B–A–C, and C–A–B patterns are considered to be equivalent sets of different objects. If novel spatial arrangements of objects are considered to be novel stimuli, then every trial tested in experiment 2 qualifies as a novel test trial because the spatial arrangement of any given set of three objects was never repeated.

First session performance with novel stimulus sets and overall accuracy were also analyzed to assess how well the dolphins' classification abilities generalized. As in experiment 1, the probability that the dolphins' choices were random was calculated using the cumulative binomial test (probability of a random response being correct was 0.5).

3.2. Results

Neither dolphin classified completely novel sets of three objects with an accuracy significantly above chance. Akeakamai correctly classified 30 of 50 (60%, $P = 0.10$) novel sets composed of three familiar objects, 12 of 20 (60%, $P = 0.25$) novel sets composed of three novel objects, and ten of 16 (63%, $P = 0.23$) novel sets composed of three familiar planar objects. She was best (80% correct) at classifying novel sets composed of three different familiar objects (see Table 1). Akeakamai classified novel sets that included two matching objects and one different object as same in 28 of 54 (52%) test trials, and as different in 26 of 54 (48%) trials. Hiapo correctly classified six of 20 (30%, $P = 0.98$) novel sets composed of three

Table 1
Choice accuracy on trials with completely novel combinations of three objects

Session type	Akeakamai		Hiapo	
	Trial type		Trial type	
	SAME	DIFFERENT	SAME	DIFFERENT
Familiar objects	14/30 (47%)	16/20 (80%)	4/16 (25%)	2/4 (50%)
Novel objects	9/16 (56%)	3/4 (75%)	11/16 (69%)	2/4 (50%)
Planar objects	9/12 (75%)	1/4 (25%)	–	–

Table 2
Choice accuracy on 'same' and 'different' trials with novel arrangements of three objects

Session type	Akeakamai		Hiapo	
	Trial type		Trial type	
	SAME	DIFFERENT	SAME	DIFFERENT
Familiar objects	38/72 (53%)	52/72 (72%)	134/224 (60%)	135/205 (66%)
Novel objects	22/40 (55%)	24/40 (60%)	29/40 (73%)	21/40 (53%)
Planar objects	28/36 (78%)	15/36 (42%)	–	–

familiar objects and 13 of 20 (65%, $P = 0.13$) novel sets composed of three novel objects. He inaccurately classified novel sets independently of whether the objects within those sets did or did not match (see Table 1).

Both dolphins classified novel arrangements of sets of three familiar objects at a level significantly above chance. Akeakamai classified 90 of 144 (63%, $P < 0.01$) arrangements of three familiar objects correctly, and Hiapo classified 269 of 429 (63%, $P < 0.01$) arrangements correctly (note that cued trials are counted as incorrect responses in Hiapo's data to preclude bias). Akeakamai failed to accurately classify either novel arrangements of three novel three-dimensional objects presented in air (46 of 80 correct, 58%, $P = 0.11$) or novel arrangements of three planar objects presented underwater (43 of 72 correct, 60%, $P = 0.06$). In contrast, Hiapo was able to classify novel arrangements of three novel objects presented in air at a level significantly above chance (50 of 80 correct, 63%, $P < 0.05$). His accuracy at classifying familiar arrangements of three familiar objects in sessions involving tests with novel objects was comparable to his prior performance (76 of 116 correct, 66%).

Table 2 describes Hiapo and Akeakamai's accuracy at classifying 'same' and 'different' arrangements for the different test conditions. Akeakamai and Hiapo both classified arrangements of three different familiar objects more accurately than arrangements in which some objects were identical. However, Hiapo classified arrangements that included identical novel objects more accurately than arrangements of all different novel objects, and Akeakamai classified arrangements that included identical planar objects more accurately than arrangements of all different planar objects. Table 3 provides details about the dolphins' performance levels for the various arrangements of 'same' object sets. Akeakamai classified 'same' arrangements most accurately when all three objects were identical. Both dolphins' choice accuracy on other types of 'same' arrangements (i.e., A–A–B, B–A–A, A–B–A) varied as a function of session type and individual.

Both dolphins' performance levels were relatively stable across test sessions (see Figs. 4 and

Table 3
Choice accuracy on 'same' trials as a function of the arrangement of objects

Session type	Akeakamai	Hiapo
Familiar objects		
A–A–B	16/30 (53%)	42/67 (63%)
B–A–A	18/30 (60%)	36/64 (56%)
A–B–A	4/12 (33%)	42/73 (58%)
A–A–A	–	14/20 (70%)
Novel objects		
A–A–B	4/12 (33%)	7/12 (58%)
B–A–A	6/12 (50%)	11/12 (92%)
A–B–A	8/12 (66%)	9/12 (75%)
A–A–A	4/4 (100%)	2/4 (50%)
Planar objects		
A–A–B	9/12 (75%)	–
B–A–A	9/12 (75%)	–
A–B–A	10/12 (83%)	–

5). Akeakamai's accuracy in her first test session with sets of three objects (75%) was comparable to her accuracy in preceding tests with planar pairs (76%). Additionally, Fig. 4 shows that her first session performance with sets of three novel objects (65%) was comparable to her most recent performance with sets of three familiar objects (63%), and her first session performance with sets of three planar objects (67%) was comparable with her most recent performance with sets of three novel objects (65%).

Hiapo's accuracy in his first test session with sets of three objects (45%) was much lower than in preceding tests with planar objects (75%). By the twelfth session, however, his accuracy was consistently near 70% correct. He maintained a similar accuracy level (67%) in his first session with sets of three novel objects (see Fig. 5).

3.3. Discussion

Experiment 2 was conducted primarily to determine how dolphins trained to classify pairs of objects as either 'same' or 'different' would react when required to classify larger sets of objects with more complex inter-object similarity relationships. Neither dolphin immediately classified novel three object sets more accurately than would be expected on the basis of chance. Both dolphins were, however, able to classify novel arrangements of three objects with an accuracy significantly above chance. Performance with

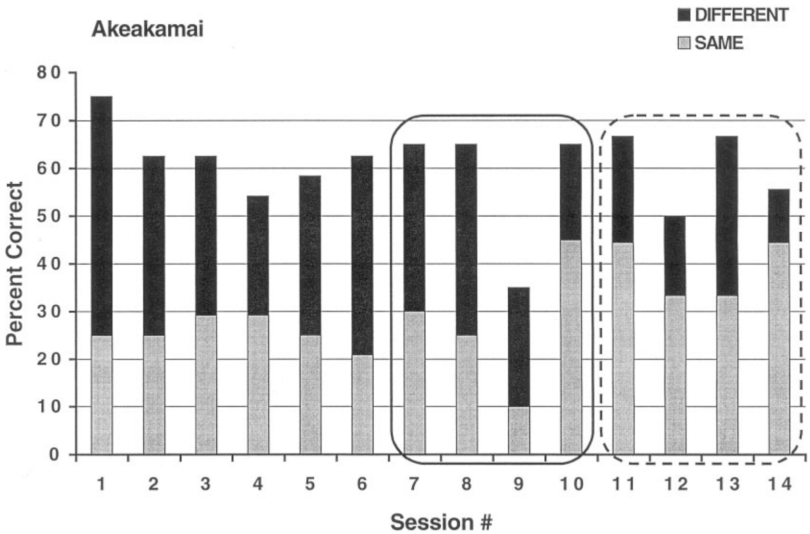

Fig. 4. Akeakamai's choice accuracy in sessions involving classification of three-object sets. Black bars show the percentage of correct 'different' responses, and gray bars show the percentage of correct 'same' responses. Sessions 1–6 are tests with familiar objects, sessions 7–10 (bounded by solid lines) are tests with novel objects, and sessions 11–14 (bounded by dotted lines) are underwater tests with planar objects.

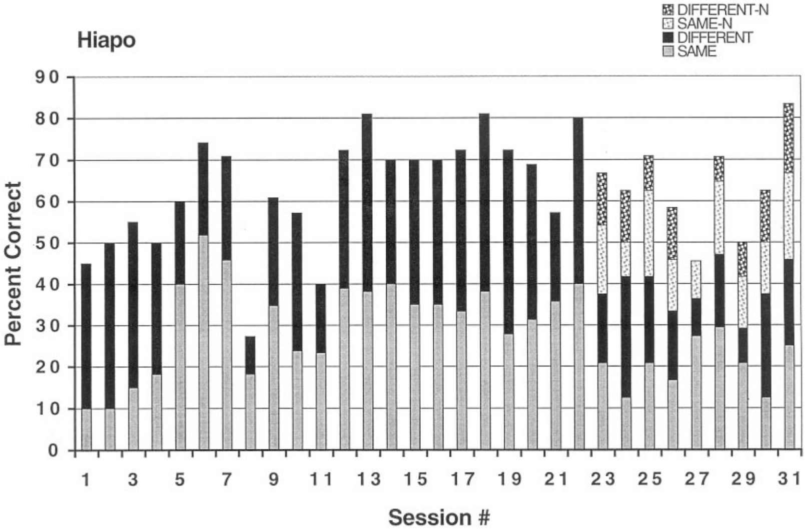

Fig. 5. Hiapo's choice accuracy in sessions involving classification of three-object sets. Black bars show the percentage of correct 'different' responses, and gray bars show the percentage of correct 'same' responses. Sessions 1–22 are tests with familiar objects and sessions 23–31 are tests with both novel and familiar objects. (DIFFERENT − N = percentage of correct 'different' responses in tests with novel objects; SAME − N = percentage of correct 'Same' responses in tests with novel objects.)

three object sets was consistently lower than in previous tests with pairs of objects.

Because the dolphins had no prior experience classifying sets that contained both identity and non-identity relationships between objects, it is not surprising that they did not classify sets of three objects as accurately as they classified sets of two objects. The more interesting result is that, despite the ambiguous nature of the 'not-all-different' sets, the dolphins were able to correctly classify a significant number of three object sets. There are two possible explanations for how the dolphins were able to do this. Either the dolphins were able to choose correctly based on their previous training, or they rapidly learned how to respond based on their experience with sets of three objects in test sessions.

The available evidence suggests that Akeakamai was not learning how to respond during test sessions. Her choice accuracy in first test sessions with novel sets of three objects was consistently as good as, or better than, her accuracy in later test sessions. Similarly, her accuracy with completely novel sets of objects was comparable to her accuracy with novel arrangements of objects. Akeakamai's choice accuracy showed no decrement after transitions from (1) sessions with two objects to tests with three objects, (2) tests with three familiar objects to tests with three novel objects, or (3) tests with three novel objects to tests with three planar objects. Thus, the preponderance of the evidence suggests that Akeakamai was classifying sets of three objects based on previously learned same–different classification strategies. Anecdotally, no behavioral differences related to the introduction of three-object sets were noted that would indicate that Akeakamai was confused by the increased number of objects (e.g. neither inspection times nor response times showed noticeable increases in duration).

It is more likely that Hiapo learned how to respond during tests because he was actively trained during test sessions using both correction trials and cued trials. Active efforts to train Hiapo were initiated after he proved to be unable to accurately classify sets of three objects in his first four test sessions (see Fig. 5). Training was not extensive; only 39 of 462 trials (8%) were either corrected or cued. Nevertheless, by the twelfth test session, Hiapo's choice accuracy had increased from 50 to 72%. Hiapo's performance does not provide strong evidence that he used previously-learned same–different classification strategies to classify sets of three objects. Although it is possible that prior training allowed Hiapo to learn the three-object classification task more rapidly than normal, the currently available data cannot address this possibility. Hiapo's performance does suggest that the strategies he used to classify sets of three objects were abstract enough to allow him to successfully apply those strategies to completely novel three-object sets.

It is not clear which aspects of Akeakamai's previous training allowed her to correctly classify sets of three objects. Her previous training in matching-to-sample tasks might have prepared her for making comparisons between sets of three objects. However, Hiapo was also skilled at matching-to-sample, but did poorly in initial test sessions with three-object sets. Akeakamai potentially could have inspected two randomly selected objects and then based her response on whether the two objects were identical or non-identical. This explanation would account for why she classified sets of three different three-dimensional objects more accurately than sets containing two identical objects, but not for why her accuracy dropped from 97% with pairs of different familiar objects to 72% with sets of three different familiar objects. This explanation also does not account for why she classified not-all-different sets of planar objects more accurately than all-different sets (recall that in experiment 1, Akeakamai's accuracy was comparable for same and different pairs of planar objects).

Akeakamai's errors can provide hints as to the strategies she used to classify sets of three objects. For example, she was just as likely to classify not-all-different sets as 'same' as she was to classify them as 'different', whereas she was much more likely to classify all-different sets as different. This suggests that she was sensitive to differences between the two types of object sets. The errors made by Akeakamai are inconsistent with random comparison of two out of three objects within a set. For example, given any set of three objects, A, A, and B, there are two ways to choose non-identical pairs, but only one way to choose an identical pair, predicting twice as many 'choose different' responses.

Akeakamai's greater success in classifying 'all-different' sets versus 'not-all-different' sets composed of familiar objects might seem to suggest that she based her decisions on dissimilarities between objects rather than on similarities. However, she was errorless at classifying sets consisting of three identical objects (see Table 3), and she classified 'not-all-different' sets more accurately in tests with planar objects, suggesting that similarity was also a salient feature. Further tests are needed to determine the relative salience of similarity versus dissimilarity for dolphins. Only a small number of tests of sets involving all-identical object were included in the current experiment because tests were designed to investigate responses to two classes of three objects sets (all-different sets and not-all-different sets) in which all sets involved novel arrangements. All-identical sets represent a special case in which objects cannot be reorganized to form novel arrangements. To include such sets without repetition would require a large number of different objects (e.g. 148 objects would be required to replicate the tests of Akeakamai conducted in experiment 2, if not-all-different sets were replaced with all-identical sets). It would be interesting to see, in future studies, whether Akeakamai or Hiapo's accuracy at classifying not-all-different sets would change when such sets are contrasted with all-identical sets.

The results of experiment 2 suggest that the same–different classification strategies learned by the dolphins are probably not limited to pairwise comparisons used to determine whether objects within a stimulus set are identical. If the dolphins had previously learned to classify pairs of objects based on the presence or absence of an identity (or non-identity) relationship between the objects, then they should have been biased towards classifying A–A–B, B–A–A, and A–B–A sets as different (i.e. not identical) because of the presence of non-identical object pairs. The fact that both dolphins classified sets of three non-identical objects less accurately than pairs of non-identical objects suggests that the classifications made by the dolphins were probably based on more holistic judgments of inter-object similarities or the relative homogeneity of features within object sets.

4. General discussion

The results of these experiments suggest that dolphins can use knowledge about similarity-based classification strategies gained from previous training to correctly classify novel stimuli in novel situations. Experiment 1 showed that both dolphins were able to classify novel pairs of planar objects without additional training, and that one dolphin, Akeakamai, was able to do so when objects were presented underwater. Experiment 2 showed that both dolphins could classify sets composed of three different objects as ‘different’ and sets composed of two identical objects and one different object as ‘same’; Akeakamai required no additional training to perform this novel task and Hiapo required only minimal training. Such capabilities have only previously been demonstrated in chimpanzees, orangutans, and humans (Robinson, 1960; King, 1973).

Comparative studies have shown that certain species learn same–different and oddity tasks much more quickly than others (Bernstein, 1961; Strong and Hedges, 1966; King, 1973; Thomas and Boyd, 1973; Wilson et al., 1985; Chausseil, 1991). More recent research has also revealed that some species (e.g. humans, chimpanzees, and dolphins) appear to be naturally attentive to similarity relationships (Oden et al., 1988, 1990; Medin et al., 1990; Tyrrell et al., 1991; Herman et al., 1993). Species have also shown varying abilities to learn particular cognitive tasks involving similarity judgments. Pigeons, for example, rapidly learn successive same–different discriminations with pairs of objects (Delius, 1994; Macphail et al., 1995; Young et al., 1997a), but have great difficulty learning simultaneous same–different tasks with stimulus sets composed of less than about eight items (Edwards et al., 1983; Santiago and Wright, 1984; Cook et al., 1995; Wasserman et al., 1995).

There is also evidence of striking parallels in the way different species judge similarity (Sands and Wright, 1980; Premack, 1983a,b; Wright et al., 1983, 1984). For example, monkeys appear to veridically represent the parametric similarity of three–dimensional objects in ways that are directly comparable to human representations (Sugihara et al., 1998). The classificatory abilities of dolphins seem more comparable to those of primates than to those of pigeons because dolphins have no problem classifying stimulus sets composed of two items and because they are able to apply learned classification strategies to a wide range of situations.

The exploratory nature of these experiments and the special limitations of dolphin research made it difficult to develop a testing situation that would allow for detailed comparisons of treatment effects. Thus, it is impossible to say with certainty what stimulus features the dolphins attended to during any given trial, or what specific factors determined the choices they made. Additionally, differences in the dolphins’ training histories make interpreting individual differences in their classification abilities difficult. Whether the dolphins used similar strategies to classify stimulus sets in these experiments cannot be ascertained from the current data. Additional research is warranted.

Studies of dolphins’ cognitive abilities provide a useful counterpoint to studies of cognition in primates. The evolutionary history, ecology, and neurology of dolphins are so different from those of primates, that demonstrations of similar cognitive capacities across these species are quite surprising. Such parallels suggest that mammalian brains analyze environmental features using ‘basic’ processes that are evolutionarily old and, therefore, likely common across a wide range of mammalian species.

Acknowledgements

This research was supported by a grant from Earthwatch. These experiments would not have been possible without the help of numerous interns, volunteers, and graduate students. We thank the following people who helped with the training involved in this experiment: Faby Art, Mark Deakos, Kitty Hudson, Perry Ragoobar, Billy Rickards, Carrie Southgate, and Erika Weber. We also thank Dennis Schork for his help in the design and construction of test objects. Patricia Murray, Ute Schroeder, Oliver Pronge, Lisbet

Kugler, Heather Brandon, Yasmin Tajik, Sandra Helmer, and Kimberly LeRoy provided extensive assistance in preparing and conducting these experiments, and in organizing and analyzing the collected data. We especially thank Toni Cimino for her help with the training and testing of Hiapo, and Amy Wood for initially training both dolphins to perform the visual same–different task. Finally, we thank Hiapo and Akeakamai for their cooperation in these experiments.

References

Bernstein, I.S., 1961. The utilization of visual cues in dimension-abstracted oddity by primates. J. Comp. Physiol. Psychol. 54, 243–247.

Chausseil, M., 1991. Visual same–different learning, and transfer of sameness concept by coatis (Nasua Storr, 1780). Ethology 87, 28–36.

Cook, R.G., Wixted, J.T., 1997. Same–different texture discrimination in pigeons: testing competing models of discrimination and stimulus integration. J. Exp. Psychol. Anim. Behav. Process. 23, 401–416.

Cook, R.G., Cavoto, K.K., Cavoto, B.R., 1995. Same–different texture discrimination and concept learning in pigeons. J. Exp. Psychol. Anim. Behav. Process. 21, 253–260.

Cook, R.G., Katz, J.S., Cavoto, B.R., 1997. Pigeon same–different concept learning with multiple stimulus classes. J. Exp. Psychol. Anim. Behav. Process. 23, 417–433.

D'Amato, M.R., Salmon, D.P., Columbo, M., 1985. Extent and limits of the matching concept in monkeys (*Cebus apella*). J. Exp. Psychol. Anim. Behav. Process. 11, 35–51.

Delius, J.D., 1994. Comparative cognition of identity. In: Bertleson, P., Eelen, P., d'Ydewalle, G. (Eds.), International Perspectives on Psychological Science. Erlbaum, Hillsdale, NJ, pp. 25–40.

Edwards, C.A., Jagielo, J.A., Zentall, T.R., 1983. 'Same/different' symbol use by pigeons. Anim. Learn. Behav. 11, 349–355.

Forestell, P.H., Herman, L.M., 1988. Delayed matching of visual materials by a bottlenosed dolphin aided by auditory symbols. Anim. Learn. Behav. 16, 137–146.

Harley, H.E., Roitblat, H.L., Nachtigall, P.E., 1996. Object representation in the bottlenosed dolphin (*Tursiops truncatus*): integration of visual and echoic information. J. Exp. Psychol. Anim. Behav. Process. 22, 164–174.

Herman, L.M., Pack, A.A., Morrel-Samuels, P., 1993. Representational and conceptual skills of dolphins. In: Roitblat, H.L., Herman, L.M., Nachtigall, P.E. (Eds.), Language and Communication: Comparative Perspectives. Erlbaum, Hillsdale, NJ, pp. 403–442.

Herman, L.M., Pack, A.A., Wood, A.M., 1994. Bottlenose dolphins can generalize rules and develop abstract concepts. Mar. Mam. Sci. 10, 70–80.

Hunter, G.A., 1988. Visual delayed matching of two-dimensional forms by a bottlenosed dolphin. Unpublished Master's Thesis, University of Hawaii, Honolulu.

King, J.E., 1973. Learning and generalization of a two-dimensional sameness–difference concept by chimpanzees and orangutans. J. Comp. Physiol. Psychol. 84, 140–148.

Macphail, E.M., Good, M., Honey, R.C., Willis, A., 1995. Relational learning in pigeons: the role of perceptual processes in between-key recognition of complex stimuli. Anim. Learn. Behav. 23, 83–92.

Medin, D.L., Goldstone, R.L., Gentner, D., 1990. Similarity involving attributes and relations: judgments of similarity and difference are not inverses. Psychol. Sci. 1, 64–69.

Oden, D.L., Thompson, R.K.R., Premack, D., 1988. Spontaneous transfer of matching by infant chimpanzees (*Pan troglodytes*). J. Exp. Psychol. Anim. Behav. Process. 14, 140–145.

Oden, D.L., Thompson, R.K.R., Premack, D., 1990. Infant chimpanzees (*Pan troglodytes*) spontaneously perceive both concrete and abstract same/different relations. Child Dev. 61, 621–631.

Pack, A.A., Herman, L.M., 1995. Sensory integration in the bottlenosed dolphin: immediate recognition of complex shapes across the senses of echolocation and vision. J. Acoust. Soc. Am. 98, 722–733.

Pack, A.A., Herman, L.M., Roitblat, H.L., 1991. Generalization of visual matching and delayed matching by a California sea lion (*Zalophus californianus*). Anim. Learn. Behav. 19, 37–48.

Premack, D., 1983. Animal cognition. Ann. Rev. Psychol. 34, 351–362.

Premack, D., 1983. The codes of man and beasts. Behav. Brain Sci. 6, 125–137.

Robinson, J.S., 1960. The conceptual basis of the chimpanzee's performance on the sameness–difference discrimination problem. J. Comp. Physiol. Psychol. 53, 368–370.

Roitblat, H.L., von Fersen, L., 1992. Comparative cognition: representations and processes in learning and memory. Ann. Rev. Psychol. 43, 671–710.

Roitblat, H.L., Herman, L.M., Nachtigall, P.E. (Eds.), 1993. Language and Communication: Comparative Perspectives. Erlbaum, Hillsdale, NJ.

Sands, S.F., Wright, A.A., 1980. Serial probe recognition performance by a rhesus monkey and a human with 10- and 20-item lists. J. Exp. Psychol. Anim. Behav. Process. 6, 386–396.

Santiago, H.C., Wright, A.A., 1984. Pigeon memory: same/different concepts in learning, serial probe recognition acquisition, and probe delay effects on the serial-position function. J. Exp. Psychol. Anim. Behav. Process. 10, 498–512.

Shettleworth, S.J., 1998. Cognition, Evolution, and Behavior. Oxford University Press, New York.

Strong, P.N. Jr, Hedges, M., 1966. Comparative studies in simple oddity learning: I. Cats, raccoons, monkeys, and chimpanzees. Psychon. Sci. 5, 13–14.

Sugihara, T., Edelman, S., Tanaka, K., 1998. Representation of objective similarity among three-dimensional shapes in the monkey. Biol. Cyber. 78, 1–7.

Thomas, R.K., Boyd, M.G., 1973. A comparison of *Cebus albifrons* and *Saimiri sciureus* on oddity performance. Anim. Learn. Behav. 1, 151–153.

Thompson, R.K.R., 1995. Natural and relational concepts in animals. In: Roitblat, H.L., Meyer, J.A. (Eds.), Comparative Approaches to Cognitive Science. MIT Press, Cambridge, MA, pp. 175–224.

Thompson, R.K.R., Oden, D.L., 1993. 'Language training' and its role in the expression of tacit propositional knowledge in chimpanzees (*Pan troglodytes*). In: Roitblat, H.L., Herman, L.M., Nachtigall, P.E. (Eds.), Language and Communication: Comparative Perspectives. Erlbaum, Hillsdale, NJ, pp. 365–384.

Tyrrell, D.J., Zingaro, M.C., Minard, K.L., 1991. Learning and transfer of identity-difference relationships by infants. Infant Behav. Dev. 16, 43–52.

Vauclair, J., 1996. Animal Cognition: An Introduction to Modern Comparative Psychology. Harvard University Press, Cambridge, MA.

Wasserman, E.A., Hugart, J.A., Kirkpatrick-Steger, K., 1995. Pigeons show same–different conceptualization after training with complex visual stimuli. J. Exp. Psychol. Anim. Behav. Process. 21, 248–252.

Wilson, B., Mackintosh, N.J., Boakes, R.A., 1985. Transfer of relational rules in matching and oddity learning by pigeons and corvids. Q. J. Exp. Psychol. 37B, 313–332.

Wright, A.A., Santiago, H.C., Urcuioli, P.J., Sands, S.F., 1983. Monkey and pigeon acquisition of same/different concept using pictorial stimuli. In: Commons, M.L., Herrnstein, R.J., Wagner, A.R. (Eds.), Quantitative Analyses of Behavior: Discrimination Processes, vol. IV. Ballinger, Cambridge, MA, pp. 295–317.

Wright, A.A., Santiago, H.C., Sands, S.F., 1984. Monkey memory: same/different concept learning, serial probe acquisition, and probe delay effects. J. Exp. Psychol. Anim. Behav. Process. 10, 513–529.

Young, M.E., Wasserman, E.A., 1997. Entropy detection by pigeons: response to mixed visual displays after same–different discrimination training. J. Exp. Psychol. Anim. Behav. Process. 23, 157–170.

Young, M.E., Wasserman, E.A., Dalrymple, R.M., 1997. Memory-based same different conceptualization by pigeons. Psychol. Bull. Rev. 4, 552–558.

Young, M.E., Wasserman, E.A., Garner, K.L., 1997. Effects of number of items on the pigeon's discrimination of same from different visual displays. J. Exp. Psychol. Anim. Behav. Process. 23, 491–501.