| Dependent Measure	Previous CMM mean±SEM (4 birds, 28 neurons)	New CMM mean ±SEM (6 birds, 20 neurons)	t-test p-value
Trials to p<0.01 criterion	1000±168	833±95	0.37
Total trials performed	48025±24415	16294±6188	0.17
Behavioral performance at end of training (d-prime)	3.88±0.49	2.25±0.29	0.02
Lateral coordinate of neurons (distance from midline, μm)	485±17	436±72	0.43

Supplemental Table 1. Comparison of behavioral data and neuron locations between the previous set of CMM data and the new CMM data. The only significant difference is in the behavioral performance at the end of training, but all birds performed well above chance.

a

b

c
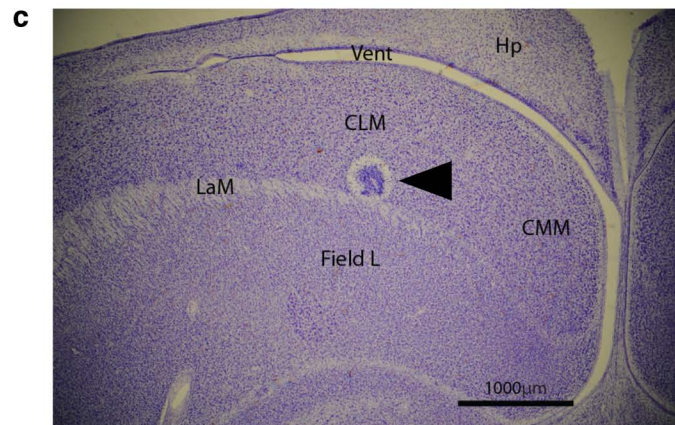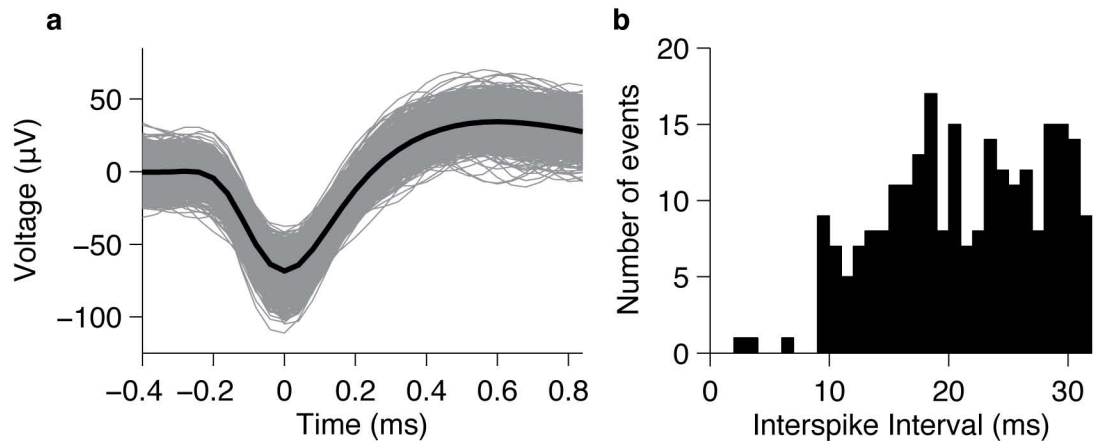
Supplemental Fig. 1. Spike sorting and histology. (**a**) Superimposed plots of extracellular action potential shape for the neuron shown in figure 2a. Light grey lines are individual action potential waveforms, black line is mean action potential shape. (**b**) Interspike interval histogram for the same neuron showing very few refractory period violations. (**c**) Photomicrograph of Nissl-stained coronal tissue section showing fiduciary electrolytic lesion in CLM (denoted by arrowhead) at the bottom of a recording penetration. Scale bar = 1000µm. LaM: Mesopallial Lamina.
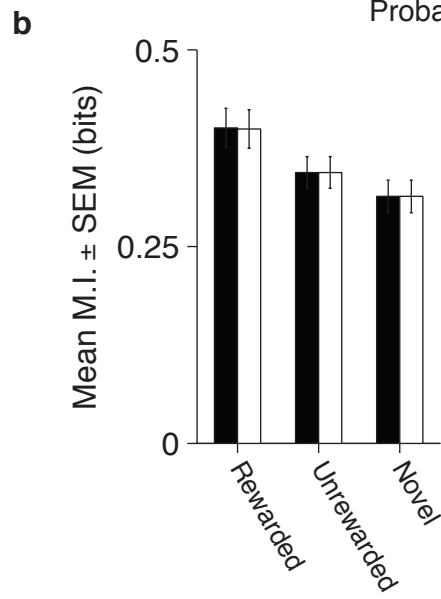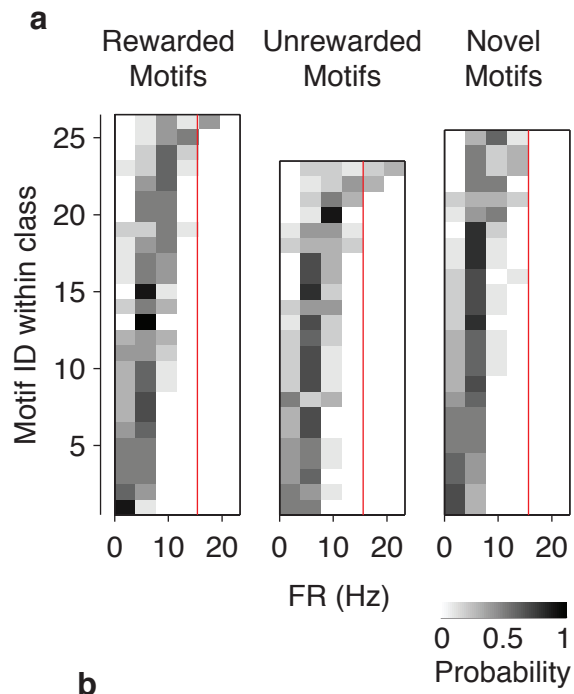
Supplemental Fig. 2
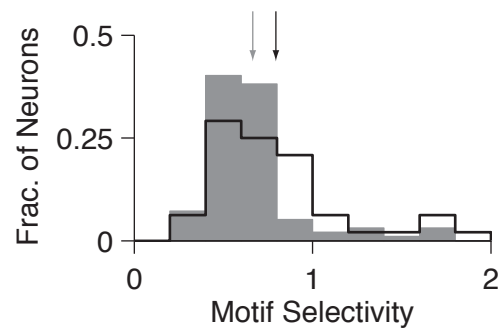
Supplemental Fig. 2. Comparison of selectivity values for CLM and CMM neurons using the entropy method (Materials and Methods). Gray bars denote CLM distribution and black outlines denote CMM distribution. On average, selectivity in CLM was 0.66±0.03 and selectivity in CMM was 0.79±0.05 (Wilcoxon rank sum test, p = 0.038).
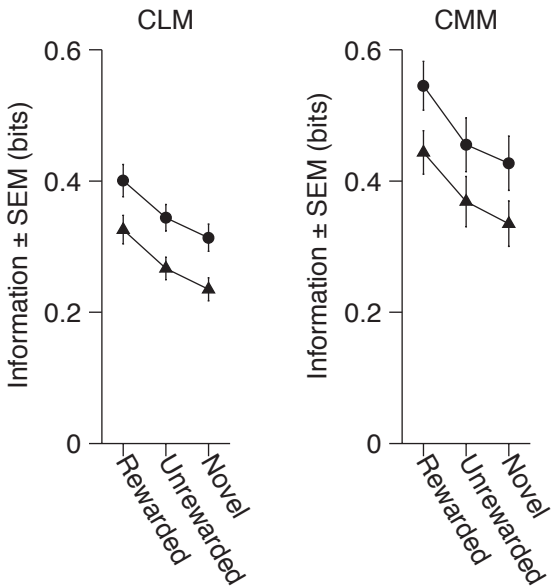
Supplemental Fig. 3
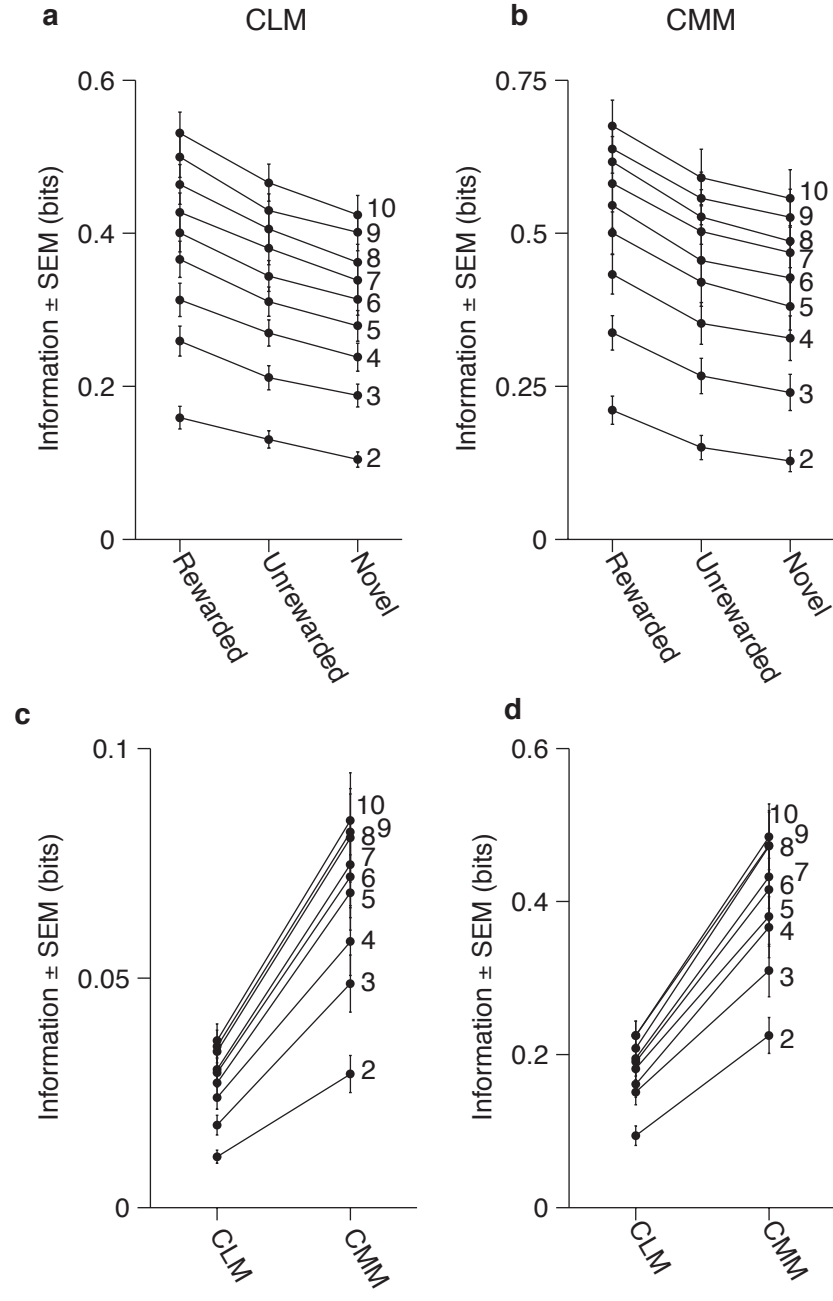
Supplemental Fig. 3. Effects of varying the number of bins in information calculations. (**a**) Information about motif identity in CLM. (**b**) Information about motif identity in CMM. (**c**) Information in motifs about category. (**d**) Information in songs about category. In all plots, the numbers to the right of each set of connected points denotes the number of bins used for that calculation. Increasing the number of bins increases the number of bits, but the overall effects of learning are not affected.

a

b

Supplemental Fig. 4. Effects of limiting firing rate range on single neuron information encoding in CLM. (**a**) Conditional probability distributions of a sample CLM neuron. Red bars indicate the maximum firing rate elicited by a novel motif. Information was computed using six bins linearly spaced from the minimum novel firing rate to the maximum novel firing rate. Responses outside this range were ignored in this analysis. (**b**) Comparison of mean (± SEM) information encoded under this restricted firing rate range (grey bars; Friedman test, $p = 1.7 \times 10^{-4}$) with the mean information encoded under the full firing rate range (black bars; Friedman test, $p = 1.7 \times 10^{-4}$). No difference is observed between the effects of learning under the two conditions (two-way repeated measures ANOVA interaction term: $p = 0.19$).

Supplemental Fig. 5

CLM

Information ± SEM (bits)

0.6

0.4

0.2

0

Rewarded
Unrewarded
Novel

CMM

Information ± SEM (bits)

0.6

0.4

0.2

0

Rewarded
Unrewarded
Novel

Supplemental Fig. 5. Effects of treating multiple renditions of the same motif type as identical motifs for the purposes of information encoding in CLM (left) and CMM (right). Circles denote information when all renditions are considered independent (as in the main text). Triangles denote information when all renditions of a given motif type are considered to be identical. No differences are found between the two calculations (2-way repeated measures ANOVA interaction; CLM: $p = 0.87$; CMM: $p = 0.42$).