Anim Cogn (2009) 12:511–526
DOI 10.1007/s10071-009-0212-5

ORIGINAL PAPER

Local rather than global processing of visual arrays in numerosity discrimination by pigeons (*Columba livia*)

Jacky Emmerton · Jennifer C. Renner

Received: 10 December 2008 / Revised: 9 January 2009 / Accepted: 9 January 2009 / Published online: 30 January 2009

Abstract A series of experiments investigated which stimulus properties pigeons use when they discriminate pairs of visual arrays that differ in numerosity. Transfer tests with novel stimuli confirmed that the birds' choices were based on relative differences in numerosity. However, pigeons differed from other species in the non-numerical cues that affected their choices. In human and non-human primates, numerical discrimination is often influenced by continuous variables such as surface area or overall stimulus brightness. Pigeons showed little evidence of using those cues, even when summed area and brightness had been correlated with numerosity differences and reward outcome. But when array-element sizes were asymmetrically distributed across numerosities, the birds readily utilized information about item sizes as an additional discriminative cue. These novel results are discussed in relation to pigeons' tendency to focus on local, rather than global dimensions when they process other non-numerical complex visual stimuli. The findings suggest there may be inter-specific differences in the type of perceptual information that provides the input stage for mechanisms underlying numerical processing.

Keywords Numerosity discrimination · Size cues · Local vs. global visual processing · Pigeon

Introduction

Numerical competence has been demonstrated in a variety of avian and mammalian species (Davis and Pérusse 1988). A few experiments have shown that animals can discriminate cardinal numbers of items and associate these quantities with symbols (chimpanzees: e.g. Boysen 1993; Matsuzawa 1985; parrots: Pepperberg 1987; pigeons: Xia et al. 2001). An ability to distinguish the absolute number of things may be adaptive in some circumstances. For instance, American coots, which are brood parasites, discriminate the cardinal number of their own eggs within a clutch (Lyon 2003). Perhaps of more widespread utility is the ability to discriminate relative differences in numerosity. For example, when foraging, an animal may not need to know the exact number of items in a food patch, but whether one patch has more in it than another one. Experimentally, both birds (e.g. Emmerton and Renner 2006; Fetterman 1993) and mammals (e.g. Jordan and Brannon 2006; Meck and Church 1983; Tomonaga 2008) discriminate differences in numerosity in a way that obeys Weber's law (i.e. their performance depends on numerosity ratios).

In an organism's natural environment, groups of items that differ in number often differ in other properties too. For instance, two or more groups of nuts may differ in the total surface area or combined volume of each set, in their overall brightness, or in the extent of each group. In this example, all of these properties can vary on a continuous scale, and so may be correlated with the number of nuts. Some have claimed that such continuous variables are more salient to animals than numerosity, and are more likely to be the stimulus attributes that determine their choices about quantity (Davis 1993; Seron and Pesenti 2001). While this view that animals discriminate number

---

Electronic supplementary material The online version of this article (doi:10.1007/s10071-009-0212-5) contains supplementary material, which is available to authorized users.

J. Emmerton (✉) · J. C. Renner
Department of Psychological Sciences, Purdue University,
703 Third Street, West Lafayette, IN 47907-2081, USA
e-mail: jemmert@psych.purdue.edu

only as a last resort (Davis and Memmott 1982; Davis and Pérusse 1988) has been disputed (Capaldi and Miller 1988; Cantlon and Brannon 2007), there is evidence from a variety of species, including humans, that certain continuous variables are either the basis for differentiating quantity, or at least influence numerosity discrimination to some extent.

Some studies with human infants have shown that they rely on differences in contour length (Clearfield and Mix 1999) or cumulated surface area (Feigenson et al. 2002), rather than numerosity per se, to respond differentially to stimuli that consist of small numbers of items. Although infants may use these continuous dimensions when all the stimulus objects are identical, they do react to numerosity with heterogeneous objects (Feigenson 2005). Even if infants fail to respond to numerosity differences when surface area is matched for small numbers of items (≤ 3), they do so when either contour length or area is controlled in displays of larger numbers of filled circles (Xu et al. 2005).

Several studies of numerical competence in non-human primates have, like ours, used artificial two-dimensional visual stimuli. Numerosity has varied over a fairly wide range, and non-numerical stimulus parameters have been manipulated. Of relevance to our own experiments are tests for the effect of stimulus area. In a matching to sample task, macaque monkeys preferentially chose stimuli that matched in numerosity rather than surface area when these cues, which were confounded in training, were pitted against each other in probe trials (Cantlon and Brannon 2007). But stimulus area served as an adjunct cue to numerosity when the monkeys chose a match and both cues were congruent. Moreover, monkeys that had no prior experience in numerosity discrimination tasks were more influenced by surface area than were number-experienced monkeys. This was particularly so when relative disparity in numerosity decreased, making numerical matching more difficult. Similarly, when chimpanzees had to choose the more numerous of paired arrays of dots, their discrimination accuracy declined more at smaller numerical difference ratios when stimulus area was equated than when numerosity and area were congruent (Tomonaga 2008). But again, this cue was additive to numerosity, rather than being the more salient cue for the apes.

These effects of continuous variables that are confounded with number occur not just with artificial stimuli. For instance, an experiment with mosquitofish utilized their natural shoaling behaviour in preference tests involving a fish's approach to one of two groups of conspecifics. The fish being tested approached the more numerous group when the groups also differed in other dimensions. But when the "stimulus" fish were of sizes that equated the two groups for surface area, the tested fish no longer showed any preference, either with small numbers (<4) or with larger numbers of fish in the comparison groups (Agrillo et al. 2008).

In contrast to these findings that continuous variables can influence or determine choices when stimuli differ in numerosity, previous studies with pigeons have so far shown relatively little effect of these types of stimulus attributes on the birds' performance in numerical tasks. For instance, Emmerton et al. (1997) found that birds' numerosity discrimination was still maintained when a variety of continuous (brightness, area) or other non-numerical (stimulus configuration, contrast, and item shape) cues were controlled in tests with novel stimuli. In that study, pigeons were presented with one array at a time in a conditional discrimination paradigm. Similarly, Honig and Stewart (1989) reported no effects on relative numerosity discrimination of altering the summed areas of the stimulus components or the types of elements when their birds were tested in a go/no-go procedure. Some of these findings have been confirmed and extended in a conditional discrimination task in which pigeons were trained with stimuli that varied across trials in the configuration and sizes of items, as well as in the overall area and brightness of array elements (Emmerton and Renner 2006). Performance was maintained when area and brightness were subsequently equated across numerosities. Also, the birds' choices in other tests were related to numerosity, and not to the summed "filled" areas occupied by spread of activity amongst neighbouring items (cf. Allik and Tuulmets 1991; Van Oeffelen and Vos 1983; Vos et al. 1988).

In another experiment (Emmerton 1998), changing the density of the stimulus array had subtle effects on the accuracy with which pigeons discriminated numerosity differences. In humans, decreasing the overall display density (or increasing the dot spacing) can lead to increases in numerosity estimates (e.g. Allik and Tuulmets 1991; Vos et al. 1988). Increasing the spacing between array items appeared to have the opposite effect on pigeons' numerosity judgements. The pattern of results in Emmerton's (1998) experiment suggested that the birds made more errors of omission when they processed arrays in which the items were farther apart. It seems that for pigeons the distance between individual items may be important, rather than the density as an overall, continuous attribute of the stimulus. Emmerton and Renner (2006) also found that pigeons tended to judge the numerosity to be less when array items were spread out than when they were densely packed. Based on these effects of item spacing it was hypothesized that pigeons might "scan" an array, and either process items one by one, or else process localized regions of the stimulus, rather than processing all the elements of an array in a global fashion, as has been suggested in some models of numerical processing (e.g. Dehaene and Changeux 1993).

Some of the effects of non-numerical stimulus properties therefore appear to differ between pigeons and other species—in particular human or non-human primates on which much of the research has been conducted. We have suggested previously (Emmerton and Renner 2006) that some of these disparate effects might be related to a tendency by pigeons to focus on local details, or elements, of a complex stimulus (cf. Cerella 1980), rather than on the overall configuration, as humans do for instance. Thus pigeons sometimes show a “local precedence effect” (Cavoto and Cook 2001; Fremouw et al. 2002; but see Goto et al. 2004) in contrast to the “global precedence effect” more commonly shown by humans (Navon 1977).

This contrast between local and global effects has usually been investigated in the context of shape perception, with stimuli in which the shape of constituent elements may be congruent or incongruent with the overall shape formed by these elements. Shape perception is not our concern here since the elements in our stimuli were all small rectangles, and the array form varied within and across trials. Instead we are interested in whether pigeons use global information, in the sense of continuous variables that must be integrated from a perception of the whole stimulus array, or local information, in the sense of attending to properties of stimulus elements, when they discriminate numerosity.

The stimuli in the current experiments were “patches” of visual items that the pigeons had to peck. Whereas most of the experiments described above on this species presented only a single array on each trial, this experiment employed a simultaneous discrimination procedure with a pair of arrays displayed on each trial. The birds had ample opportunity to view and process both stimuli. If pigeons are able to utilize continuous variables in discriminating stimuli that differ in quantity, this procedure should make it easier for them to do so. The main aim of the experiments was to investigate what types of non-numerical cues are most likely to affect pigeons’ choices when the arrays they see differ reliably in numerosity. The reasons for doing these experiments are twofold. Firstly, it is important to know which cues a species is most likely to detect if these cues are to be controlled in future studies of that species’ numerical abilities. Secondly, identifying these cues may further elucidate the perceptual and cognitive mechanisms that mediate numerical competence in different species.

We started this study by allowing overall stimulus area (and brightness) to covary, on average, with numerosity. When these variables were subsequently controlled, it emerged that other cues, related to the sizes of stimulus elements themselves, were more effective as supplemental cues for pigeons when they discriminate numerosity.

General method

Subjects

Except for the final experiment, the same four adult, mixed-breed pigeons were used as subjects. The birds were housed individually in a colony room with a 14:10-hour light:dark cycle, and constant temperature and humidity. They were maintained at about 85% of their ad-lib body-weights. Health-grit and water were available in their home-cages. The birds all had previous experimental experience, but not with the task described here.

Apparatus

The operant conditioning chamber measured 33 cm × 34.5 cm × 32 cm. Stimuli were presented on a VGA monitor (Zenith 1492 FTM, flat screen) located just behind an opening (26.5 cm × 15 cm) on one side of the chamber. The monitor was equipped with an infrared scanning touch screen (CarrollTouch) that sensed a bird’s responses. A 1 mm-thick clear plastic sheet protected the monitor screen, and cushioned a bird’s pecks. Openings (3.5 cm × 5.5 cm) on the left and right walls gave access to two grain-feeders (Colbourn Instruments, Model E14-10) which could be lit separately. An infrared light gate, added to each feeder, sensed when a bird’s head was in the opening. A dim (1.2 W) house-light, positioned towards the rear of the box, was shielded to minimize light reflection on the computer screen. A PC controlled experimental events via an interface.

Numerosity stimuli

Stimuli were generated using Borland Turbo Pascal. Numerosity arrays consisted of small rectangles in either white or grey (as described below), against a dark background. Most of the arrays consisted of elements in mixed sizes. There were three sizes of elements. Small ones were programmed to be 2 × 2 pixels, medium-sized items were 3 × 3 pixels, and large items were 4 × 4 pixels, but since the arrays were presented in EGA mode, these elements appeared as rectangles. Based on calibrations of blocks of rectangles, the on-screen sizes of the small, medium, and large items were calculated to be 1.19 mm × 1.55 mm, 1.54 mm × 2.06 mm, and 1.96 mm × 2.54 mm (W × H) respectively. Luminance calibrations (described in Emmerton and Renner 2006) showed that, for a given element size and colour, stimulus brightness increased linearly with the number of elements.

In agreement with previous reports (Maldonado et al. 1988), the birds in our experiments were observed to stay close to the monitor, facing it, so they most likely viewed the stimuli in their frontal visual field. Before pecking a stimulus, pigeons fixate it from a distance of about 6 cm (Goodale 1983; Macko and Hodos 1985). At this viewing distance, a single small array-element subtended a visual angle of about 1.1° × 1.5°, a medium-size element 1.5° × 2°, and a large element 1.9° × 2.4°.

Coding files specified the number of elements to appear in a given array, and the colour and size of each item. These files also stored the locations of the elements, chosen from an irregular matrix of 75 possible locations in the first two experiments, and 80 locations in the subsequent experiments. The configurations of items varied across arrays. In creating arrays we avoided having orphaned elements, or items that were spread out, since pigeons are more likely to underestimate the number of items if they are widely separated (Emmerton 1998). The horizontal and vertical extents of arrays were measured on screen to calculate approximate array sizes. As examples, the mean sizes of training arrays in the first two experiments were 13.8 mm × 15.0 mm (13.1° × 14.3°) for 10-item arrays, and 9.4 mm × 11.4 mm (9.0° × 10.6°) for 6-item arrays. The mean sizes of training arrays in Experiments 3–5 and Experiment 7 were 17.5 mm × 20.2 mm (16.6° × 19.1°) for 10-item arrays, and 14.7 mm × 15.7 mm (14.0° × 14.9°) for 6-item arrays. While the size of a larger numerosity array often exceeded that of a paired smaller numerosity array, this was not always the case. But tests in Experiment 7 provided a control for array size as a potential discriminative cue.

General procedure and experimental design

A simultaneous discrimination procedure was employed. Trial parameters were gradually changed in the first few training sessions of Experiment 1. Initially, a trial began with a “ready” signal (a white circle, 30 pixels in diameter, filled with a cross-hatch pattern). The bird had to peck this stimulus to switch on a pair of numerosity arrays, and a single peck to one of the arrays led to food reward or timeout. As training progressed, a “preview” phase was introduced, and the number of pecks for choosing an array was increased. The final trial structure was as follows. The trial began with a 6-s “preview” phase. During this phase a pair of arrays was shown to the left and right of the screen. Pecks to these arrays were recorded but had no consequences. Then the central “ready” signal appeared, while the arrays stayed on. A single peck to the ready signal switched it off, leaving the arrays on the screen. During the subsequent “choice” phase of the trial, a bird could switch back and forth between arrays, but pecks to each array were accumulated until the bird had made a fixed ratio of 8 pecks (FR8) to one of the arrays. If FR8 accumulated first on the S+ array, a training trial was scored as correct and ended in food reward. If FR8 accumulated first on the S− array, a training trial was scored as incorrect and ended in timeout. Reward or timeout was followed by a 10-s intertrial interval.

For a food reward, the feeder light on the side of the S+ was switched on and the hopper activated. If the bird put its head in the feeder recess within 20 s, it then had 2.5 s feeding time before the feeder and light were switched-off. If the bird did not approach the feeder within 20 s, it forfeited this reward. A timeout consisted of 5 s with the house-light off. In training sessions, a timeout also initiated a correction procedure. The trial was repeated with the same arrays until the bird correctly chose the S+. Correction trials were recorded, but not used in data analysis.

Sessions were run 5 or 6 days a week. Training sessions consisted of 48 trials. The left and right locations of the S+ array varied quasi-randomly across trials (Fellows 1967). The birds were trained to choose the array with the greater numerosity. Except for Experiment 6, there were ten elements in the S+ array, and six in the S−. Training sessions continued until performance was ≥80% correct. Then, in preparation for testing, reinforcement probability was reduced in 0.1 steps until it reached 0.5.

Testing was conducted in series of five sessions. There were 96 trials in each test session: 48 reinforced trials, with the same stimulus configurations as in training, were intermixed with 48 non-reinforced test trials in which novel arrays were presented. For most test trials, cumulative FR8 choices of the array with the greater numerosity were still scored as correct, and choices of the smaller numerosity as incorrect. No correction procedure was used during test sessions. Data analysis focused on the scores obtained on test trials.

Experiment 1

The purpose of this experiment was to train the pigeons to discriminate quantity, and to establish some baseline data. The stimulus arrays differed reliably in their numerosity, the attribute we expected to serve as a primary cue for the birds (cf. Emmerton and Renner 2006). But other parameters (namely surface area and overall brightness) were not systematically controlled in training. Area and brightness were not perfectly correlated with numerosity, but on each trial the array with the greater numerosity also had the greater summed area, as well as luminance. Thus, these cues were ones the birds could potentially associate with reward during training.

Previous experiments (Emmerton and Renner 2006), in which pigeons were presented with one array at a time, showed that numerosity discrimination in these birds obeys Weber’s law, i.e. their performance depends on relative disparities in numerosity. To check that the birds' discrimination with pairs of arrays also conforms to Weber's law, tests were conducted in which the absolute numerosities of the stimulus pairs differed, while their relative disparity remained constant. If the pigeons had learned to discriminate relative numerosity differences they should transfer to the novel stimulus arrays used in the test phase. Differences in non-numerical cues were reduced in these test trials, and were not always consistent with numerosity.

Method

Pre-training

One bird had previously been trained in a conditional discrimination procedure in this apparatus, and did not need pre-training. The other birds were first trained to peck at a stimulus located centrally on the screen, and to feed from left and right feeders, in a modified autoshaping procedure. The stimulus was similar to the "ready" signal used in discrimination training.

Autoshaping was followed by instrumental training. The same stimulus as in autoshaping was presented randomly to the left or right of the monitor. The bird had to peck the stimulus once to activate the feeder on the same side as the stimulus.

Discrimination training

In training sessions, there were 24 S+ arrays and 24 S− arrays. Each array was presented twice per session, with left and right on-screen locations counterbalanced across trials. Specific S+ and S− arrays were recombined into different pairs, both within and across sessions. Arrays consisted of mixtures of small- and medium-sized elements that were also mixed in colour (white and grey). The summed areas of elements varied across arrays, but the area (and also luminance) of an S+ array was always greater than the area (and luminance) of an S− array. Three different sequence files encoded which arrays were to be presented on a given trial, and the left or right location of the S+. These sequence files were applied in mixed order across sessions.

Test sessions

On reinforced trials, the array configurations were the same as in training, but all the elements were changed to white. Intermixed with these trials were 48 non-reinforced trials in which fixed pairs of novel arrays were presented. For these test trials, the arrays within each trial were approximately matched in area by choosing appropriate combinations of small, medium, and large sizes of white elements. (E.g. for some of the 10 vs. 6 test trials, one array consisted of five small + four medium + one large-sized element, and the other of two medium + four large-sized elements. This pair of arrays differed in area by about 1%.) On average, the summed areas of paired arrays differed by about 3%. Where there was a disparity, the array with the larger area was the one with the greater numerosity on 32 out of 48 test trials, but had the smaller numerosity on the other 16 trials.

On test trials there were three subsets of numerosity pairs, with 15 vs. 9, 10 vs. 6, and 5 vs. 3 elements, respectively. For each subset, there were eight different pairs of arrays, and each pair was shown twice per session with left and right stimulus positions counterbalanced. Choices on test trials were scored as correct if FR8 pecks accumulated first to the larger numerosity.

Results

Two birds reached the training criterion of ≥80% correct choices in 12 sessions, one in 13, and the fourth bird in 17 sessions. For each bird, reinforcement probability was gradually reduced in the next five sessions. On the final training session with partial reinforcement the mean score was 91.2% correct (SD = 2.6).

In test sessions, the mean score on the reinforced trials across all birds and sessions was 93.1% correct (SD = 4.9). Test scores were analysed with a two-way, repeated-measures ANOVA, with numerosity combinations, and sessions as the main factors. There were no significant differences in mean scores across numerosity combinations ($F_{2,6} = 1.11$, $P = 0.39$), nor across sessions ($F_{4,12} = 1.29$, $P = 0.33$). Scores with the different numerosity combinations are shown in Fig. 1. These mean scores were all above chance level. However, an ANOVA that compared test session scores on reinforced trials with scores pooled across all non-reinforced trials showed that the birds performed significantly better on the reinforced trials ($F_{1,3} = 109.13$, $P = 0.002$).

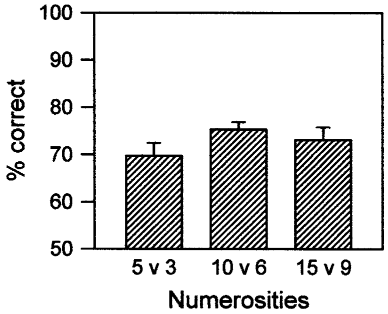

Fig. 1 Mean test scores with novel arrays of different absolute numerosities, but with the same disparity ratio. *Vertical bars* in all figures show standard errors

From trial-by-trial data, discrimination scores were cumulated across birds and sessions according to the array-pairs used on the test trials. These scores were correlated with the positive or negative area-disparity ratios for the respective arrays. The Pearson product–moment correlations were non-significant across all test trials ($r = 0.18$, $P = 0.4$, $n = 24$) as well as for each numerical subset of trials (10 vs. 6: $r = 0.33$, $P = 0.4$, $n = 8$; 15 vs. 9: $r = 0.67$, $P = 0.07$, $n = 8$; 5 vs. 3: $r = 0.02$, $P = 0.9$, $n = 8$).

Discussion

Although scores were poorer on test trials than on reinforced trials, all the birds showed significant transfer to the novel array pairs. While these arrays differed in the absolute numbers of items they contained, their disparity ratio was the same across the three test subsets. The similarity in scores across subsets, which complies with Weber's law, suggests that the birds discriminated the relative numerosity differences in the arrays.

Summed area in the arrays was not consistently related to numerosity in the novel test stimuli. (Since the stimuli were all white, overall array luminance also no longer indicated the "correct" stimulus on test trials.) Performance on test trials did not correlate with the residual differences in array area. But area and brightness differences had been confounded with numerosity differences on reinforced trials. So these additional cues might have contributed to the better scores on reinforced compared with non-reinforced trials in the test sessions. On the other hand, the mixtures of sizes that were used in test arrays were novel and included one or more large elements on most of the test trials. The addition of this new element size could have been a cue to the birds that a test trial would not be reinforced, and led to poorer performance on these trials. The role of numerical as well as non-numerical cues was examined more closely in the following experiments.

Experiment 2

This experiment was designed to test how pigeons would perform with novel array pairs in which the summed areas and luminances of the arrays were made virtually equal. If the birds had learned in the previous experiment to discriminate numerosity as the primary cue, they should still perform at an above-chance level on test trials in which the numbers of elements in the arrays were easily discriminable. But their performance should collapse on trials in which paired arrays had identical numerosities.

Method

The pigeons were retrained, in one to four sessions, to a criterion of $\geq$80% correct choices, followed by five sessions with decreasing reinforcement probability. Stimulus arrays were the same as those used for training in Experiment 1.

In test sessions, arrays on the reinforced trials were the same as those used in Experiment 1. On the non-reinforced test trials, all the arrays were novel. In subsets of 16 trials, the array numerosities were 10 vs. 6, 6 vs. 6, and 10 vs. 10. Within each subset, there were eight fixed pairs of arrays, and each pair was presented twice per session with left/right stimulus positions counterbalanced.

Three sizes of white elements (small, medium, and large) were used in test arrays, most of which included mixtures of sizes. Test stimuli were equated, within pairs, for area (and thus luminance) by choosing appropriate combinations of sizes. Areas were defined as equal if they differed by no more than 1%. Within each of the three test numerosity subsets there were four different combinations of size-mixtures that were used to produce arrays of equal areas. Thus different mixtures of element-sizes were used both within array-pairs and across trials.

On test trials with 10 vs. 6 items, choices were scored as correct if the 10-element array was chosen. Although the array pairs were equated in area as closely as possible, the 10-element array always had the slightly larger area (the residual difference of $\leq$1%). So on equal-numerosity test trials, choices were scored in two ways. One set of scores was of choices directed to the array in each pair with the residually larger area. If the birds were able to discriminate an area difference of $\leq$1%, then test scores should be at a similar, above chance level in spite of the array numerosities being indiscriminable. The other scores were of choices directed to the array presented on the left side of the screen. This was to check whether any birds adopted a position preference.

Results

The mean discrimination score on the final training session was 93.2% (SD = 5.5). The overall mean score on the reinforced trials in test sessions was still 93.2% correct (SD = 5.0).

Test data (percent correct choices on trials with unequal numerosities, and percent choices of arrays with the residually larger areas on the equal-numerosity trials) were analysed with a two-way, repeated-measures ANOVA. The main factors were numerosity combinations (10 vs. 6, 6 vs. 6, and 10 vs. 10) and sessions. There was a significant effect of numerosity ($F_{2,6} = 42.62$, $P < 0.001$; see Fig. 2), but not of sessions ($F_{4,12} = 0.10$, $P = 0.98$). Newmans–Keuls post hoc tests showed a significant difference when

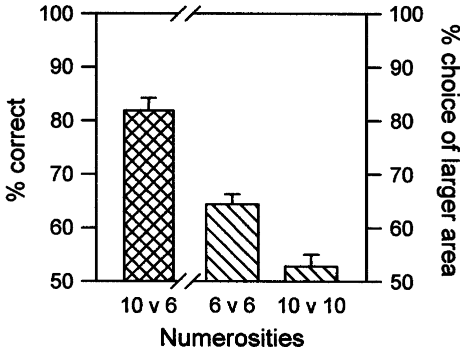

Fig. 2 Mean test scores for array pairs with unequal or equal numerosities. Stimulus areas for each pair differed by $\leq 1\%$. Percent correct choices of the greater numerosity for test values 10 vs. 6 are plotted against the *left axis*. Percent choices of the residually larger area for equal-numerosity pairs are plotted against the *right axis*

choices on 10 vs. 6 trials were compared with 6 vs. 6 trials ($P = 0.002$), and with 10 vs. 10 trials ($P < 0.001$). There was also a significant difference in the percent choices on 6 vs. 6 and 10 vs. 10 trials ($P = 0.01$). Mean choices of the arrays with the marginally larger area were barely above chance at 52.8% on 10 vs. 10 trials. The mean on 6 vs. 6 trials was higher at 64.4%. Binomial statistics on each bird's data across sessions showed that two birds' scores exceeded an above chance criterion (of 61%) on the 6 vs. 6 test trials. None of the other scores on the equal-numerosity trials were reliably above chance.

No bird showed a consistent side preference across sessions on the test trials with arrays of equal numerosity. Also, an ANOVA for all birds confirmed there was no significant difference in side-choices between the two types of equal-numerosity trials.

Discussion

Discrimination scores were much higher on test trials in which arrays differed in their numerosity than on the other two types of trials in which the numerosities were identical in paired arrays. This suggests that numerosity acted as the primary cue for the birds. This is supported by the chance level of performance seen when both arrays had the value of the training S+ (10 vs. 10). However, when both arrays had the S− value (6 vs. 6), two of the birds were above chance in their choices of the array with the marginally larger (i.e. by $\leq 1\%$) summed area. Although it is conceivable that birds integrated area (or else brightness) across the whole array, and that some birds could discriminate extremely small differences in area (or luminance), a further ad hoc assumption would have to be made. This integration would have to be restricted to relatively few (6) but not many (10) items. Although infants sometimes use area as a cue when the number of items is small (≤ 3), six does not usually count as a "small" number. So it is not entirely clear whether the birds used these small differences in integrated area or luminance as discriminative cues, or whether that interpretation would amount to a Type I error of falsely rejecting the null hypothesis with two birds' results. Since the results remain slightly ambiguous, some further experiments were conducted.

Experiment 3

The question of whether pigeons utilize the summed area of elements to discriminate arrays was the focus of the next experiment. In all the reinforced trials in the first two experiments, the area of the S+ array had been greater than that of the S−. In the next experiment, array areas were equated within stimulus pairs on training trials, but differed on some of the test trials. If area is a cue that birds discriminate preferentially, or at least one that contributes towards their choices, their discrimination scores should be higher if the S+ area is much greater than the S− area, compared with trials in which the areas of S+ and S− are virtually equal. This experiment tested that prediction.

Numerosity combinations varied across test trials as a further check for the pigeons' use of numerosity as the primary cue. Weber's law predicts that the birds' discrimination accuracy should decrease as the numerosity disparity ratio decreases.

Method

During training sessions, arrays consisted of white elements in small, medium, and large sizes. The summed areas of arrays were equated within pairs to $\leq 1\%$ difference in area. Three different combinations of sizes were used to achieve these area matches. One or both arrays in a pair contained mixtures of sizes. There were 24 fixed pairs of arrays. Each pair was shown twice per session with left and right array locations counterbalanced across trials. The quasi-random order of trials per session was controlled by four different sequence files.

The stimuli on reinforced trials of the test sessions were the same as those in the training sessions. Intermixed with these trials were novel test trials in three numerosity combinations: 8 vs. 4, 10 vs. 6, and 7 vs. 5, with 16 trials at each pair of values. On these test trials, half the novel array pairs were equated in summed area within pairs. In each of the three numerosity subsets, three different combinations of sizes were used to create these equal areas. (More details are given in Supplementary material about the combinations of sizes that were used to equate area at each numerosity combination.) On the other test trials, element sizes were the same in both arrays of a pair, but varied across trials (i.e. both arrays consisted of small, medium, or else large sized elements). This meant that the greater numerosity also had the greater area, with an exact match between the disparity ratios for numerosity and summed area. Each test array-pair was used twice in a session, with array locations counterbalanced across trials.

Choices of the greater numerosity were scored as correct on the test trials. Reinforced and non-reinforced trials occurred in quasi-random order that changed across sessions according to codes in four different sequence files.

Results

Three birds reached the training criterion within four sessions then reinforcement probability was reduced to 0.5 in a further five sessions. One bird reached criterion in six sessions, but it took another ten sessions to fully reduce reinforcement probability. Some sessions were repeated when its performance fell below criterion. On the final training session, the mean discrimination score across birds was 93.8% correct (SD = 3.0).

The overall mean score on reinforced trials in the test sessions was 96.0% correct (SD = 5.0). Test scores were analysed with a three-way, repeated-measures ANOVA, with numerosity-combinations, relative area (equal vs. unequal) and sessions as factors. There were significant main effects for numerosity ($F_{2,6} = 42.31$, $P < 0.001$), and for area ($F_{1,3} = 23.71$, $P = 0.02$), but not for sessions ($F_{4,12} = 0.57$, $P = 0.69$). There were no significant interactions. Newman–Keuls post hoc tests showed significant differences amongst the scores at all the numerosity combinations. Discrimination of 8 vs. 4 was significantly better than that of 10 vs. 6 ($P = 0.02$) and of 7 vs. 5 ($P < 0.001$). Also, discrimination scores with 10 vs. 6 were significantly better ($P = 0.001$) than with 7 vs. 5. The effects of numerosity and area are illustrated in Fig. 3.

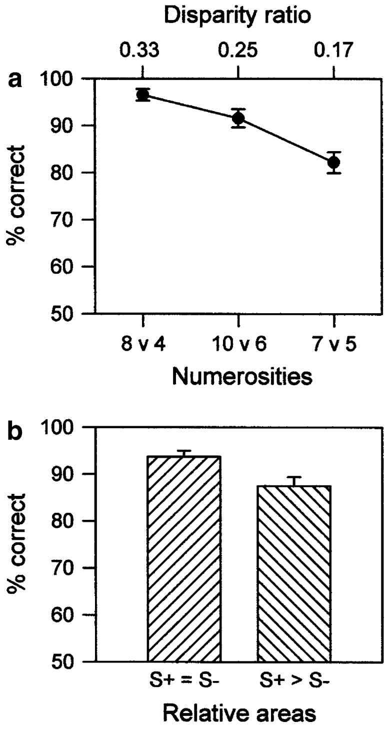

Fig. 3 Mean test accuracy as a function of **a**: disparity ratio with novel numerosity combinations, and **b**: relative summed areas of the greater (S+) and smaller (S−) numerosities. Disparity ratio was calculated as the difference divided by the sum of paired numerosities

Discussion

As predicted by Weber's law, the birds' discrimination accuracy on test trials decreased significantly as the numerical disparity ratio became smaller. Although there was a significant effect of array areas on performance, it was in the opposite direction to that predicted if the birds had used the summed area of array components as a discriminative cue. Discrimination scores were higher on trials in which array areas were equated than on trials in which the S+ area was greater than the S− area. If, on the other hand, the pigeons were not using integrated area (or correlated luminance) as a property to distinguish the S+ from the S− array, it was puzzling that differences in relative area had any effect on the birds' performance. Since the effect of relative array areas on the birds' performance was contradictory, another experiment was performed to check this finding using different stimulus sets.

Experiment 4

This experiment tested the pigeons with new stimulus pairs that differed in relative numerosity, and were either equal or unequal in the summed areas of the arrays. But, compared with the previous tests, the area relationships were reversed on those test trials in which the array areas were unequal: the area of the array with the larger numerosity was now less than the area of the smaller numerosity array. In addition, the relative inequality in area was held approximately constant across the different numerosity combinations that were tested. If the effect of relative area seen in the last experiment was reliable, then in this experiment test results should still reveal a significant difference in scores on trials with equal vs. unequal array areas. However, this time performance should be better on those trials with unequal rather than equal array areas. If the birds still discriminate numerosities, their scores on test trials should once again depend on the numerosity disparity ratios of the array pairs.

Method

The birds were retrained, on partial reinforcement, for four sessions. Training stimuli were the same as those used in Experiment 3.

On reinforced trials of test sessions the stimuli were the same as in training sessions. On intermixed test trials the stimuli consisted of fixed pairs of novel arrays in the numerosity combinations 10 vs. 6, 7 vs. 5, and 6 vs. 5. For half of these pairs, arrays were equated (within a 1% difference) in area by selecting appropriate mixtures of the three sizes of white elements. For the other pairs, combinations of element-sizes were chosen that gave an area disparity ratio of approximately 0.23 across all these stimulus pairs. (More details are provided in the Supplementary material about the size-combinations used to create these equal- and unequal-area test stimuli.) The array with the smaller numerosity was the one with the larger summed area. For each numerosity combination, there were four different array pairs with equal areas within each pair, and four pairs with unequal areas. Each pair was presented twice per session, with left and right stimulus positions counterbalanced. Trial order was controlled by four sequence files that were changed across sessions. Choices of the array with the greater numerosity were scored as correct on all trials.

Results

The mean performance on the final training session was 99.0% correct (SD = 1.0). The overall mean score on reinforced trials during the test sessions was 97.6% correct (SD = 1.9).

Data on test trials were analysed with a three-way, repeated-measures ANOVA with numerosity combinations, relative area (equal vs. unequal) and sessions as factors. There were significant effects for the main factors of numerosity ($F_{2,6} = 117.75$, $P < 0.0001$), and relative area ($F_{1,3} = 61.36$, $P = 0.004$), but not for sessions ($F_{4,12} = 1.05$, $P = 0.42$). These results are illustrated in Fig. 4. There was also a significant numerosity × area interaction ($F_{2,6} = 25.94$, $P = 0.001$). As shown in Fig. 4c, this was mainly due to deviations in the scores on trials with 6 vs. 5 arrays. When array areas were unequal, these scores were almost as high as the ones with the other numerosity combinations. Scores declined only on trials with equal array areas at numerosities 6 vs. 5.

Discussion

Although there was an overall effect in test trials of relative numerosity on discrimination scores, this was mainly due to the poorer scores that birds showed on some of the 6 vs. 5 trials. Otherwise, their performance was almost at ceiling with the other numerosity combinations. The significant differences between trials with equal vs. unequal array areas complement the findings of Experiment 3 since the mean scores for this relative area factor were now reversed, as predicted. This time, scores were higher when the S+ array (with the greater numerosity) had the *smaller* area. Throughout the birds' training in Experiments 1 and 2, the larger area (as well as the greater numerosity) had been correlated with reward on the training trials. So, once again, the differential effects of relative area are in the opposite direction to what would be expected if birds were summing the area of all the elements in an array, and then using this as a cue for choosing a "correct" stimulus. The results of this experiment and the previous one thus suggest

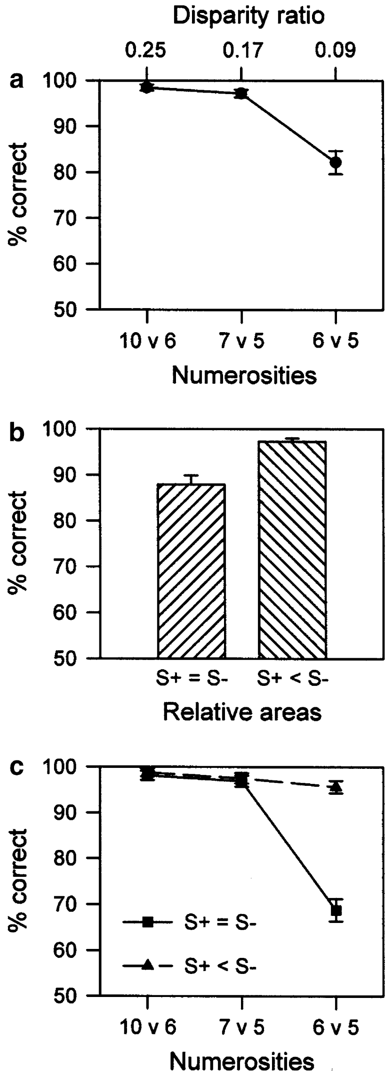

Fig. 4 Mean test accuracy as a function of **a**: numerical disparity ratio, and **b**: relative summed areas of the S+ and S− numerosities. **c**: Interaction between the effects of relative numerosity and relative array areas on test accuracy

that the pigeons did not really use area (or correlated brightness) as a cue, but rather some other stimulus property that was altered when area was manipulated. One possibility was that the birds were focusing on cues associated with the individual elements that comprised the arrays. This was investigated in the next experiment.

Experiment 5

Equating the areas of arrays that differ in numerosity unavoidably introduces another type of confound that could have been used by the birds as a discriminative cue. In spite of using a variety of size-mixtures in the elements that made up these equal-area arrays, small elements occurred more frequently overall in the S+ arrays with the greater numerosities. Conversely, large elements occurred more frequently amongst the S− arrays with fewer items in them. This unequal distribution of element sizes had in fact been exaggerated in creating the unequal-area test arrays that were used in Experiment 4.

So the aim of this experiment was to test whether pigeons focus on local stimulus features, namely the sizes of individual elements, and use these as discriminative cues. Two numerosity combinations were employed on test trials to check whether birds' performance still depended on relative numerosity differences. If it did, discrimination scores should be higher overall on those trials in which the numerosity disparity ratio was greater. In addition, different combinations of element sizes were created with the expectation that if the pigeons were using size cues, their scores should be highest if small element-size coincided with larger numerosity, and large element-size coincided with smaller numerosity. Conversely, test scores should be lowest with the opposite relationship between numerosity and element size.

Method

The birds were retrained for four sessions with the same training arrays as in Experiments 3 and 4.

On reinforced trials of test sessions, the array pairs were the same as those used in training. On intermixed test trials, the array pairs were novel. The numerosity combinations were 10 vs. 6 and 6 vs. 5. For each numerosity combination, there were three types of array pairs with respect to their element sizes. In one subset, the array pairs were equal in area and consisted of mixed element sizes. (The size-mixtures were the same as those used to create equal-area test stimuli for 10 vs. 6 and 6 vs. 5 in Experiment 4, but new array configurations were generated.) In another subset, arrays with the greater numerosities consisted entirely of small elements, whereas the smaller numerosity arrays consisted of large elements. In the remaining subset, arrays with the greater numerosities were made up of large elements, and arrays with smaller numerosities consisted of small elements. Trials were scored as correct if the array with the greater numerosity was chosen.

Results

The mean score on the final training session was 99.0% correct (SD = 1.0). In test sessions, the mean score on reinforced trials was 97.3% (SD = 2.6).

Test data were analysed with a three-way, repeated-measures ANOVA, with numerosity combinations, array-types, and sessions as factors. The main effects of numerosity ($F_{1,3} = 47.09$, $P = 0.006$) and array-type ($F_{2,6} = 50.93$, $P < 0.001$) were significant. These results are illustrated in Fig. 5. The effect of sessions was not significant ($F_{4,12} = 0.45$, $P = 0.77$). The interaction between numerosity combinations and array-types was also significant ($F_{2,6} = 5.50$, $P = 0.04$). This interaction showed that, while scores were better across all array-types when the numerosities were 10 vs. 6 rather than 6 vs. 5, the difference in scores between numerosity combinations was minimal when the larger numerosity consisted of small elements and the smaller numerosity consisted of large elements (see Fig. 5c).

Discussion

The birds' scores indicated that they were still discriminating relative numerosity since their performance was better overall when they had to choose between 10 vs. 6 items compared with 6 vs. 5. However, the sizes of the individual elements in the paired arrays obviously had a major effect on the birds' choices. Throughout training, choices to the greater numerosity were counted as correct and were rewarded. But, as predicted, the birds were more likely to choose an array if it consisted entirely of small elements, and to reject one that consisted of large elements. The highest test scores for both numerosity pairs were obtained when the greater numerosity coincided with small element size. When greater numerosity corresponded with large element size, the likelihood of the birds' selecting the greater numerosity was reduced when the relative disparity in numerosity was fairly large (10 vs. 6), and thus easier to discriminate. When the numerosity difference was very small (6 vs. 5), their discrimination scores fell below chance. When array areas were equated, the difference in the distribution of element sizes was not so pronounced, although small elements were still more frequent in the larger numerosity, and large ones in the smaller numerosity. Thus, the dual cues of greater numerosity and more of the small elements in the "correct" stimuli produced high

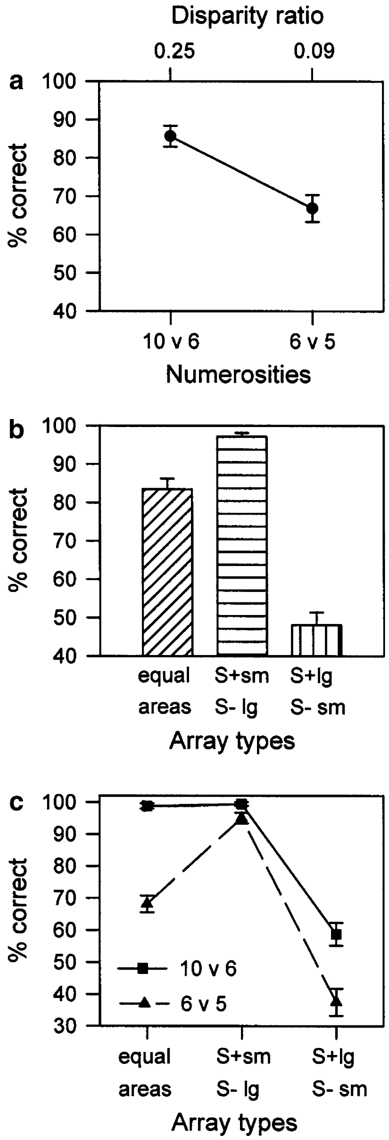

Fig. 5 Mean test accuracy as a function of **a**: numerical disparity ratio, and **b**: array types. Equal area arrays consisted of mixed element-sizes. Other types had all small (*sm*) or all large (*lg*) sized elements in the array with the greater (S+) or smaller (S−) numerosity. **c**: Interaction between the effects of relative numerosity and types of arrays on test accuracy

scores, at ceiling, for arrays with 10 vs. 6 elements. With the more difficult numerosity discrimination (6 vs. 5), the birds' scores with the mixed-size arrays were intermediate between those obtained with the other size combinations.

Experiment 6

The next experiment was to check that the birds were still primarily discriminating relative numerosity differences. The pigeons were retrained with novel numerosity combinations. The stimuli were not equated for overall area, or brightness, since these factors did not appear to have had a major influence on the birds' choices. Some array pairs were composed of elements in mixed sizes. The size mixtures varied across trials, and were chosen to avoid having any consistent skew in the size distributions. On other trials, there were no differential size cues. The elements within a stimulus pair were identical in size, but sizes varied across trials, similar to the stimulus conditions on some of the test trials in Experiment 3. So the question in this experiment was whether the pigeons would still perform well when there were no longer consistent differences in item sizes as potential cues.

Method

There were 10 sessions of 48 training trials, with 12 trials at each of 4 numerosity combinations: 12 vs. 8, 6 vs. 4, 8 vs. 6, and 4 vs. 3. For each numerosity combination, the array pairs on half the trials had novel mixtures of element sizes. (Details of the size-mixtures are given in the Supplementary material.) On the other half, the element sizes within an array pair were identical (all small, all medium-sized, or all large). The stimulus elements were white. All trials terminated in reward or timeout (with correction trials).

Results

The data were analysed with a three-way, repeated-measures ANOVA with numerosity combinations, array-type (mixed vs. homogeneous item-sizes) and sessions as main factors. Only numerosity combinations yielded a significant effect ($F_{3,9} = 7.64$, $P = 0.008$). The mean percent correct scores (±SEs) at each numerosity combination were: 81.67% (±1.71) for 12 vs. 8, 83.33% (±1.96) for 6 vs. 4, 78.75% (±1.72) for 8 vs. 6, and 71.04% (±1.83) for 4 vs. 3. Newman–Keuls post hoc tests showed that the lowest discrimination score for numerosities 4 vs. 3 differed significantly from the score with each of the other combinations ($P = 0.01$ for 12 vs. 8; $P = 0.008$ for 6 vs. 4; $P = 0.02$ for 8 vs. 6). There were no significant differences amongst the other scores.

Discussion

When item size was no longer a reliable cue, the birds still performed the discrimination task at a level well above chance. These scores are probably a more accurate reflection of the birds' ability to discriminate relative numerosity differences per se since ceiling effects were eliminated. The data are only partially in accord with Weber's law, however. There were no differences in performance with combinations 12 vs. 8 and 6 vs. 4 which have the same disparity ratio of 0.2. While the combination 4 vs. 3 yielded the lowest score overall, the score with the 8 vs. 6 stimulus pairs was slightly higher, although both combinations have the same disparity ratio of 0.14.

Experiment 7

The final experiment examined once more the role of element size as a cue in the birds' discrimination. The experiment also reassessed whether the summed area (or luminance) of stimulus elements might serve as a discriminative cue in the absence of numerosity differences since some of the test results in Experiment 2 were ambiguous with respect to area as a contributory cue to pigeons' discrimination of numerosity.

In Experiment 5 we suggested that pigeons readily learn about differences in the sizes of array items, and use this information as a supplementary cue when discriminating numerosity. This focus on local features would tally with the results of other studies that have shown a tendency of pigeons to give precedence to local over global aspects of visual stimuli when they peck at them (Cavoto and Cook 2001; Fremouw et al. 2002; but see Goto et al. 2004). However, an alternative possibility is suggested by Ginsburg and Nicholls' (1988) finding that adult humans' estimates of the number of dots in arrays are influenced by dot size. They reported an inverse relationship between numerosity estimates and item sizes, i.e. for a given numerosity value, their subjects perceived numerosity to be greater when the dots were uniformly smaller than when they were larger. They attributed this finding to Gestalt effects, rather than to local processing effects.

To test whether the pigeons had been focusing on the properties of individual items, a subset of test trials had stimulus pairs consisting of just a single item. The size of the element differed between these stimuli. If the birds had been using the size of local features as a cue, they should choose the smaller over the larger single element. But Experiment 6 showed that numerosity was still the predominant cue for the birds, and their experience in this study had been with multi-item arrays. Presentation of stimuli consisting of single elements might make these novel trials readily discriminable from the other ones and, in the absence of reinforcement, lead to rapid extinction. So in another subset of test trials arrays were equal in numerosity but both consisted of eight items, the numerosity midway between the S+ and S− training numerosities. On these 8 vs. 8 trials, element size also differed between arrays. If the birds utilized item size as a cue, they should choose the small-item array. If they used summed area as a cue, then, in the absence of numerosity differences, they should choose the large-item array.

Method

Subjects

After completion of Experiment 6, one of the birds died, so the remaining three pigeons served as subjects in this experiment. One of the birds had been used in an interim study involving a matching-to-sample procedure with colour stimuli (see Sect. "Discussion") before Experiment 7 began.

Procedure

The birds were retrained with the same stimuli used in Experiments 3–5. Training lasted for 12–16 sessions, before the test sessions were run. On reinforced trials in these test sessions, the stimuli were the same as those used in training. On non-reinforced test trials, one subset of stimuli consisted of 10 vs. 6 items, with the summed areas of the arrays equated within pairs. (Element size-mixtures were the same as those used for 10 vs. 6 test stimuli in Experiment 4.) Another subset consisted of one small item vs. one large item. The third subset consisted of eight small vs. eight large items. The average size of arrays with eight small items was 13.4 mm × 16.6 mm (or 222.4 mm^2) and the average size of the large-item arrays was 17.4 mm × 15.8 mm (274.9 mm^2).

Results

The mean discrimination score on the final training session was 93.1% (SD = 1.0). The overall mean score on the reinforced trials in test sessions was 91.8% (SD = 6.6).

The test data (percent correct choices on unequal numerosity trials, and percent choices of arrays with smaller elements on the two types of equal-numerosity trials) were analysed with a two-way, repeated-measures ANOVA. There was no significant main effect of trial-type ($F_{2,4} = 2.88$, $P = 0.17$) but a significant effect of sessions ($F_{4,8} = 8.89$, $P = 0.005$). This latter effect was mainly due to an increase across sessions in the mean choices of the small-item arrays on 8 vs. 8 trials. However, there were striking differences in performance across birds on 1 vs. 1 trials, as shown in Table 1. One bird (182) predominantly chose the small item on these trials. A dependent *t*-test of this bird's scores across sessions on these 1 vs. 1 trials showed that choice of the small element was significantly different from chance level. Choices of the other two birds on these trials did not differ from chance. Dependent *t*-tests on the scores for the other subsets of test trials showed that for each bird the scores were significantly different from chance (see Table 1).

Table 1 Test results for individual birds in Experiment 7

Bird	Test numerosities								
	10 vs. 6			8 vs. 8			1 vs. 1		
	Score	*t*	*P*	Score	*t*	*P*	Score	*t*	*P*
182	87.5	7.75	0.001	83.8	9.00	<0.001	90	12.55	<0.001
14	95	19.24	<0.001	82.5	4.19	0.01	48.8	0.41	0.70
B7045	90	16.00	<0.001	88.8	8.44	0.001	51.3	1.00	0.37

Scores are % correct choice of greater numerosity for 10 vs. 6 and % choice of small-element array for 8 vs. 8 and 1 vs. 1. Results of *t*-tests, for $N = 5$, $df = 4$, compared scores against chance

Discussion

One bird (182) completed the five test sessions quite rapidly, and on the equal-numerosity trials showed a strong preference for the arrays consisting of small items, irrespective of whether these stimuli contained just one element or eight. In the absence of numerosity differences, this bird showed an increasing tendency across sessions to choose the small-item arrays. This tendency was greater on trials with 1 vs. 1 than with 8 vs. 8 items.

The other two birds behaved differently. There was some increase in preference for small-item arrays on 8 vs. 8 trials. But choices on 1 vs. 1 trials were at chance. Also, intermittent observations of their responding via a webcam indicated that they were often slower to initiate or complete pecking on 1 vs. 1 trials. For one bird (B7045), sessions were aborted four times in all (and the partial data discarded), when it failed to complete a test session within 3 h or more. In each case, the bird had stopped responding on a 1 vs. 1 trial. So we attribute the chance scores of these two birds to extinction effects with stimuli they perceived to be novel. The bird (182) that responded steadily on these trials had received intervening training in a matching-to-sample procedure in the same apparatus. In that procedure each of the stimuli consisted of a single filled circle, 85 mm in diameter and in various colours. Since those stimuli did not resemble the arrays used here, no carry-over effects between experiments had been expected. But perhaps that bird's recent experience of being reinforced for pecking a single item led to some resistance to extinction with unitary stimuli in this test.

While the results of this bird on the 1 vs. 1 trials suggest that pigeons do indeed focus on local features of individual elements, the chance performance of the other birds does not allow us to totally exclude a global effect of uniform item size on the perception of numerosity. An effect of the type that Ginsburg and Nicholls (1988) reported would still be compatible with the results in the 8 vs. 8 trials.

It should be noted, though, that Ginsburg and Nicholls' results contrast with those of other studies in which item size per se had no effect on numerosity judgements. In experiments with children who had to judge which array had more dots in it Barth et al. (2005) found that, although dot size (amongst other factors) was negatively correlated with judgements of greater numerosity, it was actually the extent of the array (size of a virtual rectangle enclosing the array) that acted as a supplemental cue. In other experiments with children, Barth et al. (2006) manipulated array size as well as dot size in comparison arrays and again reported that children more heavily weighted array size, rather than dot size, in making judgements about the relative numerosities of dots in array pairs. In a study of adaptation effects on perceived numerosity Burr and Ross (2008) included a control condition in which the size of dots differed between adaptor and test arrays. Dot size had no effect on the shift in perceived numerosity induced by the adapting array. Furthermore, Tomonaga (2008) found that chimpanzees were more accurate at choosing the array with the greater numerosity when the more numerous array consisted of larger dots, and the less numerous of smaller dots than vice versa. This is the opposite of Ginsburg and Nicholls' (1988) report that perceived numerosity decreases with increased dot size. Instead Tomonaga concluded that the summed area of dots, as well as their greater density, promotes accuracy in the chimpanzees' choices of greater numerosity.

While Ginsburg and Nicholls' own results were consistent, their explanation for the effects of dot size on numerosity estimation is puzzling. They interpreted one of Koffka's (1935) Gestalt principles to mean that smaller items will make "better" figures than larger ones, and so, by extension, smaller items will be perceived as more numerous than larger ones. However, their citation of Koffka actually referred to the effect of the relative size of contrasting parts of a stimulus on figure-ground perception. Koffka gave the example of alternating black and white sectors of a circle. If the white sectors are narrower than the black ones, we perceive the white parts as a unified figure (like a windmill) on a black background. If the relative sizes of the black and white portions are reversed, we perceive a black figure on a white background. But it is difficult to see why it follows from this that varying the sizes of black dots upon a white background, as Ginsburg and Nicholls did, should alter their perceived numerosity. Their procedural description suggests that array extent was fairly constant so their results are unlikely to be due to differences in array sizes. Changes in dot sizes would have altered the summed areas of dots and their densities. But if these parameters had influenced perceived numerosity, the trend in Ginsburg and Nicholls' data should have been in the opposite direction, so the causal mechanisms underlying their results remain unclear.

Although the preference shown by pigeons in this experiment for the small-item arrays in 8 vs. 8 trials cannot preclude effects on numerosity estimation of the type described by Ginsberg and Nicholls it does strengthen the conclusion that summed area of array items had little effect as a cue for the pigeons. If summed area had been important, the birds should have preferred the array containing the larger items.

In addition, the birds' preference for the smaller-item array on these trials makes it unlikely they were using array extent as a cue (cf. Barth et al. 2005, 2006). Array density can affect pigeons' performance in numerosity discrimination tasks (Emmerton 1998), so we avoided having much variation in inter-item spacing across stimuli in the current experiments. However, the corollary of controlling density is that, on average, the S+ stimuli (with the larger numerosity) also had the greater extent in training. On 8 vs. 8 trials, the arrays with larger items had the greater extent, but were chosen less frequently. Thus, the birds' discrimination performance did not seem to depend on overall array size (cf. Emmerton and Renner 2006).

General discussion

Like other species that demonstrate numerical competence, pigeons are influenced by some non-numerical properties of visual stimuli when they discriminate differences in their numerosity. However, the cues that affect choices in these tasks apparently differ between pigeons and primates. The property we focused on initially was summed area since that has been shown to act as a supplemental cue for quantity in several non-avian species. But in spite of greater area (and brightness) being correlated with reward in early training (Experiments 1 and 2), pigeons showed little evidence that they used area (or overall brightness) of the arrays, either as the primary cue, or as an additional cue to numerosity in choosing the correct stimulus. Also, some of the results in the final experiment confirmed findings in a previous one (Emmerton and Renner 2006) that the extent, or overall size of an array does not serve as a prominent cue in pigeons' numerosity discrimination, whereas it can influence choices in humans (e.g. Barth et al. 2006).

Equating array areas on training trials, starting in Experiments 3, introduced asymmetries between the S+ and S− array sets in the distributions of different element sizes. Most of the training arrays consisted of mixed sizes of elements. Small items occurred more frequently across trials in the larger numerosity, and large items were more frequent in the smaller numerosity. However, there was some variability in this asymmetry in size distribution. (Numerosity was not simply inversely correlated with uniform size, so item size was not a continuous variable.) Nevertheless, the birds quickly picked up on these differential cues to reward.

This has some practical implications for the types of stimuli that should be used when pigeons are trained in numerosity discrimination tasks. While the training stimuli must include different numbers of elements, our results indicate that the size of elements should vary to discourage discrimination based on element size. Subsequent tests of the birds' relative numerosity discrimination should then control for the summed area of elements in novel arrays to check that the birds are not using area as a discriminative cue.

The birds' focus in the current experiments on the properties of component items rather than on overall characteristics of the arrays is reminiscent of the local rather than global precedence effect which, as mentioned earlier, has been reported before when pigeons process complex, multi-item visual stimuli.

We implied in the Introduction that continuous variables, such as summed area or brightness that must be integrated from the whole stimulus array, are global percepts. The perceptual mechanisms underlying the behavioural ability to discriminate summed areas of discrete items have not, as far as we know, been directly investigated. But there is evidence for global processing when humans look at other types of dot-arrays. For instance, when stimuli consist of dot-pairs of different contrast (black vs. white) that are oriented in particular directions, the structure that is perceived by humans in such Glass patterns has been attributed to parallel processing of low-level information from the whole display that is then integrated at a later stage in the visual system (Badcock et al. 2005). In this example, the perceptual data were modelled by positing that all white dot-pairs initially excite an array of filters whose output is then combined spatially, and input to global detectors sensitive only to the white stimulus components. (Similarly, a separate channel integrates input from all the black stimulus components.) Thus models of visual function in humans suggest, in principle, that distributed properties, such as brightness or area, of an array of discrete elements can be integrated to give a global percept.

Parallel uptake of visual information has also been postulated as the initial stage in an influential model of numerical processing (Dehaene and Changeux 1993). Behavioural as well as neural response measures from rhesus monkeys lend support for such parallel processing of numerosity in visual displays (Nieder and Miller 2004). By contrast, to account for previous data, Emmerton (1998) suggested that pigeons may “scan” the elements of numerosity arrays, rather than processing the whole array in parallel. The pigeons’ bias to use cues about individual items in mixed-size arrays, rather than a continuous cue such as summed area is compatible with this idea.

Pigeons and primates differ in the way they view stimulus displays of the type used here. Old World primates (including humans) usually view such stimuli, as they would other extended surfaces, at a distance of 40 cm (e.g. Tomonaga 2008) or more. At normal viewing distances, humans (own observations) can easily see a whole array (or indeed a pair of arrays) “at a glance.” But as confirmed by our observations over a web-cam, when pigeons have to peck at the stimuli, they view the arrays from a distance of a few centimetres and use frontal vision. Specialized regions of both retinae, the red fields, subserve near-field, frontal vision in pigeons, whereas they view more distance objects laterally with the yellow retinal field (see Emmerton 1983). Functionally, the two fields differ in several ways (e.g. spectral sensitivity: Remy and Emmerton 1989; motion sensitivity: Martinoya et al. 1983). The two fields project differentially within the visual system (Remy and Güntürkün 1991), and, to some extent, information from the separate visual fields is processed independently (Remy and Emmerton 1991). Frontal vision in pigeons is thought to be specialized for locating and pecking small seeds (Goodale 1983). It is notable that experiments providing the strongest evidence for a local precedence effect in pigeons (Cavoto and Cook 2001) presented stimuli, as we did, on a computer screen at which the birds pecked directly. We suggest that the tendency of pigeons in our experiments to rapidly learn about the sizes of individual array components is an outcome of this species’ specialization in its near-field visual function.

In combination with information about the number of items, local cues about stimulus elements seemed to be more salient for pigeons than continuous, global cues under the experimental conditions used here. It remains to be seen whether factors such as summed area or overall stimulus luminance are cues that pigeons could utilize if they view stimuli from a greater distance. But in operant tasks that involve pecking directly at the stimuli, properties that are associated with local features should be carefully controlled in numerosity discrimination tasks with this species.

While pigeons, like other species, can discriminate relative differences in numerosity, it is not clear whether the processing mechanisms underlying this ability are the same across species. The current results suggest there may be differences at the perceptual input stage. Dehaene and Changeux’s (1993) neural network model assumes parallel uptake of the whole stimulus array at the input stage. However, this is not a critical feature for later processing stages in their model, which also allows for input of sequentially presented stimuli. Other models, that largely agree with Dehaene and Changeux’s approach, are neutral about the input stage (Verguts and Fias 2004).

The procedures used here, in which the birds had to peck at stimuli to obtain food rewards, tap into some of the birds’ natural foraging behaviours. An ability to discriminate differences in numerosity is probably adaptive to birds when they are faced with choices amongst food patches in their natural environment. For granivorous birds, it may also be adaptive to find out about local features of the items themselves—especially ones they are going to peck at. Pigeons frequently have to decide whether small objects on the ground are edible or not. Differences in item size are one aspect of local features that seem to be especially salient for them since they easily utilized this information as a cue for reward in this artificial foraging situation in the lab. They were not simply biased to choose large items (cf. Boysen et al. 2001; Olthof and Roberts 2000). Instead, the birds used information about the different sizes of elements in a quite flexible way to help them make decisions about which “patches” yielded reward.

Acknowledgements The Department of Psychological Sciences at Purdue University is accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International. All aspects of the study complied with the various federal regulations governing use of laboratory animals in the United States. We thank Dr Elizabeth Spelke for her comments that prompted the final experiment, and other anonymous referees for their critique of earlier versions of the manuscript.

References

Agrillo C, Dadda M, Serena G, Bisazza A (2008) Do fish count? Spontaneous discrimination of quantity in female mosquitofish. Anim Cogn 11:495–503

Allik J, Tuulmets T (1991) Occupancy model of perceived numerosity. Percept Psychophys 49:303–314

Badcock DR, Clifford CWG, Khuu SK (2005) Interactions between luminance and contrast signals in global form detection. Vision Res 45:881–889

Barth H, La Mont K, Lipton J, Spelke ES (2005) Abstract number and arithmetic in preschool children. Proc Natl Acad Sci USA 102:14116–14121

Barth H, La Mont K, Lipton J, Dehaene S, Kanwisher N, Spelke E (2006) Non-symbolic arithmetic in adults and young children. Cognition 98:199–222

Boysen ST (1993) Counting in chimpanzees: Nonhuman principles and emergent properties of number. In: Boysen ST, Capaldi EJ (eds) The development of numerical competence: animal and human models. Erlbaum, Hillsdale, pp 39–59

Boysen ST, Berntson GG, Mukobi KL (2001) Size matters: Impact of item size and quantity on array choice by chimpanzees (*Pan troglodytes*). J Comp Psychol 115:106–110

Burr D, Ross J (2008) A visual sense of number. Curr Biol 18: 425–428

Cantlon JF, Brannon EM (2007) How much does number matter to a monkey (*Macaca mulatta*)? J Exp Psychol Anim Behav Process 33:32–41

Capaldi EJ, Miller DJ (1988) Counting in rats: its functional significance and the independent processes that constitute it. J Exp Psychol Anim Behav Process 14:3–17

Cavoto KK, Cook RG (2001) Cognitive precedence for local information in hierarchical stimulus processing by pigeons. J Exp Psychol Anim Behav Process 27:3–16

Cerella J (1980) The pigeon's analysis of pictures. Pattern Recognit 12:1–6

Clearfield M, Mix K (1999) Number versus contour length in infants' discrimination of small visual sets. Psychol Sci 10:408–411

Davis H (1993) Numerical competence in animals: life beyond Clever Hans. In: Boysen ST, Capaldi EJ (eds) The development of numerical competence: animal and human models. Erlbaum, Hillsdale, pp 109–125

Davis H, Memmott J (1982) Counting behavior in animals: a critical evaluation. Psychol Bull 92:547–571

Davis H, Pérusse R (1988) Numerical competence in animals: definitional issues, current evidence, and a new research agenda. Behav Brain Sci 11:561–651

Dehaene S, Changeux J-P (1993) Development of elementary numerical abilities: a neuronal model. J Cogn Neurosci 5: 390–407

Emmerton J (1983) Vision. In: Abs M (ed) Physiology and behaviour of the pigeon. Academic, London, pp 245–266

Emmerton J (1998) Numerosity differences and effects of stimulus density on pigeons' discrimination performance. Anim Learn Behav 26:243–256

Emmerton J, Renner JC (2006) Scalar effects in the visual discrimination of numerosity by pigeons. Learn Behav 34:176–192

Emmerton J, Lohmann A, Niemann J (1997) Pigeons' serial ordering of numerosity with visual arrays. Anim Learn Behav 25:234–244

Feigenson L (2005) A double-dissociation in infants' representations of object arrays. Cognition 95:B37–B48

Feigenson L, Carey S, Spelke E (2002) Infants' discrimination of number vs. continuous extent. Cogn Psychol 44:33–66

Fellows BJ (1967) Chance stimulus sequences for discrimination tasks. Psychol Bull 67:87–92

Fetterman JG (1993) Numerosity discrimination: both time and number matter. J Exp Psychol Anim Behav Process 19:149–164

Fremouw T, Herbranson WT, Shimp CP (2002) Dynamic shifts of pigeon local/global attention. Anim Cogn 5:233–243

Ginsburg N, Nicholls A (1988) Perceived numerosity as a function of item size. Percept Mot Skills 67:656–658

Goodale MA (1983) Visually guided pecking in the pigeon (*Columba livia*). Brain Behav Evol 22:22–41

Goto K, Wills AJ, Lea SEG (2004) Global-feature classification can be acquired more rapidly than local-feature classification in both humans and pigeons. Anim Cogn 7:109–113

Honig WK, Stewart KE (1989) Discrimination of relative numerosity by pigeons. Anim Learn Behav 17:134–146

Jordan KE, Brannon EM (2006) Weber's law influences numerical representations in rhesus macaques (*Macaca mulatta*). Anim Cogn 9:159–172

Koffka K (1935) Principles of Gestalt psychology. Harcourt Brace & Co, New York

Lyon BE (2003) Egg recognition and counting reduce costs of avian conspecific brood parasitism. Nature 422:495–499

Macko KA, Hodos W (1985) Near point of accommodation in pigeons. Vision Res 25:1529–1530

Maldonado PE, Maturana HR, Varela FJ (1988) Frontal and lateral visual system in birds: frontal and lateral gaze. Brain Behav Evol 32:57–62

Martinoya C, Rivaud S, Bloch S (1983) Comparing frontal and lateral viewing in the pigeon: II. Velocity thresholds for movement discrimination. Behav Brain Res 8:375–385

Matsuzawa T (1985) Use of numbers by a chimpanzee. Nature 315:57–59

Meck WH, Church RM (1983) A mode control model of counting and timing processes. J Exp Psychol Anim Behav Process 9:320–334

Navon D (1977) Forest before trees: the precedence of global features in visual perception. Cogn Psychol 9:353–383

Nieder A, Miller EK (2004) Analog numerical representations in rhesus monkeys: evidence for parallel processing. J Cogn Neurosci 16:889–901

Olthof A, Roberts WA (2000) Summation of symbols by pigeons (*Columba livia*): the importance of number and mass of reward items. J Comp Psychol 114:158–166

Pepperberg IM (1987) Evidence for conceptual quantitative abilities in the African grey parrot: labelling of cardinal sets. Ethology 75:37–61

Remy M, Emmerton J (1989) Behavioral spectral sensitivities of different retinal areas in pigeons. Behav Neurosci 103:170–177

Remy M, Emmerton J (1991) Directional dependence of intraocular transfer of stimulus detection in pigeons (*Columba livia*). Behav Neurosci 105:647–652

Remy M, Güntürkün O (1991) Retinal afferents to the optic tectum and the nucleus opticus principalis thalami in the pigeon. J Comp Neurol 304:1–14

Seron X, Pesenti M (2001) The number sense theory needs more empirical evidence. Mind Lang 16:76–88

Tomonaga M (2008) Relative numerosity discrimination by chimpanzees (*Pan troglodytes*): evidence for approximate numerical representations. Anim Cogn 11:43–57

Van Oeffelen MP, Vos PG (1983) An algorithm for pattern description on the level of relative proximity. Pattern Recognit 16:341–348

Verguts T, Fias W (2004) Representation of number in animals and humans: a neural model. J Cogn Neurosci 16:1493–1504

Vos PG, van Oeffelen MP, Tibosch HJ, Allik J (1988) Interactions between area and numerosity. Psychol Res 50:148–154

Xia L, Emmerton J, Siemann M, Delius JD (2001) Pigeons (*Columba livia*) learn to link numerosities with symbols. J Comp Psychol 115:83–91

Xu F, Spelke ES, Goddard S (2005) Number sense in human infants. Develop Sci 8:88–101