A bottlenose dolphin discriminates visual stimuli differing in numerosity

ANNETTE KILIAN
Ruhr-Universität Bochum, Bochum, Germany
and Tiergarten Nürnberg, Nürnberg, Germany

SEVGI YAMAN
Ruhr-Universität Bochum, Bochum, Germany
and Marineland Mallorca, Mallorca, Spain

LORENZO von FERSEN
Tiergarten Nürnberg, Nürnberg, Germany

and

ONUR GÜNTÜRKÜN
Ruhr-Universität Bochum, Bochum, Germany

A bottlenose dolphin was trained to discriminate two simultaneously presented stimuli differing in numerosity (defined by the number of constituent elements). After responding correctly to stimuli consisting of three-dimensional objects, the dolphin transferred to two-dimensional stimuli. Initially, a variety of stimulus parameters covaried with the numerosity feature. By systematically controlling for these stimulus parameters, it was demonstrated that some of these attributes, such as element configuration and overall brightness, affected the animal's discrimination performance. However, after all the confounding parameters were under control, the dolphin was able to discriminate the stimuli exclusively on the basis of the numerosity feature. The animal then achieved a successful transfer to novel numerosities, both intervening numerosities and numerosities outside the former range. These findings provide substantial evidence that the dolphin could base his behavior on the numerosity of a set independently of its other attributes and that he represented ordinal relations among numerosities.

The history of the study of numerical abilities in animals is almost as long as the history of experimental psychology (see Rilling, 1993, for an overview). Yet the first convincing evidence for animal numerical competence was reported by Koehler (1937, 1943), who established a number of experimental paradigms, such as simultaneous and successive stimulus presentation, as well as matching-to-sample and oddity-matching procedures.

A large body of experimental literature has accumulated, demonstrating that a variety of avian and mammalian species are able to categorize stimulus sets on the basis of their numerosity (Boysen & Capaldi, 1993). However, there is considerable controversy surrounding these findings. Some researchers (Dehaene, 1997; Gallistel & Gelman,1992; Meck & Church,1983) have suggested that numerical discrimination is a fundamental faculty and that animals may apply this ability automatically while interacting with their worlds. Others (Davis & Memmott, 1982; Davis & Perusse, 1988) have argued that number is not a salient aspect of the environment for animals; thus, they attend to it as a "last resort"—that is, only if they cannot use any other attributes as a basis for discrimination.

Recent research has produced evidence against the latter assumption. It has been demonstrated for some animals that their behavior is controlled by numerosity, rather than by any other, nonnumerical parameters of a test stimulus (e.g., for pigeons, Emmerton, Lohmann, & Niemann, 1997, and Xia, Emmerton, Siemann, & Delius, 2001; for rhesus monkeys, Brannon & Terrace, 1998). Although most investigations have been carried out with trained animals under laboratory conditions, it has also been demonstrated that noncaptive animals spontaneously discriminate food amounts on the basis of their number (Hauser, MacNeilage, & Hauser, 2000). Still, we are far from assessing the extent to which animals may apply a number sense to their natural environment. But it is conceivable that such an ability could provide ecolog-

ically relevant information—for instance, to assess food amounts or the group size of conspecifics.

The cognitive skills of bottlenose dolphins have been studied intensively (see Herman, 1986, for an overview), yet little is known about their numerical abilities. Mitchell, Yao, Sherman, and O'Regan (1985) demonstrated that a dolphin could choose correctly among objects differing in their amount of reward (number of fish) on a scale from 0 to 5. The stimuli represented different numerosities, but this feature covaried with food amount (volume and weight) and time of feeding. Moreover, the subject could have perceived the objects as representing absolute quantities or hedonic values, rather than as members of an ordinal series. This implies that the dolphin was capable of representing relations among stimuli, but not necessarily an ordinal scale among numerosities.

The aim of the present study was to demonstrate numerical competence in a bottlenose dolphin. First, we tested for its capability to discriminate stimuli on the basis of their numerosity by systematically controlling for confounding stimulus parameters. Second, we investigated the dolphin's performance in transfer to novel numerosities. Given the problems discussed above when food items are used as stimuli representing numerosity, we decided to use stimuli with an intrinsic number feature that was defined by the total number of elements of the given stimulus. In Experiment 1, we employed stimuli consisting of three-dimensional objects, where the number of objects represented the numerosity characteristic of the stimulus. Despite the fact that these stimuli raised difficulties in the control of some parameters (such as size, volume, and overall darkness) covarying with the numerosity feature, it was yet advantageous to start with this kind of stimuli.

Although bottlenose dolphins see very well (W. W. Dawson, 1980), dolphins are typically classified as acoustic specialists. But there is a growing body of data that show that their visual system is a powerful source of information for complex cognitive behaviors and mental processes. These are, for example, the successful imitation of visually observed behaviors (Bauer & Johnson, 1994), comprehension of an artificial gestural language (Herman, 1986), and concept acquisition in visual matching-to-sample tasks (Herman, Hovancik, Gory, & Bradshaw, 1989). It has become obvious that the limited performance levels in visual tasks described in earlier works (e.g., Forestell & Herman, 1988) was basically due to the choice of stimuli. Herman (1990) recommended using three-dimensional, real-world objects and ensuring a clear distinction between figure and background. Once a dolphin has learned a visual task in the three-dimensional world, a transfer to two-dimensional and even abstract signals is easily accomplished, as was demonstrated in the study with video displays by Herman, Morrel-Samuels, and Pack (1990). In our study, the procedure and stimuli were thus chosen accordingly. After successful performance in the discrimination of three-dimensional stimuli was reached in Experiment 1, we introduced stimuli consisting of two-dimensional elements (Experiment 2),

where again the numerosity feature was represented by the number of elements in a stimulus. The use of two-dimensional stimuli made it possible to eliminate the confounding of numerosity with other aspects of the stimulus arrays. This procedure enabled us to evaluate whether a dolphin could discriminate stimuli on the basis of their numerosity when all other stimulus attributes were controlled. The demonstration of this ability would meet the criterion for a numerical category proposed by Gallistel (1990).

Besides controlling for confounding stimuli parameters, validation of numerical abilities also requires demonstration of the ability to transfer from trained to novel numerosities (Davis & Perusse, 1988). If an animal is trained to discriminate different numerosities, it could either respond to them as absolute and disconnected features or else have a one-to-one mapping from distinct numerosities to mental representatives of these entities (Gallistel, 1990). If the latter is the case, the animal ought to be able to encode novel numerosities according to an internal scale. To test for this capability in the dolphin, we offered new stimuli, either intervening numerosities or numerosities outside the former number range. An immediate transfer to these novel numerosities would suggest an understanding of abstract numerosity properties.

EXPERIMENT 1

Experiment 1 was designed to investigate whether a dolphin could discriminate stimuli varying in number of items. In addition, we planned to analyze whether any other stimuli features influenced performance. Numerosity was defined as the total number of visual elements constituting a given stimulus. With the type of stimuli used during discrimination learning, a number of features covaried with numerosity. Tests with new sets of stimuli systematically controlled for these confounding features. Previous studies have shown that discrimination tasks with static two-dimensional stimuli may lead to prolonged training phases in dolphins (Forestell & Herman, 1988). We thus used three-dimensional real objects for the stimuli, which could be easily perceived and distinguished by the animal.

Method

Subject. The subject for both experiments was an experimentally naive male bottlenose dolphin, Noah, born and housed at the dolphinarium of the Nuremberg zoo (Germany). Six years of age at the start of this investigation, Noah was still living together with his mother and younger sister in the breeding facility. Experimental sessions took place indoors in a circular tank 10 m in diameter (depth, 3.5 m).

Apparatus and Stimuli. For the numerosity stimuli, we used four different types of objects. All of the objects would sink in water. These were the following: Objects A, pieces of a hosepipe 0.07 m in diameter, cut straight into pieces 0.4 m in length; Objects B, T-pieces of synthetic-coated tubes (length, 0.25 m); Objects C, pieces of a hosepipe 0.03 m in diameter with a coated wire drawn through, formed to an "8" (length, 0.3 m); and Objects D, small wooden hoops (0.18-m diameter) weighted down with screws. With

the help of thin ropes, these objects were hung to floating hoops 1 m in diameter. Depending on the length of the ropes, objects could be submerged up to 1 m underneath the water surface. All the objects were dark-colored; thus, they would stand out well from the white walls of the tank. The floating hoops with the object arrays were hung from hooks at the side of the tank (see Figure 1).

Procedure. Before starting with the discrimination training, the animal received some habituation training in the procedure. He learned to be sent to and wait at the stationing point until given his starting signal and then to swim back and touch one object under water.

In a two-choice discrimination paradigm performed afterward, the dolphin had to choose between two simultaneously presented stimuli. For each trial, two floating hoops with a defined number of objects were hung from the hooks at a fixed distance of 1 m from each other. One trial of a session went as follows. The dolphin was sent to a stationing point at the opposite side of the tank, where the animal remained at the wall, holding his head out of the water. The two stimuli were hung into the pool. A short whistle blow was the starting signal for the dolphin to swim toward the stimuli. He made his choice between the two stimuli by touching one object of one of the arrays with the tip of his snout. Thereafter, depending on the animal's choice, there were two possibilities of feedback.

1. In a correct trial, the dolphin would touch an object of the stimulus with the lower number of objects, which was followed by a continuous whistle blow (secondary reinforcer) and food (primary reinforcer).

2. An incorrect choice, in which the animal would touch an object of the stimulus with the larger number of objects, was followed by a specific noncontinuous whistle blow without food being provided. The subsequent trial was then a correction trial in which stimuli with

exactly the same numerosity were presented. Correction trials were repeated until the dolphin chose the correct stimulus.

Each of the daily sessions consisted of 10 trials, not including correction trials, which were registered but not used to score discrimination performance. The left–right positions of the correct stimuli changed in a quasirandom sequence for each session (Fellows, 1967).

To test whether the dolphin was guided by inadvertent cuing of the experimenter, four sessions were conducted by an assistant who was not informed about the task. These sessions took place after the first transfer test.

Discrimination training with numerosities of 2 versus 5 (homogeneous stimuli). During initial discrimination training, only numerosities of 2 versus 5 were presented. There were two different types of stimuli for each numerosity, each consisting of identical objects, either Object A or Object B. This yielded four different stimuli pairings (see Figure 2 for an example), which were balanced across trials. The spatial configuration of the stimuli remained fixed throughout training; that is, the objects' points of attachment to the floating hoops did not alter for the numerosities of 2 and 5, respectively. The objects were fixed to the hoops at a minimal distance of 0.25 m from one another and at a depth of 0.7 m. The accuracy criterion was set at 90% correct responses over four consecutive sessions, with chance level at 50%.

Test 1: Transfer to novel objects and heterogeneous stimuli. This test was designed to examine whether the dolphin's performance would deteriorate when novel objects formed the numerosity stimuli. Thus, the only parameter changed during this transfer test was the object composition of the stimuli. Over four sessions, novel objects (Types C and D) were used exclusively. Each stimulus was still homogeneous, and again, object type was counter-
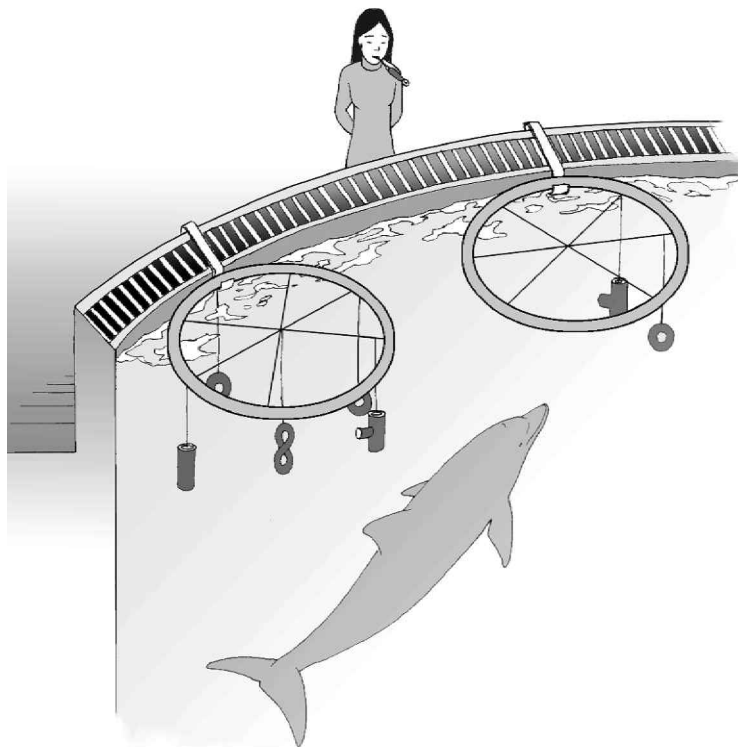

Figure 1. Schematic overview of a test situation with heterogeneous stimuli in Experiment 1.
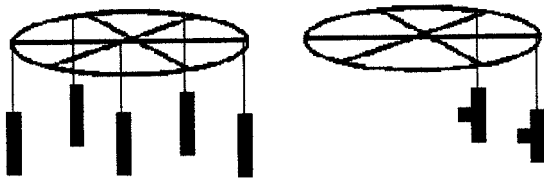
Figure 2. Example of a stimulus pairing in Experiment 1. This configuration was used during discrimination training with homogeneous stimuli (Objects A and B).

balanced with numerosity value across trials. In the next four sessions, only new heterogeneous stimuli were presented, with each numerical stimulus being composed of different objects randomly chosen from all four object types.

Test 2: Transfer to new spatial configurations. So far, each of the numerosities of 2 and 5 had been presented in the same or at least very similar spatial configurations. It was possible that the dolphin had learned to discriminate the object patterns of the stimuli, rather than their numerosities. A test with new spatial configurations was thus performed. For each trial of a session, new configurations were presented for the numerosities of 2 and 5. Again, object types varied within each numerical stimulus. The objects were fixed at random points to the floating hoops at different depths (0.5–1.0 m under water) at a minimal distance of 0.25 m. This yielded heterogeneous object arrays differing in configuration in a three-dimensional space. Depending on the spatial configuration, stimuli were classified as either *dense* or *wide*. Half of the numerosities of 5 were defined as dense, with the farthest objects being less than 0.8 m apart; the other half were defined as wide. The same applied to the numerosities of 2, but with the farthest objects having to be less than 0.5 m apart to be classified as dense. All combinations of dense and wide stimuli were evenly distributed among sessions. Sessions were continued until the animal reached criterion.

Test 3: Transfer to new numerosities of 3 versus 4. The previous transfer tests were conducted using the reward and feedback contingencies that applied during discrimination training, including a correction procedure for incorrect choices. As a preparation for the next experimental step, first one (two sessions) then two (four sessions) extinction trials were introduced per session. In these trials, independently of the animal's choice, no feedback was given by the experimenter—that is, the dolphin received neither whistle blows nor food. All other trials proceeded as before. Extinction trials were distributed randomly among a session, with two restrictions: They were never introduced as the first or the last trial of a session, and there were at least two reinforced trials between two extinction trials.

So far, the dolphin had been confronted with only two different numerosities. This test was intended to evaluate whether the animal would still be able to choose correctly the stimulus with the lower number between two new numerosities. The new stimuli pairing of 3 versus 4 was presented in extinction trials only. This made sure that no learning process for the new stimuli occurred during the presentation of these test trials. For all the trials of a session, we continued to use heterogeneous stimuli with differing spatial configurations. In a session of 10 trials, there were 9 trials with two versus five objects and 1 trial with three versus four. One of the trials with the familiar number arrays was also an extinction trial, to prevent a consistent association of novelty with nonreward, which might have disrupted discrimination. Overall, 12 test trials distributed over 12 sessions were conducted for the new numerical stimuli.

Results

During the five habituation training sessions, the animal learned the procedure of stationing at the opposite side of the tank, waiting for the starting signal, and then swimming toward an object underwater and touching it. After the dolphin was able to reliably follow the general procedure, discrimination training started.

Discrimination training of the numerosities of 2 versus 5, using homogeneous stimuli. In this first phase of training, the animal reached criterion with the 18th session. His performance for the last 4 sessions yielded an average of 92.5% correct choices. During learning, he first developed a side preference, choosing only the stimulus to his right and then, later, the one to his left throughout a session. From Session 11 on, he reduced this behavior, resulting in a continuous increase in performance level.

Test 1: Transfer to novel objects and heterogeneous stimuli. When novel objects with homogeneous stimuli were confronted, the dolphin's performance level averaged 90% over the four sessions. For the heterogeneous stimuli composed of all four object types, the animal reached 87.5% correct responses. As compared with the four final sessions of the discrimination training phase, there were no significant differences in performance [novel objects, $t(6) = 0.603$, $p > .05$; heterogeneous stimuli, $t(6) = 1.412$, $p > .05$].

In the four sessions conducted afterward by an assistant who was not familiar with the task, the dolphin made 90% correct choices.

Test 2: Transfer to new spatial configurations. With stimuli of new object patterns, the dolphin's performance level declined to an average of 67.5% over the first four sessions. This proved to be significantly different from the prior level reached during discrimination training [$t(6) = 5.345$, $p < .05$]. Until Session 14, his performance remained clearly lower than criterion level (72.1% correct choices over 140 trials).

The animal's percentage of correct choices over these first 14 sessions differed according to the four combinations of wide and dense stimuli (see Table 1). The dolphin performed best with stimuli that differed in density. The latter of these combinations (two dense vs. five wide) was most similar to the spatial configurations used during discrimination training. From the 15th session onward, the animal's performance level continuously increased and reached criterion with the 21st session (92.5% correct choices over the last 4 consecutive sessions).

Test 3: Transfer to the new numerosities of 3 versus 4. Introduction of extinction trials did not affect the behavior of the animal. His performance level over all six sessions averaged 96.7%; thus, he performed even better than he had during any previous phase of the experiment.

Table 1
Performance in Transfer to New Spatial Configurations (Test 2) Over the First 14 Sessions According to the Four Combinations of Dense and Wide Stimuli

Stimulus Combinations	Correct Choices (%)
2 wide versus 5 wide	57.1
2 dense versus 5 dense	60.0
2 wide versus 5 dense	80.0
2 dense versus 5 wide	91.4

For the 12 sessions in which the new numerical stimuli of 3 versus 4 were introduced, Noah's performance averaged 91.7%. In 11 of the test trials with the new numerosities, the animal made a correct choice. According to the binomial distribution, this result was significant ($p < .01$).

Behavioral observations. During the first sessions of discrimination learning, the dolphin approached the stimuli in two ways. Either he would swim straight toward them, stop at a distance of about 1 m and move his head slowly from one stimulus to the other before making his choice, or else he would swim along the right side of the tank and turn toward the stimuli with his right side about 1 m in front of them, then choose the first stimulus to his right or pass it by and choose the left stimulus. In later phases of the experiment, when he was performing close to criterion, the dolphin would stop under water close to the stationing point with his head turned straight toward the stimuli and then swim very fast to the stimulus of his choice. This behavior suggests that the dolphin made his choice while he was still about 10 m away from the stimuli. In the first sessions of Test 2, when new spatial configurations were presented, the animal's choice behavior was more similar to that during the first phase of discrimination learning. Overall, the dolphin's behavior did not offer any explicit indication about the sensory modalities applied for the discrimination of the stimuli. He could have used visual perception, information obtained through echolocation, or both.

Discussion

Of chief interest in Experiment 1 was how stimulus features covarying with numerosity would influence the dolphin's discrimination performance. Among the confounding parameters during the first phase of discrimination learning with homogeneous stimuli were overall brightness and density of the stimuli, object type, size, and arrangement. Thus, the dolphin could possibly have chosen among a large variety of attributes to accomplish the discrimination task. He reached the learning criterion very quickly, especially considering that he was naive to any experimental two-choice paradigm.

With the introduction of novel elements, as well as with heterogeneous stimuli, the animal's performance did not decline, as compared with prior level. Thus, it is quite unlikely that the subject based his discrimination on distinct features of single objects. Similar results in transfer to heterogeneous sets have been found for most of those animals tested—for example, pigeons (Emmerton et al., 1997), a gray parrot (Pepperberg, 1987), and rhesus monkeys (Brannon & Terrace, 1998).

Testing for stimuli with novel spatial configurations yielded a different picture. The result, a decline in performance over many sessions, suggests an influence of element pattern during discrimination. This assumption is supported by the fact that the dolphin's performance was at criterion level only for those element patterns most similar to the ones used during discrimination learning. For half of the stimuli pairings in which both of the spatial configurations were either dense or wide, his percentage of correct responses was close to chance. Together with the fact that the subject needed more sessions to reach criterion in this test than during discrimination training, these findings suggest that element pattern was the most important stimulus attribute for discrimination. However, after intensive training, the dolphin was able to choose correctly. At this point, his choice was obviously no longer influenced by the spatial configurations of the stimuli.

Comparisons with other animals tested in similar experimental paradigms are difficult, since in most of these studies, element configuration was either varied from the beginning to prevent rote learning (e.g., Brannon & Terrace, 1998; the same applies to the study by Xu & Spelke, 2000, on 6-month-old infants) or else was not explicitly tested. The data for pigeons collected by Emmerton et al. (1997) suggest that for one test group, dot pattern might have influenced discrimination performance, since transfer to novel stimuli was poor. Evidence for the irrelevance of element configuration was given by the gray parrot, Alex (Pepperberg, 1987), who succeeded in labeling quantities of objects presented in novel random arrays. We assume that for the subject in our study, the element pattern was a very strong stimulus attribute, and so he relied mostly on simple pattern recognition to accomplish the task. The animal thus got confused when confronted with novel spatial configurations.

In the third test, the animal showed successful transfer to the new intervening numerosities 3 versus 4. This demonstrates the dolphin's flexibility in accomplishing the task, since he is not fixed to the learned discrete numbers of elements but, obviously, applies a more abstract rule for the discrimination of stimuli. Yet we have to be careful about these findings, since the animal did not necessarily need to base his discrimination on numerosity. There were still some confounding attributes, such as size (volume) of the constituting objects and overall darkness of the stimuli. Although the object types differed in size, the stimuli with the higher numerosity would have a greater overall size and darkness for most stimuli pairings. Furthermore, we could also argue that a configuration of two objects is more similar to a configuration of three than of four objects, and the animal could have made his choice upon perceptual similarities. Thus, at this experimental stage, it would be jumping to conclusions to speak of an ordinal representation of amounts. To defend this position, it was necessary to control for other confounding parameters and to test for transfer to novel numerosities outside the former range. This was accomplished in Experiment 2.

EXPERIMENT 2

So far, we have demonstrated that a dolphin is able to discriminate between different numbers of objects. In the previous experiment, we could not test for all stimulus

dimensions covarying with numerosity. The overall size of the constituent objects of a given stimulus could be confounded with numerosity. New sets of stimuli were thus designed to evaluate the influence of this dimension. These stimuli, consisting of two-dimensional elements, furthermore offered the opportunity to test for transfer to purely visually perceivable stimuli in discrimination.

Method

Subject. The investigation was conducted with the same subject as that in Experiment 1, Noah, then 7 years of age. During the course of this experiment, he was moved from the breeding to the show facility, where he is living together with three adult females. Experimental sessions then took place in a circular indoor tank 14 m in diameter (depth, 4.7 m) adjacent to the show tank.

Apparatus and Stimuli. The stimuli consisted of elements (pieces of white adhesive film) that were stuck on black square panels (side lengths, 1 m; 6 mm thick) of synthetic material. Minimum distance of the elements to the edges of the panel, as well as to one another, was 0.1 m. The size and form of the elements varied according to the experimental phase (see below). The panels were hung underwater in stainless steel frames, which were fixed by steel hooks to the tank side (see Figure 3). The distance between the frames and, thus, the stimuli was fixed at 1 m.

Procedure. Apart from the apparatus, the procedure was similar to the one in Experiment 1. After the dolphin was sent to his stationing point at the opposite side of the tank, the experimenter hid behind the tank. Two stimuli were placed in the water by an assistant who kept observing the animal. The starting signal prompted the dolphin to swim toward the panels. He made his choice by touch-ing one of the panels with the tip of his snout and pushing the panel slightly upward. The chosen panel moved above the water surface, so that the experimenter could give the dolphin the correct /incorrect signal. Each session comprised 10 trials. Accuracy criterion was set at 90% correct choices over four consecutive sessions (chance level, 50%).

Before starting a session, the dolphin was separated from his group. He was given a sign to swim into the adjacent tank. When the sign had to be given three times and the animal was still not re-acting accordingly, his motivation was judged as too low, and thus, the session was postponed to the next training unit.

First, three habituation training sessions were carried out in order to acclimate the dolphin to the new apparatus. During these ses-sions, only one of the panels was placed underwater, and the animal was reinforced for touching it and pushing it upward.

Test 1: Transfer to two-dimensional stimuli. Since we did not expect an immediate transfer from the three-dimensional to the two-dimensional stimuli, we used three means to facilitate transfer.

1. Only numerosity stimuli of 2 versus 5 were used.

2. In the first three transfer sessions, we presented either three-dimensional familiar stimuli or two-dimensional novel stimuli in a trial. A session always started with four trials in which three-dimensional stimuli were used and continued with six trials of two-dimensional stimulus pairings.

3. The elements of the novel stimuli were comparable in form and size to the three-dimensional objects. Forms used were filled circles, rectangles, and Ts, with surface area varying from 0.03 to 0.04 m^2 (see Figure 4A). The novel and familiar stimuli were het-erogeneous and varied in their spatial configuration.

From Session 4 onward, only the novel two-dimensional stimuli were presented to the dolphin until performance reached criterion.
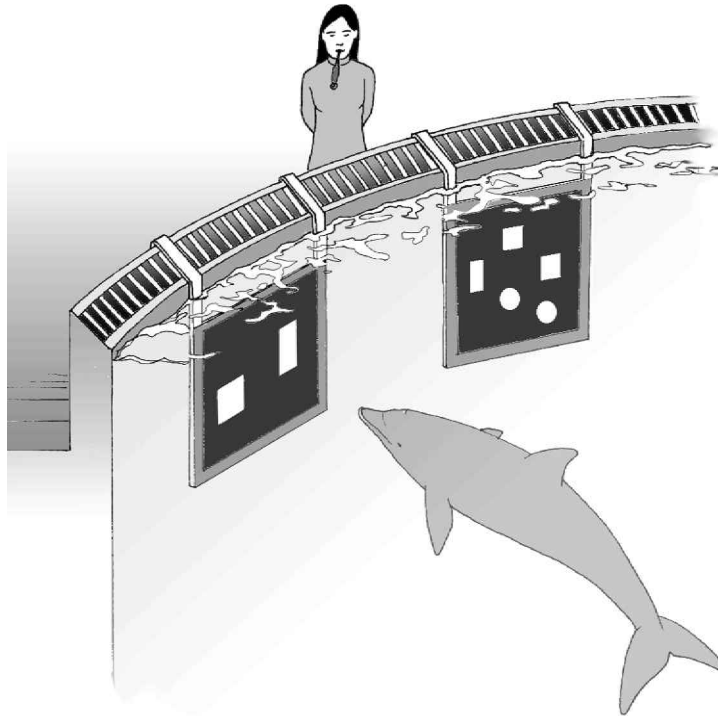

Figure 3. Schematic overview of a test situation in Experiment 2 with stim-uli with an overall equal surface area for the elements.
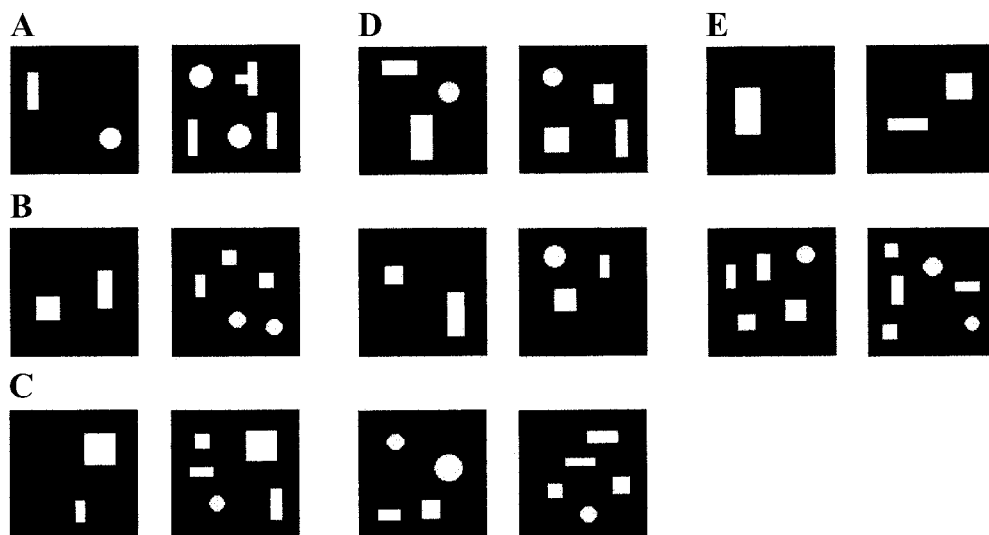
Figure 4. Examples of stimulus pairings for each test phase (Experiment 2). (A) Test 1: transfer to two-dimensional stimuli. (B) Test 2: overall equal surface area of stimulus elements. (C) Test 3: mixed sizes of stimulus elements. (D) Test 4: transfer to new intervening numerosities (3 vs. 4, 2 vs. 3, and 4 vs. 5). (E) Test 5: transfer to new numerosities outside the former range (1 vs. 2 and 5 vs. 6).

Test 2: Stimuli with overall equal surface area of elements. For all the sessions of Experiment 1 and those of Experiment 2 carried out so far, numerosity had covaried with the overall size of the stimuli's elements. This stimulus feature could have affected the dolphin's performance. To test whether this applied, we introduced stimuli that differed in numerosity but had an equal overall surface area.

Elements of three different forms (circles, rectangles, and squares) and four different sizes were used (see Figure 4B). Each stimulus was composed of elements of different forms but identical sizes, combined in such a way that, in every choice, the overall surface area of the two presented stimuli was the same (either 0.072 or 0.125 m²). Sessions were continued until criterion.

Test 3: Stimuli with mixed-sized elements. During the previous test, the dolphin could have learned to compare the sizes of single elements of two presented stimuli and to base his choice on these size differences. This test was thus designed to examine whether the animal was affected by the elements' sizes.

The size and form of the elements were the same as those in Test 2. They were randomly distributed among stimuli, so that each numerosity stimulus would comprise elements of different forms and sizes (see Figure 4C). The elements of the stimuli in a given trial thus had (1) the same overall surface area, (2) a larger surface area for the numerosity of 2, or (3) a larger surface area for the numerosity of 5. These three situations were distributed evenly but randomly among sessions.

Test 4: Transfer to new numerosities. In the previous transfer tests, the subject was always given a feedback for his choices, and incorrect choices were followed by correction trials. Most of the new number pairings in the test series described here were introduced in extinction trials only. The dolphin was familiar with extinction trials from Experiment 1. Therefore, only four retraining sessions with two extinction trials each were carried out, to make sure that the animal was still accustomed to this procedure.

For all the numerical stimuli used in these transfer tests, the elements differed in size and form but had an equal overall surface area for each pairing. Test sessions with extinction trials followed the same procedure as that in Test 3 of Experiment 1—that is, a session of 10 trials consisted of 8 trials with the usual feedback con-

tingencies and 2 extinction trials, one with familiar and the other with new number pairings.

First, we tested three new number pairings with intervening numerosities (3 vs. 4, 2 vs. 3, and 4 vs.5), depicted in Figure 4D. For each of the three new number pairings, there were 12 test trials distributed over 12 sessions—that is, 36 test trials in all.

Next, two new number pairings consisting of familiar numerosity stimuli were introduced: 2 versus 4 and 3 versus 5. They were not presented in extinction trials, since owing to their familiarity and a disparity of 2, we did not expect a decline in performance for these pairings. Four sessions were conducted, with each of the new stimulus pairings occurring two times per session. All the other trials consisted of familiar number pairings (2 vs. 5, 2 vs. 3, 3 vs. 4, and 4 vs. 5).

With all the numerical stimuli pairings tested so far, the dolphin could have based his discrimination on a simple rule that would yield a high performance level—*choose 2; if not available, choose 3*—or, at least for part of the trials, he could have relied on *never choose 5*. Another explanation for his correct choices is a reliance on a more abstract rule, *always choose the lower number* by comparing the two stimuli in a given trial.

The next transfer test was thus intended to substantiate the latter alternative. For the first time, numerosities outside the former range were introduced, presented in pairings of 1 versus 2 and 5 versus 6 (see Figure 4E). Twelve extinction trials for each of the new pairings were carried out. These were distributed over 24 sessions.

Results

During habituation training, the dolphin quickly got accustomed to the new apparatus and the procedure of pushing the panels. With the third habituation session, he was performing well and without hesitation. In the course of the experiment, there were variations in the animal's behavior before starting a session. For most of the sessions conducted, the dolphin reacted promptly after being given the sign to swim into the adjacent tank, but occasionally, it was difficult to separate him from his

group. This behavior was predominant during test sessions with the number pairing 4 versus 5. In 10 of these 12 test sessions, it took two or three attempts to separate the animal.

Test 1: Transfer to two-dimensional stimuli. The dolphin reached criterion with the ninth session. Overall performance level for the last four consecutive sessions was 90% correct responses. From the third session onward, the animal performed above chance level. A site preference occurred only in the first two sessions.

Test 2: Stimuli with overall equal surface area of elements (same brightness). The animal's performance declined to 72.5% for the first 4 sessions. As compared with the last 4 sessions of Test 1, there was a significant difference [$t(6) = 3.212, p < .05$]. The dolphin reached criterion with the 11th session.

Test 3: Stimuli with mixed-sized elements. Over the four consecutive sessions, the subject had a performance level of 87.5%. There was no significant difference from the level of the last four sessions in the prior test [$t(6) = 0.739, p > .05$].

Test 4: Transfer to new numerosities. The overall results for the transfer tests with new numerosities are depicted in Figure 5.

In the four retraining sessions with extinction trials, the animal reached 95% correct responses. As before in Experiment 1, his performance was better for sessions with extinction trials.

The first tests examined the dolphin's transfer performances for three new number pairings consisting of intervening numerosities. In 11 of the 12 test trials for the new numerosities of 3 versus 4, the dolphin made a correct choice. All 12 test trials with numerosities of 2 versus 3 were performed correctly by the animal. According to the binomial distribution, both results were significant ($p < .01$). For the pairings of 4 versus 5, the animal did not reach a significant performance. His choice was correct in 7 of the 12 test trials ($p > .05$).

The two new number pairings consisting of familiar numerosity stimuli, 2 versus 4 and 3 versus 5, were introduced in trials with a normal feedback and reward

procedure. In each of the 16 trials with these new pairings, the animal made a correct response.

Regarding the results for the new numerosities outside the former range, the dolphin achieved a significant transfer performance with both of the new number pairings of 1 versus 2 and 5 versus 6. In the 12 test trials, the animal chose correctly in 11 (1 vs. 2) and 10 (5 vs. 6) trials, with $p < .01$ and $p < .05$, respectively.

Discussion

Transfer to the new two-dimensional stimuli occurred rapidly. Although we used several means to facilitate transfer, this result cannot be explained exclusively by the transfer technique. It suggests that the dolphin had perceived the three-dimensional stimuli visually, and thus he only had to get accustomed to the new picture and the contrast reversal. However, we cannot exclude the role of echolocation in discrimination of the three-dimensional set of stimuli. A cross-modal transfer from echolocation to vision and vice versa has been demonstrated experimentally in the bottlenose dolphin (Pack & Herman, 1995). However the subject in our study managed to perceive and discriminate the novel two-dimensional stimuli, we suggest a conceptual transfer as an explanation for his high performance level, achieved after a few sessions.

Controlling for the confounding parameter of the overall surface area of the elements in the second test demonstrated that this stimulus attribute obviously influenced the animal's decision in the discrimination task. Still, the dolphin's performance over the first four sessions was clearly above chance level, which suggests that overall element size or brightness was not the only basis for discrimination. If we consider that this parameter covaried with numerosity for all the sessions prior to this test, it is not astonishing that the dolphin used it as an additional clue for choosing correctly. However, after a few sessions, he reached criterion level, demonstrating that, by then, this stimulus feature was irrelevant for his response. The results of the following test, in which elements of different sizes were mixed within each given stimulus, indicates that the dolphin had not focused ex-
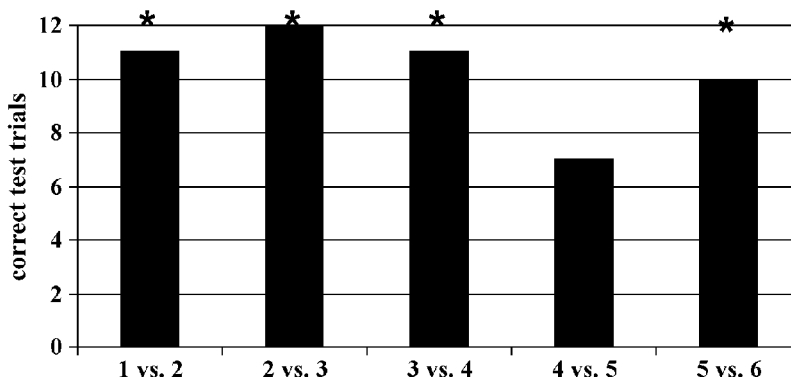

Figure 5. Overall results for transfer tests with new numerosities. Asterisks indicate significant performances according to the binomial distribution at the 5% level.

clusively on the sizes of the elements. By then, we can assume that the animal discriminated the stimuli primarily on the basis of the numerosity feature.

Do these results up to now support a strategy of "last resort," as was suggested by Davis and Perusse (1988)? Control of behavior by alternative sources of information has been found in other investigations (e.g., Breukelaar & Dalrymple-Alford, 1998; Davis & Bradford, 1986; Fetterman, 1993). Yet, in other studies, some nonnumerical cues were either controlled during habituation training or covaried during test sessions, the latter being the case for most experiments in which numerosity was defined by the number of food items (e.g., Olthof, Iden, & Roberts, 1997; Washburn & Rumbaugh, 1991).

If application of numerical abilities is not the strategy of choice for a number of animal species, then how about humans? Children do not acquire most concepts of numerosity spontaneously, and even adults may resort to visual or spatial directions or other cues (Lenneberg, 1971). Recent studies (Clearfield & Mix, 1999, 2001) have demonstrated that infants respond to changes in contour length or area of a given stimulus, rather than to changes in number. The authors suggested that estimating amount of spatial extent is less effortful; therefore, infants do so, whenever possible. This explanation may also hold for many of the findings when animal behavior is not exclusively controlled by numerosity. We suggest that animals, as well as humans, use a simple strategy to discriminate. For the dolphin in our study, the strategy was a discrimination based on the configuration of stimulus elements. The data demonstrate that the animal was able to use numerosity after confounding stimulus dimensions were controlled.

The dolphin accomplished a spontaneous transfer to most of the novel intervening numerosities, as well as to both of the new numerosities outside the former test range. For the intervening numerosities, it could be argued that the discrimination was controlled by perceptual similarities to the learned numerosities of 2 and 5 (as was pointed out earlier in Experiment 1). However, this explanation would not explain performance with the novel numerosities outside the former range—that is, the pairings of 1 versus 2 and 5 versus 6, where there was a reversal of positive to negative stimuli and vice versa. This was the critical stage of the test series, since prior to the presentation of these pairings, the numerosity of 2 was always positive and the numerosity of 5 was always negative. All other intervening numerosities could have either a positive or a negative value. The animal's significant performance with these novel pairings outside the former number range clearly demonstrates a representation of ordinal relations among numerosities.

The dolphin did not reach a significant performance level in the tests with the pairings of 4 versus 5. There are two possible explanations for this result. Either a loss in motivation was responsible for the subject's low performance, or the dolphin had arrived at the threshold of nu-

merosities he was still able to discriminate. Numerous investigations on animals (e.g., Boysen, 1993; Thomas, Fowlkes, & Vickery, 1980), as well as on human infants (e.g., Strauss & Curtis, 1981; Xu & Spelke, 2000), have suggested that discriminations of numerosity follow Weber's law: With pairings of higher number but constant absolute difference, the relative difference becomes smaller and is, thus, more difficult to discriminate. The result of the ensuing transfer test, in which the dolphin achieved a high performance level with the number pairing of 5 versus 6, clearly refutes the latter explanation.

The high performances with numerosity pairings outside the former test range provide substantial evidence that the dolphin represented ordinal relations among numerosities. It is yet unclear whether the pairing of 5 versus 6 marks the threshold for the discrimination of numerical differences. In future research, it would be interesting to examine whether accuracy follows a distance effect by testing pairings with different numerical disparities.

GENERAL DISCUSSION

According to Gallistel and Gelman (1992), the ability to base behavior on the numerosity of a set independently of its other attributes and to respond in the same way even to novel numerosities suggests a mapping from all instances of a given numerosity to a mental representative of that category of sets. We have demonstrated here that a dolphin can fulfill both of these prerequisites. Despite the finding that his behavior was also controlled by some other nonnumerical features before eliminating the covariance of these attributes, the spontaneous transfer to novel number pairings strongly suggests a representation of ordinal relations among numerosities. This conclusion is underlined by the fact that the subject had only experienced a few of the possible number pairings during training and yet was capable of generalizing its discrimination in a categorical way to new pairings.

To further investigate the richness of the dolphin's representation of number, it would be interesting to check for abilities in arithmetic processes, such as addition and subtraction. This could be eased by the use of symbols or representatives for distinct numerosities. Davis and Perusse (1988) called for the demonstration of a concept of number established by the capacity to transfer across sense modalities. This capability has been found for some animals (e.g., jackdaws, B. V. Dawson, 1961; rats, Meck & Church, 1984), as well as for small infants (Starkey, Spelke, & Gelman, 1983). Dolphins are capable of solving complex cognitive tasks, using their auditory system, either hearing or echolocation (Herman & Gordon, 1974; Roitblat, Penner, & Nachtigall, 1990), and moreover, a cross-modal transfer between vision and echolocation has been demonstrated in a matching-to-sample paradigm (Pack & Herman, 1995). Now that a bottlenose dolphin has been found to be able to solve a visual numerical task, this species would be a good sub-

ject to test for transfer across the visual and the acoustic modalities.

REFERENCES

BAUER, G. B., & JOHNSON, C. M. (1994). Trained motor imitation by bottlenose dolphins (*Tursiops truncatus*). *Perceptual & Motor Skills*, **79**, 1307-1315.

BOYSEN, S. T. (1993). Counting in chimpanzees: Nonhuman principles and emergent properties of number. In S. T. Boysen & E. J. Capaldi (Eds.), *The development of numerical competence: Animal and human models* (pp. 39-59). Hillsdale, NJ: Erlbaum.

BOYSEN, S. T., & CAPALDI, E. J. (1993). *The development of numerical competence: Animal and human models*. Hillsdale, NJ: Erlbaum.

BRANNON, E. M., & TERRACE, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science*, **282**, 746-749.

BREUKELAAR, J. W. C., & DALRYMPLE-ALFORD, J. C. (1998). Timing ability and numerical competence in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, **24**, 84-97.

CLEARFIELD, M. W., & MIX, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science*, **10**, 408-411.

CLEARFIELD, M. W., & MIX, K. S. (2001). Amount versus number: Infants' use of area and contour length to discriminate small sets. *Journal of Cognition & Development*, **2**, 243-260.

DAVIS, H., & BRADFORD, S. A. (1986). Counting behavior by rats in a simulated natural environment. *Ethology*, **73**, 265-280.

DAVIS, H., & MEMMOTT, J. (1982). Counting behavior in animals: A critical evaluation. *Psychological Bulletin*, **92**, 547-571.

DAVIS, H., & PERUSSE, R. (1988). Numerical competence in animals: Definitional issues, current evidence, and a new research agenda. *Behavioral & Brain Sciences*, **11**, 561-615.

DAWSON, B. V. (1961). Counting in jackdaws. *Behaviour*, **18**, 229-238.

DAWSON, W. W. (1980). The cetacean eye. In L. M. Herman (Ed.), *Cetacean behavior: Mechanisms and functions* (pp. 53-100). New York: Wiley.

DEHAENE, S. (1997). *The number sense: How the mind creates mathematics*. Oxford: Oxford University Press.

EMMERTON, J., LOHMANN, A., & NIEMANN, J. (1997). Pigeons' serial ordering of numerosity with visual arrays. *Animal Learning & Behavior*, **25**, 234-244.

FELLOWS, B. J. (1967). Chance stimulus sequences for discrimination tasks. *Psychological Bulletin*, **67**, 87-92.

FETTERMAN, J. G. (1993). Numerosity discrimination: Both time and number matter. *Journal of Experimental Psychology: Animal Behavior Processes*, **19**, 149-164.

FORESTELL, P. H., & HERMAN, L. M. (1988). Delayed matching of visual materials by a bottlenosed dolphin aided by auditory symbols. *Animal Learning & Behavior*, **16**, 137-146.

GALLISTEL, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.

GALLISTEL, C. R., & GELMAN, R. (1992). Preverbal and verbal counting and computation. *Cognition*, **44**, 43-74.

HAUSER, M. D., MACNEILAGE, P., & HAUSER, L. B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings of the Royal Society of London*, **267**, 829-833.

HERMAN, L. M. (1986). Cognition and language competencies of bottlenosed dolphins. In R. J. Schusterman, J. A. Thomas, & F. G. Woods (Eds.), *Dolphin cognition and behavior: A comparative approach* (pp. 221-252). Hillsdale, NJ: Erlbaum.

HERMAN, L. M. (1990). Cognitive performance of dolphins in visually-guided tasks. In J. Thomas & R. Kastelein (Eds.), *Sensory abilities of cetaceans* (pp. 455-462). New York: Plenum.

HERMAN, L. M., & GORDON, J. A. (1974). Auditory delayed matching in the bottlenosed dolphin. *Journal of the Experimental Analysis of Behavior*, **21**, 19-26.

HERMAN, L. M., HOVANCIK, J. R., GORY, J. D., & BRADSHAW, G. L. (1989). Generalization of visual matching by a bottlenosed dolphin (*Tursiops truncatus*): Evidence for invariance of cognitive performance with visual or auditory materials. *Journal of Experimental Psychology: Animal Behavior Processes*, **15**, 124-136.

HERMAN, L. M., MORREL-SAMUELS, P., & PACK, A. A. (1990). Bottlenosed dolphin and human recognition of veridical and degraded video displays of an artificial gestural language. *Journal of Experimental Psychology: General*, **119**, 215-230.

KOEHLER, O. (1937). Können Tauben "zählen"? [Can pigeons "count"?] *Zeitschrift für Tierpsychologie*, **1**, 39-48.

KOEHLER, O. (1943). "Zähl" Versuche an einem Kolkraben und Vergleichsversuche an Menschen ["Counting" study with a raven and comparative research with people]. *Zeitschrift für Tierpsychologie*, **5**, 575-712.

LENNEBERG, E. H. (1971). Of language, knowledge, apes, and brains. *Journal of Psycholingual Research*, **1**, 1-29.

MECK, W. H., & CHURCH, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, **9**, 320-334.

MECK, W. H., & CHURCH, R. M. (1984). Simultaneous temporal processing. *Journal of Experimental Psychology: Animal Behavior Processes*, **10**, 1-29.

MITCHELL, R. W., YAO, P., SHERMAN, P. T., & O'REGAN, M. (1985). Discriminative responding of a dolphin (*Tursiops truncatus*) to differentially rewarded stimuli. *Journal of Comparative Psychology*, **99**, 218-225.

OLTHOF, A., IDEN, C. M., & ROBERTS, W. A. (1997). Judgements of ordinality and summation of number symbols by squirrel monkeys (*Saimiri sciureus*). *Journal of Experimental Psychology: Animal Behavior Processes*, **23**, 325-339.

PACK, A. A., & HERMAN, L. M. (1995). Sensory integration in the bottlenosed dolphin: Immediate recognition of complex shapes across the senses of echolocation and vision. *Journal of the Acoustical Society of America*, **98**, 722-733.

PEPPERBERG, I. M. (1987). Evidence for conceptual quantitative abilities in the African grey parrot: Labeling of cardinal sets. *Ethology*, **75**, 37-61.

RILLING, M. (1993). Invisible counting animals: A history of contributions from comparative psychology, ethology, and learning theory. In S. T. Boysen & E. J. Capaldi (Eds.), *The development of numerical competence: Animal and human models* (pp. 3-37). Hillsdale, NJ: Erlbaum.

ROITBLAT, H. L., PENNER, R. H., & NACHTIGALL, P. E. (1990). Matching-to-sample by an echolocating dolphin (*Tursiops truncatus*). *Journal of Experimental Psychology: Animal Behavior Processes*, **16**, 85-95.

STARKEY, P., SPELKE, E. S., & GELMAN, R. (1983). Detection of intermodal numerical correspondences by human infants. *Science*, **222**, 179-181.

STRAUSS, M. S., & CURTIS, L. E. (1981). Infant perception of numerosity. *Child Development*, **52**, 1146-1152.

THOMAS, R. K., FOWLKES, D., & VICKERY, J. D. (1980). Conceptual numerousness judgements by squirrel monkeys. *American Journal of Psychology*, **93**, 247-257.

WASHBURN, D. A., & RUMBAUGH, D. M. (1991). Ordinal judgements of numerical symbols by macaques (*Macaca mulatta*). *Psychological Science*, **2**, 190-193.

XIA, L., EMMERTON, J., SIEMANN, M., & DELIUS, J. D. (2001). Pigeons (*Columba livia*) learn to link numerosities with symbols. *Journal of Comparative Psychology*, **115**, 83-91.

XU, F., & SPELKE E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, **74**, B1-B11.