

Matching of numerical symbols with number of responses by pigeons

Abstract Pigeons were trained to peck a certain number of times on a key that displayed one of several possible numerical symbols. The particular symbol displayed indicated the number of times that the key had to be pecked. The pigeons signalled the completion of the requirement by operating a separate key. They received a food reward for correct response sequences and time-out penalties for incorrect response sequences. In the first experiment nine pigeons learned to allocate 1, 2, 3 or 4 pecks to the corresponding numerosity symbols s_1 , s_2 , s_3 and s_4 with levels of accuracy well above chance. The second experiment explored the maximum set of numerosities that the pigeons were capable of handling concurrently. Six of the pigeons coped with an s_1 – s_5 task and four pigeons even managed an s_1 – s_6 task with performances that were significantly above chance. Analysis of response times suggested that the pigeons were mainly relying on a number-based rather than on a time-based strategy.

Key words Pigeons · Counting · Timing · Symbols · Responses

Introduction

At the beginning of the last century animal behaviour students became interested in the experimental study of numerical abilities in animals, not least because they were goaded by the apparent algebraical achievements of a horse named Kluger Hans whose performance eventually turned out to be artefactual (Pfungst 1907). Later the numerical capacities of several avian and mammalian species were examined using a wide variety of better-con-

trolled numerical tasks (Rilling 1993). At present there is little doubt that human numerical capabilities, at least at the non-verbally-aided counting level, have an evolutionary past and that animal studies can provide some comparative insights into that phylogenetic history (Dehaene 1997; Butterworth 1999). Beyond that the non-verbal proto-counting abilities of animals may play a role when they assess the novelty-familiarity of stimuli (Fersen and Delius 1989) and, more generally, when they quantify the badness-goodness of some ecological variables (Delius and Siemann 1997).

Two different general kinds of counting behaviour, i.e. “responsive” and “constructive” counting, may be distinguished. The more obvious terms “receptive” and “productive” counting are avoided because they have already been used by Boysen (1992) with a different meaning from that intended here. In responsive counting the number of some externally given set of items or events is counted at a perceptual level and translated into a corresponding numeric-symbolic motor response. In humans most often this response is the production of a numeric word or a written numeral, e.g. “the 3 black sheep that jumped the fence”. In constructive counting a perceived symbolic stimulus conveys a particular numerosity that has to be converted into an equivalent number of discrete motor responses. In humans this could for example involve the instruction “×3” printed by a door-bell next to a friend’s name being converted into three button-pushes. The responsive form of counting has, in different guises, often been studied in animals, including birds and culminating perhaps with the African grey parrot Alex who, upon being asked, for example, “how many blue cork?” were on a tray amongst other coloured bottle-corks and other items such as house-keys and toy cars, some of them also blue, was capable of correctly answering for example, “three blue cork” (Pepperberg 1994). Indeed, in another paper we will ourselves be reporting a further study concerned with such responsive, though of course non-linguistic, counting in pigeons (L. Xia, J. Emmerton, M. Siemann, and J.D. Delius, unpublished work). The constructive form of counting has received far less attention

L. Xia · M. Siemann · J. D. Delius (✉)
Allgemeine Psychologie, Universität Konstanz,
78457 Konstanz, Germany
e-mail: juan.delius@uni-konstanz.de,
Tel.: +49-7531-883563, Fax: +49-7531-883184

in both animal and human research (reviews: Davis and Pérusse 1988; Gallistel and Gelman 1992; Boysen and Capaldi 1993; Dehaene et al. 1998).

In an early study on constructive counting Mechner (1958) trained rats to press a lever a certain number of times in order to obtain reward. In half of the trials the reward was delivered as soon as they had pressed a required number of times (n). In the other half of the trials they were rewarded when they had pressed at least n times and had then pressed a second lever. If the second lever was pressed before the first lever had been pressed n times, no reward was given and the trial was repeated. If they had pressed the first lever less than n times before pressing the second one they received no reward. At different stages the animals were required to press the first lever up to 16 times (Mechner 1958). In the initial follow-up experiment of Platt and Johnson (1971) the rats signalled a response completion by putting their head into a reward-well equipped with a photo-electric sensor. In one group of rats an erroneous premature head-poke ended the trial, while another group of rats could continue to press the lever after such an incorrect poke and head-poke again after more presses. The rats received a reward if they head-poked after they had pressed the lever n or more times. An interval followed before the next trial began. With the latter group the number of times the lever was pressed bunched around the smaller number of responses required but, as this number increased, the distribution became broader, indicating a relative loss of precision with higher numerosities. With the former group the bunching around the required numbers was less pronounced, and furthermore at higher response requirements the response distribution peaks tended to be located around more presses than the minimum number of responses required. The latter finding also applied to Mechner's results, but here the number of surplus presses was fairly constant regardless of the minimum number of responses required.

Zeier (1966) reported a study with pigeons using a somewhat similar, but more exacting procedure. He trained the birds to deliver a predetermined number of pecks to one of two pecking keys and to indicate the completion of the task with a peck to the other key in order to receive a reward. The number of pecks required was increased in stages. The issue of interest was the maximum response requirement that pigeons could manage with some precision. Zeier (1966) found that about a third of the pigeons progressed to the stage where they yielded an above-chance level of 75% correct trials with a requirement of five pecks on the first key. Some pigeons, however, could only manage a task requiring three pecks, while a few birds coped with a requirement of eight pecks. Note, however, that at any given experimental stage Zeier's pigeons were only expected to cope with one, fixed numerical response requirement.

Koehler (1960) had earlier summarised several experiments by his students in which birds of various species had learned to grasp and eat a given number of food items depending on either which numerosity array or which numerosity symbol had been presented to them previously.

Some avian species, notably crows and parrots (Lögler 1959), were capable of learning to peck from one to several items, depending on which array or symbol had been shown on particular trials, a more flexible form of constructive counting than that described by Zeier (1966). Seibt (1982) criticised the inference from experiments with numerosity arrays that the animals had converted the number of items seen in arrays into the number of responses produced. She did an experiment in which pigeons had to match sample arrays consisting of two and three dots with either two pecks and three pecks respectively, or conversely had to match two and three dots with three and two pecks, respectively. Her pigeons apparently did about equally well under both conditions, leading her to conclude that the birds were not transferring the numerosities seen but rather employing the item arrays as numerical symbols that constructively elicited a specified number of responses.

These studies show that rats and birds, pigeons in particular, are capable of executing a particular number of discrete responses, quite precisely when the required number of responses is low but not so precisely when the response requirement is high. Except in the studies mentioned in the previous paragraph, experimenters did not explicitly test whether the animals could deal flexibly and concurrently with varying response requirements, i.e. whether they could adjust the number of responses emitted depending on being cued with different stimuli by, for example, delivering three pecks if a particular cue was present on one trial and delivering five pecks if another cue was present on the next trial. In the present study pigeons were trained to emit a certain number of pecks with each of six different arbitrary symbols. They were first trained to respond to one of the symbols with a single peck, and then, one after another, additional symbols were introduced until the birds finally had to deal concurrently with six symbols that required one to six pecks each.

Experiment 1: matching up to four

This experiment was designed to find out whether pigeons would be able to learn a variable constructive counting task, and also explored which of two different training schedules might be the most efficient.

Method

Subjects

Nine experimentally naive pigeons (*Columba livia*) of local homing stock were used. They were housed in individual cages in a well-ventilated room with a 12 h light/12 h darkness cycle. Throughout the experiment they were maintained deprived of food at 85% of their free-feeding body weight.

Apparatus

A conventional conditioning chamber ($30 \times 34 \times 34$ cm) was used. The back wall of the chamber bore three side-by-side pecking keys (2.5 cm diameter, 9.5 cm apart, 16 cm above the floor). In-line micro-projectors allowed the back-projection of stimuli onto the translucent keys. An automatic feeder opening was located 9 cm below the middle pecking key. The chamber was illuminated by two house-lights (3 W) located 9 cm above the middle key. The upper half of the left side-wall was one-way transparent and allowed observation of the pigeons. All experimental events were controlled and recorded by a personal computer (Commodore) furnished with interface cards (Computer Boards) and programmed in QuickBasic (Microsoft).

Stimuli

The symbol stimuli used are shown in Fig. 1. They were selected as symbols because their surface areas and form complexities were more similar than those of the corresponding arabic numerals. The first four of the six patterns shown in Fig. 1 were used in this first experiment. These stimuli could be individually projected onto the

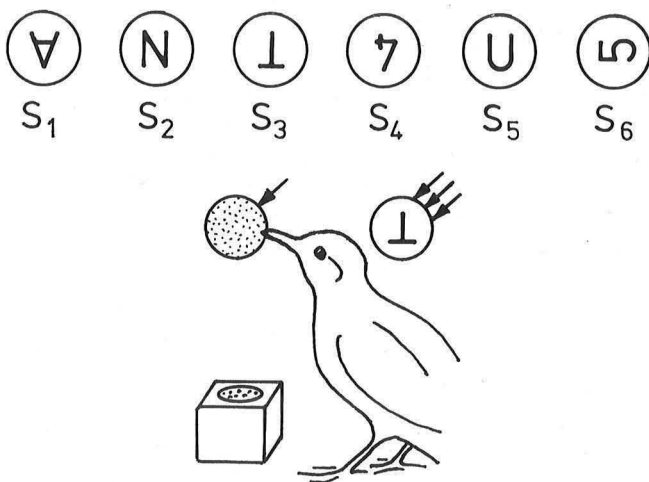


Fig. 1 Symbolic numerosity stimuli and experimental arrangement used. The first four stimuli were used in experiment 1 and all six were used in experiment 2

Table 1 Correct and incorrect response patterns associated with the various numerosity stimuli and the chance probabilities of correct sequences (s_1 – s_6 numerosity symbols, n number of s pecks required, s peck to the symbol-key, e peck to the enter-key). The s_5 and s_6 rows only apply to experiment 2

Stimuli	n	Correct response patterns	Incorrect response patterns	Probability of correct response pattern
s_1	1	se	e, ss	33%
s_2	2	sse	e, se, sss	25%
s_3	3	ssse	e, se, sse, ssss	20%
s_4	4	sssse	e, se, sse, ssse, sssss	17%
s_5	5	ssssse	e, se, sse, ssse, sssse, ssssss	14%
s_6	6	sssssse	e, se, sse, ssse, sssse, ssssse, ssssss	13%

right, *symbol-key* in white on a dark background. The key could also be fully illuminated. The middle, *enter-key* could only be fully illuminated. The left key was not used and remained inactive and dark throughout.

Shaping

The pigeons were taught to peck the right pecking key using an autoshaping procedure. A trial began with a 30-s interval after which the key was illuminated for a maximum period of 8 s. If the animal pecked the key within this period, mixed grain was offered immediately for 2 s, after which the key illumination was extinguished. If the bird did not peck the lit key the food was presented after the end of the period. When the pigeon pecked the key in more than half of the trials during a 30-min period, the same procedure was repeated with the middle key.

Training

The goal was to get the pigeons to respond to any s_n number stimulus presented on the symbol-key with n pecks on that key and an $n+1$ th peck on the enter-key. For this last peck they were rewarded with access to food. If the birds pecked the enter-key after fewer than n pecks on the symbol-key or if they pecked the symbol-key $n+1$ times they were penalised with time-out. The first four entries of Table 1 list the correct response sequences, the incorrect response patterns and the probabilities of a response sequence being correct by chance correct with the stimuli s_1 to s_4 used in this first experiment. The task was expected to be difficult for the pigeons to learn because the higher-numerosity stimuli would initially yield rewards with only low probabilities. The training therefore proceeded in stages using auxiliary procedures (assisted and consolidation sessions) to begin with. Alternative overall training designs were tried out with two different groups of pigeons.

The pigeons were therefore randomly divided into two groups, A and B. The five pigeons of group A were initially trained to cope successively with only one s_n stimulus at a time beginning with s_1 and ending with s_4 . They were exposed to four successive series of daily assisted sessions, each session lasting 200 trials. A series of sessions with a given s_n always started with assisted trials. These trials began with the display of that stimulus on the

symbol-key. Each peck on that key triggered a computer produced 36 ms, 1800 Hz audible feedback sound in addition to the mechanical clicking by the key itself. The enter-key remained unlit. When the animal had pecked the n pecks corresponding to the s_n , the symbol-key darkened and the enter-key was lit. If the pigeon now pecked the enter-key the response pattern was counted as correct and yielded a reward consisting of 2 s access to mixed grain. If, however, the pigeon pecked the dark symbol-key once more or pecked the dark enter-key before completing the n pecks on the symbol-key the response pattern was counted as incorrect and yielded a 5-s houselights-off penalty. During the reward and penalty periods the keys were unlit and inactive. The next trial began immediately after the reward or penalty period. After incorrect trials, the next trial was a correction trial, i.e. a repeat trial. These ceased after the pigeons produced a correct trial. Correction trials were disregarded for the purpose of trial counts and performance scores.

Whether assisted trials were replaced by ordinary trials was determined by the current level of performance shown by the birds. A running weighted performance index c_n was updated after each trial n according to the formula

$$c_n = c_{n-1} + w(1 - c_{n-1})$$

when the pigeon had produced a correct response or according to the formula

$$c_n = c_{n-1} - wc_{n-1}$$

when it had produced an incorrect response. The parameter w , routinely set to $w = 0.01$, determined the impact which the outcome of the n th trial had on the c_{n-1} that had accumulated up to the preceding $n-1$ th trial. If, after a given trial, c_n was below a criterion 0.3 the next $n+1$ th trial was an assisted trial, while if c_n was equal to or above that criterion the $n+1$ th trial was an ordinary trial. These latter trials began with both the symbol-key and the enter-key being lit and active. The response pattern definitions and reinforcement outcomes, including correction trials, however, were the same as with assisted trials. As the pigeons' performance improved the proportion of assisted trials decreased and that of ordinary trials increased. The current numerosity stimulus was replaced by the next higher numerosity stimulus as soon as the pigeons met a criterion of 75% correct responses within 20 consecutive ordinary trials.

After the pigeons had successively reached the learning criterion for each and all of the four numerosity stimuli, they were trained in consolidation sessions in which each and all the stimuli occurred at various times. All 200 trials of these sessions were ordinary trials. To aid learning, a given numerosity stimulus was presented repeatedly in a run of trials. A run with that particular stimulus ended when the pigeons produced three correct responses within five consecutive trials. The stimulus selected for the next run was that exhibiting the worst, i.e. the lowest c_n performance index at the time. As the pigeons' performance improved the stimulus switch criterion was re-

duced to two correct trials out of three trials and then to one correct trial out of one. The training terminated when the animals achieved 75% correct out of 20 consecutive trials within a single session under this last switch criterion.

The four pigeons in the B group were trained with the same types of sessions as the A group pigeons but following a different overall design. They began by being trained on the s_1 and s_2 stimuli in exactly the same way as group A. However, when the birds had successively achieved the 75% correct criterion on both stimuli they were trained in consolidation sessions where these two stimuli were presented in alternating runs of ordinary trials. When the pigeons met the 75% criterion with these stimuli, they were trained with the numerosity stimulus s_3 in assisted sessions. Upon reaching the criterion with stimulus s_3 they were exposed to consolidation sessions involving stimuli s_1 , s_2 and s_3 until the 75% criterion was met again. The pigeons were then trained with assisted sessions on stimulus s_4 until they reached criterion. Finally, they were trained in consolidation sessions involving all four stimuli presented in runs of ordinary trials. This terminal training was fully equivalent to the final stage of group A and also ended when the pigeons met the same final criterion of 75% correct trials in 20 consecutive trials.

Testing

To assess the pigeons' response counting in the absence of any auxiliary cues or procedures the birds of both groups were finally exposed to a test session consisting of 80 ordinary trials without correction trials. Each numerosity stimulus was presented in 20 ordinary trials but in a quasirandom sequence incorporating the restriction that a given numerosity stimulus could not appear in more than two consecutive trials. In contrast to the consolidation sessions, in this testing session incorrect trials did not lead to repeat correction trials.

Results and discussion

All nine pigeons completed all training stages by achieving the criterion of 75% correct in 20 consecutive trials. The number of trials which group A and group B pigeons needed to reach this criterion with the different stages is shown in Fig. 2. The groups differed significantly in the total number of trials they required to reach the final training stage criterion when all four stimuli were concurrently presented ($t_7 = 2.42$, $P = 0.046$). The group A pigeons required an average of 13,180 trials (10,811–17,317) to complete the last stage and the group B pigeons an average of 20,606 trials (15,527–27,781).

The results of the test session are shown in Fig. 3. The five pigeons in group A yielded a mean of 80%, 59%, 58% and 60% correct responses with stimuli s_1 , s_2 , s_3 and s_4 , respectively. These accuracies are well above the corresponding chance levels of 33%, 25%, 20% and 16%

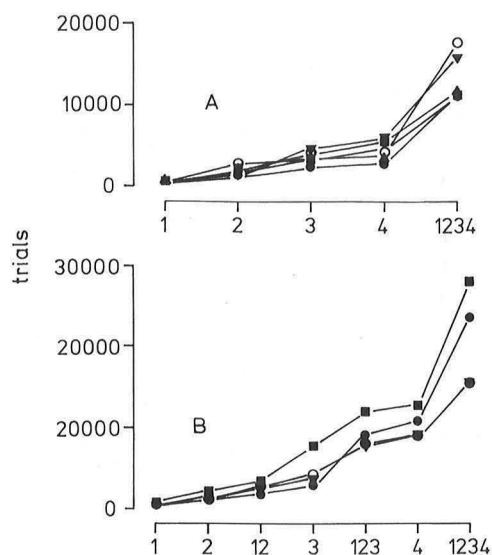


Fig. 2 Experiment 1. Cumulative number of trials that the individual pigeons of **A** group A (5 pigeons) and **B** group B (4 pigeons) needed to complete the various training stages up to stage 1234

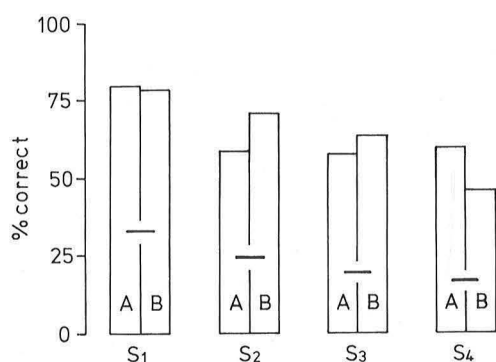


Fig. 3 Experiment 1. Mean percent correct responses shown by group A and group B pigeons with the s_1 – s_4 numerosity stimuli during the test stage. Bars indicate chance-level performances

correct trials listed in Table 1 (t -tests; s_1 : $t_4 = 17.0$, $P = 0.00007$; s_2 : $t_4 = 7.0$, $P = 0.0022$; s_3 : $t_4 = 7.0$, $P = 0.0022$; s_4 : $t_4 = 10.3$, $P = 0.00050$). The pigeons in group B also revealed an above-chance-level performance of 79%, 72%, 64% and 46% correct responses with stimuli s_1 , s_2 , s_3 and s_4 , respectively (t -tests; s_1 : $t_3 = 10.6$, $P = 0.0018$; s_2 : $t_3 = 6.2$, $P = 0.0084$; s_3 : $t_3 = 6.6$, $P = 0.0070$; s_4 : $t_3 = 12.3$, $P = 0.0011$). The accuracy differences between the two groups were not significant except for the s_4 numerosity ($t_7 = 2.65$, $P = 0.032$).

Since there was only a small difference between the test results of the two groups, the data of all nine pigeons were pooled for the purposes of a finer analysis. The total numbers of correct and incorrect responses during the final consolidation stage and again during the test stage was computed separately for each of the numerosity stimuli s_1 – s_4 . The number of correct and various kinds of incorrect responses to each of these stimuli were expressed as percentages of the corresponding totals and are shown plotted as response distributions of the responses in Fig. 4.

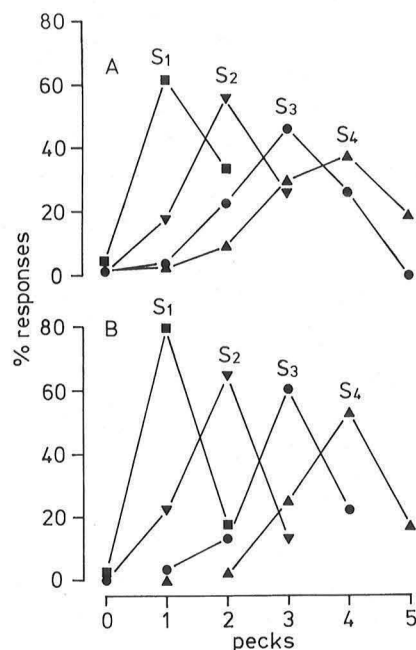


Fig. 4 Experiment 2. Mean percent distributions of responses with the s_1 – s_5 numerosity stimuli during **A** the consolidation and **B** the test stage

Both diagrams demonstrate that the pigeons most often made errors by pecking the symbol-key once more or once less than the number of pecks defining the correct response patterns. This was more pronounced during the test stage than during the consolidation stage. The latter of course includes a phase when the birds were not yet showing their optimal performance. The levelling out of the response distributions may represent a loss of precision with higher numerosities, perhaps because they had been less trained than the lower numerosities, and/or because they allowed more kinds of error response patterns than the smaller numerosities (Table 1).

The results show that pigeons are able to respond with significant accuracy to four different numerosity symbols with a corresponding number of pecks, even when the various symbols were presented in a quasi-randomised sequence during the test stage. That means that the pigeons were able to match four different numerosity symbols with the corresponding number of pecks.

Experiment 2: matching up to six

This experiment was designed to explore the upper bounds of the pigeons' ability to learn to match different numerosity symbols with a corresponding number of pecks.

Method

Seven of the nine pigeons that had participated in the previous experiment were used. Two pigeons were excluded

because they had required more than 20,000 trials to complete that experiment. The same apparatus and procedures were used. All birds, regardless of which group they had belonged to, were now trained and tested in the same way. In addition to the four numerosity stimuli used in experiment 1, experiment 2 also involved the numerosity stimuli s_5 and s_6 (Fig. 1, Table 1). The birds were first trained on stimulus s_5 using assisted sessions. When the animals reached the criterion of 75% trials correct within 20 ordinary trials with that stimulus, all five stimuli s_1 – s_5 were used in consolidation sessions. The birds were trained until they yielded 75% correct trials out of 12 trials for each stimulus within a single session, or until they had completed more than 30,000 trials without reaching this criterion, in which case they were excluded from further participation. Then the remaining pigeons were tested in a single testing session, within which all five stimuli were presented 20 times in a quasi-randomised sequence as described earlier. The next stage involving stimulus s_6 proceeded in a similar manner.

Results and discussion

Six out of the seven pigeons reached the learning criterion with the numerosity stimuli s_1 to s_5 after an average of 11,800 training trials (3,800–27,500) additional to those that they had needed to complete experiment 1. Four of these six pigeons reached the criterion with the stimuli s_1 – s_6 after an average of 19,340 additional trials (11,500–27,500). Figure 5 shows the cumulative total number of trials the six and four pigeons required to complete the various training stages.

The results of the test sessions are shown in Fig. 6. The six pigeons that were successful with the s_1 – s_5 training yielded 82%, 83%, 62%, 57% and 42% average correct responses with the s_1, s_2, s_3, s_4 and s_5 stimuli, respectively. All these scores were significantly above the chance levels (t -tests, $s_1: t_5 = 12.7, P = 0.000054; s_2: t_5 = 14.5, P = 0.000028; s_3: t_5 = 7.9, P = 0.00052; s_4: t_5 = 8.3, P = 0.00041; s_5: t_5 = 7.2, P = 0.00080$). Similarly, the four pi-

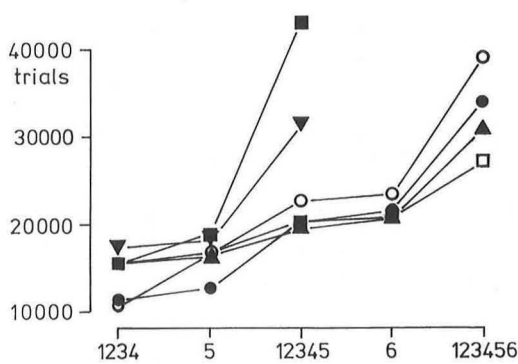


Fig. 5 Experiment 2. Cumulative number of training trials that individual pigeons needed to reach the various stages up to stages 12345 (6 pigeons) and 123456 (4 pigeons)

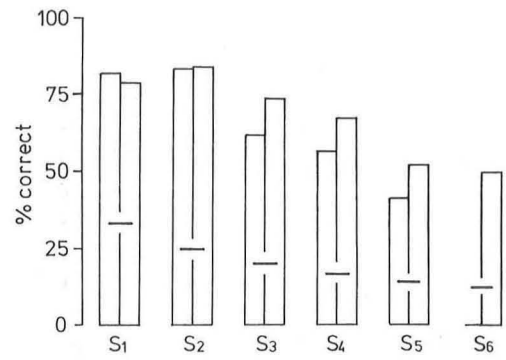


Fig. 6 Experiment 2. Mean percent correct responses during the test sessions with the s_1 – s_5 (six pigeons; left columns) and the s_1 to s_6 stimuli (four pigeons; right columns). Bars indicate chance-level performances.

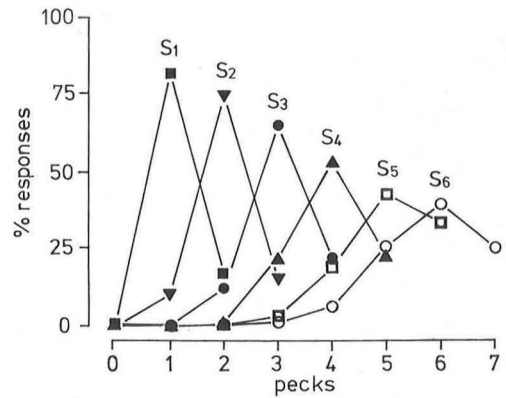


Fig. 7 Experiment 2. Mean percent distributions of responses with stimuli s_1 – s_6 during the final consolidation stage

geons that also completed the training with the s_1 – s_6 stimuli yielded an average 79%, 84%, 74%, 68%, 53% and 50% average correct trials with the s_1, s_2, s_3, s_4, s_5 and s_6 stimuli, respectively. All these averages were again significantly above the corresponding chance levels (t -tests, $s_1: t_3 = 10.6, P = 0.0017; s_2: t_3 = 11.4, P = 0.0014; s_3: t_3 = 12.5, P = 0.0011; s_4: t_3 = 15.7, P = 0.00056; s_5: t_3 = 6.4, P = 0.0077; s_6: t_3 = 10.6, P = 0.0017$).

Figure 7 shows the response distributions, calculated in the same manner as explained earlier, for the six stimuli based on the data of the final consolidation stage with stimuli s_1 – s_6 . There were not enough data for the test stage to yield meaningful response distributions. It is nevertheless apparent that, compared to the corresponding results of the first experiment, the response distributions with the numerosity stimuli s_1 – s_4 were now steeper. This might be because by now the pigeons had acquired much additional experience with these lower numerosities, so erroneous responses occurred less frequently for these stimuli than in experiment 1. As in experiment 1, however, the spread of errors was less with the lower-numerosity stimuli than with the higher ones.

General discussion

As far as we know this is the first demonstration in pigeons of a flexible constructive-counting-like behaviour that involved the production of up to six responses conditional upon the separate pre-presentation of up to six numerosity symbols. The counting behaviour was flexible in the sense that the pigeons ended up responding in a concurrent manner to four, five or six different arbitrary visual symbols each standing for the different numerosities between 1 and 6. The pigeons responded correctly to these different numerosity stimuli at levels that were significantly above chance by emitting the number of pecks the stimuli were meant to symbolise. This clearly goes beyond the constructive counting ability that Zeier (1966) had previously demonstrated in pigeons. His birds had only to respond a fixed number of times at any stage of his experiment even though the particular numerical response demanded was progressively augmented. In our case the various symbolic stimuli signalled in a random trial-to-trial order the number of responses that the pigeons had to produce to obtain reinforcement.

The pigeons undeniably required extensive training to master this variable constructive counting task. They needed, for example, an average of some 15,000 trials before they reliably counted up to four pecks at the end of experiment 1. But this training was nevertheless briefer than the average 25,000 trials or so that other pigeons had needed to reliably count four visual stimuli in a responsive counting task (L. Xia, J. Emmerton, M. Siemann, and J.D. Delius, unpublished work). Incidentally, four of these latter pigeons that had managed to reliably count four or even five stimuli in that study were then trained using the same procedures as those employed with group A group in experiment 1 of this study. These birds only took an average of some 8,000 trials compared with the about 13,000 trials that our present group A pigeons needed to reliably count up to four responses (Xia 1998). Although the speedier learning of these experienced pigeons might have been due to their pre-exposure to conditioning procedures generally, it is also possible that some of their advantage might have been due to a more specific transfer of numerical skill from the responsive to the constructive forms of counting.

Quite what information the pigeons utilised when they responded with the correct number of pecks before switching over to the enter-key remains largely uncertain, however. A hypothetical cumulative peck-counting process could, for example, have relayed either on the registration of the efferent neural peck commands or else of the reafferent sensory volleys arising from the pecking actions. Among these latter signals one might expect the tactile and auditory stimuli arising from the beak's impact on the key and from the mechanical motion of the key as well as the computer produced feedback tone to possibly play a role. But besides these exteroceptive stimuli, interoceptive mechanical and vestibular stimuli consequent on the peck motions could also have intervened (Schall

and Delius 1991). It is, however, also possible that the correct response sequences might have been assessed on the basis of the time that elapsed between the first and the last peck, since repetitive pecks of pigeons are known to be emitted at a rather constant rate of about 3–4 s⁻¹ (Delius et al. 1986). We may thus assume that the production of six pecks, for example takes about three times as long as the production of two pecks.

Roberts and Mitchell (1994) had pigeons learn to discriminate two different visual stimuli. One lasted 2 s and consisted of two light pulses, another one lasted 8 s and consisted of eight light pulses. To get a reward the birds had to peck one key if the longer stimulus with more pulses had been presented and another key if the shorter stimulus with fewer pulses had been presented. Intermediate signals were then presented. They were either of a fixed duration (e.g. 4 s) while the number of pulses varied from 2 s to 8 s or they consisted of a fixed intermediate number of pulses (e.g. 4) but their duration varied from 2 s to 8 s. In these time-based and number-based tests the relative choice frequencies of the longer/more key increased with both the number and the duration of the relevant test stimuli. This indicated that pigeons were using both counts and times to achieve correct responding during the test trials. Fetterman (1993) reviewed older data and reported additional findings that replicated this ambivalent finding with rats. A currently still popular so-called mode control model assumes that there is a trade-off between a timing mechanism and a counting mechanism operating on the basis of the accumulation of either endogenous clock pulses or exogenous perceptual item pulses (Meck and Church 1983; but see Staddon and Higa 1999). However, Emmerton et al. (1997) ran an experiment where pigeons were presented with arrays of dots. When one or two dots were presented, pigeons had to peck one key, when six or seven dots were presented they had to peck another key. Tests with intermediate arrays of three, four or five dots also resulted in an orderly graded choice of the two keys. In this instance, because of a simultaneous numerosity information presentation, the pigeons are unlikely to have estimated the number of dots on a passage-of-time basis.

The timing of the pecks onto the symbol-key were in fact recorded during the final test stages of experiments 1 and 2. A summary of this data is presented in Table 2. The table ignores the intervals from the onset of the trial to the first peck of each sequence since these latencies were too variable to convey truly meaningful information. No doubt their variability was due to episodes of inattention which caused at least some of the birds to miss the beginning of the symbol-key illumination. The table suggests that the average interval between the pecks was in principle regular enough for a timing-based production of correct responses. Interestingly, the emission of the final symbol-key peck was quite consistently associated with a longer inter-peck interval than the other symbol-key pecks. This could mean that the required switch to the enter-key was proactively interfering with the end of the symbol-key sequence. However, this effect is compatible

Table 2 Average intervals (ms) between pecks to the numerosity stimuli s_2 , s_3 and s_4 during the correct response sequences indicated. Intervals in (ms) between 1st-2nd, 2nd-3rd, 3rd-4th, 4th-5th

Stimulus Sequence	s_2 sse	s_3 ssse	s_4 sssse	s_5 ssssse	s_6 sssssse
Experiment 1	334	278	349	274 274	324 –
Experiment 2	258	227	234	215 213	251 244 245 251 289 271 232 228 239 312

and 5th-6th pecks of these sequences for final test sessions of experiment 1 (9 pigeons) and experiment 2 (4 pigeons). The closing mean intervals for each sequence are printed in **bold** numbers

Table 3 Proportion of error response times connected with under- or overproduction of number of pecks that were in agreement with a timing strategy. Based on the test session of experiment 2 (4 pigeons)

Numerosity stimulus	s_3	s_4	s_5	s_6	Total
Proportion cases where $c_{t_{n-1}} \geq c_{t_n}$	0/11	1/12	4/17	2/20	7/60
Proportion cases where $c_{t_n} < c_{t_{n-1}}$	7/10	7/12	10/17	14/15	38/54
Total agreeing with timing hypothesis	7/21	8/24	14/34	16/35	45/114

with both a counting and a timing process. Beyond this there was no indication that the pigeons developed any symbol-specific rhythmic response patterns which could have enabled them to generate the correct number of pecks in a manner that bypassed counting- or timing-like operations.

To examine whether the pigeons had in fact used a timing strategy we carried out an analysis of errors based on response times. First, the average times elapsed between the first and the last numerosity key-peck were calculated for the correct responses in the experiment 2 test session. This was done separately for each pigeon and each numerosity from s_3 to s_6 (s_2 was excluded because the corresponding single interval response times were too brief for meaningful comparisons). If the pigeons were relying on a timing mechanism the correct response times c_{t_3} to c_{t_6} should have been memorized by the pigeons as a mean criterion time indicating when it would pay to switch from pecking the symbol key to pecking the enter key. It could be then expected that erroneous sequences that involved a peck too few (we ignored the few errors involving even fewer pecks) might have arisen because the pigeons had been exceptionally slow in pecking and had reached the corresponding internalized criterion time before completing the correct number of pecks. If so the response times associated with such errors $c_{t_{n-1}}$ should, as a rule, have equalled or even exceeded the corresponding c_{t_n} criterion times. It is also possible that erroneous sequences involving a peck too many might have arisen because the pigeons had been exceptionally fast in pecking and thus not reached the corresponding criterion time after the number of pecks actually required, this causing them to produce a superfluous, erroneous peck. If so the average cumulated times c_{t_n} associated with the nominally required number of pecks (i.e. one less than that actually produced) associated with such errors should as a rule been less than the corresponding criterion time c_{t_n} . Table 3 shows that while the times associated with overcounting errors may have

partially supported the timing hypothesis, those associated with undercounting errors definitely did not do so. The total number of cases supporting the timing hypothesis was statistically indistinguishable from the 50% chance level (binomial test). Inclusion of error cases involving fewer pecks than just one, of the s_1 and s_2 numerosity trials and of the trial-onset to first peck intervals into analogous evaluations did not alter the conclusion that a timing strategy played at most a very partial role in determining the counting performance of our pigeons

Returning again to the counting aspect, it is apparent that, as shown in Figs. 4 and 7, the distribution of error responses varied systematically with the number of pecks required. In this respect the results compare with those obtained in rats by Platt and Johnson (1971). Errors mostly concerned numbers of pecks that were near rather than distant from the correct response pattern, that is, if animals were required to peck four times then they tended to erroneously peck three or five times rather than erroneously peck two or six times. Moreover, the proportions of correct response sequences decreased as the number of elementary responses required increased. In other words, the response distributions became flatter and broader when the required number of lever presses/pecks was increased. At first sight this seems to roughly obey Weber's law, i.e. the bigger the absolute number of responses demanded, the larger the just-discriminable numerical difference that could be achieved. But in our study the errors were not only downwards-censored by necessarily excluding less than zero responses, but also upwards censored by allowing no more than $n+1$ responses. This meant that, for example, with s_2 there were only two erroneous response sequences and that with s_5 there were no fewer than five such erroneous sequences among which the animals could distribute their errors (Table 1). This would necessarily tend to generate Weber-like data artefactually based on a procedural constraint rather than on any sensory mechanism. One must also consider that the pigeons had more training experience with the part-tasks involving lower numerosities. Due to these procedural constraints it is difficult to be sure whether there was a significant additional perceptual effect even though it is likely that there was. This is regrettable because much has been theoretically made of the Weber effect in responsive counting contexts where it was valid to assume that it was truly based on perceptual factors (e.g. Meck and Church 1983; Emmerton 1998). Note that applied to our counting task Weber's law would suggest that although pigeons might well have difficulties in accurately producing say, seven and eight responses to a pair of corresponding symbols, they might still be capable of accurately producing,

say, 15–17 and 18–22 responses to a pair of corresponding symbols. That is, the count resolution becomes coarser with increasing numerosities (cf. Rilling 1965). Whether this is actually so still needs to be investigated.

Dehaene and Changeux (1993) have presented a neural net model capable of a counting-like performance of this kind. Its attraction is that it offers a rigorously mechanistic account of a responsive kind of non-verbal counting. The net consists of several layers: dots of different sizes and location can be presented to an input layer, which is connected to an array of neural Gaussian filters. Through this filtering, the sizes and locations of the dots become irrelevant. The output of this so-called topographical mapping layer is transferred to a summation layer, containing a number of units that have increasing thresholds. The summed and threshold-exceeding activation of a certain unit in this layer is roughly proportional to the number of dots present in the stimulus array. These units are connected to an output layer of units within which, through lateral connections, units responsible for higher numerosities can inhibit units responsible for lower ones. This ensures that the net output layer only responds to the highest numerosity that is activated, because units responsible for lower numerosity, which are activated by subsets of the input stimulus, are suppressed. The activity distributions of the output layer units show Weber-like properties insofar as with increasing numerosity arrays the peak activities decline and the activity distributions become broader (implying that the overlap between adjacent active units increases). Clearly, as presented the model cannot produce a specified number of responses upon recognising a given numerical symbol, i.e. behave in a constructive manner, but it would be feasible to add such a module. Dehaene and Changeux (1993) have in any case sketched how the model may be easily modified to count successive events, such as the reafferent stimuli upon pecking or indeed the peck commands themselves, as demanded in our particular case. The development of a fully explicit neural model of constructive counting is desirable because it would generate more precise predictions about the neurophysiological basis of such counting (cf. Thompson et al. 1970) than purely formally descriptive models of counting.

Acknowledgements The research was supported by the Deutsche Forschungsgemeinschaft, Bonn. We are grateful to Prof. Jacky Emmerton for excellent advice while she was on sabbatical leave from Purdue University and a visiting professor in Konstanz. She also helped much by improving an earlier draft. We also thank Jessica Grante for revising the language of the final manuscript. We dedicate this paper to the memory of Otto Koehler whose pioneering work on a verbal counting in animals has in the past been frequently doubted but which when replicated has nearly invariably proven right.

References

- Boysen ST (1992) Counting as the chimpanzee views it. In: Fetterman G, Honig WK (eds) *Cognitive aspects of stimulus control*. Erlbaum, Hillsdale, pp 367–383
- Boysen ST, Capaldi EJ (eds) (1993) *The development of numerical competence: animal and human models*. Erlbaum, Hillsdale
- Butterworth B (1999) *The mathematical brain*. Macmillan, London
- Davis H, Pérusse R (1988) Numerical competence in animals: Definitional issues, current evidence, and a new research agenda. *Behav Brain Sci* 11:561–615
- Dehaene S (1997) *The number sense: how the mind creates mathematics*. Oxford University Press, Oxford
- Dehaene S, Changeux JP (1993) Development of elementary numerical abilities: a neuronal model. *J Cogn Neurosci* 5:390–407
- Dehaene S, Dahan-Lambertz G, Cohen L (1998) Abstract representations of numbers in the animal and human brain. *Trends Neurosci* 2:355–361
- Delius JD, Siemann M (1997) Transitive responding in animals and humans: exaptation rather than adaptation? *Behav Proc* 42: 107–137
- Delius JD, Lindenblatt U, Lombardi C (1986) Coupling between pecking and heartbeat in pigeons. *J Interdis Cycle Res* 17:243–251
- Emmerton J (1998) Numerosity differences and effect of stimulus density in pigeons' discrimination performance. *Anim Learn Behav* 26:243–256
- Emmerton J, Lohmann A, Niemann J (1997) Pigeons' serial ordering of numerosity with visual arrays. *Anim Learn Behav* 25: 234–244
- Fersen L von, Delius JD (1989) Long term retention of many visual patterns by pigeons. *Ethology* 82:141–155
- Fetterman JG (1993) Numerosity discrimination: both time and number matter. *J Exp Psychol Anim Behav Proc* 19:149–164
- Gallistel CR, Gelman R (1992) Preverbal and verbal counting and computation. *Cognition* 44:43–74
- Koehler O (1960) Le dénombrement chez les animaux. *J Psychol Norm Pathol* 1:45–58
- Lögler P (1959) Versuche zur Frage des "Zähl"-Vermögens an einem Graupapagei und Vergleichsversuche an Menschen. *Z Tierpsychol* 16:179–217
- Mechner F (1958) Probability relations within response sequences under ratio reinforcement. *J Exp Anal Behav* 1:109–122
- Meck WH, Church RM (1983) A mode-control model of counting and timing processes. *J Exp Psychol Anim Behav Proc* 9:320–334
- Pepperberg IM (1994) Numerical competence in an African grey parrot (*Psittacus erithacus*). *J Comp Psychol* 108:36–44
- Pfungst O (1907) *Das Pferd von Herrn Osten*. Leipzig, reprinted (1977) as: *Der kluge Hans. Ein Beitrag zur nicht-verbalen Kommunikation*. Fachbuchhandlung für Psychologie, Frankfurt
- Platt JR, Johnson DM (1971) Localization of position within a homogeneous behavior chain: effects of error contingencies. *Learn Motiv* 2:386–414
- Rilling M (1993) Invisible counting animals: a history of contributions from comparative psychology, ethology, and learning theory. In: Boysen ST, Capaldi EJ (eds) *The development of numerical competence: animal and human models*. Erlbaum, Hillsdale, pp 3–37
- Rilling M, McDiarmid C (1965) Signal detection in fixed-ratio schedules. *Science* 148:526–527
- Roberts WA, Mitchell S (1994) Can a pigeon simultaneously process temporal and numerical information? *J Exp Psychol Anim Behav Proc* 20:66–78
- Schall U, Delius JD (1991) Grasping in the pigeon: control through sound and vibration feedback mediated by the nucleus basalis. *Physiol Behav* 50:983–988
- Seibt U (1982) Zahlbegriff und Zahlverhalten bei Tieren. *Neue Versuche und Deutungen*. *Z Tierpsychol* 60:325–341
- Staddon JER, Higa JJ (1999) Time and memory: towards a pace-maker-free theory of interval timing. *J Exp Anal Behav* 71: 215–251
- Thompson RF, Mayers KS, Robertson RT, Patterson CJ (1970) Number coding in association cortex of the cat. *Science* 168: 271–273
- Xia L (1998) Numerische Fähigkeiten bei Tauben. *Tectum, Marburg*
- Zeier H (1966) Über sequentielles Lernen bei Tauben, mit spezieller Berücksichtigung des "Zähl"-Verhaltens. *Z Tierpsychol* 23:161–189