

Learning Recruits Neurons Representing Previously Established Associations in the Corvid Endbrain

Lena Veit*, Galyna Pidpruzhnykova**, and Andreas Nieder

Abstract

■ Crows quickly learn arbitrary associations. As a neuronal correlate of this behavior, single neurons in the corvid endbrain area nidopallium caudolaterale (NCL) change their response properties during association learning. In crows performing a delayed association task that required them to map both familiar and novel sample pictures to the same two choice pictures, NCL neurons established a common, prospective code for associations. Here, we report that neuronal tuning changes during learning were not distributed equally in the recorded population of NCL neurons. Instead, such learning-related changes relied almost exclusively on neurons which were already encoding familiar associations. Only in such neurons did behavioral improvements during learning of novel associations coincide

with increasing selectivity over the learning process. The size and direction of selectivity for familiar and newly learned associations were highly correlated. These increases in selectivity for novel associations occurred only late in the delay period. Moreover, NCL neurons discriminated correct from erroneous trial outcome based on feedback signals at the end of the trial, particularly in newly learned associations. Our results indicate that task-relevant changes during association learning are not distributed within the population of corvid NCL neurons but rather are restricted to a specific group of association-selective neurons. Such association neurons in the multimodal cognitive integration area NCL likely play an important role during highly flexible behavior in corvids. ■

INTRODUCTION

Association learning is a fundamental ability underlying cognitive behaviors in humans and animals alike. In the life of corvid songbirds, association learning plays a key role as a foundation for these birds' astonishing behavioral flexibility (Clayton & Emery, 2015). A neuronal correlate for the representation of arbitrary cross-temporal associations can be found in the endbrain area nidopallium caudolaterale (NCL; Moll & Nieder, 2015; Veit, Pidpruzhnykova, & Nieder, 2015). The NCL is an avian executive brain area that is involved in reversal learning (Hartmann & Güntürkün, 1998) and flexible control of behavior (Ditz & Nieder, 2015; Veit, Pidpruzhnykova, & Nieder, 2015; Veit & Nieder, 2013). NCL shares many similarities with the independently evolved PFC of mammals (Nieder, 2017; Güntürkün, 2005; Miller & Cohen, 2001), which is similarly involved in association learning (Brincat & Miller, 2015, 2016; Cromer, Machon, & Miller, 2011; Asaad, Rainer, & Miller, 1998).

Recently, we reported a neuronal correlate of association learning in NCL of crows performing a delayed asso-

ciation task (Veit, Pidpruzhnykova, & Nieder, 2015). Neurons that were selective for familiar sample-choice associations also became selective for novel samples mapped onto the same choices during the course of learning. The selectivity strength of these neurons increased in parallel with the birds' improved behavioral performance. That is, the neurons established a prospective working memory representation, which encoded not features of the remembered sample stimulus but the associated choice. This prospective representation was similar for familiar and newly learned, novel associations.

The previous study (Veit, Pidpruzhnykova, & Nieder, 2015) focused exclusively on the emergence of prospective selectivity by preselecting neurons that were tuned to familiar sample-choice associations and averaging their responses during the learning of different novel associations. Therefore, this approach was uninformed to neurons that were nonselective to familiar associations or may have responded differently to different novel associations. In the current study, we therefore consider all neurons regardless of their selectivity to familiar associations and investigate how their activity changes during learning of individual novel association blocks. The first objective was to identify which populations of neurons signal novel and familiar associations, respectively. We may expect a pool of associative neurons that is engaged in both familiar and novel associations, as suggested by the previous study. Alternatively, distinct groups of neurons could be involved in the encoding of newly learned

University of Tübingen

*Present address: Center for Integrative Neuroscience, University of California, San Francisco, CA 94158, USA.

**Present address: Center for Neuroprosthetics and Brain Mind Institute, School of Life Sciences, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland.

and familiar associations. In primate hippocampus and perirhinal cortex, for instance, neurons change their activity during the learning of particular associations, but not for new learning in general (Yanike, Wirth, Smith, Brown, & Suzuki, 2009; Wirth et al., 2003). These associative cells in the primate medial-temporal lobe seem to signal arbitrary new associations without a frame of reference such as a common associated response. Consequently, learning-related changes may remain unnoticed if analyses focus exclusively on prospective selectivity (Brincat & Miller, 2015).

Second, we investigated the time course of association selectivity throughout the trial. Because sample images become grouped with choice stimuli over a brief temporal delay, association selectivity may emerge already during sensing of the sample and continue throughout the delay. Such selectivity changes during the sample phase were reported in primates (Brincat & Miller, 2015, 2016; Hirabayashi & Miyashita, 2014; Hirabayashi, Takeuchi, Tamura, & Miyashita, 2013; Freedman, Riesenhuber, Poggio, & Miller, 2002; Rainer, Rao, & Miller, 1999; Asaad et al., 1998). Alternatively, sensory representations could remain unaffected by learning, with associative selectivity emerging only later in the delay period.

Third, we wondered how arbitrary stimulus associations are established based on success and failure of the crows' choices. During reinforcement learning, information about the animals' responses needs to be integrated with the evaluation of response feedback (Brincat & Miller, 2015; Heilbronner & Platt, 2013; Starosta, Güntürkün, & Stüttgen, 2013; Histed, Pasupathy, & Miller, 2009; Wirth et al., 2009). To test whether NCL is involved in this crucial part of learning, we analyzed neuronal activity during a feedback delay period following the response to novel and familiar associations. The current study therefore provides a detailed picture of the single-cell processes engaged in the corvid NCL during association learning.

METHODS

Subjects

Two juvenile carrion crows (*Corvus corone corone*) were used. They were housed in social groups in spacious indoor aviaries (Hoffmann, Rüttler, & Nieder, 2011). The crows were maintained on a controlled feeding protocol during the sessions and earned food during and after the daily tests. All animal preparations and procedures were approved by the local ethical committee and authorized by the national authorities (Regierungspräsidium Tübingen).

Behavioral Protocol

The crows were trained on a delayed paired association learning task (Figure 1A). Visual stimuli were displayed

on a touchscreen monitor (ART development MT1500-BS, 15 in., 60 Hz refresh rate). An infrared light barrier in combination with a reflector attached to the crow's head registered when the bird was positioned in front of the screen and facing it. Birds had to remain inside the light barrier throughout presample, sample, and delay periods. The size of sample and test images was approximately 20 × 20 mm.

A sample stimulus was presented for 500 msec, followed by a 1000-msec working memory delay. After the delay, crows chose between a red and blue response option, which appeared on randomized and balanced positions on a touch screen. The crows indicated their choices by selecting one of the test images with their beak. If no response occurred within 1700 msec, the trial was discarded. After the crows made their choices, a 300-msec feedback sound indicated trial outcome (correct or error), followed by a 500-msec feedback delay. Finally, correct trials were rewarded by an automated feeder, which made a movement with sound and light after each correct trial, and delivered food reward on approximately 60% of correct trials. Incorrect trials were followed by a flash of the screen and a short time-out (3 sec) after the feedback delay period.

Trials were presented in blocks of familiar and novel associations (Figure 1B). Each block introduced a pair of sample pictures, which were arbitrarily assigned to the two response options. The sample picture was chosen from this pair of images on randomly alternating trials. Familiar blocks lasted 60 correct trials. Crows worked a minimum of 120 correct trials on each novel association block. When the association was not learned according to an online criterion of 80% correct in the last 40 trials, novel blocks were continued until 180, 240, or 300 correct trials.

Each novel block introduced two new, arbitrary pictures for association learning, and crows had to learn the correct associations by trial and error. In familiar blocks, the same two images (bird and flower) were kept constant for several months and, therefore, had well-known associations to the choices (Figure 1C). Novel and familiar blocks were alternated throughout each recording session; typically the crows worked on two novel and two familiar blocks per day. Therefore, depending on the crow's performance and the time the neuron could be held, each neuron's response was observed to the familiar sample pair and up to two novel sample pairs. This task therefore required a many-to-one mapping of several visually distinct samples—some newly learned and others highly familiar—onto the same red or blue responses. For details, see Veit, Pidpruzhnykova, and Nieder (2015).

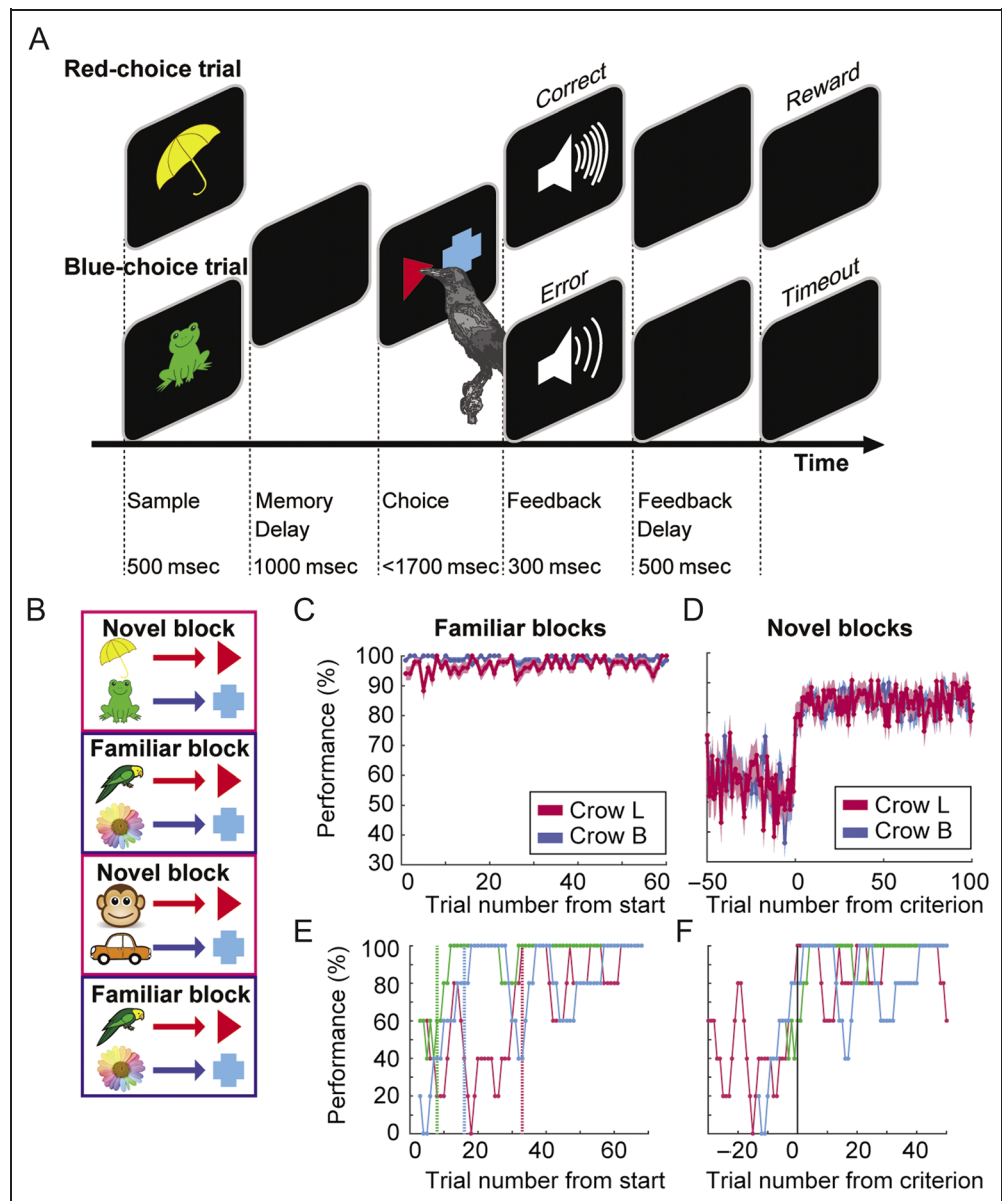
Learning Criterion

We determined a learning criterion from the behavioral data in each novel association block to divide and compare neuronal activity from the beginning of novel blocks,

Figure 1. Task design of the delayed paired association learning task. (A) Crows were working on a touchscreen monitor. Sample pictures uniquely associated with either red or blue choice were presented on randomly interleaved trials. Samples were presented for 500 msec, followed by a 1000-msec delay. In the choice period, red and blue response pictures appeared on both sides of the screen in randomized and balanced positions. The crows pecked at one picture to indicate their choice. After the choice, crows received 300 msec auditory feedback if the choice was correct, followed by a 500-msec feedback delay. After the feedback delay, reward was delivered on correct trials, whereas incorrect trials were followed by a flash of the screen and short timeout (3 sec).

(B) Novel and familiar pairs of sample pictures were presented in alternating trial blocks throughout each daily session. The association was known for familiar sample pictures, which remained constant throughout the experiment, and had to be learned by trial and error for novel sample pictures, which were exchanged for each individual novel block.

Typically, crows worked on two novel and two familiar blocks per session. (C) Behavioral performance in familiar blocks. (D) Behavioral performance in each novel block aligned by learning criterion. Performance was around chance level (50%) before the learning criterion and reliably above chance after the criterion. (E) Behavioral performance on three example novel blocks, smoothed in a five-trial window, along with the learning criterion (vertical line). (F) Behavioral performance in the same three blocks aligned by the learning criterion.



when the birds were performing around chance level (“during learning”), to the end of the block, when performance was better (“after learning”; Figure 1D; for details on behavioral performance, see Veit, Pidpruzhnykova, & Nieder, 2015). We used a state-space model of dynamic learning (Smith, Wirth, Suzuki, & Brown, 2007) to estimate the learning curve, that is, the probability of a correct response as a function of trial number, from the behavioral data. The software (www.neurostat.mit.edu/software) uses Bayesian analysis of the state-space model to determine the earliest trial in each block when the bird was performing reliably above chance. The learning curve is computed along with its upper and lower 95% confidence bounds, and we defined the learning trial as the first trial after which

there is 95% certainty that the bird is performing better than chance for the next 40 trials. For illustration, three randomly selected raw learning curves (smoothed by a five-trial window) and the criterion trials determined by the algorithm are shown in Figure 1E. Figure 1F shows the same three curves aligned by their criterion trials. During recordings, we presented a total of 149 new associations to Crow B and 103 new associations to Crow L. A learning criterion could be determined in 135 (91%) associations for Crow B and 87 (85%) associations for Crow L (successful associations). To concentrate neuronal analyses on well-controlled behavioral data in two clearly separated behavioral states (during and after learning), only successful associations were included in this data analysis.

Although behavioral performance seems to increase sharply when successful blocks are aligned by the criterion trial (Figure 1D), the actual learning behavior in individual sessions could be more variable and is described comprehensively in Veit, Pidpruzhnykova, and Nieder (2015). An association was considered not learned if the algorithm failed to converge, if it did not find a criterion (i.e., the performance of the crow was never reliably above chance), or if the criterion was found only after more than 300 completed trials. The “during learning” period includes all neural activity from Trial 1 to one trial before the criterion trial, the “after learning” period includes the criterion trial to the end of the block, or to the trial when the lower confidence bound crosses chance level.

Recording

All surgeries were performed while the animals were under general anesthesia, and the crows received postoperative analgesics. 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 & Hodos, 1967). Two custom-built microdrives with four glass-coated tungsten microelectrodes (2 M Ω impedance, Alpha Omega Ltd., Nazareth Illit, Israel) each were implanted using stereotaxic coordinates (center of craniotomy: AP 5 mm; ML 13 mm). Coordinates were obtained by identifying the NCL through immunohistochemical staining of tyrosine hydroxylase-positive fibers in brain sections of different crows (Veit & Nieder, 2013; Divac, Mogensen, & Björklund, 1985). At the start of each session, the electrodes were advanced manually until good single unit signals were obtained. Each microdrive was advanced approximately 4 mm to record across the NCL at different depths over a period of several weeks. Signal amplification, filtering, and digitizing of spike waveforms was accomplished using the Plexon system (Dallas, TX). Single-cell waveform separation was performed offline (Plexon Systems).

Data Analysis

On the basis of the birds' behavior, we determined a criterion trial for each learning session (Veit, Pidpruzhnykova, & Nieder, 2015; Smith et al., 2007), that is, the earliest trial when the bird was performing reliably above chance. This learning criterion was used to compare neuronal activity before and after the association was mastered. The criterion was reached after a median of 43 trials for Crow B (23 correct trials), and 56 trials for Crow L (32 correct trials). Details about the behavioral performance can be found in Veit, Pidpruzhnykova, and Nieder (2015).

The analysis includes all recorded neurons ($n = 342$, 177 from Crow B in 77 recording sessions, 165 from Crow L in 51 sessions) with a firing rate of at least 0.5 Hz during the trial (beginning of presample until end of delay

period) and at least 10 trials recorded for each of the two familiar sample pictures. An association was included in the analysis if the neuron was held from Trial 1 of the association until at least 10 trials after learning criterion was reached. Neuronal firing rates were calculated in a sample window starting 100 msec after sample onset, ending 100 msec after sample offset, and in a delay window starting 500 msec after delay period onset, ending 100 msec after delay period offset.

We analyzed each novel association block individually. Because each recording session presented up to two novel blocks, each neuron may contribute one or two novel association blocks to analyses, which were considered as independent measures for current analyses. Averaging individual block data from neurons contributing two novel blocks did not change prospective selectivity (see Veit, Pidpruzhnykova, & Nieder, 2015). During recording of the 342 neurons, 494 novel association blocks were recorded. Responses to a novel association were considered selective if firing rates on red and blue choice trials differed significantly ($p < .05$, ranksum test). We analyzed 188 sample-selective association blocks from 148 neurons and 128 delay-selective association blocks from 111 neurons.

We quantified the neuronal association selectivity for each neuron by calculating the area under the receiver operating characteristic curve (AUROC) between firing rate distributions elicited during the two (red and blue choice) associations in each block. AUROC is a measure of neuronal selectivity with 0.5 indicating no selectivity and both 0 and 1 indicating perfect selectivity (Green & Swets, 1966). In this case, values higher than 0.5 indicate selectivity with preference for red, whereas values lower than 0.5 indicate selectivity with preference for blue. For the comparison of AUROC values before and after learning, we calculated the mean distance from the diagonal in the direction of increased selectivity by: $(\text{AUROC_after} - 0.5) - (\text{AUROC_before} - 0.5) \times \text{sign}(\text{AUROC_after} - 0.5)$. To quantify how much the tuning of individual neurons was inverted in error trials, we defined a tuning inversion index as $\text{abs}((\text{AUROC_correct} - 0.5) - (\text{AUROC_error} - 0.5))$, so that a highly selective neuron with AUROC 1 in correct trials, which changed to 0 in error trials (or the other way around) would have a tuning inversion index close to 1. The selectivity index of Figure 4D is simply $\text{abs}(\text{AUROC} - 0.5) \times 2$, so that highly selective neurons are closer to 1. For sliding window analyses, all associations with a minimum of five trials for each condition prelearning were selected and analyzed within a 300-msec sliding window that was advanced in steps of 20 msec.

Activity during the feedback phase was analyzed in a 500-msec window during the feedback delay period, that is, starting 300 msec after the crows' choice and ending with reward delivery. All neurons with at least three errors in familiar associations were included in these analyses. Explained variance (ω^2) was used to calculate the

variance in firing rate explained by different task variables (trial outcome, trial block, interaction) calculated from firing rates in this window using a MATLAB toolbox for measures of effect size (Hentschke & Stüttgen, 2011).

RESULTS

Neurons Selective to Familiar Associations Change Activity with Learning

We recorded the activity of 342 single neurons in NCL during performance with familiar associations and during the learning process with novel associations (Figure 1A, B). Thus, the activity of each neuron was monitored in response to the familiar sample pair and either one or two novel sample pair blocks. A learning criterion was determined from the behavioral responses to split up novel blocks into trials during learning, that is, when the bird's performance was close to chance level, and trials after learning, that is, when performance was reliably above chance (Figure 1C–F; see Methods). The criterion was reached after a median of 43 trials and 56 trials for the two crows. The average performance on all trials before the criterion was 53% and 54%, and on all trials after the criterion, it was 85% and 86%, respectively.

Many NCL neurons discriminated between the two familiar sample pictures and formed a sustained working memory representation in the delay period. We have previously shown that sustained delay activity emerged in preselected familiar association-selective neurons during learning of novel associations in parallel with the crows' improved behavioral performance (Veit, Pidpruzhnykova, & Nieder, 2015). This activity favored the association of the novel sample picture with the same choice as the neuron's preferred association in familiar blocks. Therefore, a population of association-selective NCL neurons grouped distinct novel and familiar sample-choice pairs according to their meaning. These neurons formed a prospective delay representation encoding the upcoming associated choices, not the visual appearance of the individual sample pictures. For example, the neuron in Figure 2A–C responded strongest to the association of a sample with the red choice image for both the familiar and novel paired associates by increasing its selectivity during the delay with learning.

In our previous analysis (Veit, Pidpruzhnykova, & Nieder, 2015), we focused on the emergence of this prospective selectivity. We used association selectivity during familiar associations as a reference for selective responses during novel association trials in the same neurons. Therefore, we first selected neurons that discriminated between the red and blue choice trials in familiar association trials. Then, we investigated if and how such neurons would gain association selectivity in the same direction also in novel association trials. By averaging over novel trial blocks in those neurons that were recorded for

two novel trial blocks, learning-related changes could only be observed if they occurred in the same direction in both novel blocks.

In the current study, we followed a different approach, since a focus on meaningful groupings of novel and familiar associations might obscure other learning-related changes during novel blocks. For instance, the neuron shown in Figure 2D–F did not respond during familiar associations but was strongly activated during novel associations. It is therefore conceivable that single neurons show selectivity over learning of novel associations, even if they do not respond to familiar associations.

Sample Representations Are Stable while Delay Period Activity Reflects Learning and Behavior

We quantified potential learning-related changes in association selectivity in the entire recorded neuron population without considering selectivity in familiar blocks. We quantified selectivity of all neurons that significantly discriminated between the red and blue choice trials in any individual novel block using the AUROC as a bias-free measure of association selectivity (Green & Swets, 1966).

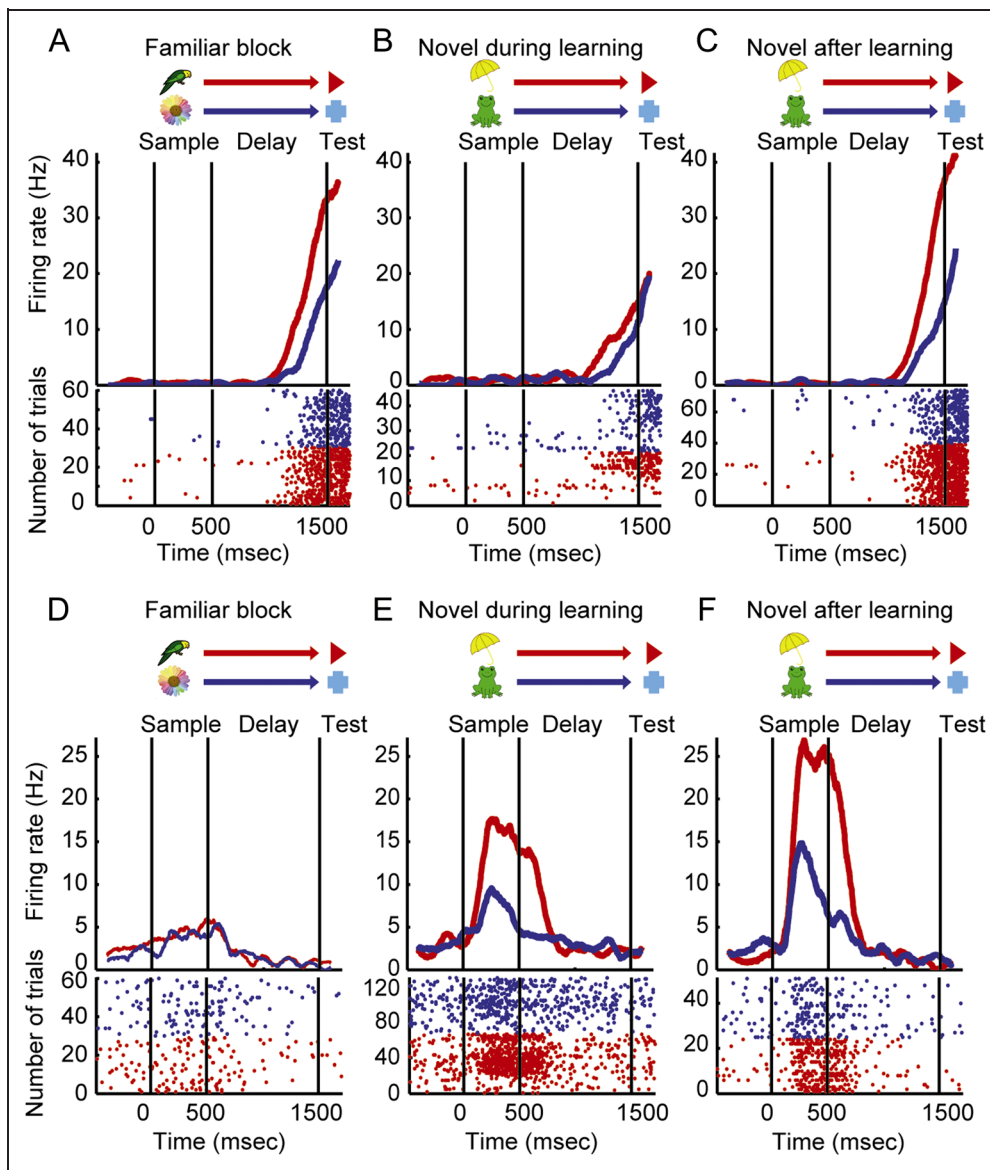
We compared selectivity for the same pictures at the beginning of novel blocks (“during learning”), before the learning criterion was passed, to the rest of the block (“after learning”), when the crow was performing reliably above chance. In the sample phase, selectivity (i.e., AUROC) for the same pictures on correct trials during learning and after learning was highly correlated (Figure 3A; $r^2 = .67$ for all selective associations, $n = 188$, $p < .001$) and fell on the diagonal (slope of total least-squares linear fit = 1.02). This indicates that neurons were selectively discriminating sample pictures in red and blue choice trials equally strong before or after learning the correct association for those pictures. Thus, neuronal selectivity during presentation of the sample pictures was not influenced by learning novel associations. In contrast, sample selectivity in the primate association cortex changed during association learning (Brincat & Miller, 2015, 2016; Hirabayashi & Miyashita, 2014; Hirabayashi et al., 2013; Freedman et al., 2002; Rainer et al., 1999; Asaad et al., 1998).

In contrast, during the memory delay, selectivity for the same pictures during and after learning was correlated less strongly than in the sample period (Figure 3B; $r^2 = .34$, $n = 128$ delay-selective blocks, $p < .001$). The preference for each picture changed in the direction of increased selectivity, away from 0.5 (i.e., neurons with AUROC values above 0.5 tended to increase, whereas values below 0.5 tended to decrease with learning). This enhancement of selectivity with learning resulted in a positive slope of 1.57 (total least-squares linear fit) when comparing selectivity during and after learning novel associations. (Without selectivity changes, a slope of 1 would have been expected.) The same effects held true when only neurons with association selectivity that occurred in both the sample and delay

Figure 2. Example neurons.

(A) Top: On separate trials, two familiar sample pictures (parrot and flower) had to be mapped to familiar test responses (red and blue choice). Middle: Spike density histogram of an example NCL neuron, obtained by averaging neuronal activity in the two conditions in 1-msec bins and smoothing with a 150-msec Gauss window. Selective delay activity was stronger for the sample associated with the red choice than for the sample associated with the blue choice (sample period AUROC selectivity value: 0.52; delay period AUROC selectivity value: 0.95). Bottom: Dot raster histogram of the neuron's discharge on individual trials. Each dot represents one action potential, colors represent red and blue choice trials. Only correct trials are shown. Trials are shown ordered by the presented sample pictures, but their actual presentation in the session was randomly interleaved. (B) During novel associations, two novel pictures (here: frog and umbrella) had to be mapped to the same two choices. The arbitrary association was acquired by trial and error learning. The response histograms show activity of the same neuron as in A on the first 46 trials of the novel association, before the bird's behavior was reliably above chance (AUROC in sample period 0.51, delay period 0.55).

(C) The same neuron's activity for the remainder of the session, after the association was learned. The neuron responded strongest to the sample image associated with the red choice again, even though the picture was visually distinct from the parrot used in the familiar association (AUROC in sample period 0.49, delay period 0.93). (D–F) As in A–C: A different example neuron did not discriminate red and blue choice for familiar sample pictures (AUROC in sample period 0.22, delay period 0.58) but responded selectively to two novel pictures in the sample period (sample period AUROC 0.69, delay period 0.49) and increased selectivity after the bird had learned the association (AUROC in sample period 0.80, delay period 0.48).



period were considered ($n = 50$ blocks, 45 neurons, $p < .05$, ranksum test), that is, when calculating the correlations with the exact same data for both periods ($r^2 = .67$ and $.37$, slope 1.03 and 1.28, in the sample and delay period, respectively, both $p < .001$). To statistically verify these observations, we calculated the mean distance of ROC values from the diagonal in the direction of increased selectivity. This distance was larger in the delay period (0.083) than in the sample period (0.046; $p < .01$, signed rank test for sample- or delay-selective associations, $n = 266$ blocks in 197 neurons).

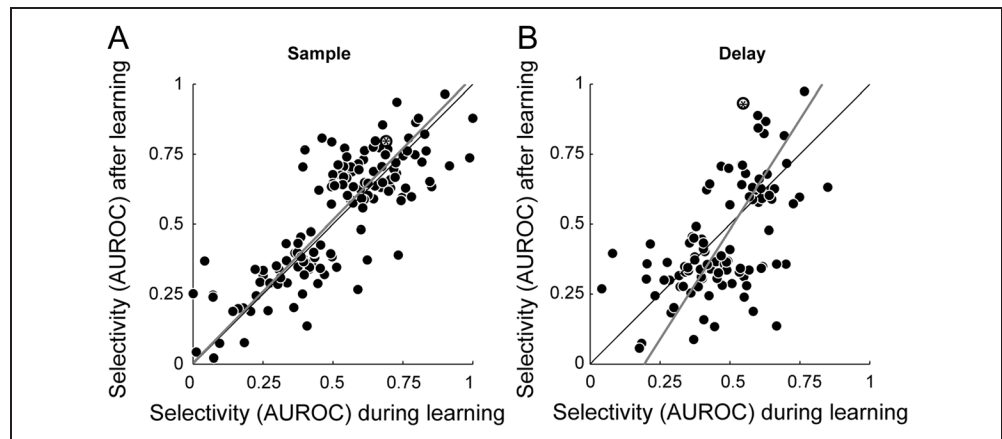
In summary, the quality of neuronal selectivity in the delay period—but not the sample period—increases with

association learning. Even when selectivity to familiar associations was not considered as an inclusion criterion and novel associations were not averaged (as in Veit, Pidpruzhnykova, & Nieder, 2015), the sample representations remained stable and did not change with association learning.

Time Point of Behaviorally Meaningful Associative Activity

We found that selectivity in the sample phase remained stable during learning, suggesting that it might reflect purely sensory information. In the delay phase, however,

Figure 3. Learning-related changes in novel association blocks. (A) Response selectivity in the sample period before the learning criterion for a novel association was passed (“during learning”) and after successful learning (“after learning”). AUROC values measure selectivity between the distribution of firing rates between the red and blue choice trials for each neuron; therefore, a value of 0 indicates strong preference for blue choice trials, a value of 1 indicates strong preference for red choice trials, whereas 0.5



indicates no selectivity. All associations with significant neuronal activity differences between red and blue choice trials are plotted ($p < .05$, ranksum test, $n = 188$). Note that each neuron may contribute only one or sometimes two data points (see Methods). Gray lines show best total least squares linear fit for the population. Only correct trials were analyzed. Asterisk marks example neuron in Figure 2D–F. (B) As in A for the delay period ($n = 128$). Asterisk marks example neuron in Figure 2A–C.

selectivity reflected the crows’ changing knowledge of the associated responses by becoming more selective over the learning process.

To explore the time point of the switch from a mostly visual to a more plastic representation during the trial, we performed a sliding window analysis. The correlation of selectivity (AUROC) values before and after learning was strongest in the sample period (Figure 4A) and the slope of the regression line was approximately 1 (i.e., no change, Figure 4B). The correlation of selectivity before and after learning became weaker in the second half of the delay, indicating modification of selectivity by learning (Figure 4A). Moreover, the slope of the regression became positive during the second half of the delay period (Figure 4B) because of more extreme AUROC values indicating enhanced selectivity in associative responses (see Figure 3B). This indicates learning-resistant activity in the sample period, which may reflect purely visual selectivity, and a more flexible task-relevant representation in the second half of the delay period.

Additional evidence for this interpretation came from comparing activity during error trials and correct trials. If neuronal activity is related to the animal’s behavior, it should change when the animal makes a mistake. We have previously reported that association selectivity in the delay period is reversed in familiar-selective neurons during error trials (Veit, Pidpruzhnykova, & Nieder, 2015), reflecting the crow’s choice and not the sample stimulus. A sliding window analysis revealed that trial outcome information (Percentage Explained Variance ω^2) was carried by the population of all neurons only toward the end of the delay period, approximately 1000 msec after sample onset (Figure 4C). Crucially, the most prominent influence of trial outcome was not a general difference in activity between correct and error trials but an interaction effect between trial outcome and sam-

ple identity. This interaction effect reflects either a breakdown of selectivity in error trials or the reversal of sample selectivity direction in error trials, that is, the prospective encoding of the crow’s choice instead of the response associated with the sample stimulus. A tuning inversion index was calculated from the delay period activity of each delay-selective neuron to show how much the tuning in error trials changed from tuning in correct trials (Figure 4D). This analysis reveals that all highly selective neurons (selectivity index close to 1) had almost completely inverted tuning in error trials (inversion index close to the diagonal). The tuning change in error trials argues that neurons were becoming behaviorally relevant for recalling and planning the response in the delay period about 500 msec before test stimulus onset.

Relating Selectivity in Novel Associations with Familiar Associations

How does the learning-related delay period increase in selectivity during novel associations relate to selectivity in familiar associations? Our task design allowed the direct comparison of each neuron’s responses to the familiar sample pair and at least one novel sample pair. In the sample period, selectivity (AUROC) for novel and familiar associations were not related (during learning: $r^2 = .02$, $p = .14$, $n = 117$ familiar-selective neurons, Figure 5A; after learning: $r^2 = .005$, $p = .43$, $n = 128$ familiar-selective neurons, Figure 5B). This was in agreement with our previous results (Veit, Pidpruzhnykova, & Nieder, 2015) that selectivity for sample pictures was not grouped according to their common choice associates in familiar-selective neurons.

Strikingly different effects were observed in the delay period. Although the selectivity values for novel and familiar associations were only weakly correlated in the

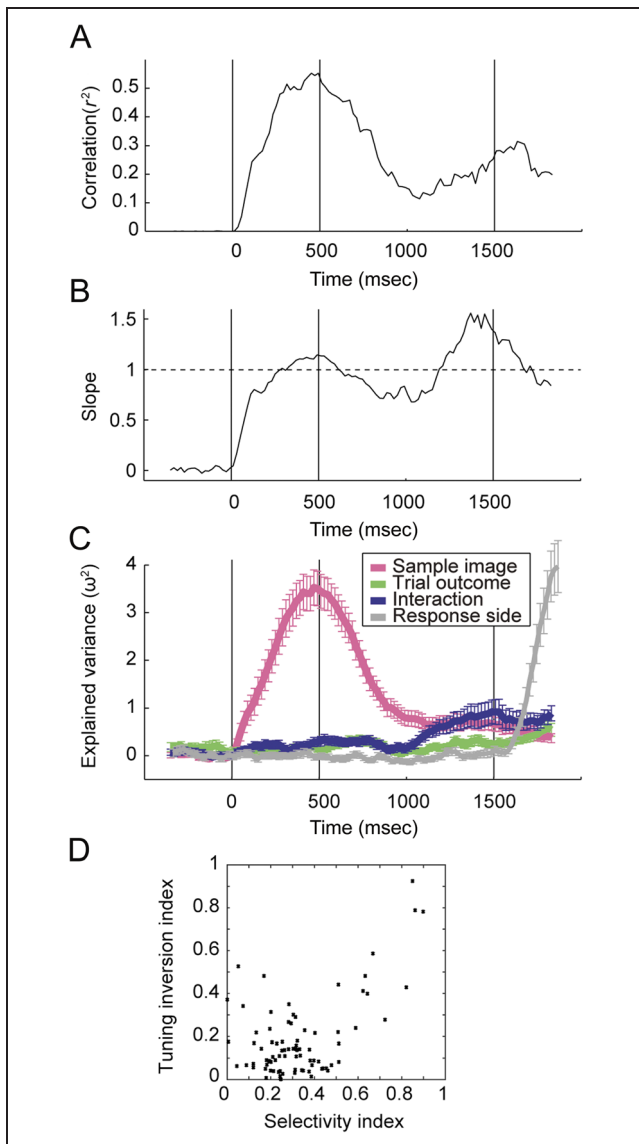


Figure 4. Temporal onset of learning-related changes. (A) Correlation of AUC value before and after learning in a 300-msec window throughout the trial. (B) Slope of total least squares linear fit to AUC values before and after learning. Compare to Figure 3, a slope of 1 indicates no changes with learning, whereas a slope larger than 1 indicates increased selectivity after learning. (C) Explained variance in firing rates by response side, sample image, trial outcome, and interaction of sample and outcome. The interaction between sample and outcome captures sample-specific changes in neuronal tuning in error trials. (D) Tuning inversion index (see Methods) for delay period firing rates for all delay-selective neurons shows that highly selective neurons tended to invert their tuning in error trials.

beginning of novel association blocks (during learning: $r^2 = .20$, $p < .001$; $n = 72$ neurons; Figure 5C), this changed after learning. After the crows mastered the novel associations, the selectivity values in novel and familiar associations were highly correlated ($r^2 = .69$ for $n = 83$ familiar-selective neurons, $p < .001$; Figure 5D). Figure 5D demonstrates that each individual neuron with strong selectivity in familiar blocks exhibited selectivity in

the same direction and, with approximately equal strength, also in novel blocks once the crows mastered the novel associations. In other words, selectivity strength in familiar association trials strongly predicts whether a neuron will become association selective during the learning of novel associations.

Moreover, the learning-related selectivity increase in the delay period seen in Figure 3B was strongest when taking neurons' selectivity for familiar associations into account. The slope for the regression line in novel associations was 1.91 ($n = 60$ associations) if the analysis was restricted to familiar-selective association neurons only. In contrast, the slope for regression based on nonselective neurons during familiar association trials was only 1.07 ($n = 68$ associations). In the sample period, the slope of the regression line did not depend on the selectivity of the neurons during familiar association trials (slope 1.07 for $n = 94$ familiar-selective associations compared with 0.97 for $n = 94$ nonselective associations). Therefore, the delay period learning effect in novel association trials was prominent in neurons that were already

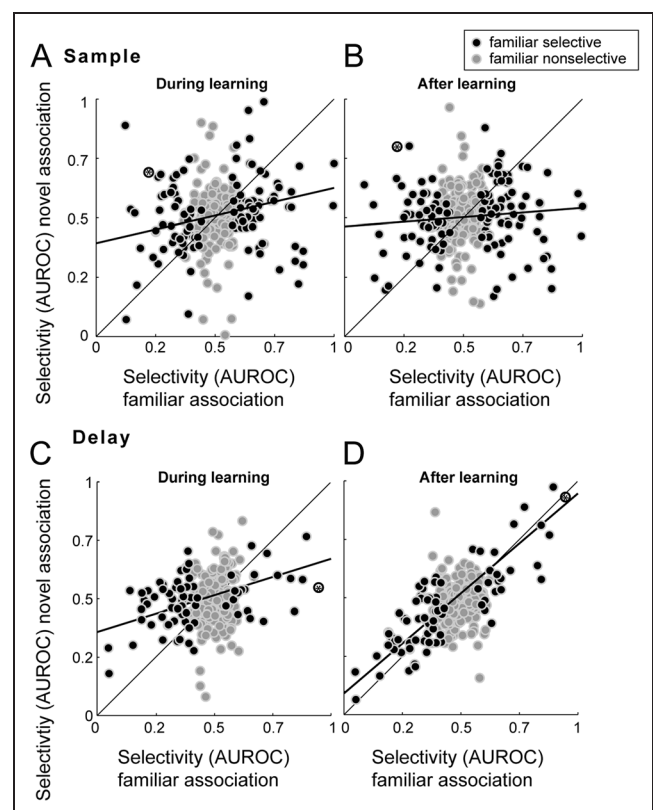


Figure 5. Relationship of selectivity between familiar and novel association blocks. (A, B) Response selectivity for familiar and novel samples on all correct trials before a novel association was learned (A) and after learning (B). Familiar-selective neurons are shown in black, familiar nonselective neurons in gray. Gray line shows best total least squares linear fit for the familiar-selective population. Asterisks mark example neuron in Figure 2D–F. (C, D) As in A and B for the delay period. Asterisks mark example neuron in Figure 2A–C.

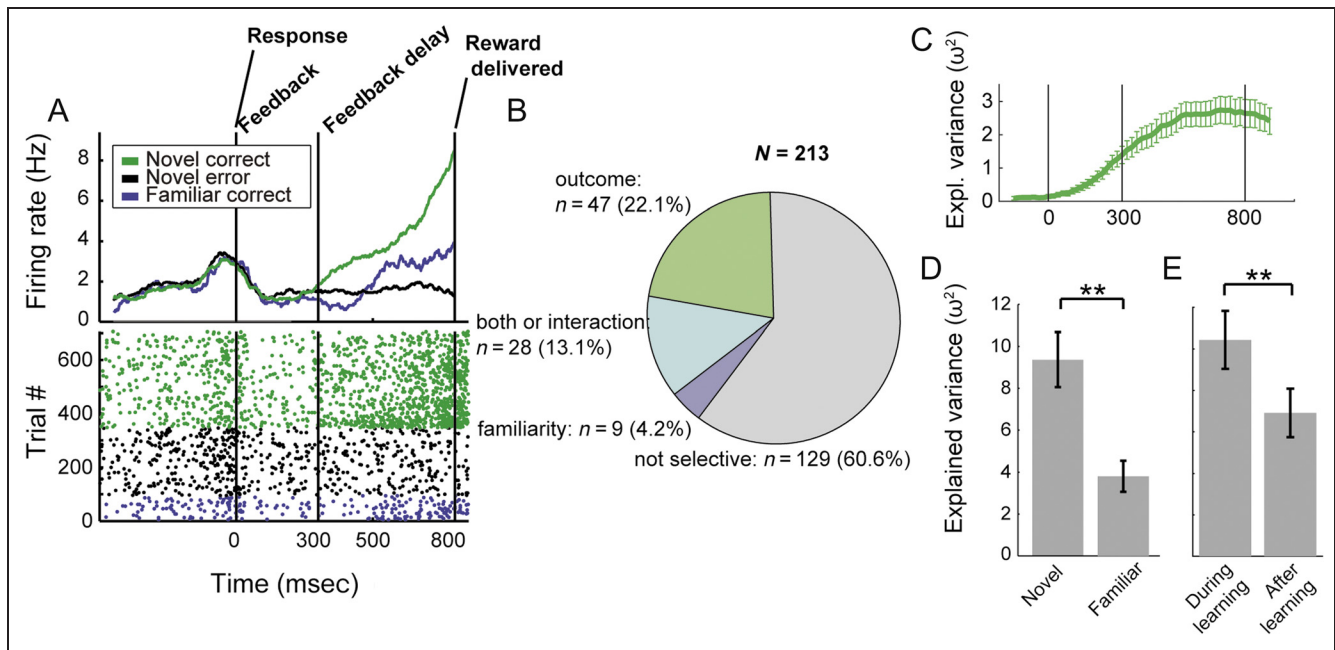


Figure 6. Evaluation of feedback. (A) Spike density histogram (top) and dot raster histogram (bottom, trials sorted according to condition) of an example neuron. A 300-msec tone delivered feedback about trial outcome immediately after the bird's response. This example neuron increased its firing rates for correct trials, but not for error trials, and this increase was stronger in novel blocks than in familiar blocks. Many neurons discriminated trial outcome in the feedback delay period following the feedback tone. (B) Percentage of neurons selective for trial outcome (correct or error), trial familiarity (novel or familiar), and interaction in the 500-msec feedback delay period following feedback. (C) Explained variance by trial outcome in a 300-msec moving window. Error bars indicate *SEM* across neurons; vertical lines indicate response (Time 0), end of feedback tone delivery (Time 300) and reward delivery (Time 800). (D) Explained variance by trial outcome for all outcome-discriminating neurons in novel and familiar blocks. (E) Explained variance by trial outcome for all outcome-discriminating neurons during and after learning in novel association blocks.

selective for previously learned associations but absent for neurons not selective in familiar association trials.

Trial Outcome Signals Are Stronger in Novel Learning than in Familiar Trials

Reinforcement learning requires that information about the animals' response is integrated with information about the success or error of that choice. To test whether NCL is involved in this crucial part of learning, we analyzed neuronal responses during a feedback period following the response. Because learning-related changes in the earlier trial periods were restricted to a particular population of neurons involved in prospective recall of learned associations during the delay period, we wondered whether that same group of neurons might be predominantly involved in evaluation of feedback at the end of the trial to flexibly adjust their tuning properties over learning.

Following the birds' response, there was a 300-msec feedback phase during which the bird received visual and auditory feedback whether the response was correct. This was followed by a 500-msec feedback delay (without any stimulation) before reward was delivered. The example neuron in Figure 6A responded more strongly for correct than for error trials in this period. Across the population, many neurons were influenced by trial outcome

in the feedback delay (25%, two-factor ANOVA, $p < .05$, $n = 75$; Figure 6B). Of these neurons, 42 responded more strongly in error trials, and 33 cells discharged more strongly following correct responses. Outcome selectivity started appearing during the feedback period and was strongest during the feedback delay (Figure 6C). In addition to the evaluation of feedback, differences between correct and error trials may reflect reward expectation, sensory, or motor factors. However, these factors should be identical during novel and familiar trials. In contrast, we found that a large percentage of neurons was influenced by familiarity of the association (17.3%, two-factor ANOVA, $n = 37$; Figure 6B), either alone or in conjunction with trial outcome.

One intriguing difference between feedback for novel and familiar trials is that information about trial outcome is crucial information required for learning in novel trials, but not for performance of highly familiar associations. Therefore, it is conceivable that a highly associative brain area like NCL is concerned more with evaluating trial outcomes during new learning and less so during recall of familiar associations. Indeed, the variance explained by trial outcome in all outcome discriminating cells ($n = 75$) was larger in novel than in familiar trials (Figure 6D, 9.4% in novel associations, 3.8% in familiar associations, $p < .0001$ signed rank test). Furthermore, even within individual novel association blocks of the outcome

discriminating cells ($n = 101$), variance explained by trial outcome was larger during learning than after learning (Figure 6E, $p < .01$ signed rank test).

How is encoding of this outcome information linked with the recall and encoding of choice information earlier in the trial? No difference in the magnitude of this feedback encoding was present between cells with sample selectivity or delay selectivity in novel associations (sample-selective: $n = 37$, 8% and 3% explained variance in novel and familiar associations, $p < .001$; delay-selective: $n = 25$, 12% and 6% explained variance in novel and familiar associations, $p < .01$). Similarly, no difference between neurons that did ($n = 21$, 11% and 4% $p < .01$) or did not discriminate familiar associations in the delay period ($n = 54$, 9% and 4% $p < .001$) were detected. Therefore, even though familiar association-selective neurons were the only group of neurons with learning-related changes in the delay period of the task, selectivity in earlier trial periods did not predict whether the neurons would participate in the evaluation of feedback for learning and coding of outcome information at the end of the trial.

DISCUSSION

Neurons in the avian cognitive integration area NCL formed a flexible prospective working memory representation of distinct sample stimuli by encoding them according to their common paired associates for upcoming behavioral choices. Neurons increased selectivity for novel associations in the delay period, but not the sample period. Learning-related changes relied almost exclusively on those neurons that were already selective for previously learned associations. A partially overlapping population of NCL neurons was involved in the evaluation of performance after the crows received feedback on their responses.

Novel Associations Map onto Existing Association Neurons

Our recordings show that the recruitment of already existing association-selective neurons for the encoding of new associations is the most important change in NCL during learning. Task-relevant neuronal selectivity changes during learning were not distributed across the population but relied almost exclusively on neurons with strong selectivity for previously learned associations, which developed equally strong selectivity for novel associations over the course of learning. In other words, as novel stimuli acquired behavioral meaning during learning, they started activating neurons that were already choice-selective. In virtually all familiar association-selective neurons, selectivity for well-known associations was highly predictive of selectivity for novel associations after learning (Figure 5D).

It has been suggested that single neurons in associative brain areas, which often represent combinations of mul-

iple task factors (Mante, Sussillo, Shenoy, & Newsome, 2013; Rigotti et al., 2013; Miller & Cohen, 2001), could accomplish the challenge of dealing with changing task demands by representing different task factors to random degrees in a population of “category-free” neurons (Raposo, Kaufman, & Churchland, 2014). By showing that learning-related changes and encoding of the crows’ choices were restricted to a specific subgroup of neurons, our data contrast the idea that flexible behavior is supported by a neuron population with randomly distributed task selectivities. Rather, our data favor a specialized and highly dedicated neuron population in an associative brain area. Whether the observed discrepancy in neural coding compared with other findings is due to differences in task demands, species differences or differences in the general endbrain organization of birds and mammals (Clayton & Emery, 2015; Güntürkün, 2005; Jarvis et al., 2005) remains to be explored. For example, our task design differs from the one used by Raposo et al. (2014) by incorporating a memory delay period and by dissociating the decision toward one of the choices from planning of a particular motor response (Merten & Nieder, 2012). Such factors of our task design could have been crucial for observing the specialized subpopulation encoding choices in the delay period.

Introducing novel samples to be associated did not recruit more neurons with selective responses in the delay period. Instead new associations were mapped onto those neurons already encoding previously learned responses. This suggests, more broadly, that highly selective delay period responses recorded in corvid NCL in various tasks (Moll & Nieder, 2015, 2017; Ditz & Nieder, 2016; Veit & Nieder, 2013) may in fact reflect the same specialized neuron population encoding behavioral output in an abstract way. Different tasks may therefore recruit highly overlapping neuron populations, with neurons dynamically adjusting their tuning properties to task demands (Veit, Hartmann, & Nieder, 2015; Stokes et al., 2013; Rao, Rainer, & Miller, 1997). Further studies interleaving several tasks while recording from the same neurons are needed to determine the extent to which the same NCL population encodes crucial cognitive operations in a range of tasks. The possibility of a specialized neuron population for encoding the output of cognitive operations suggests a promising avenue for future investigations. If not all neurons in a brain area participate in all tasks to random degrees, it might be possible to closely examine the distinct functional components of a neuronal circuit for flexible cognitive behavior in a highly associative endbrain area (Jacob & Nieder, 2014).

Prospective Memory Signals for Report-independent Decisions

In a delayed association task, selective neuronal responses for one of the associations may represent three different aspects: They could constitute retrospective

working memory that reflects properties of the remembered stimulus (Miller, Erickson, & Desimone, 1996), they could encode prospective signals that point forward to the response to be chosen (Rainer et al., 1999), or they could be selective to one particular pairing of sample and choice stimulus in memory (Wirth et al., 2003). The many-to-one task used in this study allowed disentangling these possibilities. We found that virtually all neurons with robust delay selectivity in the familiar association task also became selective toward novel associations. This argues that all neurons with strong delay selectivity encoded a prospective representation of the upcoming choice, not a retrospective representation of the remembered sample item.

Therefore, selective delay activity in NCL may represent prospective encoding of the required response whenever prospective recall is encouraged by task demands. The selective delay activity recorded so far in corvid NCL during rule-switching (Veit & Nieder, 2013), matching to numerosity (Ditz & Nieder, 2016), and cross-modal associations (Moll & Nieder, 2015) seems to support this hypothesis. In agreement with this idea, delay selectivity in tasks with simple visual matching to sample has been virtually absent (Wagener & Nieder, 2016) or much weaker (Veit, Hartmann, & Nieder, 2014). This prospective representation of the correct choice item is reminiscent of learned associations in areas of primate frontal cortex that are functionally comparable to NCL (Brincat & Miller, 2015; Rainer et al., 1999; Asaad et al., 1998; Chen & Wise, 1995). There is mounting evidence that primate PFC may not store sensory working memory per se. Rather, PFC may mainly exert top-down bias on regions that store sensory representations (Nieder, 2016a; Lara & Wallis, 2015; Jacob & Nieder, 2014).

It is important to note that the prospective representation in our task cannot reflect planning of a particular motor response, because the locations of the red and blue choices were unknown during the delay period, before the onset of the choice screen (see variance explained by response side in Figure 4C). Responses therefore seem to encode a report-independent decision toward one of the choices that is dissociated from specific motor-related activity (Nieder, 2016b; Merten & Nieder, 2012). This decision correlate also reflects how well sample stimuli and choices have been linked through association learning: Associative activity is much stronger for correct responses after learning than for correct guesses at the beginning of learning, potentially reflecting decision confidence after learning (Kiani & Shadlen, 2009). During the subsequent choice period, this activity may activate neurons that translate the decision into specific motor actions (Veit, Hartmann, & Nieder, 2015).

Sensory Representations Are Segregated from Prospective Associative Signals

In contrast to the flexibly changing delay selectivity, activity in the sample period seems to reflect mainly

visual selectivity for individual sample pictures, without relation to their meaning in the task or the crows' behavior. Specifically, sample selectivity seemed stable over learning and was the same during correct and error trials.

It is surprising that processing of visual information in the sample period was not influenced by the associated meaning of sample stimuli on behavior. Previous recordings in crows have shown that NCL activity can group distinct auditory and visual cues according to their behavioral associations during cue presentation (Moll & Nieder, 2017; Veit & Nieder, 2013). Likewise, sensory representations in songbird higher auditory areas reflect the behavioral meaning of auditory stimuli after training on auditory discriminations (Jeanne, Sharpee, & Gentner, 2013; Jeanne, Thompson, Sharpee, & Gentner, 2011; Gentner & Margoliash, 2003). Finally, in the primate PFC or temporal cortex, neuronal activity during association learning (Brincat & Miller, 2015; Pasupathy & Miller, 2005; Asaad et al., 1998) or performance of previously learned associations (Hirabayashi & Miyashita, 2014; Hirabayashi et al., 2013; Freedman et al., 2002; Rainer et al., 1999) consistently reflects learned functional categories during sample processing. Similarly, the lack of error trial effects in the sample period was unexpected. In previous studies of corvid NCL, error trial differences in the sample period were present, albeit weaker than in the delay period (Wagener & Nieder, 2016; Ditz & Nieder, 2015; Veit et al., 2014). A recent model of visual category learning suggests that the unusual absence of error trial effects could be related to the absence of tuning changes in the sample period, as choice-correlated activity differences may be required for establishing neural tuning changes over learning (Engel, Chaisangmongkon, Freedman, & Wang, 2015).

Our findings suggest a switch in coding strategy from a sensory representation in the sample period to a prospective, response-based representation in the second half of the delay. Different from studies in monkeys (Brincat & Miller, 2015; Freedman et al., 2002; Rainer et al., 1999), a clear separation in time between an exclusively sensory representation and an exclusively prospective representation was present in corvid NCL. Interestingly, a recent study of learned associations revealed that the transformation of representations from cued object to recalled object occurs in different layers of primate temporal cortex and that a subset of layer six neurons exclusively encoded the chosen target (Koyano et al., 2016), just as the NCL neurons in our study. One of the main hypotheses about the organization of the avian brain posits that different layers in mammalian cortex may correspond to different areas in the avian brain (Pfenning et al., 2014; Jarvis et al., 2005), so that it is conceivable that such layer-specific processing is regionally separated in the avian brain, and thus, it may be easier to disentangle distinct computations involved in cognitive tasks in birds.

Learning-related Feedback Signals

The learning-related tuning change in association-selective neurons requires information about trial outcome. In a feedback period following the crow's response, NCL neurons discriminated correct from error trials. The NCL of pigeons has previously been shown to encode information about upcoming rewards (Johnston, Anderson, & Colombo, 2017; Scarf et al., 2011; Kalenscher et al., 2005; Kalt, Diekamp, & Güntürkün, 1999). NCL neurons can also encode trial outcome in a task without learning requirement (Starosta et al., 2013); this study found that most NCL neurons responded more strongly to errors than correct outcomes. In agreement with this, we have found a slightly larger proportion of NCL neurons increasing their firing for errors. Such outcome/reward information may be carried to NCL through dopaminergic projection from midbrain dopaminergic nuclei (Durstewitz, Kröner, & Güntürkün, 1999).

Interestingly, trial outcome information was represented more strongly during novel than during familiar associations in our study. That way, trial outcome information was most available during learning. This difference could be explained in terms of reward prediction error (Gadagkar et al., 2016; Schultz, Dayan, & Montague, 1997). Alternatively, the difference could reflect the crows' increased attention to trial outcomes during novel blocks as opposed to familiar blocks, since trial outcome information is crucial for reinforcement learning in novel blocks, but irrelevant in well-known familiar associations.

In general, information about trial outcome is paramount for any learning organism. In the mammalian brain, neurons in a variety of associative areas have been found to reflect trial outcome information during associative learning. These areas include PFC (Cai & Padoa-Schioppa, 2014; Histed et al., 2009), the cingulate cortex (Heilbronner & Platt, 2013), as well as the hippocampus (Brincat & Miller, 2015; Wirth et al., 2009) and BG (Histed et al., 2009). Similar to all these high-level association areas in the mammalian brain, the response characteristics of the corvid NCL during learning underscore its role in mediating flexible behaviors.

Acknowledgments

This work was supported by a PhD fellowship from the German National Academic Foundation to L. V. and by DFG grant NI 618/7-1 to A. N. L. V. and A. N. designed the experiments, G. P. and L. V. performed the experiments, L. V. analyzed the data, L. V. and A. N. wrote the paper.

Reprint requests should be sent to Andreas Nieder, Animal Physiology, Institute of Neurobiology, University of Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany, or via e-mail: andreas.nieder@uni-tuebingen.de.

REFERENCES

Asaad, W. F., Rainer, G., & Miller, E. K. (1998). Neural activity in the primate prefrontal cortex during associative learning. *Neuron*, *21*, 1399–1407.

- Brincat, S. L., & Miller, E. K. (2015). Frequency-specific hippocampal-prefrontal interactions during associative learning. *Nature Neuroscience*, *18*, 576–581.
- Brincat, S. L., & Miller, E. K. (2016). Prefrontal cortex networks shift from external to internal modes during learning. *Journal of Neuroscience*, *36*, 9739–9754.
- Cai, X., & Padoa-Schioppa, C. (2014). Contributions of orbitofrontal and lateral prefrontal cortices to economic choice and the good-to-action transformation. *Neuron*, *81*, 1140–1151.
- Chen, L. L., & Wise, S. P. (1995). Neuronal activity in the supplementary eye field during acquisition of conditional oculomotor associations. *Journal of Neurophysiology*, *73*, 1101–1121.
- Clayton, N. S., & Emery, N. J. (2015). Avian models for human cognitive neuroscience: A proposal. *Neuron*, *86*, 1330–1342.
- Cromer, J. A., Machon, M., & Miller, E. K. (2011). Rapid association learning in the primate prefrontal cortex in the absence of behavioral reversals. *Journal of Cognitive Neuroscience*, *23*, 1823–1828.
- Ditz, H. M., & Nieder, A. (2015). Neurons selective to the number of visual items in the corvid songbird endbrain. *Proceedings of the National Academy of Sciences, U.S.A.*, *112*, 7827–7832.
- Ditz, H. M., & Nieder, A. (2016). Sensory and working memory representations of small and large numerosities in the crow endbrain. *Journal of Neuroscience*, *36*, 12044–12052.
- Divac, I., Mogensen, J., & Björklund, A. (1985). The prefrontal “cortex” in the pigeon. Biochemical evidence. *Brain Research*, *332*, 365–368.
- Durstewitz, D., Kröner, S., & Güntürkün, O. (1999). The dopaminergic innervation of the avian telencephalon. *Progress in Neurobiology*, *59*, 161–195.
- Engel, T. A., Chaisangmongkon, W., Freedman, D. J., & Wang, X.-J. (2015). Choice-correlated activity fluctuations underlie learning of neuronal category representation. *Nature Communications*, *6*, 1–12.
- Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2002). Visual categorization and the primate prefrontal cortex: Neurophysiology and behavior. *Journal of Neurophysiology*, *88*, 929–941.
- Gadagkar, V., Puzerey, P. A., Chen, R., Baird-Daniel, E., Farhang, A. R., & Goldberg, J. H. (2016). Dopamine neurons encode performance error in singing birds. *Science*, *354*, 1278–1282.
- Gentner, T. Q., & Margoliash, D. (2003). Neuronal populations and single cells representing learned auditory objects. *Nature*, *424*, 669–674.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1974). New York: Wiley.
- Güntürkün, O. (2005). The avian “prefrontal cortex” and cognition. *Current Opinion in Neurobiology*, *15*, 686–693.
- Hartmann, B., & Güntürkün, O. (1998). Selective deficits in reversal learning after neostriatum caudolateral lesions in pigeons: Possible behavioral equivalencies to the mammalian prefrontal system. *Behavioural Brain Research*, *96*, 125–133.
- Heilbronner, S. R., & Platt, M. L. (2013). Causal evidence of performance monitoring by neurons in posterior cingulate cortex during learning. *Neuron*, *80*, 1384–1391.
- Hentschke, H., & Stüttgen, M. C. (2011). Computation of measures of effect size for neuroscience data sets. *European Journal of Neuroscience*, *34*, 1887–1894.
- Hirabayashi, T., & Miyashita, Y. (2014). Computational principles of microcircuits for visual object processing in the macaque temporal cortex. *Trends in Neurosciences*, *37*, 178–187.
- Hirabayashi, T., Takeuchi, D., Tamura, K., & Miyashita, Y. (2013). Functional microcircuit recruited during retrieval of object association memory in monkey perirhinal cortex. *Neuron*, *77*, 192–203.

- Histed, M. H., Pasupathy, A., & Miller, E. K. (2009). Learning substrates in the primate prefrontal cortex and striatum: Sustained activity related to successful actions. *Neuron*, *63*, 244–253.
- Hoffmann, A., Rüttler, V., & Nieder, A. (2011). Ontogeny of object permanence and object tracking in the carrion crow, *Corvus corone*. *Animal Behaviour*, *82*, 359–367.
- Jacob, S. N., & Nieder, A. (2014). Complementary roles for primate frontal and parietal cortex in guarding working memory from distractor stimuli. *Neuron*, *83*, 226–237.
- Jarvis, E. D., Güntürkün, O., Bruce, L., Csillag, A., Karten, H., Kuenzel, W., et al. (2005). Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews Neuroscience*, *6*, 151–159.
- Jeanne, J. M., Sharpee, T. O., & Gentner, T. Q. (2013). Associative learning enhances population coding by inverting interneuronal correlation patterns. *Neuron*, *78*, 352–363.
- Jeanne, J. M., Thompson, J. V., Sharpee, T. O., & Gentner, T. Q. (2011). Emergence of learned categorical representations within an auditory forebrain circuit. *Journal of Neuroscience*, *31*, 2595–2606.
- Johnston, M., Anderson, C., & Colombo, M. (2017). Neural correlates of sample-coding and reward-coding in the delay activity of neurons in the entopallium and nidopallium caudolaterale of pigeons (*Columba livia*). *Behavioural Brain Research*, *317*, 382–392.
- Kalenschers, 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. *Current Biology*, *15*, 594–602.
- Kalt, T., Diekamp, B., & Güntürkün, O. (1999). Single unit activity during a go/nogo task in the “prefrontal cortex” of pigeons. *Brain Research*, *839*, 263–278.
- Karten, H. J., & Hodos, W. (1967). *A stereotaxic atlas of the brain of the pigeon: (Columba livia)* (Vol. 696). Baltimore: Johns Hopkins Press.
- Kiani, R., & Shadlen, M. N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. *Science*, *324*, 759–764.
- Koyano, K. W., Takeda, M., Matsui, T., Hirabayashi, T., Ohashi, Y., & Miyashita, Y. (2016). Laminar module cascade from layer 5 to 6 implementing cue-to-target conversion for object memory retrieval in the primate temporal cortex. *Neuron*, *92*, 518–529.
- Lara, A. H., & Wallis, J. D. (2015). The role of prefrontal cortex in working memory: A mini review. *Frontiers in Systems Neuroscience*, *9*, 829.
- Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, *503*, 78–84.
- Merten, K., & Nieder, A. (2012). Active encoding of decisions about stimulus absence in primate prefrontal cortex neurons. *Proceedings of the National Academy of Sciences, U.S.A.*, *109*, 6289–6294.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.
- Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience*, *16*, 5154–5167.
- Moll, F. W., & Nieder, A. (2015). Cross-modal associative mnemonic signals in crow endbrain neurons. *Current Biology*, *25*, 2196–2201.
- Moll, F. W., & Nieder, A. (2017). Modality-invariant audio-visual association coding in crow endbrain neurons. *Neurobiology of Learning and Memory*, *137*, 65–76.
- Nieder, A. (2016a). The neuronal code for number. *Nature Reviews Neuroscience*, *17*, 366–382.
- Nieder, A. (2016b). Representing something out of nothing: The dawning of zero. *Trends in Cognitive Sciences*, *20*, 830–842.
- Nieder, A. (2017). Inside the corvid brain—probing the physiology of cognition in crows. *Current Opinion in Behavioral Sciences*, *16*, 8–14.
- Pasupathy, A., & Miller, E. K. (2005). Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature*, *433*, 873–876.
- Pfenning, A. R., Hara, E., Whitney, O., Rivas, M. V., Wang, R., Roullac, P. L., et al. (2014). Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science*, *346*, 1256846.
- Rainer, G., Rao, S. C., & Miller, E. K. (1999). Prospective coding for objects in primate prefrontal cortex. *Journal of Neuroscience*, *19*, 5493–5505.
- Rao, S. C., Rainer, G., & Miller, E. K. (1997). Integration of what and where in the primate prefrontal cortex. *Science*, *276*, 821–824.
- Raposo, D., Kaufman, M. T., & Churchland, A. K. (2014). A category-free neural population supports evolving demands during decision-making. *Nature Neuroscience*, *17*, 1784–1792.
- Rigotti, M., Barak, O., Warden, M. R., Wang, X.-J., Daw, N. D., Miller, E. K., et al. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature*, *497*, 585–590.
- Scarf, D., Miles, K., Sloan, A., Goulter, N., Hegan, M., Seid-Fatemi, A., et al. (2011). Brain cells in the avian “prefrontal cortex” code for features of slot-machine-like gambling. *PLoS One*, *6*, e14589.
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593–1599.
- Smith, A. C., Wirth, S., Suzuki, W. A., & Brown, E. N. (2007). Bayesian analysis of interleaved learning and response bias in behavioral experiments. *Journal of Neurophysiology*, *97*, 2516–2524.
- Starosta, S., Güntürkün, O., & Stüttgen, M. C. (2013). Stimulus-response-outcome coding in the pigeon nidopallium caudolaterale. *PLoS One*, *8*, e57407.
- Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., & Duncan, J. (2013). Dynamic coding for cognitive control in prefrontal cortex. *Neuron*, *78*, 364–375.
- Veit, L., Hartmann, K., & Nieder, A. (2014). Neuronal correlates of visual working memory in the corvid endbrain. *Journal of Neuroscience*, *34*, 7778–7786.
- Veit, L., Hartmann, K., & Nieder, A. (2015). Spatially tuned neurons in corvid nidopallium caudolaterale signal target position during visual search. *Cerebral Cortex*, *27*, 1103–1112.
- Veit, L., & Nieder, A. (2013). Abstract rule neurons in the endbrain support intelligent behaviour in corvid songbirds. *Nature Communications*, *4*, 2878.
- Veit, L., Pidpruzhnykova, G., & Nieder, A. (2015). Associative learning rapidly establishes neuronal representations of upcoming behavioral choices in crows. *Proceedings of the National Academy of Sciences, U.S.A.*, *112*, 15208–15213.
- Wagener, L., & Nieder, A. (2016). Encoding of global visual motion in the nidopallium caudolaterale of behaving crows. *European Journal of Neuroscience*, *45*, 267–277.
- Wirth, S., Avsar, E., Chiu, C. C., Sharma, V., Smith, A. C., Brown, E., et al. (2009). Trial outcome and associative learning signals in the monkey hippocampus. *Neuron*, *61*, 930–940.
- Wirth, S., Yanike, M., Frank, L. M., Smith, A. C., Brown, E. N., & Suzuki, W. A. (2003). Single neurons in the monkey hippocampus and learning of new associations. *Science*, *300*, 1578–1581.
- Yanike, M., Wirth, S., Smith, A. C., Brown, E. N., & Suzuki, W. A. (2009). Comparison of associative learning-related signals in the macaque perirhinal cortex and hippocampus. *Cerebral Cortex*, *19*, 1064–1078.