

## Pigeons' serial ordering of numerosity with visual arrays

JACKY EMMERTON, ANNETTE LOHMANN, and JOACHIM NIEMANN  
*Ruhr-Universität Bochum, Bochum, Germany*

Pigeons were trained in a conditional discrimination paradigm to differentiate successively presented visual arrays according to the relative number of their elements. Transfer tests with novel stimuli demonstrated that they discriminated the categories of "many" (6 or 7) from "few" (1 or 2) items. In further tests, other new stimuli were introduced that consisted not only of these training numerosities, but also of the intervening ones (3, 4, and 5). Variations in the birds' discrimination performance corresponded to the order of stimuli on a numerosity dimension. This serial ordering was maintained when other factors such as brightness, size, shape, area, and contour of the elements were systematically controlled across tests. Smaller numerosities were somewhat better discriminated than those at the higher end of this test range.

Since the creative approach developed by Koehler (e.g., 1937, 1943) for the investigation of infrahuman numerical abilities, evidence has accumulated over recent years that a variety of animals are capable of differentiating stimuli on the basis of their number. Much of this work has been reviewed by Davis and Memmott (1982), Davis and Pérusse (1988), and in Boysen and Capaldi's (1993) book. Some researchers have argued that a sense of number is one of the fundamental faculties in animal cognition, and that there are basic and elemental mechanisms for the internal representation of numerosity (Gallistel, 1993). From this perspective, numerical discrimination, even counting, is an ability that some animals apply routinely and perhaps automatically in their behavioral repertoires (Capaldi & Miller, 1988). Others (Davis, 1993; Davis & Memmott, 1982; Davis & Pérusse, 1988), while acknowledging that a range of abilities they term "numerical competence" has been demonstrated experimentally, nevertheless hold the view that animals are not naturally attuned to number, and that the discrimination of absolute number in particular is a strategy of last resort. The extent to which animals utilize number in their natural environments remains speculative. However, it is not difficult to suggest ways in which a sense of number could provide ecologically relevant information. In foraging, for instance, an ability to estimate number would enable birds to assess the rate of return of serially gathered food items or the relative profitability of different seed patches.

These examples of potential applications of numerical abilities in fact parallel two of the major experimental approaches to studying animals' discrimination of number. In many investigations, chiefly with rats or pigeons, the animals have been trained to differentiate the numerical dimension of series of events, which have been either the animals' own responses (e.g., the number of barpresses or keypecks emitted; Fetterman, 1993; Mechner, 1958) or external stimuli (e.g., a sequence of tone pulses or light flashes; Alsop & Honig, 1991; Fernandes & Church, 1982; Roberts, Macuda, & Brodbeck, 1995). As Gallistel and Gelman (1992) have pointed out, to demonstrate that animals have number categories it must be shown that they base their behavior on the numerosity of a set of items independently of other attributes of that set. A potential confound with numerosity in sequences of events is the total duration of the sequence or of the stimulus components that comprise it. Recent research has examined the effects of both number and time on animals' behavior. From such studies, models have been developed that show how time and number, while they can be evaluated separately, may be processed by some common underlying mechanisms (Fetterman, 1993; Meck & Church, 1983; Roberts et al., 1995; Roberts & Mitchell, 1994).

The alternative approach, which we adopt here, has been to present multi-item visual displays in which the number of items varies from one array to another. The displays we used were intended to be quasi-ecological, with the arrays being composed of random arrangements of seed-like dots or small squares. With this type of stimulus, numerosity can potentially be confounded with a number of other parameters, for instance the overall brightness or area of the display, or the size or density of the constituent elements. One way of controlling for these confounds is that used by Honig, who has demonstrated that birds can discriminate the relative numerosity, or proportion, of a subset of elements in matrices consisting of a fixed total number of items (Honig, 1991; Honig &

---

This research was supported by a grant from the Deutsche Forschungsgemeinschaft ("Kognitaves" Em 41/1-1). Portions of the research were presented at the 1988 meeting of the Psychonomic Society in Chicago. We would like to thank Professor Juan D. Delius for numerous discussions during the course of this study, and Herr G. Keim for his help with photographic preparations. Correspondence should be addressed to J. Emmerton, Purdue University, Department of Psychological Sciences, West Lafayette, IN 47907-1364 (e-mail: jemmert@psych.purdue.edu).

Stewart, 1989; Honig & Stewart, 1993). We have employed a different strategy, one of having pigeons discriminate visual arrays in which numerosity is defined by the total number of elements within a given array, but then systematically controlling for confounding parameters when new sets of stimuli are presented to the birds.

If an organism really discriminates the abstract property of numerosity, then the internal representation of numerical categories ought to map systematically onto some external scale of magnitude (see Gallistel, 1990). Thus there ought to be a functional correspondence between the number of items in a given display and how the brain encodes these quantities. Furthermore, if an animal is trained to discriminate different numerosities, and these map onto separate points along a number scale, then intervening numerosities, with which the animal has not been specifically trained, should be assigned "spontaneously" to interpolated points along a numerical continuum. The alternative is that an animal could utilize number in only a limited way, responding to different numerosities as discrete and disconnected features. In this case, its categorization of novel intervening numerosities should be essentially random.

Our aim in the present study was to provide further evidence that birds can discriminate visual stimuli on the basis of their numerosity, and that their performance does not depend on, nor is it significantly influenced by, a variety of other nonnumerical stimulus attributes. Furthermore, we were interested in determining whether they categorized novel stimuli systematically, with the allocation of behavioral choices corresponding to an increasing number of items in a given array. The numerosities we used fell within the range that pigeons had been able to discriminate when they were tested with similar stimuli in previous experiments (see Emmerton & Delius, 1993). The initial training condition required the birds to differentiate "few" from "many" items, rather than responding to only fixed numerosities. This training was meant to encourage some flexibility in the animals' allocation of choices about slightly varied numbers of items.

## METHOD

### Subjects

There were two groups of subjects. In the first part of the experiment, 8 experimentally naive homing pigeons (*Columba livia*), obtained from local breeders, were used. During the first author's leave of absence, these birds were erroneously sacrificed in another experiment. Therefore, the second part of the experiment was conducted with 10 new pigeons, similarly naive and of local stock. During the experiment, the animals were housed in individual cages in a colony room with a 14:10-h light:dark cycle. They were maintained at 80% of their ad-lib body weights, but water and grit were freely available to them in their home cages.

### Apparatus

The conditioning chamber (35 × 35 × 37 cm) was constructed mainly of aluminum and was of in-house design. On one panel there were three horizontally arranged response keys (2.5 cm diam), made of matte acrylic plastic. They were situated 21.5 cm above the grid floor of the chamber and were separated by 8.5 cm, center to center. A 1-W red light could be switched on behind the right key,

and a 1-W green light could be switched on behind the left key. Baffles around the middle key shielded it from scattered colored light. Stimuli were displayed on this center key by opening a solenoid-operated shutter, located behind the key. The stimuli were back-projected onto the key with a Zeiss Ikon Unimat 2500 slide projector. A collector lens inside this projector was darkened with diffusely sprayed black paint to reduce the light intensity.

The conditioning chamber was illuminated by a 1.7-W light, set 7.5 cm above the center response key. A second light, directly beside the houselight, was switched on during food-reinforcement periods. A small ledge below these lights shielded the response keys from scattered light. Access to a solenoid-operated food hopper was via an opening 12.5 cm below the center response key.

A Commodore VC 20 computer and an interface, manufactured in-house, were used to control events in the operant chamber, to drive the projector, and to register responses. Stimuli were prepared as photographic negatives. These were mounted over a 1-cm-diam central hole in each specially constructed 5 × 5 cm slide frame. Toward the lower edge of the frame were three coding holes, each 0.5 cm in diameter. The binary state of these coding holes, that is, each either open or closed, was signaled to the computer via three photocells built into the slide projector. The stimuli consisted of varying configurations of white elements (dots, circles, squares, or triangles) on a dark background. The sizes of these elements also varied. For instance, the diameters of the smallest and largest circles, when they were projected onto the key, varied between 0.8 and 6 mm. The type and size of elements that were used depended on the phase of the experiment. In the training phase, the elements within a stimulus were all of the same size and type (e.g., all squares) but differed across stimuli. Various parameters of the stimuli were modified in subsequent novelty-habituation and test phases and will be described in later sections.

### Procedure

Each group of birds underwent initial response training, after which there were three different procedural conditions in the experiment: a prolonged conditional discrimination-training phase, incorporating sessions in which transfer to new stimulus exemplars was tested, various novelty-habituation phases, and associated test phases. The novelty-habituation and test conditions were repeated each time that new numerosities and parametric changes in the stimuli were introduced.

**Response training.** The pigeons were trained to keypeck with a modified autoshaping procedure (Brown & Jenkins, 1968). Following a 20-sec intertrial interval, the two side keys were illuminated for 6 sec with white light. A peck to either key during this stimulus period produced immediate food reinforcement for 3 sec. If the bird failed to respond, however, reinforcement was presented concurrently with the last 3 sec of the stimulus interval. If an animal showed a strong side preference, the preferred key was disabled in the next session.

After six sessions of autoshaping, an instrumental procedure was introduced. For the birds in the first part of the experiment, the center key was illuminated with white light and a single peck to it switched on the side keys with the center key remaining lit up. A response to either side key switched off the illumination on all three keys and led to food reward. These birds were trained instrumentally for eight sessions. For the birds in the second part of the experiment, an identical instrumental procedure was employed for nine sessions. During the next two sessions, the response requirement on each key was raised to a fixed ratio of two (FR2). In a further five sessions, the middle key was illuminated with a diagonal cross-hatch pattern and the response requirement per key was gradually raised to FR6.

**Conditional discrimination training and transfer sessions.** The pigeons were then trained with a conditional method to discriminate stimuli containing "many" items from those containing "few." Stimuli of the "many" type were defined as those with 6 or 7 ele-

ments. Stimuli of the "few" type consisted of 1 or 2 elements. After a bird had responded to the stimulus on a fixed-ratio schedule (increased from FR1 to FR6 early in acquisition), the right-hand key was illuminated with red light and the left with green light, while the center-key stimulus remained on.

For half of the pigeons, responses to the right key were reinforced when the center-key stimulus was of the "many" type, while pecks to the left key were correct following responses to the "few" stimulus. For the other subjects, the choice contingencies were reversed. The side-key response requirement by the end of acquisition was raised from FR1 to a modified FR10 schedule, that is, 10 consecutive pecks to the correct side key constituted a correct choice and led to food reward. However, if the animal changed sides before completing 10 pecks, these incorrect responses were subtracted from the correct count and additional responses were needed to complete the FR requirement. (In practice, however, the birds rarely switched from one key to another once responding had been initiated.) Incorrect responses were simply accumulated until the FR contingency had been met (irrespective of any key switching). A total of 10 incorrect pecks counted as an incorrect choice and led to a 3-sec time-out period. The trial was then repeated in a correction procedure until it ended in reward. These correction trials were registered, but they were not used to score discrimination performance.

With the first group of birds, 12 "few" and 12 "many" stimuli were presented in quasi-random order (Fellows, 1967) within a session for the first 10 sessions of training. Then, as an intended aid to learning, the stimuli were presented for 10 sessions in alternating blocks of 6 different "few" and 6 different "many" exemplars. The order of these blocks was changed every 2 sessions. Subsequently, a quasi-random order of stimulus presentation was reintroduced and stimulus order was changed every 4 sessions.

After the birds were discriminating at over 75% correct choices with the original 24 stimuli, 24 new stimuli were added to this set in Session 37. Starting with Session 50, partial reinforcement was introduced by not reinforcing 3 "many" and 3 "few" trials, irrespective of side-key choices. In Session 60, 12 familiar stimuli were replaced by novel ones (6 "few" and 6 "many"). Choices on half of each type of new trial were reinforced; on the other novel trials, no reinforcement was given. In Sessions 68 and 69, the brightness of the stimuli was randomly varied by affixing neutral-density filters to 16 of the 48 stimuli. These filters attenuated the already variable stimulus luminance by 0.06 to 0.72 log units. In Sessions 70 and 74, half the stimuli were replaced by new ones, half of which were presented in reinforced trials and half in nonreinforced trials. In addition, the mean luminance over all the "many" and all the "few" stimuli was equated to a level of 0.2 log cd/m<sup>2</sup> (as measured by a Mavolux electronic luxmeter) by adding neutral-density filters to individual stimuli. Finally, in Session 84, all 48 stimuli were replaced by novel ones and 12 trials were nonreinforced. Throughout this extended training and transfer phase of the experiment, the quasi-random order of stimulus presentation was changed every 4 sessions.

The stimuli and the procedure were essentially the same for the second group of animals that were trained. For these birds, "few" and "many" stimuli were presented in blocks of six during the first 18 sessions, with the order of blocks being changed every 3 sessions. After that, the stimulus order was quasi-random. By Session 36, these birds also discriminated "few" from "many" stimuli with an accuracy exceeding 75% correct. Subsequently, similar transfer sessions with new stimuli, modifications in partial reinforcement, and alterations in stimulus brightness were introduced as they had been for the first group of pigeons.

**Novelty-habituation phases.** The most critical data in the experiment were to be collected in various test phases. The main purpose of these test phases was to examine the effect on discrimination performance of introducing new numerosities. At the same time, as a check that discrimination depended on numerosity rather than on some other stimulus dimension, various other stimulus pa-

rameters that had been partly confounded with numerosity during prior discrimination training were to be altered. To obtain sufficient test data, but to avoid additional training effects, choices on critical test trials were to be nonreinforced. However, since the introduction of novelty per se, or the consistent association of novelty with nonreward, might have disrupted discrimination, each test phase in which novel stimulus features were to be introduced was preceded by a "novelty-habituation" phase in order to preclude such effects.

A novelty-habituation phase lasted for four sessions. In each habituation phase, eight familiar stimuli were replaced by new ones that featured the particular parametric change that was to be included in the new test-phase slides. On four of these novelty-habituation trials, correct choices were rewarded and incorrect ones produced a time-out period. On the other four trials, no reinforcement was given, irrespective of the bird's choices. Thus, stimulus novelty did not consistently signal nonreinforcement. The habituation stimuli were always of the type "many" and "few," that is, they consisted of the familiar training numerosities 1 or 2 and 6 or 7.

**Test-phase procedure.** Each novelty-habituation phase was followed by a test phase that lasted for six sessions. The test stimuli comprised not only new exemplars of the numerosities 1 or 2 for "few" and 6 or 7 for "many," as in the training and habituation phases, but also the interpolated numerosities 3, 4, and 5. Choice responses on test trials were never reinforced, but these trials were intermixed within a session with reinforced trials in which stimuli from the habituation phase were used again. Since birds never received feedback about the correctness of their choices on test trials, the same stimuli were shown repeatedly in the six test sessions. Trial sequences were changed every two sessions.

In the first five test phases, there were three novel stimulus exemplars for each numerosity group (1 or 2; 3, 4, 5; 6 or 7), that is, there were 15 test stimuli altogether. In addition, 16 of the "few" stimuli (1 or 2 elements) and 16 of the "many" stimuli (6 or 7 elements) were retained from the habituation phase and comprised the reinforced trials. Thus there was a total of 47 trials in a test session. Of the 47 trials in the final three test phases, 21 were test trials (3 for each separate numerosity value of 1 to 7). The remaining 13 "few" and 13 "many" trials comprised the reinforced trials. Further details about the rationale for modifying various stimulus parameters will be given below when the results of each test are described.

## RESULTS

### Acquisition Phase With Transfer Sessions

This prolonged phase provided a reliable level of baseline performance before the critical test phases began. Over Sessions 32–36 (the last five sessions before transfer sessions began), the mean level of correct choices was 80% for the first group of birds and about 78% for the second group. Transfer to new stimuli was then assessed in individual sessions among the ongoing training sessions.

For the first group of birds, new stimuli were introduced on Sessions 37, 60, 70, 74, and 84. Correlated *t* tests that compared the percentage of correct scores for the familiar versus novel stimuli within each of the first four transfer sessions also showed that, with the exception of the first of these sessions in which new stimuli were more poorly discriminated than familiar ones, there were no reliable differences ( $p > .05$ ) between data obtained with old and new configurations (see Table 1). In the third of these transfer sessions, stimuli were also equated in luminance, but discrimination performance was not adversely affected by this control. On the final transfer test, the comparison was between the discrimination data from Ses-

**Table 1**  
Mean Scores and *t* Values From Transfer Sessions During Acquisition Phase for First and Second Groups of Pigeons

Session Number	% Correct Choices		<i>t</i>
	Familiar Stimuli	Novel Stimuli	
	First Group		
37	81.2	66.1	3.84‡
60*	80.9	76.0	0.84
70*†	87.5	85.4	0.54
74*	91.1	92.1	0.27
83 vs. 84*	86.5	84.1	1.26
	Second Group		
37	78.3	77.1	0.26
54*	83.6	75.1	2.07
66*	82.9	82.9	0.002
71*†	88.1	82.9	2.16
79 vs. 80*	89.4	87.9	0.71

Note—In *t* tests,  $df = 7$  for the first group,  $df = 9$  for the second group. \*Mixture of reinforced and nonreinforced novel trials. †Novel stimuli equated in luminance. ‡ $p < .01$ .

sions 83 (all familiar slides) and 84 (all novel slides). Again, the birds' choice scores did not differ.

For the second group of birds, transfer tests were performed on Sessions 37, 54, 66, 71, and 80. The results for these animals were similar, except that none of the paired-differences *t* tests revealed reliable differences in discrimination of slides that were familiar versus novel, or else novel and also equated in brightness.

By the end of the acquisition phase, the mean discrimination scores over the last five sessions were approximately 85% correct choices for Group 1 and 88% for Group 2.

### Novelty Habituation

In the novelty-habituation phases, performance was assessed on the very first session in which a stimulus change (e.g., in size, shape, brightness) was introduced. In each of these initial sessions, the percentage of correct choices to the eight novel stimuli (on both reinforced and nonreinforced trials) was calculated for each bird. The scores from these new trials were then compared with the scores from the trials with familiar training stimuli. Correlated *t* tests performed on these data (see Table 2) did not show any significant differences ( $p > .05$ ) in mean discrimination with old and new stimuli.

### Test Phases

The specific stimulus conditions, and rationale for using them, will now be described for each of a series of test phases, together with the results for each of these phases.

**Test 1: Same-size/same-form elements.** In the first test, the stimuli were similar to the type previously employed in the training phase, that is, the elements within a particular test stimulus were of the same form and size but varied between stimuli. The component elements are shown in the inset on Figure 1. They were squares or circles, filled in or outlined, and were in one of two different

sizes. The important feature of this test was that it was the first time in the experiment that the first group of birds encountered the numerosities 3, 4, and 5.

Test scores were expressed as the percentage of choices, for each numerosity category, made to the side key that was designated as correct for a "many" stimulus. Figure 1 shows the mean scores for the 8 birds over all six sessions on nonreinforced test trials only. This score was high for test stimuli consisting of 6 or 7 elements, as would be expected from the birds' training performance. The test score was low for stimuli containing 1 or 2 items, that is, the animals correctly chose the "few" key and seldom responded to the "many" key. The percentages of "many" choices for the new numerosities 3, 4, and 5 were distributed between these two extremes.

The results of a repeated measures two-way analysis of variance (ANOVA) revealed a significant main effect of numerosity category [ $F(4,28) = 66.12, p < .001$ ]. In this analysis and in ANOVAs for all subsequent test data, there was no significant main effect of test sessions ( $p \geq .2$ ); nor was the numerosity  $\times$  session interaction significant (also  $p \geq .2$ ). Post hoc comparisons of mean percentage choices were made with Newman-Keuls tests for the five numerosity categories. Discrimination scores did not differ significantly ( $p > .1$ ) for stimuli containing 5 elements compared with those containing 6 or 7. All other comparisons between pairs of numerosity categories were significant ( $p = .01$ ).

**Test 2: Variable-size elements.** Investigations of perceptual invariance with humans have shown that the time taken to recognize that one geometric form is the same as another one varies with differences in size between the two patterns (Bundesen & Larsen, 1975). A time-dependent process of "mental zooming," by which the two patterns are brought into size correspondence, has been postulated. The next test with pigeons was designed to check if the accuracy of numerosity judgments would be impaired by a potential increase in perceptual compu-

**Table 2**  
Mean Scores and *t* Values From the First Sessions of Novelty-Habituation Tests in Which Stimulus Parameters Were Changed

Novel Stimulus Type	% Correct Choices		<i>t</i>
	Familiar Stimuli	Novel Stimuli	
Variable-size elements	87.5	89.1	0.44*
Equal-luminance stimuli	88.5	86.3	0.56
Mixed outline/filled elements	90.0	93.8	1.09
Mixed shape elements	93.3	95.0	0.58
Equal stimulus contour, same-size elements	92.5	93.8	0.33
Equal stimulus contour, variable-size elements	89.0	92.5	0.94
Equal stimulus area, same-size elements	92.3	93.8	0.42
Equal stimulus area, variable-size elements	93.5	93.8	0.15

Note—"Familiar" stimuli had same-size, same-shape elements within a stimulus. \* $df = 7$ ; for remaining *t* tests,  $df = 9$ .  $p > .05$  on all tests.

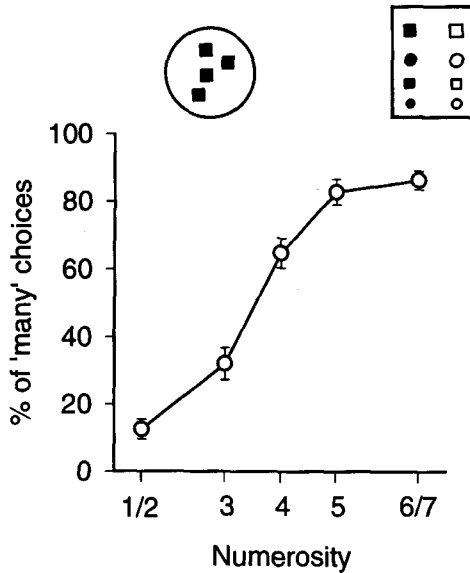


Figure 1. Mean data from Test 1, in which the elements within each stimulus were of the same size and shape. Scores are presented as percentage of choices to the side key that was correct if a "many" type of stimulus was presented. Vertical bars in this and subsequent figures show standard errors. The circular inset gives an example of a test stimulus. The rectangular inset shows the type of elements used for different test stimuli. For purposes of illustration, stimulus contrast has been reversed: the actual stimuli consisted of white elements on a dark background.

tation incurred by using stimuli with elements of variable size.

In the second test, the elements within a stimulus were of the same form (filled or unfilled squares or circles) but two or three different sizes of items were mixed within a stimulus. (Only one test slide contained a single element, that is, numerosity 1.) The squares or circles on the original stimulus drawings measured 1.5, 2.5, or 5 mm across. After photographic reduction and then subsequent projection, the corresponding dimensions on the pecking key were calculated to be 1.3, 2.1, and 4.2 mm. The results of these test trials are presented in Figure 2.

An ANOVA of these data again revealed a significant main effect only of numerosity category [ $F(4,28) = 55.90, p < .001$ ]. Post hoc comparisons of discrimination scores in Newman-Keuls tests showed no significant differences ( $p > .1$ ) for numerosities 4 versus 5, 4 versus 6 or 7, and 5 versus 6 or 7. All other comparisons for pairs of numerosity groupings were significant ( $p = .01$ ).

**Test 3: Stimuli of equal brightness.** The remaining tests were performed with the second group of 10 animals. In the first test with these pigeons, the constituent elements within each stimulus were again homogeneous in form and size. However, all the test stimuli were equated in luminance by first calibrating them with a luxmeter and then affixing the appropriate neutral-density filters to the individual slides. The mean luminance of the set of test stimuli was  $0.29 \log \text{cd/m}^2$  with a standard deviation of  $0.01 \log \text{cd/m}^2$  and a luminance range of  $0.28\text{--}0.32$

$\log \text{cd/m}^2$ . This was also the first test in which these birds encountered numerosities 3, 4, and 5. The results of this test are shown in Figure 3.

An ANOVA yielded a significant main effect of numerosity category [ $F(4,36) = 49.38, p < .001$ ]. Mean data comparisons in Newman-Keuls tests showed no significant differences ( $p \geq .1$ ) in discrimination scores for numerosity pairs 4 versus 5, 4 versus 6 or 7, and 5 versus 6 or 7. All other comparisons showed reliable differences ( $p = .05$  for 1 or 2 vs. 3,  $p = .01$  for the remaining pairs).

**Test 4: Outline and filled elements.** If birds assess the relative number of light and dark seeds (e.g., rice and linseed) lying on the ground, they could judge the numerosity of one subgroup (for instance, the lighter seeds that contrast more strongly with the background) rather than the overall amount. The next test was designed to model this situation by presenting stimuli that consisted of mixtures of outline and filled elements that varied in their relative numbers (and proportions). The mean results are presented in Figure 4.

In an ANOVA, the main effect of numerosity was significant [ $F(4,36) = 125.62, p < .00001$ ]. In Newman-Keuls tests, there was no significant difference ( $p > .1$ ) between mean discrimination scores for test stimuli with 5 elements versus those with 6 or 7. Comparisons of all other numerosity pairs were significant ( $p = .01$ ).

**Test 5: Elements of mixed shape.** Stimuli that pigeons are likely to encounter in their natural environment, such as food grains, also vary in form rather than being homogeneous in shape. The stimuli constructed for the next test mimicked that situation by including (for all numerosities greater than 1) two or three different shapes within each stimulus (i.e., some combination of circles, squares, and triangles). Once again, the question was whether birds

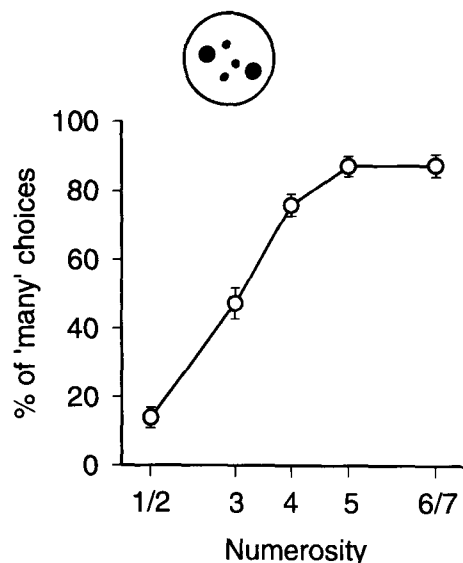


Figure 2. Mean scores and standard errors obtained in Test 2 for the effects of inhomogeneity in the size of items within a test stimulus.

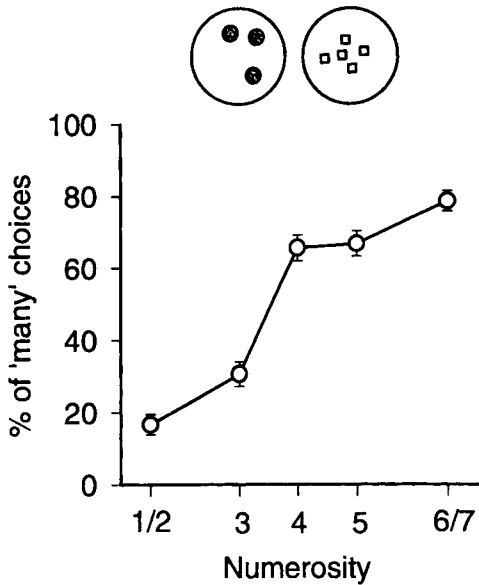


Figure 3. Results of Test 3, in which stimuli were equated in luminance by the addition of neutral-density filters to individual stimuli. Insets show the types of stimuli that were used, although luminance equality cannot be properly emulated in this figure.

would judge the numerosity of the entire array or of some subset within it. The results of testing for the effect of mixed shape on the relative numerosity judgments are shown in Figure 5.

In an ANOVA, the main effect of numerosity was significant [ $F(4,36) = 153.75, p < .00001$ ]. In Newman-Keuls tests, there was no significant difference ( $p = .1$ ) for a comparison of discrimination scores at numerosities 5 versus 6 or 7. All other comparisons were significant ( $p = .01$ ).

**Test 6: Separate end numerosities.** In all previous tests, the end numerosities 1 and 2 had been treated as a single category, as had the numerosities 6 and 7 at the other end of the range. This grouping of data corresponded to the way in which responses to training stimuli were scored. If pigeons can discriminate stimuli on the basis of the numerosity of their elements, they might also differentiate the stimuli in these end groupings. Up to this point, each test set contained only three exemplars for the category 1 or 2 and three for the group 6 or 7. To obtain sufficient data to analyze key choices with each of the numerosities from 1 to 7 separately, the size of the test set was increased. In this and the following tests, there were three new slides for each number value. Stimuli in Test 6 were composed of homogeneous elements of the same form and size. Test performance is shown in Figure 6.

An ANOVA revealed a significant effect of numerosity on percentage of "many" choices [ $F(6,54) = 133.72, p < .00001$ ]. Newman-Keuls tests showed that there were no significant differences ( $p \geq .1$ ) between the mean choice scores for numerosities 1 versus 2, 5 versus 6, and 5 versus 7. Comparisons of the means at numerosities 4 versus 6 and at 6 versus 7 revealed a significant difference

at  $p = .05$ . All other comparisons between means were significant at  $p = .01$ .

**Test 7: Stimuli with total contour equalized.** With stimulus elements that vary in size or form, a parameter that can covary with number is the amount of contour, that is, the transition edges that delineate the light from dark surfaces. In the case of outline stimuli, the elements can be regarded as consisting of contour alone. In the next test, conducted in two parts (six sessions each) with separate stimulus sets, the total amount of contour was held constant across all numerosities. The stimuli consisted of circles that were equated in total circumference. In the first set, the circles within a given stimulus were of the same diameter so that their sizes decreased with increasing numerosity. In the second set, a mixture of up to three different diameters occurred within a stimulus.

The total circumference of circles within a stimulus, when it was projected onto the pecking key, was calculated to be 18.9 mm for all the stimuli in the second set. This was also the total circumference for the majority of stimuli in the first test set. The exceptions were stimuli consisting of 5 circles, which had a total circumference of 17.7 mm, and those that consisted of 7 circles, which had a total circumference of 19.2 mm. These deviations from the norm represent error factors of 0.06 and 0.02, respectively. The results of the two parts of this test are shown in Figure 7.

ANOVAs yielded significant main effects of numerosity for both parts of this test [with fixed-diameter circles,  $F(6,54) = 109.23, p < .00001$ ; with variable-diameter circles,  $F(6,54) = 120.31, p < .00001$ ]. Subsequent Newman-Keuls tests on both sets of data revealed no significant differences between mean percentage choices of the "many" key at numerosities 1 versus 2, 5 versus 6, 5 versus 7, and 6 versus 7. Comparisons between all other

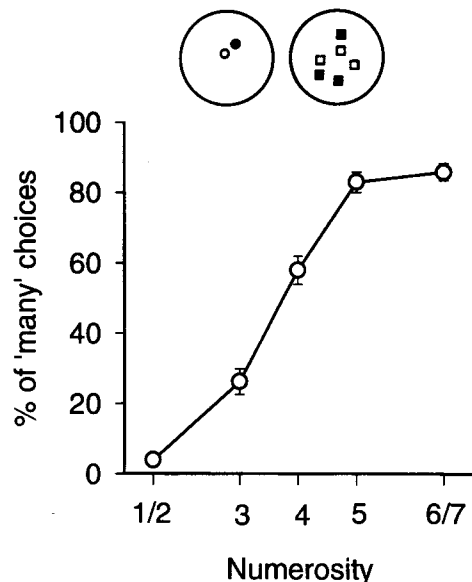


Figure 4. Data from Test 4, which investigated the effect of inhomogeneity of "dark" and "light" elements on discrimination performance.

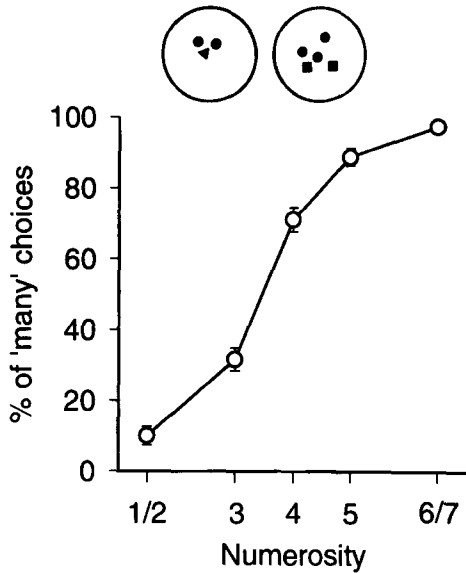


Figure 5. Results of Test 5 on the effect on performance of inhomogeneity in the shape of elements within a stimulus.

pairs of means showed significant differences ( $p = .05$  for numerosities 4 vs. 5 in the first part of the test, and  $p = .01$  for all other pairs).

**Test 8: Stimuli with total area of elements equalized.** A similar test was performed in which the integrated area of the filled white dots within a stimulus was held constant. The new stimuli in the first six sessions of this test phase consisted of dots that were of the same size for a particular numerosity. Thus the size of these dots decreased as their number increased. In a further six sessions, the test stimuli consisted of dots in mixtures of up to three different sizes.

When the stimuli were projected on the key, the mean total area of the homogeneous dots in the first test set was computed to be  $14.6 \text{ mm}^2$ . At different numerosities, the area ranged maximally  $\pm 0.4 \text{ mm}^2$  from the mean (an error factor of  $\pm 0.03$ ). In the second set of stimuli with dots of mixed sizes, the mean integrated dot area was also  $14.6 \text{ mm}^2$ . Areas across stimuli ranged maximally from  $0.4 \text{ mm}^2$  above to  $0.2 \text{ mm}^2$  below this mean (corresponding to error factors of 0.03 and 0.01, respectively). The results of the two parts of this test are illustrated in Figure 8.

In ANOVAs, the effect of numerosity on side-key choices was again highly significant in both parts of this test [for stimuli with consistent dot sizes,  $F(6,54) = 159.87$ ,  $p < .00001$ ; for stimuli with mixed dot sizes,  $F(6,54) = 181.21$ ,  $p < .00001$ ]. In both parts of the test, there were no significant differences on Newman-Keuls tests between mean percentage choices of the "many" key at numerosities 5 versus 6, 5 versus 7, and 6 versus 7 ( $p \geq .1$ ). Comparisons of all other mean data pairs revealed significant differences ( $p = .01$ ).

**Trend analysis.** Although there was a general trend for more choices to be directed to the "many" decision key as the arrays' numerosity increased, visual inspection of

the figures suggested that nonlinear, as well as linear, components might contribute to the form of the data plots. To check this, a trend analysis was applied to the test results, where appropriate.

Based on a repeated measures ANOVA with the single factor of stimulus numerosity, orthogonal trend components were computed separately for each set of data obtained in Tests 6–8, in which there were equal intervals between all the treatment levels of numerosity (Winer, 1971). For the results of Test 6, both parts of Test 7, and both parts of Test 8, only the linear trend component contributed significantly to the variance of the data. [For each of these tests, or subtests,  $F(1,54) = 13.54, 11.28, 12.14, 11.28$ , and  $16.79$ , respectively;  $p < .01$  in each case.] None of the quadratic or cubic components were significant [in each case,  $F(1,54) < 1, p > .05$ ].

## DISCUSSION

One question in this experiment concerned whether or not birds abstracted information about the relative number of items in small visual arrays, independently of other parameters that are often confounded with number. Of chief interest was how pigeons responded to the variety of novel stimuli that were introduced in transfer sessions during the acquisition phase, in novelty-habituation sessions, and in a series of test sessions.

In transfer sessions with both groups of subjects, the generally high level of discrimination performance with new stimulus exemplars indicated that the pigeons could distinguish the categories of "few" versus "many" elements, as these terms are defined within this experiment. This at least suggests that they utilized relative number information, since their discrimination performance could not have been based solely on the rote learning of dot patterns or on detection of brightness differences. The con-

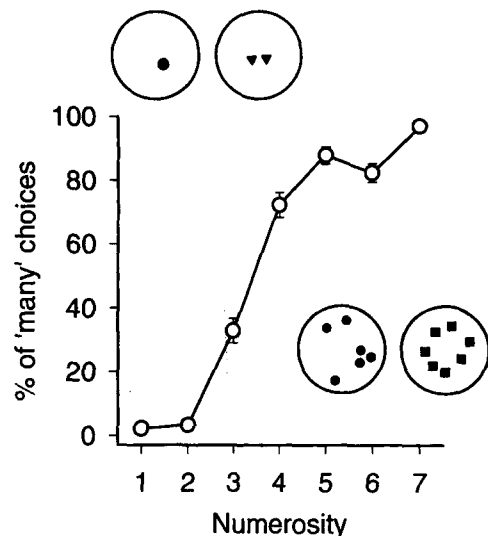
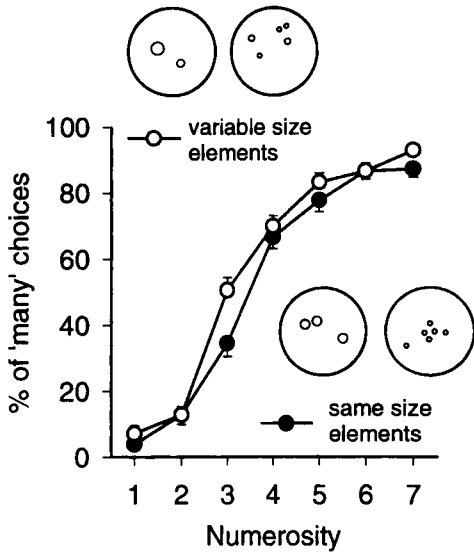


Figure 6. Results of Test 6, in which there were equal numbers of test stimuli for each numerosity from 1 to 7. (For further details, see text.)



**Figure 7.** Data from Test 7, in which the summed circumferences of circles within a stimulus were held constant across all numerosities. In one set of test stimuli, elements within a stimulus were of uniform circumference. In the other set, items within a stimulus varied in size.

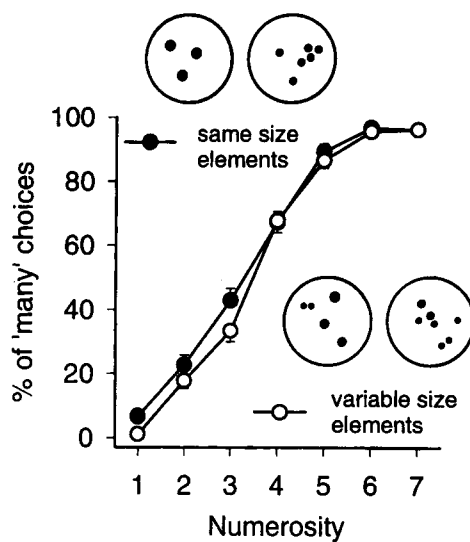
clusion that the birds' discrimination relied on the abstraction of relative number is strengthened by data from the first sessions of the novelty-habituation phase. Data shown in Table 2 give the birds' discrimination scores on first exposure to these new stimulus types. In these initial sessions, performance with novel stimulus exemplars was assessed before reinforcement contingencies with some of these stimuli could have been effective in influencing the birds' choices. Discrimination scores with new stimuli in these sessions were well above chance, and also similar to scores with familiar stimuli, despite the various stimulus modifications that were introduced with the numerosities 1 or 2 and 6 or 7. Thus, although partial reinforcement was used in this phase to preclude novelty itself from disrupting their responding, the birds' choices, when they were initially confronted with a change in stimulus type, did not depend on training effects.

The most compelling evidence that discrimination was based on the property of relative number came from the results in the test phases. The data of chief interest in these phases were derived from nonreinforced choices with further sets of novel stimuli. In addition to new stimulus configurations' being used for the anchor values (1 or 2 and 6 or 7), entirely new numerosities were introduced. The numerosities 3, 4, and 5 were presented for the first time in Tests 1 and 3 for the first and second group of pigeons, respectively. The results from the first sessions of the novelty-habituation phases had already established that the birds immediately discriminated new styles of "few" and "many" stimuli at levels well above chance. The test results showed once more that the animals responded appropriately to further novel exemplars of these familiar numerosities. More importantly, however, both groups of

birds also responded to the novel intermediate numerosities as if they judged them to lie between the "few" and "many" quantities that had been used in training. Furthermore, the data showed that as the number of constituent elements in the stimulus increased, so also did the percentage of "many" choices. This correspondence between choice performance and the serial order of stimuli on a number scale was maintained in the other tests too. Trend analyses on the data in the last tests showed that a simple linear trend accounted for most of the variance. However, many of the data plots appear to be slightly curved or flattened toward the higher end of the numerosity range tested here. This apparent flattening is reflected in the results of post hoc tests that showed, by and large, no significant differences in the birds' discrimination of adjacent numerosities at the higher end of the numerosity scale (5, 6, and 7) but reliable differences among the smaller numerosities in the range of 1 to 4. There were, however, slight variations in the form of the data plot from one test condition to another, so, without having a priori reasons for choosing a specific mathematical function, we did not attempt more precise curve fitting.

The question of whether the birds would spontaneously differentiate the numerosities 1 from 2 and 6 from 7 was addressed in Tests 6-8. If they did so, this would indicate that their choice behavior was controlled very precisely by the number of items in a stimulus. But a differentiation among these stimuli at each end of the test range would have to override the animals' prolonged equivalence training, since, in acquisition, they had to make the same response when 1 or 2 items were presented and an alternative response when 6 or 7 items were displayed.

The effects of equivalence training may indeed have predominated, since these tests yielded contradictory re-



**Figure 8.** Results of Test 8, in which the integrated area of white dots within a stimulus was held constant across all numerosities. For one set of test stimuli, the size of dots within a stimulus was uniform. For the other test set, these elements were of mixed size within a stimulus.



sults. In Test 6, the pigeons did not differentiate numerosities 1 and 2, but this was probably due to a floor effect. Choice responses for numerosities 6 and 7 were significantly different (although marginally so), but the lack of significant discrimination between numerosities 5 and 7 detracts from this finding. In Test 7, neither 1 versus 2 nor 6 versus 7 were differentiated, whereas in Test 8, the birds were more likely to choose "few" for 1 dot than for 2 dots, but were no more likely to choose "many" for 7 than for 6 dots, although in the latter case this result may be due merely to a ceiling effect. Since other numerosities were reliably differentiated in each of these tests, there does not seem to be any reason to attribute differences in discriminability of the end numbers to variations in other stimulus characteristics that were manipulated in each test.

Irrespective of the differentiation of the numerosities at the ends of the test range, discriminability was still better for the smaller numbers of items (2 vs. 3, 3 vs. 4) than for the larger numbers (4 vs. 5, 5 vs. 6). This qualitative effect is consistent with Weber's law, which seems to apply when numerical discriminanda consist of sequences of items (e.g., a series of an animal's own responses, Mechner, 1958; or of tone pulses, Meck & Church, 1983; see Gallistel & Gelman, 1992, for a review). So far, there have been no systematic studies with animals of the application of Weber's law to stimuli composed of small numbers of simultaneously presented items of the type used here. The present experiments, with a mixture of numerosity values as anchor points in training, were not originally designed to assess bisection points. Nevertheless, from each of the data curves obtained in the various test phases, it was possible to calculate a point of subjective equality (*PSE*), the numerosity value for which the birds made 50% of their choices to the "many" key and 50% to the "few" key. The mean *PSE* across all the tests is 3.41 ( $SD = 0.23$ ). A difference limen (*DL*) can also be computed as half the difference between the numerosity corresponding to 75% choices of "many" and the numerosity corresponding to 25% choices of "many." The mean *DL* is 1.10 ( $SD = 0.30$ ). The Weber fraction is then given by  $DL/PSE$ . The mean value for the Weber fraction is 0.32 ( $SD = 0.09$ ). In Meck and Church's (1983) study, the *PSE* was equal to the geometric mean, rather than the arithmetic mean, of the values they used to train rats to discriminate sequences of external events. The rats' choices of "many" or "few" thus seemed to be based on the ratio of the number of tone stimuli presented on a test trial to the number in either of the anchor stimuli to which they had been trained. So far, the data for pigeons that have been trained on bisection tasks involving sequential number are rather inconsistent. When the discrimination is based on the number of feeder light flashes, birds' *PSE* lay between the harmonic and geometric mean (Fetterman, Stubbs, & Dreyfus, 1986). When birds had to discriminate the numbers of responses they emitted, the bisection point was either at the harmonic mean (Fetterman, Dreyfus, & Stubbs, 1985) or close to the geometric mean (Fetterman, 1993). The reason for these inconsistencies is not yet clear.

The arithmetic mean of the numerosities to which the birds in our experiment were trained was 4.0. If it is assumed that the anchor values in training were 1.5 (from the mixture of stimuli with 1 or 2 dots for "few") and 6.5 (from the mixture of 6 or 7 for "many"), then the geometric mean would be at a numerosity value of 3.12 whereas the harmonic mean would lie at 2.44. (Both of these means would be lower if, instead, the anchor values in training were assumed to be 1 and 7.) The estimate for the *PSE* from the present data does not fit the harmonic mean, but whether or not the birds were also basing their choices on a ratio computation remains indeterminate. We intend to investigate this topic in the future since precise data of this type would contribute to a better understanding of the way animals scale number and of the processes underlying the discrimination of numerosity in simultaneous displays.

An alternative explanation for the slightly better discriminability of smaller test numbers, up to about 4, may also be considered. Drawing parallels with some of the early work on human subjects' quick and accurate identification of small numbers of simultaneously presented visual items, Davis and Pérusse (1988) have proposed that animals, like humans, rely on a process of subitizing to discriminate among small numbers. Although the precise nature of this process remains unclear, subitizing supposedly involves a form of pattern recognition as a means of differentiating stimuli that also happen to vary in number (e.g., 3 dots are often arranged in a triangle, 4 in a quadrilateral, etc.).

There are a number of problems in accepting subitizing as a likely mechanism for numerosity discrimination, in spite of the present finding that discriminability was better within the numerical range of 1 to 4, in which subitizing is meant to operate. First, it is difficult to see how pattern recognition could account for the data we obtained, particularly since orderly data were obtained with new numerosities, 3, 4, and 5, which would yield pattern configurations different from those that are formed with 2 dots or with 6 or 7. The other problems are independent of our results. The evidence for subitizing in humans comes from a reported change in the slope of the reaction time (RT) function. With numerosities up to about 4 items, RTs are fitted to a line of shallow slope, whereas with greater numerosities, the RT function has a steeper slope (e.g., Chi & Klahr, 1975; Klahr & Wallace, 1976). But as Miller (1993) has pointed out, no such RT data exist from animal studies. Furthermore, a more recent, detailed study of humans' RTs when they discriminate numerosity yielded no evidence for a discontinuity in the RT function and has called into question the existence of subitizing as a specialized process for assessing number (Balakrishnan & Ashby, 1992). At present, it seems more parsimonious to assume that numerosity discrimination in birds, as in humans (van Oeffelen & Vos, 1982), relies on probabilistic judgments that obey Weber's law, without there being separate underlying processes for limited parts of the number continuum.

Another question in this study was whether discrimination was influenced by stimulus parameters other than that of numerosity. Since an overall change in stimulus brightness is the most obvious potential confound with varying number, several phases of the experiment included controls for the effect of brightness variations. In transfer sessions toward the end of the acquisition phase, the discrimination performance of both groups of birds was maintained when novel stimuli of equal overall luminance were introduced so that brightness did not seem to be the essential discriminative cue.

Similarly, in Test 3, stimulus intensity was controlled by adding neutral-density filters to individual test stimuli. The pigeon's brightness discrimination threshold is  $0.11 \log \text{ cd/m}^2$  when white light on a single key is successively varied in luminance (Hodos, Bessette, Macko, & Weiss, 1985). The standard deviation of mean luminance in the test stimuli employed here was below this threshold. Affixing neutral-density filters to the stimuli had the disadvantage, however, of confounding equalized luminance with reductions in contrast between the white elements and their dark background. This confound was not present in Test 8 which provided an additional control for stimulus luminance by equating the total area of white light on the response key. In this test, the residual luminance variation (mean luminance =  $1.26 \log \text{ cd/m}^2$ ,  $SD = 0.08 \log \text{ cd/m}^2$  with the same-sized stimulus elements;  $M = 1.27 \log \text{ cd/m}^2$ ,  $SD = 0.04 \log \text{ cd/m}^2$  with mixed-sized elements) was also below the pigeon's threshold level of brightness discrimination. The differential choice of the "many" response key when test luminance was equated shows that discrimination was not dependent on this parameter.

In other tests, various aspects of stimulus inhomogeneity were modified to see if the birds based their discrimination on subgroupings of identical stimulus elements. Mixtures of outline and filled elements, presented in Test 4, were such that, if conditional discrimination had been based only on elements of the same type, then the distribution of choices on test trials could not have been so clearly related to the total stimulus numerosity as it was. None of the controls for the effect of inhomogeneity (outline vs. filled elements, elements of mixed shape, elements of variable size) produced a deterioration in discrimination (i.e., the data curves did not flatten or lose their consistent trend). This suggests that the specific type of items that comprised a stimulus configuration was not critical to the pigeon's ability to assign a numerosity value to the stimulus. The training-and-transfer phase of the experiment had in any case been designed to mitigate against the animals' relying on a discrimination of specific features. Furthermore, although some of the novel stimuli were differentially reinforced during the novelty-habituation phases, discrimination on initial exposure to various changes in stimulus characteristics was well above chance on the first sessions of those phases. In any case, since intermediate numerosities (3, 4, and 5) were never differentially reinforced in any part of the experiment, the birds' choices with these stimuli apparently de-

pendent on their judgments of the relative numerosity values of the stimuli.

Test 2 was based on the further possibility that pigeons might use a process resembling "mental zooming" in humans in order to recognize the equivalence of items that were discrepant in size. On the one hand, pigeons can discriminate very small differences in size. For instance, these birds can detect a difference of less than 2% in the lengths of two simultaneously displayed bars when the reference length is about 11 mm (Schwabl & Delius, 1984). With successively presented stimuli more similar to the components of our task, pigeons can just discriminate a 3-mm-diam dot from one that is 10% larger in size (Hodos, 1993). The relative size differences among the stimulus elements in Test 2 were greater than that since the diameters of the circles and squares that made up the stimuli in that test varied by 40%–70%. On the other hand, in spite of their ability to detect size differences, pigeons, like humans, are capable of perceptual size invariance when these birds are tested in an oddity-from-sample task (Lombardi & Delius, 1990). The mechanism underlying invariance performance may, of course, differ in pigeons and people. Be that as it may, whatever extra perceptual processing might be involved, the variation in the size of stimulus elements did not critically influence the present test results.

Consistent numerosity discrimination was also maintained in Tests 7 and 8, in spite of the elimination of other potentially confounded cues. When either the total amount of stimulus contour within each stimulus was equated across numerosities or the integrated area of the white stimulus surfaces was equalized, the test data were similar to those obtained in the other tests in which these factors were not controlled.

The overall conclusions that can be drawn from the present experiment are that these birds discriminated the visual stimuli on the basis of the numerosities of their constituent items and did not rely on any of the other parameters that were manipulated in this study. This general finding concurs with the increasing amount of evidence for some form of numerical competence shown not only by pigeons but also by a variety of other species (e.g., rats, raccoons, parrots, monkeys, chimps) that have been tested with different procedures and stimulus material (see Boysen & Capaldi, 1993; Davis & Memmott, 1982; Davis & Pérusse, 1988). The pigeons generalized their discrimination in a categorical way to novel displays of small and larger numbers of items, at least within the range of 1 to 7. To optimize the experimental situation, "seed-like" stimuli were created. Although monochrome, two-dimensional, and geometric, the forms and sizes of the constituent elements did not deviate greatly from those of normal food grains. Although this use of stimuli resembling ones encountered daily by pigeons was intended to expedite learning, and indeed to model a situation in which they might apply numerosity discrimination, these animals also display numerically based discrimination abilities when highly artificial stimulus arrays are employed, for example, sets of stickers depicting stylized birds, flowers, or even unicorns (Honig & Stewart, 1989).

Within the limits tested here, the birds could also spontaneously differentiate entirely novel numerosities. Their categorization of these stimuli corresponded to a serial order along a scale of the number of items in each display. Furthermore, the birds' categorization choices had to be based on their memory for anchor numerosities since, in the conditional discrimination paradigm we used, only one array was presented at a time, so that no direct perceptual comparison of stimuli was possible.

In the present investigation, the animals were not required to make a defined response to a specific and absolute number of elements. The behavior tested here would not provide adequate evidence of an ability to count, according to the criteria for such an ability defined by Gelman and Gallistel (1978), although the birds clearly displayed numerical competence (Davis & Pérusse, 1988). Earlier work by Koehler, in which pigeons were able to "act upon" a particular number of objects or events (such as eating only 3 seeds out of a variable number: Koehler & Wachholtz, 1936), encourage us to think that the pigeon's numerical abilities extend beyond those that could be demonstrated here.

#### 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.
- BALAKRISHNAN, J. D., & ASHBY, F. G. (1992). Subitizing: Magical numbers or mere superstition? *Psychological Research*, *54*, 80-90.
- BOYSEN, S. T., & CAPALDI, E. J. (1993). *The development of numerical competence: Animal and human models*. Hillsdale, NJ: Erlbaum.
- BROWN, P. L., & JENKINS, H. M. (1968). Auto-shaping the pigeon's key peck. *Journal of the Experimental Analysis of Behavior*, *11*, 1-8.
- BUNDESEN, C., & LARSEN, A. (1975). Visual transformation of size. *Journal of Experimental Psychology: Human Perception & Performance*, *1*, 214-220.
- 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.
- CHI, M. T. H., & KLAHR, D. (1975). Span and rate of apprehension in children and adults. *Journal of Experimental Child Psychology*, *19*, 434-439.
- DAVIS, H. (1993). Numerical competence in animals: Life beyond Clever Hans. In S. T. Boysen & E. J. Capaldi (Eds.), *The development of numerical competence: Animal and human models* (pp. 109-125). Hillsdale, NJ: Erlbaum.
- 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 & Brain Sciences*, *11*, 561-615.
- EMMERTON, J., & DELIUS, J. D. (1993). Beyond sensation: Visual cognition in pigeons. In H. P. Zeigler & H.-J. Bischof (Eds.), *Vision, brain, and behavior in birds* (pp. 377-390). Cambridge, MA: MIT Press.
- FELLOWS, B. J. (1967). Chance stimulus sequences for discrimination tasks. *Psychological Bulletin*, *67*, 87-92.
- FERNANDES, D. M., & CHURCH, R. M. (1982). Discrimination of the number of sequential events by rats. *Animal Learning & Behavior*, *10*, 171-176.
- FETTERMAN, J. G. (1993). Numerosity discrimination: Both time and number matter. *Journal of Experimental Psychology: Animal Behavior Processes*, *19*, 149-164.
- FETTERMAN, J. G., DREYFUS, L. R., & STUBBS, D. A. (1985). Scaling of response-based events. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*, 388-404.
- FETTERMAN, J. G., STUBBS, D. A., & DREYFUS, L. R. (1986). Scaling of events spaced in time. *Behavioural Processes*, *13*, 53-68.
- GALLISTEL, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.
- GALLISTEL, C. R. (1993). A conceptual framework for the study of numerical estimation and arithmetic reasoning in animals. In S. T. Boysen & E. J. Capaldi (Eds.), *The development of numerical competence: Animal and human models* (pp. 211-223). Hillsdale, NJ: Erlbaum.
- GALLISTEL, C. R., & GELMAN, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*, 43-74.
- GELMAN, R., & GALLISTEL, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.
- HODOS, W. (1993). The visual capabilities of birds. In H. P. Zeigler & H.-J. Bischof (Eds.), *Vision, brain, and behavior in birds* (pp. 63-76). Cambridge, MA: MIT Press.
- HODOS, W., BESSETTE, B. B., MACCO, K. A., & WEISS, S. R. B. (1985). Normative data for pigeon vision. *Vision Research*, *25*, 1525-1527.
- HONIG, W. K. (1991). Discrimination by pigeons of mixture and uniformity in arrays of stimulus elements. *Journal of Experimental Psychology: Animal Behavior Processes*, *17*, 68-80.
- HONIG, W. K., & STEWART, K. E. (1989). Discrimination of relative numerosity by pigeons. *Animal Learning & Behavior*, *17*, 134-146.
- HONIG, W. K., & STEWART, K. E. (1993). Relative numerosity as a dimension of stimulus control: The peak shift. *Animal Learning & Behavior*, *21*, 346-354.
- KLAHR, D., & WALLACE, J. G. (1976). *Cognitive development: An information-processing view*. Hillsdale, NJ: Erlbaum.
- 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" experiments with a raven and comparative experiments with humans]. *Zeitschrift für Tierpsychologie*, *5*, 575-712.
- KOEHLER, O., & WACHHOLTZ, R. (1936). Weitere Versuche an der Taube "Nichtweiß" zur Frage des "Zählvermögens" [Further experiments with the pigeon "Nichtweiß" concerning the question of "counting ability"]. *Verhandlungen der Deutschen Zoologischen Gesellschaft*, *38*, 211-236.
- LOMBARDI, C., & DELIUS, J. D. (1990). Size invariance in visual pattern recognition by pigeons. In M. L. Commons, R. J. Herrnstein, S. M. Kosslyn, & D. B. Mumford (Eds.), *Quantitative analyses of behavior: Vol. VIII. Behavioral approaches to pattern recognition and concept formation* (pp. 41-65). Hillsdale, NJ: Erlbaum.
- MECHNER, F. (1958). Probability relations within response sequences under ratio reinforcement. *Journal of the Experimental Analysis of Behavior*, *1*, 109-122.
- 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.
- MILLER, D. J. (1993). Do animals subitize? In S. T. Boysen & E. J. Capaldi (Eds.), *The development of numerical competence: Animal and human models* (pp. 149-169). 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.
- SCHWABL, U., & DELIUS, J. D. (1984). Visual bar length discrimination threshold in the pigeon. *Bird Behaviour*, *5*, 118-121.
- VAN OEFFELEN, M. P., & VOS, P. G. (1982). A probabilistic model for the discrimination of visual number. *Perception & Psychophysics*, *32*, 163-170.
- WINER, B. J. (1971). *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill.