

RESEARCH ARTICLE

Sensory Processing

Effect of background clutter on neural discrimination in the bat auditory midbrain

 **Kathryne M. Allen,^{1*}**  **Angeles Salles,^{1*}** **Sangwook Park,²** **Mounya Elhilali,²** and  **Cynthia F. Moss¹**

¹Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, Maryland and ²Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, Maryland



Abstract

The discrimination of complex sounds is a fundamental function of the auditory system. This operation must be robust in the presence of noise and acoustic clutter. Echolocating bats are auditory specialists that discriminate sonar objects in acoustically complex environments. Bats produce brief signals, interrupted by periods of silence, rendering echo snapshots of sonar objects. Sonar object discrimination requires that bats process spatially and temporally overlapping echoes to make split-second decisions. The mechanisms that enable this discrimination are not well understood, particularly in complex environments. We explored the neural underpinnings of sonar object discrimination in the presence of acoustic scattering caused by physical clutter. We performed electrophysiological recordings in the inferior colliculus of awake big brown bats, to broadcasts of prerecorded echoes from physical objects. We acquired single unit responses to echoes and discovered a subpopulation of IC neurons that encode acoustic features that can be used to discriminate between sonar objects. We further investigated the effects of environmental clutter on this population's encoding of acoustic features. We discovered that the effect of background clutter on sonar object discrimination is highly variable and depends on object properties and target-clutter spatiotemporal separation. In many conditions, clutter impaired discrimination of sonar objects. However, in some instances clutter enhanced acoustic features of echo returns, enabling higher levels of discrimination. This finding suggests that environmental clutter may augment acoustic cues used for sonar target discrimination and provides further evidence in a growing body of literature that noise is not universally detrimental to sensory encoding.

NEW & NOTEWORTHY Bats are powerful animal models for investigating the encoding of auditory objects under acoustically challenging conditions. Although past work has considered the effect of acoustic clutter on sonar target detection, less is known about target discrimination in clutter. Our work shows that the neural encoding of auditory objects was affected by clutter in a distance-dependent manner. These findings advance the knowledge on auditory object detection and discrimination and noise-dependent stimulus enhancement.

AQ: 5 auditory object; ~~clutter~~; echolocation; inferior colliculus; noise; sonar object

AQ: 6 INTRODUCTION

Parsing and discriminating natural sounds in complex scenes is vital for humans and other animals that rely on hearing for survival behaviors, such as communication, predator evasion, and navigation. This is particularly important for those that rely on audition for navigation, such as visually impaired **humans** or echolocating animals. Bats present a powerful model system to study neural encoding

of complex acoustic information in noisy environments. Echolocating bats are able to successfully catch prey, even when target echoes are partially masked by echoes from surrounding clutter, such as vegetation (1–4). Indeed, they can detect targets in clutter, even when the echoes are only 1 dB stronger than that of the clutter (5). They also adapt their flight trajectories to separate a target from clutter along the range axis (1, 6). Bats respond to echoes in the environment with adjustments in the interval, duration, and spectral

* K. M. Allen and A. Salles contributed equally to this work.

Correspondence: K. M. Allen (kate.allen.89@gmail.com); A. Salles (angiesalles@gmail.com).

Submitted 10 March 2021 / Revised 22 September 2021 / Accepted 12 October 2021



content of their calls to mitigate acoustic interference from clutter (1, 7, 8). Even with flight trajectory and sonar call adaptations that aid in the localization of objects that are close to clutter, the presence of masking echoes may either impair or augment target detection, depending on the temporal separation of target and clutter echoes (9). Though bats have been shown to accurately discriminate target shape and texture in open environments (10), previous studies have not addressed how bats contend with target discrimination in clutter conditions and the neural mechanisms that enable this fundamental task. Our work aims to bridge this gap by studying the neural underpinnings of sonar object discrimination in clutter.

Neural responses to sound source location and features have been characterized in the bat inferior colliculus (IC), a midbrain structure in the central auditory pathway that receives ascending input from brainstem nuclei and relays information to cortical areas, while also modulated by descending input from auditory cortex, among other areas. The midbrain IC shows many neuronal specializations in bats that point to its fundamental role in auditory processing. First, the IC comprises frequency selective neurons, which are topographically organized, with neurons tuned to lower frequencies located on the more dorsal regions and those tuned to high frequencies in more ventral regions (11–14). Second, a population of IC neurons are tuned to the duration of sounds (15, 16). In addition, some IC neurons exhibit pulse-echo delay facilitation and tuning (17–19), posited as the mechanism of sonar target ranging. Last, some neurons in the bat IC respond selectively to sweep rate and direction (14, 20). Collectively, IC tuning properties in bats operate to support sonar object discrimination by encoding information carried by returning echoes. Although much has been reported on the auditory features encoded by the IC, it is still unknown how the encoding of these acoustic features is altered by the presence of physical clutter. Likewise, behavioral detection of sonar targets in clutter has been studied in bats, but their ability to discriminate between object features in clutter remains unexplored. This study seeks to understand how the presence of temporally overlapping clutter echoes affects the encoding of sonar objects.

To advance the understanding of sonar object discrimination in the presence of background clutter, we recorded single unit response to echoes from physical objects in the IC of *Eptesicus fuscus* to explore the neural encoding and discrimination of sonar objects. We investigated the extent to which neural activity in the IC can be used to extract acoustic features that allow for sonar object discrimination, and further tested the influence of physical clutter on sonar object discrimination. We discovered that clutter can sometimes produce acoustic scattering that facilitates neural discrimination of object features.

AQ: 7

METHODS

Acoustic Stimuli

Echoes from three-dimensional (3-D) solid plastic objects (Table 1) were collected in an acoustically isolated room lined with acoustic foam. Echoes were created by broadcasting a

Table 1. Target object sizes

Object	Size, cm
Cube	3 × 3 × 3
Cylinder	2.5 base × 3 height
Large dipole	1 cm spheres connected with 3 cm hypodermic tube
Small dipole	1 cm spheres connected with 1 cm hypodermic tube
Small sphere	1 cm diameter

synthetic echolocation FM sweeps (100–20 kHz, 2-ms duration) from an ultrasonic custom-made electrostatic speaker placed 50 cm away from the object. The echoes were recorded by a Petersen ultrasonic microphone located directly above the speaker. The speaker was calibrated, and a custom filter was created in MATLAB and applied to speaker output to obtain a flat frequency response. The synthetic FM sweep mimicked an *E. fuscus* echolocation call. Each object was suspended in place by a 0.5-mm diameter hypodermic tube at a fixed angle. For all objects, this was ~5° off axis from a centered flat face (i.e., flat face of cylinder). Clutter was added to some recordings by hanging a small leafed artificial plant located either 10 cm or 20 cm behind the target (Fig. 1). The sonar broadcast (pulse) and echo recordings were then used as playback for passively listening animals.

F1

Stimuli were presented as pulse-echo pairs, with echoes delayed artificially by inserting a 10-ms interval from the initial pulse, corresponding to a target distance of 1.7 m. Several echo delays (5–15 ms) were tested beforehand, however, the 10-ms delay captured the largest number of responsive neurons and was thus selected as our stimulus. The 10-ms delay between call and echo was filled with the background room noise picked up in the sound recordings. Five objects were presented in three clutter conditions (none, +10 cm, +20 cm) for 18 total stimulus conditions. Responses to these stimuli were each recorded over 20 presentations in a randomized order. Neural responses to 5 ms pure tones were collected to build frequency tuning curves for each unit. The 5 ms pure tones were presented with 0.5 ms ramps at beginning and end, in the range of 20–90 kHz in frequency in steps of 5 kHz; and 20–70 dB SPL in steps of 10 dB. Each stimulus was generated at a sampling rate of 250 kHz using a National Instruments card (PXIe 6358) and transmitted with a calibrated custom-made electrostatic loudspeaker connected to an audio amplifier (Krohn-Hite 7500). The loudspeaker was placed 30 cm from the bat's ear, contralateral to the recording site. The frequency response of the loudspeaker was flattened by digitally filtering the playback stimuli as described by Luo et al. (21).

AQ: 8

Animals

Animals (big brown bats, *E. fuscus*) were recovered from an exclusion site in Maryland under Permit No. 554400 issued by the Maryland Department of Natural Resources. All protocols are approved by Johns Hopkins University under IACUC Protocol No. BA20A65. Bats were housed socially in flight-rooms and provided with ad libitum food and water. Four bats were used for this experiment (3 females, 1 male) and each was housed individually after head-post surgery and during the duration of the recordings.

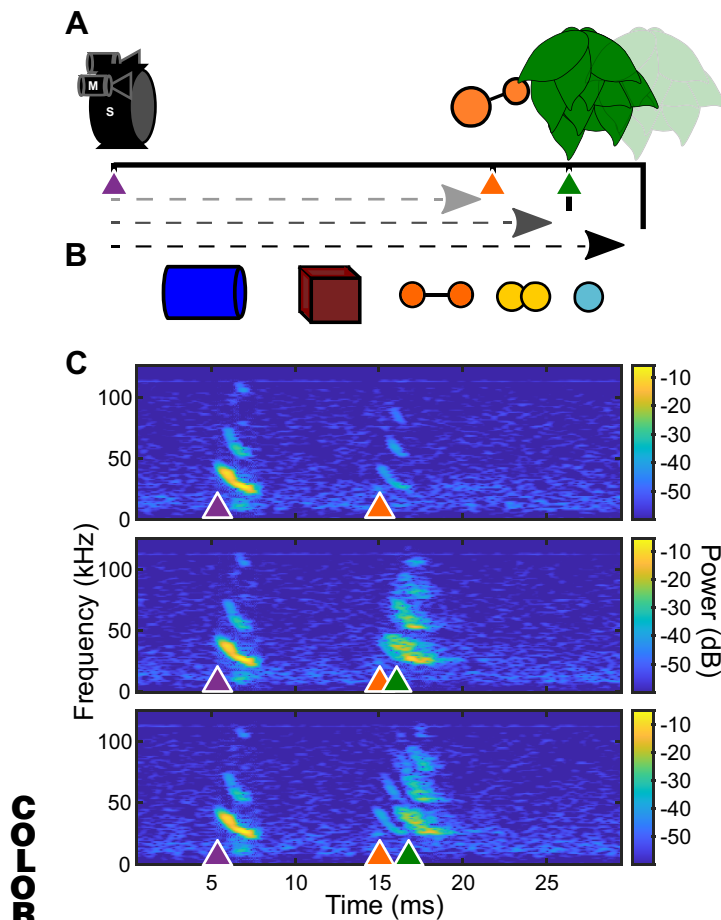


Figure 1. Schematic and example of acoustic stimuli. **A:** an ultrasound loudspeaker and two microphones were suspended 50 cm (lightest gray) from an object with artificial foliage suspended either +10 cm (middle gray) or +20 cm (black) behind the object. **B:** objects used in this study. Cylinder (blue), cube (maroon), small dipole (yellow), large dipole (orange), and small sphere (cyan). Details about object size and composition are found in Table 1. **C:** example echo spectrograms from the large dipole object. Synthetic calls with properties matched to average *Eptesicus fuscus* calls (purple arrow, 2 ms 100–20 kHz down sweep) were broadcast at objects in front of no clutter (top), +10 cm clutter behind object (middle), or +20 cm clutter behind object (bottom). Sonar sound broadcast and object echoes (orange) were recorded by two ultrasound microphones. Objects in front of clutter resulted in echoes with overlap between object echoes and clutter echoes (green).

Neurophysiology

Bats were prepared for neural recordings by securing a head post to the skull, as described by Salles et al. (14). Briefly, bats were anesthetized with 1%–3% isoflurane gas. A longitudinal midline incision was made through the skin overlying the skull and the musculature was retracted. A custom-made metal post was attached to the skull surface using dental cement (Metabond). After surgery, bats received an oral analgesic (Metacam 12.5 mg/kg), and subsequently the same dose daily after neural recordings, until the experiment was concluded. Bactrim was administered once daily for 7 days to prevent infection together with the analgesic. Animals were allowed to rest for 48 h before neurophysiological recordings. On the first day of recordings, a <1 mm craniotomy was drilled on the skull over the IC before electrode

insertion, preserving the dura. For neural recordings, each bat was placed in a foam body mold, with its head tightly fixed by the head post attached to a metal holder. All recordings were conducted in a sound attenuating and electrically shielded chamber (Industrial Acoustics Company, Inc.). Recordings from each bat were taken over multiple sessions over a period of 2–3 wk, with each session lasting a maximum of 4 h. Bats were offered water in the middle of each session. No sedative or other drugs were administered during recording sessions. A Neuronexus (impedance 1.50–1.78 MΩ) silicon probe (A1x16-5mm-50-177-A16) was orthogonally inserted into the brain using a hydraulic microdrive mounted on a micromanipulator (Stoelting Co.). The craniotomy was covered with saline solution during the experiments and care was taken to prevent desiccation. The brain surface was used as reference for depth measurement (0 μm), the electrode was advanced in steps of 10 μm and recordings were taken at least 100 μm apart. A silver wire, placed 1–2 cm rostral to the recording electrode and underneath the muscle, was used as a grounding electrode. The electrical signals from the recording probe channels were acquired with an OmniPlex D Neural Data Acquisition System (Plexon, Inc.) at a sampling rate of 40 kHz (per channel) and 16-bit precision. Synchronization between the neural recordings and acoustic stimulus broadcasts was achieved with a transistor transistor logic (TTL) pulse output from a National Instrument card and was recorded on channel 17 of the Plexon data acquisition system. The output of the broadcast was recorded on channel 18 to confirm stimulus onset timing. Each echo stimulus was broadcast 20 times, randomly selected from the entire stimulus set. Afterward each tone-intensity combination was randomly presented 15 times to generate frequency tuning curves. After each day's recording session, the craniotomy site was covered with kwik-cast gel to prevent dehydration.

Analysis

Single units were isolated following methods outlined by Quiroga et al. (22). The Wavelet transformation and the superparamagnetic clustering resulted in isolation of single-unit extracellular potentials that matched with qualitative assessments of spike waveforms and estimates of single-unit isolation based on spike refractory periods. Neural responses of each unit to all acoustical stimuli were visualized with dot raster displays and poststimulus time histograms (PSTH, 1-ms bin width). For each echo stimulus, neural responses were quantified by measuring the number of spikes in a 45-ms window from call onset. For the call discrimination analysis, the window was reduced to 10 ms to avoid overlap with responses to echo return. Neurons with fewer than five spikes in the entire 45-ms response window were not considered responsive and thus discarded from the analysis.

Population Clustering

A total of 422 responsive units were initially selected for analysis. To cluster units into groups by response profile, a 45-ms echo poststimulus window was created. Call-echo PSTHs were compared across units. Call-echo responses for each object were averaged across all 20 repetitions. The mean PSTHs were concatenated to create a single vector

PSTH for each unit. The response patterns across units were highly variable, but several groups showed similarities, such as tonic responses during stimulation, or isolated activity to pulses, echoes, or both, suggesting that neurons responded to different stimulus components. To quickly cluster groups of neurons that showed similar response properties *t*-distributed stochastic neighbor embedding (23) was performed using MATLAB. A *t*-SNE (three-dimensional) was run on average PSTHs for each unit to compute response distance to sort units into groups. Plotting the results of the *t*-SNE as a three-dimensional scatter plot with each point the average PSTH of a unit showed roughly three clusters. Individual units were assigned to each population through *k*-means (*k* = 3) clustering of the three-dimensional response distances (Fig. 2).

Discrimination Analysis

Discrimination analyses were based on methods described by Allen and Marsat (24) using a modified van Rossum spike distance metric (25) and the discrimination toolkit described by Marsat (26). The measure is based on comparing averaged neural population responses to either repeated presentations of the same echo (interecho distances) or pairwise comparisons to the other echoes (intraecho distances). The 45-ms PSTHs used to identify populations aforementioned were separated into two separate windows around the call-evoked portion of the responses [$R_c(t)$] and echo-evoked [$R_e(t)$] portion of the responses of length *L* ($L_c = 10$ ms, $L_e = 25$ ms) extracted from the sorted spike trains. [$R_c(t)$] comprised time 10–19 ms post stimulus and [$R_e(t)$] of *t* (20–45). Individual neuronal responses varied within these windows, but always fell in these time frames. The responses were convolved with a Gaussian filter with a half-maximum width of 2 ms. Two hundred randomly selected responses to a single repetition of each echo were averaged to create a sample population response for both call and echo responses (PR_c and

PR_e) using the function: $PR_z(t) = \sum_{i=1}^n \frac{R_z(t)}{n}$. Distance (*D*) was calculated for all sets of combined responses, creating an array of response distances for each comparison (inter- [D_{xx}] and intraecho [D_{xy}]) using the following function:

$D_{xy} = \frac{\sum_{t=0}^L [R_x(t) - R_y(t)]^2}{L}$. The probability distributions of the values in these arrays [$P(D)$] were used for analysis. Receiver operator characteristic curves were calculated by varying a threshold distance (*T*) to separate inter- (*xx*) and intraecho (*xy*) responses. This process was also repeated on intracall responses. For each threshold, the probability of discrimination (PD) was calculated as the sum of *P* ($D_{xy} \times T$), and the probability of false alarm (PF) as the sum of *P* ($D_{xx} \times T$). This process was repeated through 100 iterations, each time selecting a new pool of random call and echo responses to create a bootstrapped estimation of true population responses. The error level for each threshold value is $E \times PF/2 \cdot (1 \times PD)/2$. The discrimination rates reported in Figs. 3 and 5 are the inverse minimum value of *E* averaged across iterations (i.e., Accuracy = 100 – *E*) with standard error.

Discriminability Analysis with a Biomimetic Midbrain Model

Using a computational model that replicates the frequency modulation and sweep direction selectivity of the bat's IC neurons, we performed additional analyses to explore the effect of clutter on discrimination. We built an ensemble model by gathering output nodes (i.e., artificial neurons) of three biomimetic networks (300 total nodes). Each biomimetic network was constructed with the same configuration: a triple stacking network with 10% sparsity and nonlinear activation (leaky ReLU) (27).

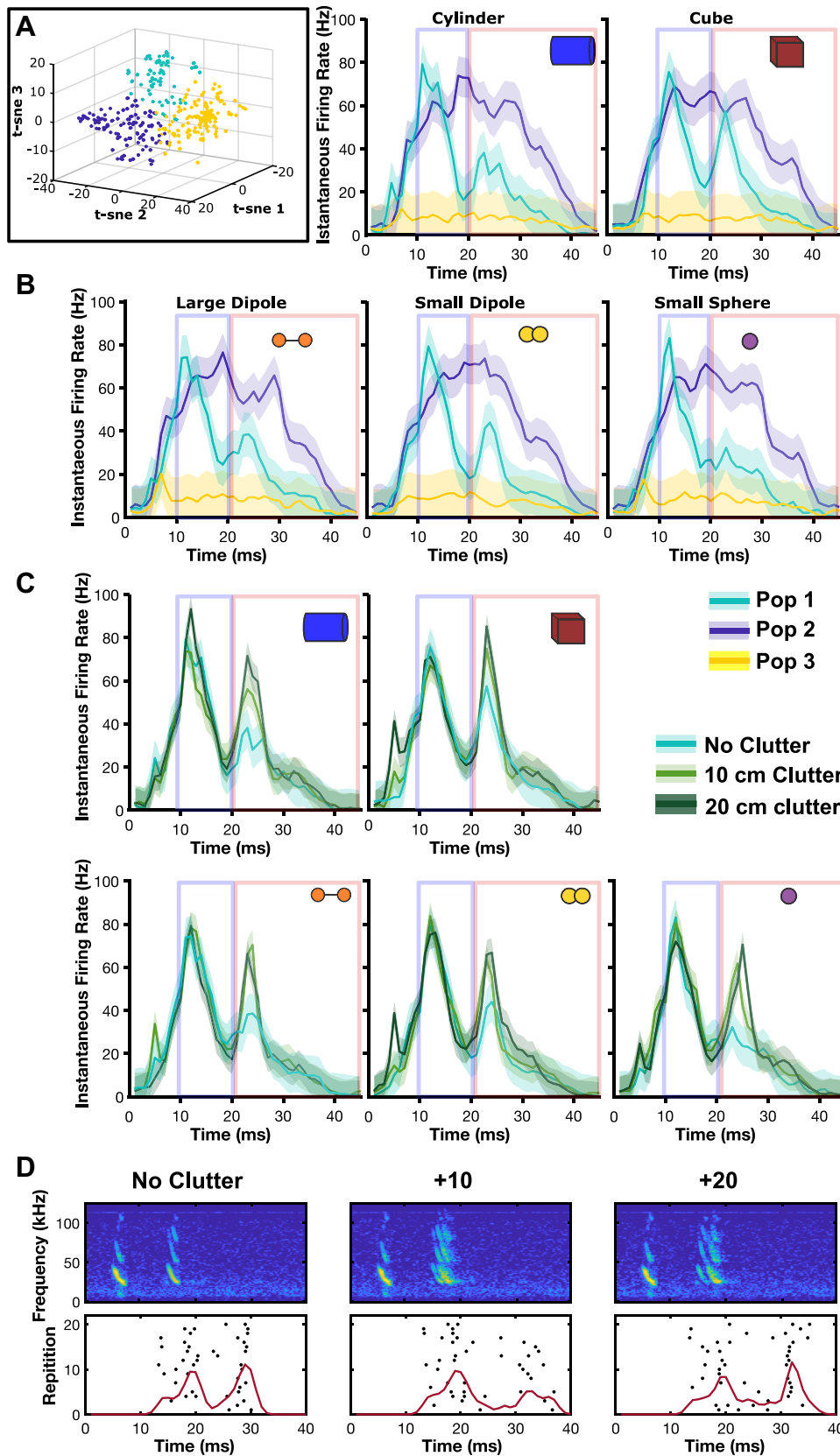
In the first analysis, we reperformed the neural discrimination of object echoes with the computational model. After upsampling to 333 kHz, the acoustic stimuli used in the physiology were converted to auditory spectrograms with 128 frequency bins (28). To obtain responses of the computational model $PR(t)$, we built a response model for node *k* as

$$PR^k(t) = \text{Bernoulli}\{\sigma(M^k(t))\}$$

$$M^k(t) = m^k(t) + \gamma g^k(t) + b^k,$$

where $M^k(t)$ is membrane potential resulting from a stimulus and disturbance. With a 32-ms sliding window in steps of 0.2 ms, stimulus potential $m^k(t)$ is calculated over the time. For disturbance, $g^k(t)$ and b^k is noise and bias designed by Gaussian and Poisson distribution, respectively. γ is a parameter to adjust the power ratio of stimulus potential and Gaussian noise. And $\text{Bernoulli}\{\cdot\}$ is the Bernoulli sampling based on activation probability represented by applying sigmoid $\sigma(\cdot)$ to the membrane potential. We separated the responses into two parts responding to call $PR_c^k(t)$ and echo $PR_e^k(t)$, respectively. As in the neural recordings, features to calculate distance were obtained by performing a convolution between the responses $PR_e^k(t)$ for 25 ms and a Gaussian filter with a half-maximum width of 1 ms. Then, we performed the pairwise discrimination with distributions of intra- and interdistance (D_{xx} and D_{xy}) on each node in the same manner as the neural recording data. Based on average discrimination rate for all discrimination pairs, 34-best informative nodes were used in the analysis of clutter effects. Note that the number of nodes was heuristically optimized near the 10% sparsity. We performed the discrimination 100 times due to randomness in the Bernoulli sampling and disturbance in the response model, and the results were summarized to mean and standard deviation over all repetitions.

Although the pairwise discrimination tests in IC and artificial neurons were performed on a single acoustic stimulus per each object, in the second analysis, we compared discriminability in various clutter conditions with multiple stimuli per each of the objects mentioned in *Acoustic Stimuli*. Instead of membrane potential over the time, we used a react feature $m = [m^k | 1 \leq k \leq 34]^T$ that represents stimulus potential across the nodes applied in the first analysis. Note that a time index is omitted in the react feature because the feature was calculated for a single rectangular patch spanning 128 frequency bins and 32 ms and encompassing echoes from object and clutter in the spectrogram. To quantify the discriminability, we defined an *f*-distance

C
O
L
O
R

AQ: 20

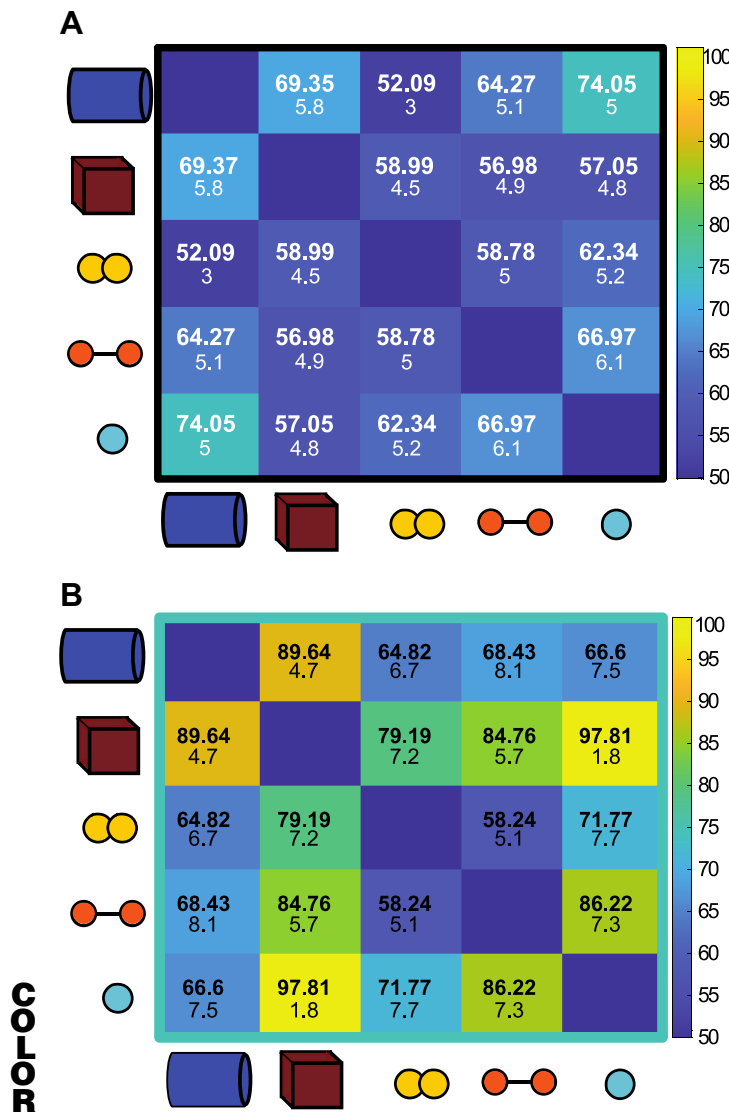


Figure 3. Pairwise discrimination of objects by the inferior colliculus (IC). **A:** population responses evoked by repeated presentations of calls. Numbers shown indicate mean percent discrimination accuracy after 100 iterations (100: perfect accuracy, 50: chance level). Numbers shown indicate mean percent discrimination accuracy after 100 iterations. Smaller numbers indicate standard deviation around that mean. Note that some amount of above-chance level discrimination accuracy is achieved. This is potentially due to variations in background noise present in the recordings themselves. **B:** population responses evoked by repeated presentations of ensorified objects are generally more discriminable than calls, with varying levels of accuracy.

that represents the ratio of interclass to intraclass variance of react features $f - dist. = trace(C_B)/trace(C_W)$, where $C_B = E_i[E_{m \in C_i}[m]E_{m \in C_i}[m]^T]$ is a covariance matrix for the mean of react features in each class and $C_W = E_i[E_{m \in C_i}[mm^T]]$ is an averaging covariance matrix of intraclass covariances. The $trace(.)$ is an operator for summation of diagonal elements that is equal to summation of its eigenvalues. If the feature distributions have small variances within class and large variance between classes, the features can be easily categorized. Thus, the larger the f -distance, the more discriminable the features.

RESULTS

A Population of IC Neurons Responds Differentially to Different Sonar Objects

We analyzed the firing rate profile in the responses of 422 single units in the bat midbrain inferior colliculus to playbacks of call-echo pairs described in METHODS. Briefly, a synthetic call was broadcast at different five simple geometric objects (Table 1) and the returning echoes were collected by microphones located right above the speaker (Fig. 1). A t -SNE analysis revealed three distinct populations of neurons based on the firing rate profile (Fig. 2). *Population I* ($n = 89$) showed clear firing rate peaks at the call and at the echo, respectively. Although the responses to the call were highly similar and on average less accurately discriminated from each other (Fig. 3A), visual inspection of the firing rate peak corresponding to the response to the distinct object echoes revealed differences among objects. *Population II* ($n = 150$) showed sustained firing rates during the whole call-echo pair stimulus presentation and did not show differences in the responses to the distinct objects. *Population III* ($n = 183$) showed very low firing rate and was excluded from further analysis.

Further analysis of *population I* revealed that the response patterns to the distinct echoes of the different objects in the no clutter condition can be correctly discriminated by a spike distance metric classifier response profile with >90% accuracy (Fig. 3B). *Population II* and *population III* showed average classification accuracy of 61.5% and 52.9%, respectively (Supplemental Fig. S1; all Supplemental material is available at <https://doi.org/10.6084/m9.figshare.14639856>) in the no clutter condition. For *population I*, the cube was most consistently well discriminated from other objects with an average discrimination accuracy of 87.8% against other objects, whereas overall discrimination accuracy was 76.7% (range for all objects 55.4%–99.6%). The most accurate discrimination was achieved in comparing the cube against the small sphere, reflecting the visually apparent differences in firing rates evoked by those objects.

Neurons Involved in Sonar Object Discrimination Show Frequency Tuning to Enhanced Spectral Content of Echoes

We evaluated the frequency tuning characteristics of the units by playing pure tones ranging from 20 kHz to 90 kHz at intensities ranging from 20 to 70 dB SPL. Candidate neurons for sonar object discrimination (*population I*) show a wide range of frequency selectivity (Fig. 4, A and B). In accordance with previously reported results, tonotopy appeared in the IC, with deeper recordings showing tuning to higher frequencies and more dorsal units being tuned to lower frequencies (Fig. 4C). Interestingly, 41.5% of units in *population I* showed tuning below 30 kHz, at which the different echoes show the greatest variability in power (Fig. 4D). We calculated the variance in the power spectrum across objects in 5 kHz bins starting at 20 kHz, we then compared the number of neurons with BF in each bin to the calculated variance and found no correlation (Supplemental Fig. S2). There are some neurons tuned to frequencies where the variance in the power of the echoes is minimal. This

F4

AQ: 10

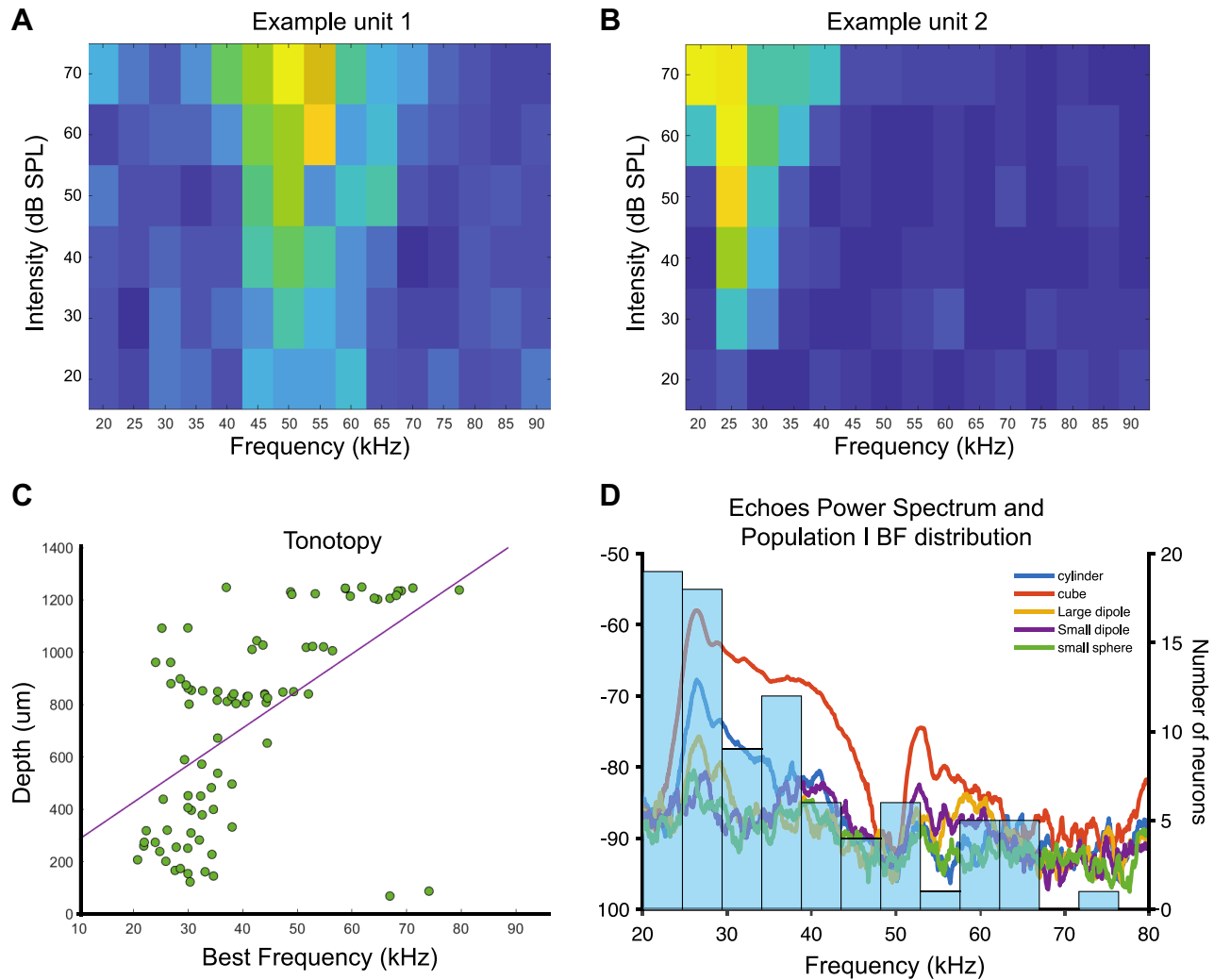


Figure 4. Frequency tuning in *population I*. *A*: example unit tuned to 50 kHz. *B*: example unit tuned to 25 kHz. *C*: tonotopy evidenced by best frequency for neurons in *population I* (overlapping data points have been offset slightly to see all units). *D*: distribution of best frequencies for *population I* units over the power spectrum of the echoes used for the playbacks.

implies that at least for some units in *population I*, object selectivity may not arise from frequency selectivity alone. The intensity of the echoes could be one of the main drivers in the differential neural responses between objects, this is supported by the fact that the cube (which shows the greatest energy across frequencies) and the small sphere (which shows the lowest energy across frequencies) are the most accurately discriminated objects.

Object Discrimination is Affected by Clutter

To understand the effects of physical clutter on object discrimination in the IC, we incorporated acoustic noise created by foliage into the object echo recordings. Plastic plants were suspended behind objects at close (+10 cm) and far (+20 cm) distances during ensonification. The echoes from the suspended foliage overlapped with the object echoes in both configurations, although the close clutter condition results in a near complete overlap between object and clutter echoes (Fig. 1). The echoes resulting from the clutter created spectral

interference with the object echoes observable in the power spectra of echoes in both close and far clutter conditions (Fig. 5). When presented with close clutter echoes, discrimination performance dropped for the majority of object pairs, and average discrimination performance was 70.7%, compared with 76.7% for the no clutter condition. Far clutter condition echoes, on average, produced higher accuracy discrimination than baseline (no clutter) echoes, up to 82.6%. For several object pairs, discrimination performance improved, but not up to the level of No Clutter echoes. However, for some object pairs, the inclusion of clutter in the recording increases object discriminability above that of the no clutter condition (Fig. 5; Supplemental Table S1). Notably, the sphere in particular was improved by distant clutter. Comparing the spectral features of the echoes with and without clutter suggests that both overall echo amplitude and the location of spectral notches in the echo profiles are altered by clutter at +10 and +20 cm.

F5

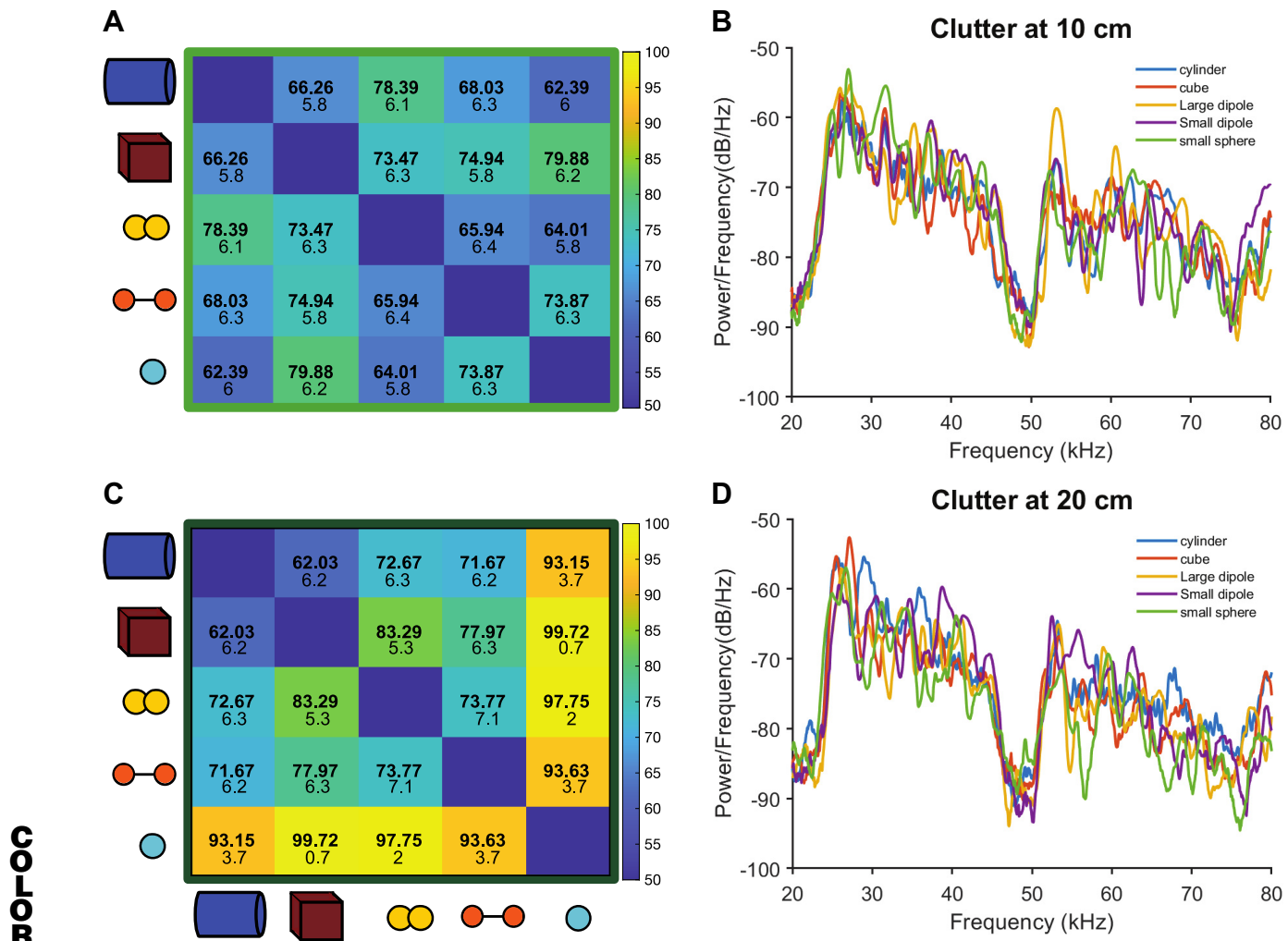


Figure 5. Clutter affects object discrimination differentially based on distance. **A:** close clutter (+10 cm) impairs pairwise discrimination of all objects compared with the no clutter condition. **B:** power spectra for echoes recorded with 10 cm clutter. **C:** far clutter (+20 cm) interferes with discrimination less than close clutter, and for some object pairs, improves discrimination over the no clutter condition. **D:** power spectra for echoes recorded with 20 cm clutter.

A Biomimetic Model of the Bat Midbrain Replicates the Effect of Clutter on Discriminability by IC Neurons

Before exploring the effect of clutter echoes, we first re-performed object discrimination with the computational model using the same stimuli presented in neural recording experiments. The modeling results are summarized in Fig. 6, alongside the results from IC neurons. The results are marked as orange for enhanced discriminability or blue for decreased discriminability compared with no clutter conditions. Green denotes no change or no value (e.g., diagonal elements). Mean values were compared to quantify enhanced, decreased, or no-change. Although the values for the neural and model discrimination are different, we find the same trends with respect to the effect of clutter on discrimination. In the no clutter condition, IC neurons show the most accurate discrimination between the cube and the small sphere, whereas the computational model shows the highest discrimination accuracy between the cube and the cylinder. Still, both biological and artificial neurons result in high discriminability (>80%)

for many pairwise comparisons. In the +20 cm clutter condition, IC and biomimetic model discrimination is enhanced for all stimuli, except for the cube and cylinder. Consistent with the neural recording results, small sphere biomimetic model discrimination was significantly enhanced compared with the no clutter condition. Although the computational model showed less improvement following the introduction of clutter than the neural data, the overall trends are similar. In the +10 cm clutter condition, object discriminability is generally decreased, except for the small dipole, which shows enhanced discriminability in all stimulus pairings in the computational model and only in pairings with the cylinder and large dipole in the neural recording data.

Next, we explored the effect of clutter on discriminability quantified in f -distance across multiple echo recordings from the objects to control for effects caused by variations in acoustic background noise. By bootstrapping with 16 samples of 20, f -distances of three different representations (computational model, IC neurons, and auditory spectrogram) are summarized as mean and standard

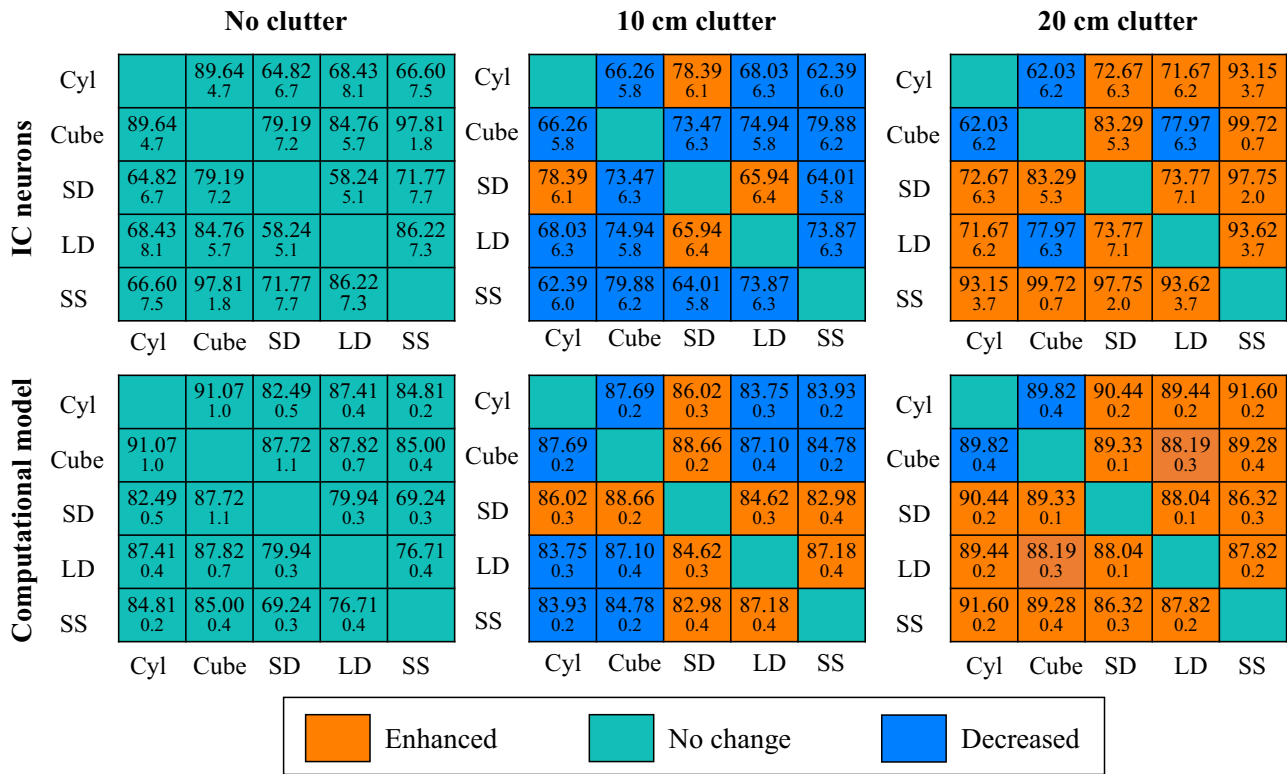


Figure 6. Pairwise discrimination rates using a biomimetic model matches inferior colliculus (IC) neurons. Comparing discrimination accuracy between IC neurons (*top*) and artificial neurons (*bottom*) in three clutter conditions: No Clutter, 10 cm, and 20 cm shows similar trends: decreased discrimination accuracy in close clutter conditions, and enhanced accuracy in far clutter conditions. Cyl, cylinder; LD, large dipole; SD: small dipole; SS, small sphere.

T2 deviation of 100 iterations (Table 2). The f -distances of each representation are normalized by the value in the no clutter condition. Note that the f -distance of IC neurons was calculated with a single set of recordings (20 stimulus presentations) for each object echo. An auditory spectrogram represents the time-frequency distribution of stimulation along the cochlea. We used the auditory spectrogram as an input to the computational model. The f -distances between spectrograms suggest that background clutter makes object discrimination more difficult, even in the +20 cm clutter condition. However, both IC neurons and the computational model show that discriminability is enhanced in the far clutter condition (+20 cm) and it is decreased when the clutter is closer to the object (+10 cm). This analysis shows that the enhancement effect of clutter is not an artifact of single-neural recordings.

Table 2. The enhancement by clutter has been reproduced in simulation with a biomimetic model for multiple echo recordings

Feature	No Clutter	20 cm Clutter	10 cm Clutter
Computational model	1.00	1.2811 ± 0.0454	0.7744 ± 0.0230
IC neurons*	1.00	1.1137 ± 0.1370	0.9235 ± 0.1242
Spectrogram	1.00	0.8742 ± 0.2153	0.8666 ± 0.2404

IC, inferior colliculus. *Performed with a single recording per each object.

DISCUSSION

Humans and other animals must often detect, discriminate, and extract meaning from target sounds under noisy conditions. A great deal of research has been devoted to characterizing the spectral and temporal features of acoustic stimuli that influence the perceptual organization of complex sounds (29–33), but comparatively little attention has been devoted to the neural basis of auditory figure-ground segregation, the separation of salient target sounds from overlapping and interfering noise (34, 35). In many other species, such noise with temporal overlap would create a backward masking effect that interferes with sonar object discrimination. Here, we report the new discovery that acoustic scattering from background clutter can increase neural discrimination of sonar objects when object and clutter echoes are separated in time.

Echolocating bats need not only detect sonar objects, such as prey, and discriminate them from other sound sources in the environment but also must separate targets from background vegetation and other clutter. Our study investigated the robustness of acoustic discrimination of neurons in the bat IC, with the specific goal of identifying principles governing sonar object representation and figure-ground segregation.

In baseline recordings, we show that a population of IC neurons reveal distinct response profiles to echoes from different objects. The frequency tuning of the neurons to pure tones did not fully account for the differential responses. In other words, even though the content of the echoes from the

sonar objects differed in energy across spectral bands, frequency tuning alone could not fully explain sonar object discrimination. We posit there are neurons in the inferior colliculus that exhibit nonlinear feature selectivity that contributes to sonar object discrimination.

When we presented echoes from sonar objects in the presence of clutter, neural discrimination depended on the spatiotemporal separation of the sonar object and background clutter echoes. When the clutter was very close to the object (10 cm) and echoes temporally overlapped, discrimination performance of the classifier decreased. This could be explained by backward masking (35), in which the detection or discrimination of a target sound is reduced when followed by noise, as it would be the case for the target echoes followed by clutter echoes. Spectral interference created by constructive and destructive interference between overlapping echoes could also play a role in this reduction in discrimination. These physical interactions from temporally overlapping echoes could result in target echoes insufficiently different from each other to evoke distinct responses from the neurons. However, when the clutter was further behind the object (20 cm), clutter and target echoes only partially overlapped, and the classifier's discrimination accuracy based on neural responses to some objects improved markedly. Due to the increased temporal separation of the target and clutter echoes, the effect of spectral interference is lessened, compared with the close clutter condition, but not eliminated entirely. The power spectra of the object echoes with distant clutter showed frequency signatures for each object that are diminished when the clutter was in close proximity. These findings were also replicated using a biomimetic model of the bat auditory midbrain. We conclude that background clutter echoes can yield additional spectrotemporal information that modulates the responses of IC neurons, contributing to object discrimination. We hypothesize that distant clutter returns acoustic reflections from the backside of the target object, providing additional information about object identity and 3-D shape.

Another potential mechanism to explain the effect we report could be stochastic resonance, whereby the signal:noise ratio (SNR) of the object echo is increased by the addition of background noise created by the clutter echoes. Classically, stochastic resonance is achieved when white noise is added to a target stimulus, which intermittently amplifies the target's SNR (36). In the case of the background clutter used in our experiment, clutter echoes interacting with the target echoes could serve to amplify acoustic energy that carries information about the object, providing the auditory system with enough signal to amplify acoustic features that enable discrimination of target echoes. Stochastic resonance has been described across different sensory modalities, including mechanoreception, electroreception, vision, and audition (36). For example, in the paddlefish, it was shown that populations of plankton produce electrical noise that enhances signals processed by the electroreceptors and allows for increased success in the detection and capture of the prey (37). In humans, psychophysical experiments showed that the addition of white noise (at optimal noise levels) improved the detection of pure tones, but noise at higher levels produced a masking effect (38). The two hypotheses described above (spectral interference and stochastic resonance) are not necessarily mutually exclusive in the interpretation of our findings.

It is possible that the detection of new spectral features created through interference is enhanced by the increased energy of high-intensity clutter echoes.

We propose that the overarching explanation for increased discriminability of target echoes in clutter is the result of "acoustic mirroring," by which additional information about objects can be gleaned from secondary reflections that arise through reverberation with the surfaces of 3-D targets. This suggests that bats could obtain information from the complex spectral interactions, which may enhance object echo signatures. Future behavioral experiments will be necessary to fully test this hypothesis, both to determine whether the changes in object representations we observe in the study can be used by bats for object discrimination, and to fully explore the extent to which secondary reflections can be leveraged by bats to explore hidden object faces and features. The acoustic features and interactions we speculate on here will also be fully explored computationally to understand the processes by which acoustic interactions from multiple reflecting surfaces produce information that can be detected and discriminated by the bat auditory system. This study sought to explore the effect of acoustic clutter on sonar object discrimination by neurons in the IC of the big brown bat. We made the surprising discovery that clutter can enhance the discriminability of sonar targets based on neural responses, but this effect is constrained by the spatiotemporal overlap between target and clutter echoes. This finding contributes to the understanding of sonar object representation, figure-ground segregation, and more broadly the neural encoding of complex acoustic stimuli.

SUPPLEMENTAL DATA

Supplemental Figs. S1 and S2 and Table S1, including discrimination analyses for *populations II* and *III*, and variance analysis of power spectra: <https://doi.org/10.6084/m9.figshare.14639856>.

AQ: 13

ACKNOWLEDGMENTS

We thank Kaylee Fu and Chris Farid for help with pre-processing of data.

GRANTS

The study is supported by the Human Frontiers Science Program long-term fellowship Award LT000220/2018 (to A.S.), National Institute of Health Training grant postdoctoral fellowship Award 5T32DC000023-35 (to K.M.A.; Pls Kathleen Cullen and Paul Fuchs). Also by Brain Initiative NCS-FO 1734744 (2017-2021), AFOSR FA9550-14-1-0398NIFTI, and ONR N00014-17-1-2736 (to C.F.M.).

AQ: 14
AQ: 15

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AQ: 16

AUTHOR CONTRIBUTIONS

K.M.A., A.S., M.E., and C.F.M. conceived and designed research; K.M.A. and A.S. performed experiments; K.M.A., A.S., and S.P. analyzed data; K.M.A., A.S., S.P., M.E., and C.F.M. interpreted results of experiments; K.M.A., A.S., and S.P. prepared figures; K.M.A. and A.S. drafted manuscript; K.M.A., A.S., S.P., M.E.,

and C.F.M. edited and revised manuscript; K.M.A., A.S., S.P., M.E., and C.F.M. approved final version of manuscript.

AQ: 17

REFERENCES

- Moss CF, Bohn K, Gilkenson H, Surlykke A. Active listening for spatial orientation in a complex auditory scene. *PLoS Biol* 4: e79, 2006. doi:10.1371/journal.pbio.0040079.
- Surlykke A, Moss CF. Echolocation behavior of big brown bats, *Eptesicus fuscus*, in the field and the laboratory. *J Acoust Soc Am* 108: 2419–2429, 2000. doi:10.1121/1.1315295.
- Simmons JA, Eastman KM, Horowitz SS, O'Farrell MJ, Lee DN. Versatility of biosonar in the big brown bat, *Eptesicus fuscus*. *Acoust Res Lett Online* 2: 43–48, 2001. doi:10.1121/1.1352717.
- Bates ME, Simmons JA, Zorikov TV. Bats use echo harmonic structure to distinguish their targets from background clutter. *Science* 333: 627–630, 2011. doi:10.1126/science.1202065.
- Stamper SA, Simmons JA, DeLong CM, Bragg R. Detection of targets colocalized in clutter by big brown bats (*Eptesicus fuscus*). *J Acoust Soc Am* 124: 667–673, 2008. doi:10.1121/1.2932338.
- Taub M, Yovel Y. Segregating signal from noise through movement in echolocating bats. *Sci Rep* 10: 382, 2020. doi:10.1038/s41598-019-57346-2.
- Mao B, Aytekin M, Wilkinson GS, Moss CF. Big brown bats (*Eptesicus fuscus*) reveal diverse strategies for sonar target tracking in clutter. *J Acoust Soc Am* 140: 1839, 2016. doi:10.1121/1.4962496.
- Wheeler AR, Fulton KA, Gaudette JE, Simmons RA, Matsuo I, Simmons JA. Echolocating big brown bats, *Eptesicus fuscus*, modulate pulse intervals to overcome range ambiguity in cluttered surroundings. *Front Behav Neurosci* 10: 125, 2016. doi:10.3389/fnbeh.2016.00125.
- Wagenhäuser PA, Wiegrebe L, Baier AL. Biosonar spatial resolution along the distance axis: Revisiting the clutter interference zone. *J Exp Biol* 223: jeb224311, 2020. doi:10.1242/jeb.224311.
- Falk B, Williams T, Aytekin M, Moss CF. Adaptive behavior for texture discrimination by the free-flying big brown bat, *Eptesicus fuscus*. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 197: 491–503, 2011. doi:10.1007/s00359-010-0621-6.
- Poon PWF, Sun X, Kamada T, Jen PH-S. Frequency and space representation in the inferior colliculus of the FM bat, *Eptesicus fuscus*. *Exp Brain Res* 79: 8391, 1990. doi:10.1007/BF00228875.
- Pollak GD, Schuller G. Tonotopic organization and encoding features of single units in inferior colliculus of horseshoe bats: functional implications for prey identification. *J Neurophysiol* 45: 208–226, 1981. doi:10.1152/jn.1981.45.2.208.
- Casseday JH, Covey E. Frequency tuning properties of neurons in the inferior colliculus of an FM bat. *J Comp Neurol* 319: 34–50, 1992. doi:10.1002/cne.903190106.
- Salles A, Park S, Sundar H, Macías S, Elhilali M, Moss CF. Neural response selectivity to natural sounds in the bat midbrain. *Neuroscience* 434: 200–211, 2020. doi:10.1016/j.neuroscience.2019.11.047.
- Casseday JH, Ehrlich D, Covey E. Neural tuning for sound duration: role of inhibitory mechanisms in the inferior colliculus. *Science* 264: 847–850, 1994. doi:10.1126/science.8171341.
- Fremouw T, Faure PA, Casseday JH, Covey E. Duration selectivity of neurons in the inferior colliculus of the big brown bat: tolerance to changes in sound level. *J Neurophysiol* 94: 1869–1878, 2005. doi:10.1152/jn.00253.2005.
- Wenstrup JJ, Portfors CV. Neural processing of target distance by echolocating bats: Functional roles of the auditory midbrain. *Neurosci Biobehav Rev* 35: 2073–2083, 2011. doi:10.1016/j.neubiorev.2010.12.015.
- Suga N. Echo-ranging neurons in the inferior colliculus of bats. *Science* 170: 449–452, 1970. doi:10.1126/science.170.3956.449.
- Macías S, Luo J, Moss CF. Natural echolocation sequences evoke echo-delay selectivity in the auditory midbrain of the FM bat, *Eptesicus fuscus*. *J Neurophysiol* 120: 1323–1339, 2018. doi:10.1152/jn.00160.2018.
- Morrison JA, Valdizón-Rodríguez R, Goldreich D, Faure PA. Tuning for rate and duration of frequency-modulated sweeps in the mammalian inferior colliculus. *J Neurophysiol* 120: 985–997, 2018. doi:10.1152/jn.00065.2018.
- Luo J, Macías S, Ness TV, Einevoll GT, Zhang K, Moss CF. Neural timing of stimulus events with microsecond precision. *PLoS Biol* 16: e2006422, 2018. doi:10.1371/journal.pbio.2006422.
- Quiroga RQ, Nadasdy Z, Ben-Shaul Y. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput* 16: 1661–1687, 2004. doi:10.1162/089976604774201631.
- Van der Maaten L, Hinton G. Visualizing data using t-SNE. *J Mach Learn Res* 9: 2579–2605, 2008.
- Allen KM, Marsat G. Task-specific sensory coding strategies are matched to detection and discrimination performance. *J Exp Biol* 221: jeb170563, 2018. doi:10.1242/jeb.170563.
- van Rossum MCW. A novel spike distance. *Neural Comput* 13: 751–763, 2001. doi:10.1162/089976601300014321.
- Marsat G. Characterizing neural coding performance for populations of sensory neurons: Comparing a weighted spike distance metrics to other analytical methods. *BioRxiv*, 2019. doi:10.1101/778514.
- Park S, Salles A, Allen K, Moss CF, Elhilali M. Natural statistics as inference principles of auditory tuning in biological and artificial midbrain networks. *eNeuro* 8: ENEURO.0525-20.2021, 2021. doi:10.1523/ENEURO.0525-20.2021.
- Chi T, Ru P, Shamma SA. Multiresolution spectrotemporal analysis of complex sounds. *J Acoust Soc Am* 118: 887–906, 2005. doi:10.1121/1.1945807.
- Bregman AS. *Auditory Scene Analysis*. Cambridge, MA: MIT Press, 1994.
- Elhilali M, Ma L, Michey C, Oxenham AJ, Shamma SA. Temporal coherence in the perceptual organization and cortical representation of auditory scenes. *Neuron* 61: 317–329, 2009. doi:10.1016/j.neuron.2008.12.005.
- Shamma SA, Elhilali M, Michey C. Temporal coherence and attention in auditory scene analysis. *Trends Neurosci* 34: 114–123, 2011. doi:10.1016/j.tins.2010.11.002.
- Bidet-Caulet A, Bertrand O. Neurophysiological mechanisms involved in auditory perceptual organization. *Front Neurosci* 3: 182–191, 2009. doi:10.3389/neuro.01.025.2009.
- Pressnitzer D, Sayles M, Michey C, Winter IM. Perceptual organization of sound begins in the auditory periphery. *Curr Biol* 18: 1124–1128, 2008. doi:10.1016/j.cub.2008.06.053.
- Michey C, Oxenham AJ. Pitch, harmonicity and concurrent sound segregation: psychoacoustical and neurophysiological findings. *Hear Res* 266: 36–51, 2010. doi:10.1016/j.heares.2009.09.012.
- Elliott LL. Backward masking: monotonic and dichotic conditions. *J Acoust Soc Am* 34: 1108–1115, 1962. doi:10.1121/1.1918253.
- Moss F, Ward LM, Sannita WG. Stochastic resonance and sensory information processing: a tutorial and review of application. *Clin Neurophysiol* 115: 267–281, 2004. doi:10.1016/j.clinph.2003.09.014.
- Russell DF, Wilkens LA, Moss F. Use of behavioural stochastic resonance by paddle fish for feeding. *Nature* 402: 291–294, 1999. doi:10.1038/46279.
- Zeng FG, Fu Q-J, Morse R. Human hearing enhanced by noise. *Brain Res* 869: 251–255, 2000. doi:10.1016/S0006-8993(00)02475-6.
- Teki S, Chait M, Kumar S, von Kriegstein K, Griffiths TD. Brain bases for auditory stimulus-driven figure-ground segregation. *J Neurosci* 31: 164–171, 2011. doi:10.1523/JNEUROSCI.3788-10.2011.

AQ: 18

AQ: 19

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

1

AQau— Please confirm the given-names and surnames are identified properly by the colors.

■ = Given-Name, ■ = Surname

AQ1— Please read entire proof (including query page), answer queries and mark corrections (using Adobe tools per online instruction), and return it within 2 business days to the online proof review site (link provided in email). IMPORTANT: Make sure query page is included. Please only make changes that are essential to correct data errors. Requests for cosmetic and other nonessential changes to figures and text at this stage will not be accommodated.

AQ2— Please check all figures, tables, equations, and legends carefully. Changes will be made only to correct scientifically relevant errors. Please confirm or correct abbreviation definitions that have been inserted in the legends by the copy editor. If changes are required to the figures themselves, please list ALL changes requested (even if you provide a new figure).

AQ3— Please note: with the updated 2021 journal format, JNP references have been renumbered according to order of first mention in the text. Please check carefully throughout that callouts link to the correct references.

AQ4— Please verify accuracy of your e-mail for correspondence. (Note: this material is listed as a footnote at bottom of left column of text on the first page.)

AQ5— Up to five key words are allowed per Journal style; please reduce your list to five.

AQ6— Please note that your manuscript has been copyedited and that style changes have been made, including those related to punctuation, abbreviations, italics, hyphens and word divisions, units of measurement, etc., to conform to Journal style preferences. Please check that no errors in meaning were inadvertently introduced. Not all of the changes are pointed out in the Author Query list, so please check all text carefully.

AQ7— Please carefully review the hierarchy of headings throughout your article.

AQ8— Please define “SPL” at first use unless this is a standard abbreviation in your field.

AQ9— Please define “t-SNE” at first use unless this is a standard abbreviation in your field.

AQ10— Please define “BF” at first use unless this is a standard abbreviation in your field.

AQ11— Please check the accuracy of the phrase “(1 + 0 cm)” in the sentence and revise as needed.

AQ12— Please specify the exact section instead of using above/below.

AQ13— Please verify that all supplemental DOI links are public and active and access all intended supplemental material.

AQ14— Please verify that any GRANTS are listed correctly.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

2

AQ15— Please define “AFOSR” and “ONR” at first use unless these are standard abbreviations in your field.

AQ16— The text in the DISCLOSURES section reflects your data entry into the Peer Review submission. Is this still complete, relevant, and accurate?

AQ17— The text in the AUTHOR CONTRIBUTIONS section reflects your data entry into the Peer Review submission. Is this still complete and accurate? If changes are needed, please follow this template.

AQ18— If possible, please provide complete information (volume, pages) for Ref. 26.

AQ19— Ref. 39 is not cited in text. Please cite in text or delete from the reference list.

AQ20— Please define “ISI” at first use unless this is a standard abbreviation in your field.
