

Pigeons (*Columba livia*) Learn to Link Numerosities With Symbols

Li Xia
Universität Konstanz

Jacky Emmerton
Purdue University

Martina Siemann and Juan D. Delius
Universität Konstanz

After responding to each element in varying, successive numerosity displays, pigeons (*Columba livia*) had to choose, out of an array of symbols, the symbol designated to correspond to the preceding number of elements. After extensive training, 5 pigeons responded with significant accuracy to the numerosities 1 to 4, and 2 pigeons to the numerosities 1 to 5. Several tests showed that feedback tones accompanying element pecks, the familiarity of element configurations, and the shape of the elements were not crucial to this performance. One test, however, indicated that the number of pecks issued to the elements was important for numerosities above 2. An additional test confirmed that the birds chose the symbol that corresponded to a particular numerosity rather than the positions that the symbols had held during training.

Numerical abilities in animals have attracted the interest of students of behavior since early on, not least because they were thought to represent evolutionary antecedents of human nonverbal counting abilities (Butterworth, 1999; Dehaene, 1997). Several avian and mammalian species have accordingly been examined with a variety of numerical tasks (e.g., Boysen, 1993; Davis & Albert, 1986; Davis & Pérusse, 1988; Koehler, 1949; see Rilling, 1993, for an overview). Some researchers have looked at animals' ability to discriminate the numbers of items in sets of simultaneously displayed elements (e.g., Emmerton, Lohmann, & Niemann, 1997; Koehler, Müller, & Wachholtz, 1935; Thomas, Fowlkes, & Vickery, 1980). In other experiments (e.g., Honig & Stewart, 1989), birds had to discriminate the proportions of different types of elements intermixed within an array. In all these studies with simultaneously displayed items, the animals did not necessarily have to discriminate the precise number of elements. Rather, they had to judge relative differences in numerosity. This was also true of studies in which animals had to make choices based on the relative numbers of sequentially presented stimuli in a series (Alsop & Honig, 1991; Fernandes & Church, 1982; Roberts, Macuda, & Brodbeck, 1995; Roberts & Mitchell, 1994). In studies that use sequential presentation, the temporal duration of

events has often been controlled to ensure that the animals discriminated on the basis of number and not just time. Other authors have investigated whether animals would discriminate relative differences in the number of responses they emit (Fetterman, 1993) or would be able to produce specified numbers of responses (e.g., Koehler, 1937; Platt & Johnson, 1971; Xia, Siemann, & Delius, 2000; Zeier, 1966).

Hooded crows were able to acquire a generalized matching or oddity concept when stimulus sets consisted of black or white cards, Arabic numerals (used as shapes), or numerosity arrays (Smirnova, Lazareva, & Zorina, 2000). In the latter stimulus sets, the color, shape, and positions of the individual items differed between the sample and matching comparison arrays while they shared the same numerosity. These birds successfully transferred their discrimination to stimulus sets consisting of novel Arabic numerals as well as to novel numerosity arrays with values of 5–8 instead of 1–4 that they had encountered earlier. Note, though, that the crows were not trained to match numerical symbols to array values.

Another kind of numerical task that more closely approximates what humans do when they count requires animals to form associations between the cardinal numbers of items in given sets and arbitrary symbols assigned to each numerosity set. Chimpanzees are capable of associating in a one-to-one fashion Arabic numerals with sets of items containing one to six elements (Matsuzawa, 1985) and even zero to eight elements (Boysen, 1993). An African Grey parrot has also learned to label sets of up to six items by vocalizing the appropriate number in an approximation of English for the quantity present in the set (Pepperberg, 1994). Extensive training was required before these animals could master these associations between symbols and numerical quantities. However, because their training with numbers was interwoven with other learning tasks, it is difficult to say how many training trials were

Li Xia, Martina Siemann, and Juan D. Delius, Allgemeine Psychologie, Universität Konstanz, Konstanz, Germany; Jacky Emmerton, Department of Psychology, Purdue University.

This research was supported by the Deutsche Forschungsgemeinschaft, Bonn, Germany. While collaborating on this research, Jacky Emmerton was on sabbatical from Purdue University and was a visiting professor at Konstanz University. We thank Manola Petrovici for assistance with running part of Experiment 2.

Correspondence concerning this article should be addressed to Juan D. Delius, Allgemeine Psychologie, Universität Konstanz, Konstanz 78457, Germany. Electronic mail may be sent to juan.delius@uni-konstanz.de.

necessary to specifically master the use of numerical symbols (see Boysen, 1993; Pepperberg, 1987).

Studies with primates have further demonstrated their ability to respond to symbols that were assigned to specific numbers of items. For instance, a chimpanzee learned to use a joystick to move the cursor around on a computer screen. If a numeral between 1 and 4 appeared on the screen, the chimpanzee first had to move the cursor to the numeral. Then it had to use the cursor to point successively at a corresponding number of boxes on the screen before pointing once more at the numeral to obtain a reward (Rumbaugh & Washburn, 1993). Rhesus monkeys have been trained to choose between two Arabic numerals to obtain a food reward containing a number of pellets corresponding to the numeral chosen. Training involved different combinations of numerals between 0 and 9, and the monkeys tended to choose the numeral representing the larger quantity of reward (Washburn & Rumbaugh, 1991). Similarly, squirrel monkeys have learned to choose Arabic numerals to obtain the corresponding number of food items. The numerals were 0, 1, 3, 5, 7, and 9. The monkeys chose the larger numerical value in pairs of numerals and also the largest value among sets of four numerals (Olthof, Iden, & Roberts, 1997). Although there is evidence that animals are capable of associating symbols with absolute numbers of items, this ability has mostly been demonstrated in primates and in an African Grey parrot, a species that is generally considered more intelligent than the pigeon.

In the experiments we now report, pigeons were trained to link sets containing from one to five visual elements with five corresponding letter symbols. Subsequent tests explored the stimulus characteristics on which the pigeons based their choices.

Experiment 1

Method

Subjects. Eight adult pigeons (*Columba livia*) of local stock were used. They were housed in individual cages in a well-ventilated and brightly illuminated (12-hr light-dark cycle) room. The pigeons had free access to water and grit but were kept food deprived at 85% to 90% of their free-feeding body weight.

Apparatus. A conventional conditioning chamber (33 × 34 × 33 cm) was used. One wall consisted of the face of a flat-screen computer monitor (28 × 20 cm, Zenith, Glencoe, IL) that was protected by a 2-mm thick glass plate and equipped with an infrared touch screen (Carroll Touch, Round Rock, TX). A 6.5 × 6.5-cm stimulus field was defined and centered within the dark monitor face 15 cm above the floor of the chamber. It was invisible during the shaping phase, but later it was shown framed by a white outline. Pecks within that field were detected and located by the touch screen with a 3 × 3-mm resolution. Two lights bulbs (2 W) that served as houselights were both located 19 cm above the floor midway on the left-hand wall. A solenoid-operated grain dispenser was situated on the roof of the chamber. When activated, it delivered millet grains into a receptacle (5.0 cm diameter, 3.5 cm high) located to the right of stimulus field (6 cm from the right-hand wall, 6 cm above the floor). All experimental events were controlled and recorded with a personal computer (Highscreen, Vobis, Munich, Germany) provided with a digital interface (Plug-In) and a program written in QuickBasic (1987).

Shaping. Using an autoshaping procedure, we first trained the pigeons to peck stimuli displayed on the monitor. Each daily session lasted 30 min. The houselights were continuously on. A single trial started with the presentation of a red-filled circle (2.7 cm diameter). If no response occurred, it extinguished after 8 s and six to eight grains of millet were

delivered. If the pigeon pecked anywhere (early sessions) or pecked the circle (later sessions) during the stimulus period, the circle was immediately removed, and the food reward was given. An intertrial interval lasting 30 s followed during which the monitor was dark. Then the next trial followed. Once a pigeon had pecked in 50% of the trials of a session, the intertrial interval was reduced to 5 s, and the stimulus presentation was increased to 100 s for the next session. Automatic rewards ceased, and only pecks to the stimulus or in its vicinity (within a tolerance of a 2-mm wide annular surround) were reinforced. Once the pigeon pecked correctly in 80% of the trials of a session, the diameter of the target stimulus was reduced by 0.5 cm for the next session. This was repeated until the pigeons fulfilled the criterion of responding accurately on 80% of the trials with a stimulus of 0.7 cm diameter. Six out of the 8 pigeons reached this final stage after 15 to 25 sessions. The remaining 2 pigeons did not reach the criterion within 50 sessions and were not tested.

Stimuli. The numerosity stimuli used for training consisted of between one and five elements displayed on the monitor within the 6.5 × 6.5 cm stimulus field now framed by a white outline. These elements were 0.7-cm diameter, red-filled circles shown on a black background (Figure 1). To begin with, each element bore a central black dot (0.3 cm diameter). For each numerosity, the circles could appear in 10 different spatial configurations within the stimulus field. The number symbols were the letters A, B, F, G, and H in red color. These letters were assigned in a one-to-one manner to the numerosities 1, 2, 3, 4, and 5 and were selected as symbols because they had more similar areas and form complexities than the corresponding Arabic numerals. They were 0.9 cm high, positioned within white outline circles of 1.5 cm diameter on a black background, and arranged within the stimulus field as shown in Figure 1. During the initial stages of the experiment, the stimulus field only included the symbols that were relevant at a particular stage.

Training. Pigeons were trained to associate the numerosity stimuli with the number symbols using a successive symbolic matching-to-sample procedure. Training was divided into blocks of trials. The number of trials per block was initially 60, later 80, and finally 100 trials. The daily sessions consisted of two blocks of trials (Table 1). A trial started with the display of a quasirandomly chosen numerosity stimulus. When the pigeon pecked an element of the array, the computer issued a brief audible feedback tone (0.1 s, 1.8 kHz), and the black central dot was removed, leaving that element filled with the red color. Any further pecks on this element had no consequences. The pigeons were required to peck each of the stimulus elements at least once, that is, until they had removed the black dot from all of them. The numerosity stimulus was then replaced by the symbol array. If the birds pecked the symbol corresponding to the number of elements it had pecked before, the symbol array extinguished and reward

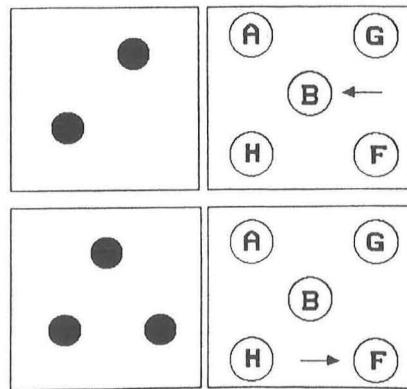


Figure 1. Examples of the numerosity and symbol stimuli used during training in Experiment 1. The arrows indicate the correct symbols for the numerosity stimuli illustrated on the left side.

Table 1
Conditions That Applied During the Four Training Stages of Experiment 1

Stage	Numerosities	Symbols	Chance level (%)	Criterion (%)	Trials/block	Blocks/session
1	1, 2	A, B	50	70	60	2
2	1, 2, 3	A, B, F	33	60	60	2
3	1, 2, 3, 4	A, B, F, G	25	52	80	2
4	1, 2, 3, 4, 5	A, B, F, G, H	20	50	100	2

was delivered. A feeding period of 3 s followed. If the pigeons pecked an incorrect symbol, they were penalized with 5 s of darkness. The subsequent trial was then a correction trial in which exactly the same numerosity configuration was presented again. Correction trials were repeated until the pigeon chose the correct symbol. These correction trials were ignored, however, when the trials were counted and discrimination accuracy was scored.

Stage 1 of training involved only the numerosity stimuli with one and two elements and the corresponding symbols A and B (Table 1). The next stage began for each animal after it responded to each numerosity with 70% or more correct trials over two successive blocks. The numerosity stimuli 1 to 3 and the corresponding symbols A, B, and C were used in Stage 2, and a criterion of 60% correct trials was set. Stages 3 and 4 followed and were run according to the entries shown in Table 1. During each stage, the various numerosity stimuli (with each numerosity in 10 different display configurations) were presented with equal frequency. Each numerosity, in all its configurations, was shown three times per block in Stage 1 and twice per block in Stages 2 to 4. Numerosities occurred in a quasirandom sequence across trials, and the same stimulus configuration was shown no more than twice in a row. The accuracy criteria (Table 1) were all chosen to guarantee a performance amply better than $p < .01$ above chance level according to the binomial distribution.

Results and Discussion

Figure 2 shows the cumulated number of trials that pigeons needed to reach the criteria associated with the various training stages. Table 2 summarizes the average number of trials to criterion for each stage as well as the percentage of correct responses for the various numerosities for the last four criterion blocks. Five pigeons met the training criterion of Stage 3 and so learned to associate the four numerosities 1, 2, 3, and 4 with the corresponding four symbols A, B, F, and G. All of the 6th pigeon's data were

excluded because this bird repeatedly failed to complete sessions within a time limit of 2 hours. Two birds out of the remaining 5 reached the criterion corresponding to Stage 4, where each of five numerosities had to be linked with one of five symbols. The other 3 pigeons were late in completing Stage 3 and so did not participate in Stage 4. Stage 4 took place after the 2 pigeons had completed most of the tests of Experiment 2.

Figure 3 shows the mean distribution of responses to the symbols A, B, F, and G as a function of the numerosity stimulus that had been presented. Means were based on data from the last four criterion blocks of Stage 3, averaged over all 5 pigeons. Such an analysis was not meaningful for Stage 4 because with only 2 birds participating in this stage, insufficient data were available. The analysis shows that when the birds erred they tended to choose the symbols that neighbored on the numerosity value of the correct symbol rather than on the spatial location of that symbol. If the birds had based their choices on spatial proximity to the correct symbol (see Figure 1) then on G(= 4) correct trials, for example, they should have distributed their choices in accordance with $G(= 4) > B(= 2) > F(= 3) \approx A(= 1)$, whereas they actually distributed them in accordance with $G(= 4) > F(= 3) > B(= 2) > A(= 1)$.

Experiment 2

This experiment consisted of a number of tests designed to examine the role of several cues that could have influenced the performance reported in connection with Experiment 1. Except for the final one, these tests immediately followed the completion of Training Stage 3 of the previous experiment.

Method

Subjects and apparatus. The same pigeons and the same chamber were used as before.

Procedure. Six tests of varying designs were run. Table 3 provides an overview; the details are explained in the following sections.

No tone. During Experiment 1, a peck on a given element of the numerosity stimulus had yielded a feedback tone in addition to the removal of a black dot from the center of the element. A test session was conducted in which the tone was omitted. This test assessed whether the pigeons were relying on counting this acoustic signal. The 5 pigeons that had completed Training Stage 3 were each exposed to a single test session consisting of 80 trials. These trials involved the presentation of the numerosity stimuli 1, 2, 3, or 4 in a quasirandom order (20 trials per stimulus) and the display of the array containing the symbols A, B, F, and G. No acoustic feedback for pecks on the numerosity stimulus elements (red circles) was given. Otherwise all conditions, including extinguishing the central black dot, were the same as in Training Stage 3.

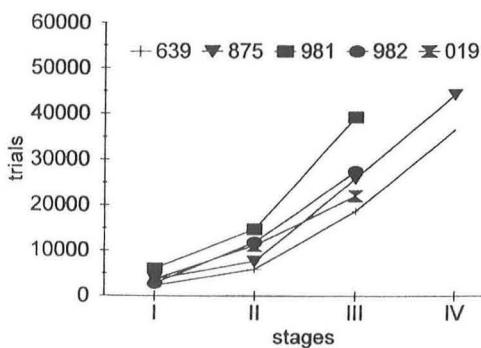


Figure 2. Experiment 1: Cumulated number of trials that the various pigeons (see legend) needed to reach the criteria associated with the four training stages.

Table 2
Average Percentage of Correct Responses During the Final Two Sessions of Each of the Training Stages of Experiment 1

Stage	<i>n</i>	Trials to criterion (<i>M</i>)	Numerosities/% correct trials				
			1 = A	2 = B	3 = F	4 = G	5 = H
1	5	3,720	74.8	75.4			
2	5	6,510	86.4	63.2	71.0		
3	5	16,260	72.0	60.0	59.0	62.0	
4	2	18,000	68.0	64.0	61.0	56.0	70.0

Note. Numbers in Columns 4–8 represent percentages. *n* = number of pigeons that reached criterion.

New configurations. So far, each of the numerosity stimuli had been presented in 10 different spatial configurations. Because the birds had extensive experience with these configurations and pigeons are known to have a capacious memory for visual patterns (Fersen & Delius, 1989), it was possible that these birds had learned to discriminate the configurations rather than their numerosities. A test with new spatial configurations was thus performed.

Before the test proper started, the pigeons were exposed to between four and six sessions that were run exactly as those of Training Stage 3 with the following exception: To familiarize the birds to a modified procedure, we did not reinforce a randomly chosen fifth of the trials. Instead, both correct and incorrect symbol choices led directly to the next trial. During these trials, acoustic feedback was given, but there were no correction trials. Each of the four proper test sessions that followed consisted of 100 trials. Of these, 80 followed the same procedure as in Training Stage 3 and involved the familiar numerosity stimulus configurations. The other 20 trials that were randomly inserted, not reinforced, and not corrected involved five new spatial configurations for each numerosity.

Novel elements. Although the pigeons were able to estimate the numerosity of the relevant stimuli, they might have been able to do this only when the stimuli involved the standard elements used during training. New numerosity elements instead of the standard circles were introduced to test whether this applied.

The test consisted of three series of four test sessions each. Each session consisted of 80 training trials and 20 unreinforced test trials as described above, but they included the standard and novel elements, respectively. The three series differed in that the first involved squares, the second triangles, and the third butterflies as novel numerosity stimulus elements. The elements are illustrated in Figure 4 and were shown in red color. They were selected to be of various areas and perimeters (0.7-cm diameter circles:

area = 290 pixels, perimeter = 50 pixels; squares: area = 289 pixels, perimeter = 64 pixels; triangles: area = 252 pixels, perimeter = 60 pixels; butterfly: area = 417 pixels, perimeter = 91 pixels; and cross [used in next test]: area = 370 pixels, perimeter = 88 pixels). Similar to the training circles, these test elements included a central black dot until the pigeons had pecked at them once. They were presented in five of the familiar configurations per numerosity. The same configurations were used for all three types of novel elements.

Heterogeneous stimuli. This test was designed to clarify why the birds' performance deteriorated when novel elements formed the numerosity stimuli. It consisted of three successive phases, Test 1, training, and Test 2. The first phase served to establish a baseline performance unaffected by any training and consisted of four sessions. There were 80 training trials and 16 unreinforced test trials per session. In the test trials, 16 novel numerosity stimuli made up of heterogeneous elements were presented (see bottom of Figure 4). For each of the numerosities 1 to 4, four different stimuli were constructed from the elements circle, square, triangle, cross, and butterfly.

In the second training phase, trials were similar to those of Stage 3 of training except that the homogeneous numerosity stimuli were composed of either the squares, triangles, crosses, or butterflies (see top of Figure 4). In each session, there were 20 trials with each element type in random order. This training lasted until the birds reached a criterion of 50% trials correct in two successive sessions on each numerosity (a significance level of $p < .01$ on binomial tests). The third phase consisted of four retest sessions that were identical to those of the first phase, including the use of the same heterogeneous test stimuli.

Abbreviated presentation. In all the previous phases, the pigeons could have counted the elements of the numerosity stimuli or the pecks they made to these elements. This test was intended to eliminate the latter alternative. It consisted of four sessions. Each session incorporated 80 training trials and 20 test trials. The training trials were conducted like those of Stage 3 training sessions in Experiment 1. Per numerosity, 10 different configurations of circle element stimuli were presented in the test trials of all four sessions. During these trials, a single peck at any element of the numerosity stimulus brought on the premature removal of the numerosity stimulus and the immediate presentation of the A, B, F, and G

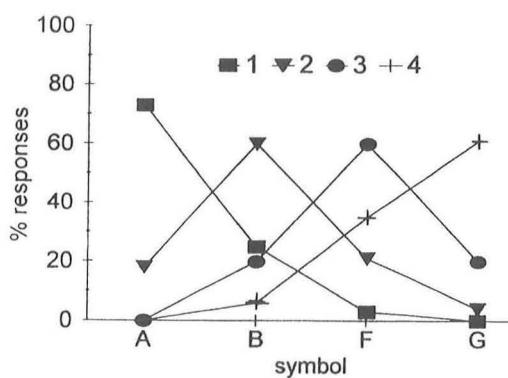


Figure 3. Experiment 1: Mean percentage of distributions of the responses to the numerical symbols A, B, F, and G as a function of the numerosity stimulus previously presented.

Table 3
Experiment 2: Summary of Tests

Test	Description	Sessions	Trials/session
1	No tone	1	80
2	New configurations	4	100
3	Novel elements	12	100
4	Heterogeneous sets	4	96
5	Abbreviated presentation	4	100
6	Symbol locations	20	48

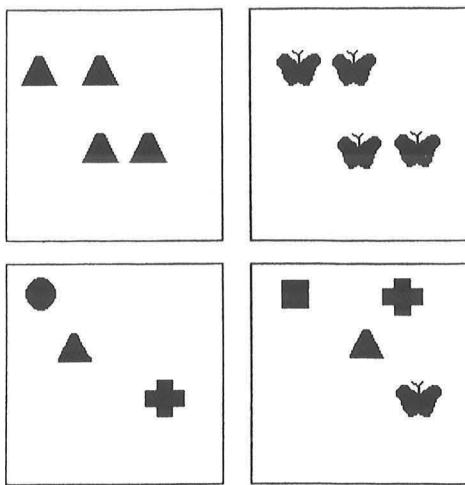


Figure 4. Examples of new homogeneous and heterogeneous configurations used during Experiment 2.

symbols. Choices of the symbols were neither rewarded nor penalized but led directly to the next trial.

Symbol locations. Previously, the locations of the number symbols within the stimulus field were always constant. To substantiate the analysis described in connection with Figure 3, we used this test to examine whether the birds had learned to associate a given numerosity with a specific symbol rather than with its standard location (cf. Figure 1). So the test checked whether the birds could still correctly choose the symbol that matched a given numerosity stimulus when a correct and an incorrect symbol were offered in varying locations. The test was carried out about 1 year after the abbreviated presentation test was completed. Only the 4 pigeons that were the first to complete that test participated. In the meantime, these 4 birds had taken part in a response production experiment described in detail by Xia et al. (2000). It involved their learning to peck a precise number of times, between one and four, as indicated by the particular symbol presented to them on a given trial. After that, the pigeons had spent about 9 months in an outside aviary before they participated in the present test.

The procedure resembled that used in Stage 2 of training, except that only six different configurations for each of the three numerosities 1, 2, and 3 were used. Also, only two of the corresponding numerosity symbols (A, B, or F) were presented simultaneously within each trial. The correct one always corresponded to the numerosity of the stimulus presented immediately before. The other, incorrect symbol corresponded to either one of the two numerosities not presented in the particular trial. The positions of the two symbols varied randomly among the nine possible fixed locations within the stimulus field. During the first five sessions, choice of the correct symbol led to food reward followed by a trial extension before the next trial began. Choice of the incorrect symbol led to timeout followed by a correction procedure in which the trial was restarted with the same stimuli and locations as before. In the subsequent 15 sessions, rewards were not immediately given for correctly choosing the symbol but rather were given when the trial extension was completed. Incorrect symbol choices had the same consequence as before. The trial terminated at this point with a timeout followed by a correction procedure. The trial extensions were unrelated to the test and were really intended to train the birds on a superimposed task related to that described by Xia et al. (2000). These extensions involved showing the correct symbol again between one and four times in succession, at randomly varying locations, but with the symbol color and the shape of its surround altered. The aim was to train the birds to peck the symbol for as many presentations as the

symbol required and then to complete the trial by pecking at an adjacent blue rectangle. Because the pigeons were largely unsuccessful at this secondary task, we do not report any results later. But note that this additional procedure is unlikely to have interfered with the initial part of each trial in which the symbol locations were varied when the birds had to choose between symbols.

Results and Discussion

No tone. The average percentages of correct responses for the separate numerosities 1, 2, 3, and 4 were 78%, 59%, 57%, and 57% respectively. All were significantly above the 25% chance level: numerosity 1, $t(4) = 8.77$; 2, $t(4) = 9.25$; 3, $t(4) = 9.44$; 4, $t(4) = 4.82$; $p < .05$ (Figure 5). Compared with the final session of Training Stage 3, there were no significant differences in performance (t tests; $p > .05$). The birds' performance with the numerosity stimuli did not appear to depend on assessing the number of feedback tones. Instead, they must have relied on either the visual cues of the numerosity stimuli or on the number of their own pecking actions.

New configurations. The average percentages of correct trials for the test trials with numerosities 1, 2, 3, and 4 were 66%, 63%, 55%, and 67%, respectively. These results were all significantly above the chance level: numerosity 1, $t(4) = 14.06$; 2, $t(4) = 8.49$; 3, $t(4) = 6.00$; 4, $t(4) = 5.59$; $p < .01$. Furthermore, there were no significant differences between these scores and the corresponding scores from training trials in the same sessions, except for the numerosity 1, $t(4) = 3.20$, $p < .05$ (Figure 6). Thus, the pigeons' performance did not rely on rote learning of the many numerosity stimulus configurations used for training. Either array configuration is unimportant to birds when they discriminate the number of elements, or the birds' having to peck each element of the numerosity stimuli may have prevented them from perceiving and encoding the stimuli as configurations. In the latter case, pecking each element might have encouraged them to attend more to the individual elements themselves.

Novel elements. The average performance for test trials during each of the three phases with novel types of elements is listed in Table 4 (see also Figure 7). The pigeons showed excellent transfer in the test trials involving numerosity stimuli composed of squares instead of circles. At all numerosities, the percentage of correct trials was well above chance level. However, with triangles and

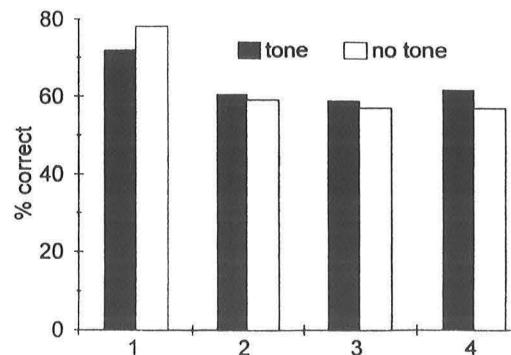


Figure 5. Experiment 2: Comparison of performance plotted as a function of the four relevant numerosities during the final Stage 3 of training and the test with no feedback tone.

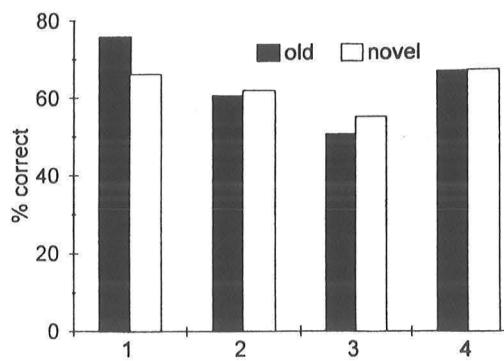


Figure 6. Experiment 2: Performance at each of four numerosities in the test with new configurations. Scores are from training trials with familiar stimulus configurations (old) and test trials with new ones (novel).

butterflies, the test performance was poor at the lower numerosities and only reached significance at the higher ones.

Previously, we argued that the pigeons probably did not learn the stimulus configurations but rather that they paid more attention to the component elements. Now it appears that at lower but unaccountably not at higher numerosities, they attended to the specific form of the elements and recognized that the triangles and butterflies were novel. Initially, the test elements were completely novel, and they were only presented 20 times during the tests. Experience in our laboratory has frequently indicated that stimulus novelty per se tends to disrupt pigeons' performance in categorization tasks. The better overall transfer to the square elements is likely to be due to a process of stimulus generalization from circles to squares.

A stimulus property that may have contributed to the generalization between circles and squares was their areas, which were very similar. However, when pigeons are trained to make relative quantity judgements, stimulus area acts only as a secondary cue for them (Emmerton, in press). Human infants who were tested in a counting task relied on the summed lengths of contours (or perimeters) in the elements of numerosity arrays (Clearfield & Mix, 1999). The perimeter lengths in the present test were most similar for squares and triangles, rather than squares and circles. Because performance levels with arrays of squares and triangles were dissimilar, it is unlikely that the birds relied on stimulus contour as a cue. The variability in their test results across numerosities (see Figure 7) suggests that they were not simply discriminating either

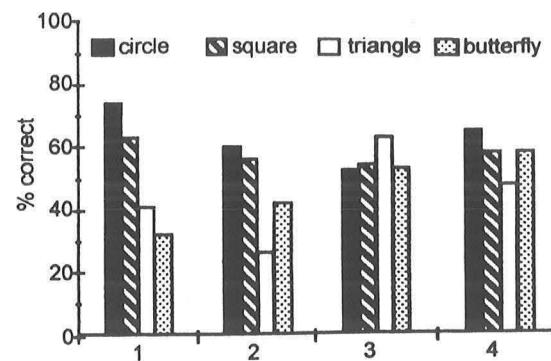


Figure 7. Experiment 2: Performance at each of four numerosities during the novel elements test. The circle was the familiar element previously used in training.

stimulus area or contour alone. However, pigeons do often utilize multiple cues in visual discrimination tasks (Chatlosh & Wasserman, 1993) so that sudden changes in stimulus properties such as area or shape may have affected stimulus generalization in this test.

Heterogeneous stimuli. The results of the three stages of this test are summarized in Figure 8. During the Test 1 phase, the pigeons responded on average with 47%, 32%, 40%, and 44% correct to the test numerosity stimuli 1, 2, 3, and 4, respectively. During the subsequent training phase they required 3 to 12 training sessions (an average of 6.8 sessions) to reach the criterion of 50% correct trials at each numerosity. Their final performance during the criterion training sessions was 68%, 64%, 52%, and 70% for the numerosities 1, 2, 3, and 4, respectively. Against a chance level of 25%, *t* test results were as follows: numerosity 1, *t*(4) = 10.59; 2, *t*(4) = 13.38; 3, *t*(4) = 7.22; 4, *t*(4) = 16.43; *ps* < .01 (Figure 8).

Retesting with the heterogeneous stimuli in Test 2 after the intervening training revealed that all the birds now responded well above chance level (*t* tests; *ps* < .05). There were no significant differences between the scores obtained during training with the homogeneous stimuli and the second test with the heterogeneous stimuli (*t* tests; *ps* > .05). Thus, even if a factor such as stimulus area had affected the birds' performance in the previous test, the

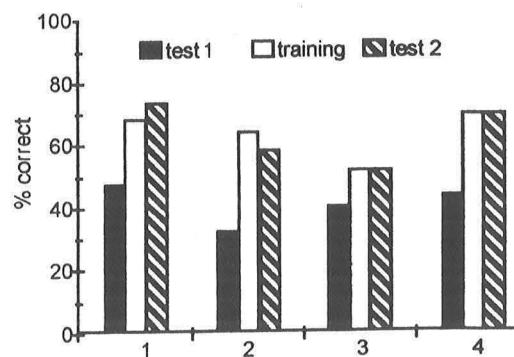


Figure 8. Experiment 2: Performance at each of four numerosities in the three phases of the heterogeneous stimuli test.

Table 4
Experiment 2: Average Percentage of Correct Scores During Test Trials With Different Novel Element Forms

Element type	Numerosities/correct trials (%)			
	1	2	3	4
Squares	63**	56*	54**	58**
Triangles	41	26	63*	48*
Butterflies	32	42	53*	58*

p* < .01. *p* < .001, *t* tests with significance levels of above-chance performance.

pigeons did not rely solely on that type of cue. The numerosity of the array seemed to be more important. This improved performance with the heterogeneous stimuli no doubt reflected the birds' increased familiarity with the novel elements through their extended training with the homogeneous stimuli. Note, however, that the birds never received reinforcement feedback with the heterogeneous stimuli because these were only presented in unreinforced test trials.

Abbreviated presentation. Figure 9 presents the results of this test. For the test trials, the percentages of correct scores were 68%, 55%, 17%, and 0% for the numerosities 1, 2, 3, and 4, respectively. The decline in accuracy to below chance level (i.e., < 25% correct symbol choices) with the two higher numerosities under the present conditions suggests that the pigeons were not relying on a holistic perception of the numerosity stimuli as arrays. Instead, they were attending to and pecking at the individual elements composing them. On the other hand, the fact that the pigeons were still responding correctly on a proportion of the 1, 2, and 3 numerosities indicates that they were not simply relying on a time-elapsed principle (i.e., how long they had spent pecking the numerosity stimulus) nor on a pecking-rhythm principle to choose the number symbol. If they had done so, they should have produced almost exclusively A-symbol choices to correspond with the pecking response they had learned to make for a numerosity of 1. Although actual response latencies were not recorded, observations made during the subsequent symbol-location test of the birds' pecking at the arrays suggested that the time taken to complete their responses was quite variable. When the same array was presented repeatedly on correction trials, for example, the bird often pecked at the circles in a different order and did not always choose the nearest one when shifting its responding from one circle to another (see also Xia et al., 2000).

Other experiments with pigeons (Emmerton, 1998) suggest that they may visually scan arrays. In those experiments, the arrays were presented in pairs on standard pecking keys and consisted of much smaller circles (2.0–3.5 mm diameter) than were used here. Also, the birds only had to emit a single peck to the correct (smaller numerosity) array. The inference that the birds scanned the arrays was based on the pattern of errors they made when the

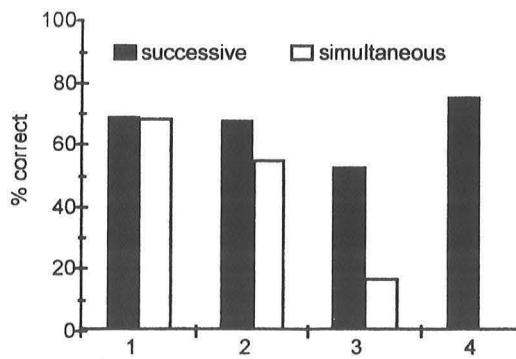


Figure 9. Experiment 2: Performance at each of four numerosities during the abbreviated presentation test. The legends *successive* and *simultaneous* refer to the birds' having to peck the numerosity elements successively or having to estimate numerosity during the elements' simultaneous, abbreviated presentations.

spacing of array elements was manipulated. The implication is that even if items in arrays are presented simultaneously, birds actually attend to the items sequentially, whether they have to peck each item or not. Even though the individual items in the arrays used here were quite large, they were also spaced out. Abbreviating the presentation time by having the first peck extinguish the array probably prevented the birds from properly scanning all items in the array. This can explain why performance accuracy decreased as a function of the number of items in a given test array. More generally, then, the question is whether birds ever process simultaneously presented arrays in a simultaneous, parallel-processing fashion or whether they effectively process the individual items sequentially. It remains to be seen whether timing plays a part in such sequential processing or whether animals can tag items one after the other and encode their number without using timing procedures. So far, no test has been proposed that can decide whether a pacemaker-accumulator model (a timing and counting model; Meck & Church, 1983; Roberts & Mitchell, 1994) or a neuronal filtering model (a counting only model; Dehaene & Changeux, 1993) is the one that most adequately accounts for animal counting performances.

Symbol locations. During the initial five sessions the 4 birds achieved means on noncorrection trials of 66%, 61%, and 54% correct choices for symbols A(= 1), B(= 2), and F(= 3), respectively. Although the means for symbols A and B were significantly above the chance level of 50% correct ($p < .01$ on a binomial test), accuracy was still relatively poor, and performance for the symbol F did not reach significance ($p > .05$, binomial test). However, these discrimination scores were obtained after a long experimental interruption and with a modified experimental procedure. So the birds were given additional training sessions. Mean choice accuracy scores calculated from noncorrection trials for the last five of these sessions were 82% (2.0), 76% (2.6), and 73% (2.6) for correct choices of symbols A, B, and F, respectively. They were all significantly above the chance level of 50% correct ($ps < .01$ on binomial tests). These scores compare well with those obtained during Stages 1 and 2 of Experiment 1, but they involved choices among 2 and then 3 symbols (Table 2). Although the symbols had always been presented in fixed positions in Experiment 1 and prior stages of Experiment 2, the pigeons apparently had learned to associate a symbol with the various 1 to 3 numerosity displays. In earlier stages of the experiments, the symbols and their locations provided redundant information that could have been linked with numerosity values. The birds required several thousand trials to achieve the acquisition criteria in Experiment 1 (see Table 2) but had regained their previous performance levels after at most 700 trials (15 sessions) in this test. It thus seems unlikely that their primary associations before the present test were between the array numerosities and the symbols' locations. Note also that the error analyses reported previously in connection with Experiment 1, Figure 3, suggested that the pigeons had not been primarily relying on the locations when choosing among the symbol stimuli.

General Discussion

Taken together, the results of Experiments 1 and 2 show that pigeons are indeed able to associate small numerosities with arbitrary numerical symbols. They do this in spite of variations in the characteristics of the numerosity stimuli, such as the shape, size,

and configuration of the elements, and in the location of the symbols representing numerosity. As we outlined above, many of the studies so far have examined birds' ability to make relative judgments about differences in numerosity (but see, e. g., Koehler, 1941; Pepperberg, 1994). The present results show that pigeons can also discriminate small sets of elements in terms of the absolute number of items these sets contain.

The ability to discriminate the absolute number of items in a set is only one of the important attributes of counting (Stevens, 1951). However, before an animal (or preverbal child) is considered to be capable of counting, other criteria have to be met. First, Gelman and Gallistel (1978) elaborated the principles of counting that were introduced by Stevens. Then the criteria for judging whether animals show evidence of counting on the basis of these principles were discussed in detail by Davis and Memmott (1982) and Davis and Pérusse (1988). According to these authors, relatively few animals to date have met these criteria (but see also Capaldi & Miller, 1988).

The results showing that pigeons can discriminate absolute numbers, at least within the small number range, accord with the cardinal principle of counting. The use of large, clearly separated visual elements together with the requirement that the birds peck directly at each element in the stimulus display ensured that they processed each item before they chose a symbol. Across varying trials, their accuracy in choosing the symbol that designated a particular quantity after they had physically tagged each item in a given set concurred with the one-to-one principle. What is lacking is a way of ascertaining whether the animals utilize a progression of "mental tags" (Thomas & Lorden, 1993) each time they physically respond to an additional item or event in a set consisting of more than one element (but see Orlov, Yakovlev, Hochstein, & Zohary, 2000).

Even for the same numerosity, the configuration of the stimulus display of simultaneously presented elements varied across trials. The pigeons maintained their accurate performance level in choosing the assigned symbol when they were tested with novel configurations for each numerosity. These novel configurations, with elements presented at locations that differed at least partially from those of items in familiar training configurations, would have interfered with any potential tendency of birds to respond to elements in a fixed order. Informal observations suggested that there was in fact no such tendency. In the event of an incorrect symbol choice, the same configuration was shown again in the next correction trial. Rather than pecking stimulus items on these correction trials in exactly the same order as they had in the preceding trial, the birds often varied the order of locations to which they responded. This is what would be expected according to the order-irrelevance principle.

The abstraction principle of counting requires that an animal can accurately discriminate the absolute number of items in a set for any type of items that may constitute the set. With some novel test elements (squares instead of the circles in Experiment 2), the birds showed immediate transfer to the new items for all four numerosities. With items that were more dissimilar to the training circles (triangles and butterflies), a high level of transfer occurred only with novel stimulus displays that contained several items. The birds' poorer transfer with just one or two of these dissimilar items was attributed to a transient disruptive effect of novel stimulus features per se. After receiving additional exposure to the novel

elements, the birds demonstrated their ability to numerically discriminate a variety of different items in heterogeneous sets, as the abstraction principle requires. It would have been interesting to investigate the pigeons' transfer performance with elements they already knew from some other nonnumerical learning context. However, the birds used here were experimentally naive at the beginning of the study. Perhaps such prior familiarity with even highly dissimilar elements would have led to immediate transfer of discrimination across all the numerosities.

In one respect, however, the results could not provide sufficient evidence to satisfy one of the counting principles. Counting includes not just the ability to label cardinal numbers of items in a set but also assumes ordinality, or the idea that numerical symbols are themselves perceived as being ordered along a scale representing small to large numerical amounts. To demonstrate ordinality, or the stable-order principle, there should ideally be some evidence that the birds judged numerosities to fall in the order $1 < 2 < 3 < 4$, and correspondingly, $A < B < F < G$. Even though the patterning of the error gradients suggests an orderly numerosity and symbol dimension, specific tests beyond those performed in the current experiments would be required to provide evidence of whether pigeons have a sense of ordinality with respect to numerical symbols. Tests with visual arrays as stimuli indicate that pigeons serially order these arrays according to their numerosity values (Emmerton et al., 1997). Specific tests of ordinality with rhesus monkeys also show that primates respond to arrays as if they are ordered along a numerosity scale (Brannon & Terrace, 2000). So far, no animal has been tested for ordinality with numerical symbols.

The training environment in the present experiments and the task of associating arbitrary symbols with computer-generated visual elements did not resemble a pigeon's normal ecological environment. This may be one reason why the birds required a large amount of training to acquire the symbolic association. However, extensive training is also generally needed before primates (Boysen, 1993), parrots (Pepperberg, 1994), or indeed young children (Gelman & Gallistel, 1978) are competent at using symbolic labels to match numerosities. The way in which pigeons (as well as other animals) may use their numerical abilities remains speculative. One possibility is that the ability to assess the number of visual items presented simultaneously or encountered in succession could contribute to foraging success. It is less obvious how they might apply an ability to associate a symbol with an absolute number of items, although possibly they could learn that some specific features of their environment are reliably correlated with amounts of food to be obtained in that part of their habitat. But perhaps the way to think about the issue is that animals like the pigeon may be able to learn to associate external numerosities like flock sizes with representational symbols with a good resolution at low numbers but with increasing fuzziness at higher numbers. Perhaps we only taught our pigeons to associate cognitive codes for perceived numerosities with the arbitrary external symbols we presented them.

Regardless of all these considerations, the present demonstration that pigeons can associate numerical symbols with numerosity sets shows that their numerical competence extends beyond the phylogenetic level that previous studies had hitherto revealed.

References

- Alsop, B., & Honig, W. K. (1991). Sequential stimuli and relative numerosity discriminations in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 17, 386–395.
- 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.
- Brannon, E. M., & Terrace, H. S. (2000). Representation of the numerosities 1–9 by rhesus macaques (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*, 26, 31–49.
- Butterworth, B. (1999). *The mathematical brain*. London: Macmillan.
- Capaldi, E. J., & Miller, D. J. (1988). Counting in rats: Its functional significance and the independent cognitive processes that constitute it. *Journal of Experimental Psychology: Animal Behavior Processes*, 14, 3–17.
- Chatlosh, D. L., & Wasserman, E. A. (1993). Multidimensional stimulus control in pigeons: Selective attention and other issues. In T. R. Zentall (Ed.), *Animal cognition: A tribute to Donald A. Riley* (pp. 271–292). Hillsdale, NJ: Erlbaum.
- Clearfield, M. W., & Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science*, 10, 408–411.
- Davis, H., & Albert, M. (1986). Numerical discrimination by rats using sequential auditory stimuli. *Animal Learning & Behavior*, 14, 57–59.
- Davis, H., & Memmott, J. (1982). Counting behavior in animals: A critical evaluation. *Psychological Bulletin*, 92, 547–571.
- Davis, H., & Pérusse, R. (1988). Numerical competence in animals: Definitional issues, current evidence, and a new research agenda. *Behavioral and Brain Sciences*, 11, 561–615.
- Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. Oxford, England: Oxford University Press.
- Dehaene, S., & Changeux, J.-P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience*, 5, 390–407.
- Emmerton, J. (1998). Numerosity differences and effects of stimulus density on pigeons' discrimination performance. *Animal Learning & Behavior*, 26, 243–256.
- Emmerton, J. (in press). Pigeons' discrimination of color proportions in computer-generated visual displays. *Animal Learning & Behavior*.
- Emmerton, J., Lohmann, A., & Niemann, J. (1997). Pigeons' serial ordering of numerosity with visual arrays. *Animal Learning & Behavior*, 25, 234–244.
- Fernandes, D. M., & Church, R. M. (1982). Discrimination of the number of sequential events by rats. *Animal Learning & Behavior*, 10, 171–176.
- Fersen, L. von, & Delius, J. D. (1989). Long term retention of many visual patterns by pigeons. *Ethology*, 82, 141–155.
- Fetterman, J. G. (1993). Numerosity discrimination: Both time and number matter. *Journal of Experimental Psychology: Animal Behavior Processes*, 19, 149–164.
- Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.
- Honig, W. K., & Stewart, K. E. (1989). Discrimination of relative numerosity by pigeons. *Animal Learning & Behavior*, 17, 134–146.
- Koehler, O. (1937). Können Tauben "zählen"? [Can pigeons "count"?]. *Zeitschrift für Tierpsychologie*, 1, 39–48.
- Koehler, O. (1941). Vom Erlernen unbenannter Anzahlen bei Vögeln [About the learning of unnamed item numbers by birds]. *Die Naturwissenschaften*, 29, 201–218.
- Koehler, O. (1949). "Zählende" Vögel und vorsprachliches Denken ["Counting" birds and preverbal thinking]. *Verhandlungen der Deutschen Zoologischen Gesellschaft*, 1949, 219–238.
- Koehler, O., Müller, O., & Wachholtz, R. (1935). Kann die Taube Anzahlen erfassen? [Can pigeons comprehend numbers of items?]. *Verhandlungen der Deutschen Zoologischen Gesellschaft*, 1935, 39–45.
- Matsuzawa, T. (1985). Use of numbers by a chimpanzee. *Nature*, 315, 57–59.
- 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.
- Olfhof, A., Iden, C. M., & Roberts, W. A. (1997). Judgments of ordinality and summation of number symbols by squirrel monkeys (*Saimiri sciureus*). *Journal of Experimental Psychology: Animal Behavior Processes*, 23, 325–339.
- Orlov, T., Yakovlev, V., Hochstein, S., & Zohary, E. (2000). Macaque monkeys categorize images by their ordinal number. *Nature*, 404, 77–80.
- Pepperberg, I. M. (1987). Evidence for conceptual quantitative abilities in the African Grey parrot: Labeling of cardinal sets. *Ethology*, 75, 37–61.
- Pepperberg, I. M. (1994). Numerical competence in an African Grey parrot (*Psittacus erithacus*). *Journal of Comparative Psychology*, 108, 36–44.
- Platt, J. R., & Johnson, D. M. (1971). Localization of position within a homogeneous behavior chain: Effects of error contingencies. *Learning and Motivation*, 2, 386–414.
- QuickBasic (Version 4.0) [Computer software]. (1987). Redmond, WA: Microsoft.
- 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.
- Roberts, W. A., Macuda, T., & Brodbeck, D. R. (1995). Memory for number of light flashes in the pigeon. *Animal Learning & Behavior*, 23, 182–188.
- Roberts, W. A., & Mitchell, S. (1994). Can a pigeon simultaneously process temporal and numerical information? *Journal of Experimental Psychology: Animal Behavior Processes*, 20, 66–78.
- Rumbaugh, D. M., & Washburn, D. A. (1993). Counting by chimpanzees and ordinality judgments by macaques in video-formatted tasks. In S. T. Boysen & E. J. Capaldi (Eds.), *The development of numerical competence: Animal and human models* (pp. 87–106). Hillsdale, NJ: Erlbaum.
- Smirnova, A. A., Lazareva, O. F., & Zorina, Z. A. (2000). Use of number by crows: Investigation by matching and oddity learning. *Journal of the Experimental Analysis of Behavior*, 73, 163–176.
- Stevens, S. S. (1951). Mathematics, measurement, and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1–49). New York: Wiley.
- Thomas, R. K., Fowlkes, D., & Vickery, J. D. (1980). Conceptual numerosity judgments by squirrel monkeys. *American Journal of Psychology*, 93, 247–257.
- Thomas, R. K., & Lorden, R. B. (1993). Numerical competence in animals: A conservative view. In S. T. Boysen & E. J. Capaldi (Eds.), *The development of numerical competence: Animal and human models* (pp. 127–147). Hillsdale, NJ: Erlbaum.
- Washburn, D. A., & Rumbaugh, D. M. (1991). Ordinal judgments of numerical symbols by macaques (*Macaca mulatta*). *Psychological Science*, 2, 190–193.
- Xia, L., Siemann, M., & Delius, J. D. (2000). Matching of numerical symbols with number of responses by pigeons. *Animal Cognition*, 3, 35–43.
- Zeier, H. (1966). Über sequentielles Lernen bei Tauben, mit spezieller Berücksichtigung des "Zähl"-Verhaltens [About sequential learning in pigeons, especially regarding counting behavior]. *Zeitschrift für Tierpsychologie*, 23, 161–189.