JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2002, **78,** 449–465 NUMBER 3 (NOVEMBER)

SEA LIONS AND EQUIVALENCE: EXPANDING CLASSES BY EXCLUSION

COLLEEN REICHMUTH KASTAK AND RONALD J. SCHUSTERMAN

LONG MARINE LABORATORY,
UNIVERSITY OF CALIFORNIA SANTA CRUZ

Experiments have shown that human and nonhuman subjects are capable of performing new arbitrary stimulus–stimulus relations without error. When subjects that are experienced with matching-to-sample procedures are presented with a novel sample, a novel comparison, and a familiar comparison, most respond by correctly selecting the novel comparison in the presence of the new sample. This exclusion paradigm was expanded with two California sea lions that had previously formed two 10-member equivalence classes in a matching-to-sample procedure. Rather than being presented with a novel sample on a given trial, the sea lions were presented with a randomly selected familiar member of one class as the sample. One of the comparisons was a randomly selected familiar member of the alternative class, and the other was a novel stimulus. When required to choose which comparison matched the sample, the subjects reliably rejected the familiar comparison, and instead selected the unfamiliar one. Next, the sea lions were presented with transfer problems that could not be solved by exclusion; they immediately grouped the new stimuli into the appropriate classes. These findings show that exclusion procedures can rapidly generate new stimulus relations that can be used to expand stimulus classes.

Key words: exclusion, fast mapping, equivalence, symmetry, differential outcomes, class-specific reinforcement, California sea lions

In the context of human language learning (semantics), appropriate responding can be facilitated by presenting new problems in the context of familiar alternatives. For example, if a child is asked "Which one is the *pafe*?" he or she may examine an array of familiar, already named objects and then select the novel item. This phenomenon, called *fast mapping*, *linguistic inference*, or the *disambiguation effect* in the field of psycholinguistics, is central to the development of language. In behavior analysis and animal cognition, this type of errorless performance is known as *exclusion* or *emergent matching* (Wilkinson, Dube, & McIlvane, 1998).

Dixon (1977) coined the term *exclusion* in a study of word learning in mentally retarded adolescents. Through reinforcement training, she taught her subjects a single response rule: When presented with a stimulus array that included the Greek letters Π and either Θ or Υ, select Π conditionally upon the spoken word *pi*. Following this training, she paired the same sets of alternatives with two new spoken words. Dixon found that when Π and Θ were paired as potential choices, the children immediately selected Θ as the match to the new word *theta*. Likewise, when Π and Υ were paired as potential choices, the children immediately selected Υ as the match to the new word *upsilon*. Because the words *theta* and *upsilon* deviated from the familiar word *pi*, the subjects apparently "responded away from" or excluded the previously positive Π and instead selected the previously negative Θ or Υ. Although the finding supported an earlier study by Vincent-Smith, Bricker, and Bricker (1974), it contradicted the commonly held belief that subjects learn the positive relation between a given sample or discriminative stimulus and the correct alternative but not the negative relation between the same sample and the incorrect alternative (see,

---

Research was supported by ONR Grant N00014-99-1064 to R. J. Schusterman. A DoD AASERT Fellowship and a GAANN Fellowship provided funding for C. R. Kastak. A portion of this manuscript was presented as an invited talk to the American Psychological Association in San Francisco, August 2001. We thank Tom Zentall for constructive comments during the preparation of this article. We greatly appreciate the encouragement given to us by David Kastak throughout the experiment, and we are particularly thankful to him for his insights during the preparation of the final manuscript. We also thank the dedicated research team at the Pinniped Cognition and Sensory Systems Laboratory, especially Kirsten Jensen and Shannon Spillman.

Correspondence and reprint requests may be addressed to either author at Long Marine Laboratory, University of California at Santa Cruz, 100 Shaffer Road, Santa Cruz, California 95060 (e-mail: coll@cats.ucsc.edu or rjschust@cats.ucsc.edu).

e.g., Berryman, Cumming, Cohen, & Johnson, 1965; Carter & Werner, 1978; Skinner, 1950).

Dixon's (1977) study demonstrated that subjects performing relational learning tasks may be capable of learning not only what is correct but also what is incorrect. Subsequent efforts to replicate and expand this finding employed variants of stimulus mapping or matching paradigms, including matching to sample (MTS). Investigators eventually converged on two contrasting principles to account for exclusion performances such as those just described (see review in Wilkinson et al., 1998). Each has been identified in the fields of linguistics and behavior analysis. The first is known as the *novel name/nameless category* principle by linguists and *sample/positive comparison control* by behavior analysts. This principle indicates an active search for a novel item in the presence of a novel sample stimulus. The alternative principle is known as *mutual exclusivity* by linguists and *sample/negative comparison control* by behavior analysts. This view advocates a process of elimination; it proposes that a subject presented with a novel cue will select a novel item after evaluating and rejecting one or more familiar items as potential matches.

In addition to establishing two potential controlling principles to account for exclusion performances, research also led to the identification of two procedures that would potentially elicit responding by exclusion. The first type is consistent with the examples described earlier. Subjects are presented with a novel sample (such as a new word or novel visual pattern), and must choose between a novel comparison stimulus and a familiar, defined comparison stimulus; that is, a comparison already associated with to a different sample. If the subjects respond by exclusion, they should avoid the familiar comparison and instead relate the novel sample to the novel comparison. Several populations of human subjects have been tested in this paradigm. Most of these subjects, including those who are normally developing, those with specific language impairments, and those with mild to severe mental retardation, perform quite well in exclusion trials involving novel samples (see review in Wilkinson, Dube, & McIlvane, 1996).

In the second type of exclusion procedure, sometimes called *Exclusion II* (see, e.g., McIlvane, Munson, & Stoddard, 1988), subjects are shown a familiar sample that is already defined through conditioning to a given comparison. One of the comparison stimuli presented is novel and the other is familiar; however, the familiar comparison has been conditioned to a sample other than the one being presented. On these trials, it is also possible for subjects to respond by exclusion. By avoiding the mismatch or incongruity between the already defined sample and comparison, subjects may avoid the familiar but incompatible comparison and instead select the novel comparison. As in the first type of exclusion described, performance on these trials depends on having learned something about what is *not* correct in the presence of a given sample. Human subjects tested in this paradigm show mixed results. Overall, it appears that exclusion performances are less likely to occur when samples have a prior experimental history (McIlvane et al., 1988). A critical factor in exclusion performances of this type, however, seems to be the nature of the familiar sample. For example, Stromer and Osborne (1982) demonstrated that mildly retarded adolescents presented with familiar, defined visual patterns as samples selected novel (rather than familiar but incompatible) comparisons as the correct match to the sample. McIlvane et al., however, found that normally developing children, aged 3 years 6 months to 5 years, often did not select a novel comparison item after hearing a familiar word. Rather, they matched the familiar word to a familiar but unrelated comparison item. These results suggest that the type of defined sample presented may interfere with or influence exclusion performances. They are also consistent with the possible confounding effects of novelty avoidance or preference for a familiar stimulus.

At the same time that researchers were exploring exclusion processes in human subjects with varying degrees of language competency, researchers in the field of animal cognition were making similar observations. For example, bottlenose dolphins and California sea lions trained in artificial language tasks were shown to immediately relate novel gestural signs to ''unnamed'' (novel) objects in the presence of familiar items (Herman, Richards, & Wolz, 1984; Schusterman & Krieger, 1984). In conditional discrimination procedures involving arbitrary or identity matching, California sea lions and harbor seals errorlessly matched novel visual samples and comparisons when the alternative was a member of an established stimulus pairing (Hanggi & Schusterman, 1995; D. Kastak & Schusterman, 1994; Pack, Herman, & Roitblat, 1991; Schusterman, Gisiner, Grimm, & Hanggi, 1993). Common chimpanzees trained to perform conditional discriminations also related novel samples to new comparisons in the contexts of language training or conditional discrimination procedures (see reviews in Cerutti & Rumbaugh, 1993; Tomonaga, 1993). Indirect evidence from pigeons tested in variants of exclusion procedures suggests that smaller brained animals may be capable of exclusion performances as well (Zentall, Edwards, Moore, & Hogan, 1981).

Only a few studies have evaluated the abilities of nonhuman animals to respond by exclusion when samples have been previously defined. Most notably, Tomonaga (1993) showed that a chimpanzee, when presented with matching problems involving stimuli differing in color and shape, often related familiar samples to novel comparisons in the presence of familiar but incompatible alternatives. Tomonaga's finding with a chimpanzee supports that of Stromer and Osborne (1982), who described similar performances for mildly retarded adolescents who were also presented with Exclusion II trials involving visual stimuli.

Early in the history of research into exclusion, it had been suggested that control by negative stimuli might depend on linguistic skills including a highly developed speech repertoire (Dixon, 1977). The successful exclusion performances just described for nonhuman animals, along with documentation of exclusion performances exhibited by profoundly retarded nonverbal individuals (see review in McIlvane et al., 1987), suggest, however, that language is not necessary for exclusion.

Learning Outcomes

Once a problem is solved by exclusion—regardless of the rule underlying the response, the type of exclusion performance, or the subject performing the response—it is still unknown whether the subject has learned an explicit relation between the newly related sample and comparison. The use of a general exclusion strategy does not require that anything new to be learned about the relations between the stimuli involved in the problem. One way of determining if a *learning outcome* has resulted from exposure to one or more exclusion trials is to present the new discrimination in the presence of novel items rather than familiar ones. If the conditional discrimination is maintained when responding by exclusion is prevented, then a learning outcome has been achieved.

Although learning outcomes do not always result from exclusion performances, they have been shown to occur. For example, Dixon (1977) found that her subjects, who had spontaneously matched Θ and Υ to novel words such as *theta* and upsilon by exclusion, did not immediately show a learning outcome; rather, their matching performance deteriorated when the familiar Π was not present as a comparison stimulus. After a history of reinforced exclusion trials, however, her subjects ultimately did show a learning outcome; they appropriately matched the words *theta* and *upsilon* to Θ and Υ in the absence of the familiar comparison stimulus Π.

Consider the implications of an individual who fails to show a learning outcome following successful exclusion performances. Such an event suggests that the individual is not learning the explicit relation between the sample and positive comparison during the exclusion trials; rather, the individual is relying only on the incongruency between the sample and the negative comparison. If a subject can successfully perform a new conditional discrimination by excluding the incorrect alternative, then to what extent is the subject learning about the relation between the sample and the correct alternative? The question of whether learning outcomes are likely to occur following successful exclusion performances was addressed by McIlvane et al. (1988) in a series of experiments with normally capable children. These investigators concluded that exclusion performances did not necessarily predict subsequent learning outcomes, even in subjects with fairly well-developed language skills. Conversely, a few studies have documented learning outcomes resulting from exclusion performances in mildly to severely retarded individuals (see review in Wilkinson et al., 1998). In situations in which learning outcomes were shown to occur, the number of exclusion trials required to produce a learning outcome varied. Young children sometimes identified the referent of a new word following a single exclusion trial (Carey & Bartlett, 1978) but other subjects, such as Dixon's (1977), required extended histories of reinforced exclusion trials prior to showing learning outcomes.

As described earlier, learning outcomes that result from exclusion procedures are revealed by the stability of the new conditional discrimination in the absence of familiar alternatives. In addition to these straightforward learning outcomes, exclusion procedures may also establish symbolic (equivalence) relations between novel sample and comparison stimuli (an outcome termed *emergent symbolic mapping*; see Wilkinson et al., 1996). Successful demonstrations of symbolic learning outcomes occur when subjects, following exposure to a new conditional discrimination in the context of exclusion, later show that the conditional discrimination that is learned has the property of symmetry, or bidirectionality. To build on a prior example, a child who learns to select an unnamed item from an array of familiar alternatives upon hearing the word *pafe* might later produce the word *pafe* when presented with the same item. This outcome has obvious implications for the learning of symbolic referents such as words (see, e.g., Samuelson & Smith, 2000). Furthermore, because symmetry itself is a property of stimulus equivalence (see Sidman, 1994), it is possible that exclusion performances may influence the formation and structure of equivalence classes, which are generally considered to be fundamental components of language.

The perceived and potential relations among exclusion performances, learning outcomes, and concepts that are closely tied to language development make the study of exclusion in language-deficient subjects a potentially fruitful area of research. As noted earlier, exclusion performances have been described for some nonhuman animals; however, learning outcomes have not yet been systematically evaluated in these subjects. There is some preliminary evidence to suggest that animals trained in language comprehension and conditional discrimination tasks require extensive histories of reinforced exclusion trials prior to the emergence of learning outcomes. For example, Schusterman et al. (1993) and Schusterman and Kastak (1995) reported that California sea lions that performed new conditional discriminations by exclusion required several hundred reinforced exclusion trials prior to achieving a stable, unidirectional learning outcome. Although the sea lions learned the new discriminations with far fewer errors in the exclusion context, the total amount of experience required to achieve a learning outcome was not significantly less than the amount required to learn new discriminations in a trial-and-error context. Symbolic or symmetrical learning outcomes derived from exclusion performances have not yet been studied in animals, which tend to have difficulty with spontaneous bidirectional performances in other contexts (see partial review in Zentall, 1998). Further research is required to characterize the learning outcomes that result from the exclusion performances of verbally limited humans and nonhuman animals.

Exclusion and Equivalence Classification

All of the exclusion experiments that we know of used a few sets of conditional discriminations to assess exclusion performance and subsequent learning outcomes, or transfer performance. We expanded these procedures by introducing exclusion problems in the context of large stimulus classes, in which new stimuli, presented with familiar stimuli in conditional discrimination trials, could be selected or excluded on the basis of class-consistent relations between the sample and positive comparison or class-inconsistent relations between the sample and negative comparison. In a previous experiment with 2 California sea lions, two 10-member equivalence classes emerged using a simple-discrimination repeated-reversal procedure and class-specific reinforcement (C. R. Kastak, Schusterman, & Kastak, 2001). Once functional classes were formed in the reversal procedure, as evidenced by a transfer of response from one member to all members of a given class, class membership immediately transferred to conditional discriminations in a matching-to-sample procedure. The conditional discriminations that emerged included all of the possible stimulus pairings within each class but not between classes. Finally, the classes were expanded through emergent equivalence relations to include (a) new stimuli that had been conditioned to at least one existing class member and (b) new stimuli that had been conditioned to the same reinforcer assigned to a given class. These experiments demonstrate emergent equivalence classification by sea lions, as shown by the transfer of responses that occurred among physically dissimilar stimuli related by common functional and reinforcer relations.

The results of this work on equivalence classification with 2 California sea lions, combined with prior successful laboratory demonstrations of exclusion (Schusterman et al., 1993), suggested that the large equivalence classes formed by the sea lions might be expanded to include several new members introduced by exclusion. This hypothesis was tested in the current experiment. Several questions were addressed: Would the sea lions spontaneously exhibit exclusion performances by responding to novel stimuli in trials in which defined sample and comparison stimuli were incompatible with respect to class membership? Would reinforced exposures to such exclusion trials establish learning outcomes? If so, how rapidly could learning outcomes be achieved? And finally, if learning outcomes did occur, would they comprise basic conditional discriminations or include more complex symmetrical relations?

METHOD

Subjects

The subjects were 2 captive female California sea lions (*Zalophus californianus*) named Rio and Rocky, 16 years old and 24 years old, respectively. Both were housed outdoors at Long Marine Laboratory at the University of California, Santa Cruz. The sea lions were tested individually in an enclosed section of the facility that contained a saltwater pool (7.5 m diameter) and a deck on which the experimental apparatus was placed. Each sea lion was fed between 5 and 7 kg of freshly thawed cut herring and capelin each day, one half of which was typically consumed during experimental sessions.

Both subjects were experienced with visual discrimination tasks including three-term contingencies (simple discriminations) and four-term contingencies (conditional discriminations). Rio had previously demonstrated generalized identity matching (D. Kastak & Schusterman, 1994) and had participated in a variety of arbitrary matching tasks (see Schusterman, Kastak, & Kastak, 2002). She had passed tests for stimulus equivalence by showing emergent transitive and symmetrical relations (Schusterman & Kastak, 1993) and had transferred those equivalence classes from MTS to sequential simple discriminations (Schusterman & Kastak, 1998). Rocky had a similar history, with a few exceptions. Unlike Rio, Rocky had extensive training in an artificial gestural language (see Gisiner & Schusterman, 1992; Schusterman & Krieger, 1984). Like Rio, she was tested in identity and arbitrary matching tasks, and although she demonstrated generalized identity matching in the context of MTS (D. Kastak & Schusterman, 1994), she did not demonstrate equivalence, including emergent symmetry, when tested with the same procedure (unpublished data). Both subjects showed similar results in their most recent experiment, which documented the development of 10-member functional classes that also comprised equivalence classes (C. R. Kastak et al., 2001). A key feature of this recent experimental history was the substitutability of the functional class members as samples and comparisons, which created omnidirectional discriminations among all the members of a given class. Because Rocky and Rio were familiar with the apparatus and general procedure, no new training was required to begin the current experiment.

Apparatus

A visual two-choice MTS apparatus was used (first introduced to the subjects in 1988; see Schusterman et al., 1993). The apparatus is shown in Figure 1. It consisted of a set of three horizontally arranged plywood panels (120 cm tall); the center panel was 120 cm wide and the two side panels were 60 cm wide. Each panel contained an enclosed stimulus presentation box that was 30 cm by 30 cm square and 10 cm deep. The center (sample) box was positioned 60 cm in front of a T-bar station, and the two side (comparison) boxes were angled toward the station such

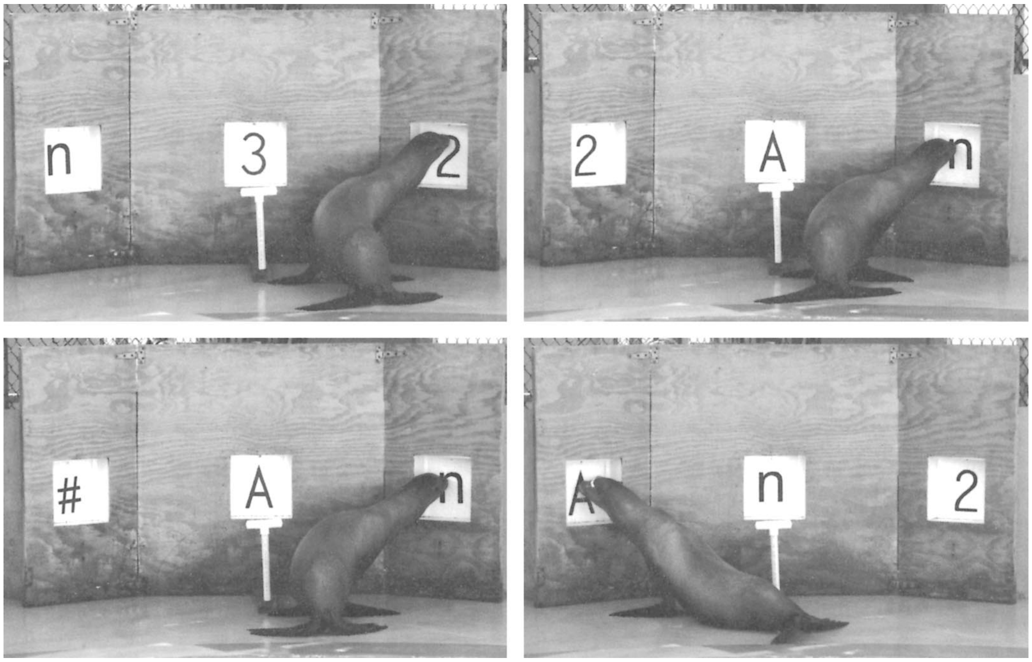

Fig. 1. The subject performing in front of the apparatus. The top two photos depict the two trial types presented in Phase 1 (exclusion). In the novel negative trial (upper left), Rocky responds to the class-consistent sample/S+ pairing and avoids the novel stimulus n. In the novel positive trial (upper right), Rocky responds to the class-inconsistent sample/S− pairing and selects n. The lower two photos show the two types of trials presented in Phase 2 (transfer). In the novel comparison trial (lower left), Rocky selects n over the other novel stimulus in the presence of A; in the novel sample trial (lower right) she performs the symmetrical relation, selecting A in the presence of n.

that each was 110 cm away from the station. All of the stimulus boxes were positioned at the eye level of the subjects, and a movable opaque door covered the front of each box. Acoustic cues used during testing were played through a speaker that was mounted approximately 3 m from the apparatus.

The stimuli used in the experiment were plywood squares (30 cm by 30 cm) with black shapes painted on white backgrounds. All of the stimuli used are shown in Figure 2. Twenty stimuli were divided into two subsets of 10 that were coded categorically as letters and numbers. These stimuli and classes were familiar to the subjects (C. R. Kastak et al., 2001). Six additional novel stimuli were presented in the current experiment. These six stimuli were divided into three pairs that were

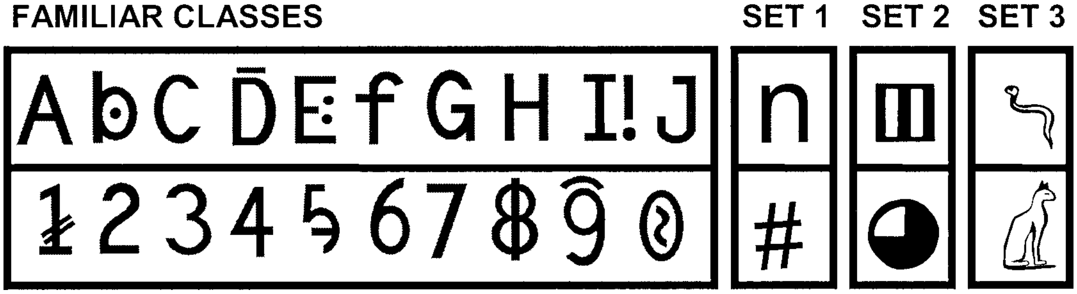

Fig. 2. The top row shows stimuli designated as letters and the bottom row shows stimuli designated as numbers. The familiar classes were established prior to the current experiment; Sets 1, 2, and 3 were introduced according to the steps shown in Figure 3. To avoid experimenter error, the assignment of stimuli to each category was the same for both subjects.

labeled Set 1, Set 2, and Set 3. The stimuli in each of these sets were configured to be roughly equal in area and brightness and were designed to be discriminable from one another as well as from the stimuli in the letter and number stimulus sets.

Experimental sessions were directed by an experimenter who observed the session in real time on video in a separate indoor room. The outdoor apparatus was operated by two assistants who were positioned behind the wooden panels. Prior to each session, a gate was opened to provide the subject access to the enclosure. At the start of the session, one of the assistants stood up and called the subject's name, and then presented a trained cue (the verbal signal ''station'' combined with a visual pointing gesture) to position the sea lion at the T-bar station in front of the apparatus. From this point until the end of the session, both assistants were seated behind the wooden panels, where they could not observe the subject or the stimuli presented on any trial. The assistants received operational instructions via headphones from the experimenter. On each trial, the experimenter instructed the assistants to select certain stimuli and place them in the appropriate stimulus boxes. Stimuli were always placed in the boxes simultaneously, so that the subject could not be cued to the correct choice by the timing of its placement. A trial began when the experimenter asked for the sample stimulus to be uncovered. After an observation interval of about 4 s, the comparison stimuli on either side were revealed. The sea lion was given 2 to 4 s to observe all three stimuli, and then a release tone played through the speaker cued the subject to respond. At this point, the sea lion moved from the center station to touch one of the two comparison stimuli; the subject was trained to hold its position on the selected stimulus until feedback was given. Correct selections were marked by a 0.5-s pure tone that served as a conditioned reinforcer. The tone was followed by a piece of fish given to the subject from behind the apparatus. Incorrect selections were marked by the vocal signal ''no,'' and reinforcement was not provided on these trials. All of the acoustic cues were triggered by the experimenter.

Procedure

Each sea lion usually completed two sessions per day, 5 days per week. All sessions were recorded on videotape. Sessions consisted of 40 conditional discrimination trials that were separated by 10- to 15-s intertrial intervals. During all sessions, the probability of the correct choice appearing in either the left or right stimulus box was .5, and no more than four left or four right trials appeared in succession. The likelihood of the correct choice appearing on the same side or the alternate side as the previous trial was also .5. Each session contained a unique sequence of trials. On each trial, the sample and the positive comparison (S+) were stimuli from the same stimulus class, and the negative comparison (S−) was a stimulus from the other class. When presented with a given sample, the subject was rewarded for selecting the S+, the stimulus that belonged to the same class as the sample. For example, A might appear as the sample, E might appear as the S+, and 3 might appear as the S−. On this trial, selection of E would be reinforced. This general procedure of reinforcing matches within and not between the letter and number classes was continued throughout the experiment.

Feedback was provided on every trial in every phase of the experiment. As in the previous study (C. R. Kastak et al., 2001), all correct responses were rewarded with class-specific reinforcement. For Rocky, correct matches within the letter class produced a 293-Hz tone followed by herring; correct matches within the number class produced a 587-Hz tone followed by capelin. For Rio, the opposite outcomes were used to reinforce correct responses.

The experiment consisted of baseline training and two test phases as depicted in Figure 3 and described in the following sections. During baseline training, the subjects were presented with familiar categorical discrimination problems. This training was conducted to ensure that the subjects were performing the MTS task and that the previously established letter and number categories were intact. In the exclusion phase, the members of the new stimulus sets were introduced to the subjects in familiar–novel conditional discriminations. In these problems, a defined sample appeared with a novel comparison and a defined comparison. These problems could potentially be solved by exclusion on the basis of class-consistent or class-inconsistent relations between the two defined stim-

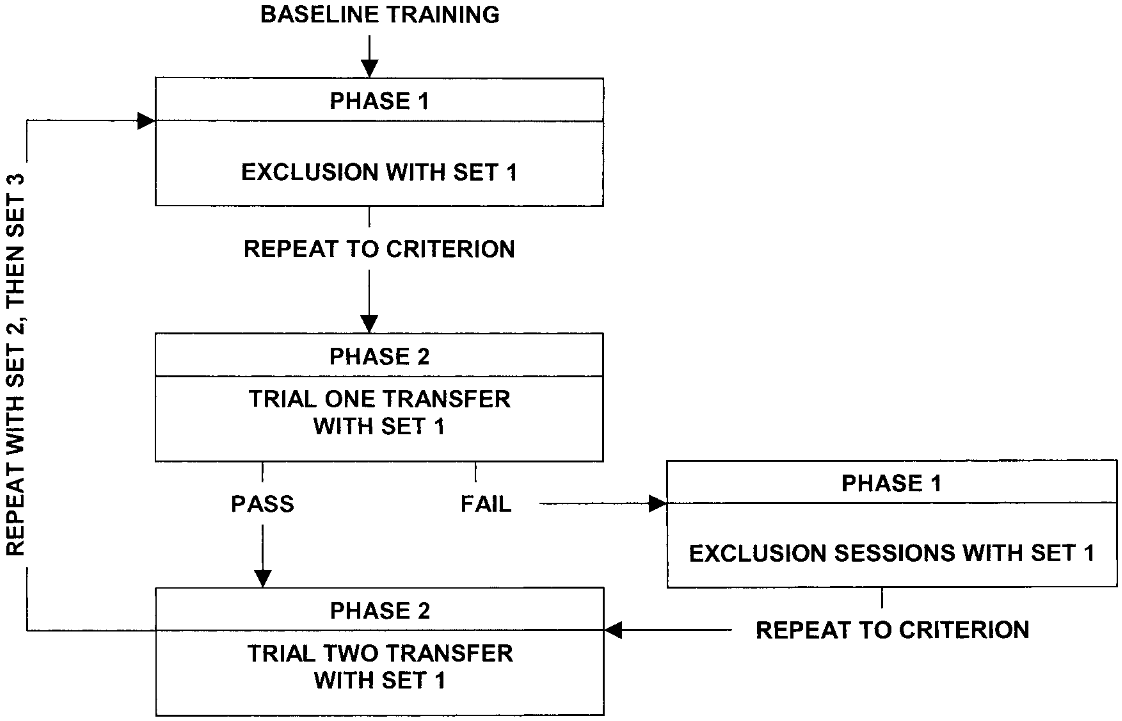

Fig. 3. Schematic of the procedure, showing the progression from baseline training with familiar classes to Phase 1 exclusion tests to Phase 2 transfer tests. Following the completion of the steps with Set 1, the sequence was repeated with Set 2 and then Set 3.

uli. In the transfer (learning outcome) phase, the subjects were presented with conditional discrimination problems that combined the new and familiar stimuli into trials that could not be solved by exclusion; these trials were used to determine whether the new stimuli had been incorporated into the existing categories following their exposure in exclusion trials.

Baseline training. Baseline trials consisted of combining the 20 familiar letter and number stimuli into 180 different conditional discriminations, with all stimuli having an equal likelihood of appearing as sample, S+, and S−. The sample and S+ were always from one class, and the S− was from the other class. All sessions contained an equal number of letter-positive and number-positive trials. Both subjects completed at least 10 baseline sessions prior to beginning the exclusion phase. In previous experiments, a standard acquisition–performance criterion of 90% was used for both subjects. For the current experiment, Rocky's criterion was lowered to 85% based on her average baseline performance; Rio's criterion remained at 90%. These acquisition criteria were used in the exclusion phase as described in the following section.

Phase 1: Exclusion. The aim of this phase was to determine whether the subjects would immediately relate new stimuli to the members of the familiar categories when presented with exclusion trials. At the start of this phase, the stimuli in Set 1 (n and #) were each assigned to one of the original categories; n was assigned to the letter class and # was assigned to the number class. The new stimuli were introduced to each subject in sessions composed of two types of trials as shown in the top photos of Figure 1. In novel negative trials, one of the new stimuli appeared as the S− with two familiar stimuli from the opposing class appearing as sample and S+. In novel positive trials, one of the new stimuli appeared as S+, with a familiar stimulus from the same class appearing as sample and a familiar stimulus from the opposing class appearing as S−. The novel negative trials, which presented familiar sample/S+ configurations, served as control trials. These trials

Table 1

Trial configurations used in Session 1 of Phase 1 (exclusion) with Stimulus Set 1. There are 40 unique sample/S+ combinations: 20 are novel negative and 20 are novel positive. The trial configurations used in Session 1 with Stimulus Sets 2 and 3 were established in the same way.

Novel negative						Novel positive					
Letter positive			Number positive			Letter positive			Number positive		
Sample	S+	S−	Sample	S+	S−	Sample	S+	S−	Sample	S+	S−
C	A	#	5	1	n	A	n	3	1	#	G
F	B	#	7	2	n	B	n	9	2	#	B
I	C	#	9	3	n	C	n	1	3	#	D
C	D	#	2	4	n	D	n	7	4	#	C
B	E	#	8	5	n	E	n	4	5	#	H
H	F	#	3	6	n	F	n	10	6	#	F
E	G	#	6	7	n	G	n	2	7	#	J
D	H	#	10	8	n	H	n	6	8	#	A
G	I	#	1	9	n	I	n	8	9	#	E
A	J	#	4	10	n	J	n	5	10	#	I

allowed the presentation of the new stimuli to be balanced in their appearance as S+ and S−, and they also served to determine whether neophobia (avoidance of new stimuli) or neophilia (attraction to new stimuli) influenced responding in this phase. The true exclusion trials were the novel positive trials, which presented novel sample/S+ configurations.

In each session, there were 40 unique trials. Of these, 20 were novel negative trials and 20 were novel positive trials. On the novel negative trials, n appeared 10 times as the S−, with 10 different sample/S+ number combinations. Stimulus # appeared 10 times as the S−, with 10 different sample/S+ letter combinations. On the novel positive trials, n appeared 10 times as the S+, with each of the familiar letters appearing as sample once and each of the familiar numbers appearing as S− once. Stimulus # appeared 10 times as the S+, with each of the familiar numbers appearing as sample once and each of the familiar letters appearing as S− once.

Session 1 of this phase comprised Trial 1 of exclusion testing; that is, it included the first exposure of each of the unique trial configurations involving n and #. For clarification, the 40 individual trial combinations presented in this session are shown in Table 1. In Session 2, the same 40 sample/S+ combinations were presented in a different sequence, this time with different negative stimuli and on alternate sides. Thus, Session 2 comprised Trial 2 of the exclusion test. Each successive session represented a single additional exposure of each of the 40 different exclusion trials involving n and #. The subjects were presented with these exclusion sessions until they reached criterion levels of performance on either one session or two consecutive sessions. At this time, they proceeded to the next phase of testing as shown in Figure 3.

Phase 2: Transfer. The transfer phase was conducted to determine if the subjects' experience with the new stimuli in exclusion trials had established learning outcomes. There were three types of trials presented in this phase. Baseline trials were conditional discriminations consisting of familiar letter and number stimuli as in baseline training. Novel comparison trials presented both n and # as comparisons, with a familiar letter or number appearing as the sample. Novel sample trials presented either n or # as the sample, with one familiar letter and one familiar number as comparisons. The two types of transfer trials are shown in the bottom photos of Figure 1. It is important to note that the transfer problems could not be solved on the basis of exclusion; rather, these trials required the subjects to relate the new stimuli directly to members of the familiar categories.

Phase 2 included two transfer tests. The Trial 1 transfer test included 20 novel sample trials, 20 novel comparison trials, and 20 baseline trials in a random sequence. The 60 trials were presented in two 30-trial sessions that were run consecutively with a short break in

Table 2

Trial configurations used in the Trial 1 session of Phase 2 (transfer) with Stimulus Set 1. There are 40 unique sample/S+ combinations: 20 are novel comparison and 20 are novel sample. The trial configurations used in the Trial 1 session with Stimulus Sets 2 and 3 were established in the same way.

Novel comparison						Novel sample					
Letter positive			Number positive			Letter positive			Number positive		
Sample	S+	S−	Sample	S+	S−	Sample	S+	S−	Sample	S+	S−
1	n	#	A	#	n	n	A	10	#	1	J
2	n	#	B	#	n	n	B	3	#	2	E
3	n	#	C	#	n	n	C	4	#	3	A
4	n	#	D	#	n	n	D	8	#	4	F
5	n	#	E	#	n	n	E	9	#	5	D
6	n	#	F	#	n	n	F	2	#	6	I
7	n	#	G	#	n	n	G	6	#	7	C
8	n	#	H	#	n	n	H	1	#	8	G
9	n	#	I	#	n	n	I	7	#	9	H
10	n	#	J	#	n	n	J	5	#	10	B

between. There were an equal number of letter-positive and number-positive trials in the test. Ten of the novel sample trials presented n as sample and 10 presented # as sample; the members of each familiar class each appeared once as S+ and once as S− during these trials. The remaining 20 novel comparison trials each presented n and # as comparisons; each of the familiar letters and numbers appeared once as sample on these trials. Correct responses were defined as class-consistent responses, for example, selection of a letter rather than a number upon presentation of n as sample. For clarification, the 40 test trials presented in the Trial 1 transfer test are shown in Table 2.

Performance on the first exposure of the transfer problems was evaluated to determine if each subject passed or failed. If the subject passed the Trial 1 transfer test, a Trial 2 transfer test was conducted the following day. If the subject failed the Trial 1 transfer test, exclusion sessions were repeated until the occurrence of two consecutive criterion sessions, and then the Trial 2 transfer test was conducted. The Trial 2 test replicated the Trial 1 test, with 20 different baseline trials and the same 40 transfer trials presented in a different order and in different configurations (with a different S−, if possible, and on alternate sides).

Additional stimulus sets. Following the completion of Phases 1 and 2 with Set 1, n and # were eliminated. The entire sequence was then repeated with Set 2 and Set 3, as shown in Figure 2. The trial configurations used with Sets 2 and 3 during the exclusion phase and the transfer phase also included all possible stimulus combinations; these trials can be inferred from Tables 1 and 2.

Data Recording and Analysis

Responses were recorded in real time by the experimenter during all sessions. Reliability assessment was conducted by two independent observers who later scored a random sample of four sessions from videotape. The two observers and the experimenter were in agreement on 99% on the trials sampled. The subjects' performance was measured as the number of correct responses out of the total number of trials completed. These data were broken down by experimental condition and evaluated using one-tailed binomial tests to determine whether performance was better than predicted by chance. This test also served as the pass–fail measure during transfer testing. Differences in performance between experimental conditions were evaluated with two-tailed Fisher's exact tests. All statistical tests were evaluated at alpha levels of .05.

RESULTS

During the 10 baseline sessions completed prior to testing, Rio averaged 98% correct responses and Rocky averaged 86% correct re-

Table 3

Phase 1 performance for each subject by trial type: novel negative, novel positive, and total exclusion (novel negative + novel positive). Performance is shown as the number of correct responses out of the total number of trials completed. The first row shows results for Trial 1 of exclusion, the second row shows Trial 2, and so on. Horizontal lines in the data columns indicate progression to transfer tests. Double lines indicate failure of Test 1 and return to exclusion sessions. Categories in which performance is significantly better than predicted by chance are shown in boldface.

		Rio			Rocky		
Session	Trial type	Set 1	Set 2	Set 3	Set 1	Set 2	Set 3
1	Novel negative	**20/20**	**19/20**	**19/20**	**17/20**	**17/20**	**17/20**
	Novel positive	14/20	14/20	**15/20**	**16/20**	**15/20**	12/20
	Total	**34/40**	**33/40**	**34/40**	**33/40**	**32/40**	**29/40**
2	Novel negative	**19/20**	**15/20**	**20/20**	14/20	**18/20**	**15/20**
	Novel positive	**18/20**	**18/20**	**16/20**	**17/20**	**16/20**	**16/20**
	Total	**37/40**	**33/40**	**36/40**	**31/40**	**34/40**	**31/40**
3	Novel negative		10/20	**19/20**	**15/20**	**16/20**	**18/20**
	Novel positive		**20/20**	**18/20**	11/20	**15/20**	**16/20**
	Total		**30/40**	**37/40**	**26/40**	**31/40**	**34/40**
4	Novel negative		11/20	**18/20**	**16/20**	**20/20**	11/20
	Novel positive		**18/20**	**20/20**	**15/20**	**20/20**	**16/20**
	Total		**29/40**	**38/40**	**31/40**	**40/40**	**27/40**
5	Novel negative		10/20	**17/20**	**17/20**		**15/20**
	Novel positive		**20/20**	**18/20**	**16/20**		**19/20**
	Total		**30/40**	**35/40**	**33/40**		**34/40**
6	Novel negative		**15/20**	**19/20**	**18/20**		**18/20**
	Novel positive		**19/20**	**20/20**	**17/20**		**20/20**
	Total		**34/40**	**39/40**	**35/40**		**38/40**
7	Novel negative		**17/40**	**19/20**	**19/20**		**18/20**
	Novel positive		**20/20**	**19/20**	**15/20**		**17/20**
	Total		**37/40**	**38/40**	**34/40**		**35/40**
8	Novel negative		**18/20**				**18/20**
	Novel positive		**20/20**				**15/20**
	Total		**38/40**				**33/40**
9	Novel negative						**19/20**
	Novel positive						**17/20**
	Total						**36/40**
10	Novel negative						**18/20**
	Novel positive						**17/20**
	Total						**35/40**

sponses. Neither subject preferred left or right responding. Rio had no bias towards responding to letters or numbers (she was correct on 195 of 200 letter-positive trials vs. 199 of 200 number-positive trials; Fisher's exact test, $p > .05$). Rocky, however, tended to preferentially respond to letters over numbers (she was correct on 183 of 199 letter-positive trials vs. 161 of 200 number-positive trials; Fisher's exact test, $p < .05$). Rocky's performance was analyzed to determine whether her errors were evenly distributed among the numbers or whether they were stimulus specific. Her errors tended to be clustered to individual stimuli when they appeared as samples. For example, Rocky had the most difficulty when 10 appeared as the sample; she made triple the number of errors expected from a rectilinear error distribution. Despite Rocky's difficulty with certain numbers as samples, she maintained stable baseline performance with the complete stimulus sets and proceeded to Phases 1 and 2 with a reduced acquisition criterion of 85% correct responses.

The performance of both subjects during each session of Phase 1 (exclusion) is shown in Table 3. The top rows of data (Session 1)

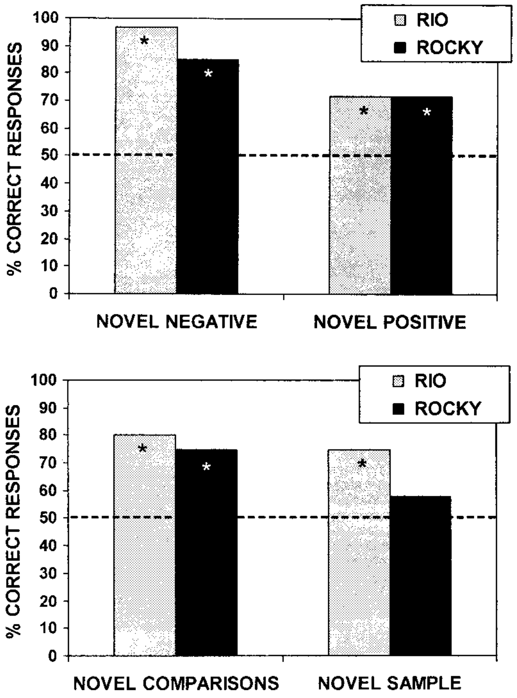

Fig. 4. Top: Trial 1 performance of both subjects during the Phase 1 exclusion tests. Each bar consists of 60 unique trial configurations involving the new stimuli introduced either in novel negative (control) trials or novel positive (exclusion) trials. Chance performance is marked by the dashed line at 50% correct responses, and categories for which performance is significantly better than predicted by chance are denoted with a star. Bottom: Trial 1 performance of both subjects during the Phase 2 transfer tests. Each bar consists of 60 unique trial configurations involving the new stimuli introduced either in novel comparisons (learning outcome) trials or novel sample (symmetry outcome) trials. Chance performance is marked by the dashed line at 50% correct responses, and categories for which performance is significantly better than predicted by chance are denoted with a star.

show performance on the first exposure of the 40 different exclusion trials with each set of new stimuli. Rio and Rocky performed better than expected by chance with every set of novel stimuli presented in the exclusion phase. When Trial 1 performance was pooled for Sets 1, 2, and 3, Rio was correct a total of 80% of the time (101 correct responses on a total of 120 unique exclusion trials); Rocky was correct 78% of the time (94 of 120 exclusion trials; binomial tests, $p < .05$ for both subjects). When those trials were broken down by type, as shown in the top graph of Figure 4, both subjects performed better on trials in which the novel stimulus was negative (control trials) than when the novel stimulus was positive. Rio scored 96% on novel negative trials and 71% on novel positive trials; Rocky scored 85% on novel negative trials and 71% on novel positive trials. Although the performance of both subjects was weaker on novel positive trials, which provided the true measure of exclusion performance, it was still much better than expected by chance (binomial tests, $p < .05$ for both subjects). After the first session of the exclusion phase, each subject was exposed to additional sessions of the same type until criterional levels of performance were established.

Rio's results for Phase 2 (transfer) are shown in the left columns of Table 4. Of the Trial 1 transfer tests with Sets 1, 2, and 3, Rio showed positive transfer in every category evaluated; that is, she responded to the newly introduced stimuli by making appropriate within-class responses more often than not. On the trials that presented familiar samples with novel comparisons, Rio's performance was better than predicted by chance on two of three transfer tests. On the trials that presented novel samples with familiar comparisons, Rio passed one of three transfer tests. When the Trial 1 transfer tests were pooled across the three new stimulus sets (see Figure 4), Rio scored 80% correct on novel comparison trials and 71% correct on novel sample trials, compared to 91% correct on familiar baseline trials. Her scores in both transfer categories are better than predicted by chance (binomial tests, $p < .05$) and are not different from one another (Fisher's exact test, $p > .05$). As data in the table show, Rio's performance in every transfer category was near perfect by the second exposure of each problem (Trial 2) and not different from baseline performance.

Rocky's results during the Phase 2 transfer tests are shown in the right columns of Table 4. Her performance was similar to Rio's, but less robust. Rocky showed positive transfer in every category tested except one. On the Trial 1 tests that presented novel comparisons, Rocky passed two of three transfer tests. On the Trial 1 tests that presented novel samples, statistically Rocky did not pass any of the three tests. When Rocky's Trial 1 test results were pooled across the three transfer tests (see Figure 4), her performance on novel

Table 4

Phase 2 performance for each subject broken down into four categories: novel comparison, novel sample, total transfer (novel comparison + novel sample) and baseline. The results for the Trial 1 test and the Trial 2 test are reported separately. Performance in each category is shown as the number of correct responses out of the total number of trials completed. Categories for which performance is significantly better than predicted by chance ($p < .05$) are shown in boldface.

		Rio		Rocky	
Set	Trial type	Trial 1	Trial 2	Trial 1	Trial 2
1	Novel comparisons	**18/20**	—[a]	**16/20**	**15/20**
	Novel sample	**17/20**	—[a]	12/20	14/20
	Total transfer	**35/40**	—[a]	**28/40**	**29/40**
	Baseline	**19/20**	—[a]	**18/20**	**17/20**
2	Novel comparisons	**19/20**	**19/20**	**19/20**	13/20
	Novel sample	14/20	**20/20**	12/20	11/20
	Total transfer	**33/40**	**39/40**	**31/40**	24/40
	Baseline	**19/20**	**19/20**	**18/20**	**15/20**
3	Novel comparisons	11/20	**20/20**	10/20	14/20
	Novel sample	12/20	**19/20**	11/20	13/20
	Total transfer	23/40	**39/40**	21/40	**27/40**
	Baseline	**17/20**	**19/20**	**19/20**	**18/20**
Total	Novel comparisons	**48/60 (80%)**	**39/40 (97%)**	**45/60 (75%)**	**42/60 (70%)**
	Novel sample	**43/60 (71%)**	**39/40 (97%)**	35/60 (58%)	**38/60 (63%)**
	Total transfer	**91/120 (75%)**	**78/80 (97%)**	**80/120 (66%)**	**80/120 (66%)**
	Baseline	**55/60 (91%)**	**38/40 (95%)**	**55/60 (91%)**	**50/80 (83%)**

[a] — indicates trial not run.

comparison trials was better than expected by chance (75% correct; binomial test, $p < .05$). On novel sample trials, however, her performance (58% correct) was not better than predicted by chance (binomial test, $p > .05$). On the pooled Trial 2 transfer tests, Rocky marginally but significantly passed both the novel comparison and novel sample trial categories, and her performance on those transfer categories was not different from her performance on baseline trials (Fisher's exact tests, $p > .05$).

Finally, during Phase 2, we observed some differences in transfer performance that correlated with a specific set of stimuli. During the exclusion phase, both subjects performed their first session at criterion between the second and seventh presentation of the 40 exclusion trials. During the transfer phase when the subjects were presented with Sets 1 and 2, they both passed the Trial 1 transfer test and immediately proceeded to the Trial 2 test. When presented with transfer problems involving Set 3, however, both subjects failed the Trial 1 transfer test and had to return to exclusion sessions before passing the Trial 2 transfer test with that stimulus set.

DISCUSSION

When presented with exclusion problems that included the new stimuli and familiar class members, both sea lions performed accurately on Trial 1. These results are consistent with earlier descriptions of marine mammals and primates performing new conditional discriminations without explicit training. As predicted, the sea lions performed better on novel negative trials (on which new stimuli had to be avoided) than on novel positive trials (on which new stimuli had to be selected), although their performance in both trial categories was significantly better than expected by chance. These trials also showed an absence of strong neophobia or neophilia. The results on novel positive trials indicated that the mechanism underlying performance was, in fact, control by the negative comparison. In these trials, the subjects rejected the defined S− as the match to the sample and instead selected the undefined S+. Because the sample was familiar, there was no opportunity for the subjects to respond based on the fact that both the sample and the comparison were novel.

Therefore, the results support the mutual exclusivity principle, or sample/negative comparison control, as the basis for the subjects' immediate accuracy on exclusion trials.

From the results of the exclusion phase of the experiment, the sea lions can be described as approaching the class-consistent relation between the sample and S+ or avoiding the class-inconsistent relation between the sample and S−. It is unknown whether the subjects responded to the class-consistent or inconsistent relations between the defined sample and comparison stimuli themselves, the class-specific reinforcers associated with each, or both. Whether the sea lions rejected defined comparisons as potential matches to defined samples on the basis of class-inconsistent stimulus relations or on the basis of class-inconsistent reinforcer relations, they clearly behaved as if using a "process of elimination" strategy (i.e., they excluded the comparison that had been conditioned to a stimulus class that was incompatible with that of the sample, or they excluded the comparison that had been conditioned to a reinforcer that was incompatible with that of the sample).

To determine if a learning outcome resulted from the sea lions' exposure to the exclusion trials, performance on novel comparison transfer trials was evaluated. These trials presented the same familiar–novel conditional discriminations used in the exclusion phase, but with alternatives that eliminated the potential for control by exclusion. Both sea lions maintained high levels of accuracy on the first exposure of these transfer trials, showing that the conditional discriminations had actually been learned during the exclusion phase. This learning occurred rapidly, following a minimum of two and a maximum of eight reinforced exposures to each of the new conditional discriminations.

The question of whether the learning outcomes generated by exclusion trials comprised the same straightforward conditional discriminations presented during the exclusion phase or whether they showed the additional property of symmetry was evaluated by performance on the novel sample transfer trials. These trials presented the symmetrical version of the conditional discriminations that had been presented in the exclusion trials (i.e., the newly introduced stimuli now served as samples for the first time, with opposing members of the familiar classes as comparison stimuli). Rio's performance revealed class-consistent matching in the context of emergent symmetry on these trials; she scored 71% correct on Trial 1 and 97% correct on Trial 2. Rocky's performance was less robust; she scored 58% correct on Trial 1 (positive but not significant transfer) and 63% correct on Trial 2 (significant transfer). Rio's high performance was consistent with her prior experimental demonstrations of symmetry (C. R. Kastak et al., 2001; Schusterman & Kastak, 1993). Rocky's performance, although not passing the Trial 1 measure of transfer, did show a trend consistent with emergent symmetry. These data show for the first time that some animals, like some human subjects, are capable of emergent symbolic mapping with visual stimuli.

The success of the sea lions in this symmetry test can be compared to the generally poor performances of nonhuman animals in most tests for emergent bidirectional relations. In doing so, it is important to consider that the sea lions had previously shown emergent symbolic performances consistent with the results obtained in the current experiment (C. R. Kastak et al., 2001; Schusterman & Kastak, 1993). Here, the previously established baseline trials presented also contained a large number of omnidirectional stimulus pairings. The subjects' experience with interchangeable sample and comparison stimuli was likely important in mitigating the effects of novelty when the newly introduced stimuli appeared as samples for the first time (see, e.g., Lionello-DeNolf & Urcuioli, 2000).

The performance of the sea lions on both types of transfer trials in Set 3 merits discussion. For ease of experimental procedures, which were not automated, both subjects were tested with Set 1, Set 2, and then Set 3 as shown in Figure 3. Both subjects failed the first transfer test only with the third stimulus set. Although it is impossible to conclude that the failure was attributable only to stimulus variables and not to set presentation order, there is evidence to suggest that the stimulus configurations of Set 3 interfered with transfer performance. The sea lions appeared to discriminate the Set 3 stimuli as "novel" during exclusion trials and responded appropriately during Phase 1. When the stimuli were presented in the transfer phase, in which the subjects could not respond by exclusion, neither subject discriminated the two novel stimuli. This lack of discrimination could be due to the configuration of the Set 3 stimuli. The outlines of both stimuli were light, and the lines that were used to create the shapes were thinner than those of typical stimuli used (see Figure 1). Sea lions have good underwater visual acuity but they are myopic in air (Schusterman & Balliet, 1971), making the discrimination of any stimulus in air more difficult than it is for humans. Although Rio discriminated the Set 3 stimuli following additional exposures of exclusion trials, Rocky showed only marginal improvement after additional training. Rocky has age-related cataracts that impair her vision, and her poor performance with the Set 3 stimuli may be attributed, in part at least, to the subtleties of the stimulus configurations. Although the Set 3 data were included in our analysis, they likely underestimate the extent to which the subjects could perform the transfer problems. Therefore, it is significant that despite their relatively poor performance with Set 3, both subjects showed positive transfer on the pooled tests with Sets 1, 2, and 3.

The primary objective of the present experiment was to determine if equivalence classes could be expanded by relations emerging from exclusion exemplars. Although the subjects did not have to learn about the specific stimulus relations presented in the exclusion phase, they appropriately classified the new stimuli into the letter and number categories during transfer testing. Thus, the exclusion procedure did expand the existing classes. This finding is consistent with that of Meehan (1995), who showed equivalence classes emerging in college students presented with an exclusion paradigm. The basis for this result is somewhat uncertain, however. Perhaps the learning outcomes were achieved for the conditional discriminations following their exposure two to eight times during two to eight exclusion sessions. Given that this would require the simultaneous learning of 20 new conditional discriminations (for each of the three sets) over the course of a few sessions, this possibility seems unlikely. An alternative is consistent with the subjects' prior demonstration of equivalence with the letter and number classes and with the potential role of class-specific reinforcement.

Given that the subjects had already formed two large equivalence classes, they may have treated the novel positive exclusion problems presented with each new stimulus set, not as 20 different discriminations, but as two. That is, the equivalencies between all members of each familiar class could have generated two general problems: [Any letter] is the sample, [any number] is the negative comparison, and novel n (for example) is the positive comparison by default; or, [any number] is the sample, [any letter] is the S−, and novel # is the S+ by default. Given this scenario, there are only two new conditional discriminations to be learned for each new set of stimuli introduced by exclusion: [Any letter] is related to n; [any number] is related to #. In this case, the sea lions would have received 20 to 80 exemplars of each of these two problems in the two to eight exclusion sessions that occurred prior to the first transfer tests. By classifying all of the familiar stimuli into categories, any class member could substitute for any other, and thus, the subjects could appropriately respond to a large number of unique problems by establishing only two general learned outcomes.

The presentation of 20 to 80 exemplars of two general ''equivalence'' exclusion problems is still a relatively small number to generate a learning outcome. Two previous studies with the same subjects introduced new sample/S+ pairings with familiar alternatives, and the results with these ostensibly simpler problems showed that learning outcomes required several hundred reinforced exclusion trials (Schusterman et al., 1993; Schusterman & Kastak, 1995). Three factors might account for the more rapid transfer observed in the present experiment. First, the sea lions were experienced in performing conditional discriminations; that is, they had likely developed a learning set as they became experienced at discounting irrelevant aspects of the MTS task (Schusterman et al., 2002). The second factor relates to the use of class-specific reinforcement, which may have facilitated the acquisition of the new conditional discriminations (Goeters, Blakely, & Poling, 1992). Finally, considering the structure of the exclusion problems presented in the earlier experiments (Schusterman et al., 1993; Schusterman & Kastak, 1995), new conditional discriminations were typically introduced in the context of not one but many different familiar S−s. These S−s were components of various familiar conditional discriminations, but they had no relation to one another. It is possible that the lack of relatedness between the S−s may have made these problems more difficult to learn; conversely, the commonality of the S−s in the current experiment may have facilitated the acquisition of these discriminations. For example, the 10 letters presented as S−s on number-positive trials were equivalent to one another; therefore, they could be chunked to represent a single general S−. The retention of a class concept versus memory for many individual discriminations likely simplified learning by exclusion of the S−, thus allowing a learning outcome to be quickly achieved.

Some may argue that the use of class-specific reinforcement throughout this study weakens the argument for emergent equivalence classification through an exclusion procedure. The procedure and results, however, are consistent with the theory that equivalence relations arise directly from reinforcement contingencies, and that equivalence classes are composed of all elements of those contingencies, including stimuli, responses, and reinforcers (Sidman, 1994, 2000). With regard to the formation of stimulus classes, it should not matter whether the relations that emerged between the newly introduced stimuli and the defined equivalence class members were based on common stimulus relations or common reinforcer relations. The potential use of both conditioning elements in this study likely strengthened and stabilized the emerging relations; a class-specific response, which was not used in the current procedure, likely would have had the same enhancing effect on stimulus classification.

These findings indicate that language skills are necessary for neither spontaneous exclusion performances nor the emergence of subsequent straightforward or symbolic learning outcomes. Perhaps most important, the results add to a growing body of literature that counters the anthropocentric notion that some cognitive abilities (e.g., equivalence and exclusion) are the sole domain of language-adapted humans (e.g., Horne & Lowe, 1996). The rudiments of these abilities in nonhumans suggest, instead, that such abilities arise from general learning and conditioning principles.

In terms of behavior in natural settings, the findings support the observation that individuals gain cognitive economy by organizing information into meaningful and useful categories. These categories appear to be structured in part by the inclusion or exclusion of newly encountered environmental stimuli. This work adds to several lines of evidence suggesting that some nonhuman animals can establish and expand equivalence classes in a variety of experimental, social, and ecological contexts, and that class members may include such disparate stimuli as individuals, events, signals, objects, responses, and outcomes (for reviews, see Schusterman, Kastak, & Kastak, in press; Schusterman, Reichmuth, & Kastak, 2000). Further experimental work is required to clarify the extent to which principles of exclusion and stimulus equivalence interact.

REFERENCES

Berryman, R., Cumming, W. W., Cohen, L. R., & Johnson, D. F. (1965). Acquisition and transfer of simultaneous oddity. *Psychological Reports, 17,* 767–775.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development, 15,* 17–29.

Carter, D. E., & Werner, T. J. (1978). Complex learning and information processing by pigeons: A critical analysis. *Journal of the Experimental Analysis of Behavior, 29,* 565–601.

Cerutti, D. T., & Rumbaugh, D. M. (1993). Stimulus relations in comparative primate perspective. *The Psychological Record, 43,* 811–821.

Dixon, L. S. (1977). The nature of control by spoken words over visual stimulus selection. *Journal of the Experimental Analysis of Behavior, 27,* 433–442.

Gisiner, R., & Schusterman, R. J. (1992). Sequence, syntax, and semantics: Responses of a language-trained sea lion (*Zalophus californianus*) to novel sign combinations. *Journal of Comparative Psychology, 106,* 78–91.

Goeters, S., Blakely, E., & Poling, A. (1992). The differential outcomes effect. *The Psychological Record, 42,* 389–411.

Hanggi, E. B., & Schusterman, R. J. (1995). Conditional discrimination learning in a male harbor seal (*Phoca vitulina*). In R. A. Kastelein, J. A. Thomas, & P. E. Nachtigall (Eds.), *Sensory systems of aquatic mammals* (pp. 543–559). Woerden, The Netherlands: De Spil Publishers.

Herman, L. M., Richards, D. G., & Wolz, J. P. (1984). Comprehension of sentences by bottlenosed dolphins. *Cognition, 16,* 129–219.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 48,* 187–208.

Kastak, C. R., Schusterman, R. J., & Kastak, D. (2001). Equivalence classification by California sea lions using class-specific reinforcers. *Journal of the Experimental Analysis of Behavior, 76,* 131–158.

Kastak, D., & Schusterman, R. J. (1994). Transfer of visual identity matching-to-sample in two California sea lions (*Zalophus californianus*). *Animal Learning & Behavior, 22,* 427–435.

Lionello-DeNolf, K. M., & Urcuioli, P. J. (2000). Transfer of pigeons' matching to sample to novel sample locations. *Journal of the Experimental Analysis of Behavior, 73,* 141–161.

McIlvane, W. J., Kledaras, J. B., Munson, L. C., King, K. A. J., de Rose, J. C., & Stoddard, L. T. (1987). Controlling relations in conditional discrimination and matching by exclusion. *Journal of the Experimental Analysis of Behavior, 48,* 187–208.

McIlvane, W. J., Munson, L. C., & Stoddard, L. T. (1988). Some observations on control by spoken words in children's conditional discrimination and matching by exclusion. *Journal of Experimental Child Psychology, 45,* 472–495.

Meehan, E. F. (1995). Emergence by exclusion. *The Psychological Record, 45,* 133–154.

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

Samuelson, L. K., & Smith, L. B. (2000). Grounding development in cognitive processes. *Child Development, 71,* 98–106.

Schusterman, R. J., & Balliet, R. F. (1971). Aerial and underwater visual acuity in the California sea lion (*Zalophus californianus*) as a function of luminance. *Annals of the New York Academy of Sciences, 188,* 37–46.

Schusterman, R. J., Gisiner, R., Grimm, B. K., & Hanggi, E. B. (1993). Behavior control by exclusion and attempts at establishing semanticity in marine mammals using match-to-sample paradigms. In H. Roitblat, L. Herman, & P. Nachtigall (Eds.), *Language and communication: Comparative perspectives* (pp. 249–274). Hillsdale, NJ: Erlbaum.

Schusterman, R. J., Kastak, C. R., & Kastak, D. (2002). The cognitive sea lion: Meaning and memory in the lab and in nature. In M. Bekoff, C. Allen, & G. Burghardt (Eds.), *The cognitive animal: Empirical and theoretical perspectives on animal cognition.* Cambridge, MA: MIT Press.

Schusterman, R. J., Kastak, C. R., & Kastak, D. (in press). Equivalence classification as an approach to social knowledge: From sea lions to simians. In F. B. M. DeWaal & P. L. Tyack (Eds.), *Animal social complexity: Intelligence, culture, and individualized societies.* Cambridge, MA: Harvard University Press.

Schusterman, R. J., & Kastak, D. (1993). A California sea lion (*Zalophus californianus*) is capable of forming equivalence relations. *The Psychological Record, 43,* 823–839.

Schusterman, R. J., & Kastak, D. (1995). There is no substitute for an experimental analysis of marine mammal cognition. *Marine Mammal Science, 11,* 263–267.

Schusterman, R. J., & Kastak, D. (1998). Functional equivalence in a California sea lion: Relevance to animal social and communicative interactions. *Animal Behaviour, 55,* 1087–1095.

Schusterman, R. J., & Krieger, K. (1984). California sea lions are capable of semantic comprehension. *The Psychological Record, 34,* 3–23.

Schusterman, R. J., Reichmuth, C. J., & Kastak, D. (2000). How animals classify friends and foes. *Current Directions in Psychological Science, 9,* 1–6.

Sidman, M. (1994). *Equivalence relations and behavior: A research story.* Boston: Authors Cooperative.

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 74,* 127–146.

Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review, 57,* 193–216.

Stromer, R., & Osborne, J. G. (1982). Control of adolescents' arbitrary matching-to-sample by positive and negative stimulus relations. *Journal of the Experimental Analysis of Behavior, 37,* 329–348.

Tomonaga, M. (1993). Tests for control by exclusion and negative stimulus relations of arbitrary matching to sample in a ''symmetry-emergent'' chimpanzee. *Journal of the Experimental Analysis of Behavior, 59,* 215–229.

Vincent-Smith, L., Bricker, D., & Bricker, W. (1974). Acquisition of receptive vocabulary in the toddler-age child. *Child Development, 45,* 189–193.

Wilkinson, K. M., Dube, W. V., & McIlvane, W. J. (1996). A crossdisciplinary perspective on studies of rapid word mapping in psycholinguistics and behavior analysis. *Developmental Review, 16,* 125–148.

Wilkinson, K. M., Dube, W. V., & McIlvane, W. J. (1998). Fast mapping and exclusion (emergent matching) in developmental language, behavior analysis, and animal cognition research. *The Psychological Record, 48,* 407–422.

Zentall, T. R. (1998). Symbolic representation in animals: Emergent stimulus relations in conditional discrimination learning. *Animal Learning & Behavior, 26,* 363–377.

Zentall, T. R., Edwards, C. A., Moore, B. S., & Hogan, D. E. (1981). Identity: The basis for both matching and oddity learning in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 7,* 70–86.

Received November 13, 2001
Final acceptance June 8, 2002