Distinct visual stimuli were mapped to the same target stimuli during associative learning in parallel with the crows’ increased behavioral performance. Thus, sustained activity in the NCL actively processes information for the upcoming behavioral choice. These data provide new insights into memory representations of behaviorally meaningful stimuli in birds, and how such representations are formed during learning. The findings suggest that the NCL plays a role in learning arbitrary associations, a cornerstone of corvids’ remarkable behavioral flexibility and adaptability.

Significance

Corvid songbirds show remarkable cognitive abilities. The executive brain area nidopallium caudolaterale (NCL) is thought to mediate flexible behavior and cognition in birds, similar to the independently evolved prefrontal cortex in mammals. Here, we show that NCL neurons in crows change their responses during associative learning tasks to signal which stimuli belong together. NCL activity mapped distinct visual stimuli to a common associated response by grouping them according to their behavioral meaning during food-caching behavior.

Behavioral Performance. During novel block, both crows started at chance performance (45% and 52% on the first trial, both $P > 0.05$, binomial test against 50% chance) and reached $\sim 75\%$ correct after 70 trials (Fig. 1C). We used a state-space model of dynamic learning (8) to estimate a learning criterion; that is, the first trial of each block after which the behavior of the crow was reliably above chance. This learning criterion was later used to compare neuronal activity during and after learning the novel associations. A learning criterion could be determined for 135 (91%) associations for crow B, and presented at randomized positions (Fig. 1A).

Two types of association tasks were shown in daily sessions in separate trial blocks: The first trial block per session was a “novel” block, in which crows had to learn by trial and error which of the two novel sample images (R1-B1) was associated with the red or blue choice, respectively (Fig. 1B, Left). The two test images were kept constant, while the sample images were exchanged for each new association block. After the crows acquired this novel association, a second, “familiar” trial block followed, during which the crows had to associate well-known paired associates (R-B) to the same two test items based on long-term memory, without a learning requirement (Fig. 1B, Center). After 60 correct trials in the familiar trial block, a further novel trial block was shown with new sample images. The crow again had to learn to associate samples with the proper choice items, followed by another familiar trial block.

Results

Two crows performed a visual delayed-association task in which one of two sample images was uniquely associated with one of two choice images (red triangle or blue cross, respectively, presented at randomized positions) (Fig. 1A).

Two types of association tasks were shown in daily sessions in separate trial blocks: The first trial block per session was a “novel” block, in which crows had to learn by trial and error which of the two novel sample images (R1-B1) was associated with the red or blue choice, respectively (Fig. 1B, Left). The two test images were kept constant, while the sample images were exchanged for each new association block. After the crows acquired this novel association, a second, “familiar” trial block followed, during which the crows had to associate well-known paired associates (R-B) to the same two test items based on long-term memory, without a learning requirement (Fig. 1B, Center). After 60 correct trials in the familiar trial block, a further novel trial block was shown with new sample images. The crow again had to learn to associate samples with the proper choice items, followed by another familiar trial block.

Behavioral Performance. During novel block, both crows started at chance performance (45% and 52% on the first trial, both $P > 0.05$, binomial test against 50% chance) and reached $\sim 75\%$ correct after 70 trials (Fig. 1C; example session in Fig. S1A). We used a state-space model of dynamic learning (8) to estimate a learning criterion; that is, the first trial of each block after which the behavior of the crow was reliably above chance. This learning criterion was later used to compare neuronal activity during and after learning the novel associations. A learning criterion could be determined for 135 (91%) associations for crow B, and presented at randomized positions (Fig. 1A).

Two types of association tasks were shown in daily sessions in separate trial blocks: The first trial block per session was a “novel” block, in which crows had to learn by trial and error which of the two novel sample images (R1-B1) was associated with the red or blue choice, respectively (Fig. 1B, Left). The two test images were kept constant, while the sample images were exchanged for each new association block. After the crows acquired this novel association, a second, “familiar” trial block followed, during which the crows had to associate well-known paired associates (R-B) to the same two test items based on long-term memory, without a learning requirement (Fig. 1B, Center). After 60 correct trials in the familiar trial block, a further novel trial block was shown with new sample images. The crow again had to learn to associate samples with the proper choice items, followed by another familiar trial block.

Behavioral Performance. During novel block, both crows started at chance performance (45% and 52% on the first trial, both $P > 0.05$, binomial test against 50% chance) and reached $\sim 75\%$ correct after 70 trials (Fig. 1C; example session in Fig. S1A). We used a state-space model of dynamic learning (8) to estimate a learning criterion; that is, the first trial of each block after which the behavior of the crow was reliably above chance. This learning criterion was later used to compare neuronal activity during and after learning the novel associations. A learning criterion could be determined for 135 (91%) associations for crow B, and presented at randomized positions (Fig. 1A).

Two types of association tasks were shown in daily sessions in separate trial blocks: The first trial block per session was a “novel” block, in which crows had to learn by trial and error which of the two novel sample images (R1-B1) was associated with the red or blue choice, respectively (Fig. 1B, Left). The two test images were kept constant, while the sample images were exchanged for each new association block. After the crows acquired this novel association, a second, “familiar” trial block followed, during which the crows had to associate well-known paired associates (R-B) to the same two test items based on long-term memory, without a learning requirement (Fig. 1B, Center). After 60 correct trials in the familiar trial block, a further novel trial block was shown with new sample images. The crow again had to learn to associate samples with the proper choice items, followed by another familiar trial block.

Behavioral Performance. During novel block, both crows started at chance performance (45% and 52% on the first trial, both $P > 0.05$, binomial test against 50% chance) and reached $\sim 75\%$ correct after 70 trials (Fig. 1C; example session in Fig. S1A). We used a state-space model of dynamic learning (8) to estimate a learning criterion; that is, the first trial of each block after which the behavior of the crow was reliably above chance. This learning criterion was later used to compare neuronal activity during and after learning the novel associations. A learning criterion could be determined for 135 (91%) associations for crow B, and presented at randomized positions (Fig. 1A).
87 (85%) associations for crow L (learned associations; see SI Materials and Methods). 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).

The birds reached a mean performance of 85% and 86% after the criterion trial, with performance for all trials before the criterion at 53% and 54%, respectively. For the 10 trials immediately before and after the criterion, performance was 53% and 54%, respectively. The criterion was determined for each session separately from the succession of correct and error responses. (E) Reaction times for crow B and crow L before and after the criterion trial. Boxes show median and second and third quartile, whiskers show range of data or 1.5-times interquartile range. Significant difference: **P < 0.01.

NCL Neurons Show Associative Selectivity. We recorded the activity of 342 neurons in the NCL of two behaving carrion crows. The block structure of the task allowed us to compare neuronal responses for well-known and for new samples associated with the same test items. Many neurons seemed to respond to sample items as a function of the associated test item. We call this prospective encoding of the upcoming paired associate “associative selectivity.” For example, the neuron in Fig. 2A selectively increased its firing rate for the sample items associated with the red test item during the sample period of the familiar block and each of the novel blocks. Fig. 2B shows a neuron which selectively increased its firing rate in the delay period following all three samples associated with the blue test item.

To evaluate the prevalence of associative selectivity in the entire population, without relying on preselection of neurons or analysis windows, we visualized all recorded neurons’ responses to the six samples using principal component analysis (PCA)
associative selectivity. We defined the reference distribution based on the first and second familiar block were pooled to determine familiar selectivity. Recordings were stable across blocks; no significant difference in firing rates \( (P = 0.12, \text{Wilcoxon signed rank test}) \) or selectivity \( (P = 0.35, \text{Wilcoxon signed rank test}) \) (Fig. S3) were detected in selective neurons contributing two familiar blocks \( (n = 69) \). In these familiar-selective neurons associative selectivity often developed during novel blocks. For example, the neuron in Fig. 3 showed selective delay activity for the sample mapped onto the blue test item in familiar trials. In both novel blocks, this neuron also showed higher delay firing rates for the sample associated with the same test item. Furthermore, the difference between the samples increased over time in both blocks (Fig. 3).

We calculated the area-under-the-receiver operating-characteristic curve (AUROC) to quantify this increase in associative selectivity by comparing firing rates in a late delay period window for the two different samples in each block. AUROC is a measure for the discriminability of two distributions, with both 0 and 1 indicating perfect separation, and values of 0.5 indicating no selectivity. We defined the reference distribution based on the preferred sample in the familiar block, so that ROC values in the familiar block were always higher than 0.5, associative neurons would have ROC values higher than 0.5 in novel blocks, and novel blocks with opposite selectivity as the familiar block would have ROC values lower than 0.5. During learning, the mean novel block AUROC value of all familiar-selective cells was 0.54 (Fig. 4A; averaged across novel blocks for neurons which were recorded for two novel blocks; individual blocks are shown in Fig. S4A). After the criterion trial, the distribution of ROC values had shifted toward the ROC values during familiar blocks [mean 0.63 (Fig. 4A), familiar mean: 0.69 (Fig. S4B)], a significant increase \( (P < 0.01, \text{Wilcoxon signed rank test}) \) (Fig. 4B). A similar, but weaker, learning-related increase in prospective selectivity is present in the entire recorded population \( (n = 309, P < 0.01, \text{Wilcoxon signed rank test}) \) (Fig. S4C). The absence of neurons tuned in the opposite direction as the familiar block after learning in Fig. 4A indicates that all neurons that showed strong selectivity in novel blocks were associative, and these neurons were the ones driving the learning effects (Fig. S4D). The increase in associative selectivity came almost exclusively from an increase in discharge rate to the preferred sample after learning, with no change in firing rate to the nonpreferred sample (Fig. 4C).

**Associative Activity Develops Rapidly at the Start of New Blocks.** The trial-by-trial course of firing rates to the preferred and nonpreferred sample revealed an absence of selectivity in the first three trials \( (P > 0.05, \text{Wilcoxon signed rank test}, n = 83) \). However, firing to the nonpreferred sample sharply decreased during the first few trials of a new block, and increased for the preferred sample (Fig. 4D). Similarly, aligning ROC values by learning criterion showed that, on average, ROC values start out at 0.5 (not selective)
but begin increasing steadily well before the criterion is reached (Fig. 4E). In comparison with behavior aligned by the same criterion trial (Fig. 1D), associative selectivity changed less abruptly over the learning process.

**Association Strength Was Weaker or Reversed in Error Trials.** During the learning of novel associations, NCL neurons developed delay selectivity with the same group preference and similar selectivity strength as in familiar blocks. These results suggest that a task-relevant representation was formed in the NCL, which could be a neural signature of the crow preparing its response to the associated red or blue test item. To evaluate the behavioral relevance of this representation, we examined associative selectivity in error trials when the crow was preparing the wrong response. Fig. 5A shows the same example neuron as in Fig. 3 during novel blocks (there were not enough error trials in familiar blocks). In both blocks, delay period activity in error trials was opposite to that in correct trials, so that the neural activity was reflecting the bird’s behavior (red or blue choice) and not the presented sample item. In all neurons, the average firing rate to the preferred sample during error trials was reduced in novel and familiar blocks, and firing rate to the nonpreferred sample was increased in familiar blocks (all P < 0.05, Wilcoxon signed rank test) (Fig. 5B).

**Sample Period Activity Did Not Show Prospective Selectivity, and Did Not Change with Learning.** In summary, these results show that NCL neurons developed associative selectivity in the delay period during the learning of new associations, forming a task-relevant representation of the upcoming behavioral choice. When does this prospective representation emerge during the trial? Sample-selective neurons did not show similar tuning in novel and familiar blocks (Fig. 6A) (mean ROC in novel blocks: 0.52) and there was no change in associative selectivity with learning (Fig. 6B and C) (P > 0.48 Wilcoxon signed rank test). Similarly, there was no significant difference between firing rates in correct and error trials in any block (Fig. S5) (all P > 0.05, Wilcoxon signed rank test). Therefore, a neural signature of response preparation appeared only in the delay period window. Examining the temporal incidence of selectivity without relying on specific analysis windows also reveals no consistent difference in the point of time when selectivity appears within the delay period as learning progressed (Fig. 6D). Therefore, only the strength but not latency of the prospective choice representation in the NCL changes with learning.

**Discussion**

We report a neuronal correlate of association learning in corvid NCL, a cognitive integration area in the avian brain. Crows performed a DPA learning task that required mapping both familiar and novel samples to the same test items. The crows quickly acquired new associations, allowing the investigation of changes to the neuronal representation in behaving animals, as novel stimuli acquired behavioral meaning over the learning process. Single neurons prospectively encoded the upcoming behavioral choice during the delay but not during presentation of the sample. The selectivity strength of this task-relevant associative representation increased during learning, in parallel with the crows’ improved behavioral performance and faster reaction times. Prospective selectivity was weaker in error trials, indicating that the birds rely on this representation in the NCL when preparing and choosing a response.

The prospective representation of the behavioral choice and learning-related changes appeared in the end of the delay period, just before choice onset. No such representation was found in the sample period. Individual neurons exhibited strong selectivity for different sample pictures in the sample period (Figs. 2A and 6A and B), in agreement with our recent reports of sample-selective activity in working-memory tasks (9, 10). However, neurons did not group samples according to functional categories, and no consistent change of this selectivity was observed with learning (Fig. 6A and B). This finding suggests that selectivity in the sample period might be purely visual selectivity without connection to the associated test item or the crows’ behavior, an interpretation that is reinforced by the lack of error trial effects in this period.

This switch from a sample-based representation in the sample and early delay periods, and a test stimulus-based representation in the late-delay period corresponds well with behavioral data. Although we did not test the behavioral strategy used by the crows, behavioral evidence in pigeons suggests that they might switch from a retrospective to a prospective coding strategy depending on delay duration (11).
Monkeys seem to follow a similar behavioral strategy depending on delay duration (12). In primate prefrontal cortex (PFC), a functional analog of NCL, one study reports that prospective selectivity for long-term paired associates does not appear until the end of the delay period, whereas visual selectivity is the only factor influencing selectivity in the sample period (13). All studies of association learning in primate PFC, however, consistently report that such selectivity shifts earlier in the trial as the monkeys learn the correct response (14–17), eventually reaching the cue period, when the animal can first start to prepare its response. This is also a typical finding in other studies of the PFC, where choice-related activity and error trial effects get stronger throughout the delay, but are also present in the sample period (18–20). Our findings strongly indicate that an associative representation in the NCL does not emerge until stimulus similarity to the top-down signal and the latency of this selectivity does not seem to change with learning, in contrast to equivalent analyses of monkey PFC (Fig. 6D). Therefore, the time point when prospective activity appears is likely to reflect either different task-specific behavioral strategies or differences in general information processing patterns between primate and corvid brains.

How is association learning accomplished in mammalian and avian brains? Learned visual associations in monkeys are reflected by association-selective representations in higher visual areas of the temporal cortex (21, 22), with top-down activation by the PFC contributing to prospective representations of the choice picture to be recalled (19). Changes to stimulus-selective representations in higher sensory areas have also been reported in anesthetized starlings trained on auditory discriminations (23–25). Single neurons as well as population activity in auditory association areas contained information about the behavioral relevance of different sounds. Because we find a similar representation of novel and familiar stimuli with the same behavioral meaning, the NCL might be a source of an abstract task-relevant signal that biases representations in higher sensory areas to encode behaviorally meaningful stimuli (26). The PFC is intimately connected to all sensory association areas, putting it in a prime position for such executive control over sensory processing and behavior (6, 26).

However, the independent evolutionary origin of the NCL and PFC leads to certain anatomical differences, in particular their connectivity to the hippocampus. Primate hippocampus is a key brain area for association learning, with frequent demonstrations of reference-dependent activity (27–30). The PFC is intimately connected to the hippocampus, and hippocampal–prefrontal interactions seem to play a major role during association learning (17, 31). In contrast, the NCL has no direct connections to the hippocampus or surrounding structures (26, 32, 33). It is likely that this major anatomical difference could lead to differences in how the cognitive integration areas of birds and mammals process information during associative learning, such as the time point when prospective representations emerge in the trial.

Further differences to the PFC exist in the responses to familiar stimuli. After learning, novel stimuli were represented in the NCL in a categorical way according to their behavioral meaning, and resembling familiar reference samples with the same common associate. The formation of new associations between novel samples and familiar test items in our task likely involves different processes than the initial learning of the structure of the DPA learning task. Our results specifically show how novel stimuli can be rapidly linked to existing representations of upcoming choices, by activating neurons, which were already selective for the same associates. Therefore, the NCL forms an orderly prospective representation of different samples in a decision network of reference, which applies to both familiar and newly learned associations. Such a task-relevant representation shows general parallels between the NCL and association cortices in the primate brain (14, 16, 34, 35), but also sets it apart from neuronal encoding in the PFC, where familiar, “overlearned” stimuli typically evoke only weak activity (14, 36; but see ref. 37). In contrast, we found the strongest discharge rates and strongest selectivity for highly familiar stimuli, with newly learned stimuli building up to this reference level over the learning process. Such a prospective representation reflecting the strength of the connection between stimuli seems ideally suited for a brain area thought to be involved in organizing goal-directed behavior (38).

In summary, we have shown that single neurons in the NCL, an avian cognitive association area, form a prospective working-memory representation of upcoming behavioral choices. These neurons discriminate both novel and familiar sample images based on their behavioral meaning. During novel blocks, delay selectivity rapidly increases over the course of ~30 trials to resemble neuronal activity following highly familiar samples. These data suggest that, in the NCL, neurons play a role in learning to make decisions, a cornerstone of corvids’ remarkable behavioral flexibility and adaptability. The data highlight important parallels, but also differences, in the processing of learning-related activity in avian and mammalian executive brain areas.

**Materials and Methods**

**Subjects and Apparatus.** Two juvenile carrion crows (Corvus corone corone) were used in these experiments. The crows were maintained on a controlled ad libitum feeding schedule and earned rewards for participating in daily tests. All animal preparations and procedures were approved by the local ethical committee (Regierungspräsidium Tübingen) and authorized by the national authorities (Regierungspräsidium Tübingen). The crows were trained in a controlled operant conditioning chamber, stimuli were presented on a touch-screen monitor, and the crows responded by selecting the appropriate item on the screen. Reward for correct trials was delivered by a custom-built automated feeder below the screen. 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. For details see SI Materials and Methods.

**Behavioral Protocol.** The crows initiated a trial (Fig. 1) by moving their head into the light barrier as a go-stimulus (white square, 11 × 11 mm) was shown on the screen. The birds had to keep their head still throughout the trial; if their head exited the light barrier, the trial was aborted. After 200 ms, the go-stimulus turned off, followed by a presample period without visual stimulation. Next, a sample stimulus was presented in the center of the screen, followed by a delay. The duration of the presample period was 500 ms, sample period 500 ms, and delay period 1,000 ms. Sample pictures were two arbitrary images (20 × 20 mm) that were exchanged for each novel block, and two arbitrary images (20 × 20 mm, bird and flower) (Fig. 1B) that were kept constant in the familiar block for several months.

In the choice period, two test items (red triangle and blue cross) were presented on the left and right side of the screen, ~66 mm apart. The side of the screen on which each test item appeared on each trial was randomized and balanced. The crows indicated their choices by selecting one of the test images with their beak. If no response occurred within 1,700 ms, the trial was marked as incorrect. Correct choices were rewarded (indicated by Pavlovian reinforcement). Reward was given when the crow reached a performance above 80% during the last 40 trials. If the association was learned within 120 correct trials, the block was completed. If it was not learned within 120 trials, the block was continued until 180, 240, or 300 correct trials. Each novel block was followed by a familiar block, lasting 60 correct trials. In familiar blocks, two well-known sample pictures (RB) (Fig. 1B) with known associations to the same two test items were presented pseudorandomly interleaved. The familiar block was followed by another novel block with two new unknown sample pictures (R2/B2) (Fig. 1B), and another familiar block. A typical session consisted of two novel blocks, lasting 120 correct trials each, and two familiar blocks, lasting 60 correct trials each (Fig. 1B and Fig. 5B).

Veit et al.
Surgery and Recordings. All surgeries were performed while the animals were under general anesthesia and the crows received postoperative analgesics (SI Materials and Methods).

We recorded from eight chronically implanted electrodes on two custom-built microdrives targeting the NCL. The analysis includes all 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.

Data Analysis. Learning criterion. We used a state-space model of dynamic learning (8) to estimate a learning criterion from the behavioral data in each individual block: that is, the earliest trial when the bird was performing reliably above chance (SI Materials and Methods). For the purpose of neuronal analyses on well-controlled behavioral data, only novel blocks in which a learning criterion could be determined (91% and 85% of all associations in crow B and L, respectively) were considered. The actual learning behavior in individual sessions could be more variable (Fig. S1).

Analysis of delay-selective neurons. Neuronal activity was analyzed using unsmoothed firing rates in a 600-ms window starting 500 ms after delay onset and ending 100 ms after choice onset. Neurons were selected for further analyses based on their selectivity in the familiar block by comparing responses from both familiar blocks and calculating a rank sum test (P < 0.05, Mann–Whitney U test). 93 neurons (27%, 53 from crow B, 40 from crow L) significantly discriminated the two sample pictures in our analysis window. An association was included in the analysis if the neuron was held from trial 1 of the association until at least 10 trials after criterion was reached. During recording of the 93 selective neurons, 136 successful associations were presented, which contributed two associations to further analyses, and 30 neurons contributed one association. All analyses were performed for individual associations and then averaged for statistics and display for those neurons which contributed two associations.

Sample-size analysis. Sample-size analysis was conducted using unsmoothed firing rate in a 500-ms window, starting 100 ms after sample stimulus onset, ending 100 ms after sample stimulus offset. We selected 138 neurons (40%) based on their selectivity in the familiar block in the sample period (P < 0.05 rank sum test); 202 successful associations were presented during the recording of these 138 neurons (74 neurons contributed two associations and 51 contributed one). All analyses were performed as described for the delay window.

Population analysis. PCA. We performed PCA of trial-averaged, smoothed, and normalized population activity to reduce dimensionality of the high-dimensional space spanned by all neurons’ firing rates in Fig. 2C (SI Materials and Methods).

Distance. We calculated Euclidean distance in the high-dimensional space (without dimensionality reduction) between all pairs of samples from different blocks to measure difference in population activity to different samples. We then averaged the distances for all samples associated with the same test item (within group distance) and all samples associated with different test items (between group distance).

ROC analysis. The quality of selectivity for each unit was quantified using ROC analysis (SI Materials and Methods). The sample stimulus (R or B) that elicited the lowest firing rate in the familiar block was used as reference (noise direction). Therefore, an unselective neuron had ROC values higher than 0.5 in familiar blocks. Selectivity in novel blocks could range from 0 to 1, with 0.5 indicating no selectivity and values higher than 0.5 indicating associative selectivity (i.e., the same preference as in the familiar blocks).

ACKNOWLEDGMENTS. This work was supported by a doctoral fellowship from the German National Academic Foundation (to L.V.); and by Deutsche Forschungsgemeinschaft Grant NI 618/3-1 (to A.N.).