Journal of Comparative Psychology
2007, Vol. 121, No. 3, 306–319

0735-7036/07/$12.00 DOI: 10.1037/0735-7036.121.3.306

Human Listeners Provide Insights Into Echo Features Used by Dolphins (*Tursiops truncatus*) to Discriminate Among Objects

Caroline M. DeLong
University of Hawai'i at Manoa and New College of Florida

Whitlow W. L. Au
University of Hawai'i at Manoa

Heidi E. Harley
New College of Florida and
The Seas, Epcot®, Walt Disney World® Resort

Herbert L. Roitblat
DolphinSearch, Inc.

Lisa Pytka
New College of Florida

Echolocating bottlenose dolphins (*Tursiops truncatus*) discriminate between objects on the basis of the echoes reflected by the objects. However, it is not clear which echo features are important for object discrimination. To gain insight into the salient features, the authors had a dolphin perform a match-to-sample task and then presented human listeners with echoes from the same objects used in the dolphin's task. In 2 experiments, human listeners performed as well or better than the dolphin at discriminating objects, and they reported the salient acoustic cues. The error patterns of the humans and the dolphin were compared to determine which acoustic features were likely to have been used by the dolphin. The results indicate that the dolphin did not appear to use overall echo amplitude, but that it attended to the pattern of changes in the echoes across different object orientations. Human listeners can quickly identify salient combinations of echo features that permit object discrimination, which can be used to generate hypotheses that can be tested using dolphins as subjects.

Keywords: echolocation, bottlenose dolphins, human listening, acoustic features, match-to-sample

Bottlenose dolphins echolocate by emitting short, broadband, high-intensity clicks and processing the echoes reflected from objects. The ability of dolphins to detect and discriminate among objects via echolocation is well documented (for a recent review, see Au, 2000). Although these capabilities have been characterized, it is still not clear which acoustic features of object echoes (e.g., amplitude, frequency, duration) convey object properties such as size, shape, and material to an echolocating dolphin.

One approach to identifying the acoustic features in echoes that dolphins may use to discriminate among objects is to present a dolphin with an echoic discrimination task, then measure the object echoes and analyze acoustic differences in the objects. These between-object differences can be examined in conjunction with the dolphin's errors to identify the acoustic features the dolphin may have used in the task (DeLong, Au, Lemonds, Harley, & Roitblat, 2006).

Another approach is to ask human listeners to discriminate among echoes and determine the relevant echo acoustic features. Research to date has indicated that the inner ear of dolphins appears to function similarly to the human inner ear (or any other

Caroline M. DeLong, Department of Psychology, University of Hawai'i at Manoa, and Division of Social Sciences, New College of Florida; Whitlow W. L. Au, Marine Mammal Research Program, Hawai'i Institute of Marine Biology, University of Hawai'i at Manoa; Heidi E. Harley, Division of Social Sciences, New College of Florida and The Seas, Epcot®, Walt Disney World® Resort, Orlando, Florida; Herbert L. Roitblat, DolphinSearch, Inc., Ventura, California; Lisa Pytka, Division of Social Sciences, New College of Florida.

Part of this article is based on a doctoral dissertation by Caroline M. DeLong at the University of Hawai'i at Manoa. This research was supported in part by grants to Caroline M. DeLong from the American Psychological Association (Dissertation Research Award), the American Association of University Women (Pacific Dissertation Fellowship), Soroptimist International (Founder Region Dissertation Fellowship), and SEASPACE (scholarship award), and grants to Heidi E. Harley from Walt Disney Company's Animal Programs.

We would like to thank S. B. Fan for writing the data collection program and D. W. Lemonds for writing the echo acoustic feature measurement program. We are grateful to the trainers and staff at Epcot's Living Seas, particularly E. Putman, A. Stamper, D. Bickel, D. Clark, J. Gory, W. Fellner, M. Muraco, M. Barringer, B. Cavanaugh, L. Davis, D. Feuerbach, C. Goonen, L. Larsen-Plott, K. Odell, C. Litz, J. Davis, T. Hopkins, B. Stevens, J. Mellen, and J. Ogden. We thank E. Hirohata and K. Pugh from the University of Hawai'i and R. Zelski from the University of South Florida for conducting the hearing tests. We thank P. E. Nachtigall, P. Couvillon, A. Peters, N. Gibbs, K. J. Benoit-Bird, J. A. Simmons, and J. Buck for their constructive suggestions on versions of this article.

Correspondence concerning this article should be addressed to Caroline M. DeLong, who is now at the Department of Neuroscience, Brown University, Box GL-N, Providence, RI 02912. E-mail: Caroline_Delong@brown.edu

mammalian ear), except for the dolphins' ability to hear much higher frequencies (Johnson, 1967). For example, both humans and dolphins can discriminate between sounds that differ in intensity by 1 dB (Evans, 1973; Green, 1993), and the frequency discrimination abilities of humans and dolphins for tonal stimuli are comparable in the range of best hearing for each species (Herman & Arbeit, 1972; Thompson & Herman, 1975; Weir, Jesteadt, & Green, 1976). There are some differences between humans and dolphins that could result in differences in the perception of echo stimuli (e.g., sharp frequency tuning and short auditory integration time in dolphins; see Supin & Popov, 1995). Even so, some insight into dolphins' use of echo features may be achieved using human listeners.

The major advantage of using human listeners to classify objects is that, unlike dolphins, they can verbalize how they performed the task. Participants can be asked to describe the acoustic features of the echoes that allowed them to discriminate among the objects. However, introspections are not always reliable or accurate indicators of how people actually perform a task (e.g., Eysenck & Keane, 1995). An analysis of the participants' errors along with an analysis of between-object differences in echo features can be used to validate their introspections and confirm participants' use of certain features. Echolocating dolphins may use complex nonlinear combinations of acoustic features, but it may be difficult to identify them using statistical analyses of the dolphins' performance. It may be easier to identify the use of complex combinations of features in the reports of human listeners.

Previous studies have shown that both blind and blindfolded listeners with normal vision can judge the distance, shape, size, and surface texture of distal objects using echolocation (Kellogg, 1962; Rice, 1967). The humans in these studies investigated objects by projecting self-generated broadband signals (e.g., tongue clicks, snaps, or hisses) and listening to the returning echoes. When using self-generated signals, dolphins are better able to discriminate smaller differences among objects than are humans. For example, a blind human could detect a change in the position of a 30-cm disk placed 0.6 m away from him when the disk was moved nearer or farther away by 10.2 cm (Kellogg, 1962). In comparison, a dolphin could detect a difference in the position of two 7.6-cm foam spheres placed 1 m away when one sphere was 0.9 cm closer to the dolphin than the other sphere (Murchison, 1980).

When presented with echoes that were generated using broadband dolphin signals instead of their own signals, humans are capable of discriminating among objects with approximately the same accuracy as dolphins (Au & Martin, 1989; DeLong, Au, & Stamper, 2007; Fish, Johnson, & Ljungblad, 1976; Helweg, Roitblat, Nachtigall, Au, & Irwin, 1995). These studies showed that humans were capable of discriminating between broadband echoes and reporting relevant echo features, but the researchers did not attempt to directly determine whether the echo features used by the humans were the ones likely to have been used by the dolphins. In the previous studies, only the overall performance of the humans and the dolphin were compared. However, an analysis of errors (i.e., object confusions) could reveal whether the humans and the dolphin may have been using similar echo features to discriminate among objects. Two objects are confused to the degree that they share similar features. If the error patterns of the humans and the dolphin match (i.e., they confuse the same objects), it would imply that they may have used the same features. Conversely, if the error patterns do not match, it would imply they may have used different features.

In the current study, a dolphin performed a cross-modal matching task, and then human participants performed the same task with echoes from the same objects used in the dolphin experiment. A comparison of errors made by the humans and the dolphin was used to determine the echo features the dolphin may have used. In Experiment 1, the participants were asked to discriminate among the objects using echoes collected from multiple orientations of the objects and to report the relevant echo features. In Experiment 2, participants were asked to discriminate between the objects in two echo conditions: "rotating" echoes that were collected from multiple orientations of the objects and "stationary" echoes that were collected from only one orientation of the objects. The goal of these experiments was (a) to determine the salient echo features used to determine object properties (size, shape, material) and (b) to determine the effect of object orientation on discrimination performance.

Bottlenose Dolphin Cross-Modal Matching Experiment

Method

Dolphin Subject

The subject was an adult male Atlantic bottlenose dolphin (*Tursiops truncatus*) housed at Disney's Epcot's Living Seas in Orlando, Florida. At the beginning of the study, the dolphin Toby was approximately 20 years old. Toby was an experienced research subject (Bauer & Johnson, 1994; Xitco, 1996) and had extensive experience with echoic matching (Xitco & Roitblat, 1996) and cross-modal matching (Harley, Putman, & Roitblat, 2003). Sessions were conducted in the main tank of The Living Seas. The tank is a circular saltwater aquarium about 67 m in diameter and 9 m deep, with a volume of 22 million liters. During sessions, Toby usually received one quarter of his daily allotment of approximately 9.5 kg of fish.

Materials and Procedure

There were six object sets, each set consisting of three objects (Figure 1). The objects within each set were selected to vary along one nominal dimension, but because they were natural stimuli, they were actually more variable. One object set was selected to vary in size (stone squares), two object sets were selected to vary in shape (stone shapes and foam cones), two object sets were selected to vary in material (figure 8s and rods), and one object set was selected to vary in surface texture (sockets).

The basic procedure involved a three-alternative identity match-to-sample task in which the dolphin was presented with a sample stimulus, the sample was removed, and then the dolphin was rewarded for selecting the identical stimulus from among three comparison stimuli. The dolphin was asked to match objects within object sets but not between object sets (e.g., if the sample was the medium stone square, the comparisons were always the three objects from the stone squares object set; see Figure 1). All stimuli in this experiment were presented in one cross-modal condition (echoic–visual) and two intramodal conditions (visual–visual and echoic–echoic). In the echoic–visual condition, the

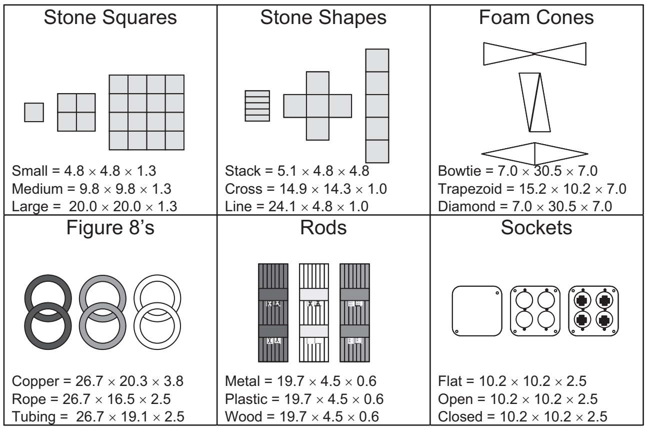

Figure 1. Object sets used in all three experiments. All measurements are height × width × depth (in centimeters).

samples were presented underwater in the echoic sample apparatus, a 0.7-m square PVC-framed box tightly covered in 10-mm black polyethylene that was echoically transparent but visually opaque. Visual comparison stimuli were presented in air, behind acrylic windows in an underwater room accessible to the public and visible from the tank. Because the dolphin could not echolocate through acrylic windows and air, the visual comparison stimuli were visually but not echoically available to the dolphin. In the echoic–echoic condition, samples were presented in the echoic sample apparatus, and choices were presented in the echoic choice apparatus, a rectangular structure covered in polyethelene (front: 2.76 m × 0.90 m; sides: 0.98 m × 0.90 m), the top of which was positioned 1 m below the water's surface. In the visual–visual condition, the samples were presented in air in front of a 0.59 m × 0.46 m black plastic screen, and the choices were hung in air behind acrylic windows.

At the beginning of each echoic–visual trial, the dolphin investigated the sample for several seconds at the underwater echoic sample apparatus, after which he swam to the visual choice array located several meters behind him. After the dolphin positioned himself in front of his object of choice for about 3 s, an assistant naive to the sample's identity reported the dolphin's choice to the trainer who blew a whistle and reinforced the dolphin's response with two small fish for a correct choice or tapped on a metal platform to recall him for an incorrect choice. The dolphin was free to echolocate the objects from several orientations (although he was not allowed to swim behind the sample or choice apparatus) and typically approached to within 0.5 m of the objects (sometimes as close as 5 cm). Echoic–echoic and visual–visual trials were identical except for method of stimulus presentation. Intertrial intervals averaged approximately 60 s (minimum 30 s).

Within a session, each stimulus was presented as the sample object an equal number of times. Each object was also presented equally often in each choice position. The order of the trials was randomized. Object sets were presented in the following order: foam cones, figure 8s, stone squares, stone shapes, sockets, and rods. For each object set, five or six 18-trial sessions were presented in the echoic–visual condition first, followed by two 18-trial sessions in the visual–visual condition, and then two 18-trial sessions in the echoic–echoic condition.

Results and Discussion

The dolphin successfully matched objects cross-modally in three of the six object sets. Figure 2 shows the dolphin's choice accuracy over all six object sets. A binomial test rather than parametric statistics was used for the dolphin because of the single subject design. Chance choice accuracy was 33% because the dolphin could choose from among three alternatives. The dolphin's choice accuracy was significantly better than chance in all three conditions on three object sets: stone squares, foam cones, and stone shapes (summed binomial test, $p < .05$). For the object sets in which the dolphin's cross-modal performance was not significantly different from chance, the dolphin performed poorly in either the visual–visual condition (figure 8s and sockets) or both intramodal conditions (rods).

Table 1 displays the dolphin's errors for each object set. In a match-to-sample task, errors occur when the dolphin chooses an object that does not match the sample object. For example, when the sample is "large" and the dolphin chooses "medium" (or vice versa), the dolphin has made a medium–large object confusion. Chi-square tests were used to determine whether the dolphin's errors were distributed uniformly among the three possible object confusions (see Table 1). For the stone squares, foam cones, and rods, the errors were not distributed uniformly among the three possible object confusions. For these object sets, the dolphin confused one of the three object pairs more frequently than the other two pairs (i.e., the predominant error, shown in bold in Table 1). For the other sets, the dolphin's errors were distributed uniformly among the three possible object confusions.

Human Cross-Modal Matching Experiment 1

Method

Participants

Fourteen participants (8 women and 6 men) with diverse ethnic backgrounds volunteered to be tested. Participants ranged in age from 21 years to 55 years ($M = 28.7$ years). Participants were

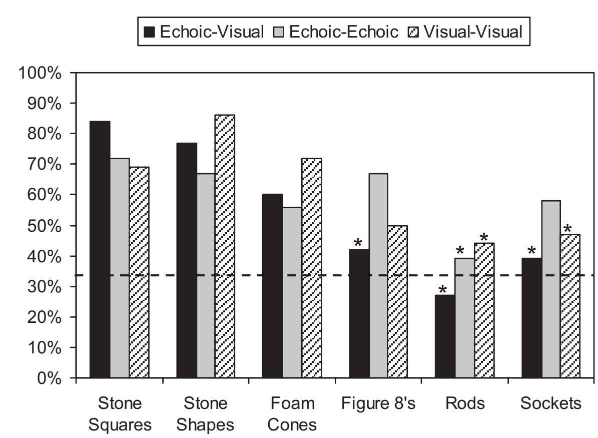

Figure 2. Dolphin's choice accuracy on all six object sets in three conditions. The line shows performance at chance (33% for a three-alternative task). *Performance was not significantly different from chance ($p > .05$).

Table 1
Errors Made by the Dolphin and the Human Participants in Experiment 1

	Human[a]		Dolphin[b]	
Object set/Error	*n*	%	*n*	%
Stone squares				
Small–medium	**24**	**96.0**	3	21.4
Medium–large	1	4.0	**10**	**71.4**
Small–large	0	0	1	7.1
Stone shapes				
Cross–line	**194**	**73.8**	9	42.8
Cross–stack	46	17.5	7	33.3
Stack–line	23	8.7	5	23.8
Foam cones				
Bowtie–diamond	**6**	**75.0**	**22**	**61.1**
Bowtie–trapezoid	1	12.5	7	19.4
Diamond–trapezoid	1	12.5	7	19.4
Figure 8s				
Rope–tubing	**114**	**80.9**	22	42.3
Copper–rope	16	11.3	13	25.0
Copper–tubing	11	7.8	17	32.7
Rods				
Metal–plastic	**85**	**98.8**	15	23.1
Metal–wood	1	1.2	19	29.2
Plastic–wood	0	0	**31**	**47.7**
Sockets				
Open–closed	**77**	**68.1**	30	46.2
Open–flat	22	19.5	16	24.6
Closed–flat	14	12.4	19	29.2

Note. Numbers in the *n* column represent the number of errors in each error type category. Numbers in the % column represent the percentage of errors for each error type (not the percentage correct for each object pair). For example, the dolphin achieved 84% correct on the stone squares (16%, or 14/90 trials were errors). Ten of 14 errors (71.4% of the errors) were confusions between the medium and large objects. Predominant errors are shown in bold type for the humans and the dolphin (i.e., when the errors were not distributed uniformly).
[a] Human errors are shown for all 14 subjects who each received 63 trials per object set. For all six object sets, the errors were not distributed uniformly among the three possible object confusions: stone squares, $\chi^2(2, N = 25) = 44.24$, $p < .001$; stone shapes, $\chi^2(2, N = 263) = 196.48$, $p < .001$; foam cones, $\chi^2(2, N = 8) = 6.25$, $p < .05$; figure 8s, $\chi^2(2, N = 141) = 143.53$, $p < .001$; rods, $\chi^2(2, N = 86) = 166.07$, $p < .001$; sockets, $\chi^2(2, N = 113) = 62.46$, $p < .001$.
[b] Dolphin errors are shown for a single subject who received 90 trials per object set (with the exception of sockets and rods, in which the dolphin received 108 trials per set). The errors were not distributed uniformly among the three possible object confusions for the stone squares, $\chi^2(2, N = 14) = 9.57$, $p < .01$; foam cones, $\chi^2(2, N = 36) = 12.5$, $p < .01$; and rods, $\chi^2(2, N = 65) = 6.40$, $p < .05$. The dolphin's errors were distributed uniformly among the three possible object confusions for the stone shapes, $\chi^2(2, N = 21) = 0.56$, $p > .05$; figure 8s, $\chi^2(2, N = 52) = 2.35$, $p > .05$; and sockets, $\chi^2(2, N = 65) = 5.02$, $p > .05$.

undergraduate or graduate students from the University of Hawai'i or residents of Honolulu. All participants were screened for normal hearing prior to the experiment at the University of Hawai'i Speech Pathology and Audiology clinic and had normal sensitivity in the frequency range of the echo stimuli.

Materials and Procedure

Echo recordings. Stimuli for this experiment included the same objects that were used in the bottlenose dolphin cross-modal matching task (see Figure 1) and the echo recordings of those objects. After the completion of the cross-modal matching experiment with the dolphin, echoes from all the objects were obtained using a representative click recorded from a different bottlenose dolphin. This dolphin click, which has been used in numerous studies, was 70 μs in duration, with a peak frequency of about 120 kHz and a 60-kHz bandwidth (see Au, 1993). This click was considered to be a typical, average click that was likely to be very similar to the clicks that were used by the dolphin Toby during this study (Au, 1993).

Echo recordings were made from more than one orientation of the objects because the dolphin could investigate the object from different angles relative to the object during the matching task, and there is evidence that dolphins attend to the pattern of changes in acoustic features as an object is scanned across a range of target orientations (e.g., Nachtigall, Murchison, & Au, 1980). To simulate the experience the dolphin could get from swimming by the object and echolocating it from different angles, echoes were measured as the object was rotated on its vertical axis. Ten echo train measurements were collected for each object. Each measurement of an object produced a 23-echo train in which 1 echo was captured for each angle (1.3° apart) between –15° and +15° (the orientation of the target that faced the front of the choice apparatus in the matching task was designated as the 0° angle). Thus, the echo trains presented to the human listeners contained echoes from a sequence of orientations (starting at –15° and ending at +15°) that simulates a scan by a moving dolphin. A detailed description of the echo recording measurement setup and procedure can be found in DeLong et al. (2006).

Echo acoustic feature measurements. Four acoustic features were measured for each individual echo: (a) target strength (the ratio in decibels of the echo intensity measured 1 m from the target to the intensity of the incident signal at the location of the target), (b) echo length, (c) peak frequency (frequency of the signal at which the spectrum has its maximum value), and (d) center frequency (the frequency that divides the power spectrum into two equal energy parts). An additional acoustic feature—*target strength bumpiness* (the number of slope direction changes in the plot of target strength across all orientations)—was measured for echo trains (23 echoes collected from –15° to +15°) instead of individual echoes. For more information on how these features were calculated and for other measurements, see DeLong et al. (2006).

Echo formatting. The stimuli were slowed to shift the spectra of the echoes into the human hearing range using CoolEdit 2000 (Syntrillium Software Corporation, Scottsdale, AZ). The original echoes were digitized at 1 MHz and had center frequencies around 120 kHz. The echoes were time stretched by a factor of 125 by converting the echoes from digital to analog at 8 kHz. The time-stretched echoes had center frequencies around 1 kHz. This factor of 125 was chosen so that the echoes presented to the human listeners would fall near their range of best sensitivity (Green, 1976) and also to allow them to pick up more time structure details in the echoes than was possible with a smaller factor (e.g., 50). The total duration of each echo train was approximately 3 s. A bandpass filter was then applied to the echoes (lowpass = 400 Hz, highpass = 1,600 Hz) to reduce background noise.

Presentation of stimuli to human participants. The human listening experimental setup consisted of a single experimenter with a laptop computer that channeled the echo stimuli via two Optimus Pro 50MX stereo headphones to both the participant and the experimenter (for the purpose of monitoring the echoes). To control for inadvertent cueing, the participant sat facing away from the experimenter as the stimuli were played. The participant sat facing the three objects from the object set being tested. A custom-written computer program (in Java) controlled the sequence of trials, stimulus presentation, and data logging.

Participants were tested individually in a quiet, sound-attenuating room by a female experimenter. The participants heard a set of instructions and read a vocabulary sheet with terms to describe the echoes. On the vocabulary sheet were the terms *loudness*, *pitch*, *duration*, and *timbre* with their operational definitions. This procedure ensured that all the participants had the same minimum set of descriptive tools with which they could describe echoic cues (although it was emphasized that they could also use terms not on the list). This vocabulary sheet was available for the participant to reference throughout the experiment.

The order of testing for the six object sets was randomized separately for each participant. For each object set session, there were three phases: training, testing, and interview. The participants completed the entire procedure for a single object set before moving to the next object set. At the beginning of the session for an object set, the three objects of the set were placed on the table in front of the participant. The position of the objects (left, center, right) was randomized separately for each participant and was not changed from trial to trial.

First, participants received training. Three different echo trains for each object were played as the participant simultaneously viewed each object (e.g., three echo trains for the left object were played, followed by three echo trains for the center object, then three echo trains for the right object). The participant could request to listen to echo trains from any of the three objects they wanted to hear again, and the number of training trials was recorded. Participants usually requested to hear the training stimuli more than 3 times for each object (M = 5.9 training trials per object). Echoes used as training stimuli were not used in the test phase. The training phase lasted approximately 2–5 min and was followed immediately by the test phase.

Each test trial consisted of presentation of a stimulus echo train, followed by the participants' response (both a verbal response and a point to the chosen object), and then feedback from the experimenter. The experimenter told the participants whether they were correct or incorrect, and then indicated the correct choice if the participants were incorrect. Participants were allowed to listen to the stimulus echo train as many times as they liked, but they usually requested to hear each stimulus once or twice. For each object set, the average number of times in a trial the participants requested to hear the stimulus echo train is as follows: stone squares = 1.08, stone shapes = 1.40, foam cones = 1.04, figure 8s = 1.24, rods = 1.19, and sockets = 1.25. Each test session consisted of three blocks, each containing 21 trials, for a total of 63 trials. In each of the three 21-trial blocks, seven different sample echo trains for each of the three objects were presented. The order of the sample echo trains was randomized separately for each participant. Participants completed the three-block test session in sequence with no pause between the three blocks. Participants completed a test session (63 trials) in 7–19 min (M = 9 min).

After the participant completed the test, participants were asked to respond to the question, "What cues did you use to discriminate among the objects?" in the tape-recorded cue-use interview which lasted approximately 5-10 min. When the interview was complete, the participant began training for the next object set. After the participants had completed the entire experiment, the experimenter asked, "Which cues in the echoes carried information about the properties of size, shape, and material in the objects?" All participants completed all six object sessions in a single experimental session that lasted approximately 2.5 hours.

Results

Performance Accuracy and Error Analysis

The participants were able to visually select the object that matched the correct echo train. Figure 3 shows the participants' choice accuracy on each object set. Chance choice accuracy was 33% because the participants could choose from among three alternatives. To determine whether the participants' performance was above chance on each object set, we performed a separate t test for each object set using one score for each participant (average performance on the set) and comparing the participants' scores to a value of .33. The participants' performance was significantly above chance for all six object sets (see Figure 3). This contrasts with the dolphin's performance, in which choice accuracy was below chance on three sets (figure 8s, rods, sockets). One potential reason for this discrepancy is that the dolphin had difficulty visually discriminating among the objects in these three sets (which were the same size and shape within sets), whereas the humans easily visually discriminated among all the objects.

A 6 (object set) × 3 (blocks) analysis of variance (ANOVA), with both factors as repeated measures, was conducted on the proportion of correct answers made by the participants. There was a significant effect of object set, $F(5, 65) = 22.52$, $p < .001$, $\eta_p^2 = .45$. Post hoc analyses revealed that choice accuracy was significantly better on the foam cones set (99.0%) and the stone squares set (97.2%) than on the other four sets. Choice accuracy was significantly worse on the stone shapes set (70.2%) than on the other five sets. Choice accuracy did not differ significantly among the figure 8s (84.0%), rods (90.2%), and sockets (87.2%; Newman–Keuls tests, $p < .05$).

There was also a significant effect of blocks (Block 1: Trials 1–21; Block 2: Trials 22–42; Block 3: Trials 43–63), $F(2, 26) = 18.57$, $p < .001$, $\eta_p^2 = .02$; and an interaction between object set and blocks, $F(10, 130) = 2.79$, $p < .01$, $\eta_p^2 = .02$. For three of the object sets (stone shapes, figure 8s, sockets), the participants' performance improved between blocks. For the stone shapes set, performance on the third block exceeded performance on the first or second block (Newman–Keuls tests, $p < .05$). For both the figure 8s and the sockets, performance on both the second and third block exceeded performance on the first block (Newman–Keuls tests, $p < .05$). There was no significant difference in performance between blocks on the other three object sets.

Table 1 shows the object confusions made by the participants. Chi-square tests were performed separately for each set to determine whether the confusions were distributed uniformly among the three possible object confusions. For all six object sets, the errors were not distributed uniformly among the three possible

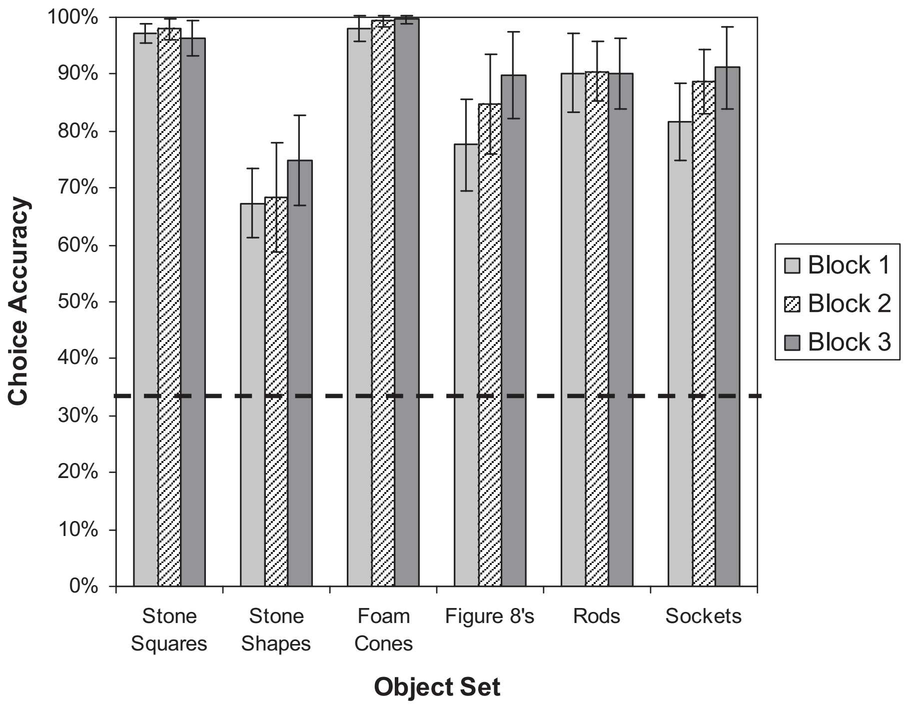

Figure 3. Average choice accuracy of the human participants on each object set in Human Cross-Modal Matching Experiment 1 shown by block (Block 1 = Trials 1–21; Block 2 = Trials 22–42; Block 3 = Trials 43–63). The line shows performance at chance (33% for a three-alternative task). Performance was significantly above chance ($p < .001$) for all six sets: stone squares, $t(13) = 90.45$; stone shapes, $t(13) = 11.39$; foam cones, $t(13) = 134.56$; figure 8s, $t(13) = 15.56$; rods, $t(13) = 21.72$; sockets, $t(13) = 21.22$. Vertical bars show the 95% confidence intervals.

object confusions. Instead, the participants confused one of the three object pairs more frequently than the other two pairs (i.e., the predominant error, shown in bold in Table 1).

Reported Echoic Cues

The discriminatory cues reported by the participants were evaluated by examining a transcript of the tape-recorded interviews. For each participant, the answer to the question "What cues did you use to discriminate the objects?" was classified into categories (e.g., overall loudness, overall duration). The answer to the post-experiment question "Which cues in the echoes carried information about the properties of size, shape, material, and texture in the objects?" was classified into the same categories. To assess reliability, a second coder coded 93% of these interview answers using the experimenter's classification system (the other 7% of the interview answers were used to train the coder). The experimenter and second coder agreed on the classification of 126 of 130 (97%) interview answers.

The echoic cues reported by the human listeners are shown in Table 2. Participants reported using between one and four cues to discriminate among the objects in each set ($M = 2.1$). Participants generally used more than one cue on five of the object sets, and different cues were used to discriminate between different object pairs. For example, in the stone shapes set, some participants reported using overall loudness to distinguish between the stack and the cross and line (stack was softest) and pitch to distinguish between the cross and line (line had a lower pitch).

The cue reported most frequently by participants on all the object sets was the pattern of change in loudness, pitch, and timbre across the echo train. For the foam cones, which varied in shape, this was the only cue they needed to discriminate between the objects. For the other sets, this cue was used in combination with other cues, sometimes for particular object pairs within the set. For the objects that varied in size (stone squares, stone shapes), one of the most salient cues reported was overall loudness. For the objects that varied in material and texture (figure 8s, rods, sockets), overall timbre, loudness, and pitch were frequently reported.

After the participants had completed the experiment, they were asked, "Which cues in the echoes carried information about the properties of size, shape, material, and texture in the objects?" All 14 participants reported that overall loudness conveyed the object size. Twelve participants reported that change in loudness or pitch across the echo train conveyed object shape. Ten participants reported that change in loudness or pitch across the echo train also conveyed object texture (texture could almost be considered "surface shape"). Ten participants reported that overall timbre conveyed material.

The errors made by the participants were examined in conjunction with the acoustic measurements of the echoes and the discriminatory cues reported by the participants to confirm that they actually used the cues they reported. Table 3 shows the acoustic

Table 2
Echoic Cues Used by Human Participants to Discriminate Among Objects in Each Object Set in Experiment 1 (N = 14)

	Object set					
Echoic cue	Stone squares	Stone shapes	Foam cones	Figure 8s	Rods	Sockets
Loudness (overall)	14	11	0	3	11	5
Pitch (overall)	3	2	0	6	3	2
Timbre (overall)	3	0	1	10	5	6
Duration (overall)	1	1	0	0	1	1
Change in loudness, pitch, or timbre across echo train	10	12	14	13	12	13
Average number of cues reported	2.6	2.1	1.1	2.1	2.2	2.1

Note. Each cell contains the number of participants who reported using each cue for each of the object sets. Participants could report multiple cues for each set.

feature measurements for each of the objects. The reported cue of loudness was considered to be correlated with the *target strength* (TS; a measure of echo intensity) of the echoes, the cue of pitch with either the peak frequency or the center frequency of the echoes, and the cue of duration with the length of the echoes. The cue of change in loudness or pitch across the echo train was considered to be partly correlated with the echo feature TS bumpiness (see Table 3). The cue of timbre was not evaluated in this analysis because it is unclear which echo features would give rise to the perception of timbre (it is a complex combination of the waveform, sound pressure, frequency, and temporal characteristics of the stimulus; American National Standards Institute, 1960).

It was possible to confirm that participants actually used the cues they reported for the majority of the object sets. Participants reported loudness as a cue to distinguish between all three objects within a set, so we would expect them to make the most errors on the pair of objects that had the most similar target strengths. This was indeed the case for five of six object sets (all but the sockets; see Table 3). Participants who reported using pitch made the most errors on the objects that were the most similar in either peak or center frequency for three of five sets (figure 8s, rods, sockets). When participants reported using the pattern of changes in the echo, their errors were correlated with the pair of objects with the most similar TS bumpiness scores for three of six sets (stone squares, foam cones, rods). However, this TS bumpiness score may not have measured the properties of the echo that the participants were hearing (e.g., it measured change in loudness but not change in pitch over the train). Finally, the echo measurements show that the object echoes differed in length, but duration was reported as a cue so rarely that it was not clear whether participants were inaccurately evaluating that feature or whether it was simply overshadowed by more salient cues.

Discussion

The human participants were able to discriminate among the objects using echoes, and they reported using multiple acoustic features of the echoes to perform the matching task. The extent to which the acoustic features reported by the humans could also have been used by the dolphin can be examined by comparing the error patterns of both species. This analysis yielded three categories of object sets: (a) those in which the error patterns of the dolphin and humans strongly matched (i.e., they had the same predominant error), (b) those in which the error patterns of the dolphin and humans moderately matched (i.e., the error type that accounted for the majority of the dolphin's errors matched the humans' predominant error), and (c) those in which the error patterns of the dolphin and humans were strongly mismatched (i.e., they had a different predominant error; see Table 1). The second category is considered to be a "moderate" match instead of a "strong" match because the dolphin's errors were more equally distributed among the three object confusions than was the case for the humans, who always showed a strong tendency to confuse one object pair more frequently than the other two pairs.

There were two object sets in which the dolphin and the humans made a different predominant error: stone squares and rods. The stone squares varied primarily in size, and all the human listeners reported using overall loudness to discriminate among them. The small and medium squares were closest in amplitude (difference of 4.5 dB), and nearly all of the humans' errors were confusions between those objects. However, the dolphin usually confused the medium and large squares, which were quite far apart in amplitude (difference of 13.1 dB). The majority of the human listeners also reported using overall loudness as one of the cues to discriminate among the rods. The humans mainly confused the two objects closest in amplitude (metal and plastic; 2.4-dB difference), whereas the dolphin mainly confused plastic and wood (7.3-dB difference).

These results suggest that this dolphin did not seem to rely on overall amplitude differences to discriminate among these objects, and it is unclear why. Another study in which the dolphin's error patterns were examined along with between-object differences in acoustic features also found that the dolphin rarely seemed to use amplitude differences to discriminate among objects (DeLong et al., 2006). Au, Schusterman, and Kersting (1980) found that dolphins are capable of discriminating between spheres and cylinders of varying size but overlapping amplitudes (i.e., amplitude could not be used as a cue, so the dolphin had to use some other echo feature). It is possible that dolphins preferentially attend to frequency cues (e.g., peak and center frequency; see DeLong et al.,

Table 3
Acoustic Feature Measurements for the Rotating Echoes in Experiments 1 and 2

	Acoustic feature				
Object	TS (dB)	PF (kHz)	CF (kHz)	Length (μs)	TS: train
Stone squares					
Small	-42.7_a	114.2_a	117.8_a	602.1_a	10.2_a
Medium	-38.2_b	120.9_b	121.4_b	275.6_b	9.4_a
Large	-25.1_c	111.3_c	114.4_c	97.2_c	7.3_b
Stone shapes					
Cross	-28.5_a	117.9_a	118.4_a	72.7_a	3.5_a
Line	-29.6_a	111.1_b	114.7_b	98.3_b	4.4_a
Stack	-33.2_b	120.9_c	118.2_a	195.6_c	4.0_a
Foam cones					
Bowtie	-43.0_a	111.6_a	110.3_a	206.7_a	6.5_a
Diamond	-43.9_a	111.4_a	109.4_a	117.0_a	6.9_a
Trapezoid	-45.1_b	108.7_b	108.4_b	498.3_b	9.6_b
Figure 8s					
Rope[a]	-33.0_a	114.3_a	113.9_a	95.3_a	8.9_a
Tubing[a]	-32.3_b	116.0_a	115.5_b	147.7_b	12.0_b
Copper	-35.3_c	111.2_b	116.6_c	230.2_c	9.9_a
Rods					
Metal	-33.1_a	118.0_a	117.9_a	145.4_a	8.4_a
Plastic	-30.6_b	118.9_a	120.6_b	73.5_b	8.8_a
Wood	-23.4_c	115.9_b	116.3_c	53.1_c	5.4_b
Sockets					
Closed	-29.1_a	120.8_a	122.3_a	115.1_a	6.2_a
Open	-32.3_b	115.9_b	122.2_a	188.3_b	11.0_b
Flat	-33.3_b	110.9_c	113.3_b	145.0_c	6.6_a

Note. The object pairs that were most often confused by the participants are shown in italics for each object set. All values are given for the original echoes presented to the dolphin. The values in each cell in the first four columns represent the mean for each feature averaged across all object orientations for 10 echo measurements. The values in the last column represent the mean for each feature averaged across 10 echo trains. Separate multivariate analyses of variance were conducted for each of the sets, and post hoc object comparisons were conducted to examine between-objects differences (see DeLong et al., 2006). Objects sharing the same subscript were not statistically significantly different from each other. TS = target strength; PF = peak frequency; CF = center frequency; TS: train = target strength bumpiness score for the echo trains.
[a] Humans cannot discriminate between sounds that are less than 1 dB apart, so rope and tubing were probably not perceived as different in loudness, even though they are statistically significantly different.

2006), given their extremely fine frequency discrimination (Supin & Popov, 1995).

There were three object sets in which the error type that accounted for the majority of the dolphin's errors matched the humans' predominant error: stone shapes (varied in size and shape), figure 8s (varied primarily in material), and sockets (varied primarily in surface texture). The figure 8s actually varied slightly in shape and size because of the different malleability of the three materials (rope, copper, tubing). For all three sets, again the human listeners reported using the pattern of change in loudness (or pitch) across the echo train. They also reported using overall loudness, pitch, and timbre. The moderate match between the error patterns of the dolphin and the humans suggests that the dolphin may have used some of the cues reported by the humans, or the dolphin may have used different cues in addition to (or in place of) the ones reported by the humans.

There was only one set in which the dolphin and the humans made the same predominant error: foam cones, which vary primarily in shape. In this set, all the human listeners reported using the pattern of change in the echo train in loudness (or pitch) as the object rotated. This was a very salient and effective cue, as evidenced by their high choice accuracy. Both the humans and the dolphin most frequently confused the diamond and the bowtie, and both split their remaining errors equally between the other two object pairs (see Table 1). The close similarity in the error patterns of the humans and the dolphin suggests that it is very likely that the dolphin used the same cue as the humans. There is other evidence that dolphins attend to the pattern of changes in acoustic features as an object is scanned across a range of target orientations (e.g., Nachtigall et al., 1980). Nachtigall et al. (1980) trained a dolphin to discriminate between a foam cube and a cylinder in an upright position. When given probe trials in which the orientation of the objects was changed, the dolphin failed to discriminate between the objects. Nachtigall et al. hypothesized that the dolphin was attending to the pattern of changes in amplitude as the dolphin scanned across the objects. The dolphin succeeded initially because the cube and cylinder had different patterns and failed in the probe condition when the objects had similar patterns.

The human listeners frequently relied on using the pattern of changes in the echo as the object rotated for all the objects. This led to the question of whether having information from multiple orientations of the object was necessary for discriminating among all the objects or was necessary for identifying specific object properties. In Experiment 2, a different echo condition was introduced in which the human listeners were asked to discriminate among the objects using echoes collected from just one object orientation.

Human Cross-Modal Matching Experiment 2

In this experiment, there were two echo conditions: rotating (echoes from multiple object orientations) and stationary (echoes from one object orientation). The human listeners were asked to perform the same matching task as in Experiment 1 in both echo conditions. If multiple object orientations are necessary for discriminating among the objects, then the participants should perform significantly better in the rotating echo condition than in the stationary echo condition. If multiple object orientations are not necessary, then there should be no difference in performance between the two echo conditions. We predicted that using echoes with multiple object orientations would be critical for objects that varied in shape (foam cones and stone shapes), but not for objects that varied in size (stone squares) or material (figure 8s).

Method

Participants

Sixteen participants (11 women and 5 men) with diverse ethnic backgrounds volunteered to be tested. Participants ranged in age from 19 years to 26 years (M = 21.1 years). Participants were undergraduate students at New College of Florida in Sarasota, Florida. All participants were screened for normal hearing prior to the experiment by audiologists from the University of South Florida and had normal sensitivity in the frequency range of the echo stimuli. No participants from Experiment 1 were used in this experiment.

Materials

Stimuli for this experiment included four of the six object sets that were used in Experiment 1: stone squares, stone shapes, foam cones, and figure 8s (see Figure 1). There were two types of echo trains for each object set: rotating and stationary. Figure 4 shows an example of a rotating versus stationary echo train. The rotating trains were the same echo trains that were used in Experiment 1 (trains of 23 echoes in which 1 echo was captured for each angle between –15° and +15°; echoes were obtained using a bottlenose dolphin click). These are called rotating trains because the echoes were captured as the object rotated, and each train included echoes from multiple orientations of the object. Ten stationary trains were created for each object using the echoes from the 0° orientation of the rotating echo trains. Each stationary echo train contained 23 echoes, the same number of echoes as in the rotating trains. For each object, there were 10 exemplars of rotating trains. Each stationary train consisted of twenty-three 0° echoes taken from those 10 exemplars (each train contained duplicates from some exemplars) in random order. These trains are called stationary

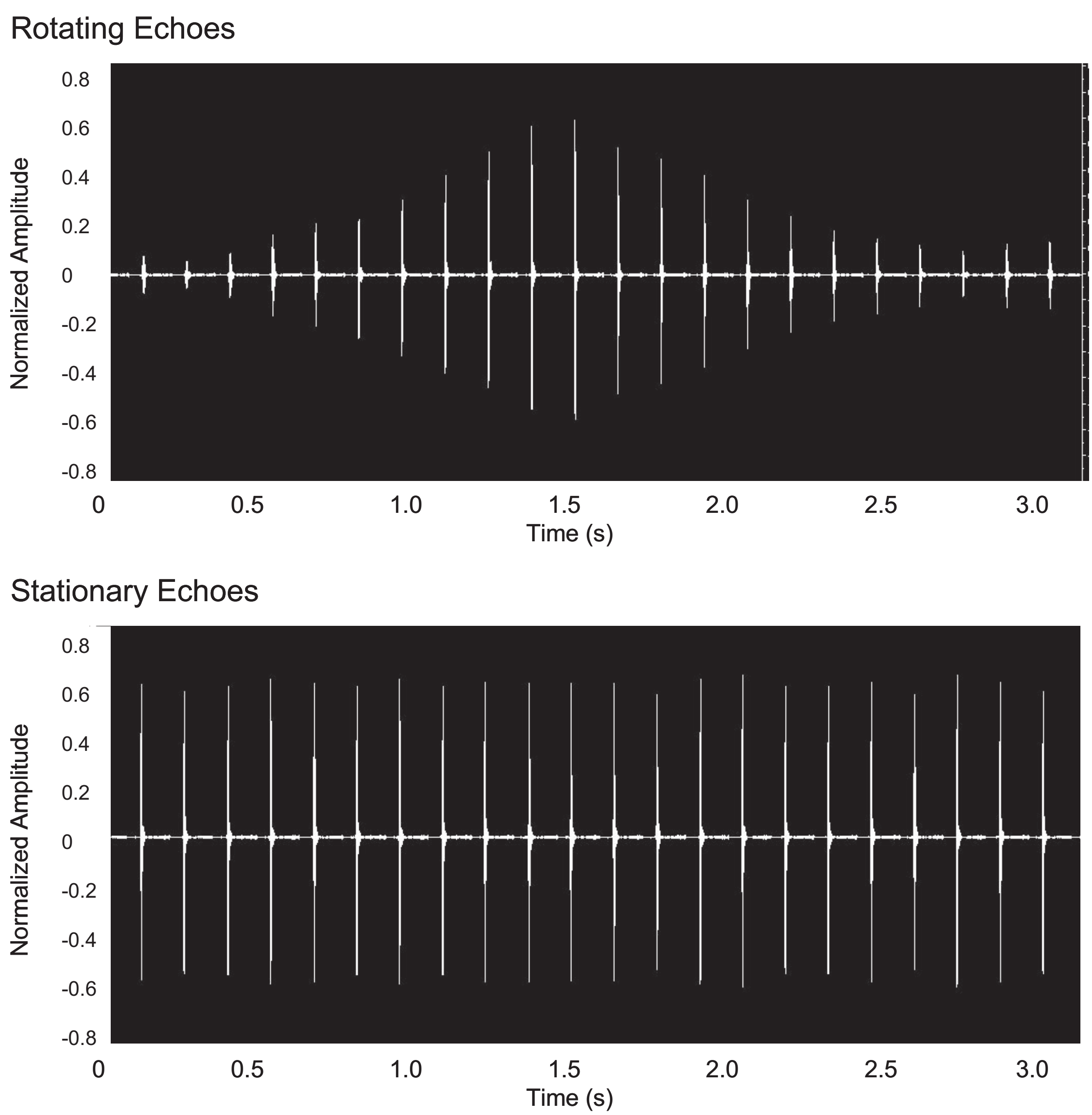

Figure 4. The two echo conditions used in Human Cross-Modal Matching Experiment 2. These example echoes come from the cross in the stone shapes set. Each graph shows a 23-echo train (each line represents a single echo). The top graph shows echoes in the rotating condition (echoes are from a range of object orientations from –15° at Time 0 s to +15° at Time 3 s). The bottom graph shows echoes in the stationary condition (each echo is from the 0° orientation). Note that not all the echoes from the 0° orientation are exactly the same amplitude: Each stationary train is made up a mixture of 0° echoes from all 10 different rotating trains for that object.

because they include echoes from only one orientation of the object, so in effect the object is stationary instead of rotating as in the rotating trains.

Procedure

The experimental setup and procedure were the same as in Experiment 1 with a few exceptions. Before the training and testing began, participants were read a set of instructions and given a vocabulary sheet with terms to describe the echoes. In Experiment 2, participants were played pure tone sounds demonstrating each of the terms as they viewed the vocabulary sheet (these sounds were not the echoes used in the experiment). They were presented with a set of sounds that differed in loudness only (same pitch and timbre), a set of sounds that differed in pitch only (880-Hz tone vs. 220-Hz tone; same volume and timbre), and a set of sounds that differed in timbre only (middle C played by a french horn, muted trumpet, and soprano sax at the same volume).

Participants were randomly assigned to one of four groups. Each group was presented with the four object sets in a different order (Group A: foam cones, stone squares, stone shapes, figure 8s; Group B: stone shapes, figure 8s, foam cones, stone squares; Group C: stone squares, foam cones, figure 8s, stone shapes; Group D: figure 8s, stone shapes, stone squares, foam cones). Within each group, half of the participants received the rotating echo condition first, followed by the stationary echo condition; the other half received the echo conditions in the opposite order.

For each object set session, there were three phases: training, testing, and interview. The participants completed the three phases for a single object set in one of the echo conditions (rotating or stationary), and then completed the procedure for the other echo condition for the same object set before moving to the next object set. The training and testing procedure was the same as in Experiment 1 with the exception of the number of trials per session. Each test session consisted of three blocks, each containing 15 trials, for a total of 45 trials. In each of the blocks, 5 different sample echo trains for each of the three objects were presented. Thus, there were 15 total echo trains presented for each object (only 10 echo trains were created for each object, so 5 echo trains were randomly selected to be reused). After the participants completed each test phase, the experimenter asked them to report the cues they used in the tape-recorded interview (same questions as in Experiment 1). All participants completed all eight sessions (4 object sets × 2 echo conditions) in a single experimental session that lasted 2–3 hours.

Results

Performance Accuracy and Error Analysis

The participants were able to visually select the object that matched the correct echo train for all four object sets. Figure 5 shows the participants' average choice accuracy on all sets in the rotating versus stationary echo condition. As in Experiment 1, we performed a separate *t* test for each object set using one score for each participant (average performance on the set) and comparing the participants' scores to a value of .33 to determine whether performance was above chance. The participants' performance was significantly above chance for all four object sets in both the

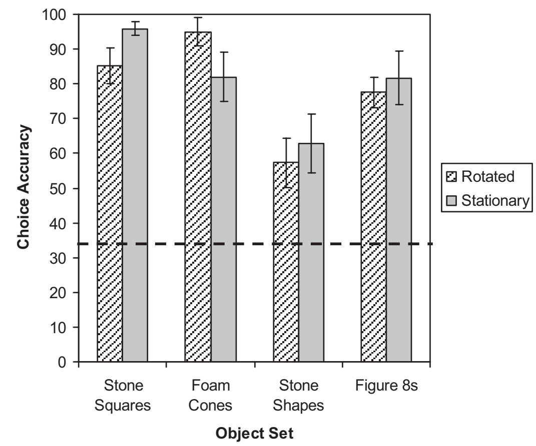

Figure 5. Average choice accuracy of the human participants in Experiment 2. The line shows performance at chance (33% for a three-alternative task). Performance was significantly above chance ($p < .001$) for all four sets in both echo conditions: stone squares, rotating, $t(15) = 21.26$; stone squares, stationary, $t(15) = 58.44$; stone shapes, rotating, $t(15) = 6.68$; stone shapes, stationary, $t(15) = 8.03$; foam cones, rotating, $t(15) = 31.39$; foam cones, stationary, $t(15) = 15.73$; figure 8s, rotating, $t(15) = 21.33$; figure 8s, stationary, $t(15) = 16.70$. Vertical bars show the 95% confidence intervals. Performance on rotating versus stationary echoes was significantly different only for the foam cones and the stone squares sets.

rotating and stationary conditions (see Figure 5). These results indicate that the participants could discriminate among the objects whether they heard echoes from multiple object orientations or from a single orientation.

A 2 (echo order: rotated first vs. stationary first) × 4 (object set: stone squares, stone shapes, foam cones, figure 8s) × 2 (echo condition: rotated vs. stationary) ANOVA, with the last two factors as repeated measures, was conducted on the proportion of correct answers made by the participants. There was a significant effect of object set, $F(3, 42) = 44.87$, $p < .001$, $\eta_p^2 = .52$; and an interaction between object set and echo condition, $F(3, 42) = 7.70$, $p < .001$, $\eta_p^2 = .07$.

The participants' performance varied between object sets. Choice accuracy was significantly better on the stone squares set (90.6%) and the foam cones set (88.7%) than on the other two sets. Choice accuracy was significantly worse on the stone shapes set (60.1%) than on the other three sets (Newman–Keuls tests, $p < .01$). This pattern of performance (best performance on stone squares and foam cones, worst performance on stone shapes) is the same as in Experiment 1.

Performance on rotating versus stationary echoes was significantly different only for the foam cones and the stone squares set (Newman–Keuls tests, $p < .05$). For the foam cones set, participants performed significantly better in the rotating echo condition (95.1%) than in the stationary echo condition (82.2%). The opposite was true for the stone squares set; participants performed better in the stationary echo condition (95.9%) than in the rotating echo condition (85.3%). There was no significant difference in performance in the rotating versus stationary condition for the figure 8s (77.6% vs. 81.7%) or the stone shapes (57.4% vs. 62.9%).

The participants' errors in the rotating versus stationary condition are shown in Table 4. The participants made the same pre-

Table 4
Errors Made by the Human Participants on Rotating vs. Stationary Echoes in Experiment 2

Object set/Error	Rotating *n*	Rotating %	Stationary *n*	Stationary %
Stone squares				
Small–medium	**101**	**95.3**	**27**	**93.1**
Medium–large	5	4.7	2	6.9
Small–large	0	0	0	0
Stone shapes				
Cross–line	**146**	**47.6**	**140**	**52.4**
Cross–stack	95	30.9	64	24.0
Stack–line	66	21.5	63	23.6
Foam cones				
Bowtie–diamond	**25**	**71.4**	8	6.3
Bowtie–trapezoid	5	14.3	5	3.9
Diamond–trapezoid	5	14.3	**115**	**89.8**
Figure 8s				
Rope–tubing	**150**	**93.2**	**123**	**93.2**
Copper–rope	8	5.0	6	4.6
Copper–tubing	3	1.9	3	2.3

Note. Numbers in the *n* column represent the number of errors in each error type category. Numbers in the % column represent the percentage of errors for each error type. Predominant errors are shown in bold type. Sixteen subjects received 45 trials per object set (720 total trials per set) for each condition (rotating echoes = 720 trials; stationary echoes = 720 trials).

dominant error on three of the object sets regardless of whether the echoes were rotating or stationary (stone squares, stone shapes, figure 8s). For these three sets, the predominant errors match those of Experiment 1. In the foam cones set, participants made a different predominant error in the rotating versus stationary condition. They confused the bowtie and diamond most often when the echoes were rotating, but they confused the diamond and trapezoid most often when the echoes were stationary.

Reported Echoic Cues

The discriminatory cues reported by the participants were evaluated by examining a transcript of the tape-recorded interviews and classifying the answer to the question "What cues did you use to discriminate the objects?" into categories that match the ones used in Experiment 1 (see Table 2). Because of equipment malfunction, 15 of 16 participant interviews were coded. To assess reliability, a second coder coded 93% of these interview answers using the experimenter's classification system. The experimenter and second coder agreed on the classification of 138 of 154 (90%) interview answers.

The cues reported by the participants in Experiment 2 are shown in Table 5. Participants used between one and five cues to discriminate between objects in each set ($M = 2.1$). The majority of the participants reported using multiple cues on each object set. As in Experiment 1, participants sometimes reported using different cues to discriminate between different object pairs within the three object sets. For example, in the figure 8s set, all 15 participants reported that copper had a different overall timbre than rope or tubing and was easy to identify, but they distinguished between rope and tubing using a variety of other cues, such as overall pitch or loudness.

For all object sets in the rotating echo condition, the majority of participants reported using the pattern of change in loudness, pitch, and timbre across the echo train, similar to Experiment 1. This cue was reported most frequently for the object sets that varied in shape (foam cones and stone shapes), but it was a salient cue for all sets. For object sets in the stationary echo condition, participants reported using differences between objects in overall loudness, pitch, or timbre (echo duration was seldom reported). These "overall" differences were reported in the rotating echo condition too.

The participants reported using different cues in the rotating echo condition versus the stationary echo condition for the two object sets that varied in shape. For both the foam cones and stone shapes, the participants reported using primarily the pattern of change in loudness, pitch, and timbre across the echo train in the rotating echo condition (secondary cues of overall loudness and pitch were reported for the stone shapes, which also varied in size). However, in the stationary echo condition (in which pattern cues were not available), more participants reported using overall loudness, pitch, and timbre than they did in the rotating echo condition.

Table 5
Echoic Cues Used by Human Participants to Discriminate Among Objects in Each Object Set in Experiment 2 (N = 15)

Echoic cues	Stone squares (size) R	Stone squares (size) S	Stone shapes (size/shape) R	Stone shapes (size/shape) S	Foam cones (shape) R	Foam cones (shape) S	Figure 8s (material) R	Figure 8s (material) S
Loudness (overall)	15	14	8	9	5	13	1	2
Pitch (overall)	5	7	6	10	0	3	6	7
Timbre (overall)	7	7	3	8	1	11	15	14
Duration (overall)	0	0	1	3	0	3	0	3
Change in loudness, pitch, timbre, or duration across echo train	9	0	13	1	15	1	8	2
Average number of cues reported	2.5	1.9	2.3	2.1	1.6	2.1	2.2	1.9

Note. Each cell contains the number of participants who reported using each cue for each of the object sets. Participants could report multiple cues for each set. R = rotated echoes; S = stationary echoes.

In the object sets that varied in size (stone squares) and material (figure 8s), the cues reported by participants were very similar in both the rotating and stationary conditions. The predominant cue for the objects that varied in size was overall loudness, whereas the predominant cue for the objects that varied in material was timbre.

After the participants had completed the experiment, they were asked, “Which cues in the echoes carried information about the properties of size, shape, and material in the objects?” Ten of 15 participants reported that overall loudness conveyed the object size. Thirteen participants reported that change in loudness, pitch, and timbre across the echo train conveyed object shape. Fourteen participants reported that overall timbre conveyed material.

Discussion

The participants were able to discriminate among all the objects whether they used echoes from multiple object orientations (rotating) or echoes from a single object orientation (stationary). They were able to find echoic cues that allowed them to discriminate among the objects in both echo conditions. However, these data show that having information from multiple object orientations was advantageous for the object set that varied primarily in shape (foam cones) but not necessary for the sets that varied primarily in material (figure 8s), shape and size (stone shapes), and size (stone squares).

The participants did not need multiple orientation information to discriminate between different-sized objects. In the stone squares set, the participants actually performed best with the stationary echoes. This may be because those echoes highlighted the predominant cue that they reported in both the rotating and stationary echo conditions: overall loudness. This cue may have been more obvious and salient in the stationary echoes because all the echoes within an object’s echo train were approximately the same amplitude (i.e., all the echoes for the small square were soft, all the echoes for the medium square were moderately loud, and all the echoes for the large square were very loud). In the rotating trains, the medium square had many soft echoes in the beginning and the end of the train, and the small square had moderately loud echoes in the middle of the train.

Echo information from multiple orientations also was not necessary to discriminate among objects made of different materials. In the figure 8s set, participants reported using timbre and pitch in both echo conditions. Approximately half of the participants reported using the cue of the pattern of changes across the echo train in the rotating echo condition, but this additional information did not lead to a significant improvement in their performance over the stationary echo condition.

Listening to echoes from a range of orientations showed a clear advantage for the different-shaped foam cones. Performance in the rotating condition was nearly perfect (95%), but performance in the stationary condition was also above chance (82%). This is because the participants were able to find other cues in the stationary echoes that allowed them to discriminate the objects (overall loudness and timbre). However, it is doubtful that these cues conveyed the shape of the objects in the way that the rotating echoes did. Shape information in aspect-dependent shapes such as the foam cones appears to be conveyed through listening to how the echoes change as the object is rotated.

Human listeners in another study also required information from multiple orientations to discriminate among aspect-dependent shapes (Helweg et al., 1995). When presented with echo trains from a rectangular prism, a pyramid, and a cube that included echoes collected from a number of different object orientations, the participants were able to discriminate among the objects. Helweg et al. (1995) then presented the participants with two transfer tests designed to determine how they would perform without multiple orientation information. In the first transfer test, participants were given single echoes as stimuli instead of echo trains. Using a single echo reduced the humans’ performance by 50%. In the second transfer test, the amplitudes of the echoes within the echo train were equated. Participants were unable to discriminate among the shapes when given amplitude-equated echo trains. These results suggest that to discriminate among aspect-dependent shapes, the participants needed to integrate changes in amplitude across successive echoes from different object orientations.

The participants’ errors confirm that they were using similar cues on the rotating and stationary conditions for all sets except the foam cones. For the stone squares, stone shapes, and figure 8s, they made the same predominant error in both conditions. This suggests that they were using the same echo features to discriminate among the objects (features that were available in the stationary echoes). However, with the foam cones, they confused the bowtie and diamond most frequently in the rotating condition, but they confused the diamond and trapezoid most frequently in the stationary condition. This suggests that they were using different cues in the two conditions.

The dolphin’s error patterns match the humans’ error patterns for the foam cones in the rotating condition, not in the stationary condition. This implies that the dolphin probably used the cues reported by the humans in the rotating condition, not the cues reported for the stationary condition. This experiment provides further evidence that dolphins attend to the pattern of changes in the echo across the echo train.

General Discussion

The human listeners in this study performed as well or better than the dolphin, and they were able to report echo features that allowed them to discriminate among the objects. An examination of the similarities and differences between the error patterns of the humans and the dolphin was used to discern which acoustic features are likely to have been used by the dolphin. When the error patterns of the humans and the dolphin matched (i.e., they confused the same objects), it implies that they may have used the same features. When the error patterns did not match, it implies they may have used different features. This error analysis reveals that the dolphin did not rely on overall echo amplitude differences between objects, but that it is likely that the dolphin attended to the pattern of changes in the echo as the object was scanned from different orientations.

This is the first human listening study to employ a comparative error analysis in an attempt to determine whether the acoustic features reported by the humans were also used by the echolocating dolphin. It was based on the hypothesis that two objects are confused to the degree that they share similar acoustic features, and that if both species confuse the same objects, they are using the same features. The possibility exists that the humans and the dolphin could have used different features that only coincidentally produced the same error patterns, or that they could have used the same features but produced different error patterns. Subsequent studies using this comparative error analysis and then directly testing the dolphins using the results of the human–dolphin comparison will be able to determine whether matching error patterns are a good indicator of use of similar features.

This comparative error analysis seems promising because these data are in agreement with the results of other studies that suggest that dolphins investigate objects from multiple orientations and use that information to discriminate among objects (DeLong et al., 2006; Helweg, Au, Roitblat, & Nachtigall, 1996; Nachtigall et al., 1980; Roitblat, Penner, & Nachtigall, 1990). In one study, a dolphin performed an echoic match-to-sample task and then the acoustic features of the object echoes were analyzed in conjunction with the dolphin's errors (DeLong et al., 2006). The object sets in that task included some of the same objects that were used in the current study. The results suggested that the pattern of changes across the echo train as a function of orientation could have been used by the dolphin to discriminate among the objects in six of the nine object sets. These sets included objects that varied primarily in shape (foam cones) but also sets that varied primarily in size, material, and texture. This could mean that this cue is useful to the dolphin when discriminating objects in a variety of situations, not just when faced with objects that vary in shape.

In this study, the human listeners were able to discriminate among the objects both when they had access to multiple object orientation information and when they could use only a single object orientation. They were able to adapt to a new context and switch the acoustic features they used. They also used a combination of features for each object set and focused on different features to discriminate between certain object pairs within a set. If dolphins behave similarly, finding an answer to the question "Which acoustic features convey object properties such as size, shape, and material?" may be complex and challenging.

Similarities and differences between human and dolphin auditory perception should be considered when deciding whether echo features reported by humans could be used by dolphins. Both dolphins and humans can discriminate sounds that vary in intensity by about 1 dB (Evans, 1973; Green, 1993), and the frequency discrimination abilities of dolphins and humans for tonal stimuli are comparable in the range of best hearing for each species (Herman & Arbeit, 1972; Thompson & Herman, 1975; Weir et al., 1976). This means that if humans report amplitude and pitch cues in the echoes, it is reasonable to assume that dolphins also have access to these cues. However, dolphins have sharper frequency tuning curves than humans (Supin & Popov, 1995), so they may be better able to interpret frequency information in echo stimuli than the human listeners.

Temporal resolution is the ability to rapidly resolve stimuli that are close together in time (e.g., to identify two stimuli as separate events rather than a single stimulus). Popov and Supin (1997) recorded the dolphin auditory brainstem response evoked by a short gap in noise and found that the dolphins' hearing is sensitive to extremely short gap durations as compared with other animals and humans. The dolphin gap detection threshold was 0.1 ms, whereas gap detection thresholds in humans were an order of magnitude longer than dolphins (about 2.2 ms; Snell, Ison, & Frisina, 1994). At first glance, this would seem to give the dolphins an advantage over the human listeners. However, the echoes in this study were time-stretched by more than 2 orders of magnitude (× 125) to bring them into the human hearing range. Because human gap detection thresholds are only 1 order of magnitude worse than dolphins, the time-stretched echoes may have provided the humans with more time cues than the dolphin.

In this and other studies, the performance of the human listeners usually exceeded that of the dolphins. There are several methodological reasons that could explain this result. First, the humans had a reduced memory load compared with the dolphin. The humans did a simultaneous cross-modal matching task (they saw the comparisons at the same time that they listened to the sample), whereas the dolphin did a successive cross-modal matching task (he ensonified the target, then made his choice approximately 30–60 s later). Other reasons involve the ease of processing the echoic sample. The echo trains presented to the human participants had a high signal-to-noise ratio, and they did not contain any extraneous echoes that had to be filtered out (e.g., echoes from the apparatus in which the objects were hung, echoes from passing fish). In addition, there was consistency between echo samples for a single object: Each 23-echo train presented to the human listeners always consisted of the same range of target orientations (±15°). However, the dolphin's sample object echoes may have contained noise, extraneous echoes, and covered different ranges of target orientations over different trials. The dolphin also may have used different numbers of echoes from trial to trial (cf. Au, 1993). In the future, it would be useful to collect the dolphins' clicks and echoes during the task (not done in this study because of the extensive dolphin training time required) and present those echoes to the human participants so that both the dolphins and the humans base their decisions on the same amount and quality of echo information.

Human listening studies are one way of elucidating how dolphins use information in echoes to discriminate among objects. They provide an idea of the potential cues in the echoes and how the use of those cues can change with different types of objects. Although caution must be used in interpreting the results, given that dolphins might have some different auditory tools to work with than humans, it is still a worthwhile endeavor because of the relative challenge of performing experiments with dolphins. Dolphin experiments can take months to years because of the extensive training time, and there are few animals at present that can undertake this work. Dolphin experiments are very expensive, so the choice of experiments is scrutinized. In human listening studies, many experiments can be performed in a short time. Another significant advantage to using human listeners is that, unlike dolphins, they can verbally report salient acoustic cues. Human listeners can quickly identify salient combinations of echo features that permit object discrimination, which can be used to generate hypotheses that can be tested using dolphins as subjects. For example, this study suggests that the dolphin, like the humans, may have used the pattern of changes across the echo train as a cue to distinguish between objects with different shapes. It would be interesting to duplicate Experiment 2 with dolphin subjects to see how their cue use changes when they have access to echo train pattern cues (in rotating echoes) versus when they do not (in stationary echoes) for object sets that vary in size, shape, material, or texture.

Human listening studies accelerate and augment the process of learning how dolphins discriminate among objects during echolocation. This study illustrates that human performance and verbal reports of conscious decision-making strategies on problems that can also be solved by other species can be used to identify potential variables at work in determining the animals' performance. The insights gained from these human performance studies can guide the design of experiments that directly test the effects of these variables on the animals' performance.

References

American National Standards Institute. (1960). *American standard acoustical terminology.* New York: Author.

Au, W. W. L. (1993). *The sonar of dolphins.* New York: Springer.

Au, W. W. L. (2000). Echolocation in dolphins. In W. W. L. Au, A. N. Popper, & R. R. Fay (Eds.), *Hearing by whales and dolphins* (pp. 364–408). New York: Springer.

Au, W. W. L., & Martin, D. W. (1989). Insights into dolphin sonar discrimination capabilities from human listening experiments. *Journal of the Acoustical Society of America, 86,* 1662–1670.

Au, W. W. L., Schusterman, R. J., & Kersting, D. A. (1980). Sphere–cylinder discrimination via echolocation by *Tursiops truncatus.* In R. G. Busnel & J. F. Fish (Eds.), *Animal sonar systems* (pp. 859–862). New York: Plenum Press.

Bauer, G. B., & Johnson, C. M. (1994). Trained motor imitation by bottlenose dolphins (*Tursiops truncatus*). *Perceptual and Motor Skills, 79,* 1307–1315.

DeLong, C. M., Au, W. W. L., Lemonds, D. W., Harley, H. E., & Roitblat, H. L. (2006). Acoustic features of objects matched by an echolocating bottlenose dolphin. *Journal of the Acoustical Society of America, 119,* 1867–1879.

DeLong, C. M., Au, W. W. L., & Stamper, S. A. (2007). Echo features used by human listeners to discriminate among objects that vary in material or wall thickness: Implications for echolocating dolphins. *Journal of the Acoustical Society of America, 121,* 605–617.

Evans, W. E. (1973). Echolocation by marine delphinids and one species of fresh-water dolphin. *Journal of the Acoustical Society of America, 54,* 191–199.

Eysenck, M., & Keane, M. T. (1995). *Cognitive psychology: A student's handbook.* Hillsdale, NJ: Erlbaum.

Fish, J. F., Johnson, C. S., & Ljungblad, D. K. (1976). Sonar target discrimination by instrumented human divers. *Journal of the Acoustical Society of America, 59,* 602–606.

Green, D. M. (1976). *An introduction to hearing.* Hillsdale, NJ: Erlbaum.

Green, D. M. (1993). Auditory intensity discrimination. In W. A. Yost, A. N. Popper, & R. R. Fay (Eds.), *Human psychophysics* (pp. 13–55). New York: Springer.

Harley, H. E., Putman, E. A., & Roitblat, H. L. (2003, August 7). Bottlenose dolphins perceive object features through echolocation. *Nature, 424,* 667–669.

Helweg, D. A., Au, W. W. L., Roitblat, H. L., & Nachtigall, P. E. (1996). Acoustic basis for recognition of aspect-dependent three-dimensional targets by an echolocating bottlenose dolphin. *Journal of the Acoustical Society of America, 99,* 2409–2420.

Helweg, D. A., Roitblat, H. L., Nachtigall, P. E., Au, W. W. L., & Irwin, R. J. (1995). Discrimination of echoes from aspect-dependent targets by a bottlenose dolphin and human listeners. In R. A. Kastelein, J. A. Thomas, & P. E. Nachtigall (Eds.), *Sensory systems of aquatic mammals* (pp. 129–136). Woerden, The Netherlands: De Spil Publishers.

Herman, L. M., & Arbeit, W. R. (1972). Frequency difference limens in the bottlenose dolphin: 1–70 KC/S. *Journal of Auditory Research, 2,* 109–120.

Johnson, C. S. (1967). Sound detection thresholds in marine mammals. In W. N. Tavolga (Ed.), *Marine bio-acoustics* (pp. 247–260). New York: Pergamon Press.

Kellogg, W. N. (1962, August 10). Sonar system of the blind. *Science, 137,* 399–404.

Murchison, A. E. (1980). Detection range and range resolution of porpoise. In R. G. Busnel & J. F. Fish (Eds.), *Animal sonar systems* (pp. 43–70). New York: Plenum Press.

Nachtigall, P. E., Murchison, A. E., & Au, W. W. L. (1980). Cylinder and cube discrimination by an echolocating blindfolded bottlenose dolphin. In R. G. Busnel & J. F. Fish (Eds.), *Animal sonar systems* (pp. 945–947). New York: Plenum Press.

Popov, V. V., & Supin, A. Ya. (1997). Detection of temporal gaps in noise in dolphins: Evoked potential study. *Journal of the Acoustical Society of America, 102,* 1169–1176.

Rice, C. E. (1967, February 10). Human echo perception. *Science, 155,* 656–664.

Roitblat, H. L., Penner, R. H., & Nachtigall, P. E. (1990). Matching-to-sample by an echolocating dolphin. *Journal of Comparative Psychology: Animal Behavior Processes, 16,* 85–95.

Snell, K. B., Ison, J. R., & Frisina, D. R. (1994). The effects of signal frequency and absolute bandwidth on gap detection in noise. *Journal of the Acoustical Society of America, 96,* 1458–1464.

Supin, A. Ya., & Popov, V. V. (1995). Frequency tuning and temporal resolution in dolphins. In R. A. Kastelein, J. A. Thomas, & P. E. Nachtigall (Eds.), *Sensory systems of aquatic mammals* (pp. 95–110). Woerden, The Netherlands: De Spil Publishers.

Thompson, R. K. R., & Herman, L. M. (1975). Underwater frequency discrimination in the bottlenose dolphin (1–140 kHz) and the human (1–8 kHz). *Journal of the Acoustical Society of America, 57,* 943–948.

Weir, C., Jesteadt, W., & Green, D. (1976). Frequency discrimination as a function of frequency and sensation level. *Journal of the Acoustical Society of America, 61,* 178–184.

Xitco, M. J., Jr. (1996). *Referential pointing by bottlenose dolphins.* Unpublished doctoral dissertation, Southern Methodist University, Dallas, TX.

Xitco, M. J., Jr., & Roitblat, H. L. (1996). Object recognition through eavesdropping: Passive echolocation in bottlenose dolphins. *Animal Learning and Behavior, 24,* 355–365.

Received January 16, 2006
Revision received April 10, 2007
Accepted April 30, 2007 ■