Functional equivalence in a California sea lion: relevance to animal social and communicative interactions

RONALD J. SCHUSTERMANN & DAVID KASTAK
Institute of Marine Sciences, University of California at Santa Cruz

(Received 2 May 1997; initial acceptance 17 May 1997; final acceptance 15 August 1997; M.S. number: x7696)

Abstract. Laboratory investigations into equivalence class formation suggest how animals in social and communicative contexts learn to place dissimilar individuals, signals, responses and social reinforcers into the same functional class. Kastak & Schusterman (1994, Anim. Learn. Behav., 22, 427–435) demonstrated that a California sea lion performed generalized identity matching-to-sample; that is, it chose visual stimulus A conditionally upon an identical sample A (AA matching), chose stimulus B conditionally upon sample B (BB matching) and chose stimulus C conditionally upon sample C (CC matching). The sea lion was later trained on 30 problems with similar stimuli to select comparison B conditionally upon sample A (AB matching), and trained on another 30 problems to select comparison C conditionally upon sample B (BC matching). Subsequently, the sea lion demonstrated trial-1 BA and CB matching and trial-1 AC and CA matching (Schusterman & Kastak 1993, Psychol. Rec., 43, 823–839). Matching of these derived relations defines the phenomenon of stimulus equivalence: when one member (A) of an equivalence class (ABC) becomes discriminative for a given behaviour, then B and C should become discriminative for the same behaviour. In the current study, we tested whether the sea lion could transfer the relations it had acquired between equivalence class members from a matching-to-sample paradigm to a simple discrimination paradigm. In 28 of 30 tests, the sea lion immediately transferred the discriminative function acquired by one member of an equivalence class to the remaining members of that class. Substitutability among members of an equivalence class is relevant to an analysis of referential communication, for example, the representational function of alarm calls.

Investigations into the social behaviour of non-human primates and other animals suggest a degree of complexity and versatility in their social and communicative interactions which until recently had been unexpected (reviewed in Johnson & Norris 1986; Cheney & Seyfarth 1990; deWaal 1991; Gritten 1992; Heyes 1994). For example, vervet monkeys, Cercopithecus aethiops, categorize two acoustically different calls along with the related referent into a functionally equivalent class (Cheney & Seyfarth 1988). Male bottlenose dolphins, Tursiops truncatus, form pairs or triplets attacking other male coalitions to sequester reproductively active females (Connor et al. 1992). Captive Java monkeys, Macaca fascicularis, classify individuals into affiliative pairs (Dasser 1988) and captive California sea lions, Zalophus californianus, are more affiliative towards relatives than non-relatives (Hanggi & Schusterman 1990). Also, non-human primates appear to switch from one behavioural social reinforcer to another, that is, exchanging a mount for tolerance at a food source or grooming for later support in an alliance (Cheney et al. 1986). At least in social and communicative contexts, many species are capable of placing dissimilar individuals, signals, responses and social reinforcers into the same functional class.

What behavioural or learning processes are responsible for these observed patterns of social and communicative interactions in non-verbal animals? Does the behaviour depend on an animals ability to learn a network of associations and form relational concepts (Cheney & Seyfarth 1990; deWaal 1991; Heyes 1994; Thompson 1995), or does the learning involved in social and communicative interactions depend on a
specialized or distinctively social cognitive process (Heyes 1994)?

Sidman (1994) suggested that both social and non-social features of the environment can become related through behavioural contingencies, becoming mutually substitutable even when sharing few or no perceptual similarities. These categories of stimuli whose interchangeability does not depend on perceptual similarity are called equivalence classes. Until recently, the investigation of the cognitive processes giving rise to the formation of equivalence classes has been tested only with human subjects, so the generality of this capability across species is as yet unknown. The ability to form equivalence relations should be adaptive, however, in that it allows an individual to relate unconnected information in novel ways in order to solve problems. Reacting to dissimilar stimuli as members of a single class and using the concept in a variety of contexts is an efficient method of grouping information and economizing cognitive effort (Schusterman & Kastak 1993; Wasserman 1993). For example, responding affiliatively or aggressively towards different individuals from the same matriline enables an animal to respond appropriately in a variety of novel social encounters.

There are several methods for empirically determining whether ordered pairs of events or interrelated stimuli make up an equivalence class. Using a matching-to-sample paradigm (MTS), relations are established with explicit reinforcement contingencies in stimulus pairs called AB and BC. The term conditional discrimination is used to describe such tasks, because responses to particular discriminative stimuli are contingent upon the sample or conditional stimuli. For example, a subject may be trained to choose comparison B conditionally upon sample A (AB matching) and later trained to choose comparison C conditionally upon sample B (BC matching). Performance on conditional discriminations, however, reveals little about the mechanisms underlying the matching behaviour; the ability to match visual stimulus A to visual stimulus B following training does not imply that A and B have become mutually substitutable, or equivalent (Sidman & Tailby 1982). The emergence of equivalence relations is tested by showing that the relations established by the original conditional discrimination procedures can be extended to stimulus pairs AA, BB, CC (reflexive), BA and CB (symmetrical), AC (transitive) and CA (equivalence) without any reinforcement contingency other than that involved in AB and BC training (Sidman 1994). These four types of stimulus pairings are termed emergent if they occur without further training. If criteria for emergence of symmetry, transitivity and equivalence are met, then stimuli A, B and C form an equivalence class.

Sidman (1971) related the emergence of novel stimulus relations to reading comprehension. Using MTS, he trained a mentally retarded child to select, for example, the printed word DOG after hearing the spoken word ‘dog’. The combination of the child’s earlier history of relating the spoken word to the picture, and the newly acquired skill of selecting the correct printed word after hearing it spoken, produced a number of novel behaviours not directly taught or originally present in the child’s repertoire. These behaviours included selecting a picture of a dog on seeing the word DOG, and vice versa. Thus, Sidman showed that reading comprehension emerged despite never having been reinforced.

In non-humans, Schusterman & Kastak (1993) reported convincing evidence of symmetry, transitivity and equivalence relations for one female California sea lion. Strong evidence for reflexivity or generalized identity matching exists for California sea lions (Kastak & Schusterman 1994), some non-human primates (e.g. Oden et al. 1988) and a bottlenose dolphin (Herman et al. 1989). Sidman & Tailby (1982) pointed out that reflexivity is not only an integral part of the definition of equivalence but it is also a behavioural prerequisite, because without it an individual would be incapable of the derived conditional discrimination required to form equivalence relations. Thus, evidence showing that the chimpanzee and the bottlenose dolphin can do generalized identity matching suggests that, like the California sea lion, these species may also be capable of forming equivalence relations with conditionally related stimuli.

Although this behavioural/mathematical definition of equivalence has become standard, it is somewhat restrictive in that it necessarily excludes non-similarity-based classes formed in procedures other than MTS. Sets of dissimilar stimuli, which for procedural reasons may not satisfy the equivalence criteria, have been termed functionally equivalent. To be functionally equivalent, all members of a class must share a common stimulus
function, such that variables applied to one class member are likely to influence the other class members without explicit conditioning (Galloway & Petre 1968; Dube et al. 1993). Thus, when a response to one member of a functional class occurs, that same response should transfer to all other members without training.

In one kind of test of functional equivalence, the stimuli in one class are discriminative for reinforcement (S+), and those in the other class are discriminative for extinction (S−). Tasks like this are termed simple rather than conditional discriminations, because the selection of a particular discriminative stimulus does not depend on a specific sample stimulus, but rather depends on contingencies that are set into place by the experimenter at the beginning of each session. For example, Vaughan (1988) taught pigeons to peck at any of 20 positive visual stimuli (slides of trees) and to withhold their response from any of a set of 20 negative visual stimuli (also slides of trees). Then the discriminations were repeatedly reversed. Once the pigeons learned that the first few stimuli in one set became discriminative for pecking and that the first few in the other set became discriminative for not pecking, responses to the remaining stimuli in each set also changed appropriately. The discrimination reversal procedure had generated two functional classes. It is unclear, however, whether functional classes (identified by their members’ common behavioural functions) and equivalence classes (identified when relations among their members meet the three defining features of equivalence relations) are behaviourally the same (Sidman et al. 1989; Hayes 1991). The question of whether members of equivalence classes also form functional classes can be asked only when subjects show equivalence class formation; this has been confirmed in humans (e.g. Lazar 1977; Wulfert & Hayes 1988). In the present study, we posed a similar question with a non-human animal: can a California sea lion that has formed equivalence classes that were created within arbitrary matching-procedures place the members of those relations into functional classes?

There has been some doubt as to whether non-similarity-based classes formed under procedures such as Vaughan’s (1988) are also equivalence classes, because reflexivity, symmetry and transitivity can be tested only in MTS. Proponents of the argument that language is a prerequisite for equivalence cite the general failure of non-human animals to satisfy the equivalence criteria as evidence that they cannot form equivalence classes (Dugdale & Lowe 1990). Successful transfer of equivalence relations from MTS to a simple discrimination, however, would indicate a link between classes formed in both procedures. Such performance would show that emergent relations formed under certain conditions can be transferred to novel contexts. It would also provide compelling evidence that functional classes in non-human animals may indeed be equivalence relations. In addition, it would imply that such classes are derived from social and/or ecological factors, rather than linguistic ones.

Using an arbitrary MTS procedure, Schusterman & Kastak (1993) first established that the California sea lion, Rio, could relate stimulus pairs AB and BC with specific reinforcement contingencies. Rio was trained to select B comparisons conditionally upon presentation of A samples and to select C comparisons conditionally upon presentation of B samples, for each of 30 potential classes. Rio could immediately extend these relations without further training to stimulus pairs BA and CB (symmetry), AC (transitivity) and CA (equivalence). Ultimately, Rio had formed 30 three-member sets of stimuli, which were mutually substitutable within a matching-to-sample procedure. Of these 30 sets, 18 were tested for emergent CA relations, and 12 were used to train symmetrical and transitive relations prior to testing. In the current study with Rio, we tested whether the discriminative function acquired in the MTS procedure by one member of an equivalence class transfers immediately and completely to the two other members of that class in a novel context. That is, if one member (A) of an equivalence class (ABC) becomes discriminative for a response, do the other members (B and C) also become discriminative for that same response? More specifically, do stimulus members of a class have the same function in a simple discrimination setting where the conditional cue, or sample, is absent, as they did in the MTS setting?

To illustrate, consider an experiment in which two equivalence classes (numbered 1 and 2), each consisting of three members (A, B and C) are tested in a simple discrimination. The subject initially receives one simple discrimination pitting the B member of class 1 (B1) against the B member of class 2 (B2). If she selects B2 rather
than B₁, establishing B₂ as the S+ and B₃ as the S−, will she then select A₂ rather than A₁ when these stimuli are first presented as alternative choices? Will she also select C₂ rather than C₁ when they also are first presented as alternative choices?

METHODS

Subject and Apparatus

The subject was Rio, a 7-year-old female California sea lion. She was housed in salt-water-filled tanks at Long Marine Laboratory in Santa Cruz, California, and tested on a haulout deck adjacent to a 7.6-m diameter tank. Rio was fed 4–6 kg of freshly thawed cut herring and capelin each day, one-third of which was typically consumed during test sessions. Usually tests were conducted 5 days per week between 0900 and 1200 hours.

Rio had been trained and tested in several conditional discrimination tasks using visual MTS procedures. In all such procedures, the sea lion received a sample stimulus followed by two comparison stimuli, and was acoustically signalled to approach and put her nose into a stimulus box containing one of the comparisons. Correct choices either duplicated the sample (identity matching) or were different from the sample (arbitrary or ‘symbolic’ matching). Earlier studies involved training and testing Rio for relational learning, learning by exclusion (Schusterman et al. 1993), generalized identity matching (Kastak & Schusterman 1994), and just prior to the current study, forming equivalence relations (Schusterman & Kastak 1993).

The apparatus was a set of three hinged plywood boards (overall size 120 cm high by 240 cm wide). Stimulus boxes, each 30 cm on a side, were built into the centre of each board. The front faces (windows) of each box were flush with the board and covered by opaque sliding doors. A stationing PVC bar was fixed upright approximately 45 cm directly in front of the centre (sample stimulus) box. Only the two side comparison boxes were used, which were equal distances to the right and left of the centre of the apparatus. The boards housing the right and left boxes were angled in approximately 25° to allow the subject to view the comparison stimuli while stationed.

To begin a trial, two assistants seated behind the boards placed the stimuli inside the two side boxes. The side doors were opened simultaneously to expose the two discriminative stimuli. Upon hearing an acoustic release signal, Rio made a choice by sticking her nose into one of the two boxes containing the discriminative stimuli. A third, concealed, assistant delivered reinforcement (a piece of fish) for correct choices. Incorrect choices were not reinforced (see also Schusterman et al. 1993).

Procedure

Rio had already demonstrated in an MTS procedure the ability to form equivalence classes, each consisting of three-member groups of non-identical stimuli (Schusterman & Kastak 1993); examples of the stimulus arrangements are shown in Fig. 1(a). In the original experiment Rio was trained to select, for example, comparison B₂₈ conditionally upon sample A₂₈, or to select B₂₉ conditionally upon A₂₉. She was further trained to select the appropriate C comparisons conditionally upon presentation of B samples. The ultimate and comprehensive tests for stimulus equivalence occurred when the experimenter interchanged the three members of each class as sample and comparison in novel configurations. That is, any of the members of a class could appear as a sample and any of the other two class members could appear as a correct comparison. Rio showed no decrement in performance related to the presentation of novel combinations of sample and comparisons (Schusterman & Kastak 1993). Rio was trained and tested in the current experiment approximately 2 weeks after demonstrating her ability to form equivalence relations.

The basic procedure was a two-choice simple discrimination, rather than a conditional discrimination. The stimuli used to test for functional equivalence were the same stimuli used to test for equivalence relations, but no sample stimulus or conditional cue was provided before the discriminative stimuli were exposed. For the simple discrimination procedure, a stimulus pairing was defined as two potential classes pitted against each other as alternative choices. Stimulus configurations for the simple discrimination were formed by pairing the A, B and C members from each of two randomly chosen three-member equivalence classes previously learned by Rio (Schusterman &
Kastak 1993). Rio’s 30 equivalence classes were divided into 15 class pairings, each consisting of three simple discriminations pitting member A versus member A, B versus B and C versus C. On the first discrimination (training phase), one of the three possible member pairings was presented, and her choice on trial 1 was reinforced, regardless of class membership. Thus, Rio’s selection on trial 1 established the reinforcement contingencies. When she had learned the training discrimination (criterion was 90% correct in a block of 10 consecutive trials), one of the two remaining stimulus pairings from the same two classes was tested (transfer test 1). The stimulus pairing was novel, in that these class members had previously been paired only in conditional discriminations and never in a simple discrimination. The third pairing from the same classes (transfer test 2) was introduced following criterion performance on the second pairing. This procedure was repeated for the remaining 14 class pairings. The sequence of class pairings for training and transfer phases were arranged so that there was a balanced order of the membership pairings for training, first transfer and second transfer phases (A-B-C, A-C-B, B-A-C, B-C-A, C-A-B, C-B-A). Our measure of transfer was Rio’s performance on trial 1 of each novel stimulus pairing, giving a total of 30 novel trials.

Figure 1 shows a training-transfer sequence of one class membership pairing (B-A-C) between equivalence classes 28 and 29. In the training phase, member B from class 28 (B28) might be pitted against member B from class 29 (B29). Rio’s choice on trial 1 then established the reinforcement contingencies for the remainder of the training phase. For example, a choice of B28 on trial 1 established B28 as the S+ and B29 as the S−. Subsequently, responses to A28 (and not A29) would be reinforced on transfer test 1 and responses to C28 (and not C29) would be reinforced on transfer test 2. Transfer tests 1 and 2 consisted of new stimulus pairings; thus, trial 1 of each transfer test was a novel two-choice discrimination. Performance on the first trial of each transfer test therefore constituted the critical measure of the equivalence response in this experiment.

Experimenters have typically used non-reinforced probes to measure the ability of their subjects (usually humans) to form equivalence classes. In the current experiment, since reinforcement plays no role in the choice of a stimulus on the first trial of a simple discrimination, we were concerned mainly with trial-1 performance as an indicator of equivalence class formation on the two transfer tests following training.

Following the equivalence transfer tests, we performed a control test consisting of a new series
of training and transfer discriminations. This test was arranged in the same fashion as the earlier functional equivalence test. The stimuli comprising each discrimination were chosen from a pool of novel objects and were assigned roles of S+ and S− on an arbitrary basis. The rationale for this test was to control for procedural artefacts by presenting the subject with problems that were impossible to solve. Better than chance performance (i.e. correctly choosing the arbitrary S+) on the first trials of these test discriminations could occur only in the presence of artefacts such as inadvertent cueing. If such responding occurred on the control tests, then correct responses on the original transfer tests could be explained by reasons other than class membership. As in the experimental condition, Rio’s initial choice on each training discrimination was reinforced, and transfer tests were presented following criterion performance. We determined, following 10 transfer tests, that additional control problems were unnecessary, because the subjects’ trial-1 performance did not deviate from chance levels over time. Experimental sessions generally consisted of 40 trials and two sessions were conducted per day.

**RESULTS**

Rio’s initial choice in the training phase was assumed to facilitate her responsivity to the designated S+. Her median number of errors to criterion was 0 (range=0–9) for the experimental condition and 1 (range=0–6) for the control condition. There was no significant difference between the number of errors to criterion in the training phases and the number of errors to criterion in the two test phases (Kruskall–Wallace ANOVA, $\chi^2=3.79$, $P>0.05$). On the first trial of each of the 30 transfer tests, Rio made 28 correct responses; such performance is significantly better than expected by chance (two-tailed binomial test, chance=0.5, $P<0.01$; Fig. 2). In the control condition, as expected, Rio made three correct responses on each of the 10 transfer tests, not significantly better than expected by chance (two-tailed binomial test, $P>0.05$). The difference between performance in the experimental and control conditions was significant (Fisher’s exact test, $P<0.01$). Rio performed at about the same levels on trials 2–10 of the experimental transfer tests as she did on trial 1 (87.3% versus 93.3% correct responses; Fisher’s exact text, $P>0.1$).

**DISCUSSION**

Twenty-eight of 30 three-member equivalence classes formed during conditional discriminations remained functional when the animal was shifted 2 weeks later to simple discriminations. Thus, like human subjects, a California sea lion that has formed equivalence classes can also form functional classes. In Sidman & Tailby’s (1982) original formulation of stimulus equivalence, the learned equivalence of dissimilar stimuli was based on the notion that equivalence relations or conditionally related stimuli (which have become interchangeable in their control over behaviour) have characteristics of reflexivity, symmetry and transitivity. Sidman (1994) stated, however, that equivalences can also be the product of simple discriminations and that responses and reinforcers, just as discriminative and conditional stimuli, can be members of equivalence classes. Our results showing that equivalence relations formed by a California sea lion under one set of conditions can be transferred to a novel set of conditions, supports Sidman’s (1994) expanded view of stimulus equivalence.

As noted earlier, in Rio’s original test for equivalence (Schusterman & Kastak 1993), the explicit training of 12 equivalence classes occurred prior to the ultimate test of equivalence using the remaining 18 sets. Following that experiment, it
was impossible to draw conclusions regarding the emergent properties of the originally trained 12 potential classes. Transfer of stimulus function from conditional to simple discriminations in the current study, however, indicates that these previously trained classes, although not explicitly tested for equivalence, had been placed by Rio into equivalence classes. Thus, equivalence probably can and does occur during the direct training of conditional discriminations if the subject has had sufficient experience.

Evidence supports the following theses: (1) current animal data support the idea that simple as well as conditional discriminations can be the source of equivalences; (2) non-similarity-based conceptual behaviour is not limited to humans and may be a precursor to human linguistic performances such as naming (but see Horne & Lowe 1996); (3) learning about the equivalence of stimuli is relevant to an analysis of an animal’s knowledge about social relations as well as its referential communication (pigeons: Vaughan 1988; Wasserman & DeVolder 1993; sea lions: Schusterman 1990; Schusterman & Kastak 1993; Schusterman et al. 1995; Reichmuth 1997).

Social and communicative interactions of California sea lions observed in the field indicate that equivalence classes may play a role in their naturally occurring behaviour. For example, the learning programme promoting imprinting by a sea lion pup of its mother’s voice (Trillmich 1981; Schusterman et al. 1992), smell and appearance may help the pup relate those cues with one another and with powerful reinforcers such as milk, suckling, protection and warmth that the pup receives from its mother. The sea lion pup’s extended period of maternal care is punctuated by its mother’s frequent trips to sea to feed. Upon returning from these foraging trips, mothers and pups locate each other primarily by means of stereotyped pup-attraction calls produced by the mother. Pup-attraction calls are often followed by interspersed with calls uttered by the pups themselves. This duetting eventually results in reunion of mother and pup (Schusterman et al. 1992). Auditory and visual input, along with proximal modalities such as olfaction and touch, probably integrate to provide both mothers and pups with accurate representations of appropriate kin (Hanggi & Schusterman 1990). The common coding of an individual’s identity through a variety of refined sensory modalities probably becomes activated in mother-pup recognition, and may be a critical developmental requirement in group-living animals such as California sea lions (Schusterman et al. 1995). Perhaps the same skills that enable the pup to recognize its mother’s odour and visual appearance when it has only heard her voice may later enable the mature sea lion to recognize its sisters through their relationship with the mother and each other (Hanggi & Schusterman 1990). We suggest that the activation of such representational skills during early development are directly reflected in the mature sea lion’s ability to pass tests of stimulus equivalence successfully. In contrast, harbour seals are notoriously poor at performing in matching-to-sample tasks (Constantine 1981; Hanggi & Schusterman 1995), and harbour seal females, unlike California sea lion females, attend their pups for a very brief period and do not produce pup-attraction calls (Lawson & Renouf 1985).

In a general way, equivalence relations are relations in which stimuli, behaviours or reinforcers are interchangeable with one another because they have previously shared a common association with other stimuli, behaviours or reinforcers. For example, when free-ranging adult female vervet monkeys are played the scream of an absent juvenile from a concealed loudspeaker for the first time, they frequently respond to the playback scream by looking at the juvenile’s mother, before the mother looks towards or approaches the speaker herself (Cheney & Seyfarth 1980). One could say that the adult females through previous experience related the sound of the juveniles scream (A) with the juvenile itself (B) and related the juvenile (B) with its mother (C) and therefore, that the first time they heard the scream of the absent juvenile vervet monkey, without any additional experience, they immediately oriented to its mother. Responsiveness to the mother by the other adult females in the group emerged in this novel context, even though these female monkeys may have never explicitly had the experience before. Equivalence learning models would argue that the existing relation between the ‘scream’ (A), the juvenile itself (B) and the frequent association between the infant (B) and its mother (C) resulted in a three-member equivalence class consisting of the scream, the juvenile and the mother (A, B and C). That is, under certain circumstances, these events and individuals may all be classified as the same.
As mentioned earlier, learned equivalence of stimuli can be applied to affiliative and aggressive behaviour of group-living individuals (e.g. Hanggi & Schusterman 1990). Following an aggressive interaction between two vervet monkeys from different genetic lineages (A1 and A2), other individuals from those lineages (B1 and B2) having observed A1 and A2 fight, are more likely to fight even though they were not involved in the original conflict (Cheney & Seyfarth 1990). The aggressive events in the description are interchangeable, not because of a physical resemblance between the individuals of each lineage, but because the individuals in each of the lineages have shared a common functional association in terms of temporal/spatial proximity and positive and negative reinforcers. In general, kinship and friendship-based coalitions seem to depend on a general learning process that cuts across a wide variety of species, but there are likely to be specializations of this ability that are applicable to social, communicative, foraging and predator-avoidance contexts. We agree with Wasserman & DeVolder (1993) and Sidman (1994), who believe that stimulus equivalence is an important and much neglected area of experimental inquiry into the realm of animal cognition.

**ACKNOWLEDGMENTS**

This research was supported in part by grants N 00014-85-K-0244 and N 00014-92-J-1970 from the Office of Naval Research and a grant from the Center for Field Research to R.J.S. The preparation of this manuscript was supported by ONR grant N 00014-96-J-0011 to R.J.S. We especially thank Colleen Reichmuth for her encouragement and help during the original writing of this manuscript as well as in its revision. Thanks also to two anonymous referees and to Michael Beecher and Con Slobodchikoff for suggesting important clarifications, particularly in the Methods section.

**REFERENCES**


Hanggi, E. B. & Schusterman, R. J. 1995. Conditional discrimination learning in a male harbor seal (Phoca