**Cortical efficient coding dynamics shape behavioral performance.**

Chris Angeloni1,2, Wiktor Mlynarski3, Aaron M. Williams2,5, Katherine C. Wood2, Linda Garami2, Eugenio Piasini5, Ann Hermundstad4, Maria N. Geffen2,5

1Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA

2Department of Otorhinolaryngology, University of Pennsylvania, Philadelphia, PA, USA

3Institute of Science and Technology Austria, Klosterneuburg, Austria

4Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA, USA

5Department of Neuroscience, Department of Neurology, University of Pennsylvania, Philadelphia, PA, USA

**Abstract**

The efficient coding hypothesis postulates that neurons shape their response properties to match their dynamic range to the statistics of incoming signals. However, whether and how the dynamics of efficient neuronal adaptation inform behavior has not been directly shown. Here, we trained mice to detect a target presented in noise shortly after a change in the noise contrast. The observed changes in cortical gain and behavior followed the predictions of a normative model of efficient cortical sound processing; specifically, target detection and sensitivity improved in low contrast relative to high contrast noise. Furthermore, the time course of target detectability varied depending on contrast, decreasing rapidly after a transition to high contrast, and increasing at a slower rate after a transition to low contrast. The auditory cortex was required for detection of targets in noise and cortical neuronal responses exhibited the predicted patterns of target detectability. The magnitude of gain modulation in cortical neurons predicted individual differences in behavioral performance. Combined, our results demonstrate that efficient neural codes in auditory cortex directly influence perceptual behavior.

**Introduction**

As we navigate the world around us, the statistics of the environment can change dramatically. The efficient coding hypothesis postulates that neurons match their limited dynamic range to the statistics of incoming sensory signals1. Thus, through adaptation of their response properties, neurons can preserve their ability to encode information within many types of environments2–4. Such neuronal adaptation to the statistics of the environment has been found throughout different sensory modalities and brain regions5–13. In the auditory system, neurons exhibit contrast gain control, adapting the gain of their response function to match the variability in level (contrast) of the incoming sounds14–18. Yet it remains unknown whether and how the dynamics of contrast gain control in the auditory system inform behavior, as a direct link between neuronal adaptation and behavior has not been previously established. The goal of our study was to test the hypothesis that the efficient coding in auditory cortex shapes behavioral performance in an auditory task.

The efficient coding hypothesis has been formally implemented through normative models of brain function3,19–22. These models allow us to assess whether and how neuronal adaptation shapes sensory information and simulate how neural function constrains behavior. There has been previous work demonstrating that efficient codes can explain psychophysical biases23 and shape the rate of information transmission when stimulus statistics change dynamically 19,21,22. These studies, which are either theoretical in nature or based in human psychophysics, rely on assumptions of neuronal mechanisms of efficient coding that were not simultaneously measured. As such, there are no studies combining behavioral psychophysics with recordings of brain activity to simultaneously test the neural assumptions and behavioral predictions of these models.

Psychophysical studies suggest that the auditory system exhibits adaptation to acoustic contrast. In humans, target volume discriminability is greater in low contrast than in high, an effect consistent with gain control observed in primary auditory cortex24. Similar effects have also been shown in ferrets performing an acoustic localization task, where it was demonstrated that neural responses in the inferior colliculus of anesthetized ferrets changed in a manner consistent with previously observed perceptual shifts10. However, it remains unclear whether the observed behavioral effects are indeed due to changes in gain in auditory neurons, as previous behavioral studies were not performed with simultaneous neural recordings, so a direct relationship between neural gain and perceptual performance has yet to be assessed.

Our first goal for the study was to build a formal framework based on efficient coding to model the dynamics of contrast gain control, allowing us to predict how behavioral performance adapts after a change in contrast. We then derived a novel generalized linear model (GLM) to estimate moment-to-moment changes in neural gain, and found that gain in auditory cortex adapted similarly to the efficient coding model predictions. Next, to directly test the role of efficient coding in auditory behavior, we trained mice to detect targets in different contrast backgrounds. Contrast-induced changes in behavioral sensitivity and detection dynamics followed the model predictions. Simultaneous chronic neural recordings found that the neural code in auditory cortex was not only predictive of individual differences in behavior, but also that cortical responses to the task stimuli adapted similarly to our model, and that variability in neural gain was predictive of individual variability in task performance. Combined, our results identify a novel relationship between efficient neuronal coding and acoustic behavior, and provide a normative framework that can be used to predict the dynamics of behavioral performance in response to changing sensory environments.

**Results**

*A novel target-in-noise detection task and normative model for task predictions.*

To assess how perceptual performance is impacted by stimulus contrast, we devised a GO/NO-GO task in which head-fixed mice were trained to detect targets embedded in different contrast backgrounds. During each trial, the mouse was first presented with 3s of dynamic random chords (DRCs) of one contrast, after which a switch occurred to the other contrast. At variable delays after the switch, broad-band target chords were superimposed on the background chords, and mice were trained to lick for a water reward upon hearing the target. Target trials were interleaved with noise-only trials, during which the mouse was trained to withhold licking, but would receive a 7s timeout for licking after the contrast switch (Figure 1a,b). To assess behavioral sensitivity to targets, we parametrically varied target volume in each contrast (Figure 1c, top panel) and to assess behavioral adaptation, we parametrically varied target timing (Figure 1c, bottom panel). This stimulus design allowed us to quantitatively test whether and how the dynamics of adaptation to background contrast affect behavioral performance.

To predict the optimal time-course of contrast gain control and its impact on target detection behavior, we developed a normative model of task performance constrained by efficient neural coding. In this model, we simulated a neuron designed to estimate the contrast of the recent stimulus by adjusting the gain of its nonlinearity. A detailed description of the model is provided in the methods, but briefly: 1) at each timestep, the model neuron encodes incoming stimuli from a background stimulus stream whose contrast varies over time, 2) the stochastic spiking response of the neuron over a brief window of time is used to estimate the variance of the background stimulus, (Supplementary Figure 1b) 3) the variance estimate is then fed back upstream to the model neuron and used to adjust the gain of its nonlinearity to improve the estimate (Figure 1d, panels 1-3). Additionally, we simulated target responses by adding targets at different volumes at each timestep after a contrast transition (Supplementary Figure 1a). This allowed us to probe the sensitivity of the model neuron to targets of varying strength over the time-course of adaption (Supplementary Figure 1c,d).

Using this framework for efficient coding of stimulus contrast, we examined how discriminable target responses were from background responses as a function of target volume and timing. First we varied target volume to generate model psychometric curves (Figure 1e). We then fit target discriminability in each contrast with logistic functions after the model had fully adapted (25 time steps after the contrast switch), finding decreased detection thresholds and steeper slopes during low contrast. Next, we selected the target volumes closest to the fitted thresholds in each contrast (target means of 1.5 and 2.25 for low and high contrast, respectively), and measured the discriminability of those targets as a function of time from the contrast switch (Figure 1f). We observed two timescales which affected the discriminability of targets: 1) the abrupt change in the background distribution generated a fast initial drop in discriminability in high contrast, and, 2) a slower change in discriminability in both contrasts, as the gain of the model neuron adapted to the background (Figure 1g). We quantified the temporal dynamics of discriminability by fitting an exponential to the model responses, finding that discriminability dropped rapidly after a switch to high contrast, but increased slowly after a switch to low contrast.

In summary, we used a model of efficient gain adaptation to simulate performance in the previously described behavioral task. This model generated three primary predictions: When adapted to low contrast, 1) target detection thresholds will be lower, and 2) sensitivity to changes in target volume will be higher; 3) Discriminability over time will be asymmetric: rapidly decreasing after a switch to high contrast, but slowly increasing after a switch to low contrast (Figure 1h).

*Estimated cortical gain dynamics follow normative model predictions.*

Most previous work on contrast gain control utilized static models of contrast adaptation, estimating the steady-state gain after the neuron has fully adapted to the new stimulus [17,24,25], but see 15,26,27. A major goal of the current study was to analyze the dynamics of gain control, so we developed a novel GLM to estimate the gain of neurons in auditory cortex at each time step following a contrast transition. This model was fit to data recorded from the auditory cortex of an untrained mouse (n = 97 neurons) presented with 3s alternations of low and high contrast noise (Figure 2a,b). A brief description of the model follows (but see *Methods* for a detailed description).

The inference model is a Poisson GLM which decomposes the relationship between spiking activity () and the presented sounds () into a stimulus component (), contrast component (), and an interaction between the stimulus and the contrast (). Note that the interaction term defines a multiplicative relationship

**Graphical user interface, diagram

Description automatically generatedFigure 1.**

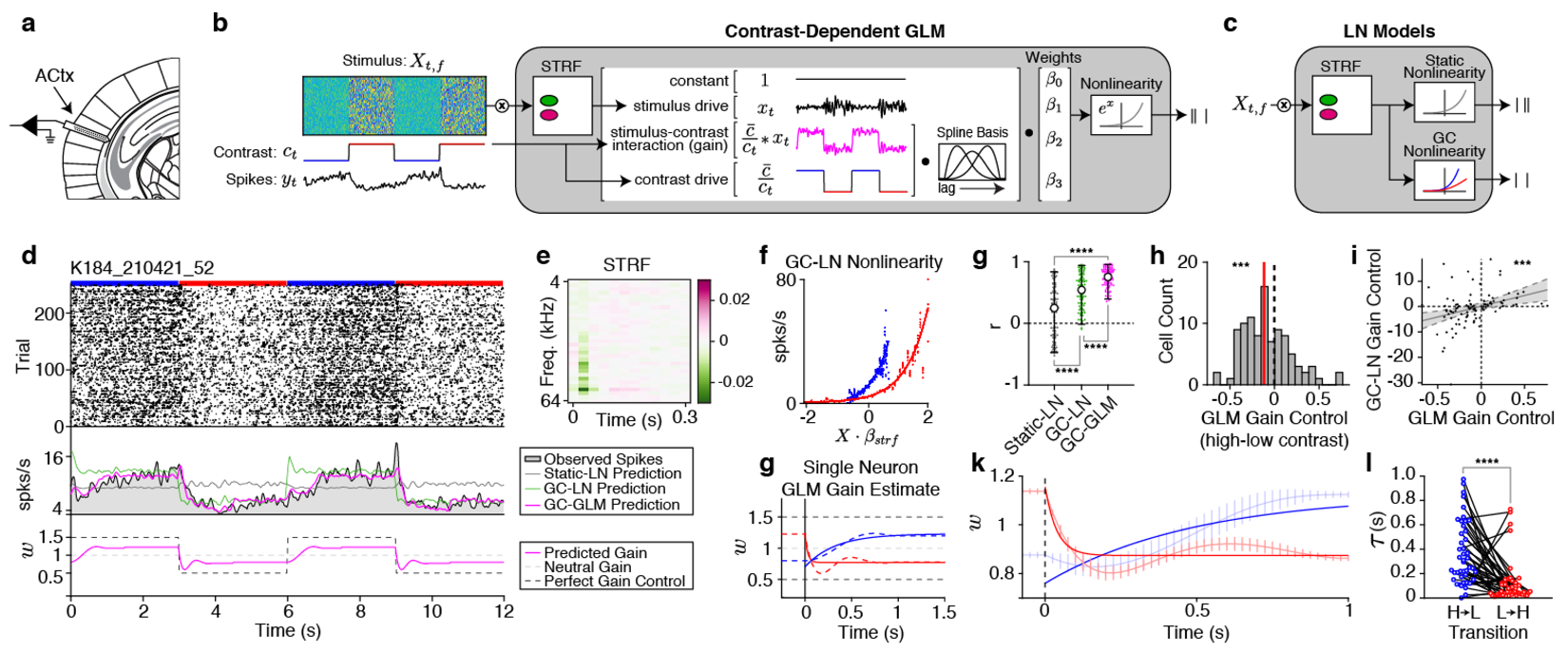
**a,** Experimental setup. Head-fixed mice are presented with sounds from an ultrasonic speaker. During behavior, mice receive water rewards through a lick spout. In a subset of mice, tetrodes were implanted in ACtx to record spiking activity. **b,** GO/NO-GO task design. *Left:* example NO-GO trials. From top to bottom: spectrogram of an example low-to-high contrast trial (colorbar indicates volume in dB SPL); waveform for sample spectrogram; example spectrogram for a high-to-low contrast trial; waveform for example spectrogram; temporally jittered response window to estimate false alarms over time; schematic lick responses in the response window; timeout of 7 seconds delivered after the first lick. Vertical red dashed line indicates the contrast switch after 3 seconds. Black horizontal scale bar indicates 1s. *Right:* example GO trials. From top to bottom: same as in left panel, except the response window immediately follows target presentation and licks within the target window trigger a ~5µL water reward. **c,** Target manipulation example waveforms. *Top:* overlaid trials where target volume differed. Volume is indicated by the amplitude and colorbar, with low volume targets shaded in cyan, and high volume targets shaded in magenta. *Bottom:* overlaid trials where target timing differed. Target timing is indicated by the colorbar, with light magenta targets occurring shortly after the contrast switch, and darker targets occurring at increasing delays. The red vertical dashed line indicates the contrast switch. **d,** Normative model of the task. *Left inset*: shows volume distributions for backgrounds (light lines) and targets (dark lines) in low and high contrast. (1) Spike generation process: a model neuron encodes stimuli sampled from a 1-dimensional sensory stimulus stream consisting of a background that transitions between low and high contrast (light lines); target stimuli (solid dots) were used to assess sensitivity. The response of the model neuron is governed by a sigmoidal function which then generates stochastic spikes. (2) The observed spike counts are integrated over a brief time window and used to estimate the current variance of the stimulus. (3) This estimate is then fed back upstream and used to adjust the gain of the model neuron to minimize the expected error in the estimate of stimulus variance at each time step. *Bottom insets*: Probability distributions of observing k spikes in response to the noise background (light lines) or targets (dark lines) at time step 7 after switches to high (red) or switches to low contrast (blue). **e,** Model psychometric functions. Discriminability between model spike rates in response to the background and targets as a function of contrast and target volume. Light dots indicate model discriminability whereas the solid lines indicate logistic fits to the data (see *Methods*). Dashed lines indicate detection thresholds (defined as from the logistic fit). **f,** Model target discrimination as a function of time and contrast. Dashed vertical line indicates the time where the background contrast changes. Light circles denote the model discriminability at each time step. Solid lines denote exponential fits to the data (see *Methods*). Each time course is the discriminability of targets at approximate threshold volume for that condition (1.5 target mean and 2.25 target mean for low and high contrast respectively).  **g,** The average change in gain of the model after each contrast transition. **h,** Model predictions for the effects of contrast on psychometric thresholds, slopes, and adaptation time constants, as assessed by logistic and exponential fits in **e** and **f**.

between the stimulus-driven response and the contrast, from which we define the gain of the neuron. The contrast and interaction components of the model were convolved with a B-spline basis set over a 1 s history window, which allowed us to model smooth changes in the underlying model components. Finally, we estimated the weights of each of these predictors for each neuron, and calculated a gain modulation index (, see *Methods*) to quantify the degree of gain control estimated by the GLM (Figure 2b). To validate the GLM, we simulated neurons with different amounts and temporal trajectories of gain control, and found that our model not only fit the simulated data well, but also accurately predicted the true gain over time under a variety of conditions (Supplemental Figure 2 and *Supplemental Methods*). For comparison, we also fit previously described linear-nonlinear models to each neuron [17,24,25], one with a static output nonlinearity (static-LN), and one with a contrast-dependent output nonlinearity (GC-LN, Figure 2c).

Model results for a representative neuron are plotted in Figure 2d-g. Qualitatively, the GLM with gain control (GC-GLM), outperforms standard LN models, capturing not only contrast-induced changes in overall firing rate, but also the time course of adaptation after the transition (Figure 2d, middle panel). Additionally, this example neuron demonstrates gain control, such that gain is greater in low relative to high contrast, and adapts faster after a switch to high contrast (Figure 2d, bottom panel; Figure 2g). Importantly, a similar pattern of gain control is also evident in the GC-LN model nonlinearities (Figure 2f), suggesting that the GC-GLM and GC-LN models are capturing similar changes in gain.

To further verify the GLM, we compared cross-validated correlations of the model predictions with the trial averaged spike count for all of our neurons. We found a significant effect of model type on goodness-of-fit (n = 97 neurons; Kruskall-Wallis test: *H*(2) = 93.61, p = 6.70e-21), and post-hoc Wilcoxon Sign-Rank tests found that the GC-GLM correlation was significantly higher (Median (*Mdn*) = 0.75, inter-quartile range (*IQR*) = 0.24) compared to the static-LN model (*Mdn* = 0.25, *IQR* = 0.73, *p* = 9.56e-10) and the GC-LN model (*Mdn* = 0.54, *IQR* = 0.49, *p* = 4.41e-6). Consistent with previous studies, we also found that the LN model with gain control outperformed the static model (*p* = 3.50e-6, Figure 2h). We then quantified whether the GLM detected significant gain control in the population by subtracting the gain estimate in low contrast from high contrast after the value of has stabilized (1s post contrast transition) and found significant gain control (*Mdn*: -0.10, *IQR*: 0.35, Wilcoxon sign-rank test: *rank* = 233, *Z* = -2.90, *p* = 0.004; Figure 2i). To further validate the GLM estimates of gain, we compared the gain control index from the GC-GLM to those of the GC-LN model, and found a significant linear prediction between the two measures (single linear regression: *F*(1,95) = 12.2, *p* = 7.33e-4, *R2* = 0.11; Figure 2j). These results demonstrate that our model more accurately predicts the variance of the neural data by incorporating gain adaptation dynamics and conclude that this method captures a similar estimate of neural gain when compared to standard LN models.

Using the GC-GLM, we estimated the time course of gain control by computing the trace of conditioned on the contrast transition, and then fitting each trace with an exponential function (Figure 2g). In neurons with gain control (gain control value less than zero) that were accurately fit by a single exponential(see *Methods*), the average time course of demonstrates asymmetric adaptation, rapidly decreasing after a switch to high contrast, and slowly increasing after a switch to low contrast (n = 45 neurons; Figure 2k). Within this same population, we quantified adaptation to each contrast using the time constant () of the exponential fit to each contrast. We found significantly longer time constants in low contrast (*Mdn* = 0.29, *IQR* = .39) relative to high contrast (*Mdn* = 0.048, *IQR* = 0.094; Wilcoxon sign-rank test: *rank* = 918, *Z*  = 4.52, *p* = 6.16e-6). This asymmetry in gain adaptation agreed with the predictions of the normative model (Figure 1g) and with previously described behavior of optimal variance estimators 19. Next, we tested whether the dynamics of gain adaptation predicted by the normative model and by the gain dynamics in auditory cortex are reflected in behavioral sensitivity to targets in noise.

**Figure 2.**

**a,** Schematic for acute recordings from auditory cortex. **b,** Schematic of generalized linear model (GLM) design. *From left to right:* The external variables considered by the model are the stimulus spectrogram, the stimulus contrast, and the observed spikes of cortical neurons. First, we estimate the spectrotemporal receptive field (STRF) of the cell. Second, we fit the parameters of a GLM with gain control (GC-GLM) designed to isolate the contributions of: 1) Pure stimulus drive; 2) The multiplicative interaction between the stimulus contrast and the stimulus drive, from which we define the gain of the neuron; 3) Pure contrast drive. To estimate smooth temporal trajectories of the gain predictor and contrast predictors, these predictors were convolved with a b-spline basis set at lags between 0 and 1s. The linear combination of the predictors and fitted weights was then passed through an exponential nonlinearity to produce spike rate predictions. **c,** Schematic of linear-nonlinear models. As in **b**, we first fit a STRF which is then passed through either a static exponential nonlinearity (static-LN) or independent nonlinearities fit separately to low and high contrast periods (GC-LN). **d,** Neuronal responses and model fits to a representative neuron. *Top*: a spike raster for the example neuron. Each period of contrast is indicated by the blue (low contrast) and red (high contrast) bars. *Middle*: PSTH of the example cell is plotted with a grey fill and black outline. The predictions of the static-LN model are plotted in grey, the GC-LN model in blue, and the GC-GLM model in red. All traces were smoothed with a 10ms wide Gaussian filter for visualization. *Bottom*: the gain estimate, w, inferred from the GLM parameters (red trace). Grey dashed line at 1 indicates the gain of a neuron with neutral gain. The dashed black line indicates the gain of a neuron with perfect, instantaneous gain control. **e,** The STRF fitted to this neuron. **f,** The nonlinearities fitted to low (blue) and high (red) contrast in the GC-LN model for the example neuron. Points indicate the mean observed firing rate (ordinate), binned according to observed filter prediction values (abscissa). Solid lines indicate exponential function fits to the underlying points. **g,** The estimate of the gain, w, for the example neuron for each contrast switch (dashed, red and blue lines). The solid red and blue lines are fits of an exponential function to the underlying traces. Dashed grey and black lines indicate neutral and perfect gain control values as in **d**. **h,** Cross-validated Pearson’s correlations between the trial-averaged firing rate trace and the model predictions. Grey, blue, and red dots indicate the correlations for each neuron (n=95) for the static-LN, GC-LN, and GC-GLM models, respectively. Open circles indicate the median correlation, and the error bars indicate 2.5-97.5 percentiles. Results of Wilcoxon Sign-Rank tests are indicated with asterisks. **i,** Distribution of gain control estimated by the GLM for the recorded population. Here, gain control is defined as after the estimate has stabilized to its final value (ie. after 1s). Dashed vertical line indicates no gain control, while the solid red line indicates the median of the distribution. Asterisks indicate the results of a Wilcoxon Sign-Rank test. **j,** Correspondence between gain control estimates from the GC-GLM model (abscissa) and the previously reported GC-LN model (ordinate). Black dots indicate the data for each neuron, while linear model fit and error are indicated by the grey line. Asterisks indicate significance of the linear fit to the data. **k,** Average time course of the gain estimate w for neurons with true gain control (ie. their gain control value is less than 0, n = 45). Light red and blue lines indicate the average value of w for transitions to high and low contrast, respectively (±SEM over neurons). Solid red and blue lines are exponential fits to the averages after the transition, which is marked by the dashed black line. **i,** Distributions of adaptation time constants of w after transitions to low, in blue, and high contrast, in red. Each dot indicates a neuron, with the black line linking within neuron measures. Asterisks indicate the results of a Wilcoxon Sign-Rank test. In all plots: ns, not significant; †p<0.1; \*p<0.05; \*\*p<0.01; \*\*\*p<0.001; \*\*\*\*p<0.0001.

*Mouse behavioral detection is modulated by background contrast.*

Mice were initially trained in a simple version of the GO/NO-GO task, where they were required to lick in response a high volume target presented on go trials, and withhold licking on trials in which only background noise was presented (Figure 1b, 3a). Mice learned this task reliably, typically reaching criterion performance of 80% correct within 2-3 weeks in either contrast, and performed this task for many weeks (Figure 3b). Observed false alarm rates were significantly larger in high contrast compared to low contrast (Supplementary Figure 3a). This result could be due either to the increased similarity between the background noise and targets (ie. a greater overlap of their volume distributions) or because the task is harder in high contrast, which we discuss next. After mice performed above criterion for at least three sessions, they moved on to psychometric testing.

By varying the volume of presented targets, we collected psychometric curves for each mouse when targets were presented in low or high contrast. Figure 3c plots the performance of an example mouse overlaid with average psychometric fits, whereas Figure 3d plots the group averages for each contrast. Across all mice, we found that targets were easier to detect in low contrast, observing significantly lower detection thresholds in low contrast (Mean (*M*)= 7.47, standard deviation (*SD*) *=* 1.59) compared to high contrast (*M =* 13.20, *SD =* 2.54; paired t-test: *t(23)* = -8.71, *p* = 9.59e-9, Figure 3e). Next, we computed how sensitive mice were to changes in target volume by calculating the maximum slope of the psychometric curve for each mouse. In the full cohort of mice, we found no significant change in the psychometric slope between low and high contrast (Supplemental Figure 3b), which contradicted our normative model predictions (Figure 1e). Investigating further, we found that the range of target volumes had a significant effect on psychometric slopes. Namely, targets drawn from a narrow range resulted in steeper psychometric slopes than targets drawn from a wide range (Supplemental Figure 3c-f), regardless of the background contrast. To test the pure influence of contrast on psychometric slope, we tested a subset of mice with target volumes matched across the contrast conditions.

In a smaller group of mice (n = 7), we presented targets of the same volume in low and high contrast (Figure 3f). As in the full cohort, we found significantly lower target thresholds in low contrast (*M* = 6.80, *SD =* 2.73) compared to high contrast (*M =* 14.96, *SD =* 3.51; paired t-test: *t(3)* = -3.59, *p* = 0.036; Figure 3g) and significantly larger psychometric slopes in low contrast (*M* = 0.051, *SD =* 0.0068) compared to high contrast (*M =* 0.042, *SD =* 0.0064; paired t-test: *t(3)* = 3.42, *p* = 0.042; Figure 3h). These results demonstrate that background contrast has a substantial impact on detection threshold, and that mice are more sensitive to changes in the volume of targets presented in low contrast.

To assess behavioral adaptation to the background contrast, we presented targets at threshold volume at variable delays following the contrast transition. We observed behavioral time courses consistent with our normative model and with gain measured in auditory cortex: in high contrast, mice initially were able to detect targets with high accuracy which decreased over time, while in low contrast we observed increasing detection rates over time (Figure 3i). We next quantified the speed of behavioral adaptation in each contrast. First, we found that in high contrast, the first significant drop in performance occurred between the first two time points, while in low contrast the first significant increase in performance occurred between the first and third time points (Figure 3i, Table 1). Then, by fitting each mouse’s adaptation time course with an exponential function and comparing time constants for each contrast, we also found that behavioral adaptation is significantly faster in high contrast (*Mdn* = 0.023) compared to low contrast (*Mdn =* 0.13; Wilcoxon sign-rank test (n = 21): *rank*  = 547, *Z* = 2.75, *p* = 0.0060; Figure 3j). Taken together, these behavioral results confirm the three predictions from our model (Figure 1h): 1) Detection thresholds are lower in low contrast; 2) Psychometric slopes are higher in low contrast; 3) Performance decreases rapidly in high contrast, while increasing gradually in low contrast.

**Diagram

Description automatically generatedFigure 3.**

**a,** Schematic of GO/NO-GO paradigm.Task details as in Figure 1 and *Methods*. **b,** Behavioral performance for the contrast in which each mouse was first trained relative to the first session of task exposure (n=12 mice were first trained in low contrast, whereas n=13 mice were first trained in high contrast). Dots indicate a session, while the traces indicate a running average using a 7 day window. Blue dots and traces indicate sessions in which mice detected targets in low contrast (ie. after high-to-low contrast transitions), while red dots and traces indicate sessions in which mice detected targets in high contrast (ie. after low-to-high contrast transitions). **c,** Psychometric functions in low and high contrast for one mouse (mouse ID indicated in the upper left). Each dot indicates percent correct for a single volume in a single session, while the solid lines indicate average psychometric fits. Colors as in **b**. **d,** Psychometric functions averaged for n=25 mice in low and high contrast. Error bars indicate ±SEM over mice at individual target SNRs, while the solid lines are logistic function fits to the average performance per contrast. **e,** Psychometric thresholds per contrast. Each dot represents a mouse, lines connect performance of individual mice on low and high contrast sessions. Bars indicate the average threshold over mice, while error bars in black indicate threshold ±SEM over mice. **f,** Behavioral psychometric functions for n=4 mice tested using the same target volumes in each contrast. Dots with error bars indicate average performance ±SEM over mice as a function of contrast and target volume. Overlaid, dark-colored curves indicate psychometric fits to the averages, with the black dot indicating the average threshold. Light-colored lines indicate the psychometric curves of individual mice. Black, dashed horizontal line indicates chance (0.5) performance. **g,** Psychometric thresholds per contrast. Each dot represents a mouse, lines indicate where mice participated in both low and high contrast sessions. Bars indicate the average threshold over mice, while error bars in black indicate threshold ±SEM over mice. **h,** Psychometric slopes per contrast. Presentation as in **g**. **i,** Behavioral performance as a function of contrast and target time relative to the switch in contrast for n=21 mice. Dots with error bars indicate average performance ±SEM over mice. Solid curves indicate exponential function fits to the average over mice. Black, dashed vertical line indicates the contrast switch. Horizontal lines at the top of the plot indicate significant changes in performance between the first target presentation time and subsequent target presentation times, as assessed by Wilcoxon Sign-rank tests with false discovery rate correction for multiple comparisons. **j,** Average time constant of exponential fits in low and high contrast. Presentation as in **h**. Unless otherwise noted, blue markers indicate data where targets were presented in low contrast and red indicates data where targets were presented in high contrast. In all plots: ns, not significant; †p<0.1; \*p<0.05; \*\*p<0.01; \*\*\*p<0.001; \*\*\*\*p<0.0001.

*Auditory cortex is necessary for detection in noise.*

Previous studies have shown that while gain control is present in many areas across the auditory pathway, it is strongest in auditory cortex 15,24. As such, we hypothesized that auditory cortex was likely to be a key brain area supporting the detection of sounds in the presence of background noise, particularly when using background sounds known to modulate neuronal gain. To test whether auditory cortex is necessary for task performance, we inactivated auditory cortex using the GABA-A receptor agonist muscimol. We first validated that muscimol disrupts cortical coding of target sounds by applying muscimol topically to the cortical surface of awake, untrained mice while recording neuronal responses during passive playback of the behavioral stimuli (Supplementary Figure 4a). We first recorded baseline responses to all stimuli, then topically applied muscimol or saline, waited 30 minutes, and recorded stimulus responses again. After muscimol application, there was a marked decrease in neural responses to targets compared to the baseline recordings (Supplementary Figure 1b, left). Notably, in our saline control, we observed little to no change in neural responses after saline application (Supplementary Figure 4b, right). We next compared how contrast, volume and muscimol or saline application changed the responses during the pre- and post-application periods, finding that muscimol drastically reduced the firing rates between pre- and post-application periods, while saline moderately increased firing rates (Supplementary Figure 4c,d, Table 1). We speculate that the small increase in firing rate between pre- and post-saline application is due to changes in recording quality or due to neural drift over the ~1 hour long recording session, and note that the effect size of saline pre-post is very small (*η2* = 0.0046) when compared to the effect size of muscimol (*η2* = 0.38). We then used a 3-way ANOVA to compare the effects of muscimol, contrast, and target volume on target responses in the saline and muscimol recording sessions. We found a significant main effect of muscimol (*F*(1) = 322.65, *p* = 4.88e-67) and volume (*F*(6) = 15.48, *p* = 1.98e-17), but no main effect of contrast (*F*(1) = 0.39, *p* = 0.53), indicating nearly complete suppression of responses to both targets and noise in high and low contrast (Supplementary Figure 4e,f). These results confirmed that muscimol effectively disrupts the cortical coding of our behavioral stimuli.

To test whether inactivation of auditory cortex affects behavioral performance, we repeated the same experiments in behaving mice, administering muscimol or saline bilaterally through chronically implanted cannulae (Figure 4a). As observed in cortical activity, there was a profound decrease in the rate of responding to both targets (hits) and noise (false alarms) in both contrasts (Figure 4b). We quantified these effects on the psychometric curve using a 3-way ANOVA with cortical intervention (muscimol or saline), contrast, and target volume as factors. We found significant main effects of cortical intervention (*F*(1,307) = 278.63, *p* = 3.83e-44), contrast (*F*(1,307) = 4.39, *p* = 0.037) and volume (*F*(6,307) = 40.90, *p =* 7.54e-36). Post-hoc tests showed that muscimol application significantly decreased hit rates by 31% (95% CI: [28,35]), whereas hit rates were significantly elevated in low contrast by 4.9% (95% CI: [2.6,7.6]). Furthermore, we observed significant interactions between target volume and cortical intervention (*F*(6,307) = 14.11, *p* = 4.47e-14), and between target volume and contrast (*F*(6,307) = 2.97, *p* = 7.87e-3), but no significant interaction between contrast and cortical intervention. This pattern of results suggests that muscimol decreases performance in a manner not dependent on the background contrast.

To better quantify the specific effects of muscimol on psychometric performance, we extracted the response rate at the maximum target volume, false alarm rates, thresholds, and slopes of psychometric functions fit to each session (see Methods). Muscimol significantly reduced every measure of psychometric performance, with the exception of behavioral threshold (Figure 4c, Table 1). From these results, we can conclude that auditory cortex is necessary for performing target in noise detection, regardless of the background contrast.

A potential alternative effect of muscimol is a general loss of function that is not specific to hearing target sounds. To control for this, we devised another task where instead of detecting targets in noise (Figure 4d), mice detected targets in silence (Figure 4e). To ensure equivalency between the two tasks, we took the highest-volume target trials in the noise task (25dB SNR in high contrast; Figure 4d, left panel), and removed the background noise during the target detection period (Figure 4e, left panel). As such, mice detected the exact same targets as in the previous task, but without the flanking noise, allowing us to test whether auditory cortex is specifically required for detection in the presence of noise.

To assess psychometric performance in this new task, we modulated detection difficulty by attenuating the volume of each target. As observed previously, inactivation of auditory cortex hindered detection in high contrast noise (Figure 4d, right panel). However, cortical inactivation had little effect on psychometric performance in silence (Figure 4e, right panel). We quantified these effects on the psychometric curve using a 3-way ANOVA with cortical intervention (muscimol or saline), task (detection in noise or silence), and target volume as factors. We found significant main effects of intervention (*F*(1,181) = 62.83, *p* = 3.62e-13), task (*F*(1,181) = 6.82, *p* =

Diagram

Description automatically generated

**Figure 4.**

**a,** Setup schematic for muscimol application in behaving mice. *Bottom*: legend indicating colors used for each background condition. **b,** Behavioral psychometric functions during muscimol or saline application for n=4 mice. Dark solid lines and filled circles indicate average performance after saline injection. Dark dashed lines and open circles indicate average performance after muscimol injection. Light solid and dashed lines are psychometric curves from individual sessions. Error bars indicate ±SEM. across sessions. **c,** Behavioral performance metrics as a function of contrast and pharmacological intervention. Open circles indicate performance in individual sessions. Colored bars indicate average performance across sessions. Bars with low transparency and solid outlines are averages after saline application, while high transparency bars with dashed outlines are averages after muscimol application. Clockwise from the upper left, are plots of the max response rate, false alarm rate, psychometric threshold, and the maximum slope of the psychometric curve. **d,** *Left*: Example stimulus spectrogram for the target-in-noise detection task with the corresponding waveform below. The scale bar indicates 1 second, and the colorbar indicates the volume for each time-frequency bin (silence is black). *Right*: psychometric performance for n=2 mice in the target-in-noise task, with target volume on the abscissa and probability of responding on the ordinate. Filled circles and dark solid lines indicate average performance after saline injection and psychometric fits to the average. Red open circles and dark dashed lines indicate average performance after muscimol injection and psychometric fits to the average. Light red solid and dashed lines are psychometric curves from individual sessions. Errorbars indicate ±SEM across sessions. **e,** *Left*: Example stimulus spectrogram for the target-in-silence detection task with the corresponding waveform below. Time scale and volume scale as in **d**. *Right*: psychometric performance for n=2 mice (same mice as in **d**) in the target-in-silence task, with target attenuation relative to the highest volume target from the target-in-noise task on the abscissa and probability of responding on the ordinate. Black filled circles and dark solid lines indicate average performance after saline injection and psychometric fits to the average. Open circles and dark dashed lines indicate average performance after muscimol injection and psychometric fits to the average. Light grey solid and dashed lines are psychometric curves from individual sessions. Errorbars indicate ±SEM across sessions. **f,** Behavioral performance metrics as a function of task type (detection in noise or detection in silence) and pharmacological intervention. Formatting and metrics as in c (with the exception of response rate at threshold). Dark and light red bars indicate performance in the detection-in-noise task, with application of saline or muscimol. Dark and light grey bars indicate performance in the detection-in-silence task, with application of saline or muscimol. In all plots: ns*p*>0.1; †*p*<0.1, \**p*<0.05, \*\**p*<0.01, \*\*\**p*<0.001, \*\*\*\**p*<0.0001.

9.86e-3), and volume (*F*(6,181) = 46.16, *p* = 1.69e-32). Post-hoc tests showed that muscimol significantly reduced hit rates by 21.8% (95% CI [15.2,25.2]). Hit rates to targets presented in silence were significantly elevated by 6.7% relative to targets presented in noise (95% CI [1.7,11.6]). Furthermore, we found significant interactions between cortical intervention and task type (*F*(1,181) = 6.36, *p* = 0.013), intervention and volume (*F*(6,181) = 3.47, *p* = 2.98e-3), and volume and task type (*F*(6,181) = 8.47, *p* = 5.43e-8). Taken together, these results show that while cortical inactivation and the presence or absence of background noise both affect behavioral performance, these effects interact: muscimol has a larger effect on performance when background noise is present.

Previously, we found significant interactions between cortical intervention, task and target volume, suggesting that our manipulations of task and cortical activity affect the shape of psychometric curves in different ways. As before, we parameterized psychometric performance by fitting each session with a psychometric curve, and extracted the response rate at maximum target volume, false alarm rate, response rate at threshold volume, and psychometric slope. During the target in noise task, we found significant effects of muscimol on the response rates at maximum volume and threshold, a moderate effect on psychometric slope, and no effect on false alarm rate. However, muscimol application had no significant effect on any of these measures in the target in silence task (Figure 4f, Table 1). Taken together, these results demonstrate that cortex is specifically necessary for detection in the presence of noise, and has no significant effect on performance during detection in silence.

An additional alternative effect of muscimol is a general loss of the ability to lick. To assess this, we monitored the lick probability of the mice throughout the trial duration, and found that muscimol specifically reduced licking responses during the period where targets were presented (Wilcoxon rank-sum test: *T* = 337, *z* = -4.23, *p* = 2.34e-5; Supplemental Figure 4g, right panel of Supplemental Figure 4h). Mice also tended to lick immediately after the trial onset (Supplemental Figure 4i, green trace), but we found that the lick rates under muscimol and saline conditions were identical during this period (Wilcoxon rank-sum test: *T* = 528, *z* = 0.23, *p* = 0.81; Supplemental Figure 4h, left panel). These results suggest that muscimol does not impair the mouse’s ability to lick in general, but results in a specific deficit in licking in response to targets.

Combined, our results demonstrate that the auditory cortex is specifically required for detection in the presence of background noise. Our next goal was to test whether neuronal activity in AC is predictive of behavioral performance.

*Population responses to targets track individual behavioral performance.*

To better understand how representations in auditory cortex could give rise to behavior, we chronically recorded from populations of neurons in auditory cortex while mice performed the psychometric task (Figure 5a; n = 12 mice overall [n = 11 mice in low contrast sessions, n = 8 mice in high contrast sessions]). In the psychometric task where we varied target volume, many cortical neurons monotonically increased their firing rate with increased target volume (example neuron, Figure 5b; simultaneously recorded populations from two example sessions, Figure 5c).

To leverage our ability to simultaneously record from multiple neurons, we adapted a population vector approach28 to generate metrics of target from noise discriminability from population activity. To do this, we estimated the coding direction in the high dimensional space of simultaneously recorded neurons by subtracting population vector responses to noise alone from population vector responses to targets (population vectors were computed from the averaged activity from 0-100ms post-target onset). The resulting coding direction vector is the direction in high dimensional space between the average response to noise and targets (Figure 5d, left panel). This vector was trained on all but one trial, and the remaining held-out trial’s population response was projected along this coding direction to generate a single projection value. This procedure was repeated for all trials. We then grouped the trial projection values into trial distributions for each target volume, and compared them to noise trial distributions by estimating a criterion projection value that best predicted whether each was a target or noise trial29 (example projection value distributions from the recording session in Figure 5c (left panel) are plotted in Figure 5d (right panel)). Using this criterion, we then computed the accuracy of the neural population in discriminating targets from noise at each volume and at each contrast.

This population decoding method allowed us to estimate neurometric functions for direct comparison to the corresponding psychometric functions of each mouse (Figure 5e). On average, neurometric and psychometric functions were qualitatively similar, with neurometric functions exhibiting slightly lower thresholds, and shallower slopes (Figure 5f). We then selected sessions in which false alarm rates were low and population representations of targets were robust (see *Methods*).Multiple regression analysis revealed that neuronal thresholds and stimulus contrast significantly predicted the observed behavioral thresholds (*F*(1,16) = 16.20, *p*

**Diagram

Description automatically generatedFigure 5.**

**a,** Experimental setup for chronic ACtx recordings from behaving mice. **b,** Example spiking responses to targets and noise in low contrast during behavior. The top portion of the plot is a spike raster ordered by target volume. Colored bars indicate the target volume, grey bars indicate noise only trials. The bottom portion of the plot contains spike rates for each target condition, averaged over trials and smoothed with a 2ms standard deviation Gaussian kernel. *Inset:* Closed circles and the solid blue line indicate the area under the ROC curve (AUC) when discriminating noise from target responses across trials (activity was averaged from 0-100ms post target to compute AUC). Circle colors indicate the presented volume. The dashed horizontal line indicates chance performance (0.5). Error bars are the bootstrapped 95% confidence interval of the AUC value (see *Methods*). **c,** Neurograms of populations of simultaneously recorded neurons during a low contrast and high contrast session from the same mouse. Neurons are plotted along the ordinate, while target volume is plotted along the abscissa. Within each plot, the shade indicates the neural response to each target volume, with the average response to noise alone subtracted. White indicates no change in firing rate, blue/red indicate increases in firing rate relative to the noise response, and cyan indicates suppression below the noise response. Asterix indicates the responses of the neuron in panel **b**. **d,** Discriminating targets from noise using population responses. *Left:* schematic of coding direction analysis. In high dimensional neural space, noise trials are represented as a gray point-cloud, while target responses are represented as a blue point-cloud. The coding direction (CD) is the vector defining the average difference between these two point-clouds as indicated by the arrow. *Right:* trial distributions of projections along the coding direction for one session (session CA118-200707, as plotted in **c**. The blue distribution is the average projection value in a 40ms window after presentation of 20 dB SNR targets (indicated by arrow in panel **e**. The gray distribution is the average projection value in the same window during noise only trials. The vertical red line is the criterion which yielded the highest performance in predicting target presence across all trials. **e,** Example neurometric and psychometric curves. *Left:* Low contrast curves. Light blue circles and solid lines indicate psychometric performance and a logistic fit, respectively. Dark blue circles and solid lines indicate neurometric performance from the session plotted in the left panel of **c**. The horizontal dashed line indicates chance performance (0.5). The arrow indicates the neural performance computed from the distributions and criterion plotted in **d**. *Right:* High contrast curves from the same mouse for the session plotted in the right panel of **c**. **f,** Average psychometric and neurometric functions across mice. Light circles indicate average behavioral performance, dark red and blue circles indicate average neural performance. Light solid curves indicate logistic fits to average behavioral performance, while vertical lines indicate the fit thresholds. Dark solid lines indicate fits and thresholds for the neural data. The dashed vertical line indicates chance performance. Shades of blue and red indicate averages over low and high contrast respectively. **g,** Relationship between behavioral and neural thresholds. Each circle represents the average behavioral and neural threshold for each mouse for each contrast (as indicated by the circle fill color). Grey lines and shaded areas indicate the linear regression fit across contrasts, ±95% confidence interval. The solid black line indicates unity. Black asterisks indicate significant multiple regression fits to the data; within that model: grey asterisks indicate that neural thresholds are significant predictors of behavior, while red asterisks indicate that contrast is a significant predictor.  **H,** Relationship between behavioral and neural slopes. Appearance as in **g**. **i,** Population decoder performance in each contrast transition, as a function of target presentation relative to the transition (indicated by the dashed vertical black line at 0s). Ticks on the abscissa indicate average target time from the transition in milliseconds. Solid lines and circles indicate the percent correct performance of a target decoder after a switch to low contrast (blue) or high contrast (red). Errorbars indicate ±SEM over sessions. Horizontal lines indicate significant changes in performance between the first target presentation time and subsequent target presentation times, as assessed by Wilcoxon Sign-rank tests with false discovery rate correction for multiple comparisons. The span of the lines indicates the target times being compared, while the color of the lines indicates whether the test was performed within high contrast (red) or low contrast (blue). **j,** Adaptation time constants of exponentials fitted to the average neural decoder performance for each mouse in each contrast. Blue and red circles indicate the adaptation time constants from neural populations for each mouse in low and high contrast respectively. Solid black lines indicate time constants from the same mouse. In all plots: ns*p*>0.1; †*p*<0.1, \**p*<0.05, \*\**p*<0.01, \*\*\**p*<0.001, \*\*\*\**p*<0.0001.

= 1.44e-4, *R2* = 0.67). Investigation of the individual predictors revealed that neuronal thresholds were significantly predictive of behavioral thresholds (*β* = 0.51, *p* = 0.027) while contrast was only marginally predictive (*β* = 2.55, *p* = 0.080; Figure 5g). We then tested whether neurometric and psychometric thresholds were similarly affected by background contrast using a two-way ANOVA, with thresholds as the dependent variable and threshold measure (psychometric or neurometric) and contrast as independent variables. We found a main effect of contrast (*F*(1) = 29.30, *p* = 5.00e-6), but no main effect of threshold measure (*F*(1) = 0.02, *p* = 0.89) or interaction between measure and contrast (*F*(1) = 0.04, *p* = 0.85), which demonstrates that behavioral and neural thresholds were similarly affected by background contrast. As expected, post-hoc t-tests found no difference between neural and behavioral thresholds (mean change [95% confidence interval] *=* 0.19 [-1.38, 1.76] dB SNR, *p* = 0.81), and that low contrast significantly decreased thresholds relative to high contrast (-4.77 [-6.34, -3.19] dB SNR, *p* = 5.43e-7). Taken together, these results demonstrate that population thresholds in auditory cortex are highly predictive of behavioral thresholds, and both behavior and neuronal thresholds are modulated by contrast as predicted by gain control.

Next, we applied the same statistical analysis to psychometric and neurometric slopes, finding that neurometric slopes and contrast significantly predicted behavioral slopes (multiple regression: *F*(1,16) = 12.00, *p* = 6.63e-4, *R2* = 0.60). Examination of the individual predictors again showed that neuronal slopes were significantly predictive (*β* = 0.67, *p* = 0.0018), while contrast was not predictive (*β* = 0.010, *p* = 0.16; Figure 4h) of behavioral slopes. We next examined whether neurometric and psychometric slopes (Figure 4h) differed using a two-way ANOVA, as described above, and found a significant main effect of slope measure (*F*(1) = 5.88, *p* = 0.021) and contrast (*F*(1) = 8.31, *p* = 0.0068), but no interaction between the two (*F*(1) = 0.18, *p* = 0.67). Post-hoc testing revealed that neuronal slopes were significantly shallower than neuronal slopes (-0.015 [-0.027, -0.0024] PC/dB, *p* = 0.021). We also found that low contrast slopes were significantly shallower overall (-0.018 [-0.030, -0.0052] PC/dB, *p =* 0.0068), conflicting with our previous behavioral findings (Figure 3h). However, if we averaged only the sessions where mice were presented matched target volumes in low and high contrast, we find a significant main effect of contrast (*F*(1) = 5.98, *p* = 0.028), and, as before, a significant main effect of slope measure (*F*(1) = 10.62, *p* = 0.0057; Supplementary Figure 5a). In summary, we find that while neuronal sensitivity was predictive of behavioral slopes, neurometric curves tended to have shallower slopes. The observed discrepancy between behavioral and neural slopes may be due to the limited number of cells we recorded from, which would introduce noise in our population discriminability metric. However, when matching the target volumes presented in each contrast, we observe that neural and behavioral slopes are shallower in high contrast, as expected from our previous results.

Taken together, these results demonstrate that parameters of neurometric and psychometric functions are affected by contrast in similar ways, consistent with our normative model. We also find that individual variation in psychometric performance is predicted by population activity in auditory cortex, independently of the effect of contrast.

*Asymmetric neural adaptation to targets.*

Using the same population decoding approach described above, we measured how cortical discriminability of target from noise trials evolved as a function of time and contrast during sessions where mice heard targets at threshold volume at different offsets relative to the contrast switch. While, neuronal adaptation was qualitatively similar to behavioral performance and model predictions, we quantified adaptation using the procedure applied to the behavioral time courses described above (Figure 3i). As in behavior, we found that in high contrast the first significant drop in performance occurred between the first two target times, while the first significant drop in low contrast occurred between the first and third target times. In high contrast, we also observed that discriminability of later targets (>0.25s) increased over time, becoming significantly larger than initial target discriminability ~1s after the contrast switch (Figure 5i, Table 1).Finally, to quantify the speed of neural adaptation, we fit the average neural discrimination time course for each mouse with an exponential function (*Methods*). Consistent with the behavioral results, we found asymmetric adaptation in the neural responses, with larger adaptation time constants in low contrast (*Mdn* = 0.14, *IQR* = 0.21) relative to high contrast (*Mdn* = 0.033, *IQR =* 0.16;Wilcoxon sign-rank test (n = 8): *rank*  = 28, *Z* = nan, *p* = 0.016).

*Cortical gain tracks individual behavioral performance.*

Our behavioral results and model provide strong evidence that gain control in the auditory system shapes patterns of behavioral performance. To directly assess the role of gain control in auditory cortex in shaping

behavior, we leveraged the design of our background sounds to estimate the gain of cortical neurons using a linear-nonlinear model (LN). Briefly, we estimated the spectro-temporal receptive fields (STRFs) of individual cortical neurons using normalized reverse correlation and convolved them with the stimulus spectrogram to generate linear predictions of cortical activity (Figure 6a, panels 1-3). We could then compare the linear prediction to the observed firing rate elicited by the background to estimate neural gain control by fitting the nonlinearity of each neuron in high and low contrast using an exponential function (GC model, Figure 6a, panel 4). For comparison, each neuron was fit in a similar fashion using a single, static nonlinearity across the two gain conditions (static model, Figure 6a, panel 4). In either model, the gain of the neuron is defined as the shape parameter of the exponential fit to the nonlinearity (*Methods*). This modeling strategy allowed us to estimate cortical gain during different periods in the task, and assess how gain is related to behavioral performance.

Figure 6b-d plots data from a representative neuron recorded during behavior. Within each session, the background noise was randomly drawn from one of five frozen noise scenes, which allowed us to observe stimulus aligned spike patterns across repeats of each scene (Figure 6b, spike raster). The estimated STRF for this example unit is shown in Figure 6c, along with the nonlinearities estimated for low and high contrast in Figure 5d, and the GC model fit to the data in the bottom of Figure 6b. We first compared the cross-validated performance of a model with a static nonlinearity versus a gain control model, and observed a significant enhancement in the model’s correlation to the observed spikes when modelling gain control (Median (*Mdn*)*:* 0.815), relative to the static model (*Mdn*: 0.645; Wilcoxon sign-rank test (n = 1,535 neurons): *rank* = 64,346, *Z* = -30.23, *p* = 9.26e-201; Figure 5e). After pooling all of the neurons recorded across all mice and sessions, and including only neurons with high reliability in both contrasts (defined as units with noise-to-signal ratio (NR) < 100 in low and high contrast, see *Methods*), we observed significantly higher gain in low contrast (*Mdn*: 0.093) than in high contrast (*Mdn*: 0.050; Wilcoxon sign-rank test (n = 1,535 neurons): *rank* = 972,742, *Z* = 22.07, *p* = 6.46e-108; Figure 5f). These results demonstrate that models incorporating contrast gain control are better predictors of cortical activity, and confirm previous reports of robust gain control in ferret and mouse auditory cortex14,17.

We next assessed whether gain in auditory cortex reliably predicts how well a mouse is able to detect targets in each contrast. To do so, we averaged the gain of target-selective neurons (see *Methods*) during the target period of the task (ie. the time window after the contrast switch) for each mouse. We then compared the gain for each mouse to the behavioral thresholds and slopes collected in the psychometric task. We found a significant negative relationship between cortical gain and behavioral threshold across contrasts (single linear regression: *F*(1,12) = 24.2, *p* < 0.001, *R2* = 0.67), suggesting that increased gain yielded greater sensitivity to lower target volumes. However, we didn’t observe this relationship when only including low contrast sessions (*F*(1,7) = 1.42, *p* = 0.27, *R2* = 0.17) or high contrast (*F*(1,3) = 2.02, *p* = 0.25, *R2* = 0.40), so we cannot definitely conclude that contrast-independent fluctuations in gain predict behavioral thresholds (Figure 5g). We conducted the same analysis between gain and psychometric slopes, and found significant positive relationships across contrasts (single linear regression: *F*(1,12) = 12.0, *p* = 0.005, *R2* = 0.50) and within low contrast (*F*(1,7) = 6.96, *p* = 0.034, *R2* = 0.50), suggesting that individual differences in cortical gain influence behavioral sensitivity to changes in volume, independently of contrast gain control (Figure 6f).

**Graphical user interface, diagram

Description automatically generatedFigure 6.**

**a,** Schematic of the linear nonlinear models fit to behavioral recordings. 1) Schematic of a spectro-temporal receptive field (STRF). 2) Example stimulus spectrogram of low and high contrast. 3) The gray trace is the filter response when convolving the STRF with the spectrogram. The black trace is the observed spike rate during the same stimulus period. 4) Schematized nonlinearities fit separately to low and high contrast periods in a gain control (GC) model or fit to all data in a static model. **b,** Example background-locked responses from a well-tuned cortical unit across the trial duration. The top portion of the plot is a spike raster sorted by the frozen noise scenes (Scene 1-5) of the background. The bottom portion of the plot is a PSTH of the observed spiking, binned every 25ms (black trace). The colored traces are the GC model predictions in each contrast (red trace uses the red nonlinearity in **d**, blue trace uses the blue nonlinearity in **d**). **c,** STRF for this example neuron. STRF values are indicated by the colorbar. **d,** Estimated nonlinearities for this example neuron. Points indicate the mean observed firing rate (ordinate), binned according to observed filter prediction values (abscissa). Solid lines indicate exponential function fits to the underlying points. Each line is a fit to the test set in 10 cross-validation runs (see *Methods*). **e,** Correlation coefficients between the observed trial-averaged spike rate and the model prediction for the static model and the gain control model. Each dot is the average correlation across 10 cross-validation folds for each neuron, where black dots are high stimulus locking neurons with low noise-to-signal ratios (NR < 100) and grey dots are neurons with low stimulus locking (NR > 100). For the remaining figures, only neurons with NR < 100 are included. The solid red line indicates unity. The red “x” indicates the median correlation for each model. **f,** Gain control in auditory cortex during the task. Each histogram is the distribution of gain values in high and low contrast across neurons with NR below 100, recorded during behavior. Dashed vertical lines indicate the median of each distribution. **g,** Relationship between gain and behavioral threshold. Each circle represents the average gain and behavioral threshold for each mouse for each contrast (as indicated by the circle fill color). Gain values were averaged over target selective neurons. Grey lines and shaded areas indicate the linear regression fit across contrasts, ±95% confidence interval. **h,** Relationship between gain and behavioral slope. Appearance as in **g**. In all plots: ns*p*>0.1; †*p*<0.1, \**p*<0.05, \*\**p*<0.01, \*\*\**p*<0.001, \*\*\*\**p*<0.0001.

**Discussion**

On a daily basis, we navigate through many auditory environments, each defined by different statistical properties. The dynamic range, or contrast, of acoustic inputs poses a challenge to the auditory system, which is composed of neurons with limited dynamic range in their response. The efficient coding hypothesis predicts that as acoustic contrast shifts, neurons throughout the auditory pathway adjust their sensitivity, so as to match the dynamic range of their response to that of the stimulus distribution30. This process allows auditory neurons to encode sound-level information within each contrast, despite changes in the dynamic range. Multiple studies have demonstrated that indeed, neurons throughout the auditory pathway exhibit contrast adaptation14–17. Whereas recent work has demonstrated a link between efficient cortical codes and human psychophysical performance 24, whether neuronal contrast adaptation plays a role in auditory perception has not been previously examined simultaneously with behavior. In this study, we directly linked neuronal contrast gain control to auditory behavior through the use of a theoretical model of efficient coding, behavioral psychophysics, and simultaneous manipulation and recordings of cortical activity.

The goal of this study was to test the hypothesis that efficient coding at the neuronal level in the auditory cortex shapes auditory behavior. To tackle this complex question, we first developed a normative framework22 that allowed us to make specific predictions for behavioral performance as expected by contrast gain control. The model predicted that (1) The detection threshold of a target in low contrast noise should be lower than in high contrast; (2) Sensitivity to changes target volume should be increased in low contrast relative to high contrast; and (3) Upon a shift in the background noise contrast, detection improves slowly after a switch to low contrast, but decreases rapidly after a switch to high contrast (Figure 1). First, we showed that gain adaptation in auditory cortex followed the predictions of the efficient coding model using a novel GLM. Then, to test the model predictions behaviorally, we trained mice to detect a target in background noise, as noise contrast shifted from low to high or from high to low. As predicted by the normative model, mice had lower detection thresholds and were more sensitive to changes in target volume during low contrast. Over time, we observed a decrease in tone detection after a switch to high contrast, and an increase in tone detection after a switch to low contrast. Behavioral adaptation was faster in high contrast, as compared to low contrast backgrounds, in agreement with our model and previous theoretical models19 (Figure 3). We furthermore found that AC is necessary specifically for this detection-in-noise task (Figure 4). When recording in AC, we observed that population thresholds for sound detection were greater in high than in low contrast and that as neurons adapted to transitions in background contrast, the time course of neuronal target discriminability adapted similarly to the behavior and model predictions (Figure 5). Additionally, a direct comparison of neurometric and behavioral performance revealed a significant correlation between the thresholds and slopes of the psychometric curves (Figure 5). Finally, we found correlations between cortical gain and behavioral thresholds and slopes (Figure 6), supporting our hypothesis that efficient coding at the neuronal level predicts auditory behavior.

*The role of cortex in behavior.*

The role of auditory cortex in auditory behavior has been subject of debate. A number of prior studies found that auditory cortex was not required for relatively simple behavioral tasks such as frequency discrimination or detection31,32. Rather, many studies found that auditory cortex is primarily involved in more complex behaviors, such those requiring temporal expectation33, localization34, or discrimination of more complex sounds35–37. Consistent with previous findings38, we found that AC inactivation selectively impaired the detection of targets in a noisy background, but did not impair detection of targets in silence (Figure 4). Furthermore, on a subject-by-subject basis, neuronal activity in AC was correlated with behavioral performance of the subject (Figures 5, 6). This set of results establishes that AC is necessary for the detection of targets in background noise and supports the more general notion that AC is required for more complex auditory tasks, but is not required for simpler tasks.

While the previous work demonstrates the necessity of auditory cortex in behavioral performance, the brain areas and mechanisms supporting the transformation from stimulus to decision are an active field of study39,40. By recording during the task, we were able to leverage behavioral variability to show that behavioral performance covaried with representations of targets within small neural populations (Figure 5), and with cortical gain (Figure 6). There is a large body of literature relating cortical codes to behavioral variability: early studies in the visual system suggested that information from relatively small numbers of neurons was sufficient to match or outperform animal behavior in psychophysical tasks41–43 and that behavioral choice can be predicted from activity in sensory areas29,43. These accounts suggest that variability in bottom-up sensory encoding drives the variability in behavioral output. However, more recent work suggests that variability in sensory areas is driven by top-down influences44–47, which are modulated by attention and learning48–51. Interestingly, a recent study imaging tens of thousands of neurons in the visual cortex demonstrated that cortical representations had higher acuity than mouse behavioral output, yet did not correlate with behavioral performance, suggesting that perceptual discrimination depends on post-sensory brain regions52. Our results suggest that bottom-up adaptation to stimulus statistics shapes behavioral output, as we observed very stereotyped patterns of behavioral adaptation (Figure 3) qualitatively consistent with the predictions of our efficient coding model (Figure 1), patterns of gain adaptation in auditory cortex in the absence of behavior (Figure 2), and patterns of target-driven activity in auditory cortex (Figure 5). Indeed, there have been other studies demonstrating that individual differences in sensory-guided behaviors are reflected in cortical activity53,54, are bidirectionally modulated by cortical manipulation55,56, and can be predicted from tuning properties in auditory cortex57,58. While our results cannot rule out top-down input as the causal driver of sensory decisions, they do support the notion that the sensory information upon which decisions are made is shaped by neuronal adaptation, which thus affects behavioral outcomes.

*Adaptation in the auditory system.*

Neurons throughout the auditory system adapt to the statistics of the acoustic environment, including the distribution of stimuli over time59,60, more complex sound patterns26,61, and even ongoing behavioral and attentional demands62–67. Inspired by the latter studies, we intentionally designed our stimuli using unbiased white-noise backgrounds, which allowed us to fit encoding models to our data. One particular benefit of these models is the ability to define abstract relationships between stimulus variables that are difficult to quantify directly from the raw data. Using these methods, we focused on contrast gain control as a fundamental statistical adaptation that relates to efficient coding14,17,18,24. Contrast gain control is present at multiple stages in the auditory system, increasing in magnitude from the inferior colliculus to the auditory thalamus and auditory cortex15,24 with slower adaptation speeds in auditory cortex than subcortical areas24. In this study, we developed a novel form of Poisson GLM that allowed us quantify the contribution of multiplicative interactions between the stimulus and stimulus contrast to the activity of neurons in auditory cortex. Reasoning about the fitted parameters of the model, we were able to accurately estimate neuronal gain as a function of time. This approach allowed us to verify that gain adaptation in auditory cortex is asymmetric (Figure 2), as predicted from efficient coding theory19 and as shown in previous work25.

Furthermore, we found that behavioral detection of targets adapted asymmetrically (Figure 3), as predicted by our efficient coding model (Figure 1) and from the gain time courses observed in AC of non-behaving mice (Figure 2). This suggested that cortical gain adaptation shaped task performance. To test this prediction, we leveraged linear-nonlinear to estimate cortical gain during the task, finding that cortical gain was predictive of behavioral performance (Figure 6). In general, the combination of carefully designed task stimuli and simultaneously recorded neural activity allowed us to probe gain as a neuronal mechanism underlying behavior, and highlights the utility of encoding models for linking neural codes to behavioral performance68–71.

*Cellular mechanisms of gain control.*

Whereas this study demonstrates the necessity of auditory cortex for detection in noise, the neuronal mechanisms driving contrast gain adaptation at a cellular level remain unclear. Additionally, while we observed theoretically optimal asymmetric adaptation to changes in contrast, the neural circuits driving these temporal asymmetries are unknown. In the current study, we have likely recorded from a mixed population of excitatory and inhibitory neurons. Different inhibitory neuronal subtypes exhibit specific roles in adaptation72,73. Although specific inhibitory neuronal subtypes facilitate divisive or subtractive control of excitatory responses in visual74,75 and auditory cortex76,77, the role of these interneurons in contrast gain control has been inconclusive18. By combining previously mentioned optogenetic methods with behavioral tasks, future studies may explore and test the specific role of inhibitory neurons in gain control and behavior.

*The missing link between efficient coding and behavior.*

Combined, our results develop a framework and provide support for the role of efficient neuronal coding in behavior. The efficient coding hypothesis has emerged as one of the leading principles in computational neuroscience that has shaped our understanding of neuronal coding, architecture and evolution30,78–81. Extensive prior research found that human behavior follows principles of efficiency23,24. Our work now provides a framework for linking the principles of neuronal coding with behavioral performance. Additionally, we have introduced a novel form of Poisson GLM designed to detect multiplicative interactions between presented stimuli and other variables. While in this study we focused on the multiplicative effect of contrast, this approach could in theory be applied to any other time-varying signal that modulates neuronal gain, such as movement82,83, arousal84,85, or targeted experimental interventions74–77. In summary, we expect the theoretical frameworks and modelling methods applied here to have broad utility in the study of neuronal adaptation, a fundamental function of the nervous system.

**Methods**

*Animals*.

All experiments were performed in adult male (n = **xxx**) and female (n = **xxx**) mice (The Jackson Laboratory; age 12-15 weeks; weight 20-30g; **STRAINS**, etc.), housed with, at most, five mice per cage, at 28°C on a 12-h light:dark cycle with food provided ad libitum, and a restricted water schedule (see *Water Restriction*). All experiments were performed during the animals’ dark cycle. All experimental procedures were in accordance with NIH guidelines and approved by the Institutional Animal Care and Use Committee at the University of Pennsylvania.

*Surgery*.

Mice were anesthetized under isoflurane (1-3%). Prior to implantation, all mice were administered subcutaneous doses of buprenorphine (Buprenex, 0.05-0.1 mg/kg) for analgesia, dexamethasone (0.2 mg/kg) to reduce brain swelling, and bupivicane (2 mg/kg) for local anesthesia. In mice undergoing electrophysiological experiments, two ground screws attached to ground wires were implanted in the left frontal lobe and right cerebellum, with an additional skull screw implanted over the left cerebellum to provide additional support. A small craniotomy was performed over the target stereotactic coordinates relative to bregma, -2.6mm anterior, -4.3mm lateral. Either custom 16-channel microdrives, 32-, or 64-channel shuttle drives (cite) holding tetrode bundles of polyimide-coated nichrome wires were chronically implanted over auditory cortex, and tetrodes were lowered 800um below the pial surface. The exposed tetrodes were covered with GelFoam (Pfizer) or sterile silicone lubricant and sealed with Kwik-Cast (World Precision Instruments). The plastic body of the microdrive and a custom stainless-steel headplate were secured to the skull using dental cement (C&B Metabond) and acrylic (Lang Dental). Mice undergoing only behavioral experiments were implanted with two skull screws in the cerebellum, and a headplate was mounted on the skull as previously described. An antibiotic (Baytril, 5mg/kg) and analgesic (Meloxicam, 5mg/kg) were administered daily (for 3 days) during recovery.

*Water restriction*.

Following surgical recovery (3 days post-operation), each mouse’s weight was monitored for three additional days to establish a baseline weight. Over the next seven days, mice were water deprived, beginning with a daily ration of 120uL/g and gradually decreasing their ration to 40-50uL/g. During the task, if mice did not receive their full ration, the remainder of their ration was provided in their home cage. Mouse weight relative to baseline was monitored during all stages of water restriction. Additional health signs were used to determine a health score and subsequent treatment plan if a mouse lost more than 20% of baseline weight, as described by previously published methods86 and approved by the Institutional Animal Care and Use Committee at the University of Pennsylvania.

*Behavioral apparatus*.

During the GO/NO-GO task, the mouse was head-fixed in a custom-built, acoustically isolated chamber. A capacitive touch sensor (AT42QT1010, SparkFun) soldered to a lick spout monitored lick activity. Water rewards were dispensed from a gravity fed reservoir, controlled by a solenoid valve (161T011, Neptune Research) calibrated to deliver approximately 4-5uL of water per reward87. Low-level task logic – such as lick detection, reward and timeout delivery, and task timing intervals – was directly controlled by an Arduino Uno microprocessor running custom, low-latency software routines. High-level task logic, such as trial randomization, stimulus buffering and presentation, and online data collection and analysis were controlled by custom MATLAB (Mathworks) software communicating with the Arduino over a USB serial port. Digital waveforms were converted to analog signals via a soundcard (Lynx E44, Lynx Studio Technology, Inc.) or a National Instruments card (NI PCIe-6353) and delivered through an ultrasonic transducer (MCPCT-G5100-4139, Multicomp). The transducer was calibrated to have a flat frequency response between 3 kHz and 80 kHz using a 1/4-inch condenser microphone (Brüel & Kjær) positioned at the expected location of the mouse’s ear, as described previously88,89. During electrophysiological recording sessions, licks were detected using an optical interrupt sensor (EE-SX771, Omron Automation), to prevent lick-related electrical artifacts introduced by contact with the capacitive sensor.

*Behavioral timeline*.

Each mouse underwent four stages in the behavioral task: 1) water restriction and habituation, 2) behavioral training, 3) psychometric testing, and, 4) offset testing. During the induction of water restriction, mice were simultaneously habituated to head-fixation in the behavioral chambers and receiving water through the lick spout, by providing a water reward for any licks separated by more than 2 s. After the mouse began to receive its entire ration by licking in the booth, behavioral training was initiated (typically 1 week). Each mouse was initially trained and tested in one contrast condition (see *Stimuli*), with the initial training condition counterbalanced across mice. Behavioral performance was monitored during training, and mice were considered trained after completing at least three consecutive sessions with over 80% percent correct (~2-3 weeks). After completing training, behavioral thresholds were measured during at least three sessions in which psychometric stimuli were presented (see *Stimuli*). After estimating the behavioral threshold for each mouse, offset stimulus sets were generated using threshold-level targets. After completion of at least three sessions in the offset task, each mouse was then retrained on the remaining contrast condition. Upon reaching the training criterion of 80% in the new contrast condition, mice were then tested in the psychometric and offset tasks as previously described. For mice in electrophysiological experiments, this sequence of training and testing was continued until the recording site yielded less than three units, or until the mouse stopped performing in the task.

*Stimuli*.

All stimuli were created in MATLAB and sampled at 192 kHz or 200 kHz and 32-bit resolution. A set of dynamic random chords (DRCs) were created with different contrasts, similarly to those described in previous studies14,17,24. This stimulus was used 1) to measure the spectrotemporal receptive fields of neurons by fitting a linear-nonlinear model, and 2) to modulate the gain of auditory neurons by manipulating stimulus contrast. To construct a DRC, amplitude modulated pure tones were generated at multiple frequencies and then superimposed to create a chord. In some experiments, 34 frequencies were sampled between 4 and ~40kHz in 1/10 octave steps, in the remaining experiments, 33 frequencies were sampled between 4 and 64kHz in 1/8 octave steps. The amplitude envelope of each tone was generated as follows: every 25 ms, amplitudes for each frequency were sampled from a uniform distribution with a mean of 50 dB and a width of ±5 dB in low contrast or ±15 dB in high contrast. Between each 20 ms chord, the amplitude envelope of each frequency band was linearly ramped over 5 ms to the amplitude value for the next chord, such that the total duration of each chord and its ramp was 25 ms. To synthesize the stimuli, amplitude envelopes were multiplied by a sine wave of their respective frequencies, and summed to produce the final waveform. Each time a set of DRCs was generated, 5 unique random number generator seeds were used to restrict the background noise to 5 distinct scenes (see raster in Figure 6 for an example of spike-locking to the repeated scenes).

In all stages of behavioral training and testing, stimuli created for each trial consisted of a DRC background containing a change in contrast, and the presence or lack of a target at a delay after the change in contrast. Each trial was initialized with 3 seconds of DRC noise of one contrast, followed by a switch to the other contrast. Targets consisted of a fixed chord composed of 17 frequencies pseudo-randomly sampled from the frequencies contained in the DRC background, such that the target frequencies were uniformly distributed across the frequency range of the background. To add targets to the background noise, the target amplitude at each target frequency was simply added to a single chord in the amplitude envelope of the background, and ramped as described previously; this procedure ensured that target timing was perfectly aligned to changes in the background noise, removing asynchronous timing cues that could be used by the animal to detect the target. Target amplitudes are described in values of signal-to-noise ratio (SNR) relative to the average level of the background noise (ie. a 50 dB target embedded in 50 dB noise would have an SNR of 0 dB). We note that because the targets only contained power in half of the frequency bands used to construct the noise background, target SNRs were typically above 0 dB (see Table 2 for a breakdown of SNRs used across all mice). In all trials, targets were embedded after a change in the background contrast, with a delay and volume dependent on the current training or testing stage (see *Behavioral Task*).

*Efficient coding model.*

We simulated a model neuron that encodes incoming stimuli via an adapting neural nonlinearity. Stimuli were drawn from a Gaussian distribution whose mean was fixed over time but whose standard deviation could switch over time between a low and a high value ( and , respectively). At each time , a stimulus was drawn from the distribution , transformed via a saturating nonlinearity of the form , distorted by Gaussian noise with variance , and finally discretized into discrete levels to generate a response . This discrete response was linearly decoded to extract an estimate of the current stimulus: . The recent history of stimulus estimates was used to update an estimate of the underlying standard deviation: . The estimate was then used to select the parameters of the encoder () and the decoder () on the next timestep. The encoding and decoding parameters were chosen to minimize the expected error in decoding stimuli given the neuron’s current estimate of the underlying standard deviation: [1,2].

The parameters of the encoder and decoder were adapted based on a background stimulus with a mean that was fixed over time and a standard deviation that switched between low and high values and , respectively. We used this adapting nonlinearity to determine how well this model neuron could discriminate target stimuli from this background. Target stimuli were sampled from a Gaussian distribution with a fixed mean and with a variance that was scaled in proportion to the variance of the background ( and , respectively). At each timestep, we computed the Bhattacharyya coefficient () of the response distributions produced by background versus target stimuli: . We used as our measure of discriminability.

We simulated the behavior of this model using a background “probe” stimulus whose standard deviation switched every timesteps. We simulated cycles of this probe stimulus, where each cycle consisted of timesteps in the low state, followed by timesteps in the high state. This yielded timeseries of the gain and offset of the adapting nonlinearity, as well as distributions of the neural response to the background and target stimuli at each timepoint following a switch in standard deviation. We averaged the gain and offset across cycles to obtain the average properties of the encoder at each timepoint following a switch. We used the distribution of responses to target and background stimuli, measured across cycles, to compute the discriminability at each timepoint following a switch.

All simulations were performed with the following values: **, ,,,,,,.**

*Behavioral task*.

We employed a GO/NO-GO task to measure the ability of mice to detect targets in noise. In this task, each trial consisted of a noise background with a contrast shift, along with the presence or absence of a target after the change in contrast. Mice were trained to lick when they detect a target (hit), or to withhold licking in the absence of a target (correct reject). This behavior was reinforced by providing a 4-5 uL water reward when the mouse performed a hit, and by initiating a 7-10 s timeout when the mouse licked in the absence of a target (false alarm). Any licks detected during the timeout period resulted in the timeout timer being reset. In a subset of mice, we introduced an additional trial abort period coincident with the first part of the contrast background, before the contrast switch. Any licks detected in this abort period resulted in the trial being reset after a 7-10 s timeout, until the mouse withheld from licking during this period. In this task, misses and correct rejects were not rewarded or punished. Trials were separated by a minimum 1.5s inter-trial-interval (ITI). To discourage spontaneous licking, licks were monitored during this period, and if any licks occurred the ITI timer was reset.

All of the behavioral tasks varied the timing of the target relative to the contrast shift, which required a method for estimating hit rates and false alarm rates at different times during each trial, and to reward and punish the animal during these times in an unbiased manner. To approach this issue, we considered licks as responses only during a 1 s response window after a target presentation in the trial (eg. if a target was presented 500 ms post-contrast-switch, the response window persisted from 500 to 1500 ms post-contrast-switch). To apply this method to noise-only trials, in which no targets were presented, we considered noise trials as target trials containing infinitely small target amplitudes. For each noise trial, we assigned a response window with equiprobable delay matched to the target conditions and considered only licks within those “target” response windows. Thus, over the course of a session, we randomly sampled lick probabilities in noise trials during the same temporal windows as those licks considered during target trials. Using this scheme, we treated target and noise trials identically, and estimated hit rates and false alarm rates over time in an unbiased manner.

Each mouse performed three stages in the behavioral task: training, psychometric testing, and offset testing. During the training task, trials consisted of two types, noise trials or target trials presented with equal probability. To facilitate learning, we selected target SNRs at the highest end of the range described previously: in low contrast training sessions, targets were 16 dB SNR, and in high contrast training sessions, targets were 20 dB SNR. To prevent response bias as a function of target timing, we randomly varied the target delay between 250, 500, 750 and 1000 ms after the contrast change in each trial. During the psychometric testing task, there were 7 trial types consisting of noise trials and target trials spanning six different SNRs (Table 2). Based on behavioral piloting, we presented high SNR trials with a greater probability, to prevent mice from giving up during the task. In low and high contrast psychometric sessions, the probability of a noise trial was 0.4, the probability of the four lowest target SNRs was 0.05 each, and the probability of the two highest target SNRs was 0.2 each. As in training, target timing was varied randomly between 250, 500, 750 and 1000 ms after the contrast change in each trial. After completing at least three sessions of the psychometric task, stimuli were generated for the offset testing task. This task consisted of 15 unique trial types: 3 target volumes (noise trials, threshold target trials, and high SNR target trials), and 5 target delays relative to the contrast change (25, 75, 225, 475, 975 ms delay). Threshold target amplitudes were determined individually for each mouse by fitting psychometric curves averaged over several sessions with a psychometric function, and extracting the volume at which the slope of the psychometric curve was steepest (see *Behavioral and Neural Detection Performance*). Based on behavioral piloting, noise trials, threshold target trials, and high SNR target trials were presented with probabilities of 0.4, 0.2, and 0.4, respectively. Target delay on each trial was selected with equal probability. In all behavioral stages, trial order was pseudorandomly generated, such that there were no more than three target or noise trials in a row.

A subset of mice (n = 2), were presented targets in the presence of noise (Figure 4). To generate this stimulus set without changing the basic structure of the task or stimuli, we simply took the spectrograms of all stimuli containing 25 dB SNR targets from the low-to-high contrast stimulus sessions, and set the stimulus power flanking each target to zero. This manipulation was only performed in the target period, and the low contrast adaptation period of the trials remained the same. Thus, the targets and adaptation periods were identical to those presented in the target in noise task (see Figure 2e). To vary the difficulty of the task, the volume of the target was manipulated using the following values: -75, -60, -45, -30, -15, and 0 dB attenuation relative to the 25 dB SNR target. Mice were previously trained in the target in noise task prior to performing the target in silence task. Before psychometrically varying the target attenuation, mice were trained in the new task to criterion performance. Mice generalized very rapidly to the new task, reaching 97% and 94% training accuracy on the first day of exposure to targets in silence (performance for mouse CA124 and CA125, respectively).

*Chronic muscimol application*.

A separate cohort of mice (n = 4) were bilaterally implanted with 26 GA guide cannulae (PlasticsOne, C315GMN-SPC mini, cut 5 mm below pedestal) in auditory cortex. The surgery was performed as described previously with the following modifications. After the skull was leveled using a stereotax, two small craniotomies were performed -2.6 mm anterior, ±4.3 mm lateral from bregma, over auditory cortex. The guide cannulae along with dummy infusion cannulae (PlasticsOne, C315DCMN-SPC mini, cut to fit 5 mm C315GMN with a 0.5 mm projection depth) were sterilized in an autoclave. The dummy cannulae were partially screwed into the guide cannulae and placed in a stereotaxic clamp. After zeroing the tip of the guide cannula to the brain surface, the cannula was lowered to 500 μm below the cortical surface. This depth was chosen because the infusion cannulae (PlasticsOne, C315LIMN-SPC mini) project 500 μm from the end of the guide cannulae when screwed in completely, leading to a final depth of 1000 μm – the location of auditory cortex. The dummy cannulae were then fully screwed down and this procedure was repeated for the next cortical hemisphere.

Prior to injecting, two injection syringes (Hamilton Syringe, 10μL Gaslight #1701) and tubing (C313CT tubing 023x050 PE50) were backfilled with mineral oil. Sterilized infusion cannulae were then attached to each syringe and ~500nL of muscimol (diluted with 1x PBS to .25 mg/mL; Sigma Aldrich, M1523) or 0.9% sterile saline was drawn up into the injection cannulae using a dual injector (Harvard Apparatus, Pump 11 Pico Plus Elite). The mouse was then headfixed and the dummy cannulae were removed and sterilized. The loaded infusion cannulae were then screwed all the way into the guide cannulae and 400 nL of muscimol or saline was infused bilaterally at a rate of 250 nL/minute. The infusion cannulae were then replaced with the dummy cannulae and the mouse rested in its home cage for 30-45 minutes before beginning the behavioral session.

*Acute electrophysiological recordings.*

For acute recordings used to fit the GC-GLM model (Figure 2), neuronal signals were recorded from n = 1 awake, untrained mouse. Prior to the recording session, the mouse was anesthetized and a headpost and ground pin were implanted on the skull (see *Surgery*). On the day of the recording, the mouse was briefly anesthetized with 3% isoflurane and a small craniotomy was performed over auditory cortex using a dental drill or scalpel (~1 mm x 1 mm craniotomy centered approximately 1.25 mm anterior to the lambdoid suture along caudal end of the squamosal suture). A 32 channel silicon probe (Neuronexus) was then positioned perpendicularly to the cortical surface and lowered at a rate of 1-2 μm/s to a final depth of 800-1200 μm. As the probe was lowered, trains of brief noise bursts were repeated, and if stimulus locked responses to the noise bursts were observed, the probe was determined to be in auditory cortex. The probe was then allowed to settle for up to 30 minutes before starting the recording. Neuronal signals were amplified and digitized with an Intan headstage (RHD 32ch) and recorded by an openEphys acquisition board90 at a rate of 30 kHz.

For this experiment, the mouse was presented with 3 s DRCs alternating between low and high contrast (uniform distribution with a mean of 50 dB and a width of ±5 dB in low contrast or ±15 dB in high contrast at a chord rate of 25 ms, as described in *Stimuli*). In order to accurately fit the GLM in an unbiased manner, these stimuli were highly random, composed of 100 unique chord patterns for each contrast (Supplementary Figure 2i,j**)**. For each of the two recording sites, 5 repeats of this stimulus set were played.

*Behavioral electrophysiological recordings*.

Neural signals were acquired from awake, behaving mice as they performed the psychometric and offset testing tasks described previously. Chronically implanted, 16-, 32-, or 64-channel microdrives90,91 were connected to one or two 32 channel Intan amplifier headstages. Amplified signals were recorded at 30 kHz using an openEphys acquisition board via an SPI cable, where the signals were digitized. Spikes were then sorted using Kilosort, as described previously.

For all recordings, broadband signals were filtered between 500 and 6000 Hz, offset corrected, and re-referenced to the median across all active channels. The preprocessed data was then sorted using KiloSort92 or KiloSort2 and the resulting clustering was manually corrected in phy2 according to community-developed guidelines. The resulting units were labelled as single units if they exhibited a clear refractory period and did not need to be split. Splitting assessments were made through manual examination of principle component features for the two best channels of a cluster. If two noticeable clusters in feature space were evident in a unit, the unit was either manually split, or classified as a multiunit.

*Generalized linear model.*

To justify the form of GLM used here, we discuss a how a model neuron could implement gain control in the simplest terms, and then structure our inference model to extract the parameters of this model neuron. We will assume that the activity of the model neuron is driven by three sources: 1) stimulus drive, 2) stimulus contrast, and 3) the multiplicative interaction between the two, which we use to define the gain (for a formal definition of this forward model and the inference model, see *Supplementary Methods*).

As discussed previously, the stimulus used in our experiments is a one-dimensional signal that evolves in discrete time steps:

where is the stimulus spectrogram that varies as a function of time and frequency . Each time and frequency bin of is sampled from a uniform distribution defined by an average value and contrast (see *Stimuli* and *Acute Electrophysiological Recordings*).

We assume that the hypothetical neuron responds selectively at some frequency and time lag, defined by a filter, or STRF with history and frequency components. Given , we can define the stimulus drive as

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

where at each time , is a row vector of size frequencies by lags (ie. a matrix of the lagged stimulus spectrogram) and is the STRF collapsed to a single row vector of the same size.

In the spirit of efficient coding theory, and as shown in previous work, we assume that the gain of the neuron should be inversely proportional to the contrast, such that (ie. when contrast is low gain should be high, and vice-versa). We also define “neutral” gain to be the average of the gain of the neuron in low and high contrast. Putting these two features together, we can summarize the gain of the neuron as

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

where is the harmonic mean of the contrast between the low and high conditions (see *Supplemental Methods*). In the case of a 3-fold change in contrast, this function constrains the gain of the neuron between 1.5 and 5, with a neutral value of 1. As mentioned previously, we consider gain to be the multiplicative interaction between the stimulus drive and the contrast, such that the contribution of gain control to the response of the neuron is related to .

To summarize, we considered a hypothetical neuron driven by the stimulus according to a STRF and by the interaction between the stimulus drive and the contrast . To infer the relative weights of each of these components of the neural response, we defined a Poisson GLM with an intercept term and the following predictors:

|  |  |  |
| --- | --- | --- |
|  |  | (3) |

In other words, the model is composed of a stimulus predictor , a contrast predictor , and their interaction. Therefore, the GLM models the firing rate at time as a Poisson distribution with the following mean:

|  |  |  |
| --- | --- | --- |
|  |  | (4) |

where are the parameters to be inferred. Based on our behavioral data (Figure 3) and the predictions of the efficient coding model (Figure 1), we expected the influence of contrast on neural gain to be asymmetric and smooth. To enforce both of these qualities in the GLM, we first defined the contrast predictors from a set of cubic B-spline temporal basis functions, then defined separate contrast predictors for transitions to low and high contrast. Incorporating these changes, we can redefine equation 4 above as

|  |  |  |
| --- | --- | --- |
|  |  | (5) |

where denotes element-by-element “broadcasting” multiplication and is a matrix of contrast predictors convolved with a set of basis functions and separated by contrast transitions (see *Supplemental Methods*). For the sake of clarity, note that in the expression above, is a number, is a column vector of length , is a number, is a -by- matrix, and and are column vectors of length , where is the number of splines.

So far, we outlined a hypothetical neuron which implements gain control, and a GLM with which we can approximate the behavior of this neuron. Next, we describe how to use the fitted parameters to quantify the gain of the neuron. Conceptually, an increase or decrease in the gain of the neuron is analogous to more or less sensitivity to small changes in the stimulus. Based on this intuition, we focused on how the response of the neuron (as modelled by a fitted GLM) is expected to change between conditions where the gain is expected to contribute (ie. in the presence of gain control) and where it is not (ie. in the absence of gain control, where gain is “neutral”). Following this logic, we derived a definition for gain as the ratio between the sensitivity of the fitted model with changes in contrast, compared to the sensitivity of the same model when the contrast is at a reference value, which we defined previously as :

|  |  |  |
| --- | --- | --- |
|  |  | (6) |

where is the estimated gain at time , and is a reference contrast design matrix identical to except that all non-zero elements are set to 1 (see *Supplementary Methods* for full derivation of ).

To fit the model, we implemented a two-step procedure. In the first step, the STRF of the neuron was estimated according to the model

|  |  |  |
| --- | --- | --- |
|  |  | (7) |

For the second step, we calculated the stimulus drive as described in equation 1, and then fit equation 5 to the data for each neuron using in MATLAB. This entire fitting procedure was 10-fold cross-validated with folds stratified across trials of each contrast. In the first step, we fit STRF with frequency bins according to the stimulus spectrogram ( = 33 or 34, see *Stimuli*) and a history window of 300 ms ( = 12). When fitting the full model, we defined the contrast design matrix to capture 1000 ms of contrast history around each transition ( = 40), convolved with a set of B-spline temporal basis functions (here, we used B-splines with a degree of 3 and 7 knots such that = 8; **CITE)**.

To validate the model, we first simulated neurons according to the forward model outlined above (Supplementary Figure 2a), while varying the amount of gain control and the temporal trajectory of gain in different simulation runs. We found that the GLM accurately predicted the STRF shape, spike rates and gain trajectories across a variety of simulation parameters (Supplementary Figure c, e-h). For a detailed description and discussion of the simulation results, see *Supplementary Methods* and Supplementary Table 3.

*Behavioral and Neural Detection Performance.* To calculate performance in target-in-noise detection task we adopted commonly used signal detection theory methods41,93 to estimate the ability of an ideal observer to discriminate between two sensory distributions: in our case, a distribution for target trials and a distribution for noise trials. When analyzing behavior, we computed the percent correct performance of an ideal observer94 as a function of the probability of hits and false alarms:

|  |  |  |
| --- | --- | --- |
|  |  | (8) |

where is cumulative probability of the normal distribution ( in MATLAB), is the inverse of the normal distribution (ie. the z-score, in MATLAB), is the hit rate, and is the false alarm rate. For psychophysical performance, hit rates and false alarm rates near 0 and 1 were adjusted using the log-linear rule95, to reduce biases in performance estimation caused by low numbers of trials.

To calculate neural performance in the same reference frame as the behavior, we employed similar ideal observer techniques. First, neuronal responses (either spike rates or single units, or population projection values), were averaged in a 100ms window post target onset (for noise trials, this window was randomly chosen on each trial to coincide with target presentation times on target trials). Then, using the distributions of responses during target and noise trials, we computed receiver-operating-characteristic curves and took the area under the curve (AUC) as the percent correct of an ideal observer discriminating between the target and noise distributions. To determine whether the AUC value for a given set of trial distributions was significantly different from chance, we performed a bootstrap procedure where we sampled from all the trials with replacement 500 times and recomputed AUC for each sample. If the 95% confidence intervals for this bootstrapped distribution did not include chance (.5), we defined that AUC value as significant. For population analyses which generated single-trial predictions, neural hit and false alarm rates were transformed to percent correct as described above.

To characterize performance, psychometric curves were fit with a logistic function:

|  |  |  |
| --- | --- | --- |
|  |  | (9) |

where is the x-offset of the function, determined the slope, or sensitivity of the function, determined the guess rate (lower bound), determined the lapse rate (upper bound) and was stimulus volume. determined the threshold of this function, defined as the volume corresponding to the steepest part of the curve. This function was fit to behavioral or neural performance using constrained gradient descent ( in MATLAB) initialized with a 10x10 grid-search of parameters and .

To characterize adaptation time constants, adaptation curves were fit with an exponential function

|  |  |  |
| --- | --- | --- |
|  |  | (10) |

where determined the y-offset of the function, was a multiplicative scaling factor, and was the time constant of the exponential in units of time . This function was fit to behavioral or neural responses using constrained gradient descent initialized with a 10x10x10 grid search across all three parameters.

*Population response metrics.*

On sessions where three or more neurons were simultaneously recorded, we used a coding direction technique28 to estimate the ability of neural populations to discriminate targets from noise. First, target and noise spike rates for each neuron were averaged in a 100ms window post-target onset. Then, using a leave-one-out procedure, we computed a trial averaged population vector for target trials, , and a separate average population vector for noise trials, . We then estimated the coding direction in high dimensional neural space that best separated the target and noise responses: The held out trial was then projected along this dimension, by taking the population response vector on that trial and projecting it along the estimated coding direction using the dot product: . This procedure was repeated holding out each trial, and estimating the coding direction from the remaining trials. For psychometric testing sessions, the target responses from the two loudest target volumes were used to estimate coding direction, and in offset testing sessions the target responses from the high SNR target trials were used. After computing projections for every trial, the resulting matrix was normalized between 0 and 1.

*Population classifier*.

Based on previously described methods29, we used a criterion-based decision rule to estimate how a hypothetical down-stream neuron may read out the neural activity of a population of neurons. As before, trial distributions of neural responses to targets and noise were created from the average activity in a 100ms window post-target. Then, we sampled 100 criterion values between the minimum and maximum response, and for each criterion estimated the proportion of correct trials under two decision rules: 1) report target present if the response is greater than the criterion, or, 2) report target present if the response is less than the criterion. By assessing these two decision rules, responses that may be suppressed by target presence were treated equally to those in which target presence enhanced the neural response. Finally, we chose the criterion and decision rule that yielded the highest proportion of correct trials, and computed neural hit rates and false alarm rates for each target level, and noise-only trials. These hit rates and false alarm rates were then transformed to percent correct according the formula above, to ensure equivalency with the behavioral metrics.

*Linear-nonlinear model*.

First, we selected only neurons in the dataset which had reliable stimulus responses (noise ratio < 100). The linear nonlinear model was then composed of two main parts, a spectrotemporal receptive field (STRF) and a set of rectifying nonlinearities. The STRF was fit using normalized reverse correlation

|  |  |  |
| --- | --- | --- |
|  |  | (11) |

where is the stimulus design matrix defined in equation 1 and is the spike count in each 25 ms bin of the DRC stimulus. When defining , we used a history window of 300 ms ( = 12) and frequency bins corresponding to the frequencies composing the dynamic random chords (see *Stimuli*). After fitting the STRF, we fit the nonlinearities of the neuron. This two-step fitting procedure was repeated using 10 fold cross-validation, as described below.

For each fold, we selected 90% of the trials for training, leaving the remaining 10% to be held out for testing. Within each trial, we excluded neuronal responses around transitions from silence, or transitions in contrast, to prevent the model from overfitting strong transients in the neural response. Additionally, we excluded neural responses within a 50ms window after target presentations, to prevent overfitting to target responses. Given these exclusion criteria, we calculated the duration of stimulus sampled in the shorter target period for each trial, and, for that trial, sampled the same duration of stimulus within the adaptation period. This procedure ensured that the model was fit to the same amount of high and low contrast stimulation, to minimize overfitting to one condition. Then, a stimulus design matrix was made using these stimulus periods, and the STRF was fit using equation 11. During an initial pilot experiment, we tested whether STRF properties were affected by stimulus contrast, and found STRFs to be largely stable when estimated separately for each contrast (*Supplementary Methods* and Supplementary Figure 5b-g). Therefore, we used both periods of contrast to estimate .

Using the STRF fit to the training data, we computed the linear drive by convolving the STRF with the lagged spectrogram of the training stimulus (equation 1). We then separated the linear predictions into low and high contrast periods. For each contrast period, we generated a histogram of the linear prediction values (50 bins), and for each bin, computed the mean spike rate of the neuron when the linear prediction fell within those bin edges (Figure 6d, scatter points). The resulting set of linear prediction values and average spike rates were fit with an exponential function:

|  |  |  |
| --- | --- | --- |
|  |  | (12) |

where determined the minimum firing rate, was a multiplicative scaling factor, determined the gain of the exponent, and determined the x-offset, or firing threshold of the neuron. This function was fit to each cell using constrained gradient descent ( in MATLAB), using a 10x10 grid search for parameters and . These fits were determined for each contrast, and the gain for each contrast for each neuron was estimated using . This entire process was repeated for each cross-validation fold, and the final parameter estimates for the STRF and nonlinearities were taken as the average over the 10 runs.

To determine the relationship between neuronal gain and behavioral performance, we examined our dataset of neurons collected during the psychometric task. First, we selected all of the neurons with significant AUC values to at least two of the six targets to ensure that we were sampling from neurons with information about target volume. Then, across all of these neurons for each mouse, we computed the average neural gain for each contrast. We then compared neural gain for each mouse to corresponding average psychometric thresholds and slopes to assess the relationship between neural gain and behavioral performance.

*Inclusion criteria*. Unless otherwise noted, behavioral sessions in which the false alarm rate exceeded 50% were discarded from analysis. One mouse (ID: CA122) had consistently high false alarm rates in the high contrast condition, so we excluded high contrast sessions from this mouse from all analyses. For Figure 4g-i, sessions with stable population decoding performance were included (defined as sessions where more than half of the target volumes or times elicited significant population AUC values, as determined by the bootstrap procedure described previously).

1. Barlow, H. Possible principles underlying the transformations of sensory messages. *Sensory communication* **6**, 57–58 (1961).

2. Brenner, N., Bialek, W. & De Ruyter Van Steveninck, R. Adaptive rescaling maximizes information transmission. *Neuron* **26**, 695–702 (2000).

3. Bharioke, A. & Chklovskii, D. B. Automatic Adaptation to Fast Input Changes in a Time-Invariant Neural Circuit. *PLoS Comput Biol* **11**, 1004315 (2015).

4. Borst, A. & Theunissen, F. E. Information theory and neural coding. *Nature Neuroscience* **2**, 947–957 (1999).

5. Maravall, M., Petersen, R. S., Fairhall, A. L., Arabzadeh, E. & Diamond, M. E. Shifts in Coding Properties and Maintenance of Information Transmission during Adaptation in Barrel Cortex. *PLoS Biol.* **5**, e19 (2007).

6. Baccus, S. A. & Meister, M. Fast and slow contrast adaptation in retinal circuitry. *Neuron* **36**, 909–919 (2002).

7. Gutnisky, D. A. & Dragoi, V. Adaptive coding of visual information in neural populations. *Nature* **452**, 220–224 (2008).

8. Clemens, J., Ozeri-Engelhard, N. & Murthy, M. Fast intensity adaptation enhances the encoding of sound in Drosophila. *Nat. Commun.* **9**, 1–15 (2018).

9. Clarke, S. E., Longtin, A. & Maler, L. Contrast coding in the electrosensory system: Parallels with visual computation. *Nature Reviews Neuroscience* **16**, 733–744 (2015).

10. Dahmen, J. C., Keating, P., Nodal, F. R., Schulz, A. L. & King, A. J. Adaptation to Stimulus Statistics in the Perception and Neural Representation of Auditory Space. *Neuron* **66**, 937–948 (2010).

11. Wen, B., Wang, G. I., Dean, I. & Delgutte, B. Time course of dynamic range adaptation in the auditory nerve. *J. Neurophysiol.* **108**, 69–82 (2012).

12. Dean, I., Harper, N. S. & McAlpine, D. Neural population coding of sound level adapts to stimulus statistics. *Nat. Neurosci.* **8**, 1684–1689 (2005).

13. Wen, B., Wang, G. I., Dean, I. & Delgutte, B. Dynamic range adaptation to sound level statistics in the auditory nerve. *J. Neurosci.* **29**, 13797–13808 (2009).

14. Rabinowitz, N. C., Willmore, B. D. B., Schnupp, J. W. H. & King, A. J. Contrast Gain Control in Auditory Cortex. *Neuron* **70**, 1178–1191 (2011).

15. Rabinowitz, N. C., Willmore, B. D. B., King, A. J. & Schnupp, J. W. H. Constructing Noise-Invariant Representations of Sound in the Auditory Pathway. *PLoS Biol.* **11**, e1001710 (2013).

16. Willmore, B. D. B., Cooke, J. E. & King, A. J. Hearing in noisy environments: noise invariance and contrast gain control. *J. Physiol.* **592**, 3371–3381 (2014).

17. Cooke, J. E., King, A. J., Willmore, B. D. B. & Schnupp, J. W. H. Contrast gain control in mouse auditory cortex. *J. Neurophysiol.* **120**, 1872–1884 (2018).

18. Cooke, J. E. *et al.* Contrast gain control occurs independently of both parvalbumin-positive interneuron activity and shunting inhibition in auditory cortex. *J. Neurophysiol.* **123**, 1536–1551 (2020).

19. DeWeese, M. & Zador, A. Asymmetric Dynamics in Optimal Variance Adaptation. *Neural Comput.* **10**, 1179–1202 (1998).

20. Wen, B., Wang, G. I., Dean, I. & Delgutte, B. Time course of dynamic range adaptation in the auditory nerve. *J Neurophysiol* **108**, 69–82 (2012).

21. Młynarski, W. & Hermundstad, A. M. Adaptability and efficiency in neural coding. *bioRxiv* 669200 (2019). doi:10.1101/669200

22. Młynarski, W. F. & Hermundstad, A. M. Adaptive coding for dynamic sensory inference. *Elife* **7**, (2018).

23. Wei, X.-X. & Stocker, A. A. A Bayesian observer model constrained by efficient coding can explain ‘anti-Bayesian’ percepts. *Nat. Neurosci.* **18**, 1509–1517 (2015).

24. Lohse, M., Bajo, V. M., King, A. J. & Willmore, B. D. B. Neural circuits underlying auditory contrast gain control and their perceptual implications. *Nat. Commun.* **11**, 1–13 (2020).

25. Rabinowitz, N. C., Willmore, B. D. B. B., Schnupp, J. W. H. H. & King, A. J. Contrast Gain Control in Auditory Cortex. *Neuron* **70**, 1178–1191 (2011).

26. Pennington, J. R. & David, S. V. Complementary effects of adaptation and gain control on sound encoding in primary auditory cortex. *eNeuro* **7**, 1–17 (2020).

27. Rabinowitz, N. C., Willmore, B. D. B., Schnupp, J. W. H. & King, A. J. Spectrotemporal contrast kernels for neurons in primary auditory cortex. *J. Neurosci.* **32**, 11271–11284 (2012).

28. Li, N., Daie, K., Svoboda, K. & Druckmann, S. Robust neuronal dynamics in premotor cortex during motor planning. *Nature* **532**, 459–464 (2016).

29. Christison-Lagay, K. L., Bennur, S. & Cohen, Y. E. Contribution of spiking activity in the primary auditory cortex to detection in noise. *J. Neurophysiol.* **118**, 3118–3131 (2017).

30. Barlow, H. B. Possible Principles Underlying the Transformations of Sensory Messages. in *Sensory Communication* 216–234 (1961). doi:10.7551/mitpress/9780262518420.003.0013

31. Talwar, S. K., Musial, P. G. & Gerstein, G. L. Role of mammalian auditory cortex in the perception of elementary sound properties. *J. Neurophysiol.* **85**, 2350–2358 (2001).

32. Gimenez, T. L., Lorenc, M. & Jaramillo, S. Adaptive categorization of sound frequency does not require the auditory cortex in rats. *J. Neurophysiol.* **114**, 1137–1145 (2015).

33. Jaramillo, S. & Zador, A. M. *Auditory cortex mediates the perceptual effects of acoustic temporal expectation*. *Nature Neuroscience* (2010).

34. Wood, K. C., Town, S. M., Atilgan, H., Jones, G. P. & Bizley, J. K. Acute inactivation of primary auditory cortex causes a sound localisation deficit in ferrets. *PLoS One* **12**, (2017).

35. Kato, H. K., Gillet, S. N. & Isaacson, J. S. Flexible Sensory Representations in Auditory Cortex Driven by Behavioral Relevance. *Neuron* **88**, 1027–1039 (2015).

36. Ceballo, S., Piwkowska, Z. & Bourg, J. Targeted Cortical Manipulation of Auditory Perception In Brief. *Neuron* **104**, 1168-1179.e5 (2019).

37. Li, Z. *et al.* Corticostriatal control of defense behavior in mice induced by auditory looming cues. *Nat. Commun.* **12**, 1–13 (2021).

38. Town, S., Wood, K. & Bizley, J. Signal processing in auditory cortex underlies degraded speech sound discrimination in noise. *bioRxiv* 833558 (2019). doi:10.1101/833558

39. Musall, S., Urai, A. E., Sussillo, D. & Churchland, A. K. Harnessing behavioral diversity to understand neural computations for cognition. *Current Opinion in Neurobiology* **58**, 229–238 (2019).

40. Shadlen, M. N. & Kiani, R. Decision making as a window on cognition. *Neuron* **80**, 791–806 (2013).

41. Newsome, W. T., Britten, K. H. & Movshon, J. A. Neuronal correlates of a perceptual decision. *Nature* **341**, 52–54 (1989).

42. Britten, K. H., Shadlen, M. N., Newsome, W. T. & Movshon, J. A. The analysis of visual motion: A comparison of neuronal and psychophysical performance. *J. Neurosci.* **12**, 4745–4765 (1992).

43. Shadlen, M. N., Britten, K. H., Newsome, W. T. & Movshon, J. A. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J. Neurosci.* **16**, 1486–1510 (1996).

44. Nienborg, H. & Cumming, B. G. Decision-related activity in sensory neurons reflects more than a neurons causal effect. *Nature* **459**, 89–92 (2009).

45. Cumming, B. G. & Nienborg, H. Feedforward and feedback sources of choice probability in neural population responses. *Current Opinion in Neurobiology* **37**, 126–132 (2016).

46. Tsunada, J., Liu, A. S. K., Gold, J. I. & Cohen, Y. E. Causal contribution of primate auditory cortex to auditory perceptual decision-making. *Nat. Neurosci.* **19**, 135–142 (2015).

47. Steinmetz, N. A., Zatka-Haas, P., Carandini, M. & Harris, K. D. Distributed coding of choice, action and engagement across the mouse brain. *Nature* **576**, 266–273 (2019).

48. Cohen, M. R. & Newsome, W. T. Context-Dependent Changes in Functional Circuitry in Visual Area MT. *Neuron* **60**, 162–173 (2008).

49. Cohen, M. R. & Newsome, W. T. Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. *J. Neurosci.* **29**, 6635–6648 (2009).

50. Ni, A. M., Ruff, D. A., Alberts, J. J., Symmonds, J. & Cohen, M. R. Learning and attention reveal a general relationship between population activity and behavior. *Science (80-. ).* **359**, 463–465 (2018).

51. Downer, J. D., Niwa, M. & Sutter, M. L. Task Engagement Selectively Modulates Neural Correlations in Primary Auditory Cortex. *J. Neurosci.* **35**, 7565–7574 (2015).

52. Stringer, C., Michaelos, M., Tsyboulski, D., Lindo, S. E. & Pachitariu, M. High-precision coding in visual cortex. *Cell* (2021). doi:10.1016/j.cell.2021.03.042

53. Hires, S. A., Gutnisky, D. A., Yu, J., O’Connor, D. H. & Svoboda, K. Low-noise encoding of active touch by layer 4 in the somatosensory cortex. *Elife* **4**, (2015).

54. Hobbs, J. A., Towal, R. B. & Hartmann, M. J. Z. Spatiotemporal patterns of contact across the rat vibrissal array during exploratory behavior. *Front. Behav. Neurosci.* **9**, 356 (2016).

55. Aizenberg, M. & Geffen, M. N. Bidirectional effects of aversive learning on perceptual acuity are mediated by the sensory cortex. *Nat. Neurosci.* **16**, 994–996 (2013).

56. Aizenberg, M., Mwilambwe-Tshilobo, L., Briguglio, J. J., Natan, R. G. & Geffen, M. N. Bidirectional Regulation of Innate and Learned Behaviors That Rely on Frequency Discrimination by Cortical Inhibitory Neurons. *PLOS Biol.* **13**, e1002308 (2015).

57. Briguglio, J. J., Aizenberg, M., Balasubramanian, V. & Geffen, M. N. Cortical neural activity predicts sensory acuity under optogenetic manipulation. *J. Neurosci.* **38**, 2094–2105 (2018).

58. Wood, K. C., Angeloni, C. F., Oxman, K., Clopath, C. & Geffen, M. N. Neuronal activity in sensory cortex predicts the specificity of learning. *bioRxiv* 2020.06.02.128702 (2020). doi:10.1101/2020.06.02.128702

59. Ulanovsky, N., Las, L. & Nelken, I. Processing of low-probability sounds by cortical neurons. *Nat. Neurosci.* **6**, 391–398 (2003).

60. Natan, R. G., Carruthers, I. M., Mwilambwe-Tshilobo, L. & Geffen, M. N. Gain Control in the Auditory Cortex Evoked by Changing Temporal Correlation of Sounds. *Cereb. Cortex* **27**, 2385–2402 (2017).

61. Espejo, M. L., Schwartz, Z. P. & David, S. V. Spectral tuning of adaptation supports coding of sensory context in auditory cortex. *PLoS Comput. Biol.* **15**, e1007430 (2019).

62. Fritz, J., Shamma, S., Elhilali, M. & Klein, D. Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat. Neurosci.* **6**, 1216–1223 (2003).

63. Mesgarani, N., Fritz, J. & Shamma, S. A computational model of rapid task-related plasticity of auditory cortical receptive fields. *J. Comput. Neurosci.* **28**, 19–27 (2010).

64. David, S. V., Fritz, J. B. & Shamma, S. A. Task reward structure shapes rapid receptive field plasticity in auditory cortex. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 2144–2149 (2012).

65. Yin, P., Fritz, J. B. & Shamma, S. A. Rapid spectrotemporal plasticity in primary auditory cortex during behavior. *J. Neurosci.* **34**, 4396–408 (2014).

66. Niwa, M., Johnson, J. S., O’Connor, K. N. & Sutter, M. L. Active Engagement Improves Primary Auditory Cortical Neurons’ Ability to Discriminate Temporal Modulation. *J. Neurosci.* **32**, 9323–9334 (2012).

67. Fritz, J. B., Elhilali, M. & Shamma, S. A. Adaptive changes in cortical receptive fields induced by attention to complex sounds. *J. Neurophysiol.* **98**, 2337–46 (2007).

68. Simoncelli, E. P., Paninski, L., Pillow, J. & Schwartz, O. *Characterization of Neural Responses with Stochastic Stimuli*. (MIT Press, 2004).

69. Paninski, L., Pillow, J. & Lewi, J. Statistical models for neural encoding, decoding, and optimal stimulus design. *Progress in Brain Research* **165**, 493–507 (2007).

70. Park, I. M., Meister, M. L. R., Huk, A. C. & Pillow, J. W. Encoding and decoding in parietal cortex during sensorimotor decision-making. *Nat. Neurosci.* **17**, 1395–1403 (2014).

71. Eggermont, J. J., Johannesma, P. I. M. & Aertsen, A. M. H. J. Reverse-correlation methods in auditory research. *Q. Rev. Biophys.* **16**, 341–414 (1983).

72. Natan, R. G. *et al.* Complementary control of sensory adaptation by two types of cortical interneurons. *Elife* **4**, (2015).

73. Natan, R. G., Rao, W. & Geffen, M. N. Cortical Interneurons Differentially Shape Frequency Tuning following Adaptation. *Cell Rep.* **21**, 878–890 (2017).

74. Atallah, B. V., Bruns, W., Carandini, M. & Scanziani, M. Parvalbumin-Expressing Interneurons Linearly Transform Cortical Responses to Visual Stimuli. *Neuron* **73**, 159–170 (2012).

75. Wilson, N. R., Runyan, C. A., Wang, F. L. & Sur, M. Division and subtraction by distinct cortical inhibitory networks in vivo. *Nature* **488**, 343–348 (2012).

76. Seybold, B. a, Phillips, E. a K., Schreiner, C. E. & Hasenstaub, A. R. Inhibitory Actions Unified by Network Integration. *Neuron* **87**, 1181–1192 (2015).

77. Phillips, E. A. K. & Hasenstaub, A. R. Asymmetric effects of activating and inactivating cortical interneurons. *Elife* **5**, e18383 (2016).

78. Attneave, F. Some informational aspects of visual perception. *Psychol. Rev.* **61**, 183–193 (1954).

79. Simoncelli, E. P. & Olshausen, B. A. Natural image statistics and neural representation. *Annual Review of Neuroscience* **24**, 1193–1216 (2001).

80. Simoncelli, E. P. Vision and the statistics of the visual environment. *Current Opinion in Neurobiology* **13**, 144–149 (2003).

81. Młynarski, W., Hledík, M., Sokolowski, T. R. & Tkačik, G. Statistical analysis and optimality of neural systems. *Neuron* **109**, 1227-1241.e5 (2021).

82. Schneider, D. M., Nelson, A. & Mooney, R. A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature* **513**, 189–94 (2014).

83. Schneider, D. M., Sundararajan, J. & Mooney, richard. A cortical filter that learns to suppress the acoustic consequences of movement. *Nature* (2018). doi:10.1038/s41586-018-0520-5

84. McGinley, M. J., David, S. V. & McCormick, D. A. Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. *Neuron* **87**, 179–192 (2015).

85. Reimer, J. *et al.* Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nat. Commun.* **7**, 1–7 (2016).

86. Guo, Z. V. *et al.* Procedures for behavioral experiments in head-fixed mice. *PLoS One* **9**, (2014).

87. Isett, B. R., Feasel, S. H., Lane, M. A. & Feldman, D. E. Slip-Based Coding of Local Shape and Texture in Mouse S1. *Neuron* **97**, 418-433.e5 (2018).

88. Carruthers, I. M., Natan, R. G. & Geffen, M. N. Encoding of ultrasonic vocalizations in the auditory cortex. *J Neurophysiol* **109**, 1912–1927 (2013).

89. Carruthers, I. M. *et al.* Emergence of invariant representation of vocalizations in the auditory cortex. *J. Neurophysiol.* jn.00095.2015 (2015). doi:10.1152/jn.00095.2015

90. Voigts, J. *et al.* An easy-to-assemble, robust, and lightweight drive implant for chronic tetrode recordings in freely moving animals. *J. Neural Eng.* **17**, 26044 (2020).

91. Voigts, J., Siegle, J., Pritchett, D. L. & Moore, C. I. The flexDrive: An ultra-light implant for optical control and highly parallel chronic recording of neuronal ensembles in freely moving mice. *Front. Syst. Neurosci.* **7**, 8 (2013).

92. Pachitariu, M., Steinmetz, N., Kadir, S., Carandini, M. & Harris, K. *Fast and accurate spike sorting of high-channel count probes with KiloSort*.

93. Stanislaw, H. & Todorov, N. *Calculation of signal detection theory measures*. *Behavior Research Methods, Instruments, and Computers* **31**, (1999).

94. Rocchi, F. & Ramachandran, R. Neuronal adaptation to sound statistics in the inferior colliculus of behaving macaques does not reduce the effectiveness of the masking noise. *J. Neurophysiol.* **120**, 2819–2833 (2018).

95. Hautus, M. J. *Corrections for extreme proportions and their biasing effects on estimated values of d’*. *Behavior Research Methods. Instruments. & Computers* (1995).