CrossMark

RESEARCH ARTICLE

Neurons in the crow nidopallium caudolaterale encode varying durations of visual working memory periods

Konstantin Hartmann[1] · Lena Veit[1] · Andreas Nieder[1]

Abstract Adaptive sequential behaviors rely on the bridging and integration of temporally separate information for the realization of prospective goals. Corvids' remarkable behavioral flexibility is thought to depend on the workings of the nidopallium caudolaterale (NCL), a high-level avian associative forebrain area. We trained carrion crows to remember visual items for three alternating delay durations in a delayed match-to-sample task and recorded single-unit activity from the NCL. Sample-selective delay activity, a correlate of visual working memory, was maintained throughout different working memory durations. Delay responses remained selective for the same preferred sample item across blocks with different delay durations. However, selectivity strength decreased with increasing delay durations, mirroring worsened behavioral performance with longer memory delays. Behavioral relevance of delay activity was further evidenced by reduced encoding of the preferred sample item during error trials. In addition, NCL neurons adapted their time-dependent discharges to blocks of different memory durations, so that delay duration could be successfully classified based on population activity a few trials after the delay duration switched. Therefore, NCL neurons not only maintain information from individual trials, but also keep track of the duration for which this information is needed in the context of the task. These results strengthen the role of corvid NCL in maintaining working memory for flexible control of temporally extended goal-directed behavior.

Keywords Delayed match-to-sample · Corvid · Avian endbrain · Single-cell recording · Songbird

Introduction

The temporal integration of information for the attainment of prospective goals is important for adaptive behavior. Flexible sequential behaviors require bridging of temporally separate events and task components. The telencephalic nidopallium caudolaterale (NCL) is an avian associative forebrain area thought to enable corvids' remarkable behavioral flexibility (Veit and Nieder 2013; Moll and Nieder 2014). The NCL is considered to be a functional equivalent of primate prefrontal cortex (PFC) (Nieder 2017). This analogy is based on the NCL's dopaminergic innervation (Divac et al. 1985; Durstewitz et al. 1999), its connectivity (Güntürkün 2005), lesion studies (Mogensen and Divac 1993; Hartmann and Güntürkün 1998), and its activity during working memory (Diekamp et al. 2002; Veit et al. 2014), decision (Lengersdorf et al. 2014; Veit et al. 2017a, b), learning (Veit et al. 2015, 2017a, b), and association tasks (Moll and Nieder 2015, 2017).

Recently, we have demonstrated that NCL neurons encode information about visual stimuli categorically (Ditz and Nieder 2015, 2016a, b; Wagener and Nieder 2017) and temporarily maintain this information after the stimulus disappeared by sustained delay activity (Veit et al. 2014). Such sustained neuronal activity throughout the delay period is a hallmark of working memory and reminiscent of PFC neurons in primates (Fuster and Alexander 1971; Goldman-Rakic 1995; Miller et al. 1996;

Konstantin Hartmann and Lena Veit have contributed equally to this work.

✉ Andreas Nieder
andreas.nieder@uni-tuebingen.de

[1] Animal Physiology, Institute of Neurobiology, University of Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany

Miller and Cohen 2001; Fuster 2008; Nieder 2016). This sustained (or persistent) activity has been shown to often vary systematically with time during delay periods (Fuster 1973). It has, therefore, been suggested that these discharges constitute a potential timing signal that would allow decoding of temporal information (Brody et al. 2003). Indeed, several neurophysiological studies indicated coding of temporal information, such as forthcoming stimuli (Fuster 2001; Goldman-Rakic 1995; Miller et al. 1996; Rainer et al. 1999), motor timing (Niki and Watanabe 1979), anticipation of reward (Roesch and Olson 2005; Schultz 2006), or durations of temporal intervals (Sakurai et al. 2004; Genovesio et al. 2006). In addition, in pigeons, NCL activity has been indicated in the temporal organization of choice behavior (Kalenscher et al. 2005, 2006).

The goal of this study was to investigate whether single neurons of the corvid NCL carry information about remembered content across varying working memory delay durations. Moreover, we examined whether the neuronal population is adapting its discharge rate to the changing durations of the delay itself, keeping track of the structure of the task that the crow is currently engaged in. We, therefore, investigated the single-neuron activity in the NCL as a function of delay duration in behaving crows. We used a delayed match-to-sample task (DMS) in which crows were cued with a sample object and then, after a delay of variable duration, selected that object from a display. To solve the task, the sensory properties of a sample needed to be maintained throughout the delay. We studied the impact of changing delay durations on sample-selective persistent activity and explored a potential time-dependent neuronal signature in the delay activity.

Materials and methods

Subjects

Two hand-raised, adult carrion crows (*Corvus corone corone*), one male (crow B) and one female (crow I), were used in these experiments. All crows were obtained from the institute's breeding facilities. The birds were housed in social groups in spacious indoor aviaries. They were maintained on a controlled feeding protocol during the sessions, and earned food during and after the daily tests. All procedures were carried out according to the guidelines for animal experimentation and approved by the responsible national authorities, the Regierungspräsidium Tübingen. For details on the birds' housing and diet, see Hoffmann et al. (2011).

Apparatus

The crows were attached to a wooden perch by a leather jess and placed in an operant conditioning chamber in front of a touchscreen monitor (3M Microtouch, 15″, 60 Hz refresh rate). All stimuli were displayed on this monitor. Reward [birdseed pellets or mealworms (Tenebrio molitor larvae)] was delivered by a custom-built automated feeder below the screen. The CORTEX program (National Institute of Mental Health) was used for experimental control and behavioral data acquisition. An infrared light barrier, in combination with a reflector attached to the bird's head, registered when the bird was positioned in front of the screen and facing it (see details of apparatus in Brecht et al. 2016).

Behavioral protocol

They were trained on a visual delayed match-to-sample (DMS) task with variable delay times. The crows initiated a trial (Fig. 1a) by moving their head into the light barrier when a go-stimulus (white square, 11×11 mm) was shown on the screen. The crows had to keep their head still throughout the trial; if they moved their head prior to the response period (as detected by the light barrier), the trial was aborted. After 200 ms, the go-stimulus turned off, followed by a 500 ms presample period without any stimulus on the screen. The sample stimulus was presented in the center of the screen for 500 ms and was randomly chosen from a set of three photographs (2×2 cm), which were exchanged every day. Following the sample, the screen remained black during the delay, in which the bird had to remember the sample stimulus. The length of the delay varied in a prefixed order, depending on the current block. We used three different delays, a short delay with 1500 ms, intermediate delay with 3000 ms, and a long delay with 4500 ms delay time (Fig. 1b). Each block consisted of 60 trials with the same delay length. The order of the blocks was always the same in a given session, starting with the short delay, followed by the intermediate and finally the long delay. From the long delay, it went back to the short, then back to the long delay and decreasing from there over the intermediate to the short delay (Fig. 1b).

The choice period consisted of two sample images (one match and one nonmatch, 2×2 cm each) displaced 6.6 cm apart to the left and right of the center of the screen. The location of the match image was randomized and balanced in both tasks. The birds indicated their choices by pecking at the appropriate stimulus. If their choice was correct, the automated feeder delivered feedback via light and sound, followed by an inter-trial interval of 1500 ms. The male crow B was rewarded with food for each correct trial. The female crow I received food reward randomly for approximately 2/3 of correct trials. The probabilistic reward schedule increased
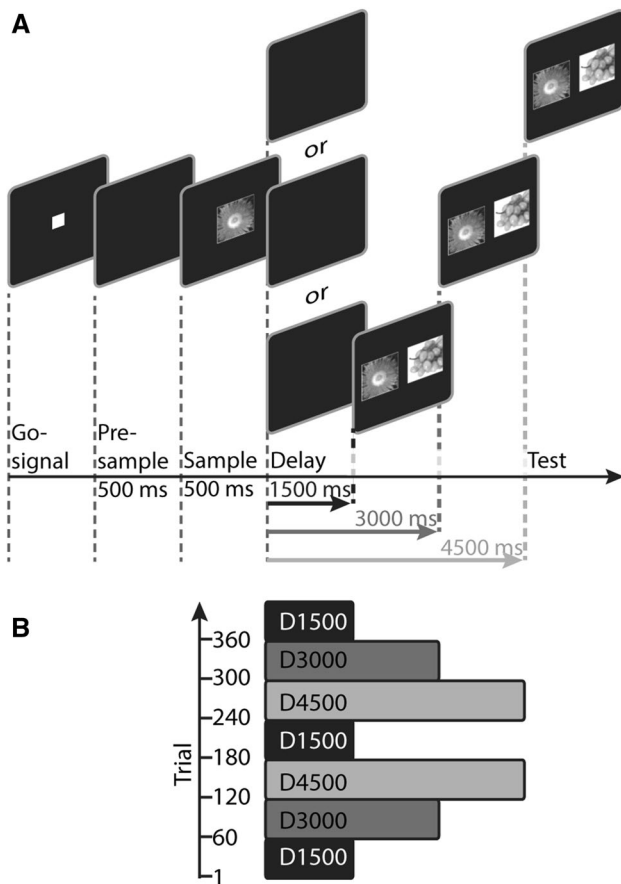
Fig. 1 Task protocol and stimuli. **a** Crow initiated a trial by keeping its head still in front of the monitor (automatically detected) to activate a go-stimulus. After a 500 ms presample period, a sample stimulus was presented for 500 ms, followed by a delay. Three blocks of delay times were applied. In the short delay, the crow had to wait for 1500 ms, in the intermediate delay for 3000 ms, and in the long delay for 4500 ms, respectively. In the choice period, the bird had to chose the image identical to the sample item from two items to receive a reward. All relevant task parameters were balanced. **b** Order of blocks with different delay durations. Each block consisted of 60 trials with a fixed delay. The order of delays was fixed and all possible transitions between delay lengths were presented each session

the total amount of trials performed while preventing reward expectation for specific delay duration or sample items. If the birds chose incorrectly, the trial was aborted and a short time-out (3 s) was presented before the start of the next trial. If no response occurred within 1800 ms, the trial was dismissed. All relevant task parameters were balanced.

Surgery and recordings

All surgeries were performed under sterile conditions, while the animals were under general anesthesia. Crows were anaesthetized with a ketamine/rompum mixture (50 mg ketamine, 5 mg xylazine/kg initially and supplemented by hourly 17 mg ketamine, and 1.7 mg xylazine/kg i.m.). The head

was placed in the stereotaxic holder that was customized for crows with the anterior fixation point (i.e., beak bar position) 45° below the horizontal axis of the instrument (Karten and Hodos 1967). Using stereotaxic coordinates (center of craniotomy: AP 5 mm; ML 13 mm), we chronically implanted two microdrives with four electrodes each, a connector for the headstage, and a small headpost to hold the reflector for the light barrier. After the surgery, the crows received analgesics (Butorphanol (Morphasol®), 1 mg/kg i.m.).

We recorded from eight chronically implanted microelectrodes on two custom-built microdrives, which were implanted in the left hemisphere in bird I and in the right hemisphere in bird B. We used glass-coated tungsten microelectrodes with 2 MΩ impedance (Alpha Omega LTD, Israel). The electrodes targeted the nidopallium caudolaterale (NCL). For details on the coordinates and electrode tracks, see Veit et al. (2014). We recorded a total of 103 neurons from both crows that could be analyzed, because they satisfied the inclusion criteria (see below).

At the start of each recording session, the electrodes were advanced manually until a high-quality neuronal signal was detected on at least one of the channels of each microdrive. Neurons were not preselected for involvement in the task. Each microdrive had a range of approximately 5 mm which was exploited to record from the NCL across different depths over a period of several weeks. Signal amplification, filtering, and digitizing of spike waveforms was accomplished using the Plexon system (Dallas, TX, USA). For each recording session, the birds were placed in the recording setup, a headstage containing an amplifier was plugged into the connector implanted on the bird's head and connected to a second amplifier/filter and the Plexon MAP box outside the setup by a cable above and behind the bird's head (all components by Plexon, Dallas, TX, USA). Spike sorting into single-unit waveforms was performed manually offline using the Plexon system. Each recording session lasted between 200 and 500 correct trials in approximately 2 h.

Data analysis

For behavioral analysis, we included all data collected during neuronal recordings. Therefore, the behavioral results reflect the performance during the recordings. Behavioral parameters encompassed percent correct responses and reaction times.

For neuronal analyses, we included all neurons that were recorded for at least three different delay blocks. A block was only considered if the cell was recorded for at least 16 trials. Only neurons that had a firing rate of at least 0.5 Hz during the combined presample, sample, and delay periods were considered. To account for delay length adjustments of the crows after a delay block switch, the first six trials of each block were omitted in analyses of sample selectivity.

To test for image selectivity in the memory delay phases, we used the firing rates in the last two-thirds of the delay period, i.e., a time window that scaled relative to delay duration to obtain a maximum of useful data. This resulted in an analysis windows of 1000, 2000, and 3000 ms for the 1500, 3000, and 4500 ms delay periods, respectively. A two-factor ANOVA with the factors "delay length" (three modifications) and "sample item" (three modifications) was used to determine whether the discharge rates for the three different sample images differed significantly ($p < 0.05$). Delay duration selectivity was analyzed accordingly. All neurons that discriminated between the sample items in the delay periods were additionally analyzed for image selectivity in the sample period. To that aim, the firing rate in an analysis window of 500 ms was analyzed, which was shifted by 100 ms from sample onset to account for the visual latency (Veit et al. 2014).

To construct tuning curves, the firing rates for each neuron were normalized to units of standard deviation from baseline. All selective neurons were included in the analysis of firing rates in error trials. Firing rates during correct and error trials were normalized to standard deviation from the baseline during correct trials before pooling. Statistical comparisons between the discharge rates in correct and error trials were done with the Wilcoxon sign rank test.

Decoding of block identity

We performed a decoding analysis using a k-Nearest-Neighbor classifier (Cover and Hart 1967; Veit and Nieder 2013) to determine whether information about block identity was present in the neuronal population activity patterns in different time windows throughout the task. A classifier uses "training data", i.e., a set of labeled data (in this case a set of trials with known block identities) to learn which features in the data can correctly identify the label. It is then applied to a set of new data, the "test data" (i.e. trials with block identity unknown to the classifier). The classifier's estimate of the class of the test data (i.e., block identity of the test trials) is then compared to their true classes to calculate the performance of the classifier. The k-Nearest-Neighbor algorithm classifies each trial based on the class of its closest neighboring trials in the feature space of the firing rates of all recorded neurons. Therefore, it uses actual examples from the training data and does not make assumptions about the underlying distribution of data points. We used all cells that were recorded twice for each delay duration and had at least ten trials for each sample in each delay block. The first six trials of each block were excluded for these 45 neurons. The firing rates for each neuron were normalized to units of standard deviations from baseline. We then created a pseudo-simultaneous population (i.e., combined trials from neurons which were recorded in separate sessions) by

combining trials with the same sample picture and block identity. Thus, we conserved information about sample picture, delay duration, and block order in the pseudo-population, while losing information on trial order within the remaining 54 trials of each block (as sample pictures were presented in a random order within each block).

We wanted to rule out misleadingly good decoding performance from changing neuronal baseline firing rates, either through recording quality or changes in neuronal response properties over time without specific connection to block transitions. Therefore, the classifier was trained with the first repetition of each block and was tested on the second repetition, to get a very conservative estimate for the amount of information about delay duration contained in neuronal firing rates. Specifically, we selected neurons which were recorded from block number 2 through block number 7, and thus experienced delay durations in the following order: (short)–middle–long–short–long–middle–short. Thus, to evaluate decoding performance for the long block, training data were taken from the last 54 trials of block three, and test data from the last 54 trials of block five. Each trial in the test set was classified as either long, middle, or short delay depending on the majority vote of its $k = 25$ nearest neighbors in the test set. Nearest neighbors were those trials with the smallest Euclidean distance in the high-dimensional space spanned by the firing rates of all neurons.

For the permutation test, we repeated the procedure with shuffled labels 1000 times. This classifier was used on seven time windows throughout the task, each 500 ms long. The first window started 1000 ms before the presample phase started (ITI early time −2000 to −1500), followed by a second window 500 ms before the presample (ITI late time −1500 to −1000). The presample (time −1000 to −500) and sample windows (time −500 to 0) started at the onset of the presample and sample period, and covered the entire period. The delay was divided into three windows, one starting with the delay phase (early delay time 0 to 500), the second 500 ms after the start of the delay (delay middle time 500 to 1000), and the last 1000 ms after the start of the delay phase (delay late time 1000 to 1500). All analyses were performed using the MATLAB statistics toolbox.

Block transition analysis

For the block transition analysis, we focused on the transition from the longest to the shortest delay and vice versa, since differences in population activity and decoding performance were greatest for these transitions. For the transition from long delay to short delay (Fig. 7a, b), we selected all neurons which were recorded for the entire 60 trials of each block (trial 120–240, Fig. 1a, $n = 88$). For the transition from short delay to long delay (Fig. 7c), we selected all neurons which were recorded for the entire 60 trials of each block

(trial 180–300, Fig. 1a, $n = 83$). Neuronal discharge rates in a window 1000–1500 ms after delay onset (late delay window in Fig. 6e) were normalized to units of standard deviation from baseline (see above) and principal component analysis (PCA) was applied. We illustrate the population discharge pattern using the first three principal components. The ellipsoids in Fig. 7a were generated using the mean and standard deviation of the last 40 trials of each block (trial 141–180 for the yellow ellipsoid and trial 201–240 for the blue ellipsoid). The first 20 trials of the short block (trial 181–200) were smoothed with a three trial boxcar window and color-coded according to trial order.

In addition, we performed a decoding analysis to determine whether the discharge patterns during transitions resembled patterns in the long or short delay blocks. For this analysis, a pseudo-simultaneous population was created by preserving information about trial and block order, but disregarding information on sample identity. In other words, we simply aligned blocks three through five (long–short–long, trials 120–300) for all neurons which were recorded in these blocks in separate sessions. We used the middle 40 trials of each block as training set to train the classifier, and the first 10 trials and last 10 trials of each block as test set. Each test trial was classified as either long delay or short delay depending on the majority vote of its $k = 25$ nearest neighbors in the training set. Note that decoding performance in Fig. 7b (95%) and Fig. 7c (77.5%) is better than the 56% reported for the late delay window in Fig. 6e for several reasons: Chance level for the transition analysis is 50%, as we are classifying each trial as either long or short delay, ignoring the middle delay which is most difficult to classify. In addition, Fig. 6e provided a very conservative estimate of decoding performance as training and test data were taken from separate blocks. For the transition analysis, the training and test data are taken from different trials within the same block, and this analysis is, therefore, less sensitive to overall changes in firing rate in individual neurons over time throughout the recording session.

Results

Behavioral results

Two carrion crows were trained on a visual working memory task with varying delay durations to test the effect of changing memory delays on single-neuron activity. The birds had to memorize a given sample image over a varying delay and then choose the same image out of a set of two images during a test period (Fig. 1a). Delay durations of either 1500, 3000, or 4500 ms, each presented in blocks of 60 trials, were used (Fig. 1b).

Both crows reached a performance well above the chance level of 50% and performed the task proficiently in every recording session. Crow I reached $93.4 \pm 0.4\%$ correct and crow B reached $83.7 \pm 1.0\%$ correct. The performance (percent correct) differed significantly between the three delay conditions (crow I: $p < 0.001$, $n = 25$, crow B: $p < 0.001$, $n = 23$; Friedman test) (Fig. 2a). The performance of crow I dropped significantly from $98.4 \pm 0.2\%$ for the short delay to $93.5 \pm 0.7\%$ for the intermediate delay ($p < 0.001$, Wilcoxon test). For the longest delay (4500 ms), the performance further decreased to $86.8 \pm 0.8\%$ ($p < 0.001$, Wilcoxon test). The same performance decline with delay duration was observed in crow B, for which the performance dropped from $89.8 \pm 1.2\%$ (short delay) to $82.2 \pm 1.3\%$ (intermediate delay) and down to $77.9 \pm 1.2\%$ (long delay). Both differences were significant ($p < 0.05$, Wilcoxon test).

Similar to the performance, the reaction times to respond increased significantly with the different delay durations for both crows (each bird: $p < 0.001$, Friedman test) (Fig. 2b). From short to intermediate delay, the reaction times medians
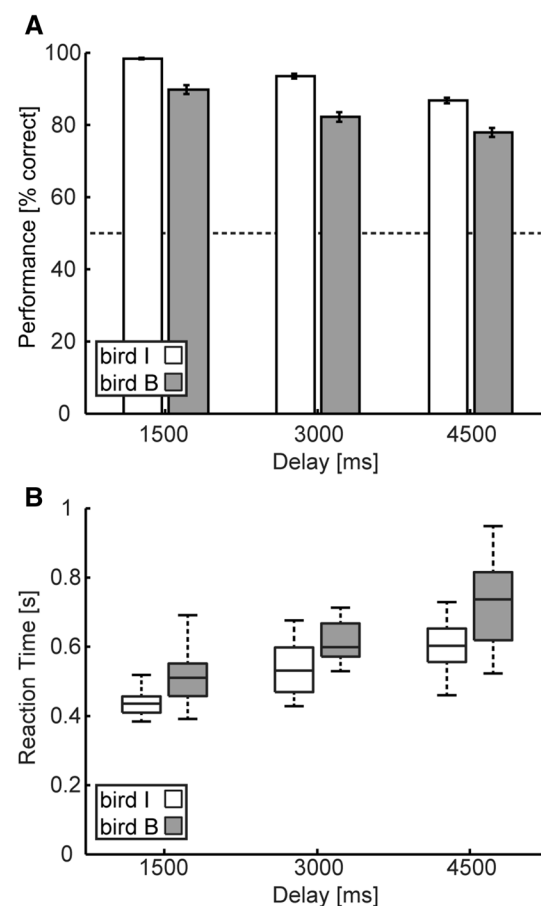
Fig. 2 Behavioral performance during recordings. **a** Performance (percent correct) of both crows separately for three different delay durations. **b** Boxplot of reaction times of both crows separately for the three different delay durations

increased from 436 to 531 ms in crow I, and from 511 to 599 ms in crow B (both crows, $p < 0.001$, Wilcoxon ranksum test). For the longest delay period, the reaction times further increased from 531 ms (short delay) to 603 ms (long delay) in crow I ($p < 0.01$, Wilcoxon ranksum test) and from 599 ms (intermediate delay) to 737 ms (long delay) in crow B ($p = 0.0265$, Wilcoxon ranksum test). Both the performance and the RTs indicate that longer memory delays were more demanding for the crows.

Neuronal responses

We recorded the activity of 103 single NCL neurons in both crows during the above-described delayed match-to-sample task with three different delay durations. To investigate whether these neurons showed sustained activity related to working memory, we focused on neuronal activity in the delay period. We analyzed the last 1000 ms before the test items were turned on for the 1500 ms delay, the last 2000 ms for the 3000 ms delay, and the last 3000 ms for the 4500 ms delay. Neurons contributing to working memory were expected to show differences in response to specific sample items. A fraction of 19.4% of all neurons (20/103) changed their firing rates significantly in the last two-thirds of the delay period as a function of the preceding sample item ($p < 0.05$, two-factor ANOVA). On average, these neurons showed firing rate differences over a time span of up to 3000 ms (the length of our analysis window in the longest delay period). Figure 3a–c depicts the response of the same single neuron to the three different sample images (indicated by different colors) and the three different delays of 1500 ms (Fig. 3a), 3000 ms (Fig. 3b) and 4500 ms (Fig. 3c). Irrespective of the delay duration, the cell showed highest sustained delay activity to the same preferred sample item (color-coded in dark blue). In addition to the 20 neurons that differentiated between sample items in the delay periods, we found that 62 out of the total 103 neurons showed differences in firing rate for the different delay durations ($p < 0.05$, two-factor ANOVA) irrespective of sample identity.

Next, we pooled the time course of activity of all neurons that showed item-selective firing during the delay period to obtain a population response. We first normalized each neuron's firing rate by its baseline firing rate (see "Materials and methods"). In addition, the responses were rank-sorted according to each neuron's sample preference. The sample item that provoked the strongest neuronal response was termed the preferred item. The sample that resulted in the lowest firing rate was called the non-preferred item. The firing rates of the population to the preferred item remained elevated throughout the delay periods (Fig. 3d–f).

Our population of cells that showed a sample-selective delay response (20/103) also showed increased activity during sample presentation. Shortly after sample onset, the activity increased when the preferred sample was presented (Fig. 3d–f). This effect can be also be seen in the example neuron (Fig. 3a–c). The preferred sample item during the delay period was also the preferred sample item during sample presentation in 55% of the selective cells, as opposed to 33.3% which would have been expected by chance. Thus, the relative order of preference for sample image in the delay and sample period was correlated.

To quantify and compare sample-selective delay discharges, we calculated normalized tuning curves over the last two-thirds of the duration of the delay period for the three different delay durations and the rank-ordered sample preferences. Figure 4a shows that the difference in selectivity between the sample items decreased for the longer delays. Nevertheless, the order of selectivity for a certain sample remained the same in every delay. A repeated-measures two-way ANOVA with the delay duration as first main factor and the sample item as second main factor showed a strong item selectivity effect ($p < 0.001$) with no main effect of the delay. An interaction between both factors was identified ($p < 0.01$). This interaction is most likely due to the convergence of selectivity indices with delay duration (Fig. 4a).

In addition, we repeated the same selectivity analysis for the sample period with the 19.4% of neurons that showed item selectivity in the delay period (Fig. 4b). The neurons' preference rankings in the delay period were taken as references. Interestingly, the order of sample preferences in the sample period reflected the order found in the delay period. Similar to the delay period, the item had a strong effect on selectivity ($p < 0.001$) and the delay had no effect when tested with a repeated-measures two-way ANOVA. There was no interaction between both factors.

To determine whether the selective activity had behavioral relevance for the crows' behavior, we analyzed the delay activity of selective cells in trials when the crows chose a wrong test item, e.g., made an error (Fig. 5). Due to the small number of errors in the shorter delays, we combined the different delays for this analysis. We found that the normalized firing rate in error trials was significantly lower for the preferred sample item ($p < 0.01$, Wilcoxon's signed-rank test). Discharge rates dropped from 0.36 SD ± 0.13 above baseline to 0.11 SD ± 0.09 above baseline, i.e., decreased by 70%. The intermediately preferred sample item showed no change ($p = 0.22$, Wilcoxon's signed-rank test). The discharge rates to the non-preferred item qualitatively tended to be higher in error trials, but did not reach significance.

We wondered whether the time course of delay-selective responses would adapt and change with switches of blocks with different delay duration. The single cell example in Fig. 6a indicates that this was the case. As evidenced by the single-trial discharges shown in the dot-raster plots (Fig. 6a), this neuron exhibited a systematically increased activation ('ramping activation') towards the end of the delay period.
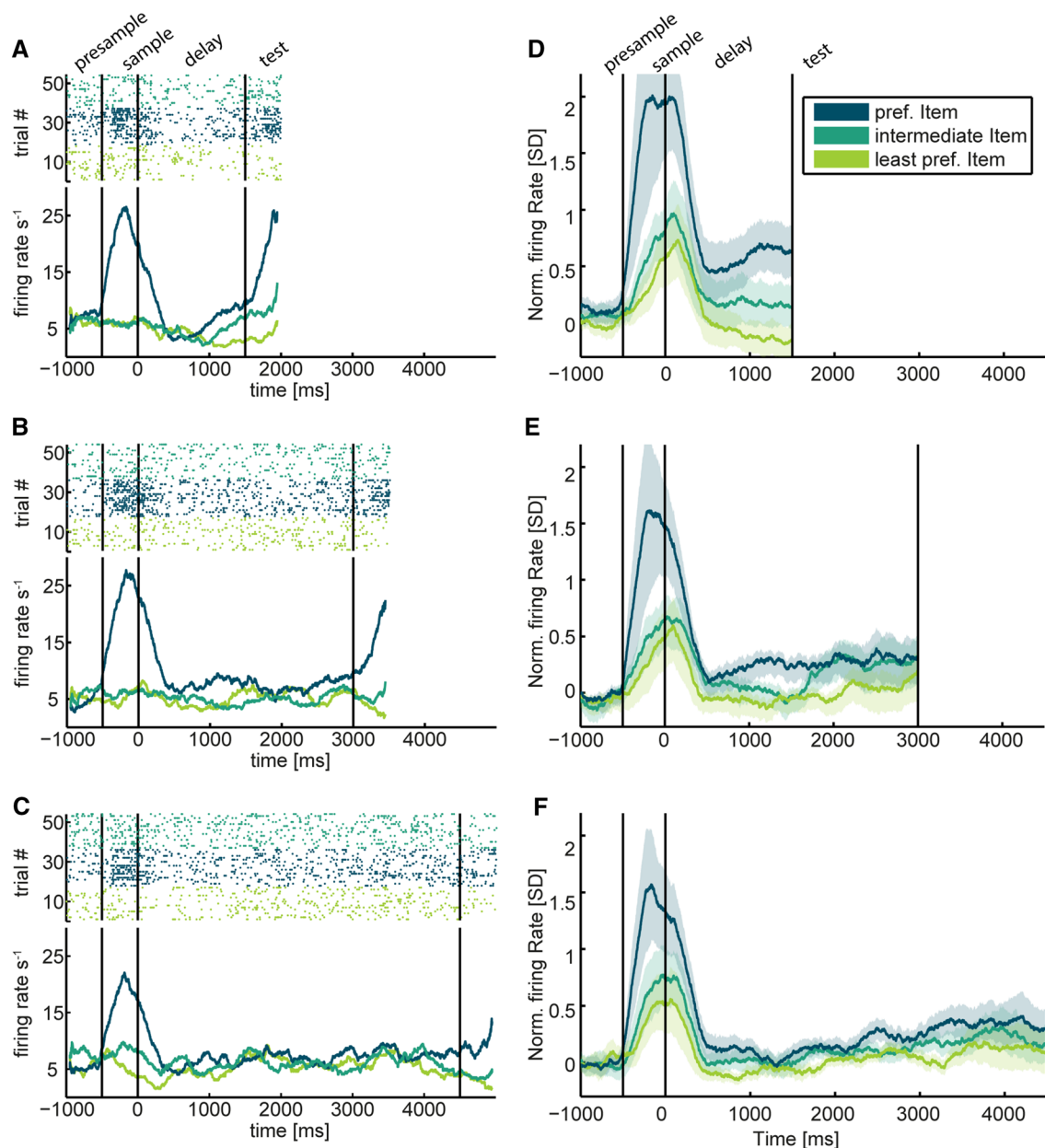
Fig. 3 Sample-selective delay activity in the NCL during different delay durations. **a–c** Example neuron. The top panel shows dot-raster histograms (each dot represents an action potential) for the three different sample items. The bottom panel shows the spike-density histograms for the same cell, smoothed with a 300 ms boxcar window for illustration. Vertical lines mark the beginning and end of the sample and delay periods. Colors represent responses to the three different sample images. **d–f** Neuronal population activity. Firing rates were normalized to units of SD over baseline before averaging across all selective neurons. Shadows show standard error of the mean, and vertical lines mark the beginning and end of the sample and delay periods, respectively. **a, d** Short delay (1500 ms), **b, e** intermediate delay (3000 ms), and **c, f** long delay (4500 ms)

Interestingly, this ramping activity adapted to, and re-scaled with, the duration of the delay; irrespective of the duration of the delay in the respective blocks, the highest discharges occurred towards the end of the delay.

Since the duration of the delay was not indicated to the crow but systematically changed after 60 correct trials, we wondered whether the temporal response pattern of the neurons might re-scale with the crows getting behaviorally adjusted to the new delay length after a block switch. Figure 6b shows the details of the single-trial discharges of the same neuron as depicted in Fig. 6a when delay duration was switched from 3000 to 1500 ms. The ramping activity is absent for the first trials and only occurs after about 5 trials into the new, shorter delay block. The converse effect
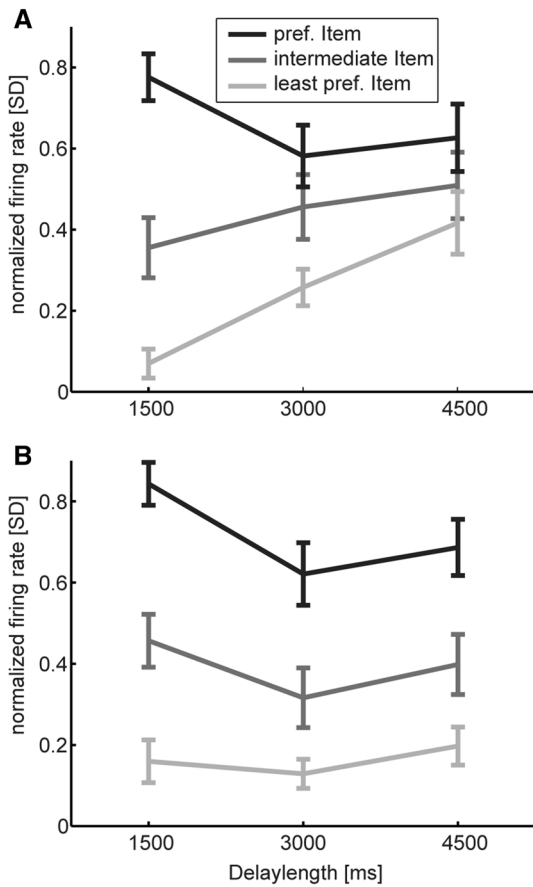
Fig. 4 Sample preference of delay-selective neurons as a function of delay duration. **a** Normalized tuning curves for the differently preferred items during the three delay durations. Tuning curves were derived for the last two-thirds of the delay period and normalized to units of standard deviation from baseline. **b** Tuning curves of the same delay-selective neurons for the sample period
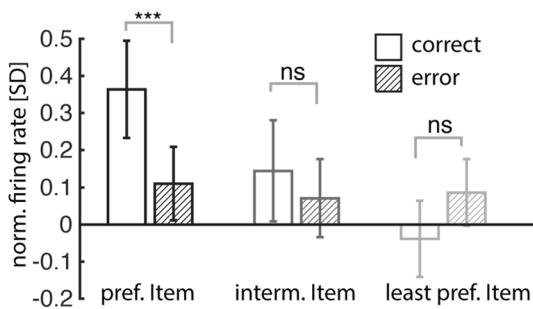

Fig. 5 Comparison of firing rates in correct versus error trials for all delay-selective neurons ($n = 20$). Neurons show reduced firing to the preferred sample during error trials. The normalized firing rate is depicted in standard deviations from baseline. Each pair of bars shows the normalized firing rate in correct (blank bars) and error trials (striped bars)
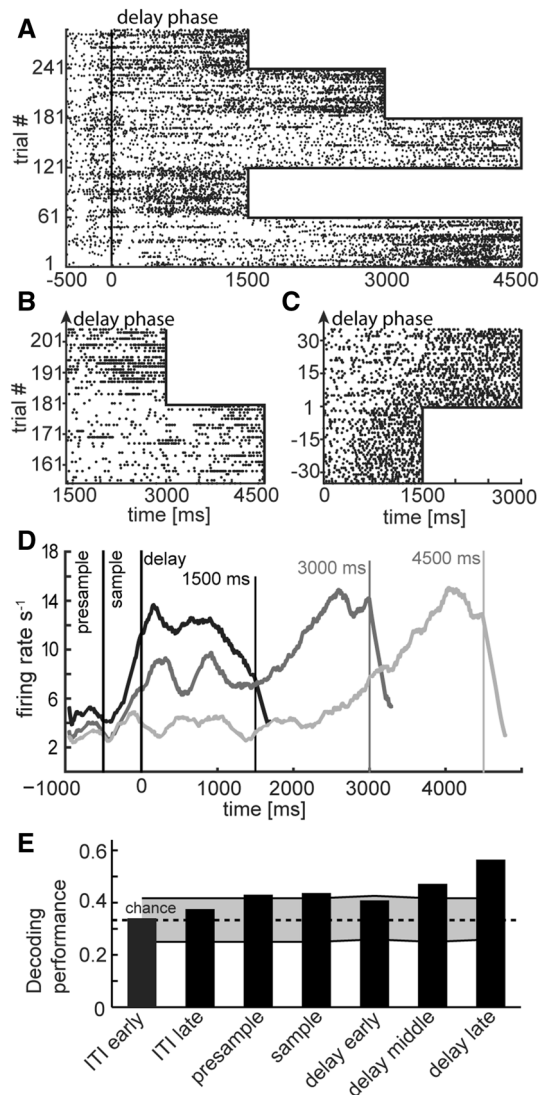
Fig. 6 Block transitions lead to changes in the firing pattern during the delay. **a** Spiking activity of an example neuron for several blocks of different delay durations. The black line depicts the end of the delay phase in each trial. Each line represents a trial; each dot indicates an action potential. Towards the end of each delay, the spiking activity increases around 1000 ms before sample onset. When the delay duration changes, ramping activity shifts in time. **b** Example of rapid adaptation of this example neuron from **a** during transition from an intermediate to a short delay block. This is a magnified section of the neuron in **a** around trial 181. **c** A different single unit showing block transition from a short delay to an intermediate delay. Block transition is at trial #1; the ramping activity shifts slowly over several trials. **d** Spike-density histograms of the cell in **a** and **b**. Ramping activity adapts to each delay duration and ramps towards the end of the delay. The data were smoothed with a 300 ms boxcar window for illustration. Vertical lines mark the beginning and end of the sample and delay periods. Shades of grey represent responses to the three different delay durations. **e** Decoding performance of a classifier during different task windows of 500 ms duration. The grey area shows 95% of all classifications with shuffled labels for 1000 times
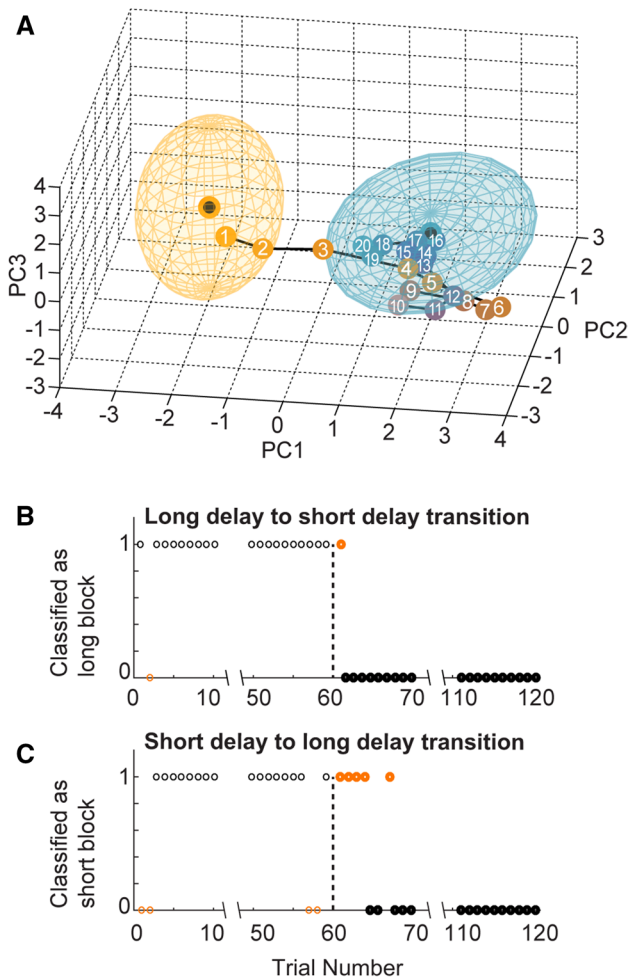
Fig. 7 Classification of delay duration based on the neuron population ($n = 88$ units). **a** Principal component analysis (PCA, first three PCs shown) depicts the high-dimensional space of firing rates during the long and short delay blocks. Each ring and ellipsoid depict the mean and standard deviation of population activity during the last 40 trials of the long (yellow) and short (blue) delay blocks. Small circles illustrate activity patterns during the first 20 trials of the short delay block, color-coded, and labeled by trial order. **b** Decoding performance for transition trials during the long-to-short block change. A classifier was trained on the middle 40 trials of each block and used to classify the first 10 trials and last 10 trials of the long block (trial 1–10, trial 50–60) and short block (trial 61–70, trial 111–120). Misclassified trials are marked in red. **c** Decoding performance for transition trials during the short-to-long block change

can be seen for a second neuron, as shown in Fig. 6c; here, the ramping activity gradually became delayed during the first few trials after the delay block was switched from short (1500 ms) to intermediate (3000 ms). Figure 6d shows the PSTH of the same neuron depicted in A and B. It can be seen that the ramping activity of the cell re-scales with the delay duration. An early response shortly after the start of the delay gets weaker for longer delays. Similarly, the ramping starts later for longer delay durations.

To evaluate at what time points throughout a trial the neuronal discharge patterns varied with the delay duration of the current block, we performed a decoding analysis in successive 500 ms time bins throughout the trial. To control for arbitrary changes in neuronal activity over time, we used the first rendition of each block to train the classifier and the second rendition to test the classifier. The data in Fig. 6e, therefore, provide a conservative estimate of the actual information about the current block present in neuronal activity. Delay length could be reliably decoded (criterion: better than 95% of decoding with shuffled data) from neuronal activity in the presample, sample, and the end of the delay period. Since decoding performance was highest in the third 500 ms time bin in the delay, i.e., the end of the delay for short delay blocks, we focus our analysis of block transitions on this time bin.

To investigate how activity changed during block transitions on the population level, we selected all neurons (irrespective of selectivity) that were recorded for the transition from the long delay to the short delay (trial 121–240, Fig. 1b, $n = 88$). We focused on discharge rates in a 500 ms window at the end of the short delay (1000 to 1500 ms after delay onset). The decoding of the delay duration was strongest at this time (Fig. 6e), because firing rates, for example by ramping activity (Fig. 6a, b), differed strongly. We performed a principal component analysis (PCA) to depict the high-dimensional space of firing rates of all neurons. Figure 7a shows the average population activity during the last 40 trials in the long delay block (yellow) and short delay block (blue). In addition, the first 20 trials of the short delay block are shown individually to illustrate the evolution of population activity during the transition period. The first two trials of the short delay block lie close to the average of the long delay block, indicating that neuronal population activity during these two trials was more similar to activity in the long delay than in the short delay. For example, neurons which exhibit ramping activity before the onset of the match period might require two trials to shift their ramping activity to the earlier point in the short delay. This is confirmed by a decoding analysis using the middle 40 trials of each block as training data and then classifying the first 10 trials and last 10 trials of each block as either long or short delay trials. Figure 7b shows that the first trial (trial # 61) of the short delay block was classified as belonging to the long delay block based on their population activity, while all the following trials as well as the ten final trials in this block were classified correctly as belonging to the short delay. Applying the same population analyses to the transition from short delay to long delay (trial 181–300, Fig. 1b, $n = 83$) revealed that population activity patterns needed approximately four trials to adjust to the discharge patterns of the long delay (Fig. 7c).

Discussion

We trained two carrion crows to remember visual items for varying delay durations in a delayed match-to-sample (DMS) task. Sample selectivity of NCL neurons was maintained throughout different working memory durations but selectivity strength decreased with increasing delay durations, which mirrored worsened behavioral performance for longer memory delays. Furthermore, the neurons encoded the length of the current delay by adjusting their temporal activity profiles to varying delay durations. These changes could be read out of the neuronal population already a few trials after the delay duration was changed.

Behavioral performance declined with memory duration

The carrion crows in the present experiment showed a systematic decline in proportion of correct responses with increasing delay intervals in the DMS tasks. Such a decreasing performance pattern was in agreement with a previous study in jungle crows (*Corvus macrorhynchos*) (Goto and Watanabe 2009). These authors reported the memory retention of jungle crows with three comparison stimuli to be above chance level, even for delay periods that lasted up to 64 s, but with reaction times that were about four times longer than those of our carrion crows. In pigeons, performance in the DMS task with two comparison stimuli drops to chance level at a delay interval of approximately 20 s (White 1985). We did not test the maximum retention time in our crows, but linear extrapolation of performance data to longer delays would indicate more reminiscence of the pigeon data. However, such comparison has to be treated carefully, since the experimental protocols varied in important parameters. For example, the number of trials per session in our experiment was close to an order of magnitude higher compared to both aforementioned studies. The combination of a high number of trials and short inter-trial intervals could lead to interference effects from the previous trials or reduced motivation to perform trials with comparatively long delays. Furthermore, birds seem to perform better in similar tasks when the delay durations are not blocked (White 1985), as was the case for the jungle crows but not for our carrion crows. Since our task was not designed to test the maximal retention time for visual working memory, but to optimize parameters for neurophysiological recording, we cannot give conclusive evidence for the behavioral limits of working memory in carrion crows.

Timing

In addition to information about the currently remembered sample stimulus, NCL neurons stored information about the current delay duration. In our task, there was no cue to indicate the delay duration. However, since delay durations were presented in blocks of 60 trials, the crows could know the current delay duration after the first trial of a new block. Block identity could be read out from population activity during the presample, sample, and delay periods, but not during the inter-trial interval. Note that the decoding performance in Fig. 6e is a conservative estimate for the amount of information stored about block identity, since we took great care to exclude effects by changes in baseline firing rate over time that were unrelated to delay blocks. Therefore, the neuronal population seemed to slightly shift its baseline activity state in different blocks even before the sample stimulus was presented. This change in activity could be related to reward expectation or could be storing information about the current block to adapt the temporal activity profiles over the trial. Remarkably, the neurons maintained their selectivity for the same sample item over the entire delay, as well as for different delay durations. This indicates that individual neurons overall do seem to participate in similar ways in the encoding of stimulus identity for different delay durations, in spite of dramatic differences in their temporal response profiles and subtle changes to their baseline state.

In the delay periods, more pronounced changes to the temporal response profile of individual units in different blocks could be observed. We frequently noticed nonspecific ramping activity at the end of the delay, and this activity switched to different time points after delay onset depending on delay duration. Interestingly, the activity profile required several trials to gradually adjust to a new delay duration, even though the crow would have been able to infer the current block after the first trial in that block.

In primates, neurons in higher sensory or association cortices are also known to show anticipatory ramping activity preceding predictable sensory stimuli or planned movements. This anticipatory activity may reflect the time of a planned saccade in the posterior parietal cortex (Janssen and Shadlen 2005), the appearance of a forthcoming vibrotactile stimulus in the prefrontal cortex (Brody et al. 2003), or the temporal evolution of attention to an upcoming visual stimulus in V4 of the extrastriate visual cortex (Ghose and Maunsell 2002). The influence of delay duration on such PFC activity was evaluated in rhesus macaques (Kojima and Goldman-Rakic 1982), in which the majority (62%) of neurons had a firing pattern that was unchanged by the delay length, but 33% of neurons responded to changed delay durations. The highest firing rates were found in the early part of the delay, while other studies report predominantly ramping activity at the end of the delay (Brody et al. 2003). Such temporal sensitivity requires access to an internal clock, but the nature of this clock (or multiple clocks) is unknown (Bermudez and Schultz 2014). Time-dependent delay signals, therefore, need to be taken into account when

designing models of persistent activity as a basis for short-term memory, both in mammals and in birds.

Encoding of the sample image decreased with memory duration

Consistent with the decline in behavioral performance, the selectivity of single neurons decreased for longer delays. This finding emphasizes the behavioral relevance of the observed delay activity. However, while behavioral performance remained far above chance for long delays, neuronal selectivity seemed to drop off quite sharply. While sustained activity during delay periods is the most accepted neuronal code for short-term memory, other mechanisms are also discussed. For instance, just as assumed for long-term memory, information that must be held for seconds in short-term memory might also be encoded in a distributed pattern of transiently modified synaptic weights, a code that would not be detectable by activity-based measurements. Such processes have been postulated from multivariate studies of extracellular activity in monkey PFC (Barak et al. 2013, 2010; Stokes et al. 2013). For PFC activity recorded during monkeys' performance of a delayed paired-association task, Stokes et al. (2013) suggested that patterned activity may leave behind a patterned change in the synaptic weights of the network (i.e., hidden state). As a consequence, subsequent stimulation will be patterned according to the recent stimulation history of a network. Any driving input to the system will, therefore, trigger a systematic population response that might be exploited to decode the recent stimulation history of the network, which would then constitute a short-term memory signal. In a similar vein, Sugase-Miyamoto et al. (2008) have proposed that neuronal networks in inferior temporal cortex support visual short-term memory for objects by acting as 'matched filters' that hold a 'static copy' of the target in a pattern of synaptic weights. Future studies will have to investigate if short-term memory signals in NCL networks depend on transiently configured patterns of synaptic weights in addition to persistent activity.

The decrease in sustained activity with time could also reflect a devaluation of reward with time. NCL neurons are known to be influenced by reward expectancy and subjective reward value (Kalenscher et al. 2005; Browning et al. 2010), and the crows could be argued to expect less reward for long delays, either because their performance is worse or because reward is discounted after the longer delays. In the pigeon NCL, Kalenscher et al. (2005) have shown that the reward-anticipating activity of single forebrain neurons is modulated by a combination of delay duration and reward amount. While we cannot rule out block-specific influences of reward expectation on neuronal activity in our task, it is interesting to note that the decline in selectivity does not result from an overall reduction in firing rate in long delays

or similar nonspecific factors (Fig. 4a). Rather, there is a decrease in firing to the preferred sample, and increase in firing to the non-preferred sample, leading to a specific decrease in sample selectivity without overall change in firing rate. On the other hand, encoding of the sample item is weaker even during sample presentation in the long delay blocks (Fig. 4b), indicating that the decline in selectivity is not simply a result of passive deterioration of the code over time either.

We have shown that single neurons in corvid NCL, an associative brain structure involved in cognitive behavior in birds, temporally maintained stimulus information in working memory over varying delay durations of several seconds. The sensitivity of NCL neurons to different delay durations shows that NCL stores behaviorally relevant information on several stages. NCL maintained the currently required information in working memory, i.e., the identity of the sample stimulus, but also kept track of slower changing contingencies in the task, like the current block, both within trials and even outside of individual trials. Therefore, NCL neurons seem to be involved in the organization of behavior on several levels, from maintaining information from individual trials to keeping track of the task which the bird is currently engaged in.

Author contribution KH and AN designed experiments, KH performed experiments, KH and LV analyzed data, and LV, KH, and AN wrote the paper.

References

Barak O, Tsodyks M, Romo R (2010) Neuronal population coding of parametric working memory. J Neurosci 30:9424–9430

Barak O, Sussillo D, Romo R, Tsodyks M, Abbott LF (2013) From fixed points to chaos: three models of delayed discrimination. Prog Neurobiol 103:214–222

Bermudez MA, Schultz W (2014) Timing in reward and decision processes. Philos Trans R Soc B 369:20120468

Brecht KF, Wagener L, Ostojić L, Clayton NS, Nieder A (2016) Comparing the face inversion effect in crows and humans. J Comp Physiol A. https://doi.org/10.1007/s00359-017-1211-7

Brody CD, Hernández A, Zainos A, Romo R (2003) Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex. Cereb Cortex 13:1196–1207

Browning R, Overmier JB, Colombo M (2010) Delay activity in avian prefrontal cortex—sample code or reward code? Eur J Neurosci 33:726–735

Cover T, Hart P (1967) Nearest neighbor pattern classification. IEEE Trans Inf Theor 13:21–27. https://doi.org/10.1109/TIT.1967.1053964

Diekamp B, Kalt T, Güntürkün O (2002) Working memory neurons in pigeons. J Neurosci 22:RC210

Ditz HM, Nieder A (2015) Neurons selective to the number of visual items in the corvid songbird endbrain. Proc Natl Acad Sci USA 112:7827–7832

Ditz HM, Nieder A (2016a) Numerosity representations in crows obey the Weber–Fechner law. Proc Biol Sci 283(1827):20160083

Ditz HM, Nieder A (2016b) Sensory and working memory representations of small and large numerosities in the crow endbrain. J Neurosci 36:12044–12052

Divac I, Mogensen J, Björklund A (1985) The prefrontal "cortex" in the pigeon. Biochemical evidence. Brain Res 332:365–368

Durstewitz D, Kröner S, Güntürkün O (1999) The dopaminergic innervation of the avian telencephalon. Prog Neurobiol 59:161–195

Fuster JM (1973) Unit activity in prefrontal cortex during delayed-response performance: neuronal correlates of transient memory. J Neurophysiol 36:61–78

Fuster JM (2001) The prefrontal cortex—an update: time is of the essence. Neuron 30:319–333

Fuster JM (2008) The prefrontal cortex, 4th edn. Academic Press, New York

Fuster JM, Alexander GE (1971) Neuron activity related to short-term memory. Science 173:652–654

Genovesio A, Tsujimoto S, Wise SP (2006) Neuronal activity related to elapsed time in prefrontal cortex. J Neurophysiol 95:3281–3285

Ghose GM, Maunsell JHR (2002) Attentional modulation in visual cortex depends on task timing. Nature 419:616–620

Goldman-Rakic PS (1995) Cellular basis of working memory. Neuron 14:477–485

Goto K, Watanabe S (2009) Visual working memory of jungle crows (*Corvus macrorhynchos*) in operant delayed matching-to-sample. Jpn Psychol Res 51:122–131

Güntürkün O (2005) The avian "prefrontal cortex" and cognition. Curr Opin Neurobiol 15:686–693

Hartmann B, Güntürkün O (1998) Selective deficits in reversal learning after neostriatum caudolaterale lesions in pigeons: possible behavioral equivalencies to the mammalian prefrontal system. Behav Brain Res 96:125–133

Hoffmann A, Rüttler V, Nieder A (2011) Ontogeny of object permanence and object tracking in the carrion crow, *Corvus corone*. Anim Behav 82:359–367

Janssen P, Shadlen MN (2005) A representation of the hazard rate of elapsed time in macaque area LIP. Nat Neurosci 8:234–241

Kalenscher T, Windmann S, Diekamp B, Rose J, Güntürkün O, Colombo M (2005) Single units in the pigeon brain integrate reward amount and time-to-reward in an impulsive choice task. Curr Biol 15:594–602

Kalenscher T, Ohmann T, Windmann S, Freund N, Güntürkün O (2006) Single forebrain neurons represent interval timing and reward amount during response scheduling. Eur J Neurosci 24:2923–2931

Karten HJ, Hodos W (1967) A stereotaxic atlas of the brain of the pigeon: (*Columba livia*). Johns Hopkins Press, Baltimore

Kojima S, Goldman-Rakic PS (1982) Delay-related activity of prefrontal neurons in rhesus monkeys performing delayed response. Brain Res 248:43–49

Lengersdorf D, Pusch R, Güntürkün O, Stüttgen MC (2014) Neurons in the pigeon nidopallium caudolaterale signal the selection and execution of perceptual decisions. Eur J Neurosci 40:3316–3327

Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. Annu Rev Neurosci 24:167–202

Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. J Neurosci 16:5154–5167

Mogensen J, Divac I (1993) Behavioural effects of ablation of the pigeon-equivalent of the mammalian prefrontal cortex. Behav Brain Res 55:101–107

Moll FW, Nieder A (2014) The long and the short of it: rule-based relative length discrimination in carrion crows, *Corvus corone*. Behav Process 107:142–149

Moll FW, Nieder A (2015) Cross-modal associative mnemonic signals in crow endbrain neurons. Curr Biol 25:2196–2201

Moll FW, Nieder A (2017) Modality-invariant audio-visual association coding in crow endbrain neurons. Neurobiol Learn Mem 137:65–76

Nieder A (2016) The neuronal code for number. Nat Rev Neurosci 17:366–382

Nieder A (2017) Inside the corvid brain—probing the physiology of cognition in crows. Curr Opin Behav Sci 16:8–14

Niki H, Watanabe M (1979) Prefrontal and cingulate unit activity during timing behavior in the monkey. Brain Res 171:213–224

Rainer G, Rao SC, Miller EK (1999) Prospective coding for objects in primate prefrontal cortex. J Neurosci 19:5493–5505

Roesch MR, Olson CR (2005) Neuronal activity dependent on anticipated and elapsed delay in macaque prefrontal cortex, frontal and supplementary eye fields, and premotor cortex. J Neurophysiol 94:1469–1497

Sakurai Y, Takahashi S, Inoue M (2004) Stimulus duration in working memory is represented by neuronal activity in the monkey prefrontal cortex. Eur J Neurosci 20:1069–1080

Schultz W (2006) Behavioral theories and the neurophysiology of reward. Ann Rev Psychol 57:87–115. https://doi.org/10.1146/annurev.psych.56.091103.070229

Stokes MG, Kusunoki M, Sigala N, Nili H, Gaffan D, Duncan J (2013) Dynamic coding for cognitive control in prefrontal cortex. Neuron 78:364–375

Sugase-Miyamoto Y, Liu Z, Wiener MC, Optican LM, Richmond BJ (2008) Short-term memory trace in rapidly adapting synapses of inferior temporal cortex. PLoS Comput Biol 4(5):e1000073

Veit L, Nieder A (2013) Abstract rule neurons in the endbrain support intelligent behaviour in corvid songbirds. Nat Commun 4:2878

Veit L, Hartmann K, Nieder A (2014) Neuronal correlates of visual working memory in the corvid endbrain. J Neurosci 34:7778–7786

Veit L, Pidpruzhnykova G, Nieder A (2015) Associative learning rapidly establishes neuronal representations of upcoming behavioral choices in crows. Proc Natl Acad Sci USA 112:15208–15213

Veit L, Hartmann K, Nieder A (2017a) Spatially-tuned neurons in corvid nidopallium caudolaterale signal target position during visual search. Cereb Cortex 27:1103–1112

Veit L, Pidpruzhnykova G, Nieder A (2017b) Learning recruits neurons representing previously established associations in the corvid endbrain. J Cogn Neurosci 29:1712–1724

Wagener L, Nieder A (2017) Encoding of global visual motion in the Nidopallium caudolaterale of behaving crows. Eur J Neurosci 45:267–277

White KG (1985) Characteristics of forgetting functions in delayed matching-to-sample. J Exp Anal Behav 45:161–174