Abstract

The aim of this paper is to discuss the important role that behavioral mechanisms, such as contingency learning and equivalence class formation, play in the production and comprehension of auditory signals in the context of mammalian social communication. Observations and experiments on vocal communication in mammals have often emphasized the importance of learning either from the perspective of the signaler or from the perspective of the receiver. It is our goal to discuss the roles and potential mechanisms of learning on both sides of communication. While marine mammals are notable in their capacity for complex learning in their vocal communication, until now, the major emphasis has been on the study of cetaceans. In the current paper, we focus primarily on the pinnipeds (seals, sea lions, and walruses) as a source for insight into how the learned aspects of auditory signaling and receiving may be acquired. We find that the results from carefully designed laboratory experiments can aid in the interpretation of field observations of communicative behavior. The complementary use of both of these approaches improves our understanding of the cognitive operations being carried out by animals in their natural environment.

Key Words: pinnipeds, learning, vocalizations, cetaceans, communication, equivalence, mammals, primates

Investigators of animal communication are generally interested in determining how individuals of a given species influence and are influenced by others. Observational field studies conducted with a variety of different species tend to emphasize actions on the part of either signalers or receivers, with the end result being studies that focus on either the production or the comprehension aspects of communication. It is not always easy to reconcile these perspectives to build a better understanding of how conspecifics interact in complicated social and ecological contexts. An integrative approach requires that knowledge of sensory and cognitive capabilities be applied to real-world observations of both signalers and receivers, with the goal of gaining insight into the generalized and specialized characteristics that influence communication as a whole. Such an approach has revolutionized the study of communication in some primates, for which long-term, detailed observational studies of social behavior have been carefully evaluated in light of controlled field and laboratory experiments dealing with topics such as sensory perception and the development of vocal communication, learning, and social attachment (e.g., Cheney & Seyfarth, 1990; Hauser, 1996; Owren & Rendall, 1997).

One of the central topics emerging from integrated behavioral studies of acoustic communication is that of the role of learning in both signal production and signal comprehension. Observations and experiments have shown that in some avian species, vocalizations must be learned over time; while in others, these responses appear to be less influenced by experience (Kroodsma & Miller, 1996). Among terrestrial mammals, including non-human primates, a smaller body of data suggests that while learning plays a role in shaping how individuals respond to auditory signals, it appears to have little or no influence on how calls are structured and used by individuals (e.g., Seyfarth & Cheney, 1997). In contrast, a variety of field and laboratory studies indicate that some marine mammals may be exceptional in the extent to which learning influences their vocal behavior (see reviews in Tyack & Sayigh, 1997; Schusterman, in press). Studying the learning mechanisms involved in marine mammal communication may prove to be useful in gaining a better understanding of the proximal causes of a range of responses involved in individual recognition, reproductive behavior, and the acquisition of dialects and cultural transmission of songs (Schusterman, in press). The aim of this paper is to briefly consider how auditory, social, and environmental experiences modify the
production and comprehension of auditory signals in a particular group of marine mammals—the pinnipeds. Laboratory experiments and field observations which provide complementary evidence are used to examine the effects of learning on auditory signaling in these animals.

To emphasize a change “in gears” (aka topic), Janik & Slater (2000) have stated that learning can modify the emission of vocalizations in one of two ways: (1) by influencing the context in which a particular signal is used and/or (2) by altering the acoustic structure of the call itself. A topic of special interest is the extent to which the vocalizations of nonhuman primates and other terrestrial mammals may be modified in these ways. A variety of observations support the traditional premise that most mammalian vocalizations are primarily influenced by genetic, structural, and emotional features, with social or environmental experience playing little if any role in modifying vocal behavior (see review in Schusterman, in press). For example, studies of captive nonhuman primates that were raised in isolation, deafened at birth, or cross-fostered with mothers of a different species revealed that individuals without normal auditory experiences still produced species-typical vocalizations and emitted these vocalizations in appropriate situations (e.g., Owren et al., 1993; Hammerschmidt et al., 2001). Likewise, hybrid gibbons emit vocalizations that are intermediate to the calls of either parent species. This suggests strong genetic influences on vocal behavior as well as a possible lack of learned components in these animals’ vocalizations (Tenaza, 1985).

Further, attempts to condition the vocalizations of nonhuman primates using food rewards have had only limited success, leading some primatologists to conclude that their vocal emissions are not under voluntary control (e.g., Myers, 1976). A handful of studies, however, do provide some evidence that nonhuman primates can learn to control their vocalizations under specific conditions (Pierce, 1985; see review by Adret, 1993). Similar observations which demonstrate contextual control of vocal responses in other mammals, including dogs, cats, and rodents, combined with only sparse evidence showing learned modification of call structure, indicate that it is likely easier for mammals to learn when to use a particular call than to alter the structure of the call itself (see reviews by Adret, 1993; Janik & Slater, 1997).

Several investigators have pointed out that some marine mammals appear to have an extraordinary capacity to alter both the contextual and structural features of their vocalizations as a result of experience (e.g., Janik & Slater, 1997; Tyack & Sayigh, 1997; Schusterman, in press). A variety of long-term field studies describing dolphin whistles and whale songs show that at least some cetacean species are capable of learning to reproduce seemingly arbitrary acoustic call contours and patterns (e.g., Payne et al., 1983; Ford, 1990; Tyack & Sayigh, 1997). These observations are supplemented by highly structured laboratory experiments with bottlenose dolphins demonstrating that these animals can be trained to produce arbitrary whistle contours following exposure to playbacks of recorded models (e.g., Richards et al., 1984; Sigurdson, 1993). In contrast to most terrestrial mammals which show some ability to alter their call usage, very limited ability to alter their call structure, and no ability to imitate auditory models, the acoustic behavior of some cetaceans demonstrates an ability to modify acoustic emissions as a result of auditory experience.

While cetaceans are usually considered the most suitable subjects for studies of vocal learning among mammals, an analysis of the acoustic behavior of pinnipeds offers a different view into how learning may modify vocal signaling. Like cetaceans, pinnipeds are highly social and emit sounds under water where other sensory modalities are often less useful as modes of intra-specific communication. In addition, their amphibious lifestyle requires that some socially relevant sounds also be produced in air. Underwater and aerial sound production in a few pinniped species has been described and examined in a series of detailed experimental laboratory studies, the results of which clearly show that call usage as well as call structure can be modified by learning. This literature has been reviewed in detail by Schusterman (in press) and is briefly summarized below.

In captive studies dating back to the 1960s, it was demonstrated that male California sea lions (Zalophus californianus) can learn to inhibit their barking in the presence of any male dominant to them, but vocalize normally when dominant males are absent (Schusterman & Dawson, 1968). Sensory and cognitive experiments involving vocal conditioning procedures revealed that the vocal behavior of male and female sea lions can be placed under the discriminative control of a variety of different arbitrary stimuli, including visual cues differing in size, shape, and grating density and auditory cues differing in frequency and amplitude (see Schusterman, 1978). Further, acoustic features of the vocalizations were shown to be modifiable with experience—not as a result of exposure to auditory models but, rather, as a result of selective reinforcement using fish rewards (Schusterman & Feinstein, 1965). The results of these experiments are consistent with recent work on gray seals (Halichoerus grypus) and walruses (Odobenus rosmarus) showing that
different call types can be selectively conditioned and placed under the discriminative control of different cues in these species (Shapiro et al., 2004; Schusterman, in press). The use of food reinforcement to systematically modify vocal emissions is a technique that has recently been applied with birds—specifically, budgerigars (*Melopsittacus undulatus*), whose vocal structure, like other parrots and songbirds, had previously been thought to be influenced only by auditory and social feedback (Manabe et al., 1997). Thus, it appears that, like motor responses, some aspects of vocal signaling are subject to general principles of reinforcement that do not require specific auditory exposure.

Recently, conditioning experiments with captive harbor seals (*Phoca vitulina*), as well as walruses, have shown that reinforced variability in sound emissions leads to broadened vocal response repertoires and the production of novel sounds (see Schusterman, in press). While these experiments with pinnipeds are the first to use this technique in the vocal domain, the findings are consistent with new ideas about nonvocal, voluntary responses drawn from psychological studies, which suggest that individuals can learn to vary their behavior under the control of reinforcing consequences and that such variability can lead to diverse and novel responses (Neuringer, 2004).

Even so, all of the experimental evidence showing the effects of reinforcement training on the auditory responses of pinnipeds pale in comparison to a single, but convincing case of vocal mimicry in a captive male harbor seal named “Hoover.” Hoover was hand-reared in Massachusetts by human caretakers before being transferred to an aquarium. As he matured, he spontaneously emitted a repertoire of about 12 English words delivered in a clearly recognizable New England accent (Ralls et al., 1985). To date, Hoover remains the only non- cetacean mammal (other than humans) to unequivocally demonstrate vocal mimicry. While similar observations have not been reported since, Hoover’s behavior shows that, under the right circumstances, pinnipeds may use auditory experience, in addition to environmental consequences such as food reinforcement and social feedback, to modify their vocal emissions.

There is an obvious relationship between laboratory studies and field observations of acoustic behavior. While field studies can describe how individuals interact in complex communicative situations, their laboratory equivalents can help to identify the conditions under which learning may modify their vocal responses. For example, field observations showing that nonterritorial sub-adult male sea lions, who frequently vocalize as they move through a crowded rookery, will sometimes “sneak” through a bull’s territory by moving low to the ground and inhibiting their vocalizations while in the bull’s territory (Schusterman et al., 2003). These findings have been corroborated by laboratory studies demonstrating that male sea lions can voluntarily control their auditory signals and learn the contexts in which to produce or suppress them (Schusterman & Dawson, 1968). Laboratory work, then, may serve to inform field studies seeking to better understand the interplay between distal and proximal causes of vocal communication. The results of this type of complementary research may help to identify learned components of other natural phenomena involving vocal behavior such as the geographic and temporal variability exhibited by the dialects in some seals (see Janik & Slater, 1997) and year-to-year variability in the structure of individual walrus songs (Sjare et al., 2003).

The aspect of vocal communication where the influence of learning tends to be most obvious is in the responses of receivers to auditory signals. As is true for their vocal emissions, the behavioral responses that individuals make in response to acoustic signals can be influenced by “hard-wired” as well as learned components (through imitation, contingency learning, or some as yet unknown mechanism). As compared to the genetic and behavioral constraints on production of acoustic signals, the ways in which individuals respond to acoustic signals appears to be more labile. For example, in the primate studies described earlier, offspring raised by foster mothers of a different species developed appropriate species-typical vocal behavior despite inappropriate auditory exposure; however, the same individuals were shown to be able to learn to respond appropriately to the vocalizations of the foster species (Owren et al., 1993). Although the remainder of this paper will focus primarily on operant methods of response acquisition and the learning mechanisms that may influence comprehension of auditory signals, there is a literature on the influence of classical conditioning that is also relevant to this topic (e.g., Rendall & Owren, 2002).

An acoustic signal can become meaningful to a receiver through the reinforcement contingency to which the receiver is exposed. For instance, when a naïve animal hears the sound of an approaching predator, it may sniff the air and visually scan the environment for the predator. If these responses result in detection and avoidance of the predator, the potential prey may have learned that the sound of a predator is a meaningful cue, thereby learning that an appropriate response to the auditory signal is the same as a response to the visual or olfactory stimulus. The auditory signal now can become referential—it can serve to “represent” the predator.
We suggest that animals are able to learn referents for a signal by the mechanism of stimulus equivalence. An equivalence class is a group of perceptually dissimilar stimuli whose relations to each other emerge because they share a common spatio-temporal or functional relationship (Schusterman et al., 2003). Equivalence is demonstrated by “interchangeability” of stimuli such as an individual’s call, its visual profile, and its odor. Through the formation of equivalence classes, receivers can more efficiently respond to signals; further, the association of a novel signal with one of the members of an already established class can lead to emergent relationships with other class members, thereby increasing response efficiency even more.

The mechanism of stimulus equivalence was first described in the psychological literature by Sidman (1971), who used an auditory-visual equivalence paradigm to demonstrate emergent reading comprehension in a mentally retarded child. The subject was taught to relate pictures of objects to corresponding spoken words, and then to match the same spoken words to printed words. Following the acquisition of those trained relationships, the child spontaneously matched the pictures to the printed words even though he had never done so before. This example shows how a printed word, a spoken word, and a picture of a particular object, despite their lack of physical similarities, can all be used interchangeably to represent the same object. The power of such a process is that it allows information from different sensory modalities to be integrated into meaningful and useful units.

An assortment of field observations and playback studies suggest that many animals are able to form cross-modal categories and referential associations, allowing them to respond appropriately in a variety of novel situations. In terms of auditory-visual equivalence, a probable example can be derived from an experimental field study conducted with vervet monkeys (Cercopithecus aethiops) by Cheney & Seyfarth (1990). The experimenters played the call of a juvenile vervet monkey through a speaker to its mother and several other females in the juvenile’s absence. When the call was played, the females responded by looking directly to the juvenile’s mother rather than to the speaker from which the sound was coming. Since the call had previously been associated with the juvenile, and the juvenile with its mother, the responses of the other females seem to indicate comprehension of a relationship between the juvenile’s vocalization (an auditory cue) and the presence of its mother (a visual cue). Referential use of auditory cues by vervet monkeys can also be seen in their use of alarm calls. These vocalizations have been classified into three distinct call types: (1) an “eagle” call, (2) a “snake” call, and (3) a “leopard” call. Each of these calls is correlated with the presence of different predators (Struhsaker, 1967; Seyfarth et al., 1980). What makes these calls particularly interesting is not only that they are acoustically distinctive, but that the vervet monkeys receiving these vocal signals respond differentially dependent upon the call type emitted (Seyfarth et al., 1980). For example, if a signaler produces an “eagle” alarm call, the receivers will run down from treetops and into dense vegetation—an appropriate response as eagles prey from above, swooping into trees to catch monkeys. Thus, these monkeys have formed an equivalence class, which includes the alarm call and the signal of the predator.

In pinnipeds, referential learning can be illustrated by individual recognition between sea lion mothers and their pups. Female sea lions typically nurse their pups over a period of about 8 mo, and during that time, they alternate nursing periods of 2 to 3 d on shore with foraging trips lasting 3 to 5 d at sea (Reidman, 1990). When mothers return to the rookery following these foraging trips, they must identify and locate their pups in what are often crowded and chaotic conditions. When searching for her pup, an individual female usually scans the shoreline and repeatedly emits a unique vocalization, which has been termed a pup attraction call; upon hearing its mother’s distinctive voice, a pup responds with its own unique call (see Insley et al., 2003). The mother and pup exchange these signature calls to locate one another on the rookery, and once they have established contact, they confirm their mutual identification through sight and smell (Dobson & Jouventin, 2002; Insley et al., 2003). The importance of auditory signals in this process has been confirmed through playback experiments conducted with several species, showing that mothers and their pups respond preferentially to one another’s vocalizations (see review by Insley et al., 2003) and that these preferences may persist well beyond the period of maternal dependency (Insley, 2000). The finding that auditory cues can represent individuals in the absence of other temporal or sensory cues suggests that meaningful cross-modal associations have been formed.

A variety of other behavioral interactions observed in naturalistic settings support the premise that auditory signals may be referentially learned by some pinnipeds. For example, a “dear enemy” effect is exhibited by male Stellar sea lions (Eumetopias jubatus) during the breeding season (Gisiner, 1986). The males are territorial and will fight other males to defend the borders of their individual breeding areas; however, there appears to be less combat between familiar rivals (neighbors) than between unfamiliar rivals.
(strangers). This does not appear to be due to habituation effects as differential responses to these two groups of conspecifics are consistent over many breeding seasons (Schusterman et al., 2000). Thus, territorial males appear to save their energy for fighting new opponents. Recent playback experiments conducted with male Australian fur seals (Arctocephalus pusillus doriferus) during the breeding season indicate that this neighbor/stranger recognition, following a period of exposure to neighboring males, can be accomplished through vocal recognition alone (Tripovich et al., 2005). In California sea lions, another type of recognition has been demonstrated in captivity wherein females have been shown to interact more affiliatively with their kin and more aggressively with unrelated individuals (Hanggi & Schusterman, 1990). Such differential social responses are likely based in part on multimodal individual recognition combined with learning about the common spatial-temporal relationships among individuals.

An interesting study of harbor seal responses to acoustic cues emitted by potential predators has been recently reported by Deecke et al. (2002). In a series of acoustic playbacks, these investigators tested the ability of harbor seals to discriminate and classify different call types from three different populations of killer whales (Orcinus Orca): (1) familiar (resident) fish-eaters, (2) unfamiliar (nonresident) fish-eaters, and (3) transient mammal-eaters that frequently prey upon harbor seals. They found that escape and avoidance responses by local harbor seals were significantly more likely to occur to sounds emitted by unfamiliar and transient killer whales than to the sounds of familiar fish-eating killer whales, suggesting that the seals can learn to appropriately classify the acoustic signals emitted by these potential predators and respond accordingly. Such behavior likely confers a selective advantage upon individuals who, by classifying predators according to call type, improve efficiency by reducing responsivity to harmless acoustic cues while responding appropriately to similar sounds that may indicate potential danger.

Field observations of pinniped behavior serve to show that auditory signals can become meaningful to receivers through experiences that allow individuals to relate information across the senses. Laboratory studies are required to better understand the potential learning mechanisms that underlie these behaviors, however. In a laboratory setting, it is possible to test hypotheses about how individuals form relationships between different stimuli. For example, in the laboratory, we can investigate the arbitrary conditions under which relational learning—including the formation of equivalence classes within and across the sensory modalities—occurs.

Some of the cognitive skills used by pinnipeds have been evaluated in a variety of basic laboratory experiments which demonstrate their abilities to solve problems based on perceptual and arbitrary rule-based learning (see Schusterman & Kastak, 2002, for a review). These studies also show good working- and long-term memory for specific events, learned relationships, and previously established concepts (Reichmuth Kastak & Schusterman, 2002). It is probable that such cognitive skills provide the proximal basis for problem solving in the wild.

Equivalence class formation as a special case of relational learning is an ability that may have particular relevance for communicative behavior as well as other types of problem solving. This ability has been investigated in a series of laboratory experiments with California sea lions, the first of which required a young sea lion named Rio to learn 60 different relationships among visual stimuli that differed only in shape (Schusterman & Kastak, 1993). In this task, a total of 90 stimuli were randomly divided into 30 groups of three, with each group having an “A,” “B,” and “C” as designated by the experimenter. Rio was trained using fish reinforcement and trial and error learning to relate the A member of each group to the B member (A→B training), and then to relate the B member to the C member (B→C learning). Following her acquisition of these 60 trained relationships, she demonstrated the emergence of new relationships in the absence of any further training. These were symmetrical relationships (B→A and C→B matching), transitive relationships (A→C matching), and a combination of the two (C→A matching). Rio’s performance on visual matching tests mirrored that of the child tested by Sidman (1971) on auditory-visual tests of reading comprehension, and it showed that a nonhuman animal was capable of the same sort of complex relational learning used by humans to solve novel problems. Subsequent experiments with Rio and another sea lion, Rocky, confirmed that sea lions could derive new, untrained relationships among groups of visual stimuli on the basis of information acquired through previous training (Reichmuth Kastak et al., 2001).

As noted previously, most of the cognitive research with sea lions and other pinnipeds has focused on relational learning within a sensory modality, and it is clear that these animals live in a multimodal world. It therefore follows that animals should be able to form meaningful relationships within and across sensory modalities as many of the field observations we have described would suggest. These field observations allow for
the formation of hypotheses about the mechanisms by which pinnipeds derive meaning from auditory cues and how they attach visual referents to acoustic signals. Therefore, another laboratory experiment with the sea lion Rio was recently undertaken to examine the process by which visual referents could be mapped onto arbitrary auditory cues so that the formation of auditory-visual equivalence classes could be assessed under controlled conditions. In this experiment, Rio was taught using food reinforcement to relate different arbitrary acoustic signals with individual members from each of two large classes of visual stimuli (see Reichmuth Kastak et al., 2001). Once these explicit relationships were acquired between the acoustic signals and the individual class members, Rio was tested to determine if untrained transitive relationships would emerge between the acoustic signals and the remaining members of each of the corresponding visual classes. She responded to this testing by correctly and immediately relating auditory signals to visual shapes on the basis of their mutual relationships with other class members (Lindemann et al., 2004; Reichmuth Kastak et al., 2004). Rio’s successful performance on this auditory-visual transitivity task is similar to that shown by human subjects in auditory-visual word learning and reading comprehension tasks (Sidman, 1971) and demonstrates the process through which pinnipeds and other mammals may come to associate visual referents with auditory signals.

From the receiver’s viewpoint, learning how to respond appropriately to acoustic signals probably depends on basic skills involving relational learning as well as more complex cognitive abilities such as the formation of multimodal stimulus classes. Collectively, experimental research such as that described here explains how information from multiple sensory modalities may become interrelated into equivalence classes, and how animals, such as sea lions, may use these classes to attach meaning to auditory signals.

The laboratory experiments and field observations described in this paper support the premise that learning plays a significant role in the production, as well as the comprehension, of acoustic signals by pinnipeds. Like whales and dolphins, pinnipeds appear to be especially well-suited subjects for studies seeking to identify the learning mechanisms that are involved in acoustic communication and other aspects of complex behavior. The vocal emissions of pinnipeds, while largely constrained by biological factors, can also be modified by experience to an extent not reported for other terrestrial mammals. Understanding how to respond appropriately to acoustic cues likely relies to a larger extent on cognitive abilities such as the formation of multimodal equivalence classes. Contrary to the idea that vocal behavior is best studied in the field in order to avoid the constraints imposed by an artificial laboratory environment (Snowdon & Hausberger, 1997), we suggest that the unification of field observations with complementary laboratory experiments provides a more complete understanding of the behavioral bases of acoustic communication. The use of carefully designed and implemented laboratory studies can improve interpretation of real-world behavioral responses and guide the formulation of research questions that can help to determine the extent to which these responses are learned.

Literature Cited


Sjare, B., Stirling, I., & Spencer, C. (2003). Structural variation in the songs of Atlantic walruses breeding in the...
Canadian High Arctic. *Aquatic Mammals*, 29(2), 297-318.


