Emergence of learned categorical representations within an auditory forebrain circuit

James M. Jeanne[1,2,3], **Jason V. Thompson**[1,3], **Tatyana O. Sharpee**[1,2], and **Timothy Q. Gentner**[1,3,*]

[1]Neurosciences Graduate Program, University of California San Diego, La Jolla, CA 92093, USA

[2]Computational Neurobiology Laboratory, Salk Institute for Biological Studies, La Jolla, CA 92037, USA and the Center for Theoretical Biological Physics, University of California, San Diego, La Jolla CA 92093

[3]Department of Psychology, University of California San Diego, La Jolla, CA 92093, USA

Abstract

Many learned behaviors are thought to require the activity of high-level neurons that represent categories of complex signals, such as familiar faces or native speech sounds. How these complex, experience-dependent neural responses emerge within the brain's circuitry is not well understood. The caudomedial mesopallium (CMM), a secondary auditory region in the songbird brain, contains neurons that respond to specific combinations of song components and respond preferentially to the songs that birds have learned to recognize. Here, we examine the transformation of these learned responses across a broader forebrain circuit that includes the caudolateral mesopallium (CLM), an auditory region that provides input to CMM. We recorded extracellular single-unit activity in CLM and CMM in European starlings trained to recognize sets of conspecific songs and compared multiple encoding properties of neurons between these regions. We find that the responses of CMM neurons are more selective between song components, convey more information about song components, and are more variable over repeated components than the responses of CLM neurons. While learning enhances neural encoding of song components in both regions, CMM neurons encode more information about the learned categories associated with songs than CLM neurons. Collectively, these data suggest that CLM and CMM are part of a functional sensory hierarchy that is modified by learning to yield representations of natural vocal signals that are increasingly informative with respect to behavior.

Keywords

Auditory; Categorical representation; Vocal recognition; Songbird; Experience-dependent plasticity; Hierarchical coding; Communication

Introduction

Faced with an immense quantity of sensory input, individuals learn to identify sensory features relevant to behavioral goals and ignore other features. The representation of objects by high-level sensory neurons depends heavily on this form of learning (e.g. Rolls et al., 1989; Sigala and Logothetis, 2002; Gentner and Margoliash, 2003) and some neurons

*To whom correspondence should be addressed. tgentner@ucsd.edu.

represent learned categories of objects rather than the objects themselves (Freedman et al., 2001; Sigala and Logothetis, 2002; Freedman and Assad, 2006). These kinds of complex representations likely emerge from processing pathways in which higher-order neurons integrate convergent input from lower-order neurons (Felleman and Van Essen, 1991; Binder et al., 2000; Kaas and Hackett, 2000; Riesenhuber and Poggio, 2000; Chechik et al., 2006; Rauschecker and Scott, 2009). However, very few studies have examined changes in the encoding of natural stimuli along these processing pathways (Chechik et al., 2006; Rust and DiCarlo, 2010) or how learning mediates this encoding (Freedman et al., 2003; Freedman and Assad, 2006), particularly in the auditory domain.

Songbirds serve as an excellent model system to study the learning-dependent neural encoding of natural signals because they easily learn to identify conspecific songs (Gentner and Margoliash, 2003) and have well-defined neural circuitry specialized for processing songs (Hsu et al., 2004; Woolley et al., 2005; Woolley et al., 2009). The caudomedial mesopallium (CMM), a secondary auditory area, contains some of the most complex neurons in the avian auditory system (Meliza et al., 2010), which respond stronger to learned songs than to novel songs (Gentner and Margoliash, 2003). As with high-level brain regions in mammals, however, the circuitry that produces the experience-dependent CMM responses is not well understood. CMM receives indirect input from Field L (the analogue of mammalian primary auditory cortex) via bidirectional connectivity with the adjacent caudolateral mesopallium (CLM) and caudomedial nidopallium (NCM; Fig 1a; Vates et al., 1996). The responses of CLM neurons are less well predicted by linear receptive field models (Sen et al., 2001) and are more sharply tuned to the statistics of song (Hsu et al., 2004) than many Field L neurons. Although processing in CLM may contribute to the learning-dependent representations of song in CMM, it is unknown whether the encoding of songs in CLM differs from CMM or whether CLM encodes learning-dependent information.

To understand the functional relationship between CLM and CMM, we compare their neural encoding properties in European starlings that have learned to recognize sets of conspecific songs and ask whether learning modifies the responses of neurons differently between CLM and CMM. We show that the responses of CMM neurons are more selective between song components, convey more information about song components, and respond with greater variability to repeated song components. In both regions, learning increases the information encoded about song components, but CMM neurons encode more information about behaviorally defined song categories. These results are consistent with a model of neural processing in which information about natural vocal signals flows through CLM to CMM, giving rise to complex representations of acoustic signals and their behavioral relevance.

Materials and Methods

All experiments were performed in accordance with the Institutional Animal Care and Use Committee of the University of California, San Diego.

Subjects

Nineteen adult European Starlings (*Sturnus vulgaris*) were wild-caught in southern California, and housed in aviaries with free access to food and water until the commencement of behavioral training. At the start of training, subjects were naïve to all experimental procedures and stimuli. Thirteen starlings were used for CLM experiments, five starlings were used for CMM experiments, and one starling was used for both CLM and CMM experiments. Data from CMM experiments were combined with a subset of previously published data (Gentner and Margoliash, 2003) from an additional four starlings, yielding a total of 23 subjects. Very few differences in behavioral training were observed

between the two sets of CMM data (supplemental table 1). 28 of the 48 CMM neurons reported here were from this previously published data.

Stimuli

Six starling song stimuli were created from a collection of songs previously recorded from four adult male European starlings. Each stimulus was a unique section of continuous song (durations ranging from 9.1s – 10.7s) from a single male and shared no motifs with any other stimulus. The six song stimuli were divided into three sets of two song stimuli each. For each experimental subject, the three stimulus sets were assigned as "rewarded," "unrewarded," and "novel" stimuli to reflect the subject's experience with those stimuli during behavioral training. Across subjects, this assignment was counterbalanced such that the same stimuli were used for different conditions in different birds. Nearly identical stimuli were used for CLM and CMM experiments.

Behavioral training

After acclimation to individual housing in a sound-attenuating chamber (Acoustic Systems, ETS-Lindgren, Austin, TX), each subject was trained on a standard go/no-go operant-conditioning procedure to classify two of the song stimulus sets (Gentner and Margoliash, 2003). For each bird, the rewarded songs were assigned as the "go" stimulus set and the unrewarded songs were assigned as the "no-go" stimulus set. A subject started a trial by inserting its beak into a small hole on a response panel inside the sound-attenuating chamber (Fig 1b), which initiated the playback of one of the four stimuli, chosen at random. After stimulus playback ended, a subject had two seconds to report its classification decision by either inserting its beak again (a "go" response) or by refraining from inserting its beak (a "no-go" response). Go responses to the go stimulus set were rewarded with two-second access to food and go responses to the no-go stimulus set were punished with variable-duration darkness (range, 10s – 90s) during which no food was available and trials could not be initiated. In all cases, no-go responses were neither rewarded nor punished. Water was provided *ad libitum*, but food was only available from correct go responses.

Neurophysiology

After achieving satisfactory classification performance (see results), each subject underwent surgery under isoflurane anesthesia (1.5–2% concentration) to prepare for recording. The top layer of skull was removed from the region above CLM or CMM and a small metal pin was affixed to the skull just caudal to the opening. Each subject recovered for 12–24 hours before neurophysiological recordings began. On the morning of the recording day, the subject was anesthetized with urethane (20% by volume, 7mL/kg), and head-fixed in a stereotactic apparatus inside a sound-attenuating chamber. The subject was situated 30 cm from a speaker through which the song stimuli were presented (all normalized to 95dB peak SPL).

Extracellular electrical activity of single neurons in CLM or CMM in response to 5–50 repetitions of all six song stimuli (presented in random order) was recorded using glass-coated platinum-iridium wire electrodes (1–3MΩ impedance) inserted through a small craniotomy directly dorsal to CLM or CMM. The extracellular waveform was amplified (5,000X – 50,000X gain), filtered (high pass 300Hz, low pass 3kHz), sampled (25kHz), and stored for offline analysis (Cambridge Electronic Design, Cambridge, UK). At the end of the recording session, one to three fiduciary electrolytic lesions (10–20µA, 10–20s) were made to facilitate recording site localization using standard histological techniques.

All recording sites were confirmed to be within CLM or CMM (supplemental Fig. 1c). All CLM neurons were located between 1200µm and 1650µm from the midline; all CMM

neurons were located between 0µm and 1000µm from the midline (Fig. 1d). Putative action potentials from single neurons were identified by amplitude and sorted offline using principal component analysis on waveform shape (Cambridge Electronic Design, Cambridge, UK). Only action potential waveforms with a single obvious cluster in principal component-space and with very few refractory-period violations (supplemental Fig. 1a,b) were considered to be from a single neuron. Only 4.8% (7/145) of all neurons exhibited any inter-spike intervals (ISIs) shorter than 1ms, and within that subset of neurons, ISIs shorter than 1ms accounted for less than 0.1% of all ISIs in each neuron.

Only neurons that responded significantly to any part of our stimulus set were included in our data analysis. Response significance was determined quantitatively, following previously described methods (Gentner and Margoliash, 2003). Briefly, the mean response to each song was divided into 500ms segments, and the variance of the mean response over each segment was computed. The variance for each song was normalized by the variance of spontaneous firing. To be considered auditory, the largest normalized variance value needed to be greater than $(1.96 \times 1 \text{ S.E.}) + 1$, where S.E. is the standard error of the normalized variance values for the remaining songs, and 1 is the normalized variance that would be expected for a non-auditory neuron.

Data Analysis

Analyses were performed using custom-written MATLAB (Natick, MA) software. Behavioral performance was evaluated using d prime (Macmillan and Creelman, 2005),

$$d' = zscore(Hit\ rate) - zscore(False\ alarm\ rate) \tag{1}$$

a measure of discriminability between two distributions. Values of d prime were computed in non-overlapping blocks of 100 trials.

For most analyses of neural activity, responses to full song stimuli were segmented into the responses to each constituent motif (Fig. 2). The starting time for a motif was defined as the onset of power for that motif, and the ending time for a motif was defined as the onset of power for the following motif. Thus, a neuron's response to a motif consisted of the neural activity both during that motif and during the subsequent silent period following that motif. Each individual song stimulus contained multiple renditions of some motifs, but because of subtle acoustical variability between renditions, each rendition was considered to be distinct for all analyses except the variability analysis of repeated motifs (Fig. 4).

Motif selectivity analysis—The non-parametric selectivity of each neuron was calculated over all motifs (rewarded, unrewarded, and novel) collectively for each neuron (Rolls and Tovee, 1995; Vinje and Gallant, 2000):

$$S = \left(1 - \frac{(\sum r_i/n)^2}{\sum (r_i^2/n)}\right) \Big/ \left(1 - \frac{1}{n}\right), \tag{2}$$

where r_i is the mean firing rate in response to the ith motif, and n is the total number of motifs. This measure ranges from 0 to 1, with 0 representing the minimum motif selectivity (responses to all motifs are identical) and 1 representing the maximum motif selectivity (response to only one motif). Although this measure includes all responses from each neuron, it emphasizes the larger values in the response distribution. Therefore, we also used the entropy method (Lehky et al., 2005), which equally considers selectivity of both excitatory and suppressive responses. Both measures yielded the same result (Supplemental

Fig. 2). Mean spontaneous firing rates did not differ significantly between CMM (3.24±0.59 Hz) and CLM neurons (3.37±0.34 Hz ; Wilcoxon rank sum test, p = 0.41).

Information analysis—Firing rates for each neuron's evoked response to each motif were divided into six equally spaced bins ranging from the lowest firing rate to the highest firing rate elicited by that neuron. Six bins were chosen to balance the need to capture the dynamic range of responses for each neuron with the need to appropriately sample the conditional probabilities. For each neuron, identical firing rate bins were used for all information calculations. Control analyses in which the number of bins varied from 2–10 yielded changes in the absolute number of bits but did not alter the effect of learning on information or the differences between CLM and CMM (Supplemental Fig. 3). In all cases, mutual information was calculated as

$$I(s;r) = \sum_{s,r} p(s,r) \log_2 \left(\frac{p(s,r)}{p(s)p(r)} \right).$$

(3)

Where s indexes the stimulus and r indexes the bin of the firing rate response (Brenner et al., 2000; Cover and Thomas, 2006).

For mutual information between motif identity and motif firing rate, $p(r,s)$ represented the empirical joint probability distribution of motif firing rates and motif identities. Because multiple renditions of single motifs were considered to be distinct, the distribution of motif identities, $p(s)$, was always uniform. The distribution of firing rates, $p(r)$, was computed by averaging across the conditional distributions for each motif. Information about motif identity was computed in two ways: across responses to all motifs regardless of stimulus class (rewarded, unrewarded, and novel), and separately for the motifs in each stimulus class. For information encoded about all motifs (Figure 3), $p(r)$ was found by averaging the conditional distributions across all motifs presented to each neuron. For the information encoded about motifs within each stimulus class (Figures 5,6), the conditional distributions were averaged across motifs within each class to obtain separate $p(r)$ distributions for each class.

For mutual information between motif class (rewarded, unrewarded, novel) and motif firing rate, $p(r,s)$ represented the empirical joint probability distribution of motif firing rates and each motif's associated stimulus class. This distribution was determined by averaging the response distributions for all motifs within each stimulus class. The distribution of stimulus classes, $p(s)$, was uniform and the distribution of firing rates, $p(r)$, was found by averaging the class-conditional distributions.

For mutual information between song class and song firing rate, calculations were exactly analogous to calculations for motif class, except that firing rate responses were averaged over the entire duration of the presented song.

The significance of information encoded about learned motif and song categories was evaluated relative to information about randomly shuffled categories. For shuffled motif category information, each motif was randomly assigned to one of three categories such that the behavioral meaning of the categories was lost but the association between motif identity and firing rate was preserved (because single trials were not shuffled). This procedure was repeated for 100 random permutations to generate a distribution of shuffled information values, providing an estimate of the information encoded about arbitrary categories. The distribution of shuffled values was then used to determine the p-value for the category information for each neuron. Significance was evaluated at p < 0.05. For song category

information, shuffling was conducted similarly. However, because there are only two songs per category, there are only eight distinct permutations of category assignment that disrupt all three original category boundaries. Thus, information about learned categories was considered significant when it was greater than the information about all eight shuffled information values for each neuron. Importantly, because neurons could encode any arbitrarily-defined categories, positive information about shuffled categories does not imply the presence of residual bias in the estimation of information about behaviorally-defined categories. Rather, neurons may "categorically" encode other features of song, such as a particular spectrotemporal pattern that only appears in some motifs. Making a comparison with the shuffled categories thus provides a test of whether the information encoded about the learned categories is greater than would be expected by chance.

Two control analyses for the mutual information were carried out. First, to ensure that mutual information values were not solely dependent on variations in a neuron's dynamic range, we limited responses by the maximum and minimum single-trial firing rate elicited by a novel motif (Supplemental Fig. 4). In this case, all firing rates from rewarded or unrewarded motifs that were outside this range were ignored (i.e. considered never to have occurred). Second, to test our assumption that responses to multiple renditions of acoustically similar motifs are independent, mutual information was recomputed with multiple renditions considered as a single motif. Both controls yielded qualitatively similar results to those reported in the text (Supplemental Fig. 5)

In addition, because estimates of information from limited samples are inherently biased upwards (Treves and Panzeri, 1995), bias was corrected by extrapolating the information estimate for each neuron to an infinite data size (Strong et al., 1998; Brenner et al., 2000). Variability in the estimate of mutual information was determined by a jackknife resampling of the data. The standard deviations of the information estimations were very low (median for estimates over all motifs: 0.014 bits, median for estimates within motif classes: 0.022 bits), which indicates highly stable estimations. Subtracting analytical estimates of the bias (Panzeri et al., 2007) instead yielded qualitatively similar results to the extrapolation method.

Entropy values were computed using the same bins as were used for information calculations. Total entropy was calculated as

$$H_{Total}(r) = -\sum p(r)\log_2 p(r). \tag{4}$$

Noise entropy was calculated as

$$H_{Noise}(r) = -\sum_s p(s) \sum_r p(r|s)\log_2 p(r|s). \tag{5}$$

Because estimates of the entropy are subject to similar biases as estimates of the mutual information, bias in the computed entropy was corrected by using methods analogous to those described for mutual information.

Repeated motif analysis—Four independent observers classified motifs based on visual inspection of the spectrogram and listening to the waveform. Repeated motifs were judged to be renditions of the same type only when all four observers agreed on the classification. Motifs of the same type were then considered to be identical for the purposes of the repeated motif analyses. Variability of responses within sequences of repeated motifs was measured

by computing the coefficient of variation (CV) of trial-averaged firing rates in response to each repeated motif sequence presented to each neuron. The CV values for all sequences for a given neuron were averaged to find the mean variability for that neuron.

Statistics—All data were tested for normality using the Lilliefors test with $p < 0.05$. Non-parametric tests were applied when data were not normal. Central tendencies are reported as means ± standard errors, unless otherwise noted.

Results

To compare how learning affects neural encoding in CLM and CMM, we began all experiments by training European starlings (*Sturnus vulgaris*) to recognize four different conspecific songs using an established operant procedure (Fig. 1b) (Gentner and Margoliash, 2003). Birds learned to peck in response to one pair of songs ("rewarded" songs) to obtain a food reward and to withhold pecks to the other pair ("unrewarded" songs) to avoid a mild punishment (Materials and Methods). After birds learned this task (Fig. 1c), we analyzed the activity of single neurons (supplemental Fig. 1) in either CLM (n = 97 neurons) or CMM (n = 48 neurons) (Fig. 1d) in response to the rewarded and unrewarded training songs and to two songs with which the birds had no prior experience ("novel" songs; Fig. 2; Materials and Methods).

The birds used for CLM recordings learned at similar rates and ultimately reached similar levels of performance as the birds used for CMM recordings. Song recognition performance (measured by d-prime; Materials and Methods) exceeded chance by a significant margin (p < 0.01) after a mean of 814±100 trials in CLM birds and a mean of 900±87 trials in CMM birds (Fig. 1c; t-test, p = 0.55). CLM birds performed a mean of 22,639±8,713 trials and CMM birds performed a mean of 28,987±10,913 trials (t-test, p = 0.65). By the end of training, both sets of birds recognized the training songs with high accuracy (mean d' over the last 500 trials: 2.68±0.20 in the CLM birds and 2.90±0.36 in the CMM birds, t-test, p = 0.57). Thus all birds had learned to recognize all the training songs with high proficiency prior to the neural recording.

Motif Selectivity in CLM and CMM

We first sought to characterize functional differences between neurons in CLM and CMM. Because starlings compose their songs from stereotyped clusters of notes called motifs (Chaiken et al., 1993) that are thought to be perceived as discrete auditory objects (Gentner and Hulse, 2000; Gentner, 2008; Seeba and Klump, 2009), we analyzed the neural responses to each motif. In some CMM neurons, small subsets of motifs elicit high firing rates, while many other motifs elicit low firing rates (Gentner and Margoliash, 2003; Meliza et al., 2010), a characteristic known as lifetime sparseness or non-parametric selectivity (Willmore and Tolhurst, 2001; Lehky et al., 2005). Here we refer to this characteristic as "motif selectivity," and compare this measure between the responses of neurons in CMM and CLM to the motifs that made up all four training and the two novel songs. In the representative CLM neuron shown in figure 2a, a large number of motifs elicited high firing rates. In the representative CMM neuron (Fig 2b), however, a smaller number of motifs elicited high firing rates, while most motifs elicited low firing rates. Differences in the distribution of each neuron's firing rate response can be directly compared between the sample CLM and CMM neurons by rank-ordering the responses for each motif (Fig 3a).

For each neuron in CLM and CMM, we computed the non-parametric motif selectivity (Vinje and Gallant, 2000; Materials and Methods). This measure quantifies the relative extent of the positive tail of the distribution of mean firing rates in response to motifs (Franco et al., 2007). If all motifs elicited the same firing rate, the motif selectivity would be

0; if only one motif elicited a positive firing rate, the motif selectivity would be 1. The CLM neuron in Fig 2a had a motif selectivity of 0.24, whereas the CMM neuron in Fig 2b had a motif selectivity of 0.49. Although both regions exhibited a large range of motif selectivity values, we found that, on average, neurons in CMM had higher motif selectivity values (0.40±0.04) than neurons in CLM (0.26±0.02; Wilcoxon rank sum test, p = 1.6×10^{-4}; Fig 3b). These motif selectivity differences (and those observed using other measures of selectivity; Supplemental Fig. 2) reflect the observations that CMM neurons responded with higher firing rates to a smaller subset of motifs than CLM neurons.

Motif information encoding in CLM and CMM

Responding selectively to a small subset of motifs is one effective way to encode information, but responding to many motifs can also encode a substantial amount of information, provided the response to each motif is distinct. Mutual information captures all the differences between the responses to different motifs (Cover and Thomas, 2006) and is thus agnostic to the actual method of encoding. We computed mutual information between firing rate and motif identity from the probability distribution of firing rate conditioned on motif identity (Materials and Methods). The conditional firing rate distributions of the representative neurons from Fig 2 show that the CMM neuron exhibited a greater diversity of firing rates than the CLM neuron, and that this diversity was more closely tied to motif identity for the CMM neuron than for the CLM neuron (Fig 3a). Accordingly, the CMM neuron encoded 0.76 bits of information about motif identity, whereas the CLM neuron encoded only 0.38 bits of information. Across all neurons, we observed a broad distribution of information values in both CLM and CMM (Fig 3c). On average, neurons in CMM encoded significantly more information (0.55±0.03 bits) about motif identity than neurons in CLM (0.38±0.02 bits; Wilcoxon rank sum test, p = 3.5×10^{-5}; Fig 3c).

In both CLM and CMM, high motif selectivity and high motif information did not coexist in the same neurons (Fig 3d). This is expected because each of these measures mutually constrains the other. Motif selectivity measures the distinctness of a neuron's response to a small number of motifs, while mutual information measures the diversity of a neuron's response to many motifs. The response of a highly selective neuron can effectively distinguish between a small number of motifs, yet has little ability to distinguish between the majority of motifs. High motif selectivity thus constrains the amount of information that can be conveyed about the whole set of motifs. Similarly, neurons that encode a large amount of information must necessarily respond to many motifs, but in a manner that maps different responses to different motifs. The populations of neurons from both CLM and CMM range from low motif selectivity but high information, to high motif selectivity but low information (Fig 3d). This pattern suggests that neurons in both regions possess a continuum of sensory encoding properties. In addition, because CMM contains fewer neurons with both lower motif selectivity and lower information values than CLM (Fig. 3d), CMM neurons encode motifs in a manner that is closer to the constraints set by information and selectivity.

Responses to repeated motifs in CLM and CMM

Over the course of a song, starlings typically sing multiple renditions of one type of motif before switching to a different type of motif (Eens, 1997). We examined whether CLM and CMM neurons elicited variable responses to repeated motifs of the same type. Evidence of such variable responses was found in both regions. The responses of the representative CLM and CMM neurons in Fig 4a and b show change across each rendition of the repeated motif types in the sequence. To quantify this observation, we computed the coefficient of variation (CV) of mean firing rates in response to each motif within a repeated sequence. For the sample CLM neuron in Fig 4a, CV values were 0.40, 0.22, and 0.30 for motif types A, B, and C, respectively. For the sample CMM neuron in Fig 4b, CV values were 0.54, 1.15, and

0.65, for motif types A, B, and C, respectively. We observed instances of firing rates increasing over motif repetitions (e.g. motif type A in Fig 4a), as well as instances of firing rates decreasing over motif repetitions (e.g. motif type C in Fig 4b). We computed a measure of each neuron's overall variability to repeated motifs by averaging all the CV values for each sequence of repeated motifs presented to that neuron. On average, the neurons in CMM had higher mean CV values (0.40±0.03) than the neurons in CLM (0.32±0.02; Wilcoxon rank-sum test, p = 0.0040; Fig 4c). These results suggest that neurons in CMM are more variable within a motif type than are neurons in CLM. Together, the results of motif selectivity, mutual information, and variability analyses indicate that song-evoked neuronal responses increase in complexity between CLM and CMM.

Learning increases information encoded about motifs in CLM

One way in which learning might act on CLM neurons is to modify their encoding of individual motifs. To test this idea, we compared the responses of CLM neurons to the motifs that were paired with reward during training ("rewarded motifs"), the motifs not paired with reward during training ("unrewarded motifs"), and the motifs not used for training ("novel motifs"). In the representative CLM neuron in Fig 2a, neural activity was more variable among the rewarded motifs than among the unrewarded or novel motifs. We quantified these differences by computing the mutual information between firing rate and motif identity separately for each class of motifs (Materials and Methods). The strength of the association between firing rates and motif identity within each stimulus class corresponds directly to the amount of information each neuron encodes about the motifs in that class. The conditional probability distributions of firing rates for each set of motifs presented to the example neuron (Fig 5a) shows that the firing rate diversity was more closely tied to motif identity for familiar motifs than between the unrewarded or novel motifs. Accordingly, this neuron encoded more information about rewarded motifs (0.66 bits) than about unrewarded motifs (0.16 bits) and novel motifs (0.26 bits).

Across the population of CLM neurons, learning significantly increased the amount of information encoded by individual neurons (Friedman test, p = 1.7×10^{-4}). The mean information encoded about rewarded motifs was 34.5% higher than that for novel motifs (Wilcoxon signed-rank test, p=1.6×10^{-5}; Fig 5b,c). This learning effect was observed in most neurons: 69 of 97 neurons (71%) encoded more information about rewarded motifs than novel motifs (χ^2 test: p= 0.0001; Fig. 5c). The mean information encoded about unrewarded motifs was comparable to that for novel motifs (Wilcoxon signed-rank test, p=0.091; Fig. 5c). The proportion of neurons encoding more information about unrewarded motifs than about novel motifs was not greater than that expected by chance (56/97; χ^2 test, p = 0.13). Thus, the association of songs with reward was necessary to induce significant changes in encoding by single CLM neurons.

The observed effects of learning on information encoding could arise from two forms of firing rate variability. First, the diversity of responses to all motifs ("the total entropy") could increase, which could potentially allow for a greater number of motifs to be represented. Second, the diversity of responses to repeated presentations of the same motif ("the noise entropy") could decrease, which could allow for greater discriminatory power between responses to different motifs (Strong et al., 1998). Across all CLM neurons, learning increased the total entropy (repeated measures ANOVA, p = 2.1×10^{-4}), but had no effect on the noise entropy (repeated measures ANOVA, p = 0.13; Fig 5d). Learned motifs thus elicited a greater diversity of responses than novel motifs without compromising the reliability of responses, which increased the capacity of CLM neurons to convey information about learned stimuli.

We then asked what drives the increase in the total entropy of the neural response distribution. In principle, this change may be due to an increase in the total range of responses or to an increase in the number of distinct spike rates observed within a fixed range. Across all CLM neurons, we observed a slightly larger range of firing rates for rewarded motifs (16.5±0.9 Hz) than for unrewarded (16.2±0.8 Hz) or novel motifs (15.2±0.9 Hz; Friedman test: $p = 0.005$). This increased range, however, did not fully account for the increased information encoding. The main effect of learning was unaltered in a control analysis where we omitted any response that fell outside the range of firing rates elicited by the novel motifs (Supplemental Fig. 4). Learning, therefore, increased the amount of information encoded by CLM neurons primarily by increasing the effectiveness with which this range was used.

Learning increases information encoded about motifs in CMM

Because of the substantial effects of learning on motif encoding in CLM, we next investigated whether learning also modified information encoding about individual motifs in CMM. Fig 6a shows the conditional probability distributions of firing rates for each set of motifs presented to the CMM neuron illustrated in Fig 2b. As with most CLM neurons, the diversity in this CMM neuron's firing rates was more closely tied to rewarded motifs than to unrewarded or novel motifs. Accordingly, this neuron encoded more information about rewarded motifs (0.91 bits) than about unrewarded motifs (0.60 bits) and novel motifs (0.09 bits). Over our entire sample, learning significantly modulated the information encoded by CMM neurons (repeated measures ANOVA, $p = 0.016$; Fig 6b,c). The mean amount of information encoded about rewarded motifs was 27.7% higher than that for novel motifs (paired t-test, $p = 0.012$; Fig 6c). Like CLM, the information encoded about unrewarded motifs was comparable to that encoded about novel motifs (paired t-test, $p = 0.41$). Although CMM encoded more information than CLM on average (Fig 3c), the effects of learning on neural encoding in both regions were comparable (mixed model ANOVA, interaction term, $p = 0.65$).

Across all CMM neurons, learning increased the total entropy (repeated measures ANOVA, $p = 0.019$), but had no significant effect on the noise entropy (Friedman test, $p = 0.09$; Fig 6d). Therefore, as in CLM neurons, learned motifs elicited a greater diversity of responses than novel motifs in CMM neurons, but this greater diversity did not substantially decrease the reliability of responses. We observed no significant learning-dependent increase in the range of firing rates evoked by single motifs (repeated measures ANOVA, $p = 0.08$). Thus, learning and reward enhanced neural encoding of motifs in similar ways in both CLM and CMM. In both regions, the effect of learning on information encoding was not dependent on our assumption that responses to repeated motifs are independent (Materials and Methods); similar differences were observed when repeated motifs are considered to be identical in the mutual information analysis (Supplemental Fig. 5).

Learning increases information encoded about motif categories

In addition to encoding more information about the identity of learned motifs than about novel motifs, CLM and CMM might also specifically encode information about the behaviorally relevant categories for motifs acquired through training (i.e rewarded, unrewarded, and novel). Such encoding could appear as any consistent difference in responses to motifs from different categories and thus is distinct from the foregoing analysis of information about motif identity. To explore this possibility, we formed firing rate distributions from responses to all motifs from these three categories (Fig 7a). From these distributions, we computed the information encoded by single neurons in CLM and CMM about the behaviorally-defined category of each motif. Because the behaviorally-defined categories are just one way that groups of motifs might be represented, we compared the

information about learned categories to the distribution of information values when the category membership of each motif was randomly shuffled into other groupings (Materials and Methods). The example CMM neuron depicted in Fig 7a,b encoded 0.22 bits of information about motif category and only 0.03±0.03 (mean ± s.d.) bits about the randomly shuffled categories. Because different motifs can elicit very different firing rates in the same neuron, the information about motif category is small relative to information about motif identity. Nonetheless, the information about learned categories was significantly greater than the information about shuffled categories (evaluated at $p < 0.05$) in 45.8% (22/48) of CMM neurons and in 28.9% (28/97) of CLM neurons. These proportions are larger than would be expected by chance. On average, the information about behaviorally relevant categories of motifs was larger than the mean information about shuffled categories for neurons in both CLM (0.023±0.002 bits vs. 0.014±0.001 bits; Wilcoxon signed rank test, CLM: $p = 7.3 \times 10^{-8}$; Fig. 7b) and CMM (0.072±0.009 bits vs. 0.023±0.002 bits; Wilcoxon signed rank test, $p = 1.2 \times 10^{-8}$; Fig. 7c). Correspondingly, neurons in CMM encoded significantly more information about behaviorally relevant categories than neurons in CLM on average (Wilcoxon rank sum test: $p = 1.95 \times 10^{-7}$; Fig 7d). Thus, while both CLM and CMM encode information about the learned motif categories, neurons in CMM encode significantly more of this information than neurons in CLM.

Learning increases information encoded about song categories

Because the firing rate of CLM and CMM neurons typically varies substantially over the motifs within a song (e.g. Fig 2) while the behavioral category (i.e. rewarded, unrewarded, or novel) remains unchanged, we reasoned that the firing rate averaged over the course of the song may better represent the behavioral category. As for motifs, we formed firing rate distributions from responses to each song from the three behavioral categories as well as for all permutations of shuffled categories (Fig 8a; Materials and Methods). From these distributions, we computed the information encoded by single neurons in CLM and CMM about the learned categories and shuffled categories. Because there are only two songs per category, there are only eight distinct permutations in which all categories are shuffled. Thus, we compared the information about the learned categories with the information about the shuffled set of categories that encoded the maximum information. The example CMM neuron depicted in Fig 8a encoded 0.99 bits of information about song category and a maximum of 0.68 bits about randomly shuffled categories. We found that the information about learned categories was significantly greater than the maximum information about shuffled categories in 37.5% (18/48) of CMM neurons but in only 22.7% (22/97) of CLM neurons. These percentages reflect significant differences between the populations of CLM and CMM neurons. In CMM, on average, information about the song category (0.42±0.04 bits) was significantly greater than the mean information about shuffled categories (0.28±0.03 bits; paired t-test: 4.7×10^{-4}; Fig 8c) In CLM, in contrast, information about the song category (0.19±0.03 bits) was similar to the mean information about shuffled categories (0.17±0.02 bits; paired t-test: $p = 0.11$; Fig 8c). Correspondingly, neurons in CMM encoded more category information than neurons in CLM on average (Wilcoxon rank sum test: $p = 7.4 \times 10^{-7}$; Fig 8d). This difference can also be observed by comparing the mean firing rates between the rewarded and novel songs (Fig 8e) and between the rewarded and unrewarded songs (Fig 8f). The average firing rate differences were slightly greater in CMM than in CLM, but variance of these differences was much greater in CMM than in CLM (chi-square variance test: rewarded vs. novel, $p = 2.1 \times 10^{-11}$; rewarded vs. unrewarded, $p = 2.3 \times 10^{-4}$). Learning therefore strongly modulates (by either increasing or decreasing) the average firing rate responses of CMM neurons to enhance the encoding of song category, but such modulation is much less pronounced in CLM neurons.

Discussion

The complex (Meliza et al., 2010), learning-dependent encoding of song by CMM neurons (Gentner and Margoliash, 2003) suggests that these representations are the product of an extensive neural processing network. Our results reveal some of the functional characteristics of that network by highlighting multiple differences between the encoding properties of individual CLM and CMM neurons, and by demonstrating that learning modifies these encoding properties. Together, these results suggest that CLM and CMM are part of a functional sensory circuit across which representations of natural vocal signals become increasingly informative with respect to behavior.

Coding along the avian auditory processing pathway

CLM and CMM sit near the top of a sensory processing pathway along which neural responses get progressively more complex. Within Field L, neurons in the L1 and L3 sub-regions selectively encode species-specific vocalizations more than neurons in the thalamorecipient Field L2 (Bonke et al., 1979; Langner et al., 1981). Linear spectrotemporal receptive field (STRF) models of neurons in Field L2a predict neural responses substantially better than the same models for neurons in CLM, indicating that response nonlinearities increase from L2a to CLM (Sen et al., 2001). Nonlinear stimulus transformations, such as the spectrotemporal "surprise," substantially improve the predictive power of STRF models for CLM neurons, but only moderately for Field L neurons, again highlighting the increase in nonlinear processing between CLM and Field L (Gill et al., 2008). In addition, some neurons in CLM show a moderate preference to respond to the bird's own song over other conspecific songs (Bauer et al., 2008), a hallmark of neural complexity (Margoliash, 1983) not observed in Field L (Amin et al., 2004; Shaevitz and Theunissen, 2007).

The differences in neural processing between CLM and CMM resemble those within known hierarchical circuits. First, neurons in CMM have higher motif selectivity than neurons in CLM. Selectivity often increases along ascending hierarchical circuits, including pathways in the visual (Maunsell and Newsome, 1987; Rust and DiCarlo, 2010) and auditory (Janata and Margoliash, 1999; Kikuchi et al., 2010) systems. Second, neurons in CMM encode more information about motif identity than neurons in CLM. In many sensory processing pathways, neurons at higher levels encode abstract concepts such as object identity whereas neurons at lower levels process the physical components of those objects (Nelken, 2004; Winer et al., 2005; Chechik et al., 2006; Nahum et al., 2008; Russ et al., 2008). Like visual objects, motifs are high-level concepts that are abstracted from the physical combinations of sounds from which they are composed (Gentner and Hulse, 2000; Gentner, 2008; Seeba and Klump, 2009). Third, neurons in CMM exhibit more variability in their responses to repeated motifs of the same type than CLM neurons. Because these repeated motifs have subtle acoustic differences and different positions within the song, we cannot attribute the increased sensitivity of CMM neurons exclusively to either feature. Nonetheless, sensitivity to subtle differences in complex stimuli, such as faces, is a hallmark of responses at high levels in known hierarchical circuits (Desimone et al., 1984), and temporal context-sensitivity increases between the auditory thalamus and auditory cortex in mammals (Asari and Zador, 2009). Collectively, all three of these coding differences—motif selectivity, information, and variability across motif renditions—suggest that neural representations in CMM are more complex than in CLM, and thus support the hypothesis that CLM and CMM are a part of functional hierarchical neural circuit.

Even with the evidence provided here, there are several reasons to be cautious of drawing too strict a conclusion about hierarchical processing across CLM and CMM. First, connectivity between the two regions is reciprocal (Vates et al., 1996), which could make the precise flow of information multifaceted and complex (but does not necessarily preclude

hierarchical processing; Van Essen et al., 1992). Second, CMM shares a strong reciprocal connection with the caudomedial nidopallium (NCM), another secondary auditory forebrain region that receives input from Field L (Vates et al., 1996). Responses of NCM neurons are also modified by song-recognition learning (Thompson and Gentner, 2010), and thus may also contribute directly to the emergence of complex, learning-dependent responses in CMM. Finally, the response properties of neurons in both CLM and CMM are heterogeneous and partially overlapping, suggesting that multiple pathways of information flow may be present. Regardless of the specific underlying architecture, however, our data show significant functional differences between CLM and CMM.

Learning modifies information encoding in CLM and CMM

Our results suggest that learning acts on CLM and CMM neurons in at least two ways: by increasing the information about motif identity and by increasing the information about behaviorally defined song categories. We found that both CLM and CMM neurons encoded more information about the identity of the learned motifs than about the identity of novel motifs. Behavioral experiments suggest that starlings recognize conspecifics by memorizing the motifs that compose their repertoires (Gentner and Hulse, 2000). The preferential encoding of the learned motifs by neurons in CLM and CMM may be a part of this stored memory. Alternatively, because of the rich acoustical structure of starling songs, identification could be achieved by learning a subset of the motifs that the bird finds particularly useful for recognition. The additional information encoded by CLM and CMM neurons about the learned motifs may therefore reflect changes in the representations of the most useful motifs. Consistent with this, the strongest effects of learning occur for the motifs paired with reward during training (Figs. 5–6; Gentner and Margoliash, 2003), pointing to a role for positive reinforcement in shaping the neural codes in both regions. Similar effects have been reported in primary cortical areas in mammals (Blake et al., 2006; Polley et al., 2006). Because learning increases the information encoded about motifs similarly in CLM and CMM, at least some of the learning-dependent representations in CMM (Gentner and Margoliash, 2003) may be inherited from responses in CLM.

Neurons in CLM and CMM also encoded information about the learned behavioral categories and this information was substantially larger in CMM than in CLM. Given that each bird's task was to distinguish between rewarded and unrewarded songs, we hypothesize that neural activity in both regions contributes to this cognitive process and supports categorical processing in post-synaptic targets (Prather et al., 2009). Phenomenologically, our results are similar to the processing of learned categories along the primate dorsal and ventral visual pathways (Freedman et al., 2001, 2003; Freedman and Assad, 2006). The neural encoding of behaviorally relevant categories (i.e. the grouping of signals that share behavioral meanings with similar neural representations) may be a general adaptive principle of cortical sensory processing to organize the complexity of sensory input (Merzenich and deCharms, 1996; Freedman and Miller, 2008; Hoffman and Logothetis, 2009; Seger and Miller, 2010). To date, however, the behavioral modulation of categorical processing has been studied primarily in the primate visual system and the underlying circuitry remains poorly understood. The emergence of categorical representations between CLM and CMM provides an excellent opportunity to study the encoding of natural acoustic categories at the cellular and circuit level.

Multiple pathways are likely to be involved in the transformation of information between CLM and CMM. For example, because CLM neurons encode relatively small amounts of information about learned categories, a single CMM neuron that processes convergent input from many CLM neurons could amplify this effect substantially. Recent results that the responses of some CMM neurons to whole motifs are well-modeled by a combination of the responses to motif components (Meliza et al., 2010) are consistent with a general pattern of

convergence into CMM. Furthermore, synaptic input from NCM neurons, which elicit weaker responses to learned songs than to novel songs (Thompson and Gentner, 2010), likely contributes to the encoding of learned categories in CMM. Because CMM contains large numbers of GABA$_A$-positive neurons (Pinaud et al., 2004), signals from NCM may specifically suppress the activity in CMM for novel songs. Additional studies will be necessary to compare the roles of CLM and NCM in shaping CMM responses.

This circuit will also be highly valuable for tracking changes in neural encoding over the course of learning, which is extremely difficult to do in primate models because of the large amounts of time required to train monkeys (Hoffman and Logothetis, 2009; c.f. Messinger et al., 2001). In contrast, starlings can learn to recognize songs very quickly (unpublished observations). With awake, behaving recording techniques, the modification of neural responses to encode newly learned categories could be readily observed. Thus, our identification of the emergence of behaviorally relevant information about songs along the CLM to CMM pathway highlights it as an especially valuable model for studying the circuit and plasticity mechanisms that underlie the selective neural processing of learned signals.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Amin N, Grace JA, Theunissen FE. Neural response to bird's own song and tutor song in the zebra finch field L and caudal mesopallium. J Comp Physiol A Neuroethol Sens Neural Behav Physiol. 2004; 190:469–489. [PubMed: 15064966]

Asari H, Zador AM. Long-lasting context dependence constrains neural encoding models in rodent auditory cortex. J Neurophysiol. 2009; 102:2638–2656. [PubMed: 19675288]

Bauer EE, Coleman MJ, Roberts TF, Roy A, Prather JF, Mooney R. A synaptic basis for auditory-vocal integration in the songbird. J Neurosci. 2008; 28:1509–1522. [PubMed: 18256272]

Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET. Human temporal lobe activation by speech and nonspeech sounds. Cereb Cortex. 2000; 10:512–528. [PubMed: 10847601]

Blake DT, Heiser MA, Caywood M, Merzenich MM. Experience-dependent adult cortical plasticity requires cognitive association between sensation and reward. Neuron. 2006; 52:371–381. [PubMed: 17046698]

Bonke D, Scheich H, Langner G. Responsiveness of units in the auditory neostriatum of the guinea fowl (Numida meleagris) to species-specific calls and synthetic stimuli. J Comp Physiol A. 1979; 132 1432-1351.

Brenner N, Strong S, Koberle R, Bialek W. Synergy in a neural code. Neural Comput. 2000; 12:1531–1552. [PubMed: 10935917]

Chaiken M, Bohner J, Marler P. Song Acquisition in European Starlings, Sturnus vulgaris - a comparison of the songs of live-tutored, tape-tutored, untutored, and wild-caught males. Anim Behav. 1993; 46:1079–1090.

Chechik G, Anderson MJ, Bar-Yosef O, Young ED, Tishby N, Nelken I. Reduction of information redundancy in the ascending auditory pathway. Neuron. 2006; 51:359–368. [PubMed: 16880130]

Cover, T.; Thomas, J. Elements of Information Theory. 2 Edition. Hoboken, New Jersey: John Wiley & Sons, Inc; 2006.

Desimone R, Albright TD, Gross CG, Bruce C. Stimulus-selective properties of inferior temporal neurons in the macaque. J Neurosci. 1984; 4:2051–2062. [PubMed: 6470767]

Eens M. Understanding the Complex Song of the European Starling: An Integrated Ethological Approach. Adv Study Behav. 1997; 26:355–434.

Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex. 1991; 1:1–47. [PubMed: 1822724]

Franco L, Rolls ET, Aggelopoulos NC, Jerez JM. Neuronal selectivity, population sparseness, and ergodicity in the inferior temporal visual cortex. Biol Cybern. 2007; 96:547–560. [PubMed: 17410377]

Freedman DJ, Assad JA. Experience-dependent representation of visual categories in parietal cortex. Nature. 2006; 443:85–88. [PubMed: 16936716]

Freedman DJ, Miller EK. Neural mechanisms of visual categorization: insights from neurophysiology. Neurosci Biobehav R. 2008; 32:311–329.

Freedman DJ, Riesenhuber M, Poggio T, Miller EK. Categorical representation of visual stimuli in the primate prefrontal cortex. Science. 2001; 291:312–316. [PubMed: 11209083]

Freedman DJ, Riesenhuber M, Poggio T, Miller EK. A comparison of primate prefrontal and inferior temporal cortices during visual categorization. J Neurosci. 2003; 23:5235–5246. [PubMed: 12832548]

Gentner TQ. Temporal scales of auditory objects underlying birdsong vocal recognition. J Acoust Soc Am. 2008; 124:1350–1359. [PubMed: 18681620]

Gentner TQ, Hulse SH. Perceptual classification based on the component structure of song in European starlings. J Acoust Soc Am. 2000; 107:3369–3381. [PubMed: 10875382]

Gentner TQ, Margoliash D. Neuronal populations and single cells representing learned auditory objects. Nature. 2003; 424:669–674. [PubMed: 12904792]

Gill P, Woolley SMN, Fremouw T, Theunissen FE. What's that sound? Auditory area CLM encodes stimulus surprise, not intensity or intensity changes. J Neurophysiol. 2008; 99:2809–2820. [PubMed: 18287545]

Hoffman KL, Logothetis NK. Cortical mechanisms of sensory learning and object recognition. Philos Trans R Soc Lond B Biol Sci. 2009; 364:321–329. [PubMed: 18977728]

Hsu A, Woolley SM, Fremouw TE, Theunissen FE. Modulation power and phase spectrum of natural sounds enhance neural encoding performed by single auditory neurons. J Neurosci. 2004; 24:9201–9211. [PubMed: 15483139]

Janata P, Margoliash D. Gradual emergence of song selectivity in sensorimotor structures of the male zebra finch song system. J Neurosci. 1999; 19:5108–5118. [PubMed: 10366643]

Kaas JH, Hackett TA. Subdivisions of auditory cortex and processing streams in primates. Proc Natl Acad Sci U S A. 2000; 97:11793–11799. [PubMed: 11050211]

Kikuchi Y, Horwitz B, Mishkin M. Hierarchical auditory processing directed rostrally along the monkey's supratemporal plane. J Neurosci. 2010; 30:13021–13030. [PubMed: 20881120]

Langner G, Bonke D, Scheich H. Neuronal discrimination of natural and synthetic vowels in field L of trained mynah birds. Exp Brain Res. 1981; 43:11–24. [PubMed: 6265257]

Lehky SR, Sejnowski TJ, Desimone R. Selectivity and sparseness in the responses of striate complex cells. Vision Res. 2005; 45:57–73. [PubMed: 15571738]

Macmillan, NA.; Creelman, CD. Detection Theory, A User's Guide. Mahwah, NJ: Lawrence Erlbaum Associates, Inc; 2005.

Margoliash D. Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. J Neurosci. 1983; 3:1039–1057. [PubMed: 6842281]

Maunsell JH, Newsome WT. Visual processing in monkey extrastriate cortex. Annu Rev Neurosci. 1987; 10:363–401. [PubMed: 3105414]

Meliza CD, Chi Z, Margoliash D. Representations of conspecific song by starling secondary forebrain auditory neurons: toward a hierarchical framework. J Neurophysiol. 2010; 103:1195–1208. [PubMed: 20032245]

Merzenich, MM.; deCharms, RC. Neural representations, experience, and change. In: Llinas, R.; Churchland, PS., editors. The Mind-Brain Continuum. Cambridge, MA: MIT Press; 1996. p. 61-81.

Messinger A, Squire LR, Zola SM, Albright TD. Neuronal representations of stimulus associations develop in the temporal lobe during learning. Proc Natl Acad Sci U S A. 2001; 98:12239–12244. [PubMed: 11572946]

Nahum M, Nelken I, Ahissar M. Low-level information and high-level perception: the case of speech in noise. Plos Biol. 2008; 6:e126. [PubMed: 18494561]

Nelken I. Processing of complex stimuli and natural scenes in the auditory cortex. Curr Opin Neurobiol. 2004; 14:474–480. [PubMed: 15321068]

Panzeri S, Senatore R, Montemurro MA, Petersen RS. Correcting for the sampling bias problem in spike train information measures. J Neurophysiol. 2007; 98:1064–1072. [PubMed: 17615128]

Pinaud R, Velho TAF, Jeong JK, Tremere LA, Leão RM, von Gersdorff H, Mello CV. GABAergic neurons participate in the brain's response to birdsong auditory stimulation. Eur J Neurosci. 2004; 20:1318–1330. [PubMed: 15341603]

Polley DB, Steinberg EE, Merzenich MM. Perceptual learning directs auditory cortical map reorganization through top-down influences. J Neurosci. 2006; 26:4970–4982. [PubMed: 16672673]

Prather JF, Nowicki S, Anderson RC, Peters S, Mooney R. Neural correlates of categorical perception in learned vocal communication. Nat Neurosci. 2009; 12:221–228. [PubMed: 19136972]

Rauschecker JP, Scott SK. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat Neurosci. 2009; 12:718–724. [PubMed: 19471271]

Riesenhuber M, Poggio T. Models of object recognition. Nat Neurosci. 2000; 3 Suppl:1199–1204. [PubMed: 11127838]

Rolls ET, Tovee MJ. Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. J Neurophysiol. 1995; 73:713–726. [PubMed: 7760130]

Rolls ET, Baylis GC, Hasselmo ME, Nalwa V. The effect of learning on the face selective responses of neurons in the cortex in the superior temporal sulcus of the monkey. Exp Brain Res. 1989; 76:153–164. [PubMed: 2753096]

Russ BE, Ackelson AL, Baker AE, Cohen YE. Coding of auditory-stimulus identity in the auditory non-spatial processing stream. J Neurophysiol. 2008; 99:87–95. [PubMed: 18003874]

Rust NC, DiCarlo JJ. Selectivity and tolerance ("invariance") both increase as visual information propagates from cortical area V4 to IT. J Neurosci. 2010; 30:12978–12995. [PubMed: 20881116]

Seeba F, Klump GM. Stimulus familiarity affects perceptual restoration in the European starling (Sturnus vulgaris). PLoS One. 2009; 4:e5974. [PubMed: 19551146]

Seger CA, Miller EK. Category learning in the brain. Annu Rev Neurosci. 2010; 33:203–219. [PubMed: 20572771]

Sen K, Theunissen FE, Doupe AJ. Feature analysis of natural sounds in the songbird auditory forebrain. J Neurophysiol. 2001; 86:1445–1458. [PubMed: 11535690]

Shaevitz SS, Theunissen FE. Functional connectivity between auditory areas field L and CLM and song system nucleus HVC in anesthetized zebra finches. J Neurophysiol. 2007; 98:2747–2764. [PubMed: 17898149]

Sigala N, Logothetis NK. Visual categorization shapes feature selectivity in the primate temporal cortex. Nature. 2002; 415:318–320. [PubMed: 11797008]

Strong SP, Koberle R, van Steveninck RRD, Bialek W. Entropy and information in neural spike trains. Phys Rev Lett. 1998; 80:197–200.

Thompson JV, Gentner TQ. Song recognition learning and stimulus-specific weakening of neural responses in the avian auditory forebrain. J Neurophysiol. 2010; 103:1785–1797. [PubMed: 20107117]

Treves A, Panzeri S. The Upward Bias in Measures of Information Derived from Limited Data Samples. Neural Comput. 1995; 7:399–407.

Van Essen DC, Anderson CH, Felleman DJ. Information processing in the primate visual system: an integrated systems perspective. Science. 1992; 255:419–423. [PubMed: 1734518]

Vates GE, Broome BM, Mello CV, Nottebohm F. Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches. J Comp Neurol. 1996; 366:613–642. [PubMed: 8833113]

Vinje WE, Gallant JL. Sparse coding and decorrelation in primary visual cortex during natural vision. Science. 2000; 287:1273–1276. [PubMed: 10678835]

Willmore B, Tolhurst DJ. Characterizing the sparseness of neural codes. Network. 2001; 12:255–270. [PubMed: 11563529]

Winer JA, Miller LM, Lee CC, Schreiner CE. Auditory thalamocortical transformation: structure and function. Trends Neurosci. 2005; 28:255–263. [PubMed: 15866200]

Woolley SM, Fremouw TE, Hsu A, Theunissen FE. Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. Nat Neurosci. 2005; 8:1371–1379. [PubMed: 16136039]

Woolley SM, Gill PR, Fremouw T, Theunissen FE. Functional groups in the avian auditory system. J Neurosci. 2009; 29:2780–2793. [PubMed: 19261874]
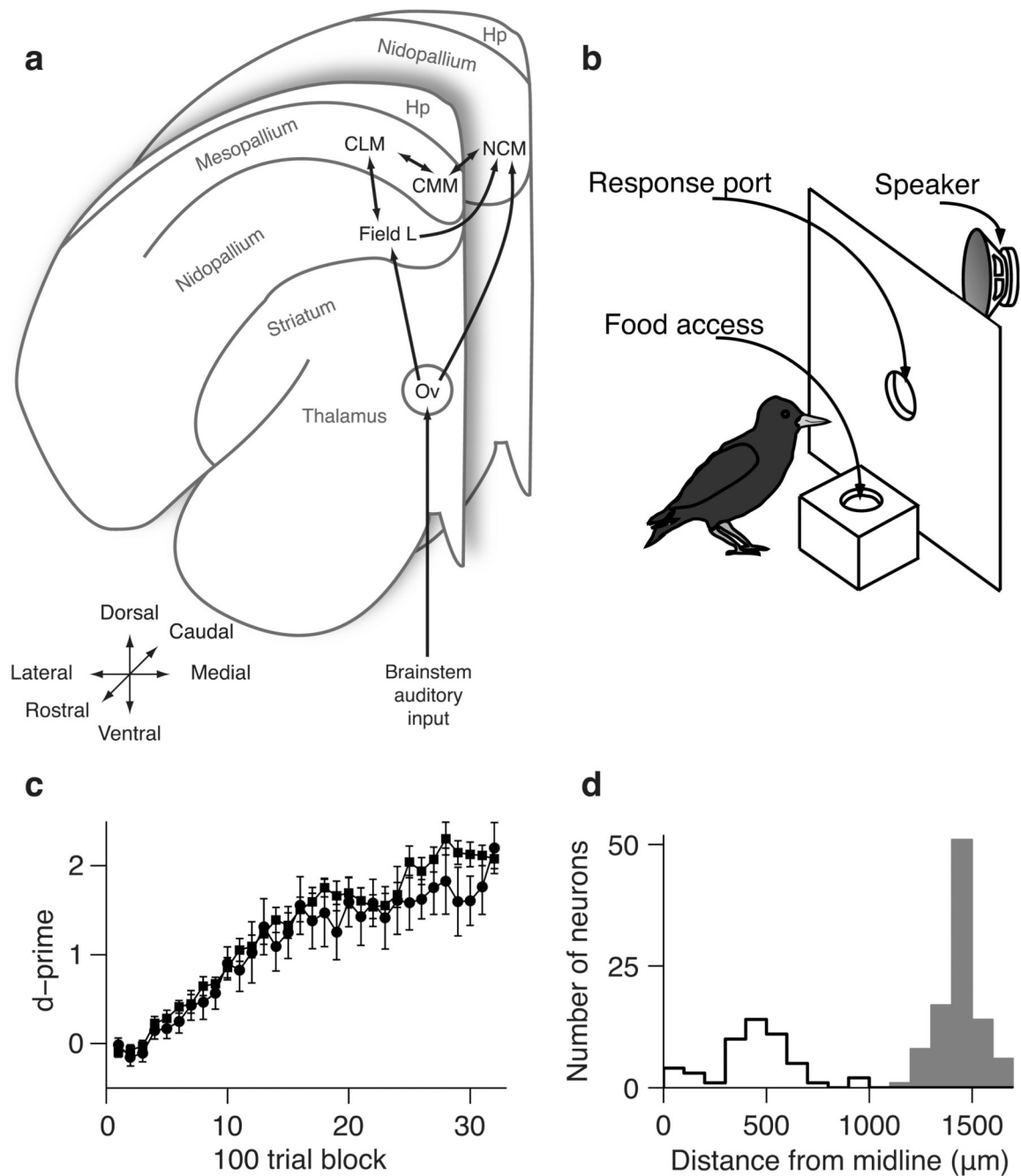
Fig. 1. Behavioral training and starling neuroanatomy
(**a**) Schematic diagram of the connectivity of the songbird auditory system across two coronal planes of one hemisphere. (**b**) Schematic of operant apparatus used for behavioral training. (**c**) Mean (± SEM) behavioral performance (d-prime) during learning for subjects used in CLM experiments (squares) and subjects used in CMM experiments (circles). Error bars are across subjects. (**d**) Distributions of recording locations for CLM neurons (black) and CMM neurons (white). Hp, Hyperpallium; CLM, Caudolateral Mesopallium; CMM, Caudomedial Mesopallium; NCM, Caudomedial Nidopallium; Ov, Ovoidalis.
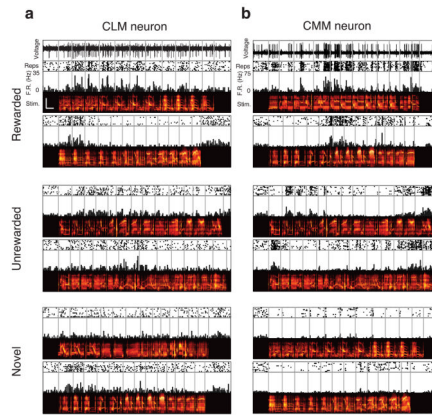
Fig. 2. Responses from two example neurons
Extracellular electrical activity recorded from a sample CLM neuron (**a**) and CMM neuron
(**b**) in response to rewarded motifs (top), unrewarded motifs (middle), and novel motifs
(bottom). Scale bar in upper left spectrogram denotes 0.5s and 5kHz. Scales are identical for
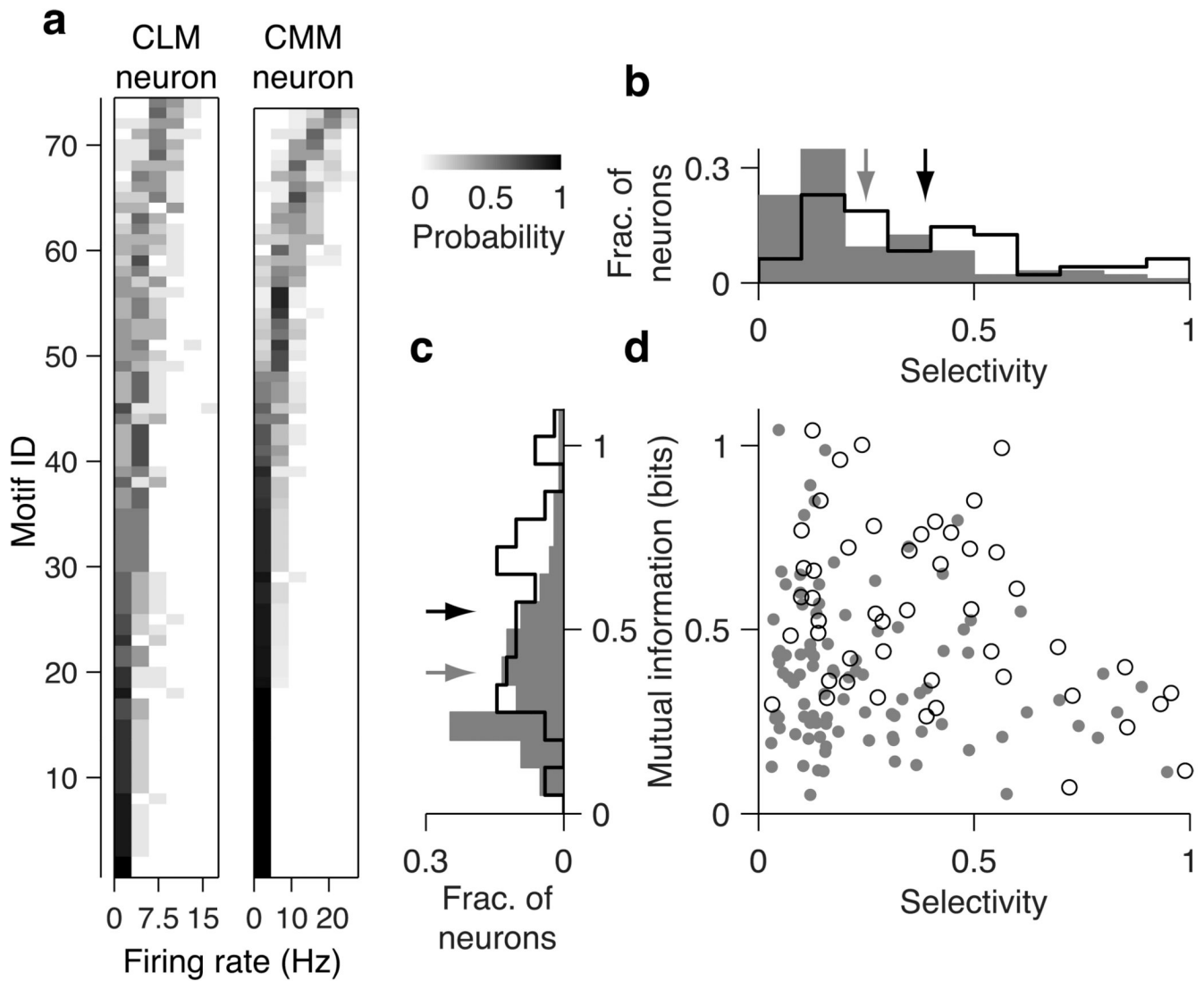all spectrograms.

Fig. 3. Motif selectivity and information encoding in CLM and CMM neurons
(**a**) Distributions of firing rates conditional on motif identity for the CLM neuron (left panel)
and CMM neuron (right panel) shown in Fig 2a and 2b, respectively. Each row in each panel
shows the firing rate distribution for a single motif. Motifs are arranged in order of
ascending mean firing rates and the conditional probability is encoded in grayscale. (**b**)
Distribution of motif selectivity values for CLM (grey bars) and CMM (black line) neurons.
Grey arrow denotes mean for CLM neurons, black arrow denotes mean for CMM neurons.
(**c**) Distribution of motif information values for CLM (grey bars) and CMM (black line)
neurons. Arrows are as in (b). (**d**) Relation between mutual information and motif selectivity
in CLM (dots) and CMM (circles) neurons. CMM neurons tend to have either low motif
selectivity or low mutual information, but not both. CLM neurons can have both low motif
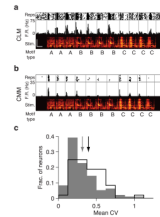selectivity and low mutual information.

Fig. 4. Response variability of CLM and CMM responses to repeated motifs
(**a**) Response of an example CLM neuron to a single song illustrating response variability to repeated motifs. Motif types are denoted as letters below the spectrograms. Motifs of the same type were judged to be acoustically similar for the purposes of the response variability analysis (Materials and Methods). (**b**) Response of an example CMM neuron to the same song, as in (a). (**c**) Distribution of mean coefficient of variation (CV) across repeated motifs for all CLM and CMM neurons. The CV is computed from the firing rates across each sequence of repeated motifs and averaged for each neuron. Higher CV values indicate greater response variability. Scale bar in upper spectrogram denotes 0.5s and 5kHz. Scales are identical for both spectrograms.
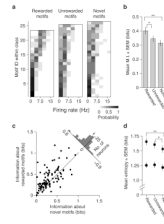
Fig. 5. Effects of learning on encoding of motif identity in CLM
(**a**) Distributions of firing rates (F.R.) conditional on motif identity for rewarded (left), unrewarded (middle) and novel (right) motifs for the neuron shown in Fig. 2a. Each row in each panel shows the firing rate distribution for a single motif. Motifs are arranged in order of ascending mean firing rates and the conditional probability is encoded in grayscale. The firing rates of this neuron allow greater disambiguation of motif identity for rewarded motifs than for novel motifs. (**b**) Comparison of mean (± SEM) mutual information across all CLM neurons for rewarded, unrewarded, and novel motifs. Wilcoxon signed-rank test: * $p < 0.01$; ** $p < 10^{-5}$. (**c**) Scatter plot illustrating distributions of information values for novel motifs and rewarded motifs. Each point represents a single neuron. Upper right, histogram of differences between mutual information values for rewarded vs. novel motifs for all neurons. The arrow denotes the mean. (**d**) Mean (± SEM) total entropy (squares) and noise entropy (circles) values over all CLM neurons. Paired t-test: * $p < 0.01$; ** $p < 0.005$.
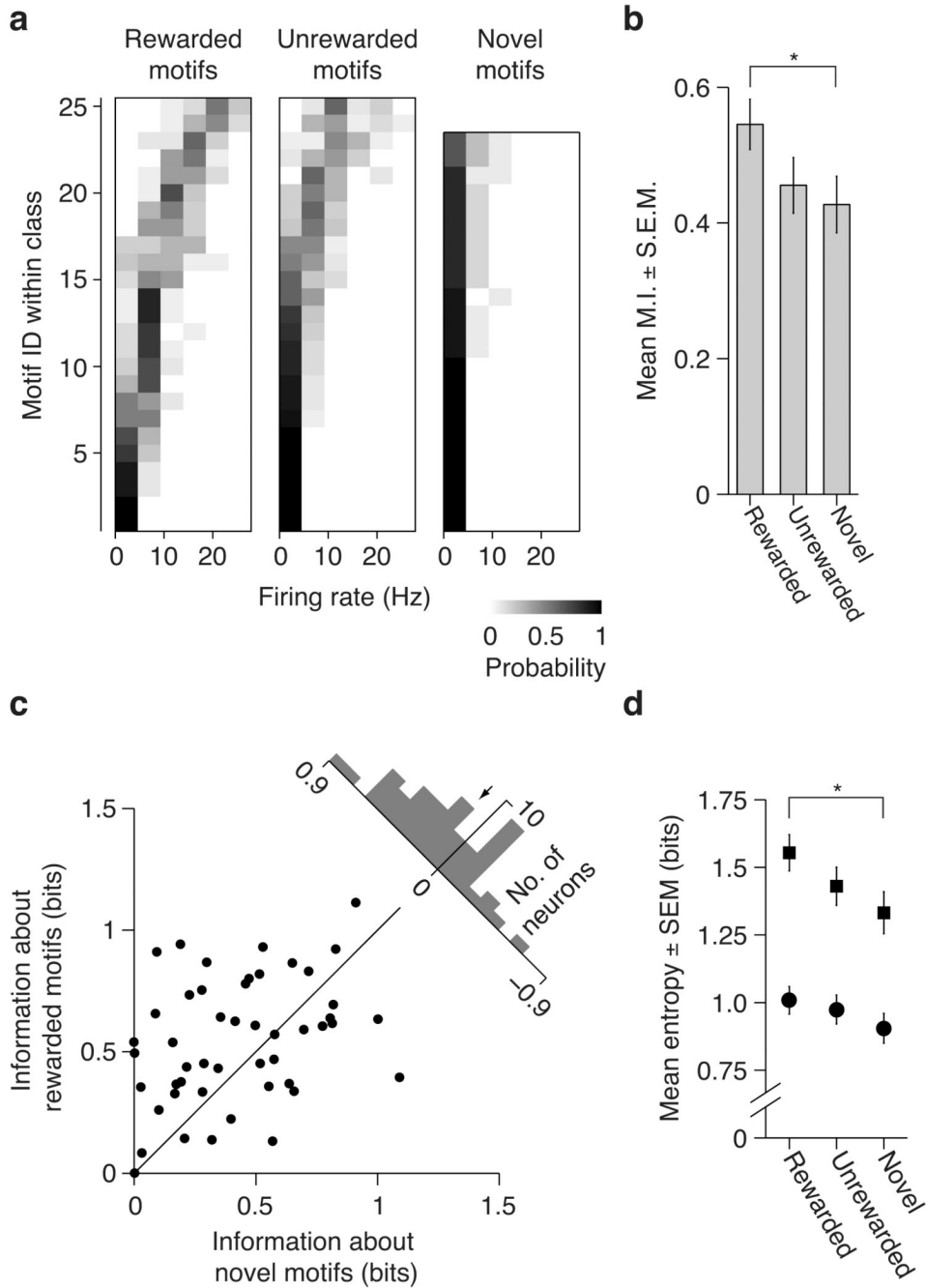
Fig. 6. Effects of learning on encoding of motif identity in CMM
(**a**) Distributions of firing rates (F.R.) conditional on motif identity for rewarded (left), unrewarded (middle) and novel (right) motifs for the neuron shown in Fig. 2b. Conventions are the same as in Fig 5a. (**b**) Comparison of mean (± SEM) mutual information across all CMM neurons for rewarded, unrewarded, and novel motifs. Wilcoxon signed-rank test: * p < 0.05. (**c**) Scatter plot illustrating distributions of information values for novel motifs and rewarded motifs. Each point represents a single neuron. Upper right, histogram of differences between mutual information values for rewarded vs. novel motifs for all neurons. The arrow denotes the mean. (**d**) Mean (± SEM) total entropy (squares) and noise entropy (circles) values for all CMM neurons. Paired t-test: * p < 0.05.
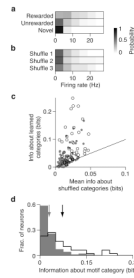
Fig 7. Effects of learning on encoding of motif category in CLM and CMM
(**a**) Probability distribution for the sample CMM neuron depicted in Fig 2b of firing rates in response to motifs, conditional on behavioral category. (**b**) Probability distribution of the same CMM neuron conditional on a randomly shuffled set of categories. (**c**) Comparison of information encoded about the learned motif categories (rewarded, unrewarded, novel), and the mean information encoded about 100 permutations of randomly shuffled categories for all CLM (gray dots) and CMM (open circles) neurons. Black line is the unity line. (**d**) Distributions of amounts of information about motif category encoded by CLM neurons (gray bars) and CMM neurons (black outline). Arrows denote means for CLM (gray) and CMM (black).
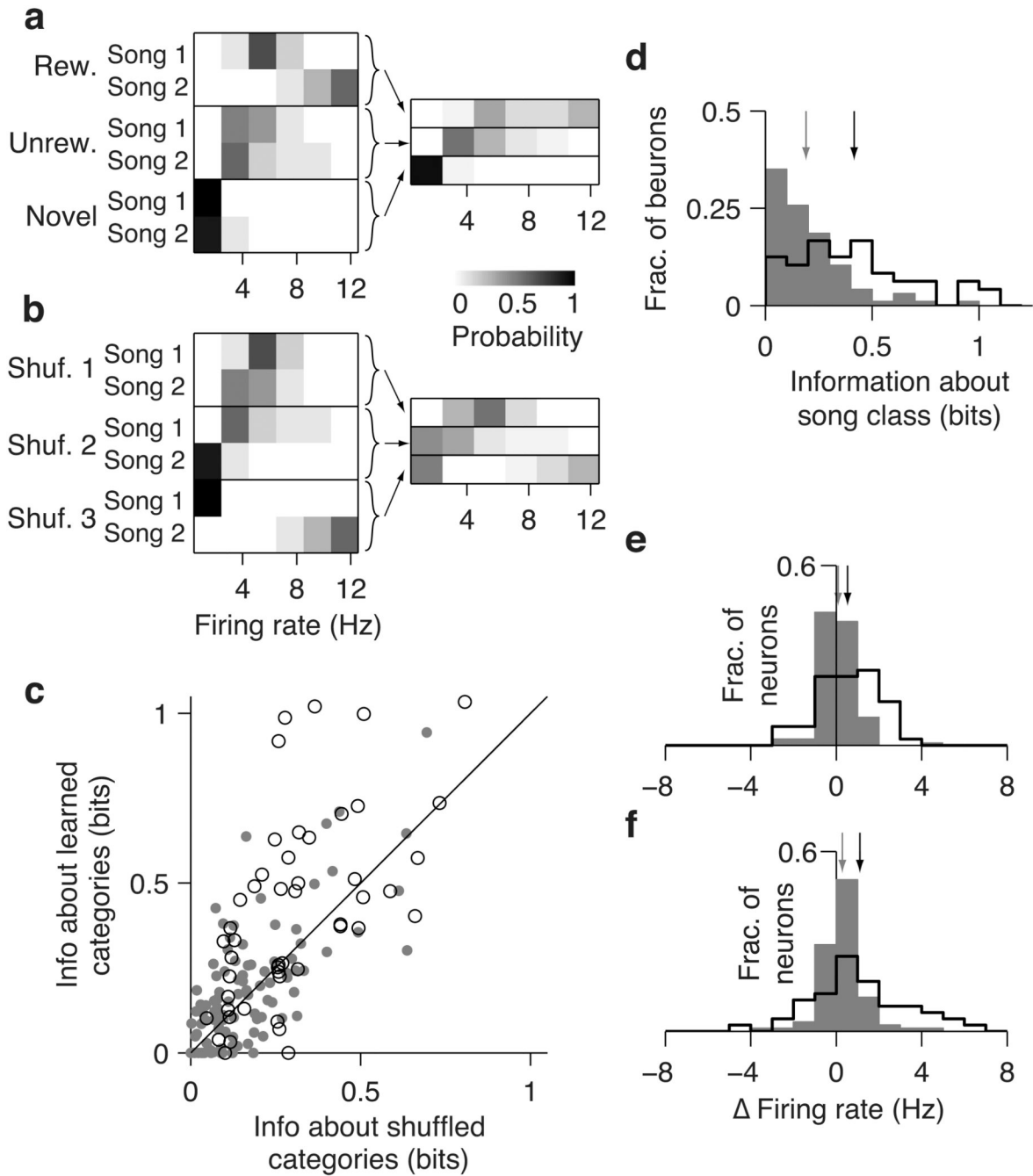
Fig 8. Effects of learning on encoding of song category in CLM and CMM
(**a**) Probability distributions of firing rates in response to songs for the sample CMM neuron depicted in Fig 2b, conditional on song identity (left) and behavioral category (right). Arrows depict the construction of category-conditional distributions. (**b**) Probability distributions as in (a) but for the randomly shuffled category with the highest information. (**c**) Comparison of information encoded about the learned categories (rewarded, unrewarded, novel), and the mean information about randomly shuffled categories for all CLM (gray dots) and CMM (open circles) neurons. Black line is the unity line. (**d**) Distributions of category information values for CLM neurons (gray bars) and CMM neurons (black outline). Arrows denote means for CLM (gray) and CMM (black). (**e**) Distribution of the

change in average firing rate between novel songs and rewarded songs for CLM neurons (gray bars) and CMM neurons (black outline). Positive values indicate higher firing rates for rewarded songs. Arrows denote means for CLM (gray) and CMM (black). (**f**) Distribution of the change in average firing rate between unrewarded songs and rewarded songs for CLM neurons (gray bars) and CMM neurons (black outline). Positive values indicate higher firing rates for rewarded songs. Arrows are as in (e).