ORIGINAL ARTICLE

Focal Suppression of Distractor Sounds by Selective Attention in Auditory Cortex

Zachary P. Schwartz[1] and Stephen V. David[2]

[1]Neuroscience Graduate Program, Oregon Health and Science University, OR, USA and [2]Oregon Hearing Research Center, Oregon Health and Science University, OR, USA

Address Correspondence to Stephen V. David, Oregon Hearing Research Center, Oregon Health and Science University, 3181 SW Sam Jackson Park Road, MC L335A, Portland, OR 97239, USA. Email: davids@ohsu.edu

Abstract

Auditory selective attention is required for parsing crowded acoustic environments, but cortical systems mediating the influence of behavioral state on auditory perception are not well characterized. Previous neurophysiological studies suggest that attention produces a general enhancement of neural responses to important target sounds versus irrelevant distractors. However, behavioral studies suggest that in the presence of masking noise, attention provides a focal suppression of distractors that compete with targets. Here, we compared effects of attention on cortical responses to masking versus non-masking distractors, controlling for effects of listening effort and general task engagement. We recorded single-unit activity from primary auditory cortex (A1) of ferrets during behavior and found that selective attention decreased responses to distractors masking targets in the same spectral band, compared with spectrally distinct distractors. This suppression enhanced neural target detection thresholds, suggesting that limited attention resources serve to focally suppress responses to distractors that interfere with target detection. Changing effort by manipulating target salience consistently modulated spontaneous but not evoked activity. Task engagement and changing effort tended to affect the same neurons, while attention affected an independent population, suggesting that distinct feedback circuits mediate effects of attention and effort in A1.

Key words: behavioral state, ferret, listening effort, neural coding, task engagement

Introduction

Humans and other animals are able to focus attention on 1 of multiple competing sounds in order to resolve details in behaviorally important signals (Cherry 1953; Dai et al. 1991; Shinn-Cunningham and Best 2008). Studies in humans have found that when attention is directed to 1 of 2 competing auditory streams, local field potential (LFP), MEG, and/or fMRI BOLD responses to task-relevant features are enhanced, and responses to distractor stimuli are generally suppressed (Ding and Simon 2012; Mesgarani and Chang 2012; Da Costa et al. 2013). However, little is known about the effect of attention when distractors compete in the same spectral band as target sounds, an important problem for hearing in natural noisy environments (Shinn-Cunningham and Best 2008). Human behavioral studies suggest auditory attention does not uniformly suppress all distractor sounds, but instead preferentially suppresses distractors near the locus of attention (Greenberg and Larkin 1968; Kidd et al. 2005).

A small number of studies in behaving animals have found that attention improves coding of task-relevant versus irrelevant features at the population level, observed through changes in multiunit and LFP synchrony (Lakatos et al. 2013) and interneuronal correlations (Downer et al. 2017). As with the human studies, this work did not distinguish between features near and far from the locus of attention, and it is not known if the same

mechanism operates in both cases. Sounds of similar frequency are encoded by topographically interspersed neurons in cortex (Bizley et al. 2005), and analysis at the level of single-neuron responses is important for understanding their representation.

Several studies have identified changes in single-neuron activity in primary auditory cortex (A1) during behavior. Neurons can enhance or suppress responses to task-relevant spectral or spatial sound features. However, most of this work has relied on comparisons between passive listening and behavioral engagement (Fritz et al. 2003; Atiani et al. 2009; Otazu et al. 2009; Lee and Middlebrooks 2011; David et al. 2012; Niwa et al. 2012; Kuchibhotla et al. 2016) or between behaviors with different structure (Fritz et al. 2005; David et al. 2012; Rodgers and DeWeese 2014). It is difficult to attribute changes in neural activity to selective attention because other aspects of internal state also change between conditions, including arousal, effort, rules of behavior, and associated rewards. Moreover, the specific acoustic task has differed between studies, ranging among tone detection, modulation detection, tone discrimination, and tone-in-noise detection. Changing task sound features could recruit different feedback systems or require different auditory areas with specialized coding properties (Tian et al. 2001; Bizley et al. 2005), making a comparison between these studies difficult.

To link specific aspects of behavioral state to changes in auditory coding, we developed an approach that isolates the effects of selective attention to sound frequency from general task engagement (Fritz et al. 2003; Otazu et al. 2009) and behavioral effort (Atiani et al. 2009). We trained animals to perform a tone-in-noise detection task in which they heard the same sound sequences but switched attention between target tones at different frequencies. To compare effects of attention and listening effort, we manipulated target salience while requiring detection of the same target tone (Atiani et al. 2009). We recorded single-unit activity in A1 during behavior. During manipulation of selective attention, we observed suppression specifically of responses to distractors near the target frequency rather than a generalized suppression of all distractors. When task difficulty was varied, we saw changes in tonic spike rate, rather than sound-evoked activity.

The noise stimuli developed for these behaviors contain natural temporal dynamics and were designed to permit characterization of the sensory filter properties of neurons reflected in their sound-evoked activity (David and Shamma 2013). We used these spectro-temporal receptive field (STRF) models to compare the encoding properties of A1 neurons between behavior conditions. This analysis revealed changes in neuronal filter properties consistent with changes in the average spontaneous and sound-evoked activity.

Materials and Methods

All procedures were approved by the Oregon Health and Science University Institutional Animal Care and Use Committee and conform to the National Institutes of Health standards.

Animal Preparation

Young adult male ferrets were obtained from an animal supplier (Marshall Farms). A sterile surgery was then performed under isoflurane anesthesia to mount a post for subsequent head fixation and to expose a small portion of the skull for access to auditory cortex. The head post was surrounded by dental acrylic or Charisma composite, which bonded to the skull and to a set of stainless steel screws embedded in the skull. Following surgery,

animals were treated with prophylactic antibiotics and analgesics under the supervision of University veterinary staff. The wound was cleaned and bandaged during a recovery period. After recovery (~2 weeks), animals were habituated to a head-fixed posture for about 2 weeks.

Auditory Selective Attention Task

Behavioral training and subsequent neurophysiological recording took place in a sound-attenuating chamber (Gretch-Ken) with a custom double-wall insert. Stimulus presentation and behavior were controlled by custom software (Matlab). Digital acoustic signals were transformed to analog (National Instruments), amplified (Crown), and delivered through 2 free-field speakers (Manger, 50-35 000 Hz flat gain) positioned ±30 degrees azimuth and 80 cm distant from the animal. Sound level was equalized and calibrated against a standard reference (Brüel & Kjær).

Three ferrets were trained to perform an auditory selective attention task modeled on studies in the visual system (Moran and Desimone 1985; McAdams and Maunsell 1999; David et al. 2008), in which they were rewarded for responding to target tones masked by 1 of 2 simultaneous, continuous noise streams and for ignoring catch tones masked by the other stream (Fig. 1). The task used a go/no-go paradigm, in which animals were required to refrain from licking a water spout during the noise until they heard the target tone (0.5 s duration, 0.1 s ramp) at a time randomly chosen from a set of delays (1, 1.5, 2, … or 5 s) after noise onset. To prevent timing strategies, the target time was distributed randomly with a flat hazard function (Heffner and Heffner 1995). Target times varied across presentations of the same noise distractors so that animals could not use features in the noise to predict target onset.

Noise streams were constructed from narrowband noise (0.25–0.5 octave, 65 dB peak SPL) modulated by the envelope of 1 of 30 distinct ferret vocalizations from a library of kit distress calls and adult play and aggression calls (David and Shamma 2013). The envelope fluctuated between 0 and 65 dB SPL, and its modulation power spectrum was low-pass with 30 dB attenuation at 10 Hz, typical of mammalian vocalizations (Singh and Theunissen 2003). Thus the spectral properties of the noise streams were simple and sparse, while the temporal properties matched those of ethological natural sounds. To maximize perceptual separability (Shamma et al. 2011), the streams were generated using different vocalization envelopes, centered at different frequencies (0.9–4.3 octave apart) and presented from different spatial locations (±30 degrees azimuth).

In a single block of behavioral trials, the target tone matched the center frequency and spatial position of 1 noise stream. It was switched to match the other stream in a subsequent block. The majority of trials (80–92%) contained a "cue tone" target with relatively high signal-to-noise ratio (SNR, 5 to −5 dB peak-to-peak, relative to reference noise). For all tones, SNR was calculated locally in period of the noise that overlapped the tone. This definition of SNR produced relatively stable performance at a given SNR. The remainder of trials contained "probe tone" targets with low SNR (−7 to −12 dB), requiring focused attention on the target stream. The exact target SNR was adjusted for each animal and target frequency so that the cue tone was super-threshold (>90% hit rate) and the probe tone was closer to threshold (70% hit rate). The large number of high-SNR cue targets provided a cue for attention and were important to maintain motivation, which flagged if animals were subjected to a large number of low-SNR probe targets. A random subset of trials (8–15%) included a "catch tone" in the nontarget stream, occurring before the target and identical to the probe
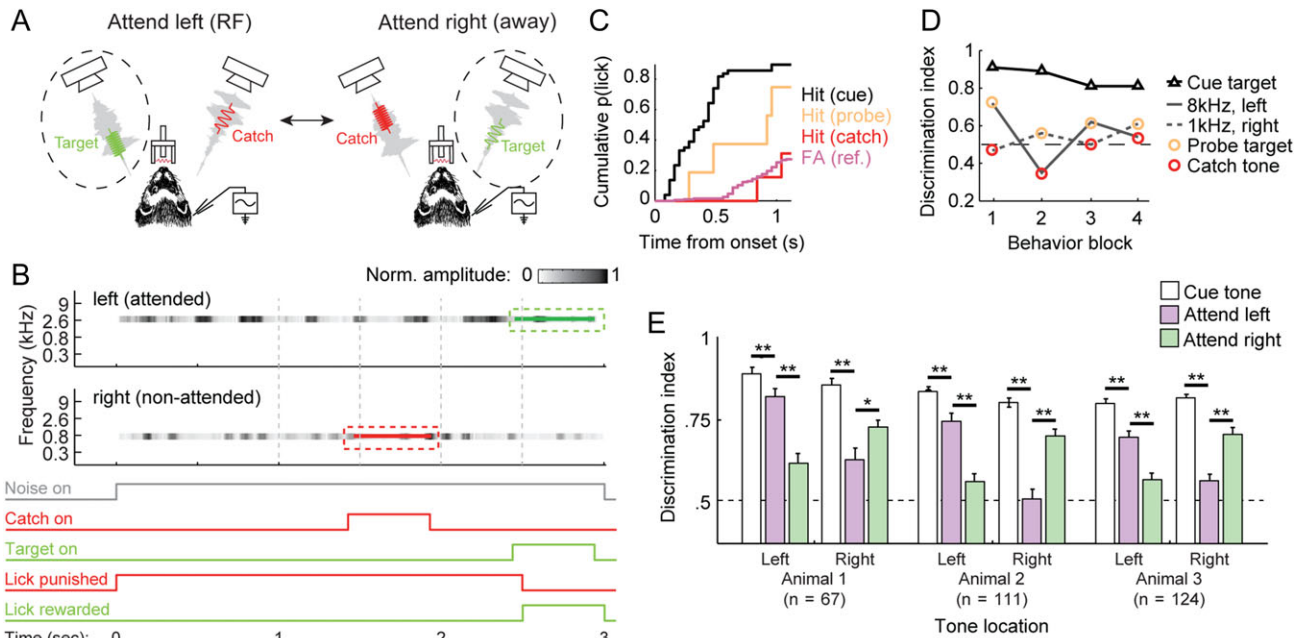
Figure 1. Selective attention behavior in ferrets. (*A*) Configuration of selective attention behavior. Head-fixed ferrets were presented with simultaneous, continuous noise streams from 2 locations 30 degrees left and right of midline. A target tone was presented at the center frequency and location of the attended noise stream 1–5 s after noise onset. Animals were rewarded for licking following target onset. Responses prior to target onset were punished with a short timeout. (*B*) Top shows spectrogram representation of stimuli during a single selective attention trial. Two narrowband (0.25 octave) noise streams with different center frequency were presented from different spatial locations (3000 Hz, left; 700 Hz, right). Noise level fluctuated from quiet (white) to loud (black) with dynamics drawn from natural vocalizations. The target tone (green) appeared at a random time, centered in the attended stream (in this case the high frequency stream). To control for selective attention, a catch tone (red) appeared in the nonattended stream on a minority (8–15%) of trials. Bottom shows the time course of noise, catch and target stimuli, as well as the time course of false alarm and hit windows during the same trial. (*C*) Cumulative probability of hit or false alarm following onset of each task stimulus during 1 behavioral block. False alarms were measured during windows when a target could occur (vertical lines in B). (*D*) DI was computed as the area under the ROC, calculated from lick probability curves in D for each target and catch tone during 4 sequential behavioral blocks. Between blocks, the low-SNR probe target alternated between 8 kHz/contra and 1 kHz/ipsi, and was reversed with the catch tone. DI was higher for the probe than the catch tone in all blocks, consistent with shifts in selective attention. Trial block order (attend RF versus attend away) was randomized during neurophysiology experiments to prevent any bias from changes in motivation, which could produce a gradual decrease in DI across blocks. (*E*) Bar plot shows average DI for each tone category (cue, probe, catch), broken down by animal and attended location, indicating the accuracy with which animals reported the presence of a tone versus the noise. DI was consistently highest for a tone with high signal-to-noise ratio (SNR, –4 to 5 dB peak relative to noise) used to cue animals to the attended location. DI was lower for low-SNR probe targets (–7 to –12 dB), but consistently greater than for low-SNR catch tones, indicating that animals consistently allocated attention to the target stream (**$P < 0.01$, sign test).

tone in the opposite trial block (–7 to –12 dB SNR). To avoid perceptual grouping of the target and catch tone, the interval between the 2 tones was jittered on each trial that contained both a target and catch. Responses to the cue and probe tones were rewarded with water (response window 0.1–1 s following tone onset). Responses to the catch tone or to the reference before the target resulted in no reward and were punished with a brief timeout (5–10 s) before the next trial. A preferential response to the probe versus catch tone indicated selective attention to the target stream (Moran and Desimone 1985; McAdams and Maunsell 1999; David et al. 2008). Trial blocks began by requiring correct behavior on 5 trials with only the cue target to direct attention to a single stream. Cue trials were discarded from subsequent analysis.

Behavioral performance was quantified by hit rate (correct responses to targets vs. misses), false alarm rate (incorrect responses prior to the target), and a discrimination index (DI) that measured the area under the receiver operating characteristic (ROC) curve for hits and false alarms (Fig. 1C–D, Yin et al. 2010; David et al. 2012). To compute DI, each time when a target could occur was identified in each trial. The first lick during each trial was treated as a hit (response following target onset) or false alarm (response to noise at a time when a target could occur), depending on whether it fell in a bin before or after tone onset. Trials with licks prior to any possible target window were punished as false alarms but classified as invalid and

excluded from behavioral analysis. The probability of a hit was computed as a function of the latency after tone onset, and the probability of a false alarm was computed for latency relative to times when targets could occur. These probabilities were then used to construct the ROC curve. A DI of 1.0 reflected perfect discriminability and 0.5 reflected chance performance.

Behavioral statistics were computed separately for the 3 tones (cue, probe, and catch). Target responses at the last possible time (5 s) were discarded, as catch responses and false alarms never occur at that latency. Only sessions with above-chance performance on cue tones (DI > 0.5, P > 0.05, jackknife t-test) were included in analysis of neurophysiology data. During about 30% of behavior blocks, animals' behavioral state lapsed at some point (indicated by missing 5–10 cue targets in a row, usually later in the day), and the experimenter provided a reward manually following a target tone to re-engage behavior. Typically, a single reminder trial was adequate. If animals failed to re-engage after multiple reminders (up to 10), the behavioral block was ended. These reminder trials were excluded from the analysis of neurophysiology data. Truncating all data acquired after reminder trials to control for long-term effects of the reminder increased noise in the analysis of neurophysiology data but did not significantly change any of the results observed across the neural population. Selective attention to the target stream was confirmed by larger DI for the probe tone than the catch tone across 2 behavioral blocks with reversed probe and catch tones.

As expected, animals were better able to report the cue tone, but also responded preferentially to the probe tone versus the catch tone (Fig. 1E). The combination of spatial and spectral streaming cues maximized behavioral attention effects. Animals were also able to perform tasks with only spectral separation between streams, but behavioral effects were weaker (data not shown). If only spatial cues were used, the 2 streams were fused, producing a strong percept of a single noise stream moving in space. This effect produced a spatial release from masking, increasing the salience of probe and catch tones and eliminating the need for selective attention. Animals were not tested on a spatial-only task because it appeared that attention would not be required for probe detection.

Variable Effort Task

Two animals were trained on a variant of the tone-in-noise task, in which the frequency and location of the target tone was fixed, but target SNR was varied between blocks. For this task either 1 or 2 noise streams were used (defined as above), but the target was always masked by the same stream. In the high-SNR condition, the target tone level was well above threshold on 80–90% of trials (10 to −5 dB SNR, measured by peak-to-peak amplitude). For a small number of trials (10–20%), a low-SNR probe target (0 to −15 dB SNR, 10 dB below high SNR) was used to measure behavioral sensitivity to a near-threshold target. Detection threshold varied according to the frequency of the target tone, and probe targets were chosen to be 10 dB above threshold. In the low-SNR condition, the majority of trials (80–90%) used low-SNR targets, and the remainder used high-SNR targets. While the exact level of high- and low-SNR targets varied across days, they were always fixed between blocks on a given day, so that performance could be compared between identical stimuli.

Behavioral performance was assessed using DI, as for the selective attention task. A change in effort was indicated by comparing DI for the low-SNR target between low-SNR and high-SNR conditions. Greater effort in the low-SNR condition was indicated by higher DI, reflecting adjustment of behavioral strategy to detect the difficult, low-SNR targets more reliably.

Neurophysiological Recording

After animals demonstrated reliable selective attention behavior (DI > 0.5 for at least 3 successive sessions), we opened a small craniotomy over primary auditory cortex (A1). Extracellular neurophysiological activity was recorded using 1–4 independently positioned tungsten microelectrodes (FHC). Amplified (AM Systems) and digitized (National Instruments) signals were stored using open-source data acquisition software (Englitz et al. 2013). Recording sites were confirmed as being in A1 based on tonotopy and relatively reliable and simple response properties (Shamma et al. 1993; Atiani et al. 2014). Single units were sorted offline by bandpass filtering the raw trace (300–6000 Hz) and then applying PCA-based clustering algorithm to spike-threshold events (David et al. 2009).

Upon unit isolation, a series of brief (100-ms duration, 100-ms interstimulus interval, 65 dB SPL) quarter-octave noise bursts was used to determine the range of frequencies that evoked a response and the best frequency (BF) that drove the strongest response. If a neuron did not respond to the noise bursts, the electrode was moved to a new recording depth. For the selective attention task, 1 noise stream was centered at the BF and at the preferred spatial location (usually contralateral).

The second stream was positioned 2 octaves above or below BF, usually outside of the tuning curve. Occasionally, neurons responded to noise bursts across the entire range of frequencies measured, and a band with a very weak response (<1/2 BF response) was used for the second stream. Thus, task conditions alternated between "attend RF" (target at BF and preferred spatial location) and "attend away" (target in the nonpreferred stream). For the variable effort task, the noise configuration was the same, but the target was always centered over neuronal BF.

The order of behavior conditions (attend RF and attend away blocks for selective attention, low- and high-SNR blocks for variable effort) was randomized between experiments to offset possible bias from decreased motivation during later blocks (note gradual decrease in DI across blocks in Fig. 1D). We recorded neural activity during both behavior conditions and during passive presentation of task stimuli pre- and post-behavior. Thirty identical streams, with frozen noise carriers were played in all behavior conditions. Of the 30 noise streams, 29 were repeated 1–2 times in each behavior block. One stream was presented 3–10 times, permitting a more reliable estimate of the PSTH response during neurophysiological recordings. Noise stimuli presented on incorrect (miss or false alarm) trials were repeated on a later trial, and a repetition was complete only when all the noise stimuli were presented on correct (hit) trials. Data from the repeated stimuli were used as the validation set to evaluate encoding model prediction accuracy (see STRF analysis, below).

Evoked Activity Analysis

Peri-stimulus time histograms (PSTHs) of spiking activity were measured in each behavioral condition, aligned to reference and target stimulus onsets (Fig. 2). Because target and catch tones were embedded in the reference sound at random times, reference responses were truncated at the time of tone onset. We compared the PSTH during 3 epochs (Fig. 3): spontaneous (0–500 ms before reference onset), reference-evoked (0–2000 ms after reference onset, minus spontaneous) and target-evoked (0–400 ms after target onset minus 0–500 ms before target onset). A longer target window did not affect changes measured in the target response, but the short target window minimized potential confounds from motor signals associated with licking, which typically had latency >400 ms for probe tones.

To measure changes in mean spontaneous and evoked activity for single neurons, we measured a "behavior modulation index" (BMI), the fraction change (Otazu et al. 2009) in spontaneous activity and evoked responses between behavior conditions:

$$d = \frac{\bar{r}_A - \bar{r}_B}{\bar{r}_A + \bar{r}_B} \tag{1}$$

The subscripts A versus B refer to experimentally controlled behavioral conditions (e.g., attend RF vs. attend away, active vs. passive, low SNR vs. high SNR). Significant differences between behavior conditions were assessed by a jackknife t-test, and significant average changes across a neural population were assessed by a Wilcoxon sign test. We compared results of the sign test for population data to a jackknife t-test, and found similar results.

To measure changes in baseline rate and response gain, we modeled the time-varying response to a stimulus in behavior condition A, $r_A(t)$, as the response in condition B, $r_B(t)$, scaled by
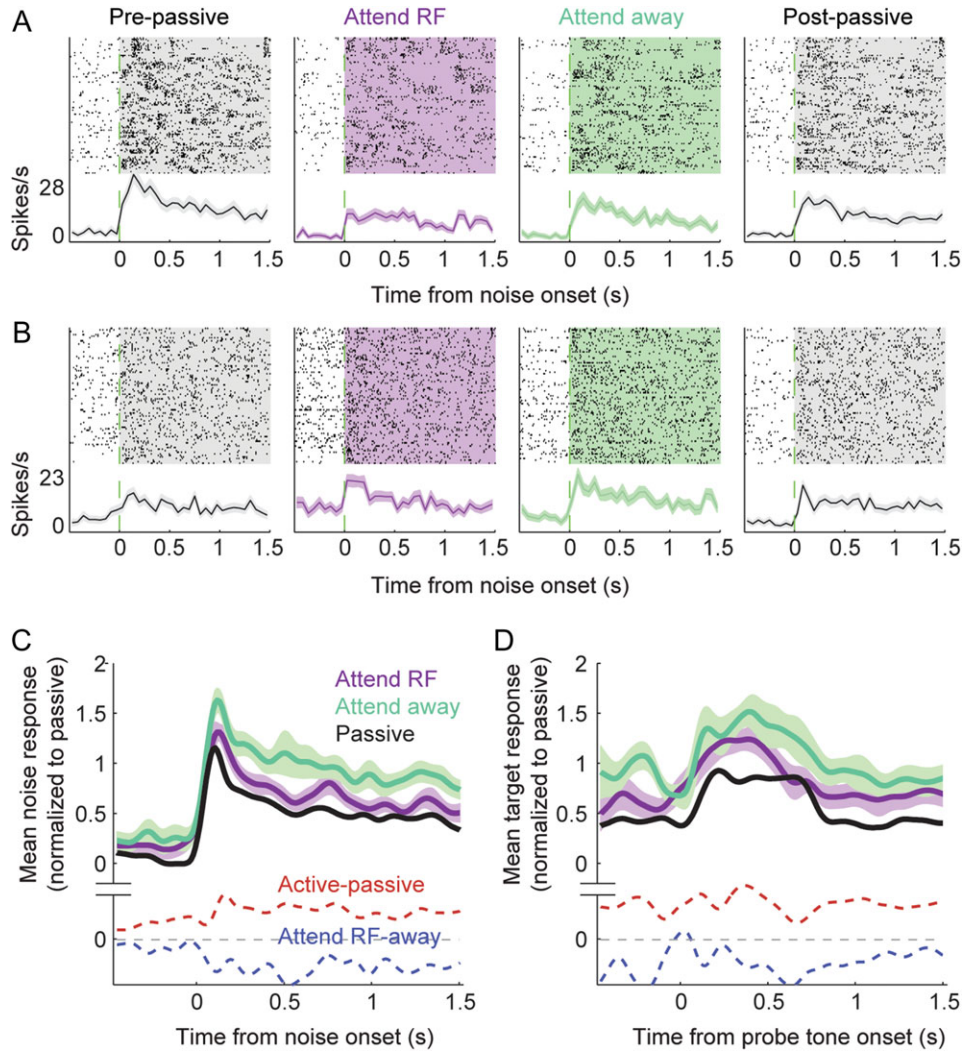
Figure 2. Sound-evoked activity during selective attention behavior. (A) Comparison of spike raster plots and average PSTH responses to reference noise for 1 neuron, aligned to trial onset in different behavior conditions. Evoked activity was suppressed when the animal engaged in the task (passive vs. attend RF or attend away). When attention was directed toward the noise stream that fell in the neuron's receptive field (attend RF) responses were further suppressed, relative to attention directed to the stream out of the RF (attend away). BMI (Eq. 1) for attend RF versus away is −0.17. (B) Responses of a second neuron are enhanced during task engagement, but again show relative suppression in the attend RF versus attend away condition (BMI −0.24). (C) Average PSTH response to distracter noise across significantly modulated units (34/54 units) in the different behavior conditions (P < 0.05 spontaneous or noise response change, Bonferroni corrected jackknife t-test). Dashed lines show the average PSTH difference between active and passive conditions (red) and between attend RF and attend away (blue). Spontaneous rate increases during task engagement; reference-evoked activity is suppressed for attend RF versus attend away conditions. (D) Average PSTH response aligned to probe target onset in the 3 conditions reflects pre-target differences in noise-evoked activity, but the magnitude of the target-evoked response does not change.

a constant gain, g, plus a constant offset, d (Slee and David 2015),

$$r_A(t) = g[r_B(t) - r_{B,0}] + r_{B,0} + d \qquad (2)$$

The gain term was applied after subtracting the spontaneous rate, $r_{B,0}$, so that it impacted only sound-evoked activity. We used least-square linear regression to determine the optimal values of d and g that minimized mean squared error over time:

$$d, g = argmin \sum_t [r_A(t) - g[r_B(t) - r_{B,0}] - r_{B,0} - d]^2 \qquad (3)$$

Identical noise stimuli were presented in each behavioral condition, but variability in performance did not always permit presentation of a complete stimulus set in all behavioral

conditions. To control for any possible difference in sound-evoked activity, these analyses were always applied to the subset of data with identical stimuli in both behavior conditions for a given experiment. For comparison of average PSTH (Figs 2C, 6C), responses were normalized by mean noise-evoked activity in the passive condition before averaging.

We used signal detection analysis to measure neural discriminability (d') of task-relevant sounds, based on single-trial responses to tone and reference noise stimuli (Niwa et al. 2012),

$$d' = \frac{\bar{r}_T}{\sigma_T} - \frac{\bar{r}_N}{\sigma_N} \qquad (4)$$

where $\bar{r}_T$ and $\bar{r}_N$ were the average response to tone and noise stimuli, respectively, and σ_T and σ_N were the standard deviation of responses across trials. For the selective attention data,

d' was computed for the probe tone in the attend RF condition and compared to *d'* for the catch tone in the attend away condition. For the variable SNR data, *d'* was compared between the low-SNR target in the low-effort (high-SNR) and high-effort (low-SNR) blocks. Neural responses to probe and catch tones were computed as the mean spike rate during 0–400 ms following tone onset. This relatively short window avoided possible motor artifacts from licking during behavior. Reponses to noise were computed during 400 ms windows prior to probe or catch, during which time a tone could occur on a different trial.

Only data from correct trials were analyzed, although no significant differences were observed for data from incorrect trials. The majority of incorrect trials were false alarms with relatively short duration. Because stimuli were terminated immediately after a false alarm, their relative contribution to the overall data set size was limited. Probe and catch tones were relatively rare during selective attention behavior (8–20% of trials). This limitation, combined with their low SNR, made it difficult to measure reliable probe responses in some neurons. Only 43/54 neurons with at least 5 presentations of probe and target tones during behavior were included in the *d'* analysis. In addition, during early experiments, probe tones were not presented during passive listening (7/43 neurons), and these were also excluded from the target behavioral modulation index (BMI) analysis (Fig. 3C).

STRF Analysis

Vocalization-modulated noise was designed so that the random fluctuations in the 2 spectral channels (noise stream envelopes) could be used to measure spectro-temporal encoding properties. The linear STRF is a current standard model for early stages of auditory processing (Aertsen and Johannesma 1981; Theunissen et al. 2001; Machens et al. 2004; David et al. 2009; Calabrese et al. 2011). The linear STRF is an implementation of the generalized linear model (GLM) and describes time-varying neural spike rate, *r(t)*, as a weighted sum of the preceding stimulus spectrogram, *s(x,t)*, plus a baseline spike rate, *b*,

$$r(t) = \sum_{x,u} h(x, u)s(x, t - u) + b \qquad (5)$$

The time lag of temporal integration, *u*, ranged from 0 to 150 ms. In typical STRF analysis, the stimulus is broadband and variable across multiple spectral or spatial channels, *x*. Here, the stimulus is composed of just 2 time-varying channels, and a spectrally simplified version of the STRF can be constructed in which *x* = 1…2 spans just these 2 channels. The encoding model for a single spectral band is referred to simply as a temporal receptive field (TRF, Ding and Simon 2012; David and Shamma 2013), but because the current study included 2 bands, we continue to refer to these models as STRFs. Analytically, this simplified STRF can be estimated using the same methods as for standard STRFs. For the current study, we used coordinate descent, which has proven effective for models estimated using natural stimuli (David et al. 2007; Willmore et al. 2010; Thorson et al. 2015). Spike rate data and stimulus spectrograms were binned at 10 ms before STRF analysis.

The ability of the STRF to describe a neuron's function was assessed by measuring the accuracy with which it predicted time-varying activity in a held-out validation data set that was not used for model estimation. The "prediction correlation" was computed as the correlation coefficient (Pearson's *R*) between the predicted and actual average response. A prediction correlation of *R* = 1 indicated perfect prediction accuracy, and a value of *R* = 0 indicated chance performance.

To measure effects of behavioral state on spectro-temporal coding, we estimated "behavior-dependent STRFs", by estimating a separate STRF for data from each behavioral condition (attend BF vs. attend away or high-SNR vs. low-SNR). Thus 2 sets of model parameters were estimated for each neuron, for example, in the case of selective attention data, filters $h_{RF}(x,u)$ and $h_{away}(x,u)$, and baseline rates b_{RF} and b_{away}. Predication accuracy was assessed using a validation set drawn from both behavior conditions, using each STRF to predict activity in their respective behavioral state. Significant behavioral effects were indicated by improved prediction correlation for behavior-dependent STRFs over a "behavior-independent STRF" estimated using data collapsed across behavior conditions. Behavior-dependent changes in tuning were measured by comparing STRF parameters, *h(x,u)* and *b*, between conditions.

Competing behavior-dependent and -independent models were fit and tested using the same estimation and validation data sets. Significant differences in prediction accuracy across the neural population were determined by a Wilcoxon sign test.

Results

Ferrets can Selectively Attend Among Competing Auditory Streams

We trained 3 ferrets to perform an auditory selective attention task requiring detection of a tone in one of 2 simultaneous noise streams. Task stimuli were composed of 2 simultaneous tone-in-noise streams (Fig. 1A–B). During a single block of trials, animals were rewarded for responding to tones masked by one stream (the "target stream") and punished for responding to the other (the "nontarget"). Reward contingencies were reversed between blocks. The task therefore allowed comparison of neural responses to identical sensory stimuli while animals reported targets occurring in only 1 of the 2 streams (Moran and Desimone 1985).

Spectral, spatial, and temporal cues were used to maximize perceptual separability of the streams. Both streams consisted of tones embedded in narrowband noise (0.25–0.5 octave) modulated by the temporal envelope of natural vocalizations (David and Shamma 2013). To facilitate perceptual segregation, the streams differed in center frequency (0.9–4.3 octave separation), location (±30 degrees azimuth) and temporal envelope dynamics (1 of 30 envelopes from a vocalization library, chosen randomly on each trial).

The task employed a go/no-go paradigm (David et al. 2012; Slee and David 2015). Ferrets initiated each trial by refraining from licking a water spout for a random period (1–3 s). They were then simultaneously presented with the target and nontarget noise streams (Fig. 1B). On each trial, a target tone at the center frequency and spatial location of the target stream was presented at a random time after noise onset (1–4 s). The majority (80–92%) of trials contained a high signal-to-noise ratio (SNR, 5 to −4 dB) *cue tone,* while the remaining trials contained a low-SNR *probe tone* (−7 to −12 dB). On a random subset of trials (8–15%), the target was preceded by a low-SNR "catch tone" at the frequency and location of the nontarget stream, with SNR matched to the probe tone. The specific SNR was manipulated between experiments to produce near-threshold behavior, but probe and catch tones were presented at identical SNRs within a single experiment. Licks that occurred before

target tone onset (including catch tone periods) resulted in termination of the trial and punishment with a brief timeout (5–10 s). Licks that promptly followed target tone onset were rewarded with water (0.1–1 s following target onset).

A behavioral session consisted of 2 blocks, with the attended stream switching between blocks. The frequency, location, and level of the noise bands and the probe and catch tones remained the same across blocks. Only the reward contingencies, indicated by the frequency of the cue tone, were reversed: the target stream became the nontarget, and the behavioral meaning of the probe and catch tones was reversed. The task therefore allowed comparison between responses to identical noise and tone stimuli under different internal states, a key requirement of a selective attention task (Moran and Desimone 1985; McAdams and Maunsell 1999).

We verified selective attention by comparing behavioral responses to the probe and catch tones. In humans, knowledge of the location or frequency of an attended sound affects the speed and accuracy with which the sound is detected relative to distractors (Greenberg and Larkin 1968; Scharf et al. 1987; Dai et al. 1991; Spence and Driver 1994; Woods et al. 2001; Kidd et al. 2005). We expected a similar improvement in behavioral discriminability for the target over catch tone. To test this prediction, we calculated a DI, which quantified the area under the ROC curve for discrimination of each tone class from the narrowband noise background (Fig. 1C–D, Yin et al. 2010; David et al. 2012). A DI of 1.0 reflected perfect discriminability and 0.5 reflected chance performance. As expected, animals were able to report the cue tone more reliably than the low-SNR tones, but they also responded preferentially to the probe tone versus the catch tone. Mean performance of all 3 ferrets showed significantly greater DI for the probe tone versus catch tone and no difference between attention to left and right streams (Fig. 1E).

The order of behavioral conditions was randomized across days, but we considered the possibility that increased satiety might lead to a decrease in DI over the course of behavior during a single day. We did observe a decrease in overall DI between the first and second behavioral block for one animal (mean DI, animal 1:86.4 vs. 82.7**, animal 2:78.0 vs. 78.3, animal 3:76.1 vs. 74.8, **$P < 0.01$, jackknife t-test, $p > 0.05$), but the random ordering of behavioral blocks controlled for this trend. We also considered the possibility of changes in DI over the course of a single behavioral block. One animal showed a trend toward decreased DI in the second half of each block (animal 1:86.7 vs. 83.2), while 2 showed a trend toward increased DI (animal 2:77.1 vs. 80.6, animal 3:74.6 vs. 76.7). However, none of these within-block changes was significant (jackknife t-test, $p > 0.05$), and performance was broadly stable over time.

A1 Single-Unit Responses Are Selectivity Suppressed for the Attended Distractor Stream

We recorded single-unit activity in primary auditory cortex (A1) of the 3 trained ferrets during selective attention behavior to determine how neural activity changed as attention was switched between noise streams. For each unit, one noise stream was centered over its BF and was presented from a location contralateral to the recording site. The other stream was presented in a frequency band outside the neuron's frequency tuning curve (0.9–4.3 octaves from BF) and ipsilateral to the recording site. Thus, the task alternated between an "attend receptive field (RF)" condition (target stream in the RF) and "attend away" condition (target stream outside the RF). We

recorded activity during both behavior conditions and during passive presentation of task stimuli ($n = 54$ neurons). The order of attend RF and attend away blocks was randomized across experiments to avoid bias from changes in overall motivation over the course of the experiment.

Changes in behavioral state can influence spontaneous spike rate and/or the sound-evoked activity of single-units (Ryan et al. 1984; Rodgers and DeWeese 2014). We measured the effects of selective attention by comparing spontaneous and noise-evoked activity (after subtracting spontaneous rate) between attend RF and attend away conditions. We measured the effects of task engagement by comparing activity between behaving and passive conditions. Because identical noise stimuli were played in all behavioral conditions, this controlled for any difference in sound-evoked responses. We measured noise-evoked responses only during the periods prior to the occurrence of target and catch tones.

When animals attended to a noise stream in a neuron's RF, the response of many neurons to the noise stimuli decreased (Fig. 2A–B). Of the 54 neurons with data from both attend-RF and attend away conditions, 16 showed significant changes in baseline rate and 26 showed significant changes in noise-evoked response ($P < 0.05$, jackknife t-test). The average PSTH response computed from the activity of neurons that underwent a change in either spontaneous or evoked activity followed a pattern consistent with the examples, showing no consistent change in spontaneous rate but a decreased evoked response (Fig. 2C). The change in evoked activity was roughly constant over time, occurring with about the same latency as the sensory response itself.

We quantified behavior-dependent changes in neural activity using a BMI (Eq. 1), computed as the ratio of the difference in neural activity between conditions to the sum of activity across conditions (Otazu et al. 2009). For selective attention comparisons, BMI was calculated as the difference between attend RF and attend away. Thus, BMI greater than zero indicated greater neural activity in the attend RF condition, and negative BMI indicated greater activity in the attend away condition. A value of 1 or –1 indicated complete suppression of responses in the attend away or attend RF condition, respectively. For the 16 neurons showing significant changes in spontaneous activity, median BMI was not significantly different from zero ($P > 0.5$, sign test, Fig. 3A,D, shaded bars). However, for the 26 neurons showing significant changes in noise-evoked response, BMI was significantly decreased in the attend RF condition (–0.12, $P = 0.005$, approximately 20% suppression, Fig. 3B, E). A decrease in the noise-evoked response was also observed across the entire selective attention data set (median BMI –0.06, $P < 0.001$, $n = 54$). Because the stimuli were designed so that only the RF stream evoked a neural response, this change reflected suppression of responses to the noise stream that masked the attended target. In addition to measuring changes in the mean evoked response, we also compared changes in the gain of noise-evoked responses (Slee and David 2015), which showed a similar suppression in the attend RF condition (median gain change –14%, $P < 0.001$, Fig. 4).

There was no systematic effect of selective attention on responses to the probe tone (target during attend RF vs. catch during attend away), indicating that suppression was selective for the masking noise in the attended stream (Fig. 2D, 3C). Across the entire set of recordings with a sufficient number of probe and catch tones to measure responses in both conditions ($n = 36$ neurons with at least 5 presentations of probe and catch tones in passive and both active conditions), the tone response
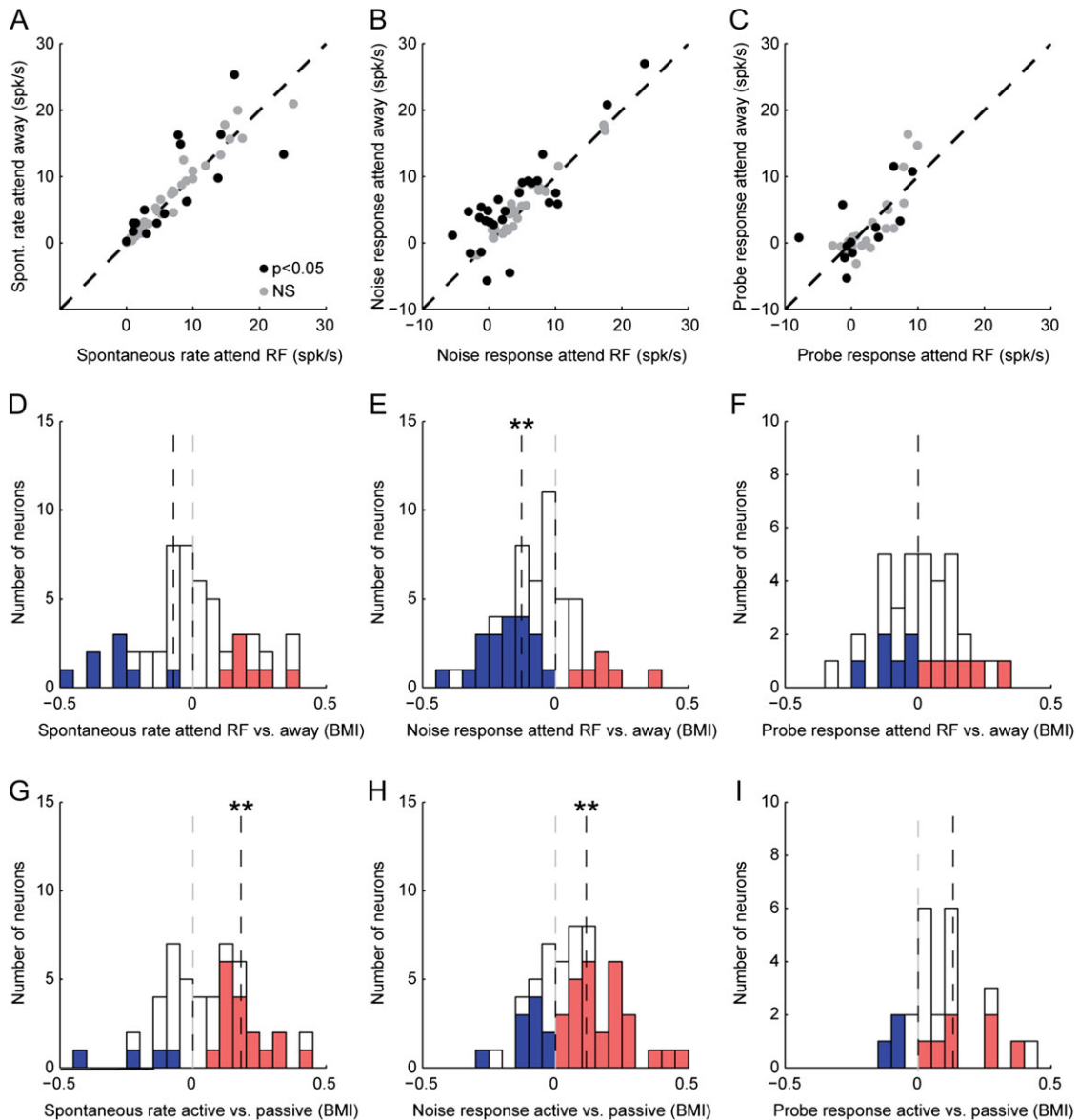
Figure 3. Summary of selective attention and task engagement effects on A1 evoked activity. (A–C) Scatter plot of average spontaneous rate (A), noise-evoked response (spontaneous rate subtracted, B), and target-evoked response (noise-evoked response subtracted, C) for each A1 neuron between attend RF and attend away conditions. Black circles indicate neurons with a significant difference between attention conditions for any of the 3 statistics (P < 0.05, jackknife t-test with Bonferroni correction). Target data are shown for the subset of neurons with adequate presentations of probe and catch tones (n = 36/54). (D–F) Histograms show BMI (Eq. 1), reflecting fractional change in spontaneous rate, noise-evoked response, and target-evoked response between attention conditions for each neuron. Neurons with a significant increase are indicated in red, and those with a significant decrease in blue (P < 0.05, jackknife t-test). Median BMI was significant only for noise-evoked responses; spontaneous rate: −0.14 (P = 0.38, n = 16/54, sign test), noise-evoked response: −0.12 (**P = 0.005, n = 26/54), target-evoked response: 0 (P > 0.5, n = 12/36). (G–I) Histograms comparing changes between active and passive conditions for the same neurons, plotted as in D–F. For the active condition, data were combined from the 2 attention conditions. In all comparisons, responses to an identical set of stimuli were used to compute the difference between behavior conditions. Median BMI increased for spontaneous rate and noise-evoked response: spontaneous rate: 0.17 (**P = 0.002, n = 27/54), noise-evoked response: 0.11 (**P = 0.005, n = 38/54), target-evoked response: 0.14 (P = 0.09, n = 13/36).

changed significantly between attention conditions in 12 neurons, but median BMI was not significantly different from zero (P > 0.5, Fig. 3F).

These attention-related effects were distinct from changes in neural activity related to task engagement. When compared with passive listening, the median spontaneous spike rate and noise-evoked response both increased significantly (Fig. 3G–I, spontaneous: median BMI 0.17, P = 0.002; noise-evoked: median 0.11, P = 0.005; probe-evoked: median 0.14, P = 0.09; sign test). Because the magnitude of spontaneous rate changes was the

same for both attention conditions, there was no difference between attend RF and attend away conditions (Fig. 5A). Thus, the suppression of noise-evoked activity between selective attention conditions contrasted with a general increase in spontaneous activity and excitability during task engagement.

The effects of selective attention and task engagement varied substantially across neurons. To better understand the interplay of these effects across the population, we compared the BMI for the noise-evoked response between behavior (averaged across attention conditions) and passive listening to the
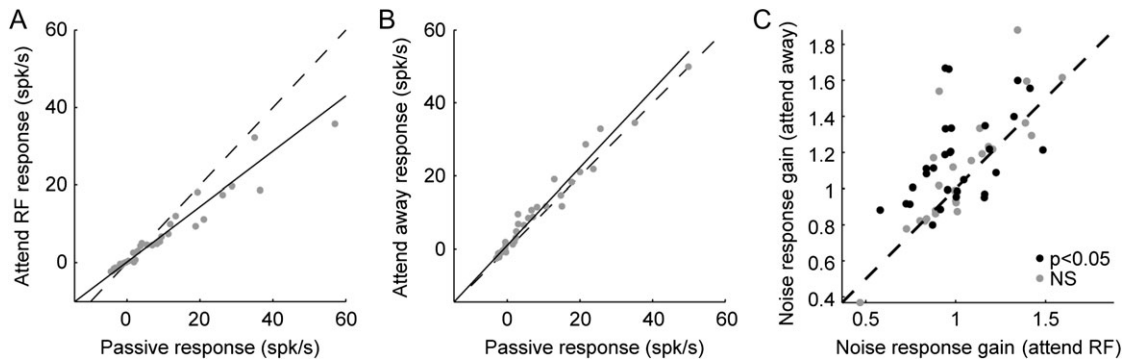
Figure 4. Effect of selective attention on A1 response gain. (*A*) Example of noise-evoked gain changes between the passive and attend RF conditions (same unit as Fig. 2*A*). Scatter plot compares responses to identical epochs of the noise stimulus between behavior conditions. The majority of points fall below the line of unity slope (dashed) indicating a general suppression when attention was directed toward a target in the neuron's RF. The slope of the line fit to the scatter plot (solid) is 0.76, indicating suppression of 24% during the attend RF condition. (*B*) Comparison between passive and the attend away condition for the same neuron reveals a slight gain enhancement (slope 1.04). (*C*) Each point compares noise-evoked response gain for a single neuron between attend RF and attend away conditions, computed relative to passive listening. Across all neurons, mean gain was higher during attend away than attend RF (1.11 vs. 0.99, P < 0.001, sign test). For neurons showing significant changes in mean response (black points, see Fig. 3*E*), mean gain was also different (1.11 vs. 0.97, P < 0.001).
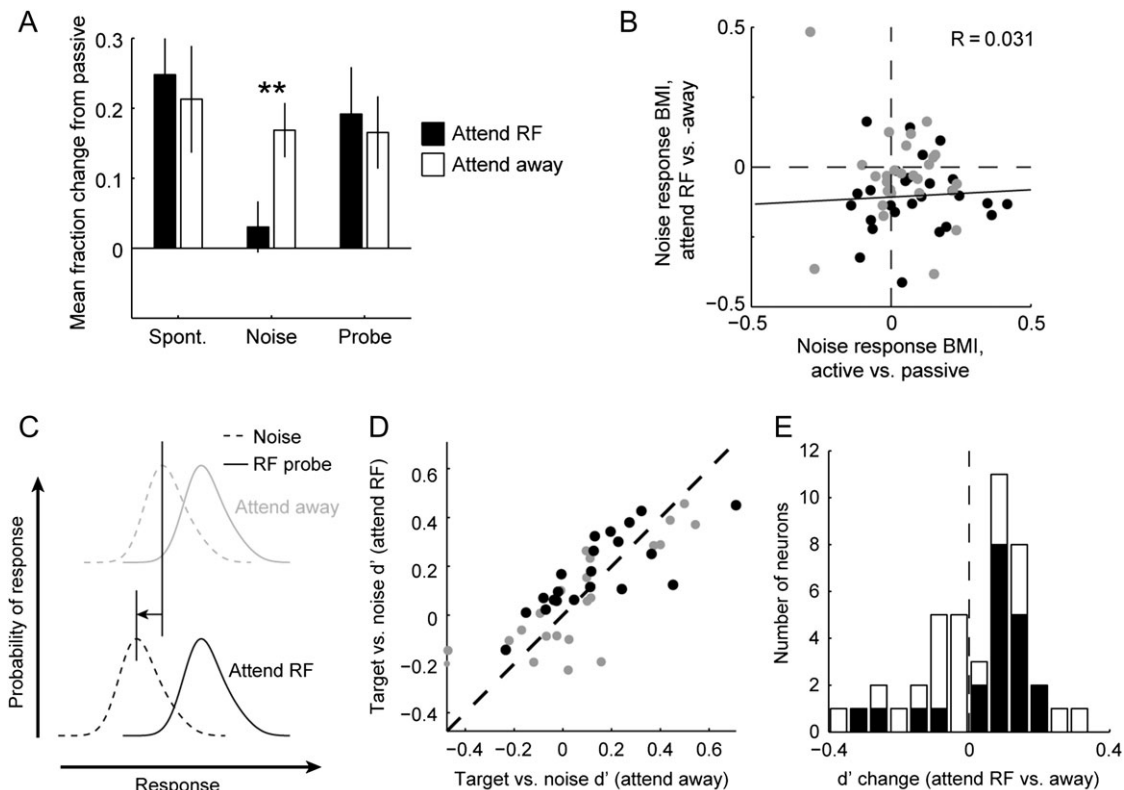

Figure 5. Effect of selective attention on A1 neural discriminability. (*A*) Mean fraction change in spontaneous rate, noise-evoked response and target-evoked response between the passive listening and attend RF (black) or attend away condition (white), for neurons showing significant effects of attention (*n* = 34, Fig. 3*D*–*F*). Error bars indicate one standard error (**P = 0.005, sign test). (*B*) Scatter plot compares BMI for the noise-evoked response of each neuron between active and passive conditions and between attend RF and attend away conditions. Most neurons fall in the lower right quadrant (increased response during behavior and decreased response during attend RF), but there is no correlation between the magnitude of the effects across cells (*r* = 0.031, P > 0.5, permutation test). (*C*) Model for enhanced discriminability following suppression of noise response at attended location. Curves represent the distribution of single-trial responses to a tone at the neuronal BF embedded in noise (solid line) and to noise alone (dashed lines). If noise responses are suppressed when attention is directed to BF, then the distributions become more discriminable, as measured by *d'*. (*D*) Scatter plot compares *d'* for the discriminability of neural responses to probe or catch tones (the same stimulus, respectively, in attend RF or attend away conditions) and the reference noise during time windows when the target could occur. Only units with at least 5 probe and catch tone presentations were included in order to obtain stable *d'* measures (*n* = 43). Filled circles indicate experimental sessions when DI for probe tones was significantly greater than for catch tones (P < 0.05, *n* = 21/43, permutation test). (*E*) Histogram of change in *d'* between attend RF and attend away conditions. For sessions in which DI was significantly greater for probe targets (filled bars), mean *d'* increased (0.10, P = 0.02, permutation test). The mean change for all sessions was smaller (0.03, P = 0.02).

BMI between attention conditions (Fig. 5B). We found no correlation between these effects (r = 0.031, P > 0.5, jackknife t-test), consistent with a system in which effects of selective attention and task engagement operate on different subsets of A1 neurons. We also considered the possibility that the magnitude of behavior effects might depend on tuning bandwidth or the recording depth in cortex. However, no significant relationship was observed (data not shown).

Distractor Suppression Enhances Neural Discriminability of Target Versus Distractors

We hypothesized that if responses to the distractor noise were suppressed relative to the probe tone when attention was directed to the neuronal BF, then the neural response to the noise versus tone-plus-noise stimuli should be more discriminable when attention was directed to the RF (Fig. 5C). To assess discriminability of these responses, we computed a neurometric d' for discrimination of the RF (probe/catch) tone from distractor noise, based on single-trial neural responses during attend RF versus away conditions (Fig. 5D–E). Previous work has used this approach to measure improvements in neural discriminability following engagement in a temporal modulation detection task (Niwa et al. 2012). Across all behavioral sessions, neural d' increased slightly between the 2 attention conditions (median change 0.03, P = 0.02, sign test). The

selective attention task was difficult for animals to perform during neurophysiological recordings, and behavioral performance varied between experiments. If we considered only sessions in which animals showed a significantly greater DI for the probe target than for the catch tone (P < 0.05, jackknife t-test), this subset of neurons showed a greater increase in d' in the attend RF condition (median change 0.1, P = 0.02, sign test, Fig. 5E). This larger change in d' suggests that selective attention improves neural discriminability in A1 during the tone-in-noise detection task, and the degree of improvement depends on the animal's performance.

A1 Spontaneous Activity, But Not Response Gain, is Consistently Modulated by Changes in Task Difficulty

Previous work has suggested that behavioral effort, driven by changing task difficulty, can also modulate neuronal activity in A1 (Atiani et al. 2009) and other sensory cortical areas (Chen et al. 2008). To assess the effects of changing effort in the current tone-in-noise context, we recorded A1 single-unit activity during a variant of the tone-detection task. In this case, the target stream was fixed, but the task varied from easy to hard by changing the SNR of the target tone relative to the noise stream from high (10 to −5 dB) to low (0 to −15 dB) between behavioral blocks (Fig. 6A). To probe differences in effort associated with changes in difficulty, the high-SNR condition included a small number of
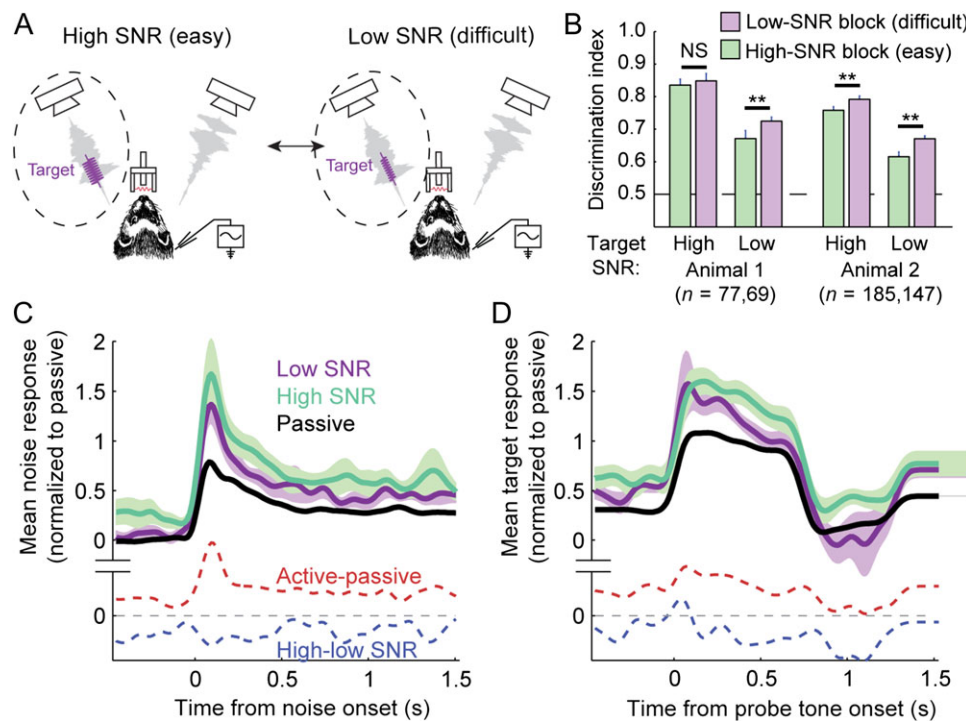

Figure 6. Effects of varying target SNR on behavioral effort and A1 evoked activity. (A) Configuration of variable SNR behavior for manipulating listening effort. Ferrets were presented with vocalization-modulated noise streams, configured as for the selective attention task. A target tone appeared at a random time, always in a stream contralateral to the recording site and centered over BF of the recorded neuron. In high-SNR blocks, the target tone was presented at high SNR (10 to −5 dB relative to the noise) for 80–90% of trials and at low SNR (10 dB lower) for the remaining trials. Conversely, in low-SNR blocks, the target tone had low SNR for 80–90% of trials. Tone levels were adjusted between experiments to span super- and near-threshold behavior, which varied across target frequency and animals. SNR values were always identical during all recordings from a single neuron. (B) To test for changes in effort, behavioral performance was compared for identical target tones between behavior conditions. Bar chart compares mean DI for high- and low-SNR targets, broken down between high- and low-SNR blocks and between animals. DI was significantly greater for the low-SNR target in low-SNR blocks when the more difficult target was likely to occur, consistent with an increase in effort (**P < 0.01, permutation test). DI did not improve for high SNR targets for Animal 1, possibly due to a ceiling effect in performance for the easy target. (C–D) Average population PSTH for noise-evoked (C) and high-SNR target-evoked responses (D) compared between passive listening, high-SNR and low-SNR behavior conditions, for all neurons showing any significant difference between SNR conditions (n = 65/88, P < 0.05, jackknife t-test, Bonferroni correction).

low-SNR targets (10–20%, Spitzer et al. 1985). We could then assess listening effort by comparing performance for identical low-SNR targets between the 2 conditions. In both animals, DI was consistently higher when the low-SNR target was more likely to occur, consistent with greater effort during blocks when target detection was more difficult (Fig. 6B Spitzer et al. 1985).

We recorded from 88 neurons from 2 ferrets in both effort conditions and from 122 neurons in at least one effort condition. Mean PSTHs showed an increase in spontaneous activity during low- versus high-SNR blocks but no change in the average response evoked by noise or targets (Fig. 6C–D). The

spontaneous rate of 47/88 units was significantly modulated by task difficulty ($P < 0.05$, jackknife t-test). Among these cells, the spontaneous rate showed a significant decrease during low-SNR blocks (median BMI −0.16, $P = 0.008$, sign test, Fig. 7A). Many neurons also showed an effect of task difficulty on the evoked responses to noise ($n = 51$) or tones ($n = 35$). However, there was no consistent average change in either sound-evoked response (Fig. 7B–C, $P = 0.4$ and 0.09, respectively, for noise and tone responses). As in the case of selective attention, we observed increases in spontaneous activity and sound-evoked responses during behavior relative to passive listening
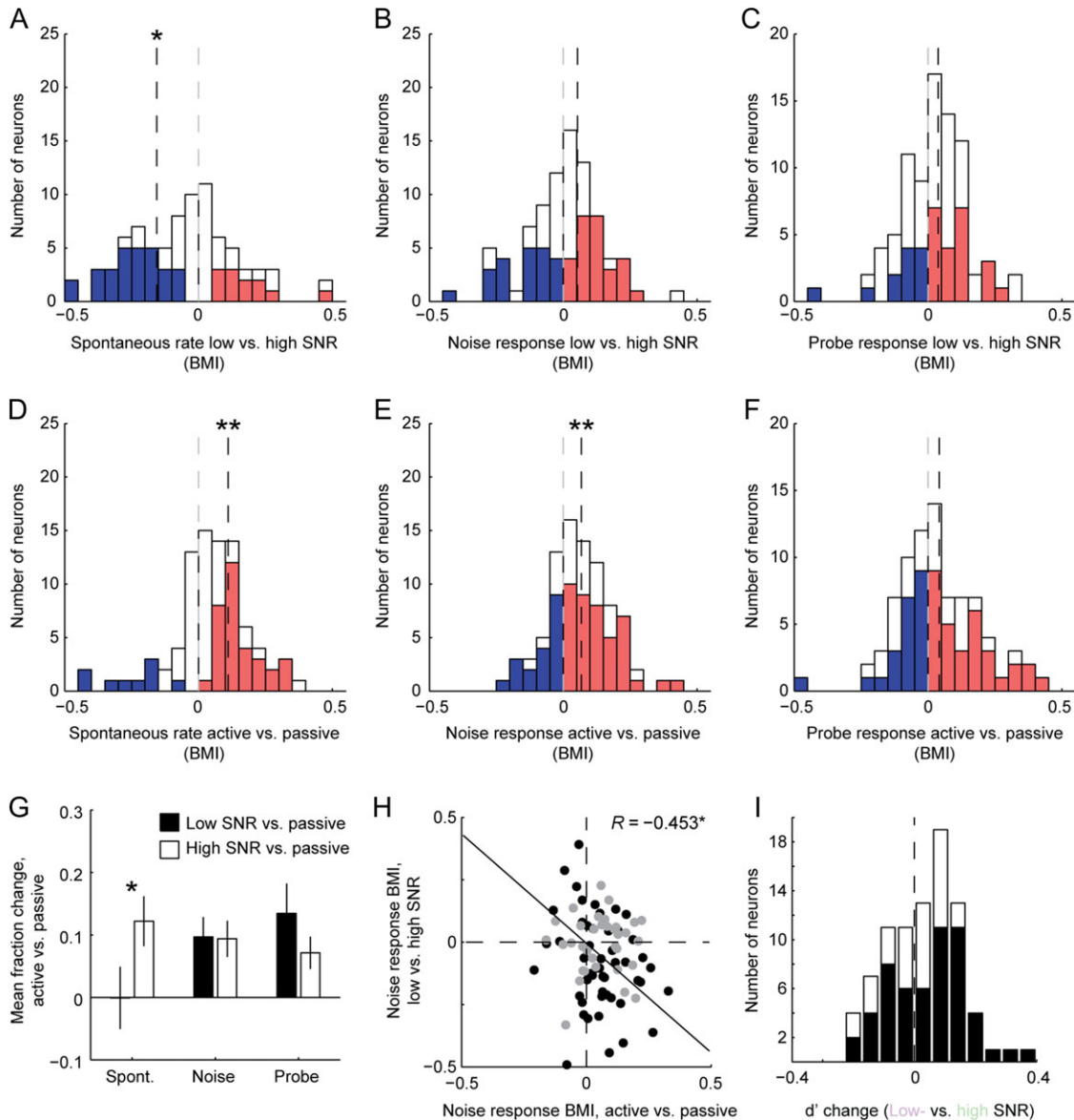

Figure 7. Summary of variable SNR effects on A1 evoked responses. (A–C) Histograms of BMI for spontaneous rate (A), noise-evoked response (B) and target-evoked response (C) between high-SNR and low-SNR blocks, plotted as in Figure 3D–F. Median spontaneous rate is suppressed (−0.16, $P = 0.008$, $n = 47/88$ significantly modulated neurons, sign test), but there is no change in noise-evoked (0.05, $P = 0.4$, $n = 51$) or target-evoked responses (0.04, $P = 0.09$, $n = 35$) D–F. Histograms comparing active versus passive BMI for the variable SNR task. High- and low-SNR data were combined for the active condition. Task engagement leads to increased spontaneous activity (0.11, $P < 0.0001$, $n = 43$) and noise-evoked responses (0.06, $P = 0.005$, $n = 63$) but no change in target-evoked response (0.04, $P = 0.11$, $n = 57$). (G) Bar chart shows mean fraction change in activity between passive listening and low-SNR (black) and high-SNR conditions (white), for neurons showing significant change in the corresponding statistic in A–C. (H) Scatter plot compares change in noise-evoked response between active and passive conditions (horizontal axis) and the change in spontaneous rate between high- and low-SNR behaving conditions ($r = −0.53$, $P < 0.001$, permutation test). (I) Histogram comparing the change in d' between low SNR and high SNR conditions. The mean d' across all sessions showed a small increase in the low SNR condition (0.04, $P = 0.05$, $n = 88$, sign test). For sessions in which DI was significantly greater for low SNR targets in low SNR blocks (filled bars), the change was not significant (0.07, $P = 0.06$, $n = 55$).

(Fig. 7D–F). Thus, task engagement effects were similar to those of the selective attention task, but unlike selective attention, changing task difficulty impacted average spontaneous rate rather than the noise-evoked response (Fig. 7G).

As in the case of selective attention, effects of task difficulty varied substantially across neurons. To better understand the interplay of task engagement and effort across the neural population, we compared the change in noise-evoked activity during task engagement (active vs. passive) to the change between difficulty conditions (Fig. 7H). In this case, we observed a significant negative correlation in BMI ($r = -0.45$, $P = 0.004$). Neurons whose response increased during task engagement tended to be suppressed in the more difficult, low-SNR condition. Conversely, neurons whose response decreased during engagement tended to produce an enhanced response in the low-SNR condition. Thus, although no consistent effects of effort on average evoked activity were observed, neurons affected by changes in difficulty were the same as those affected by task engagement.

The pattern of changes in neural activity associated with changes in task difficulty did not suggest an obvious impact on neural discriminability. For a complete comparison with the selective attention data, however, we also tested whether engaging in the more difficult, low-SNR task increased neural discriminability of target versus noise (Fig. 7I). We used the same measure of d' as for the selective attention data. Across the entire set of 88 neurons in the variable SNR data set, we observed a small increase in d' (median change 0.04, $P = 0.05$, sign test). When we considered only the subset of neurons for which DI for the low-SNR tone significantly improved in the low-SNR versus high-SNR condition, we observed a trend toward a greater increase in d', but this change was not significant (median change 0.07, $P = 0.06$). Thus, the change in effort may be accompanied by increased neural discriminability between target and noise, but any effects were weaker than in the selective attention data.

Selective Attention and Task Difficulty Effects are Reflected in Neuronal Filter Models

To investigate the effects of changing behavioral state on sensory coding in more detail, we computed STRF models for each single-unit, fit using activity evoked by the distractor noise under the different behavioral conditions (Fig. 8A). The linear STRF typically describes spectro-temporal tuning as the weighted sum over several channels of a broadband stimulus spectrogram (see Eq. 5 and Aertsen and Johannesma 1981; Thorson et al. 2015). For vocalization-modulated noise, we employed a spectrally simplified version of the STRF that summed activity over just the 2 spectral channels that comprised the noise stimulus. Because one noise band was positioned outside of the neuron's RF, the STRF typically showed tuning to only one of the 2 stimulus channels (see examples in Fig. 8A).

To investigate the interaction between attention and spectro-temporal tuning, we calculated separate STRFs for data from each behavior condition. Model performance was assessed by measuring the accuracy (correlation coefficient, Pearson's R) with which the STRF predicted the neuron's time-varying spike rate. Effects of behavioral state were identified by comparing the predictive power of these behavior-dependent STRFs to performance of behavior-independent STRFs, for which a single model was fit across all behavior conditions. A neuron was labeled as showing a significant effect of attention

if the behavior-dependent STRF predicted neural responses significantly better than the behavior-independent STRF ($P < 0.05$, permutation test). Prediction accuracy was measured using a held-out validation data set, so that any difference in prediction accuracy could not reflect overfitting to noise (Wu et al. 2006; David et al. 2009). Identical noise sounds were presented during passive listening and different task conditions. Thus, each STRF for a single neuron was estimated using identical stimuli, avoiding potential stimulus-related bias between estimates (Wu et al. 2006; David et al. 2007).

For the selective attention data, the behavior-dependent STRF performed significantly better in 26/54 neurons, and the mean prediction correlation was significantly greater for the behavior-dependent model (mean $R = 0.31$ vs. 0.25, $n = 54$, $P < 0.0001$, sign test, Fig. 8B). We then compared the fit parameters of STRFs for neurons showing a significant improvement for the behavior-dependent model. For these neurons, peak gain of the STRF's temporal filter was lower for the attend RF versus attend away condition (Fig. 8C–D, $P = 0.007$, permutation test). Baseline firing rate, which would reflect a change in spontaneous rate, did not change between attention conditions ($P = 0.3$). Thus, we observed changes in STRFs that were consistent with changes in mean spontaneous and evoked firing rates (Fig. 3E). Moreover, there was no consistent change in the shape of the temporal response (Fig. 8C), indicating that selective attention primarily affected the magnitude of A1 noise-evoked responses, but not temporal tuning.

For the variable SNR task, we also observed improved prediction accuracy for the behavior-dependent model in a subset of neurons (51/88), and an overall improvement in mean prediction accuracy (mean $R = 0.34$ vs. 0.29, $n = 88$, $P < 0.0001$, permutation test). For neurons with significantly better behavior-dependent STRFs, baseline firing rate increased during the less difficult, high-SNR task ($P = 0.02$, permutation test), and there was no change in the STRF's temporal filter ($P > 0.5$, Fig. 8E–F). Thus, as in the case of the selective attention data, the behavior-dependent changes observed in mean spike rate were reciprocated by changes in the STRF baseline, again with no systematic change in temporal filter properties.

Discussion

This study demonstrates that auditory selective attention suppresses the responses of neurons in primary auditory cortex (A1) to distractor sounds that compete with an attended target. In nearly 50% of A1 neurons, responses to distractor noise centered at the frequency of a target tone were suppressed relative to distractors in a different frequency band. Responses to the target tone were not suppressed, leading to an improvement in its neural detection threshold. Selective attention effects differed from those of task engagement and changing task difficulty, which produced more systematic changes in spontaneous rather than evoked activity and tended to affect a different population of A1 neurons.

A primary goal of this study was to isolate mechanisms producing task-related plasticity previously reported in A1 when animals engaged in auditory behavior (Fritz et al. 2003; Lee and Middlebrooks 2011; David et al. 2012; Niwa et al. 2012). The current data demonstrate that task-related effects can in fact be broken down into components that reflect task engagement, selective attention, and effort. These behavior-related changes have a similar total magnitude to those reported following task engagement in A1 (Fritz et al. 2003; Niwa et al. 2012). The same approach can be applied to measure the composition of behavior-related changes in more central belt and parabelt
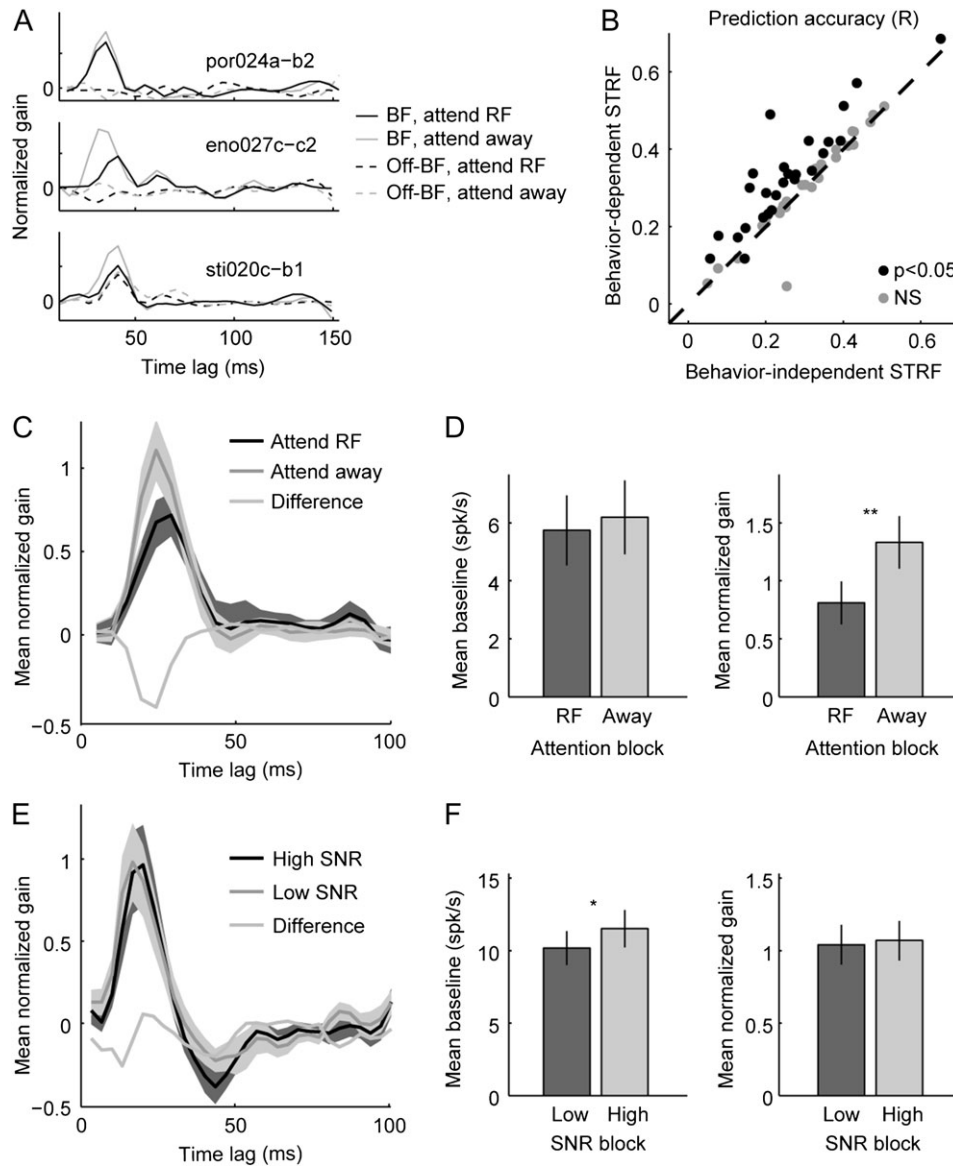
Figure 8. Effects of attention and variable SNR on neuronal filter properties. (A) Comparison of behavior-dependent STRFs for 3 neurons from selective attention experiments, estimated separately using data from attend RF and attend away conditions. In both attention conditions, the temporal filter for the RF channel (solid lines, for the stimulus channel centered in the neuron's RF) showed excitatory tuning with 0–50 ms time lag. Gain was lower for this channel in the attend RF condition (black vs. gray lines). The non-RF channel (dashed lines) typically showed weaker gain, if any, and no consistent change between attention conditions. (B) Scatter plot comparing prediction accuracy of behavior-independent versus behavior-dependent STRFs across the selective attention data set. Filled dots correspond to neurons with a significant difference in prediction accuracy between models ($P < 0.05$, sign test). Mean prediction accuracy was higher for the behavior-dependent model (mean $r = 0.24$ vs. 0.28 for behavior-dependent vs. independent, $P < 0.001$, $n = 54$, sign test). (C) Average temporal response function at BF for behavior-dependent STRFs in the attend RF versus attend away condition ($n = 26/54$ neurons with significant improvement for behavior-dependent model). (D) Mean baseline and peak gain for behavior-dependent STRFs in attend RF versus attend away conditions show a significant difference in gain (**$P = 0.007$, sign test). (E) Average temporal response functions of behavior-dependent STRFs estimated separately for low- and high-SNR conditions ($n = 51/88$ with significant benefit for behavior-dependent model). (F) Mean baseline and peak gain for behavior-dependent STRFs in low-SNR and high-SNR conditions shows a significant difference in response baseline (*$P = 0.02$, sign test).

areas, where overall behavior effects are generally larger (Niwa et al. 2013; Atiani et al. 2014; Tsunada et al. 2015).

Mechanisms of Auditory Selective Attention

Distractor suppression may be a general strategy used by the auditory system to perform a variety of tasks, including those that require selective attention. At face value, these results are inconsistent with human studies of attention during streaming of simultaneously presented sounds. When human subjects

attend to one of 2 speech or non-speech streams, the neural representation of the attended stream is enhanced over the nonattended stream (Ding and Simon 2012; Mesgarani and Chang 2012; Da Costa et al. 2013). LFP and fMRI recordings in nonhuman primates show similar enhancement for selective attention between 2 auditory streams (Lakatos et al. 2013) or auditory versus visual streams (Rinne et al. 2017).

Several factors could explain this difference from the suppression observed in the current study. First, previous studies used field recordings that sum the activity of large neural

populations, and the contribution of single-neuron activity to these signals is not well understood. For example, the magnitude of evoked field potentials can be influenced the synchrony of local neural populations, independent of their firing rate (Telenczuk et al. 2010). Thus, a change in the amplitude of one signal does not necessitate a change of the same sign in the other. Second, several of the human studies focused on non-primary belt and parabelt areas of auditory cortex that may not undergo the same changes as A1. Finally, and perhaps most importantly, these differences may be explained by task structure. In the tone-in-noise task used in the current study, the noise does not provide useful information to the subject performing the task. A strategy of suppression thus may be helpful for enhancing contrast with the masked target tone (Durlach 1963; Dai et al. 1991; de Cheveigné 1993). Streaming speech, in contrast, requires a much different listening strategy in which the semantic content of the signal must be encoded rather than suppressed. An alternative task in which animals must detect specific features in the noise stream (e.g., modulation patterns) may reveal enhanced responses similar to those observed in the speech studies. Neural populations encoding targets and spectrally similar maskers are difficult to isolate in large-scale field recordings, which may explain why the effects of attention on masker stimuli have not been characterized previously.

The distractor suppression we observed may be a neural correlate of a phenomenon observed in psychophysical studies of tone-in-noise detection. Humans can attend to a narrow frequency band surrounding a tone target: when listeners are cued to expect a tone at a given frequency, their ability to detect tone targets more than one critical bandwidth from the expected frequency dramatically decreases, effectively attenuating off-target sounds by 7 dB (Greenberg and Larkin 1968; Scharf et al. 1987; Dai et al. 1991). More generally, the hypothesis that the auditory system suppresses neural responses to predictable distractors in order to amplify target signals is widespread in studies of psychoacoustics (Durlach 1963; de Cheveigné 1993; Shinn-Cunningham 2008). We observed suppression of responses to a ½-octave noise distractor during a tone detection. This suppression may reflect sharpening of frequency tuning curves when attention was shifted into the neuron's RF. If so, then the noise suppression would increase the neural representation of the tone target at the expense of sounds at nearby frequencies.

Some aspects of our experiment suggest caution before adopting this interpretation. First, our acoustic stimuli sampled frequency tuning very sparsely and did not permit direct measurement of changes in frequency tuning bandwidth across attention conditions. Second, it is not certain that animals used frequency as the dominant cue to direct their attention, since our stimuli also provided spatial and envelope cues to distinguish the attended and nonattended stream (Nelson and Carney 2007). Specific interactions between attention and spectral coding can therefore only be resolved by future studies that require attention exclusively to spectral cues and probe a larger range of stimulus frequencies.

Separability of Engagement and Attention Effects

Previous studies of tone-detection behavior have identified changes in the selectivity of A1 neurons specific to the frequency of a target tone (Fritz et al. 2003; Atiani et al. 2009; David et al. 2012; Kuchibhotla et al. 2016), but it has remained unclear how much these changes in neural tuning reflect selective attention to the target frequency versus more global processes of task engagement. Here, we have isolated these effects and shown that selective attention produces frequency-specific suppression of responses evoked by distractors. Task engagement also produced changes that were independent of the locus of attention. Engagement was equally likely to produce enhancement or suppression of sound-evoked activity. This overall stability of average evoked activity suggests that changes in cortical network activity are tempered by a homeostatic mechanism that maintains stability in the level of spiking activity across auditory cortex. The selective suppression of noise responses at the locus of attention, thus, may be accompanied by enhancement at nonattended locations (Fritz et al. 2005).

Changing behavioral effort also influenced noise-evoked responses in A1. However, effects were not consistent, and a change in effort was equally likely to increase or decrease responses. Instead, greater behavioral effort lead to a decrease in spontaneous spike rate. The distinct effects of attention and effort on evoked versus spontaneous activity suggest that different modulatory circuits mediate these changes. Task engagement and adjusting effort could reflect large-scale changes in brain state that do not depend on the acoustic features of task stimuli. The influence of these global state variables may arise from circuits that mediate effects of arousal (Issa and Wang 2008; McGinley et al. 2015). In contrast to changes in global state, selective attention requires differential processing of acoustic features, and its effects cannot be uniform across the auditory system. Consistent with a system containing distinct global versus local modulatory top-down circuits, the neuronal populations affected by selective attention and task engagement are not correlated.

While we did observe a correlation in the magnitude of effects of task engagement and effort, there was substantial additional variability of behavioral effects across neurons that could not be explained by tuning properties or recording depth. Future studies that identify the location of neurons in the cortical circuit more precisely, either by genetic labels (Natan et al. 2015; Kuchibhotla et al. 2016) or network connectivity (Schneider et al. 2014) may explain more of this variability. These approaches may also be used to confirm whether engagement and effort effects derive from the same source.

Effects of selective attention and effort could emerge at different stages of the auditory network. Multiple populations of inhibitory interneurons have been implicated in behavioral state modulation in A1, and signals reflecting different aspects of behavioral state could arrive through distinct inhibitory subpopulations (Pi et al. 2013; Kuchibhotla et al. 2016). These signals could also arrive in different auditory brain areas. Engaging in a tone detection task changes activity in the inferior colliculus (IC), an area upstream from A1 (Slee and David 2015), but it is not known if selective attention modulates IC activity. Studies comparing the same tasks across brain areas are critical for determining where behavior-mediated effects emerge in the auditory processing network. More generally, these results indicate that multiple aspects of task structure can influence activity in sensory cortical areas. Control and monitoring of behavioral state (arousal, reward, motor contingencies, attention) is required to assess the effects of a desired behavioral manipulation (David et al. 2012; Baruni et al. 2015; Luo and Maunsell 2015; McGinley et al. 2015).

Impact of Behavioral State on Neural Coding

The changes in sound-evoked activity associated with selective attention support enhanced neural discriminability of target tone versus distractor noise in A1. The absence of significant suppression in the target response alone does not imply a

selective suppression of distractor responses. However, the increase in neural discriminability when attention is directed into the RF does indicate that any suppression is stronger for the distractor, increasing the difference in neural response between the 2 sound categories. Improvements in neural discriminability with similar magnitude have been observed previously in A1 following engagement in auditory detection tasks (Ryan et al. 1984; David et al. 2012; Niwa et al. 2012). Some early single-unit studies in monkey auditory cortex also support the view that enhanced responses occur selectively for stimuli that carry task-relevant information. In tasks that required discrimination between spectral (Beaton and Miller 1975) or spatial sound features (Benson and Hienz 1978), neural responses were enhanced to the sound requiring a behavioral response. Moreover, activity in auditory cortex can explicitly encode behavioral choice (Niwa et al. 2012; Bizley et al. 2013; Tsunada et al. 2015). Thus, across several studies, a model has emerged in which neural discriminability of task-relevant stimulus features increases, and that emergent representation feeds directly into behavioral decisions. The parallel between neural signaling and behavioral output encourages a straightforward conclusion that changes in behavioral state serve primarily to enhance coding of behaviorally relevant categories. While a change in d' in the range 0.05–0.1 is not extremely large for a single neuron, this value represents the average increase per A1 neuron. Effects of this size can be substantial when compounded across an entire neural population (Shadlen and Newsome 1998).

However, in the current study, the benefit of changes in neural coding is not always so clear. The shift in spontaneous rate associated with listening effort did not produce significant enhancement in neural discriminability. Changes in spontaneous activity that do not enhance discriminability have been reported previously, following switches between tasks that vary in difficulty (Rodgers and DeWeese 2014). Studies involving switching targets between auditory and other sensory modalities have also reported mixed results, sometimes finding enhanced coding of the attended modality (O'Connell et al. 2014) and sometimes not (Hocherman et al. 1976; Otazu et al. 2009). The cortex contains a rich diversity of circuits for learning new behavioral associations and adapting to new contexts (Fritz et al. 2010; David et al. 2012; Jaramillo et al. 2014). An architecture that supports flexibility and multiplexing of behaviors likely imposes additional constraints on behavior-related changes beyond enhanced sensory discriminability.

The analysis of behavior-dependent STRFs revealed task-related effects consistent with the changes in PSTH response gain between selective attention conditions and changes in spontaneous activity between variable SNR conditions. The STRF analysis also allowed us to identify any possible changes in temporal filter properties between behavior conditions. However, temporal response properties were largely stable. Thus, while we do observe changes in spectral tuning, top-down behavioral signals do not affect temporal tuning in the current task. It should be noted that the current task did not require attention to specific temporal features, and a task requiring discrimination of temporal features, such as modulation detection or discrimination, might have a different effect (Fritz et al. 2010; Niwa et al. 2012).

Comparison to Studies of Visual Selective Attention

Noise suppression may be viewed as a mechanism to bias competition between neural representations of the tone target and masking noise. Analogous effects have been observed in the primate visual system (Spitzer et al. 1985; Desimone and Duncan 1995; Connor et al. 1996; Reynolds et al. 1999). When macaques attend to one of 2 stimuli in the RF of a visual cortical neuron, the neural response shifts to resemble the response to the attended stimulus presented in isolation (Connor et al. 1996; Reynolds et al. 1999). Thus, when attention is directed within a visual RF, the spatial RF effectively shrinks, and responses to stimuli outside the locus of attention are attenuated. Similar effects are observed across visual cortex, growing progressively greater in magnitude across areas V1, V2, and V4 (Motter 1993; Luck et al. 1997). This narrowing of tuning has been modeled as enhanced surround inhibition (Sundberg et al. 2009). In the auditory system, stimulus bandwidth can be viewed as a dimension in sensory space (bandpass noise vs. very narrowband tones), analogous to retinotopic space (Schreiner and Winer 2007). By the logic of the current task, attention within the A1 RF is directed to narrowband versus broadband stimuli. A narrowing of spectral tuning bandwidth that would produce the distractor suppression reported here may be analogous to the shrinking of visual spatial RFs around the locus of spatial attention. The average BMI of −0.12 for noise responses in A1 falls between the magnitude of spatial attention effects in areas V2 and V4 measured using a similar statistic (Luck et al. 1997).

Simultaneous population recordings from V4 during selective attention behavior have revealed a decrease in noise correlations between pairs of neurons that encode stimuli at the locus of attention (Cohen and Maunsell 2009), and similar effects were recently reported in A1 (Downer et al. 2017). Because the current data were collected serially from single neurons, it was not possible to measure interneuronal correlations. The effect of auditory selective attention on neural population activity in this sensory context remains an open question for future studies.

Notes

References

Aertsen AM, Johannesma PI. 1981. The spectro-temporal receptive field. A functional characteristic of auditory neurons. Biol Cybern. 42:133–143.

Atiani S, David SV, Elgueda D, Locastro M, Radtke-Schuller S, Shamma SA, Fritz JB. 2014. Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. Neuron. 82: 486–499.

Atiani S, Elhilali M, David SV, Fritz JB, Shamma SA. 2009. Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. Neuron. 61:467–480.

Baruni JK, Lau B, Salzman CD. 2015. Reward expectation differentially modulates attentional behavior and activity in visual area V4. Nat Neurosci. 18:1656–1663.

Beaton R, Miller JM. 1975. Single cell activity in the auditory cortex of the unanesthetized, behaving monkey: correlation with stimulus controlled behavior. Brain Res. 100: 543–562.

Benson DA, Hienz RD. 1978. Single-unit activity in the auditory cortex of monkeys selectively attending left vs. right ear stimuli. Brain Res. 159:307–320.

Bizley JK, Nodal FR, Nelken I, King AJ. 2005. Functional organization of ferret auditory cortex. Cereb Cortex. 15:1637–1653.

Bizley JK, Walker KMM, Nodal FR, King AJ, Schnupp JWH. 2013. Auditory cortex represents both pitch judgments and the corresponding acoustic cues. Curr Biol. 23:620–625.

Calabrese A, Schumacher JW, Schneider DM, Paninski L, Woolley SMN. 2011. A generalized linear model for estimating spectrotemporal receptive fields from responses to natural sounds. PLoS One. 6:e16104.

Chen Y, Martinez-Conde S, Macknik SL, Bereshpolova Y, Swadlow HA, Alonso JM. 2008. Task difficulty modulates the activity of specific neuronal populations in primary visual cortex. Nat Neurosci. 11:974–982.

Cherry EC. 1953. Some experiments on the recognition of speech, with one and with two ears. J Acoust Soc Am. 25: 975–979.

Cohen MR, Maunsell JHR. 2009. Attention improves performance primarily by reducing interneuronal correlations. Nat Neurosci. 12:1594–1600.

Connor CE, Gallant JL, Preddie DC, Van Essen DC. 1996. Responses in area V4 depend on the spatial relationship between stimulus and attention. J Neurophysiol. 75: 1306–1308.

Da Costa S, van der Zwaag W, Miller LM, Clarke S, Saenz M. 2013. Tuning in to sound: frequency-selective attentional filter in human primary auditory cortex. J Neurosci. 33: 1858–1863.

Dai HP, Scharf B, Buus S. 1991. Effective attenuation of signals in noise under focused attention. J Acoust Soc Am. 89:2837.

David SV, Fritz JB, Shamma SA. 2012. Task reward structure shapes rapid receptive field plasticity in auditory cortex. Proc Natl Acad Sci USA. 109:2150–2155.

David SV, Hayden BY, Mazer JA, Gallant JL. 2008. Attention to stimulus features shifts spectral tuning of V4 neurons during natural vision. Neuron. 59:509–521.

David SV, Mesgarani N, Fritz JB, Shamma SA. 2009. Rapid synaptic depression explains nonlinear modulation of spectrotemporal tuning in primary auditory cortex by natural stimuli. J Neurosci. 29:3374–3386.

David SV, Mesgarani N, Shamma SA. 2007. Estimating sparse spectro-temporal receptive fields with natural stimuli. Network. 18:191–212.

David SV, Shamma SA. 2013. Integration over multiple timescales in primary auditory cortex. J Neurosci. 33: 19154–19166.

de Cheveigné A. 1993. Separation of concurrent harmonic sounds: Fundamental frequency estimation and a time-domain cancellation model of auditory processing. J Acoust Soc Am. 93:3271–3290.

Desimone R, Duncan J. 1995. Neural mechanisms of selective visual attention. Annu Rev Neurosci. 18:193–222.

Ding N, Simon JZ. 2012. Emergence of neural encoding of auditory objects while listening to competing speakers. Proc Natl Acad Sci USA. 109:11854–1859.

Downer J, Rapone B, Verhein J, O'Connor KN, Sutter ML. 2017. Feature selective attention adaptively shifts noise correlations in primary auditory cortex. J Neurosci. 37(21):5378–5392.

Durlach NI. 1963. Equalization and cancellation theory of binaural masking-level differences. J Acoust Soc Am. 35:1206.

Englitz B, David SV, Sorenson MD, Shamma SA. 2013. MANTA-an open-source, high density electrophysiology recording suite for MATLAB. Front Neural Circuits. 7:69.

Fritz JB, David SV, Radtke-Schuller S, Yin P, Shamma SA, Elhilali M. 2010. Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. Nat Neurosci. 13:1011–1019.

Fritz JB, Elhilali M, Shamma SA. 2005. Differential dynamic plasticity of A1 receptive fields during multiple spectral tasks. J Neurosci. 25:7623–7635.

Fritz JB, Shamma SA, Elhilali M, Klein DJ. 2003. Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. Nat Neurosci. 6:1216–1223.

Greenberg GZ, Larkin WD. 1968. Frequency-response characteristic of auditory observers detecting signals of a single frequency in noise: the probe-signal method. J Acoust Soc Am. 44:1513.

Heffner HE, Heffner RS. 1995. Conditioned avoidance. In: Klump GM, Dooling RJ, Fay RR, Stebbins WC, editors. Methods in comparative psychoacoustics. Basel: Birkhauser.

Hocherman S, Benson DA, Goldstein MH Jr, Heffner HE, Hienz RD. 1976. Evoked unit activity in auditory cortex of monkeys performing a selective attention task. Brain Res. 117:51–68.

Issa EB, Wang X. 2008. Sensory responses during sleep in primate primary and secondary auditory cortex. J Neurosci. 28: 14467–14480.

Jaramillo S, Borges K, Zador AM. 2014. Auditory thalamus and auditory cortex are equally modulated by context during flexible categorization of sounds. J Neurosci. 34: 5291–5301.

Kidd G, Arbogast TL, Mason CR, Gallun FJ. 2005. The advantage of knowing where to listen. J Acoust Soc Am. 118:3804–3815.

Kuchibhotla KV, Gill JV, Lindsay GW, Papadoyannis ES, Field RE, Sten TAH, Miller KD, Froemke RC. 2016. Parallel processing by cortical inhibition enables context-dependent behavior. Nat Neurosci. 20:62–71.

Lakatos P, Musacchia G, O'Connel MN, Falchier AY, Javitt DC, Schroeder CE. 2013. The spectrotemporal filter mechanism of auditory selective attention. Neuron. 77:750–761.

Lee C-CC, Middlebrooks JC. 2011. Auditory cortex spatial sensitivity sharpens during task performance. Nat Neurosci. 14: 108–114.

Luck SJ, Chelazzi L, Hillyard SA, Desimone R. 1997. Neural mechanisms of spatial selective attention in areas V1, V2 and V4 of macaque visual cortex. J Neurophysiol. 77:24–42.

Luo TZ, Maunsell JHR. 2015. Neuronal modulations in visual cortex are associated with only one of multiple components of attention. Neuron. 86:1182–1188.

Machens CK, Wehr MS, Zador AM. 2004. Linearity of cortical receptive fields measured with natural sounds. J Neurosci. 24:1089–1100.

McAdams CJ, Maunsell JHR. 1999. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. J Neurosci. 19:431–441.

McGinley MJ, David SV, McCormick DA. 2015. Cortical membrane potential signature of optimal states for sensory signal detection. Neuron. 87:179–192.

Mesgarani N, Chang EF. 2012. Selective cortical representation of attended speaker in multi-talker speech perception. Nature. 485:233–236.

Moran J, Desimone R. 1985. Selective attention gates visual processing in the extrastriate cortex. Science. 229:782–784.

Motter BC. 1993. Focal attention produces spatially selective processing in visual cortical areas V1, V2 and V4 in the presence of competing stimuli. J Neurophysiol. 70:909–919.

Natan RG, Briguglio JJ, Mwilambwe-Tshilobo L, Jones SI, Aizenberg M, Goldberg EM, Geffen MN. 2015. Complementary control of sensory adaptation by two types of cortical interneurons. Elife. 4:e09868.

Nelson PC, Carney LH. 2007. Neural rate and timing cues for detection and discrimination of amplitude-modulated tones in the awake rabbit inferior colliculus. J Neurophysiol. 97:522–539.

Niwa M, Johnson JS, O'Connor KN, Sutter ML. 2012. Active engagement improves primary auditory cortical neurons' ability to discriminate temporal modulation. J Neurosci. 32:9323–9334.

Niwa M, Johnson JS, O'Connor KN, Sutter ML. 2013. Differences between primary auditory cortex and auditory belt related to encoding and choice for AM sounds. J Neurosci. 33:8378–8395.

O'Connell MN, Barczak A, Schroeder CE, Lakatos P. 2014. Layer specific sharpening of frequency tuning by selective attention in primary auditory cortex. J Neurosci. 34:16496–16508.

Otazu GH, Tai LH, Yang Y, Zador AM. 2009. Engaging in an auditory task suppresses responses in auditory cortex. Nat Neurosci. 12:646–654.

Pi H-J, Hangya B, Kvitsiani D, Sanders JI, Huang ZJ, Kepecs A. 2013. Cortical interneurons that specialize in disinhibitory control. Nature. 503:521–524.

Reynolds JH, Chelazzi L, Desimone R. 1999. Competitive mechanisms subserve attention in macaque areas V2 and V4. J Neurosci. 19:1736–1753.

Rinne T, Muers RS, Salo E, Slater H, Petkov CI. 2017. Functional imaging of audio–visual selective attention in monkeys and humans: How do lapses in monkey performance affect cross-species correspondences? Cereb Cortex. 1–14.

Rodgers CC, DeWeese MR. 2014. Neural correlates of task switching in prefrontal cortex and primary auditory cortex in a novel stimulus selection task for rodents. Neuron. 82:1157–1170.

Ryan AF, Miller JM, Pfingst BE, Martin GK. 1984. Effects of reaction time performance on single-unit activity in the central auditory pathway of the rhesus macaque. J Neurosci. 4:298–308.

Scharf B, Quigley S, Aoki C, Peachey N, Reeves A. 1987. Focused auditory attention and frequency selectivity. Percept Psychophys. 42:215–223.

Schneider DM, Nelson A, Mooney R. 2014. A synaptic and circuit basis for corollary discharge in the auditory cortex. Nature. 513:189–194.

Schreiner CE, Winer JA. 2007. Auditory cortex mapmaking: principles, projections, and plasticity. Neuron. 56:356–365.

Shadlen MN, Newsome WT. 1998. The variable discharge of cortical neurons: Implications for connectivity, computation, and information coding. J Neurosci. 18:3870–3896.

Shamma SA, Elhilali M, Micheyl C. 2011. Temporal coherence and attention in auditory scene analysis. Trends Neurosci. 34:114–123.

Shamma SA, Fleshman JW, Wiser PR, Versnel H. 1993. Organization of response areas in ferret primary auditory cortex. J Neurophysiol. 69:367–383.

Shinn-Cunningham BG. 2008. Object-based auditory and visual attention. Trends Cogn Sci. 12:182–186.

Shinn-Cunningham BG, Best V. 2008. Selective attention in normal and impaired hearing. Trends Amplif. 12:283–299.

Singh NC, Theunissen FE. 2003. Modulation spectra of natural sounds and ethological theories of auditory processing. J Acoust Soc Am. 114:3394–3411.

Slee SJ, David SV. 2015. Rapid task-related plasticity of spectrotemporal receptive fields in the auditory midbrain. J Neurosci. 35:13090–13102.

Spence C, Driver J. 1994. Covert spatial orienting in audition: exogenous and endogenous mechanisms. J Exp Psychol Hum Percept Perform. 20:555–574.

Spitzer H, Desimone R, Moran J. 1985. Increased attention enhances both behavioral and neuronal performance. Science. 240:338–340.

Sundberg KA, Mitchell JF, Reynolds JH. 2009. Spatial attention modulates center-surround interactions in macaque visual area V4. Neuron. 61:952–963.

Telenczuk B, Nikulin VV, Curio G. 2010. Role of neuronal synchrony in the generation of evoked EEG/MEG responses. J Neurophysiol. 104:3557–67.

Theunissen FE, David SV, Singh NC, Hsu A, Vinje WE, Gallant JL. 2001. Estimating spatial temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. Netw Comput Neural Syst. 12:289–316.

Thorson IL, Liénard J, David SV. 2015. The essential complexity of auditory receptive fields. PLoS Comput Biol. 11:e1004628.

Tian B, Reser D, Durham A, Kustov A, Rauschecker JP. 2001. Functional specialization in rhesus monkey auditory cortex. Science. 292:290–293.

Tsunada J, Liu ASK, Gold JI, Cohen YE. 2015. Causal contribution of primate auditory cortex to auditory perceptual decision-making. Nat Neurosci. 19:135–142.

Willmore BDB, Prenger RJ, Gallant JL. 2010. Neural representation of natural images in visual area V2. J Neurosci. 30:2102–2114.

Woods DL, Alain C, Diaz R, Rhodes D, Ogawa KH. 2001. Location and frequency cues in auditory selective attention. J Exp Psychol Hum Percept Perform. 27:65–74.

Wu MC-K, David SV, Gallant JL. 2006. Complete functional characterization of sensory neurons by system identification. Annu Rev Neurosci. 29:477–505.

Yin P, Fritz JB, Shamma SA. 2010. Do ferrets perceive relative pitch? J Acoust Soc Am. 127:1673–1680.