



Journal of Articles in Support of the Null Hypothesis

Vol. 9, No. 1

Copyright 2012 by Reysen Group. 1539-8714

www.jasnh.com

Do jackdaws have a memory for order?

Gerit Pfuhl

Robert Biegler

Psykologisk Institutt, Norges teknisk-naturvitenskapelige universitet (NTNU)

When something happened is an essential feature of episodic memory. There are different ways to remember time of occurrence: absolute time, elapsed time, the order of events. We studied jackdaws' memory for the order of events by presenting random sequences of sample pictures against a consistently ordered sequence of backgrounds. In the choice phase, the birds had to identify the sample picture that had been presented against a specific background, or to identify which of the samples had been presented first, second or third. The birds used a background to identify the corresponding sample picture, but failed using backgrounds as retrieval cues for the ordinal position of a sample in the sequence. We found no evidence that order was encoded.

Introduction

Tulving (1972) defined episodic memory as an integrated memory of what happened where and when. He argued this was a uniquely human ability. The argument raises interesting questions. Could a complete memory system have evolved in the 5 – 8 million years since the last common ancestor between humans, chimpanzees and bonobos (Patterson et al. 2006) despite speed limits on evolution (Worden, 1995)? Or could episodic memory have emerged from adding a critical component to existing systems or integrating them in a new way? If so, those systems should exist in other species. In the last decade, evidence mounted that some animals can show episodic-like memory (corvids: Clayton & Dickinson 1998; primates: Schwartz, Colon, Sanchez, Rodriguez & Evans 2002; rodents: Babb & Crystal 2005, Eacott & Norman 2004, Ferkin, Combs, del Barco-Trillo, Pierce & Franklin 2008, Kart-Teke, Silva, Huston & Dere 2006) although some tests yielded negative results (monkeys: Hampton, Hampstead & Murray 2005, rats: Bird, Roberts, Abroms, Kit & Crupi 2003).

Some work related to episodic memory has focussed specifically on the temporal aspects, on the grounds that episodic memory is often not like a static snapshot of a moment, but more like a short video clip, a dynamic process where components of the event have an order (Bekarian & Bowers, 1983; Burt, Watt, Mitchell & Conway 1998; Ferbinteanu, Kennedy & Shapiro, 2006). The temporal aspect of episodic memory could be subserved by three different temporal codes: an absolute code (date and/or time of day), an elapsed time code (how long ago did this happen?) and an order code (Friedman, 1993). The work of Clayton and colleagues has focused on elapsed time. A bird chooses which cached item to retrieve based on how long ago it was cached. Similarly in a study to disentangle whether rats use an absolute or an elapsed time code, rats relied only on elapsed time (Roberts, Feeney, MacPherson, Petter, McMillan & Musolino 2008). However, elapsed time is not always the best temporal code. The time between elements of a sequence may be in the range of seconds, while the time elapsed between encoding and retrieving an episode may range from days to years. Under those conditions, even a small amount of noise in the elapsed time signal would make elapsed time useless for reconstructing the order of events. Remembering the order of short sequences seems to be a critical component of episodic memory.

Memory for order has been studied in other species in connection with hippocampal function (in humans, a functioning hippocampus is needed for episodic memory). The hippocampus was found to be needed to remember trial-specific sequences. In that literature, experiments on temporal order memory have so far focused on comparing two locations or items, one of them encountered earlier than the other (Agster, Fortin & Eichenbaum 2002, Hannesson, Vacca, Howland & Phillips 2004, Hoge & Kesner 2007, Manns, Howard & Eichenbaum 2007). At short delays between encoding and retrieval, such a binary comparison can be solved by familiarity, which is correlated with elapsed time. One exception is a study by Maki, Beatty and Clouse (1984), in which one group of rats could solve both item and order discriminations by always choosing the less familiar of two offered arms in a radial maze. Another group of rats in this study learned to avoid the first arm, though that was less familiar in the order discrimination and more familiar in the item discrimination. That group must have used information other than familiarity, but it

is not clear whether that is order information. Terrace, Chen and Jaswal (1996) argued that the beginning and end of sequences have distinct features in addition to ordinal position.

In the context of episodic memory, it is important to avoid tasks that can be solved using familiarity or using features based on the beginning and end of a sequence. We are more interested in whether an animal either encodes the ordinal positions of elements in a sequence, or encodes the whole sequence and is able to retrieve a specified element of that sequence, thus providing flexible retrieval of potentially relevant information. Shimp and Moffitt (1974) and Shimp (1976) developed an experimental design that addresses this exact issue. Pigeons pecked at one of two side keys that were illuminated with red, blue, or white. The colours were always in the order red, blue, white, but the left-right sequence was chosen at random. For example, one sequence might be red left, blue right and white right, another sequence red right, blue left and white right. When a third key in the middle lit up in one of the three colours, the pigeons first pecked that. Then the side keys lit up in the same colour, and the pigeons' task was to choose the response key on the side that had been illuminated in this colour. Because the colour sequence was consistent, red also meant the first stimulus was the correct match, blue directed a bird to the second stimulus, and white to the third. However, as long as the sample stimuli were coloured lights (Shimp & Moffitt, 1974), the colour of the middle key could serve as a retrieval cue to where that same colour had been shown, without remembering whether that colour had been shown first, second or third. To overcome this limitation, Shimp (1976) changed the procedure. The sample stimulus was always the same, a white X projected onto the right and left response keys. The idea was that the solution would depend on remembering the sequence of spatial positions of the X, and matching the remembered spatial sequence to the colour signal for the first, second or third ordinal position. With only two response keys, there were only eight sequences: LLL, LLR, LRL, LRR, RLL, RLR, RRL, RRR. Multiplied by the three cues to the correct choice, that gives 24 possible combinations. Shimp (1976) argued that it was unlikely that the pigeons had solved the task by learning all possible combinations. However the pigeons received on average 267 training trials per combination (395 trials per day for 16 days with 24 combinations) before testing started with the shortest retention interval. Before testing with the last retention interval began, the birds had already received an average of 818 trials per combination of sequence and retrieval cue. Furthermore, there are two additional strategies that could lift the pigeons' performance above chance level without using the colour as a retrieval cue for ordinal position in a sequence. A very strong recency effect is consistent with a pigeon having learned that white means it should peck where it last saw the white X. When a pigeon was shown the red or blue light, pecking the side where a stimulus had been shown more often would be correct in 75% of trials; always for sequences LLL and RRR, and in two thirds of cases for the other six sequences. Close examination of Shimp's table 2 shows evidence that these strategies did contribute to the pigeons' performance. Thus the conclusion that pigeons had learned to retrieve the specific item they were asked to retrieve from a sequence may not be reliable. However, the principle of the procedure is sound, so with some modifications, the experiment seemed worth replicating and extending.

To rule out rote learning and use of familiarity, we took advantage of the technical advances since 1976 to improve on the details of the procedure. Using more varied stimuli, we could generate 6840 different 3-item sequences of stimulus pictures and 20520 possible combinations of picture sequence and retrieval cue. Like Shimp, we could ask our birds

“Which was first?”, “Which was second?” and so on, or more generally, “Which was Nth?” Finally, the procedure can be modified to require recall, not recognition, of the locations of sample items (Figure 1).

Experiment 1: matching-to-sequence and matching-to-background with ordered sequence

Method

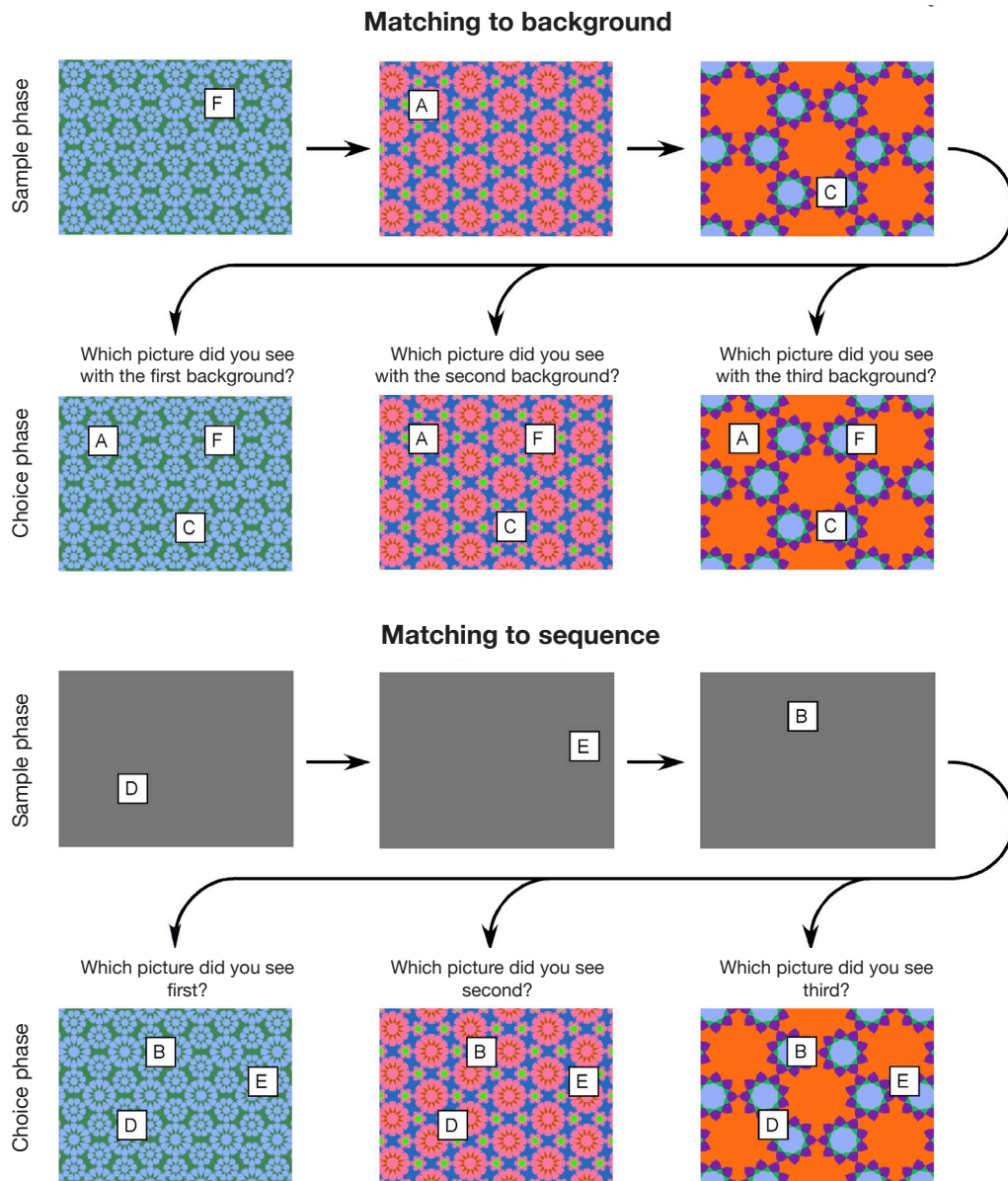
Subjects were two hand-raised jackdaws (*Corvus monedula*), approximately 2 years old at the start of the experiments. They live in an indoor aviary consisting of two separable rooms (4 m × 4 m and 5 m × 3 m), with shelves, branches, perches, toys, regulated lighting (by natural daylight until minimum day length was below 9 hours then controlled by clock, L:D 8am:5pm) and ad lib access to water. The rooms were connected by a door with a wire mesh window to allow the birds visual contact with each other when separated. Given the neophobia of adult jackdaws, we introduced changes gradually to avoid a break-down of work on the computers.

During experimental sessions, each bird could reach an infrared touch screen (Elo touch) connected to a PC. The screen was 32.8 cm wide and 24.8 cm high. Food pellets were delivered through feeders (designed and built by D. Dye) down a tube. The computer screen for Bird S and the joint sessions was in a window in a wall of room 1. In room 2, the computer, touch screen, and feeders were housed on a trolley covered by a box. Only Bird A is willing to work on the trolley system. The birds were familiar with pecking at pictures on the screen from previous experiments.

The diet was rolled pellets (ca. 0.5 g) made of a mixture of bird food (Orlux insect mix, Orlux eggfood, Nutribird A21 handraising formula) with additional calcium, vitamins and egg yolk. Peanuts, egg shells and fresh fruits were given or hidden in room 1 after the birds finished the individual sessions and the connecting door between the rooms had been opened. Outside the training sessions, the computer of Bird S was still switched on and its feeders stocked with pellets, so that even after the supplementary food had been consumed, there was food available, though the birds had to work for that, practicing the same task as during experimental sessions. The birds were never food deprived. After experimental sessions, a basin, 85 cm × 55 cm was partly filled with water, allowing the birds to take a bath.

Apparatus

The program was written in Labview 6.1 and displayed stimuli on 17" (32.8 cm × 24.6 cm) monitors with screen resolution set to 1024 by 768 pixel. The visual cues in the experiment were 7 unique periodic backgrounds (see Figure 1 for examples), created with the taprats applet (<http://www.cgl.uwaterloo.ca/~csk/washington/taprats/>) and coloured in Adobe Photoshop, and 20 sample pictures measuring 2.7 cm × 2.7 cm. The pictures were of objects that the experimenters judged as distinct, such as marbles, a waterfall, a sunflower, a Scottish tartan, mushrooms, etc. We believe it is important that the birds should not reach good performance based on treating each combination of sequence and retrieval cue as a separate problem. Using 20 different pictures, we can generate $20 \times 19 \times 18 = 6840$ possible 3-item sequences of stimulus pictures and $20 \times 19 \times 18 \times 3 = 20520$ possible combinations of picture sequence and retrieval cue. If samples are identified by location

**Figure 1.**

Basic design of the “which of the N pictures did you see with this background?” (matching-to-background) and “which was N th?” (matching-to-sequence) procedures. In the sample phase, a series of N (here 3) pictures is presented. In matching-to-background, the pictures are shown against a series of backgrounds with a consistent order across all trials (except in the experiment with shuffled backgrounds). The order and identity of the pictures changes from one trial to the next. In the choice phase, all N sample pictures are presented, in their original spatial locations, against one of the backgrounds. The correct choice is the picture that, in the sample phase, was seen against the same background shown in the choice phase.

In matching-to-sequence, the backgrounds are omitted in the sample phase. The correct choice is the picture in the ordinal position signalled by the background. This is a version of symbolic matching-to-sample. It is possible to turn matching-to-sequence into a recall task by omitting the sample items in the choice phase. Then the task can only be solved by recalling the item locations.

We do not show here the actual sample pictures due to the relatively low resolution in this image. In the figure, we replaced the sample pictures by letters on white squares, with the squares drawn larger than to scale. The backgrounds are drawn to scale.

alone, and we assume the resolution of jackdaws' spatial memory is good enough that we have 20 distinct locations on the screen, we have the same number of spatial sequences and combinations of spatial sequence and retrieval cue. We think 20 different locations is a realistic estimate because coal tits and great tits were able to distinguish neighbouring locations in a 4×5 grid of positions on a smaller screen than used in this experiment (Biegler, McGregor, Krebs & Healy 2001). The pictures were randomly distributed over the screen. If we assume that each combination of picture and location is distinct, we have 46785600 possible item sequences, and 140356800 possible combinations of item sequences with cues for the correct choice. We did not explicitly arrange for trial-unique sequences, but believe repetitions are rare enough we can consider the sequences trial-unique.

General testing procedure

Matching-to-background

The experimental task combines elements of matching-to-sample, paired-associate learning, and sequence learning, and is modelled on the procedures of Shimp and Moffitt (1974) and Shimp (1976). In the sample phase of the matching-to-background version of the task, we displayed a series of pictures against a sequence of backgrounds. The pictures were randomly chosen from the set of 20, with the constraint that the programme avoided the pictures used in the last trial. The sequence of backgrounds was always the same in each trial (for an exception see the experiment with shuffled backgrounds). For example, in one trial picture F was shown against background 1, picture A against background 2, and picture C against background 3. In the choice phase, all three pictures were shown, on one of the backgrounds. The correct picture was the one that had been shown with the background present in the choice phase. In this case, if background 2 was shown, the correct choice would be picture A. In the next trial, the pictures might be N, C and P, but the backgrounds were again 1, 2 and 3. The locations of the pictures were chosen at random, with the constraint that the distance between the centres of two pictures was at least twice the size of a picture, i.e. at least 5.4 cm. In the choice phase, the pictures were in the same location as in the sample phase, so the correct item could be associated with its background either on the basis of location or visual features or both. This task is an extension of that of Shimp and Moffitt (1974).

Matching-to-sequence

In this version of the task, we omitted the backgrounds in the sample phase. Because each background retained its consistent position in the sequence, presentation of the background in the choice phase functioned to probe which item had been in the corresponding position. Omitting the backgrounds during the sample phase changed the meaning of the background shown in the choice phase from "which of the N pictures did you see with this background?" to "which of these N pictures did you see first (or second, or N th) in the sequence?" or for short, "Which was N th?" (see Figure 1). This task is an extension of that of Shimp (1976).

Training procedure

Habituation

To overcome the birds' neophobia, we gradually habituated them to the eight backgrounds by training them to peck at a single picture in the centre of the screen, displayed against a very dim version of the background (barely visible), then progressively brighter versions. When a bird was willing to peck at one brightness level, we went to the next of the five levels. Bird A needed 5 days to peck at pictures displayed against the brightest backgrounds, Bird S 63 days (this bird is consistently more neophobic towards new objects). Later, we inserted a day with this no-sample training each time before we increased the number of samples.

Pretraining

We originally intended to use matching-to-background only as a pretraining procedure because Shimp (1976) had used the same three pigeons for his experiment as Shimp and Moffitt (1974). Thus our plan was to use matching-to-background to establish the principle that a background signalled which of the pictures in a previously shown sequence was the correct choice. Matching-to-background should be relatively simple because the same stimulus picture and background can be paired in both sample phase and choice phase. We hoped the birds would register the consistent sequence of the backgrounds, so we could gradually transfer control over behaviour to sequence (see matching-to-sequence below). Therefore we trained the birds first on 2-item sequences in matching-to-background. The criterion for going from 2 items to 3 items was either at least 70% correct choices over three days or at least 75% correct over one day (chance level = 50% correct). The criterion for going to the next stage, when training with 3-item sequences was at least 60% of choices correct over three days or at least 67% correct over one day (chance level = 33.3% correct). The interstimulus interval (ISI) was four seconds, measured from the offset of a stimulus picture a bird had pecked at to the onset of the next picture. The retention interval (RI) was one second, measured from the offset of the last sample picture to the presentation of the choice. The intertrial interval (ITI) was 60 seconds.

We tried two ways of going from matching-to-background to matching-to-sequence. If the order of events were registered automatically (Gallistel, 1990), we expected that we could first leave out one randomly chosen background during the sample phase, then two, then three. That is the missing backgrounds method. The alternative fading backgrounds method reversed the initial habituation procedure, by using progressively dimmer backgrounds in the sample phase, while still showing the brightest backgrounds in the choice phase, until the backgrounds in the sample phase had faded away to nothing. Neither of these methods worked, so we switched to training matching-to-sequence first with a 2-item sequence, then a 3-item sequence.

Data recording and statistical procedure

The programme saved the identity of the sample pictures, the correct choice, the choice made by the bird, the latencies to peck at a sample, and the latency to make a choice.

We are primarily interested in performance as a function of the ordinal position of the correct choice.

The proportion of correct choices to an ordinal position in a sequence can be distorted by choice biases. Assume a bird always chooses the first of two samples. If the first sample is correct, the bird's choice is registered as correct. If the second sample is the correct choice, the bird's erroneous choice of the first sample is registered as an error for the second ordinal position. For this example case of complete bias, with the bird choosing only position 1, all successes are attributed to position 1. The failures are attributed to position 2, because a choice of position 1 is wrong when position 2 is correct, and the outcome is attributed to whichever ordinal position is correct in this trial. Therefore the proportion of correct choices when the first choice was correct would be 1, but that would not indicate the bird remembered anything, because the proportion correct would be equal to bias (the proportion of choices, correct or wrong, of that ordinal position), and the bias towards an ordinal position is the chance level for that position. Accordingly, we show both bias and proportion correct, we calculated statistical significance using the bias towards an ordinal position as the chance level for that position, but we also calculated a measure that is independent of bias.

We treated the choice of a stimulus at ordinal position N as a yes/no decision. Taking again the case of two samples, if the first picture is the correct choice, and the bird chooses that picture, that is a hit for position 1 and a correct rejection for position 2. Choosing the second picture would be a miss for position 1 and a false alarm for position 2. From that, we calculated d' for each ordinal position for a measure of performance independent from bias. In the case of 2-item sequences, the d' values for each ordinal position must be identical, because our calculation treats the decision as a discrimination of the two ordinal positions. For 3-item (or longer) sequences, the calculation treats each decision as a discrimination between the one ordinal position and all others. The values of d' are no longer constrained to be identical. The supplementary material contains a table showing example data and a calculation of d' .

For analysis of performance at each serial position within an experiment, we used the binomial test (one-tailed). We used a chi-square test to analyse detailed patterns of performance within conditions, i.e. the missing background condition and the shuffled background condition. In the other cases we calculated the d' for each day and position and used an ANOVA, with condition as independent factor, and d' as dependent factor. Comparison between experiments is based on the last days including all 9 RI-ISI combinations (see below) and the first days including all RI-ISI combinations. The performance of the two birds is analysed separately. Where appropriate, we used Bonferroni correction, i.e. for 3 positions $p(\text{significant}) \leq 0.017$.

Results and discussion

A) Pretraining, matching-to-background

The birds quickly learned to peck at the correct picture when matching-to-background in the 2 sample and 3 sample conditions. As soon as a bird reached criterion, it proceeded to the next stage. The ultimate aim was not to train them to associate backgrounds to pictures but to remember in which ordinal position a stimulus occurred. With Bird A, which was 58 days ahead in training, we tried two ways of going from matching-to-

background to matching-to-sequence, the missing background and fading background conditions.

B) Pretraining, missing backgrounds

In the sample phase, one randomly chosen background of the three was omitted, and the sample picture shown on an otherwise dark screen. The question was whether the bird could recall the missing background. A bird might do that during the sample phase, associating the sample picture with the background expected at this point in the sequence. Alternatively, a missing background could be recalled during the choice phase, when recalling the sample pictures in sequence. Finally, instead of recalling the background, a bird might reason by exclusion, avoiding a mismatch between the chosen picture and the background shown in the choice phase. That would exclude two of the sample pictures, leaving only the one shown without background.

Bird A's performance showed no sign of either a background recall or an exclusion strategy. When testing for memory of a background present in both the sample and choice phase, the proportion correct was 0.53, 0.65 and 0.77 for pos 1, pos 2 and pos 3, respectively, all $P < 0.0001$. But when shown a background in the choice phase that had not been present in the sample phase, the performance was at chance level (all binomial tests $P > 0.05$). Performance was significantly lower when a background was missing compared to when it had been presented in the sample phase (pos 1: $\chi^2_{.05,1} = 7.96$, $P = 0.005$; pos 2: $\chi^2_{.05,1} = 19.38$, $P = 0.000011$; pos 3: $\chi^2_{.05,1} = 35.97$, $P < 10^{-8}$).

C) Pretraining, fading of background

Another attempt was to make the backgrounds dimmer, i.e. reverse the habituation procedure, but only for the backgrounds in the sample phase. In the choice phase, the backgrounds remained at full brightness. When making the background dimmer, Bird A's proportion of correct choices dropped drastically, to 0.32 ($P = .06$) for position 1, 0.31 ($P = .25$) for position 2 and 0.45 ($P = .06$) for position 3.

Experiment 1: matching-to-sequence and matching-to-background with ordered sequence

Because missing and fading backgrounds failed to teach Bird A matching-to-sequence, we decided to try neither of these methods on Bird S. We trained both birds directly on a sample phase with no backgrounds, the matching-to-sequence task.

Methods

After 19 days (Bird A) and 18 days (Bird S) of matching-to-sequence without signs of improvement in performance, we changed the training procedure: a bird was consistently rewarded for choosing the first item until it had reached a criterion of at least 10 correct choices out of the last 12 choices. Then the reward contingency was reversed, and choosing the second item was rewarded until the bird again reached criterion, and so on. The birds received 10 days of training with this successive reversals procedure, after which the standard procedure was reinstated where the correct ordinal position was selected randomly in each trial.

Experiment 1 began after the 10 days with successive reversals, when the birds returned to 2-item matching-to-sequence, with the correct ordinal position chosen randomly. After reaching a criterion of either 70% correct choices on three successive days or 75% correct on one day, the birds started matching-to-sequence with 3-item sequences.

During the following 13 days with 3-item sequences, there was no sign of improvement for either bird, so we gave them 5 days of successive reversals, then went back to the standard procedure. At that point, we also varied retention interval (RI) and interstimulus interval (ISI) between days, in case there was a better combination than the 1 second RI and 4 second ISI we had used until then. The RI was one, four or eight second, and the ISI one, four or eight seconds, giving 9 combinations. The interval between trials (ITI) remained always at 60 seconds.

The chance level of success at an ordinal position is defined by the bird's bias towards that position (see Data recording and statistical procedure section).

Results and discussion

Both birds learned to match to sequence when the sequence was two items long. Bird S reached the criterion of 75% correct on the first day after the successive reversals training, but had a second day of training because of an experimenter's error. Bird A needed 12 days to reach criterion. Analysing the last two days of training for both birds, we found that Bird A reached 80% correct choices against a chance level of 58% for ordinal position 1 ($P < 10^{-7}$, $d' = 1.23$) and 65% correct against a chance level of 42% for ordinal position 2 ($P < 10^{-6}$, $d' = 1.23$). Bird S reached 75% against a chance level of 46% for ordinal position 1 ($P < 10^{-10}$, $d' = 1.56$) and 81% correct against a chance level of 54% for ordinal position 2 ($P < 10^{-11}$, $d' = 1.56$).

During the first two days of 3-item sequence training for Bird S, the third item was appended to the sequence after the second item. We found that when the new background 3 signaled which sample was correct in the choice phase, errors were more frequent to the second than to the first sample (binomial test $P < 10^{-5}$). For the backgrounds 1 and 2, the errors were evenly distributed towards the remaining two positions ($P = 0.46$ and $P = 0.31$ for backgrounds 1 and 2, respectively). The patterns of errors suggested that further learning would be easier if background 2 retained the status as last in the sequence and the new element was inserted between the first and last instead of appended at the end. When bird A started training with a 3-item sequence two days later, we inserted, right from the beginning, the extra item between the first and last item of the 2-item sequence.

Both birds were trained with the additional item inserted into the sequence for 13 days, during which there was no sign of performance improving. The birds received 5 days of reversal training, then were returned to the standard procedure in which the correct ordinal position was chosen randomly for each trial. We also started varying RI and ISI. We were interested in RI and ISI primarily to see whether there was a combination that would allow our birds to solve this task. Bird A, who showed an effect of RI and ISI combination ($F_{8,40} = 4.09$, $P < 0.01$, $\eta^2 = 0.45$), was the one that could solve the task with any of the combinations. Bird S, who showed no effect ($F_{8,40} = 1.56$, $P > 0.1$, $\eta^2 = 0.23$), could not solve the task with any combination. Therefore we will pool across RI/ISI combinations from here on.

Bird A's final overall performance in matching-to-sequence was above chance. However, more detailed analysis showed that the bird exceeded chance level only when

asked to identify the first ($P < 10^{-7}$, $d' = 0.5$) or the last sample ($P < 10^{-16}$, $d' = 0.76$). When asked to identify the second item, the bird was at chance ($P = 0.13$, $d' = 0.08$; see Figure 2). This pattern of performance is consistent with Terrace's (1993) proposal that animals may identify stimuli related to the beginning and end of a sequence without a representation of the whole sequence. If that learning is the basis of Bird A's performance in matching-to-sequence, we should expect significant improvement when the bird can identify the second sample by matching-to-background instead of matching-to-sequence. When the bird was switched back to matching-to-background, performance increased on ordinal position 1 ($F_{1,16} = 20.69$, $P = 0.0003$, $\eta^2 = 0.56$) and position 2 ($F_{1,16} = 15.74$, $P = 0.0011$, $\eta^2 = 0.50$), but not on position 3 ($F_{1,16} = 4.33$, $P = 0.054$, $\eta^2 = 0.21$).

Bird S failed to exceed chance level in any serial position (all $P > 0.017$). Bird S was above chance level in the first and second ordinal positions when switched back to matching-to-background (pos 1: $F_{1,16} = 8.34$, $P = 0.01$, $\eta^2 = 0.34$; pos 2: $F_{1,16} = 8.9$, $P = 0.009$, $\eta^2 = 0.36$; pos 3: $F_{1,16} = 6.9$, $P = 0.018$, $\eta^2 = 0.30$).

We originally planned to use matching-to-background merely as pretraining for matching-to-sequence, and hoped to proceed to the recall test version of matching-to-sequence (Figure 1). The matching-to-sequence task we used to test our birds is still a recognition task. Recognition performance is ambiguous as a measure of processes related to episodic memory, because it can be based on implicit familiarity judgments (Huppert & Piercy, 1978). Furthermore, Voss, Baym and Paller (2008) recently dissociated yes-no recognition and forced choice recognition. As expected for explicit memory, Voss et al. found that yes-no recognition in human subjects was impaired by encoding under divided attention, and performance was better when subjects were confident. In contrast, under time pressure, performance in forced choice recognition was better after encoding under divided

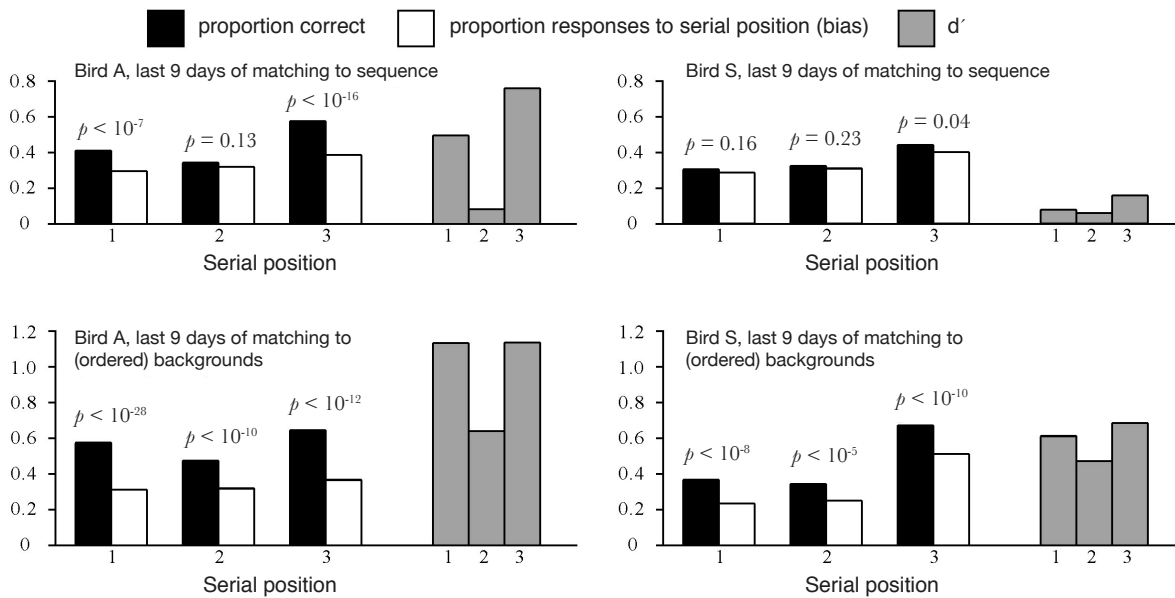


Figure 2: Performance as a function of ordinal position during the last days of matching-to-sequence (top row) and the first days of matching-to-background including all 9 RI-ISI combinations. The p values are the results of binomial tests comparing the proportion of correct choices against the bias, which defines chance level.

attention and when subjects felt they were guessing. Voss et al. argue that their data indicate the existence of a recognition mechanism based on perceptual fluency. It is unclear whether such a mechanism could support performance in our tasks. We did not impose any time pressure on our birds. We presented all sample pictures in both sample and choice phase, and the correct sample had to be identified by its association with either a background or a serial position. We do not know whether the perceptual fluency of combinations of stimuli can increase. Nevertheless, it seems advisable to use yes-no recognition or recall tasks when attempting to study processes related to episodic or episodic-like memory.

Unfortunately, our birds' failure to learn matching-to-sequence meant that instead of proceeding to the recall task, we tried to find out why they had failed. Did they learn nothing about ordinal position, or does the problem lie in our way of asking them what they know? If the latter, a different way of testing may indicate knowledge of ordinal position.

Experiment 2: Matching-to-background: transition from ordered to shuffled backgrounds

Methods

The subjects, apparatus and housing conditions were the same as in Experiment 1.

In Experiment 2, the birds first went back to 3-item matching-to-background with a consistent order of backgrounds, then the order of the backgrounds was randomly shuffled, so that only the background identity was informative. We predicted that if the birds had learned to associate each background with its ordinal position, then putting the ordinal position in one specific trial in conflict with the well-learned ordinal position should result in characteristic errors. If we call the standard sequence 1-2-3, and the shuffled sequence were 1-3-2, there should be no problem if we present background 1 in the choice phase, asking the birds to choose the first picture from the sequence. However, if we present background 3, which in this trial was presented *second*, any errors should be biased towards choosing the third ordinal position rather than the first. The supplementary material contains a table of the errors predicted if there is long term memory for the standard sequence.

Although Bird A only learned to identify the first and last sample in the matching-to-sequence procedure, and Bird S failed almost completely, it is conceivable that this was less a problem with the birds' ability to acquire this information, but more a problem with how we asked them. After all, both birds reached criterion with 2-item sequences. And when Bird S was transferred from 2-item to 3-item sequences in matching-to-sequence, the bird's preference for having the extra item inserted rather than appended suggested at least implicit knowledge of order. Therefore we wanted to see whether implicit knowledge of order might be demonstrated by the effects of conflicting information. After the birds were retrained on matching-to-background with the same sequence of backgrounds, consistent across trials, they were tested with the sequence of backgrounds during the sample phase being shuffled from one trial to the next.

Results and discussion

First, when we compared the last nine days with ordered backgrounds to the first nine days with shuffled backgrounds, we found that performance (measured as d' calculated for each day and ordinal position) decreased. More specifically, bird A performed worse in positions 1 ($F_{1,16} = 20.9, P = 0.0003, \eta^2 = 0.57$) and 3 ($F_{1,16} = 22.7, P = 0.0002, \eta^2 = 0.60$) but

not in position 2, ($F_{1,16} = 1.38$, $P = 0.26$, $\eta^2 = 0.02$). Bird S performed worse only in position 1 ($F_{1,16} = 8.97$, $P = 0.009$, $\eta^2 = 0.36$), but not in positions 2 ($F_{1,16} = 2.54$, $P = 0.13$, $\eta^2 = 0.14$) and position 3 ($F < 1$; see Figure 3). This performance decrement might have been due to the conflict between the memory of the ordered sequence and the now shuffled sequences. However, more detailed analyses failed to support this conjecture.

During the first nine days of training with shuffled backgrounds, performance with the original sequence was no better than with the five shuffled sequences (bird A: $\chi^2_{0.05,1} = 2.9$; $P = 0.09$, bird S: $\chi^2_{0.05,1} = 1.07$; $P = 0.3$). In case that analysis missed a short, transient effect, we also checked the first day only, finding again no better performance with the original order of backgrounds compared to the shuffled orders (bird A: $\chi^2_{0.05,1} = 0.03$; $P = 0.87$, bird S: $\chi^2_{0.05,1} = 0.12$; $P = 0.73$).

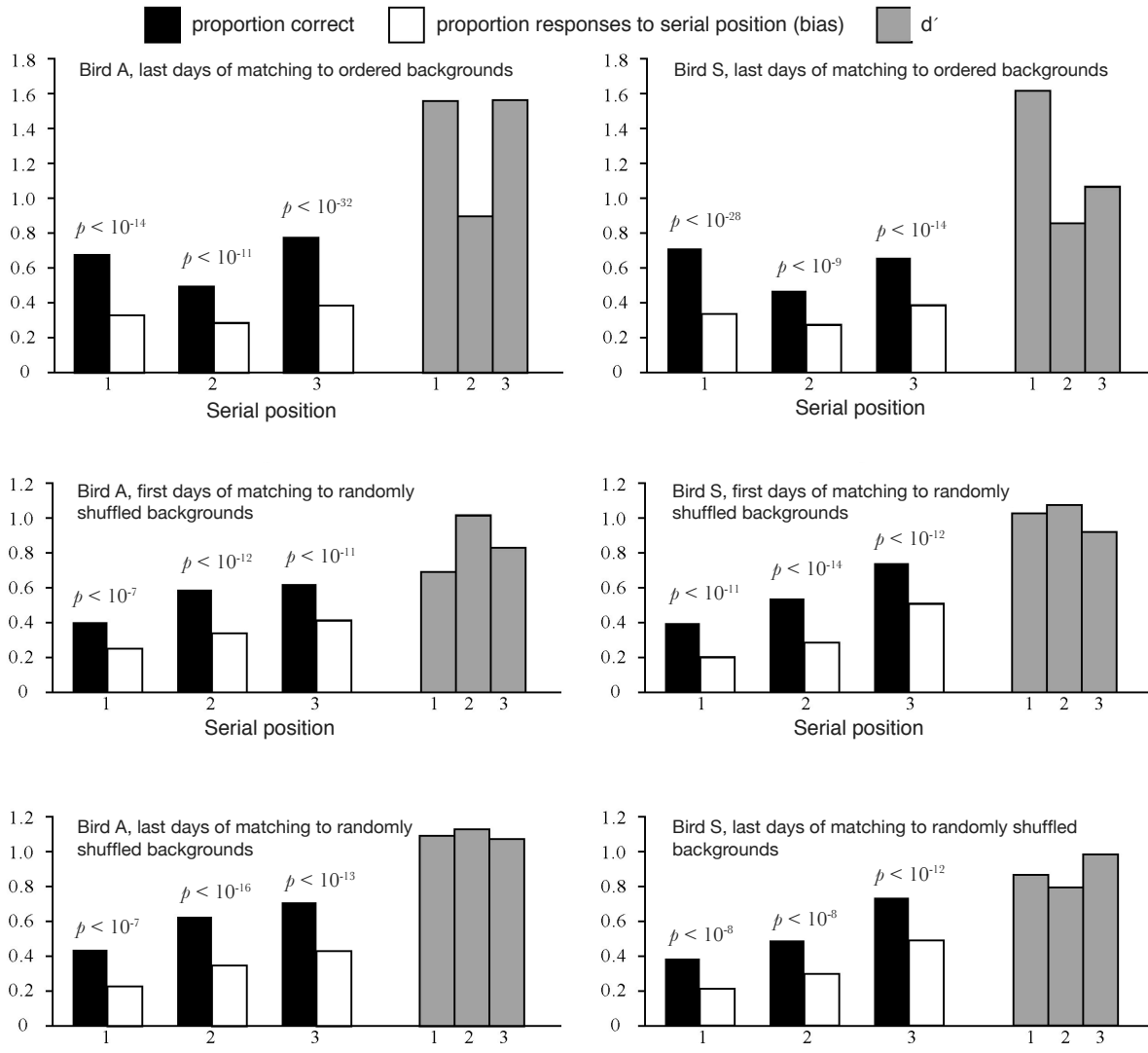


Figure 3: Performance as a function of ordinal position during the last days of matching-to-ordered backgrounds (top row), as well as the first and the last days of matching-to-randomly shuffled backgrounds. The p values are the results of binomial tests comparing the proportion of correct choices against the bias, which defines chance level. Performance with shuffled backgrounds quickly reached a steady state.

If there were implicit knowledge of sequence, then shuffling the sequence should create conflicts that lead to specific error patterns (see above and supplementary material). We found no evidence of any such biases in the errors. Analysis of error patterns during the first nine days showed for bird A 157 errors in the predicted direction out of a total of 371, less than half the total and therefore a bias opposite to that predicted. Bird S made 181 errors in the predicted direction out of a total of 375, showing again a bias opposite to that predicted. We also analysed the first day. Neither of the birds made significantly more errors consistent with the standard sequence than errors in conflict with the standard sequence (bird A: 18 out of 30, bird S: 18 out of 42, binomial test P 's > 0.1).

The performance drop when we shuffled the backgrounds suggests the birds noticed a change, specifically that the backgrounds related to the beginning and end of sequences had changed. The lack of the specific predicted error pattern indicates that they did not know the ordinal positions of the backgrounds.

Experiment 3: Matching-to-background with increasing sample size.

A difference in memory load is one conceivable reason for the birds' success in matching-to-background contrasting with their failure to match-to-sequence. In matching-to-background, a picture/background combination might be encoded as a single image, which in the choice phase can be directly matched to one sample item and its immediate surroundings. In contrast, matching-to-sequence could be solved by recalling the background during either sample or choice phase. If sample picture and background have to be encoded separately, that would double working memory load. Did the birds fail to match to sequence merely because they hit a working memory limit?

Methods

To address that question, in Experiment 3 we returned to “which of the N pictures did you see with this background?”, with a consistent sequence of backgrounds. However, whenever a bird reached an overall performance criterion, we increased the number of samples, until we reached seven samples. We report only the results from training with seven samples.

Results and discussion

In matching-to-background with seven samples, performance was above chance level for all positions for both birds (all P 's $< 10^{-11}$, binomial tests; see Figure 4). They have therefore enough working memory capacity to solve a 3-item matching-to-sequence task if that requires remembering backgrounds and samples separately.

General discussion

The matching-to-background and matching-to-sequence tasks are logically equivalent except that matching-to-sequence requires the birds to learn two additional pieces of information: they must associate sample items with ordinal positions during the sample phase. We are unable to draw any conclusions about that because the birds failed the other requirement, that they associate backgrounds with ordinal position throughout the experiment.

We found that our two jackdaws performed well on matching-to-background with up to 7 samples, but on matching to a 3-item sequence one bird's performance was entirely at chance, the other could identify only the first and last item of the sequence of pictures, but not the middle item. The birds' problem with matching-to-sequence is specific to the 3-item sequence. Both learned a 2-item sequence.

Terrace et al. (1996) argued that 2-item sequential problems can be solved through representations that do not have any information on the sequence of events, and that any test of sequence processing must use sequences at least three items long. The argument implies that Bird S either did not learn to remember trial-specific 3-item sequences for a few seconds, or that our way of testing their knowledge was unsuitable. Bird A learned to identify the beginning and end of a sequence, but the failure to identify the second item indicates that either the bird did not remember the whole sequence, or that the bird learned to remember cues associated with the beginning and end of a sequence, or that this method of testing was too difficult. We believe that Bird A relied on cues associated with the beginning and end of sequences because that bird's performance decreased only in the first and last position when we randomly shuffled the order of the backgrounds, thus destroying the previously reliable relationship between the beginning and end of the sequence and

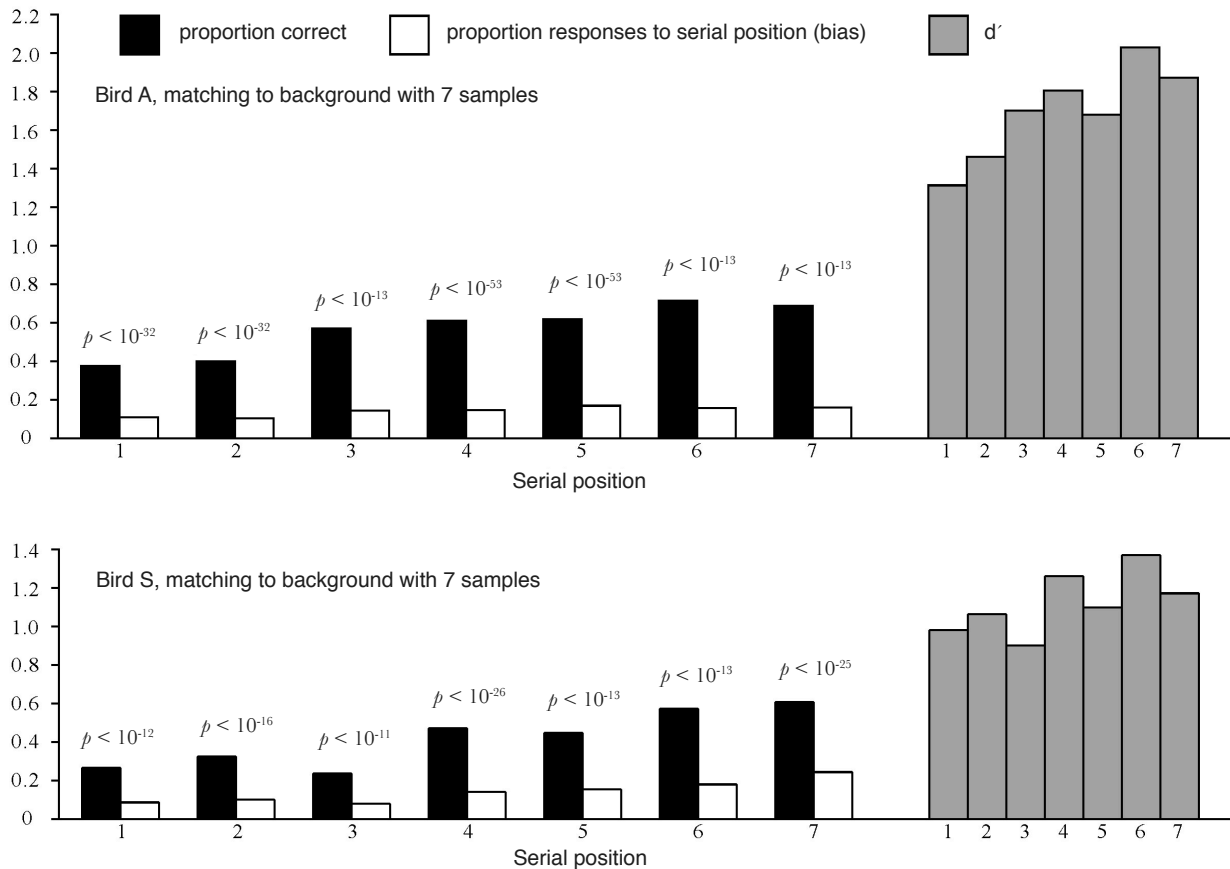


Figure 4: Performance as a function of ordinal position during the last days of matching-to-ordered backgrounds with seven samples. The p values are the results of binomial tests comparing the proportion of correct choices against the bias, which defines chance level.

the corresponding backgrounds. We cannot count Bird A's partial success in matching-to-sequence as evidence that he remembered sequences. Thus we conclude that both birds failed to learn to match to sequence.

The birds' success in matching-to-background shows that they can associate a sample picture with an external stimulus, and can identify the sample when presented with the external stimulus as a retrieval cue. The birds' difficulty in matching-to-sequence should therefore not lie in the association as such. Before examining failure modes, we should look at possible ways in which the birds might have solved the problem of matching-to-sequence, assuming that they know the fixed order of the backgrounds. That knowledge is always a necessary prerequisite. We can classify possible solutions according to whether the association between the sample item and its ordinal position is prospective (Roitblat, 1980) or retrospective, and whether the link between the ordinal positions of the sample item and the background is prospective or retrospective.

First, the birds might prospectively link the sample item to its ordinal position, then recall, still during the sample phase, the background that would be there in matching-to-sequence. They could then prospectively associate each sample item with the background that has the same ordinal position. They would have to create a separate item-background association for each sample item.

Second, the birds might prospectively associate each trial-specific sample item with ordinal position, and retrospectively, in the choice phase, use the ordinal position associated with a background as a retrieval cue for the sample item associated with the same ordinal position.

Third, on seeing the background that indicates the requested ordinal position, the birds might mentally go through the sequence of sample items until reaching the appropriate ordinal position, retrospectively link item and ordinal position, and retrospectively match the ordinal positions of background and sample item. That would certainly be mental time travel into the past, even if only for a few seconds. In that case, we would expect longer response latencies when asking for later ordinal positions (Guyla & Colombo, 2004).

The common requirement of knowing the ordinal position of the backgrounds did not seem so unlikely. In fact, the same birds later learned the ordinal positions of items in many three-item sequences (Pfuhl & Biegler, 2012). The idea that the birds would register the ordinal positions of the backgrounds without explicitly being tested on that information was the rationale for our missing backgrounds method for a gradual transition from matching-to-background to matching-to-sequence. The total failure of that method makes it unlikely there was any automatic encoding of ordinal position.

Shimp and Moffitt (1974) reported that pigeons could perform an equivalent of matching-to-background, and Shimp (1976) that they performed matching-to-sequence. Our jackdaws failed to replicate Shimp's (1976) result. However, we must examine the interpretation of Shimp's results. Shimp's interpretation depends on his pigeons having learned which colour signaled that the first, second or third ordinal position was the correct choice. For Shimp's interpretation to be true, the pigeons had to know that sequence. Shimp did not test whether his pigeons had learned the sequence. We tested whether our jackdaws had learned the equivalent sequence of backgrounds by testing how they responded to the backgrounds being shuffled in the matching-to-background task. That test revealed no knowledge of the sequence of the backgrounds. That lack of knowledge would prevent the birds from solving the task in any of the three ways we outlined above. Given the technical limitations of Shimp's apparatus and the resulting alternative strategies that contributed to

his three pigeons' performance (see introduction), we are not convinced that the ability of birds to report the first, second or third stimulus of a sequence has been established.

We believe we can discard two other possible explanations for our birds' failure to learn to match to sequence. There is the possibility that a picture/background combination might be encoded as a single image, but when mentally filling in the background, picture and background have to be encoded separately, doubling working memory load. The birds' success in matching-to-background with a 7-item sequence, more than doubling the working memory load compared to the 3-item sequence, rules out a simple working memory capacity limit as an explanation for the jackdaws' trouble with matching-to-sequence.

The other explanation we find implausible is the notion that the birds simply cannot represent any information about sequences. The extensive literature on sequence learning makes that unlikely (for example, Terrace 1993, Terrace et al. 1996, Bond, Kamil & Balda 2007, Pfuhl & Biegler, 2012).

The finding that our jackdaws performed well in matching-to-background with up to 7 samples persuades us that our procedure was not just so much more difficult, or failed to motivate the jackdaws, that this could explain the difference. Shimp (1976) argued that it was unlikely that the pigeons had learned each of the 24 possible combinations of sequence and choice cue as a separate task. Since then, it has been shown that this amount of information is within the capacity of pigeons' memory (Vaughan & Greene, 1984; Fagot & Cook, 2006). We speculate that the technical limitations of Shimp's (1976) apparatus may have allowed the pigeons to treat each of the 24 combinations as a separate problem, and that this accounts for performance beyond that supported by the pigeons choosing the side where a stimulus had been shown most recently or most often. If the performance of Shimp's pigeons was based on these alternative strategies, then the performance of our jackdaws may be more representative of birds' capacity in matching-to-sequence. If, in turn, that is true, it should be interesting to repeat our version of Shimp's experiment in a species like the jays that showed episodic-like memory and future planning beyond the current need state (Clayton & Dickinson, 1998; Raby, Alexis, Dickinson & Clayton, 2007). If these birds, also in the corvid family, showed the same failure, that would give some indications about which presumed features of episodic memory are *not* found, and so give more information about the possible architectures of memory systems.

Acknowledgements

G. P. was funded by the Norwegian Research Council.

We thank T. F. Ramstad and R. J. Sutherland for valuable comments on a previous draft.

A	B	C	D	E	F	G	H	I	J	K
	Target	Choice		Hit = 1	Hit = 1	Hit = 1		Corr. Rej. = 1	Corr. Rej. = 1	Corr. Rej. = 1
	serial	serial		False alarms = 0	False alarms = 0	False alarms = 0		Miss = 0	Miss = 0	Miss = 0
Trial	pos	pos		pos1	pos2	pos3		pos1	pos2	pos3
1	1	1		1					1	1
2	3	2			0			1		0
3	3	3				1		1	1	
4	1	1		1					1	1
5	3	1		0					1	0
6	1	3				0		0	1	
7	2	1		0					0	1
8	2	2			1			1		1
9	3	3				1		1	1	
10	3	3				1		1	1	
11	3	1		0					1	0

Hits	2	1	3	Corr. Rejections	5	8	4
False alarms	3	1	1	Misses	1	1	3
H	0,667	0,500	0,500				
F	0,375	0,111	0,200				
z(H)	0,431	0	0				
z(F)	-0,319	-1,221	-0,842				
d'	0,749	1,221	0,842				

Supplementary Table 1.

Calculation of d' , example data from 11 trials of Bird S. Taking trial 1 as the first example, we see that the correct choice was in serial position 1 (column B), the bird chose serial position 1 (column C). That meant a hit for serial position 1 (column E) and correct rejections for serial positions 2 and 3 (columns J and K). In trial 2, the correct choice was in serial position 3, but the bird chose the picture that had been presented in serial position 2. That meant a false alarm for serial position 2 (column F), a correct rejection for serial position 1 (column I) and a miss for serial position 3 (column K).

We then count hits, misses, correct rejections and false alarms for each serial position. H is then the number of hits divided by the sum of hits and misses. F is the number of false alarms divided by the sum of false alarms and correct rejections. $z(H)$ is the z-score of H and likewise $z(F)$. $d' = z(H) - z(F)$. We calculate d' for each serial position.

Examples Sequence Target bold	Error consistent with using sequence	Other error
1-2-3	Not applicable	Not applicable
1- 2 -3	Not applicable	Not applicable
1-2- 3	Not applicable	Not applicable
1-3-2	Not applicable	Not applicable
1- 3 -2	position 3, background 2	position 1, background 3
1-3- 2	position 2, background 3	position 1, background 1
2 -1-3	position 2, background 1	position 3, background 3
2- 1 -3	position 1, background 2	position 3, background 3
2-1- 3	Not applicable	Not applicable
2 -3-1	position 2, background 3	position 3, background 1
2- 3 -1	position 3, background 1	position 1, background 2
2-3- 1	position 1, background 2	position 2, background 3
3 -1-2	position 3, background 2	position 2, background 1
3- 1 -2	position 1, background 3	position 3, background 2
3-1- 2	position 2, background 1	position 1, background 3
3 -2-1	position 3, background 1	position 2, background 2
3- 2 -1	Not applicable	Not applicable
3-2- 1	position 1, background 3	position 2, background 2

Supplementary Table 2.

If we define the consistent sequence of backgrounds in Experiment 1 as 1-2-3, and if we assume that the birds have learned that sequence, then randomly shuffling the sequence should lead to distinguishable and predictable errors in $\frac{2}{3}$ of trials. If a bird makes an error, is the error biased towards the serial position that used to be signaled by the background shown in the choice phase?

For a specific example, take the last sequence, 3-2-1. The background signaling the correct sample item is background 1. Here it is in the third ordinal position, but it used to be in the first. If the bird makes an error, the error is consistent with previous sequence training if the bird chooses the sample item that had been presented in the first ordinal position, which here was associated with background 3.

References

- Agster, K. L., Fortin, N. J., & Eichenbaum, H. (2002). The hippocampus and disambiguation of overlapping sequences. *Journal of Neuroscience*, 22(13), 5760-5768.
- Babb, S. J., & Crystal, J. D. (2005). Discrimination of what, when, and where: Implications for episodic-like memory in rats. *Learning and Motivation*, 36(2), 177-189.
- Bekerian, D. A., & Bowers, J. M. (1983). Eye witness testimony: Were we misled? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 9(1), 139-145.
- Biegler, R., McGregor, A., Krebs, J. R. & Healy S. D. (2001). A larger hippocampus is associated with longer-lasting spatial memory. *Proceedings of the National Academy of Sciences of the United States of America*, 98(12), 6941-6944
- Bird, L. R., Roberts, W. A., Abrams, B., Kit, K. A., & Crupi, C. (2003). Spatial memory for food hidden by rats (*Rattus norvegicus*) on the radial maze: Studies of memory for where, what, and when. *Journal of Comparative Psychology*, 117(2), 176-187.
- Bond A. B., Kamil A. C., & Balda R. P. (2007). Serial reversal learning and the evolution of behavioral flexibility in three species of North American Corvids (*Gymnorhinus cyanocephalus*, *Nucifraga columbiana*, *Aphelocoma californica*). *Journal of Comparative Psychology*, 121(4), 372-379.
- Burt, C. D. B., Watt, S. C., Mitchell, D. A., & Conway, M. A. (1998). Retrieving the sequence of autobiographical event components. *Applied Cognitive Psychology*, 12(4), 321-338.
- Clayton N. S., & Dickinson A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, 395, 272-274
- Eacott, M. J., & Norman, G. (2004). Integrated memory for object, place, and context in rats: A possible model of episodic-like memory? *Journal of Neuroscience*, 24(8), 1948-1953.
- Fagot J., & Cook R. G. (2006). Evidence for large long-term memory capacities in baboons and pigeons and its implications for learning and the evolution of cognition. *Proceedings of the National Academy of Sciences of the United States of America*, 103(46), 17564-17567
- Ferbinteanu, J., Kennedy, P. J., & Shapiro, M. L. (2006). Episodic memory – From brain to mind. *Hippocampus*, 16(9), 691-703.
- Ferkin, M. H., Combs, A., del Barco-Trillo, J., Pierce, A. A., & Franklin, S. (2008). Meadow voles, *Microtus pennsylvanicus*, have the capacity to recall the “what”, “where”, and “when” of a single past event. *Animal Cognition*, 11(1), 147-159.
- Friedman W. J. (1993). Memory for the time of past events. *Psychological Bulletin*. 113(1), 44-66.
- Gallistel C. R. (1990). *The Organization of Learning*. Cambridge, MA: Bradford Books/MIT Press
- Guyla, M., & Colombo, M. (2004). The ontogeny of serial-order behavior in humans (*Homo sapiens*): *Representation of a list*. *Journal of Comparative Psychology*, 118(1), 71-81.
- Hampton, R. R., Hampstead, B. M., & Murray, E. A. (2005). Rhesus monkeys (*Macaca mulatta*) demonstrate robust memory for what and where, but not when, in an open-field test of memory. *Learning and Motivation*, 36(2), 245-259.
- Hannesson, D. K., Vacca, G., Howland, J. G., & Phillips, A. G. (2004). Medial prefrontal cortex is involved in spatial temporal order memory but not spatial recognition memory in tests relying on spontaneous exploration in rats. *Behavioural Brain Research*, 53(1), 273-285.
- Hoge J., & Kesner R. P. (2007). Role of CA3 and CA1 subregions of the dorsal hippocampus on temporal processing of objects. *Neurobiology of Learning and Memory*, 88, 225-231.
- Huppert, F. A., & Piercy, M. (1978). The role of trace strength in recency and frequency judgments by amnesic and control subjects. *The Quarterly Journal of Experimental Psychology*, 30, 346-354.
- Kart-Teke, E., Silva, M., Huston, J. P., & Dere, E. (2006). Wistar rats show episodic-like memory for unique experiences. *Neurobiology of Learning and Memory*, 85(2), 173-182.
- Maki W. S., Beatty, W. W., & Clouse B. A. (1984). Item and order information in spatial memory. *Journal of Experimental Psychology - Animal Behaviour Processes*, 10(4), 437-452
- Manns, J. R., Howard, M. W., & Eichenbaum, H. (2007). Gradual changes in hippocampal activity support remembering the order of events. *Neuron*, 56(3), 530-540.
- Orlov, T., Yakovlev, V., Hochstein, S., & Zohary, E. (2000) Macaque monkeys categorize images by their ordinal number. *Nature* 404(6773), 77-80.
- Patterson, N., Richter, D. J., Gnerre, S., Lander, E. S., & Reich, D. (2006). Genetic evidence for complex speciation of humans and chimpanzees. *Nature*, 441(7097), 1103-1108.

- Pfuhl, G., & Biegler, R. (2012). Ordinality and novel sequence learning in jackdaws. *Animal Cognition*. Online available DOI: 10.1007/s10071-012-0509-7
- Raby, C. R., Alexis, D. M., Dickinson, A., & Clayton, N. S. (2007). Planning for the future by western scrub-jays. *Nature*, 445(7130), 919-921.
- Roberts W. A., Feeney M. C., MacPherson K., Petter M., McMillan N., & Musolino E. (2008). Episodic-like memory in rats: is it based on when or how long ago?, *Science*, 320, 113-115
- Roitblat H. L. (1980). Codes and coding processes in pigeon short-term memory. *Behavioural Brain Sciences*, 5, 353-406
- Schwartz B. L., Colon M. R., Sanchez I. C., Rodriguez I A., & Evans S. (2002). Single-trial learning of “what” and “who” information in a gorilla (*Gorilla gorilla gorilla*): implications for episodic memory. *Animal Cognition*, 5, 85-90
- Shimp, C. P. (1976). Short-Term-Memory in Pigeon – Relative Recency. *Journal of the Experimental Analysis of Behavior*, 25(1), 55-61.
- Shimp, C. P., & Moffitt, M. (1974). Short-Term-Memory in Pigeon – Stimulus-Response Associations. *Journal of the Experimental Analysis of Behavior*, 22(3), 507-512.
- Terrace H. S. (1993). The phylogeny and ontogeny of serial memory: list learning by pigeons and monkeys. *Psychological Science*, 4(3), 162-170
- Terrace H. S., Chen S., & Jaswal V. (1996). Recall of three-item sequences by pigeons. *Animal Learning and Behavior*, 24(2), 193-205
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving and W. Donaldson (Eds.) *Organisation of Memory*, Academic Press, London
- Vaughan, W., & Greene, S. L. (1984). Pigeon Visual Memory Capacity. *Journal of Experimental Psychology – Animal Behavior Processes*, 10(2), 256-271.
- Worden, R. P. (1995). A Speed Limit for Evolution. *Journal of Theoretical Biology*, 176(1), 137-152.
- Voss, J. L., Baym, C. L., & Paller, K. A. (2008). Accurate forced-choice recognition without awareness of memory retrieval. *Learning & Memory*, 15, 454-459.

Received: 1.10.2012

Revised: 5.18.2012

Accepted: 5.19.2012

