

UC Davis

UC Davis Previously Published Works

Title

Ability of primary auditory cortical neurons to detect amplitude modulation with rate and temporal codes: neurometric analysis.

Permalink

<https://escholarship.org/uc/item/91h6k3h1>

Journal

Journal of neurophysiology, 107(12)

ISSN

0022-3077

Authors

Johnson, Jeffrey S
Yin, Pingbo
O'Connor, Kevin N
et al.

Publication Date

2012-06-01

DOI

10.1152/jn.00812.2011

Peer reviewed

Ability of primary auditory cortical neurons to detect amplitude modulation with rate and temporal codes: neurometric analysis

Jeffrey S. Johnson, Pingbo Yin, Kevin N. O'Connor and Mitchell L. Sutter

J Neurophysiol 107:3325-3341, 2012. First published 14 March 2012; doi:10.1152/jn.00812.2011

You might find this additional info useful...

This article cites 113 articles, 64 of which can be accessed free at:

<http://jn.physiology.org/content/107/12/3325.full.html#ref-list-1>

Updated information and services including high resolution figures, can be found at:

<http://jn.physiology.org/content/107/12/3325.full.html>

Additional material and information about *Journal of Neurophysiology* can be found at:

<http://www.the-aps.org/publications/jn>

This information is current as of June 15, 2012.

Ability of primary auditory cortical neurons to detect amplitude modulation with rate and temporal codes: neurometric analysis

Jeffrey S. Johnson, Pingbo Yin, Kevin N. O'Connor, and Mitchell L. Sutter

Center for Neuroscience, University of California at Davis, Davis, California

Submitted 6 September 2011; accepted in final form 8 March 2012

Johnson JS, Yin P, O'Connor KN, Sutter ML. Ability of primary auditory cortical neurons to detect amplitude modulation with rate and temporal codes: neurometric analysis. *J Neurophysiol* 107: 3325–3341, 2012. First published March 14, 2012; doi:10.1152/jn.00812.2011.—Amplitude modulation (AM) is a common feature of natural sounds, and its detection is biologically important. Even though most sounds are not fully modulated, the majority of physiological studies have focused on fully modulated (100% modulation depth) sounds. We presented AM noise at a range of modulation depths to awake macaque monkeys while recording from neurons in primary auditory cortex (A1). The ability of neurons to detect partial AM with rate and temporal codes was assessed with signal detection methods. On average, single-cell synchrony was as or more sensitive than spike count in modulation detection. Cells are less sensitive to modulation depth if tested away from their best modulation frequency, particularly for temporal measures. Mean neural modulation detection thresholds in A1 are not as sensitive as behavioral thresholds, but with phase locking the most sensitive neurons are more sensitive, suggesting that for temporal measures the lower-envelope principle cannot account for thresholds. Three methods of preanalysis pooling of spike trains (multiunit, similar to convergence from a cortical column; within cell, similar to convergence of cells with matched response properties; across cell, similar to indiscriminate convergence of cells) all result in an increase in neural sensitivity to modulation depth for both temporal and rate codes. For the across-cell method, pooling of a few dozen cells can result in detection thresholds that approximate those of the behaving animal. With synchrony measures, indiscriminate pooling results in sensitive detection of modulation frequencies between 20 and 60 Hz, suggesting that differences in AM response phase are minor in A1.

phase locking; synchrony; neuronal pooling

IN ADDITION to their spectral content, natural sounds typically include a time-varying nonspectral amplitude envelope that modulates the spectral carrier. The importance of the envelope is underscored by the fact that neurons are responsive to amplitude modulation (AM) in a wide range of natural sounds and natural acoustic environments (Attias and Schreiner 1998; Chandrasekaran et al. 2010; DiMattina and Wang 2006; Kajikawa et al. 2008; Nagarajan et al. 2002; Nelken et al. 1999; Singh and Theunissen 2003). AM plays a notable role in speech perception (Delgutte et al. 1998; Drullman et al. 1994; Shannon et al. 1995; Steinschneider et al. 2003; Young 2008). The amplitude envelope is also thought to be of particular use in segregating sound sources during auditory scene analysis (Bregman 1990; Fishman and Steinschneider 2010; Grimault et al. 2002; Yost 1991), in musical perception (Fishman et al.

2001; Helmholtz 1954), and in the identification of pitch (Burns and Viemeister 1976, 1981).

In natural environments most sounds are only partially (<100%) modulated, and in noisy environments the noise will act to reduce the effective depth of modulated sounds. Despite the behavioral relevance of partially modulated sounds, most studies of AM have used envelopes with 100% (peak to trough) modulation depths (see Joris et al. 2004 for a review). Of the few studies that have varied modulation depth, some have shown increases in synchronization correlated to modulation depth but little change in firing rate (Bieser and Müller-Preuss 1996; Müller-Preuss et al. 1994) while others see changes in both synchronization and firing rate (Eggermont 1994; Liang et al. 2002; Malone et al. 2010; Middlebrooks 2008a; Nelson and Carney 2007).

The ability to detect whether a sound is amplitude modulated as a function of modulation depth has been characterized in humans and other primate species (e.g., O'Connor et al. 2000, 2011), and comparison of this ability to single-unit or multiunit detection thresholds can provide insight on the coding strategies used by the auditory system. Compared with rate codes, temporal codes are often found to more accurately represent sounds for both pure AM and natural vocalizations with large AM components (e.g., Huetz et al. 2009; Narayan et al. 2006; Schnupp et al. 2006; Walker et al. 2008; Wang et al. 2007). Modeling approaches to AM detection have been employed with simulated auditory nerve (Goldwyn et al. 2010) and inferior colliculus (IC) (Lorenzi et al. 1995) neurons, but neurometric approaches, which use actual physiological data, have recently been attempted as well. Using a signal detection neurometric to examine the ability of rabbit IC neurons to detect AM tones, Nelson and Carney (2007) found median thresholds for spike rate measures near 30% modulation depth, while synchrony-based measures had a lower (that is, more sensitive) median. Malone et al. (2010) used a spike train classification method on macaque primary auditory cortex (A1) responses to AM tones and also found that spike timing information was more sensitive than firing rate in the detection of AM. However, both of these neurometric studies used stimuli whose duration (5–10 s) is much longer than the required temporal integration window for AM detection—improvement in peak AM sensitivity is mostly complete in both humans and macaques for 800-ms stimuli (O'Connor et al. 2011).

It is difficult to interpret how a single neuron's response relates to behavior because the auditory system has access to large numbers of neurons. In both the visual system of macaques (Shadlen et al. 1996) and the auditory system of ferrets (Bizley et al. 2010) neuronal pooling models have provided satisfactory

Address for reprint requests and other correspondence: M. L. Sutter, Center for Neuroscience, Univ. of California at Davis, 1544 Newton Ct., Davis, CA 95618 (e-mail: mlsutter@ucdavis.edu).

explanations of psychophysical results. In the auditory system, the analysis of data collected during long AM tone presentations (as in Malone et al. 2010; Nelson and Carney 2007) is somewhat analogous to pooling because the individual time epochs extracted from the long record are similar to individual neurons. Another recent study using simulated pooling of AM responses of auditory cortex units in gerbils suggests that indiscriminate pooling, rather than a most-sensitive-neuron approach, may best explain improved AM tone detection in adults relative to juveniles (Rosen et al. 2010).

In this study, we provide a detailed look at single-unit responses in awake macaque A1 to AM noise stimuli across a range of modulation depths, under conditions that can be readily compared with prior behavioral results in both macaques and humans. We selected auditory cortex for this study because lesion, psychophysical, and physiological studies support its role in the perception of temporal sound properties. Auditory cortical lesions produce deficits in simple serial discriminations (Carmon and Nachshon 1971; Fitch et al. 1994) and impairments in temporal resolution (Auerbach et al. 1982; Efron et al. 1985; Ison et al. 1991; Ison and Bowen 2000; Kelly et al. 1996; Phillips and Farmer 1990), including higher temporal fusion thresholds (Lackner and Teuber 1973), indicating that cortex is necessary for normal temporal event processing. In addition, auditory cortical lesions can result in difficulty in extracting signals from noise (Heilman et al. 1973; Olsen et al. 1975), which is conceptually similar to detecting modulation in a partially modulated sound. AM noise has the advantage of eliminating spectral cues, a possible confound in AM tone detection, and of having been used extensively in psychophysical studies (Burns and Viemeister 1976). We applied signal detection methods to determine neural AM noise detection thresholds. We also compared single-unit and pooled thresholds, using both rate and temporal (vector strength based) coding schemes in order to determine whether pooling mechanisms can result in a good approximation to behavior.

METHODS

Subjects and recording. All procedures conformed to U.S. Public Health Service policy on experimental animal care and were approved by the UC Davis institutional animal care and use committee. Animal care and surgical procedures were similar to previously described studies (O'Connor et al. 2005; Yin et al. 2011). A summary, including any changes from those procedures, is provided for brevity. One male (*monkey Y*) and one female (*monkey V*) rhesus monkey (*Macaca mulatta*) weighing 6–8 kg served as subjects. Monkey letter identifiers are consistent across all publications from M. L. Sutter's lab. Monkeys were implanted with a head post and a recording chamber for chronic access to auditory cortex. A plastic grid (Crist Instruments) was fit into the recording chamber to guide the electrode penetrations. For each recording session, a remotely controlled hydraulic microdrive (FHC) was used to insert a high-impedance tungsten microelectrode (FHC) into the cortex through guide tubes held by the plastic grid. Unit recordings were made while the monkeys sat quietly under head restraint in an acoustically "transparent" primate chair (custom made, Crist Instruments) in a double-walled, sound-attenuated, foam-lined booth (IAC: 9.5 ft. × 10.5 ft. × 6.5 ft.), with diluted juice or water administered intermittently.

All recording sites in both animals were located in the right hemisphere of A1 (as detailed in Yin et al. 2011; all 123 cells recorded here were also members of the 182-cell data set of Yin et al.). Briefly, electrode locations were confirmed by histology (border points of recording marked with biotinylated dextran amine) in one monkey

(*monkey V*) and tonotopic and latency gradients (5 repeats of 150 tone pips, 0.1-s duration, 10 intensities between 15 and 78 dB SPL, 15 frequencies spanning 3 octaves around a hand-tuned center frequency) in both monkeys. The extent of A1 was conservatively estimated at the low-frequency border between A1 and more rostral areas.

Stimulus generation and data collection. Sinusoidal amplitude-modulated (AM) stimuli were created with a "frozen" broadband noise carrier and were 400 ms in duration. AM stimuli were created at a variety of modulation frequencies (5, 10, 15, 20, 30, 60, and 120 Hz) and modulation depths (6%, 16%, 28%, 40%, 60%, 80%, and 100% depth). The modulation envelope was determined by

$$1 - m + m \times \left(\left(\frac{\sin(2\pi f_m t - \pi/2)}{2} \right) + 0.5 \right) \quad (1)$$

where m is the modulation index (ranges from 0 to 1; % modulation depth is $m \times 100$), f_m is the modulation frequency in hertz, and t is the waveform time in seconds.

The sound signals were generated by a digital signal processor (AT&T DSP32C) and a digital-to-analog converter (TDT Systems DA1) and then passed through programmable and passive attenuators (TDT Systems PA4, Leader LAT-45). The signal was amplified (Radio Shack MPA-200) before being delivered to a speaker (Radio Shack PA-110, 10-in. woofer and piezo-horn tweeter, 10-dB cutoff: 0.038–27 kHz) positioned at ear level 1.5 m in front of the subject.

The stimuli were presented at a sampling rate of 50 kHz (2.5- to 25-kHz digital bandwidth, analog bandwidth limited by 27-kHz 10-dB cutoff point of the speaker) and were cosine ramped at onset and offset (5.0-ms rise/fall time). Stimulus intensity was adjusted to ~65 dB SPL (<2-dB variation). Extracellular potentials were amplified and filtered (0.3–5 kHz; A-M Systems 1800), sampled at 50 kHz, and stored on hard disk for later analysis. Spikes were resorted off-line with Spike2 software (CED). All numerical analysis was done with custom software written for MATLAB (MathWorks). For all calculations, only spikes occurring between 70 and 400 ms after the onset of the stimulus were included, to eliminate any contribution of an onset response. Including the onset response did not substantially alter the results.

At each recording site, neurons were assessed with at least two different batteries of stimuli. To determine the best modulation frequency (BMF) of the multiunit, 100% depth AM stimuli were presented at each modulation frequency and modulation transfer functions (MTFs) were created by taking the mean spike count (SC) or phase-projected vector strength (VS_{PP}) across all trials. MTFs were calculated separately for rate and temporal measures. We defined the temporal BMF (tBMF) as the point with the highest mean VS_{PP} . Because of the number of cells that decreased their activity in response to AM relative to their noise response, the rate BMF (rBMF) was defined as the modulation frequency that evoked the mean SC furthest from the noise response [as measured by the distance of the receiver operator curve (ROC) area from 0.5; ROC area is described in the next section], regardless of whether that modulation frequency resulted in an increase or decrease relative to the noise response. A second battery of stimuli was used to determine the depth sensitivity of the same cells for a single modulation frequency. During each recording we attempted to measure depth sensitivity at several modulation frequencies: the multiunit (MU) tBMF, the MU rBMF, and 15 Hz. For some units we were able to obtain sensitivity functions at all three modulation frequencies, but for many we were not able to maintain a stable recording and only obtained data for one or two modulation frequencies. Because the BMF was not identical for all single units (SUs) within a MU and because most cells were tested at 15 Hz, a substantial number of recordings were not at BMF.

Depth sensitivity: neurometric analysis. One hundred twenty-three unique cells were tested with a neurometric analysis for AM depth sensitivity. Because many cells were tested at more than one modulation frequency, we obtained a total of 249 measurements of depth

sensitivity. Cells were tested at modulation depths of 0%, 6%, 16%, 28%, 40%, 60%, 80%, and 100%. Modulation depths were chosen to conform to the log-scaled depths in O'Connor et al. (2000), with points added near the expected psychophysical threshold (28%) and at high depths (60%, 80%) in case we encountered cells with high thresholds. For most cells, 50 blocks of trials were presented, with each block consisting of 8 stimuli, one presentation of each of the 8 modulation depths in a randomized order. (For 2 of 249 recordings there were 100 blocks, and for 2 of 249 recordings there were only 30 blocks.)

At each nonzero modulation depth, we calculated the area under the receiver operating curve (ROC area; Green and Swets 1966), for both SC and VS_{pp} (defined in the next section), comparing responses to the modulated stimulus to responses to the unmodulated stimulus. The ROC area measures how well the neural response on a trial-by-trial basis can distinguish between two stimuli (in our case a noise at a particular modulation depth vs. the unmodulated noise) for a given metric (SC or VS_{pp}). An ROC area of 1 means that for every stimulus presentation (trial) the value of the metric was larger for the modulated than the unmodulated sound—simply by observing the value of the metric on a given trial, an ideal observer would predict with 100% accuracy whether a noise was modulated or unmodulated. An ROC area of 0.5 indicates that an ideal observer would perform at chance. ROC area is symmetrical around 0.5, such that an ROC area of 0 indicates that for every trial the value of the metric was smaller for the modulated than the unmodulated sound. This will also result in 100% accuracy in predicting which of the two sounds was presented, as long as the ideal observer associates small values of the metric with the modulated sound.

The ROC is a plot of hit probability (y-axis) against false alarm probability (x-axis), which we calculated at each of 100 equally spaced decision criteria ranging between the lowest and highest observed response values (SC or VS_{pp}) for the two stimuli being compared. At each criterion level the proportions of hits and false alarms were calculated from the neural responses, where hits are modulated-sound trials on which the metric gives a greater value than the criterion, and false alarms are unmodulated-sound trials on which the metric gives a greater value than the criterion. ROC area was calculated as the trapezoidal area (MATLAB: "trapz") under this ROC. Mathematically, the ROC area is equivalent to the probability that a randomly selected trial from the modulated-sound distribution will have a response value (SC or VS_{pp}) larger than a randomly selected trial from the unmodulated-sound distribution. In Fig. 1 we plot the single-trial distributions of both SC and VS_{pp} values for unmodulated, 16% depth, and 100% depth stimuli from a single example recording (rasters for this recording are shown in Fig. 2A).

ROC area was plotted as a function of modulation depth for each unit. Neurometric depth sensitivity functions were created with an automated curve-fitting procedure on the obtained ROC area vs. depth functions. Because some cells were nonmonotonic to depth (i.e., showed higher firing rates or better phase locking for intermediate modulation depths than for fully modulated stimuli; cf. Middlebrooks 2008a) our curve-fitting procedure included both logistic (Eq. 2) and Gaussian (Eq. 3) functions:

$$y = a + \left(\frac{b}{1 + e^{-\frac{x-\mu}{s}}} \right) \tag{2}$$

$$y = a + b \times e^{-\frac{(x-\mu)^2}{2s^2}} \tag{3}$$

Both curves used four free parameters determining the y-offset (*a*), the height (*b*), the x-center (*μ*), and the slope (*s*). Fitting was performed with MATLAB's "fmincon" function, constraining the slope (*s*) of the logistic between 2 and 20 and constraining the height (*b*) of the Gaussian to six times the difference between the ROC area at the lowest modulation depth and the ROC area at the most responsive modulation depth. These constraints were sufficient to eliminate the small number of spurious fits found by manual examination in the unconstrained case. To avoid overfitting with the Gaussian, we calculated an absolute ROC value (IROC area - 0.5) for each modulation depth. If the value at 100% depth was more than 7/8 of the cell's maximal value, only the logistic function was used in fitting. In the case where both Gaussian and logistic fits were attempted, the final fit was chosen on the basis of the highest correlation coefficient between the two curve fits and the data. Threshold was taken as the point where the fitted curve crossed an ROC area of 0.75 (or 0.25 for a declining function). In the case of a Gaussian curve, only the first threshold crossing was used. Using the relationship between the ROC area and the Mann-Whitney *U* (e.g., Hand and Till 2001; Hanley and McNeil 1982), we determined the *P* value corresponding to our ROC area threshold of 0.75 (0.25) where the null hypothesis is that the test statistic (SC or VS_{pp}) does not allow us to distinguish between modulated and unmodulated trials. Our threshold corresponds to *P* = 8.3 × 10⁻⁶ for cases where 50 trials were collected, and for the rare cases (*n* = 4) where either 30 or 100 trials were collected the threshold corresponds to *P* = 4.5 × 10⁻⁴ and *P* = 5.1 × 10⁻¹⁰, respectively.

Phase-projected vector strength. VS_{pp} is a trial-by-trial vector strength measure that compares the mean phase angle on each trial with a global reference phase angle and penalizes single trials that are not in phase with the global response by multiplying the vector strength value by the cosine of the phase difference (Yin et al. 2011). It is calculated as follows:

$$VS_{pp} = \cos(\varphi - \varphi_c) \times \frac{\sqrt{(\sum_{i=1}^n \cos\theta_i)^2 + (\sum_{i=1}^n \sin\theta_i)^2}}{n} \tag{4}$$

where *n* is the number of spikes, *θ_i* is the phase of each spike in radians, and *φ* and *φ_c* are the trial-by-trial and global mean phase angles, respectively. For each trial, the resulting VS_{pp} value is always less than or equal to a non-phase-corrected vector strength value. Except for trials with low spike counts, VS and VS_{pp} are typically similar. We applied VS_{pp} specifically to reduce spuriously high VS values that can occur for trials with low spike counts, so that we could use all trials in neurometric analysis rather than discard them. Trials with no spikes were assigned a VS_{pp} of zero.

Increasing and decreasing spike count functions. At a given modulation frequency, cells could either increase or decrease their responses with increasing modulation depth. To classify depth sensitivity curves as increasing (SC_{inc}) or decreasing (SC_{dec}), we took the

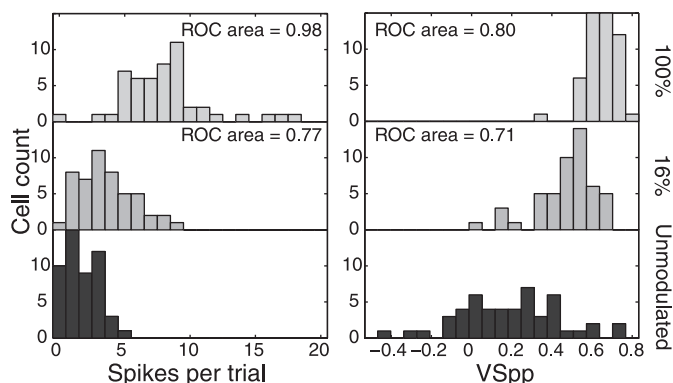


Fig. 1. Illustration of distributions underlying receiver operator (ROC) area calculations. *Left*: spike count. *Bottom*: histogram of spike count values for 50 trials of unmodulated noise for a sample recording. *Middle*: 16% modulation depth. *Top*: 100% modulation depth. *Right*: same as *left*, using phase-projected vector strength (VS_{pp}). Sample distributions are taken from cell V2475-2, which is fully illustrated in Fig. 2A.

average of the calculated ROC areas at each modulation depth. If this average was >0.5 the cell was classified as increasing and if it was <0.5 the cell was classified as decreasing, regardless of whether the cell ever reached the 0.75 (or 0.25) threshold.

Weighted mean of ROC area. To determine the effect of testing away from the BMF we calculated the mean ROC area for VS_{pp} , SC_{inc} , and SC_{dec} as a function of distance (in octaves) from the cell's BMF. Because this octave space was not evenly sampled, the number of observations at each distance from the BMF was quite variable. Consequently, ROC area as a function of distance from the BMF was smoothed as follows:

$$y_j = \frac{\sum_i w_i \mu_i}{\sum_i w_i} \quad (5)$$

where w_i is a Gaussian weighting function

$$w_i = n_i e^{-\frac{(x_i - x_j)^2}{2\sigma^2}} \quad (6)$$

and x_i is a discrete list of sampled points on the x -axis, x_j is the x -location of the point in question, n_i is the number of observations at each x value, and μ_i is the mean ROC area at each x value. We selected σ to be 1 octave. For this analysis, points at extreme distances from the BMF were considered to be outliers and not included if they were at least 1 octave away from the next-nearest data point. Three of 249 measurements were excluded from the VS_{pp} functions, 2 of 140 from the increasing SC functions, and 4 of 109 from the decreasing SC functions, and the points included in the plots reflect the extent of the nonoutlying points.

The general shape resulting from this weighted mean of ROC area analysis was Gaussian-like. Therefore, to assess whether these curves showed significant structure in the data (i.e., whether they were different from a flat function), we designed the following ad hoc Monte Carlo analysis. Each resulting curve was fit with a Gaussian function, and the height parameter of the curve fit was recorded. Then, for each of 10,000 repeats we randomly scrambled the relationship between the ROC area value and the distance from BMF value and repeated the calculation in Eq. 5, again fitting with a Gaussian and recording the height parameter of the curve fit. The resulting P value was taken as the probability that the curve fit of the scrambled data produced a larger height on the Gaussian curve fit than that produced by the actual data.

Trial pooling: within cells. For some analyses, we pooled trials to simulate the convergence of multiple presynaptic cells with similar response properties onto a single postsynaptic cell (cf. Schneider and Woolley 2010). For within-cell pooling, trials from each cell's set of responses were pooled together under the premise that reduced noise in pooled trials might result in greater discriminability. Within-cell pooling results in an inherent trade-off between the number of trials available for ROC analysis and the amount of pooling—if the cell starts with 50 trials, pooling by twos will result in 25 trials, pooling by threes will result in 16 trials, and so on.

A pooled trial was defined as the union of the spike times recorded in two or more individual trials:

$$P_x = \bigcup_i T_i \quad (7)$$

where P_x is a list of spike times in pooled trial x and T_i is a list of spike times in individual trial i . For within-cell pooling each individual trial was used exactly once without replacement. Trials were distributed in the temporal order of collection—conceptually equivalent to dealing out a deck of cards (individual trials) to each player (pooled trial) until the deck is out. The pooling number n_p then corresponds to the smallest number of individual trials dealt to any pooled trial—some pooled trials will have one extra individual trial if the number of pooled trials does not divide evenly into the number of individual trials collected. This arrangement ensures that the trials comprising

each pooled trial are evenly spaced over the time of collection as nearly as possible. Pooled trials were then analyzed in the same manner as individual trials. This method is similar to that used in Middlebrooks (2008b), where four trials per data point were pooled in order to mitigate low spike count issues associated with vector strength analyses.

Trial pooling: across cells. In addition to within-cell pooling, we also implemented an across-cell pooling technique that may model the convergence of multiple presynaptic cells with different properties. For across-cell pooling we performed an unbiased sampling from (almost) all collected cells, which might accurately describe in vivo connection patterns if cells are connected in a nonselective fashion. Because of the low number of cells tested at 5 Hz (15 total) and 120 Hz (7 total), these modulation frequencies were excluded from across-cell pooling. Pooling was carried out in a fashion similar to within-cell pooling:

$$P_x = \bigcup_{i=1}^{n_p} C_{ix} \quad (8)$$

where a total of n_p cells are randomly sampled, with replacement, from the population of all cells tested at a particular modulation frequency. After random selection of a cell, the collected trial order at each modulation depth for that cell was randomly scrambled to generate cell C_{ix} , which contained 50 trials (x ranging between 1 and 50) at each modulation depth. (For the 2 cells with 100 repetitions, after randomizing the order only the first 50 were taken; for the 2 cells with only 30 repetitions, the final 20 trials were taken from a rerandomization of the first 30.) Thus, even if the same cell were selected twice within a pool, the "trial order" for that cell would not be the same and the probability that two identical trials would be pooled together in any individual P_x would be low. Each P_x was therefore a list of 50 pooled trials (pooled across n_p cells) at a particular modulation depth. The ROC area was calculated for each P_x from nonzero modulation depths relative to the P_x from responses to unmodulated stimuli, and a threshold was calculated by fitting Eq. 2 and Eq. 3 to these ROC areas as a function of modulation depth as above. This procedure was repeated 1,000 times for each n_p cells (n_p sampled between 1 and 50), and the mean numbers of pools reaching threshold and mean threshold were calculated for each n_p and modulation frequency. The data presented here have been subsequently weighted and averaged across modulation frequency to provide a single summary value at each n_p (but note that no 5-Hz or 120-Hz test frequencies were included).

RESULTS

Depth sensitivity functions. We tested 123 cells, often at multiple modulation frequencies (resulting in 249 tests overall), for sensitivity to depth of sine-modulated noise. For each modulation frequency, neural responses were collected at depths of 0%, 6%, 16%, 28%, 40%, 60%, 80%, and 100%. At each depth the ROC area was calculated for both spike count (SC) and phase-projected vector strength (VS_{pp} , see METHODS for stimulus and calculation details) without regard to whether a neuron's response was synchronized to the AM. The ROC area corresponds to the probability that, given a random draw of one spike train from each stimulus type (unmodulated, modulated at $x\%$), we would find that the SC (or VS_{pp}) value is higher for the modulated trial than the unmodulated trial. ROC area measurements were then fitted with a logistic (or in some cases Gaussian, see METHODS for details) curve. We chose an ROC area of 0.75 on our fitted curves as a threshold value indicating the modulation depth at which the cell was reliably able to distinguish modulated from unmodulated stimuli.

We found cells with both increasing and decreasing SC sensitivity functions, as well as differing degrees of synchronization as measured with VS_{PP} . Example raster plots and depth sensitivity functions are shown in Fig. 2. Figure 2A depicts the responses to 60-Hz modulation of a cell that reaches threshold for both rate (SC) and temporal (VS_{PP}) measures. In

this example, both SC and VS_{PP} are quite sensitive to modulation, with neurometric thresholds better than typical macaque behavioral thresholds of $\sim 20\text{--}30\%$ to 60-Hz AM noise (O'Connor et al. 2011, Fig. 2). Figure 2B depicts a cell that increases phase locking to 30-Hz AM, but decreases SC, as modulation depth increases. Cells with decreasing spike counts

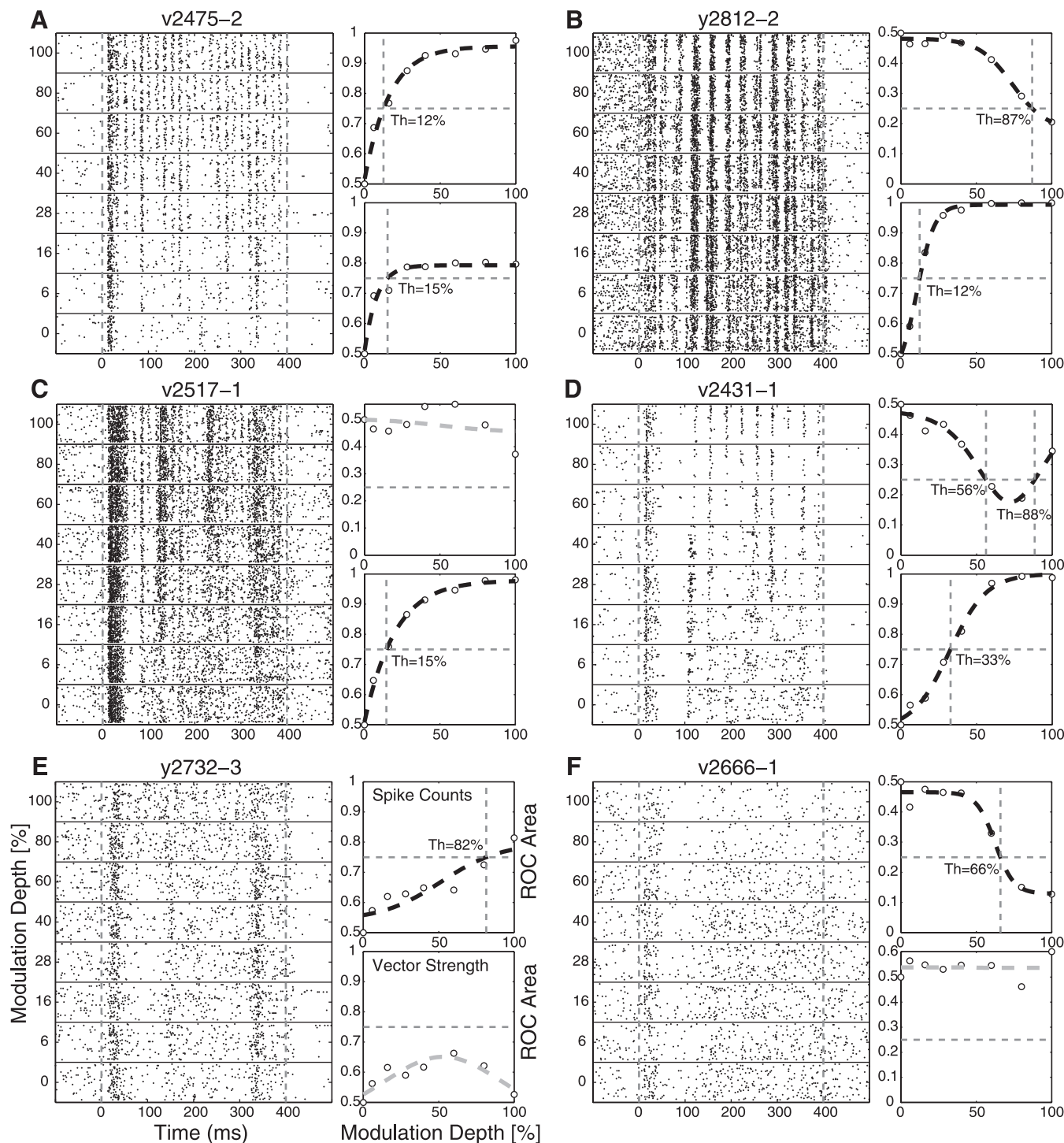


Fig. 2. Example neurons. *Left:* raster plots show responses to modulation depths between 0% (unmodulated) and 100% (fully modulated). *Right:* ROC area plots (*right*) show the ROC area measured for each modulated vs. unmodulated comparison. *Top:* spike count. *Bottom:* VS_{PP} (see E for labels). Heavy dashed line is best sigmoid or Gaussian fit to the data, marked in black if the fit is significant at the 0.01 level and gray if not. Threshold (0.75, or 0.25 for declining functions) is indicated by horizontal dashed lines. Vertical dashed lines mark threshold level. Th, % modulation at threshold. Test frequencies: 60 Hz (A), 30 Hz (B), 10 Hz (C), 30 Hz (D), 15 Hz (E), and 120 Hz (F).

to increased modulation frequency were relatively common and are discussed later. This cell (as well as many others) has an intrinsic temporal structure in its response to unmodulated noise, and the effect of increasing modulation depth appears to be the removal of extraneous spikes that are not in phase with the stimulus modulation. The intrinsic temporal structure of the response to unmodulated noise may be due to the use of frozen noise stimuli. Figure 2C shows a cell that does not change its firing rate in response to changes in modulation depth but becomes more phase locked to 10-Hz AM as the depth of modulation increases. Since SC does not change, as modulation depth increases the only change is in the temporal structure. Figure 2D shows a cell that exhibits a decreasing non-monotonic SC depth sensitivity function while having an increasing temporal function. This cell has the same general pattern as the cell in Fig. 2B except that it does not have an intrinsic temporal structure in its response to unmodulated noise. Figure 2E shows a cell that increases its SC with increasing modulation depth but does not significantly change the phase locking of its spikes. This nonsynchronized type of response has been hypothesized as important in the neural transformation of modulation encoding (Liang et al. 2002; Lu et al. 2001; Lu and Wang 2004). Figure 2F shows a cell that decreases its SC with increasing modulation depth but also does not show a change in phase locking. These examples are representative of the variety of paired rate and temporal response characteristics we found in our population of neurons.

Depth sensitivity functions: population statistics. Overall, 84 of our 249 depth sensitivity functions (34%) reached the ROC area threshold of 0.75 (or 0.25 for decreasing functions) for SC, while 128 (51%) reached threshold for VS_{PP} (0.75; only increasing functions were observed). It is important to note that each depth sensitivity function was tested with both SC and VS_{PP} measures—56 depth sensitivity functions (22%) reached threshold for both measures, while 93 (37%) did not reach threshold for either measure. A significantly higher proportion of depth sensitivity functions reached threshold for VS_{PP} than SC measures ($P = 6.7 \times 10^{-5}$, z -test for 2 independent proportions). Figure 3 shows the distribution of depth sensitivity functions reaching threshold, broken down by the modulation frequency tested. At most tested modulation frequencies, VS_{PP} resulted in a greater percentage of depth sensitivity functions reaching threshold than SC, although there is a notable exception at a test frequency of 120 Hz (VS_{PP} 0/7,

compare SC at 4/7), which is expected because of the decreased ability of A1 neurons to strongly phase lock to higher modulation frequencies and the increased percentage of nonsynchronized responders at these frequencies (Lu et al. 2001; Yin et al. 2011). As an estimate of exclusively nonsynchronized responders, we found that 28 of our 84 depth sensitivity functions reached threshold for SC but not for VS_{PP} (11% of all functions).

Many cells in A1 respond to increasing modulation depth by decreasing rather than increasing SC. Both increasing and decreasing rate functions for modulation depth encoding have been reported previously for AM tones in gerbil IC (Krishna and Semple 2000). We attempted to test each unit at the modulation frequency that best distinguished AM from unmodulated noise, regardless of whether that was due to an increase or a decrease in SC. We classified each of our depth sensitivity functions as increasing or decreasing on the basis of the mean ROC area across all modulation depths tested (depth sensitivity functions with a mean ROC area >0.5 were increasing and <0.5 decreasing). Of 140 increasing functions, 55 (39%) reached threshold (threshold ROC area = 0.75). Of 109 decreasing functions, 29 (27%) reached threshold (threshold ROC area = 0.25). The proportion of increasing depth sensitivity functions (55/140) reaching threshold significantly differs from the proportion of decreasing functions (29/109) reaching threshold (z -test for 2 independent proportions, $P = 0.002$), suggesting that increasing SC functions are more sensitive to modulation depth than decreasing SC functions.

However, increasing and decreasing depth sensitivity functions do not appear to differ in mean threshold (Fig. 4A). There was no significant difference between the mean thresholds of increasing and decreasing functions (increasing 65%, decreasing 62%, 2-sided t -test, $P = 0.45$, not significant). Accordingly, all analyses include both increasing and decreasing depth sensitivity functions except where noted.

The sensitivities of the A1 neurons that reached threshold for SC and VS_{PP} are compared in Fig. 4, B and C, left. If we look at all cells that reached threshold regardless of the modulation frequency of the test stimulus, we find a mean threshold of 64% depth for SC, while for vector strength the mean threshold is 52% (2-sample t -test, $P = 7 \times 10^{-4}$). These mean neuro-metric thresholds are relatively high compared with the macaque's psychophysical threshold of ~ 20 –30% across the range of frequencies tested. If we look only at SC sensitivity

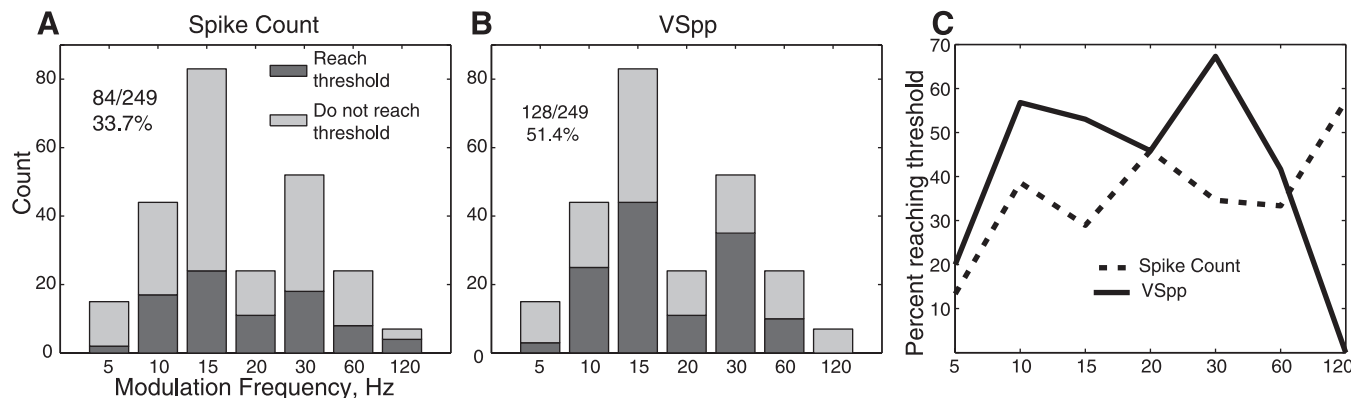


Fig. 3. Population statistics. A: total number of cells reaching detection threshold (ROC area >0.75 or <0.25), broken down by amplitude-modulated (AM) test frequency, with spike count as a test measure. B: as in A, with VS_{PP} as the test measure. C: data in A and B presented in terms of % cells reaching threshold.

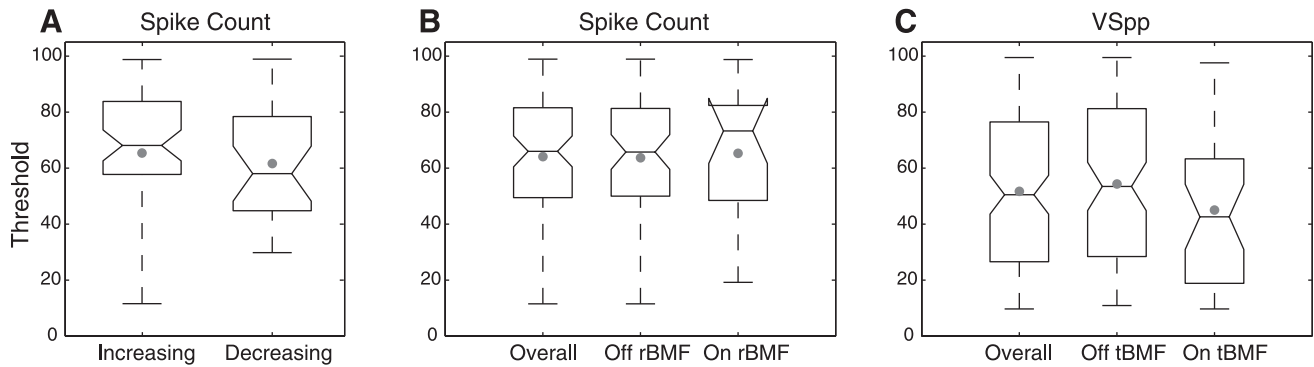


Fig. 4. Threshold distributions. Gray dots indicate mean thresholds. Central horizontal lines of box-and-whisker plots indicate medians, upper and lower edges of boxes indicate quartiles, and whiskers indicate most extreme data points. The notch is a comparison interval; 2 medians are significantly different if the notches do not overlap; the notch may extend beyond the quartile. **A:** spike count thresholds of increasing ($n = 55$) and decreasing ($n = 29$) cells. **B:** spike count thresholds for all ($n = 84$) recordings, recordings made off the rate best modulation frequency (rBMF) ($n = 63$) and on the rBMF ($n = 21$). **C:** VS_{pp} thresholds for all ($n = 128$) recordings, recordings made off the temporal BMF (tBMF) ($n = 36$) and on the tBMF ($n = 92$).

for cells that do not reach threshold for VS_{pp} (an estimate of the exclusively nonsynchronized population), we find that the mean threshold is 70%, which is not statistically different from the mean SC threshold of 61% for the cells that also reached threshold for VS_{pp} ($P = 0.06$, t -test), suggesting that SC sensitivity is not strongly different in cells that synchronize and cells that do not.

Depth sensitivity functions: relation to modulation transfer functions. It is reasonable to expect that neurons might be more sensitive to depth at the modulation frequency to which they respond most strongly. To examine this, we looked separately at cells that were recorded at their best modulation frequency (BMF), and off their BMF. A total of 46 recordings were made at the rate BMF (rBMF), and 203 recordings were made off the rBMF. On-rBMF recordings showed 21 cells (46%) reaching threshold, while off-rBMF recordings revealed 63 cells (31%) reaching threshold. These proportions are not significantly different (z -test for independent proportions, $P = 0.06$). The mean threshold measured off rBMF was 64%, while the mean on-rBMF threshold was 65% (Fig. 4B). These mean values are not significantly different (2-sample t -test, $P = 0.78$). Looking at temporal measures, 36/51 (71%) recordings made at the temporal BMF (tBMF) reached threshold while only 92/198 (46%) recordings off the tBMF did. These proportions are

different (z -test, $P = 0.002$); therefore cells are more likely to reach threshold for depth encoding with VS_{pp} if they are tested at their tBMF. Despite the fact that significantly more cells reach threshold when measured at their tBMF, the mean thresholds in the on-tBMF and off-tBMF cases (Fig. 4C) do not significantly differ (54% off-tBMF vs. 45% on-tBMF, 2-sample t -test, $P = 0.09$).

Similar population thresholds for on- and off-BMF recordings initially suggest that cells have little advantage for encoding depth if the modulation frequency of the signal is at their preferred frequency. However, an alternate possibility is a recruitment effect—it might be that some cells that would not reach threshold if tested off BMF (so they would not contribute to the off-BMF threshold calculation) do reach threshold on BMF. Because depth encoding is weaker in these cells, their on-BMF thresholds would be relatively high and counteract on-BMF improvements in cells that also reach threshold off BMF. If this were the case, we would expect that a paired within-cell comparison would show better thresholds for on-BMF recordings. Figure 5 shows a scatterplot of thresholds taken from on-BMF recordings against thresholds taken off BMF in the same cells. There were 41 cells (Fig. 5A) that had depth thresholds measured both at the rBMF and at least one other frequency (47 total paired thresholds because some cells

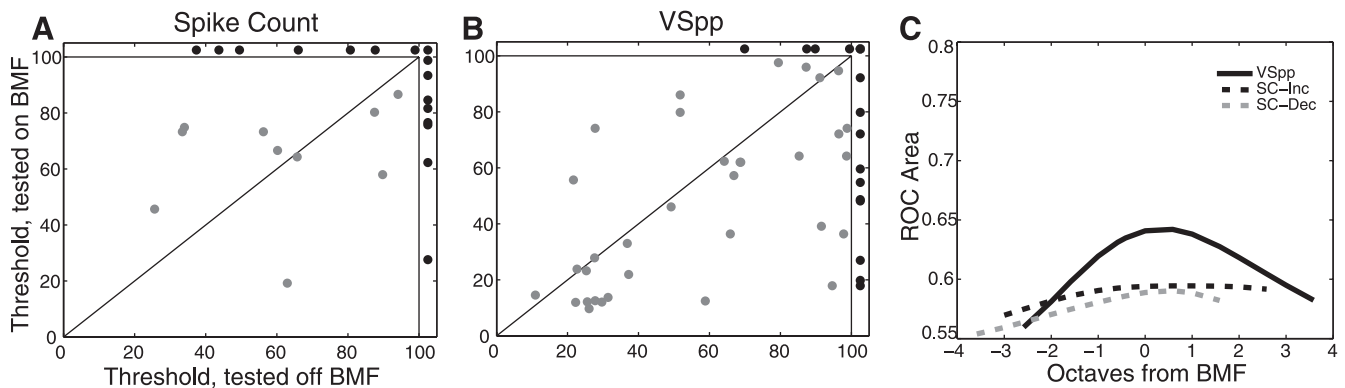


Fig. 5. Changes in threshold for recordings made at and away from BMF. For **A** and **B**, only cells that had recordings made both at the BMF and away from the BMF are included. **A:** spike count thresholds, 47 pairs. **B:** VS_{pp} thresholds, 67 pairs. Cells that lie above the horizontal line at 100% did not reach threshold when tested on BMF. Cells that lie to right of the vertical line at 100% did not reach threshold when tested off BMF. **C:** mean ROC area measured for the comparison of 40% modulation depth against the unmodulated case as a function of distance of test frequency from BMF in octaves. Lines are weighted running averages (see METHODS). Values for decreasing spike count curves (SC-Dec) have been reflected about ROC area = 0.5 to bring them into register with the other 2 measures. BMFs for SC-Dec correspond to the minimum of the modulation transfer functions. SC-Inc, increasing spike count curve.

were recorded at more than one off-rBMF frequency) and 51 cells (Fig. 5B) that had depth thresholds measured both at the tBMF and another frequency (67 total paired functions). Points plotted outside and to the right of the box in Fig. 5 represent cells that reached threshold for the on-BMF test but not for the off-BMF test; points plotted outside and above the box represent cells that reached threshold for the off-BMF test but not for the on-BMF test (note that 20 recordings for rBMF and 14 recordings for tBMF did not reach threshold for either test and are plotted on top of each other at top right). If there were a trend toward lower thresholds at the BMF, it would result in points being clustered below the unity line. In the case of “recruited” cells that reach threshold at the BMF but not at another test frequency, the points would cluster to the right of the box. We performed a binomial sign test with the null hypothesis that the median change in threshold is not different from zero in the on-BMF and off-BMF cases. A binomial test (in contrast to, for instance, a paired *t*-test) allowed us to consider cells that only reached threshold in one of the two cases as having a threshold improvement (or decrement) without assigning an arbitrary threshold value to cells that did not reach threshold. For SC we were unable to reject the null hypothesis ($P = 0.70$), but for VS_{PP} cells showed a clear improvement ($P = 8 \times 10^{-4}$) when tested on rather than off BMF.

Several factors could lead to a weak SC effect. One possibility is an unintentional sampling bias since rBMFs tend to be at higher modulation frequencies than tBMFs. A second possibility is that the available statistical power is reduced because we have few neurons that reach threshold for both rBMF and tBMF. To alleviate these problems and to look at the effect of testing on or off BMF across our entire data set, we recoded the test modulation frequency in terms of octave distance and direction from the BMF for all neurons and then calculated the mean ROC area for the 40% depth stimuli for each distance from the BMF, smoothing with a Gaussian-weighted average (see METHODS); 40% depth stimuli were chosen for this analysis because they were slightly below the population threshold and above typical behavioral thresholds and we were looking for effects that resulted in an improvement of threshold. The result is shown in Fig. 5C, with separate results for cells with increasing or decreasing SC functions. For decreasing SC functions only, the value plotted is $1 - \text{ROC area}$ in order to bring the values into register with the other functions and the BMF was considered to be the minimum, rather than the maximum, in the modulation function. Significance of the curve fits (i.e., whether there was a dependence on distance from BMF) was assessed by a Monte Carlo analysis (see METHODS). For VS_{PP} , mean ROC area at 40% depth shows a significant peak value near the BMF ($P < 0.02$) and rolls off on either side over a range of several octaves, suggesting that there is a decline in depth encoding as test frequencies move away from the BMF. However, the SC functions are quite flat compared with the VS_{PP} function (neither SC function significant; increasing: $P > 0.2$, decreasing: $P > 0.05$), suggesting that testing at or near the BMF should result in more threshold improvement relative to off-BMF tests for temporal measures than for SC measures, just as we find in Fig. 5, A and B. Together these results indicate that for VS_{PP} there is decreased neural sensitivity to modulation off BMF that declines as a function of distance. Little or no BMF relationship is found with SC.

Depth sensitivity functions: pooling/multiunit population data. We have shown, using SC and VS_{PP} metrics, that mean individual single units are less sensitive to depth than the behaving animal. Since the animal has a large pool of neurons to draw on to detect modulation, we examined the effects of several types of pooling. In the first case we looked at the multiunit (MU) rather than the single-unit (SU) response. In a typical recording we found two to four isolatable cells. Although they did not always have identical properties, the columnar organization of A1 suggests that they should make similar connections, so the MU recording may effectively simulate one type of (postsynaptic) neural pooling. Alternatively, the auditory system may pool neurons that are more alike in response properties than those found in a MU cluster. To simulate this possibility, we took multiple spike trains produced on different trials by the same neuron and pooled them (“within-cell pooling,” see METHODS). Indiscriminate pooling (“across-cell pooling”) will be considered later.

Both MUs and within-cell pooling result in significant increases in the proportions of units reaching threshold relative to single units (SUs) for both rate (z -test: multiunit, $P = 0.01$; pooled, $P = 0.002$) and temporal (z -test: multiunit, $P = 0.002$; pooled, $P = 0.005$) measures. This can be seen by comparing Fig. 6 to Fig. 3. For VS_{PP} the increase in the number of units reaching threshold from SU to MU can be seen at every modulation frequency except 120 Hz.

A higher percentage of pooled-unit responses reached threshold for VS_{PP} than for SC, as was the case with SUs. The ROC area threshold (0.75) was reached for 42 of 85 MUs (49%) with SC, while 60 MUs (71%) reached threshold for VS_{PP} . The numbers of MUs reaching threshold for the two measures are significantly different ($P = 0.005$, z -test for two independent proportions). Figure 6, *left*, shows the distribution of MUs reaching threshold, as a function of modulation frequency. Aside from 120 Hz, where none of the three MUs reached threshold for VS_{PP} (and all 3 reached threshold for SC), temporal pooling resulted in a higher percentage of cells reaching threshold than rate pooling, a pattern that is similar to that seen for SUs. The same trend holds for cells with trials pooled in pairs (Fig. 6, *center*). Here, 118 of 249 pooled-trial responses (47%) reached threshold for SC and 159 (64%) reached significance for VS_{PP} . Again, temporal measures result in significantly higher numbers of pooled-trial responses reaching threshold than rate measures ($P = 0.0002$, z -test for 2 independent proportions), with similar advantages for all modulation frequencies except for 120 Hz.

The mean MU SC threshold was 61%, and the mean MU VS_{PP} threshold was 45% (Fig. 6, *right*). As for single units, the VS_{PP} mean is significantly lower than the SC mean (2-sample *t*-test, $P = 0.002$). For pooled trials, the mean SC threshold was 61% and the mean VS_{PP} threshold was 49%. Again, the means were significantly different (2-sample *t*-test, $P = 4 \times 10^{-4}$). However, none of the three SC thresholds (SU, MU, pooled) is significantly different from any other, nor are any of the three VS_{PP} thresholds different from each other.

Thus, although threshold is more commonly reached under these two pooling methods, this pooling does not result in a lower mean threshold. To determine whether pooling results in lower thresholds in individual cells, we created paired-threshold plots in Fig. 7. Figure 7, *left*, shows the MU threshold plotted against each SU threshold for simultaneously recorded

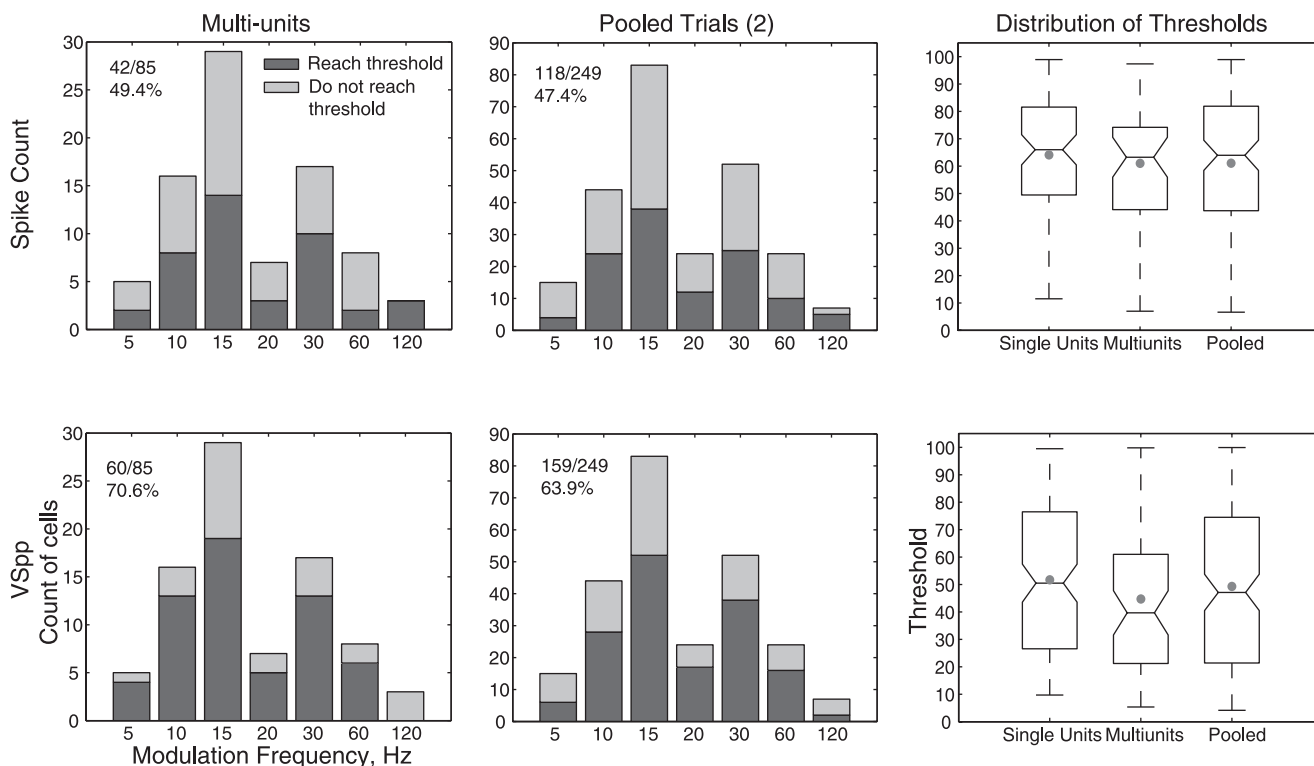


Fig. 6. Population statistics (left and center) and thresholds (right) for 2 types of pooling. Details of population histograms are as in Fig. 3. Details of threshold distributions are as in Fig. 4. Top: spike count. Bottom: VS_{pp} . Multiunits are the combined activity of all isolated cells (usually 2 or 3) recorded simultaneously at a single site. Pooled trials represent the combined activity of 2 separate trials taken without replacement from the same cell (see METHODS).

units. Here gray dots indicate cells/MUs that reached threshold for both SU and MU analysis, and black dots indicate cells that only reached threshold in one of the two cases (cells to the right of the graph only reached threshold in the MU case; cells above the graph only reached threshold in the SU case). For both SC and VS_{pp} , the majority of points lie below the unity line (or on the right side of the graph), indicating that thresholds of individual SUs improve when pooled as MUs. This improvement is statistically significant for both rate and temporal measures (Wilcoxon signed-rank test: SC, $P = 9.0 \times 10^{-7}$; VS_{pp} , $P = 3.5 \times 10^{-12}$).

Threshold improvement for individual cells is even more apparent for two-trial within-cell pooling (Fig. 7, right). Almost all points lie below the unity line for both SC and VS_{pp} , and the improvement compared with SUs was significant (Wilcoxon signed-rank test: SC, $P = 1.3 \times 10^{-13}$; VS_{pp} , $P = 6.3 \times 10^{-20}$). These results demonstrate that pooling methods do improve thresholds relative to those of individual cells on a cell-by-cell basis despite the lack of an improvement in overall mean threshold across all cells. As seen in the comparison of on-BMF and off-BMF thresholds above, the lack of improvement in overall mean threshold is due to a recruitment effect: Many cells that did not reach threshold (and were therefore not included in mean threshold calculations) in the SU case did reach threshold and were included in the mean threshold calculations after pooling. The addition of these pools, which generally have high thresholds, counteracts the general improvement seen for individual cells, washing this improvement out of the overall average.

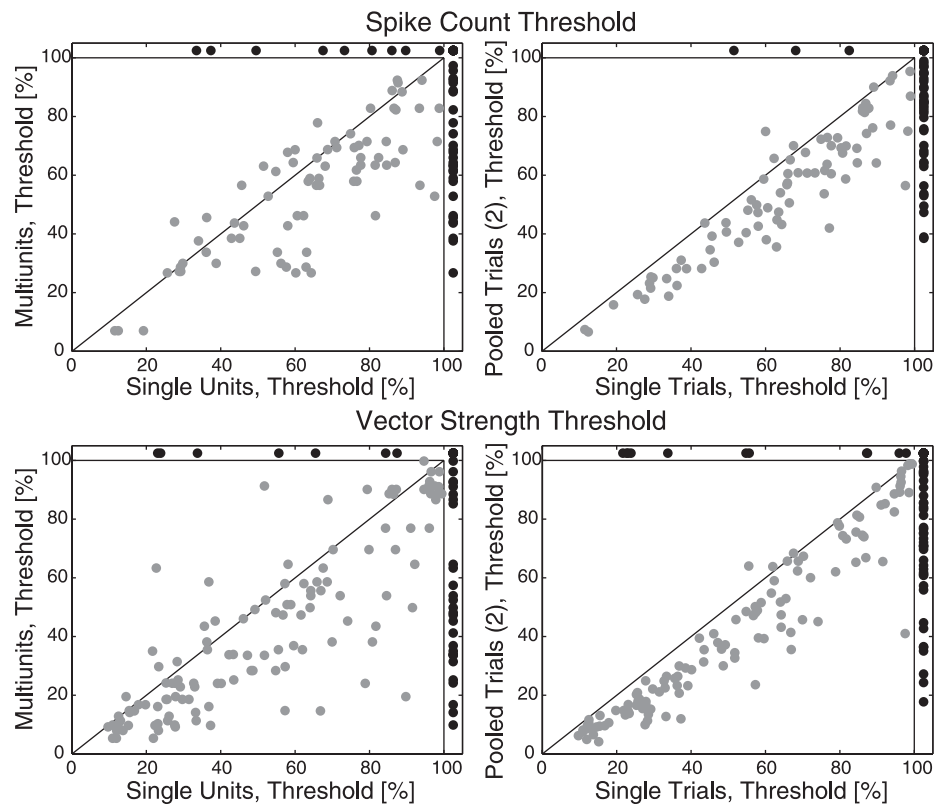
Depth sensitivity functions: limits of within-cell pooling. Because pooling resulted in an improvement of threshold on a

cell-by-cell basis, we decided to test the limits of our within-cell pooling method by increasing the number of pooled trials from the original 2 to a maximum of 25, and found improvement approaching behavioral thresholds with increased pooling. We found that as the number of trials pooled increased both the percentage of pooled responses reaching threshold and the mean threshold improved (Fig. 8A).

Trials were pooled without replacement, so as the number of trials pooled increased the number of “trials” available for ROC area calculations decreased. Data showing the number of pooled responses reaching threshold as a function of the number of trials pooled together are plotted in Fig. 8A, left. The unconnected square (SC) and circle (VS_{pp}) represent the number of responses reaching threshold without pooling. Both SC and VS_{pp} measures show drastic increases in the number of cells reaching threshold as the number of trials pooled increases to about six to eight and a plateau thereafter (although SC peaks at 12 trials pooled). The decrease after 15 trials pooled for SC probably results from the reduction of “trials” available to carry out the ROC area calculations (3 “trials” when pooling 16 trials together and only 2 when pooling 25 together). Although only 34% (SC) and 51% (VS_{pp}) of our cells reach threshold in the unpooled conditions, on average between 65% and 75% of cells reach threshold with a reasonable amount of within-cell pooling (equivalent to 6–12 cells with similar responses).

In addition to investigating the effect of within-cell pooling on the number of pooled responses reaching threshold, we also asked what effect such pooling had on mean thresholds. These data are shown in Fig. 8A, right. We limited the mean threshold analysis to only those cells that reached threshold without

Fig. 7. Threshold scatterplots. Plot conventions are similar to Fig. 5, *A* and *B*. *Left*: multiunit threshold plotted against single-unit threshold. *Right*: 2-trial within-cell pooled threshold plotted against unpooled threshold. *Top*: spike count. *Bottom*: VS_{PP} . Gray dots, cases in which both measures reached threshold; black dots, cases in which only 1 or neither measure reached threshold.



pooling (SC: $n = 84$, VS_{PP} : $n = 128$) to avoid recruitment effects that would obscure improvement in the mean threshold (cf. Fig. 6 and Fig. 7). Although we found that the majority of recruitment of previously nonsensitive cells occurs within the first 5–10 trials pooled (Fig. 8*A*, *left*), individual cells continue to see threshold improvements across the range of pooled trials tested—suggesting that recruitment is not limited by the pooling process but by exhausting cells whose sensitivity will improve with pooling from our sample. It is notable that as pooling increases VS_{PP} thresholds remain better than SC thresholds throughout. The mean VS_{PP} threshold drops below 20% for a pooling of 25 trials, demonstrating that modest pooling of responsive cells (i.e., those reaching threshold without pooling) results in neuronal thresholds roughly equivalent to psychophysical thresholds in the behaving macaque (~ 20 – 30% modulation depth across a relatively wide bandwidth; O'Connor et al. 2011). Thus pooling may provide an effective strategy of increasing behavioral sensitivity for measures based on either rate or temporal (so long as phase relationships are consistent) aspects of the neural response in A1.

Depth sensitivity functions: across-cell pooling. While within-cell pooling assumes that neurons with similar response properties converge to segregated postsynaptic neurons, it is possible that this assumption is not justified and the code for modulation might rest with a less selective model. Under the assumption that convergence mechanisms might not be highly specific with respect to response properties, we developed an across-cell pooling technique that modeled the indiscriminate convergence of between 2 and 50 cells as described in METHODS. Cells tested at modulation frequencies between 10 and 60 Hz (5 and 120 Hz were excluded because of low numbers of tested cells) were grouped by the modulation frequency of the test

and were randomly sampled for each pool size for 1,000 iterations. Mean threshold and percentage of pools reaching threshold were calculated for each set of 1,000 iterations and then averaged across tested modulation frequencies (weighted by the number of cells at each modulation frequency) to arrive at a single summary value for each pool size. An important property of this model is that responses from different neurons are pooled (i.e., the union of all spike times is collected into a single pooled spike train) before the pooled spike train undergoes neurometric analysis (as opposed to taking a measurement from each cell and then pooling the analyzed metric).

For across-cell pooling, as the number of cells pooled increased almost all pools reached threshold (Fig. 8*B*, *left*). Despite the fact that the population of cells tested were not likely to have identical or even similar temporal characteristics (particularly in the phase of response), we found that the percentage of pools reaching threshold for VS_{PP} exceeded 90% for pools of 20 cells or larger. Slightly higher percentages of pools reached threshold for SC, as long as cells with increasing functions and cells with decreasing functions were kept segregated. If increasing and decreasing cells were not kept segregated, the percentage of pools reaching threshold for SC measures dropped drastically to between 50% and 60% for pool sizes of 20 and above (this means it might be necessary to have separate postsynaptic targets for neurons with increasing and decreasing functions).

Across-cell pooling improved VS_{PP} -based thresholds, but not as much as within-cell pooling, while across-cell pooling improved SC thresholds (for increasing-function cells) more than within-cell pooling (Fig. 8*B*, *right*). When fewer than 20 cells are pooled, VS_{PP} provides more sensitivity than spike count, but temporal pooling thresholds appear to asymptote at $\sim 30\%$. Rate measures, particularly those from

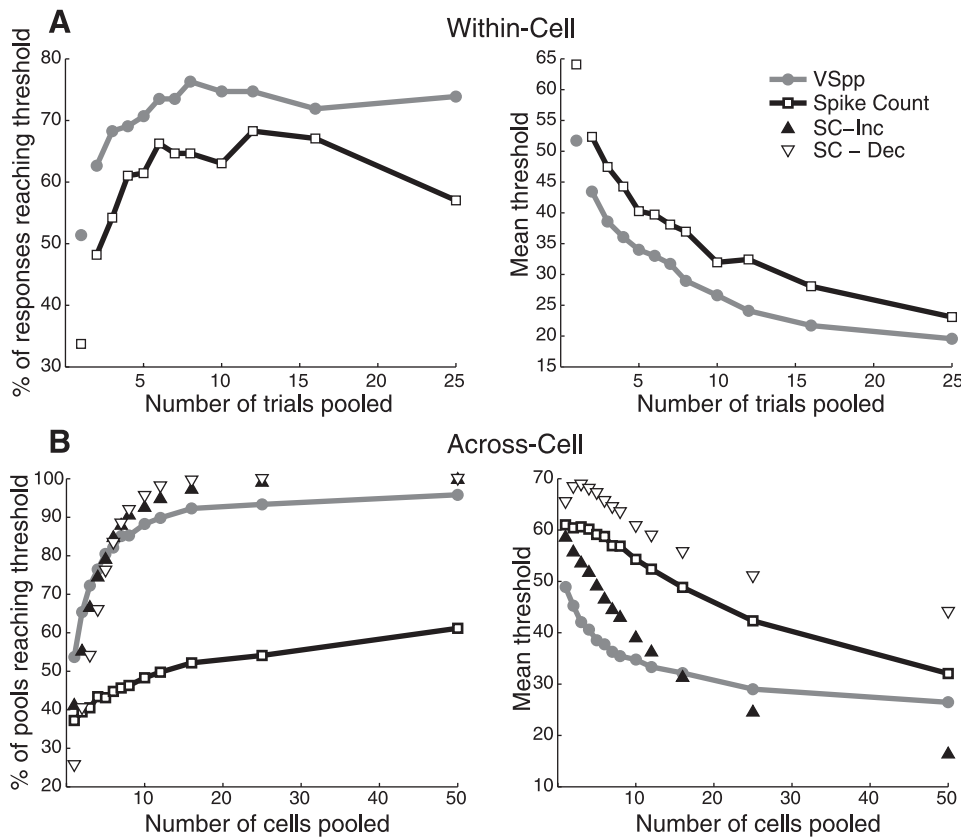


Fig. 8. Within-cell and across-cell pooling. *A*: within-cell pooling; pooling is done without replacement (see METHODS). All values are averaged across all tested modulation frequencies. *Left*: % of cells reaching threshold as a function of number of trials pooled. Unconnected symbols are unpooled values. Connected symbols are unpooled values. Connected symbols are unpooled values. *Right*: mean threshold as a function of number of trials pooled. For the mean threshold plot, only cells that reach threshold in the unpooled case (unconnected symbols) are used to calculate subsequent thresholds to avoid recruitment effects (see text). *B*: across-cell pooling; pooling is done with replacement (see METHODS). All values are averaged across all modulation frequencies used in the analysis (10–60 Hz). For VS_{PP} and spike count, pooling was done from a pool of all cells. For SC-Inc and SC-Dec, spike count pooling was done from a pool of only cells with increasing (decreasing) spike count functions. *Left*: % of pools reaching threshold as a function of number of cells pooled. *Right*: mean threshold as a function of number of cells pooled.

segregated increasing-function cells, continue to improve their sensitivity up to 50 cells per pool. The mean SC threshold for nonsegregated cells starts high with low numbers of pooled cells but approaches the sensitivity of VS_{PP} by 50 pooled cells.

Comparison to behavioral data. In Fig. 9 we highlight some of the behavior/neurometric comparisons that have been made above while partitioning the averaged data by tested modulation frequency (these data were presented as an experiment-wide average in Fig. 4). Macaque behavioral thresholds for AM noise detection using 400-ms stimuli are replotted (thin dashed line, threshold is $d' \geq 1.0$) from O'Connor et al. (2011). For VS_{PP} measures, the range of single-cell neurometric thresholds (individual cells plotted with dots) extends below the behavioral threshold. Lower-envelope models propose that the most sensitive cells in a

population underlie behavioral performance, but these data imply that if the macaques were using a lower-envelope model in combination with temporal coding, they would be more behaviorally sensitive than they are. However, pooled VS_{PP} values (for across-cell pooling of 25 cells) approximate the animal's behavioral performance at modulation frequencies of 20 Hz and above, while for 10–15 Hz pooling does not suffice to explain behavior. For SC, both lower-envelope and pooling (across cell, 25 cells) models appear to provide good approximations to the behavioral performance of the animal. Our across-cell pooling simulations suggest that the simple, indiscriminate pooling of responses recorded in A1 is sufficient to allow rate-based codes—and for higher modulation frequencies, temporal-based codes—to approximate the behavioral performance of the animal while pooling across ~25 cells.

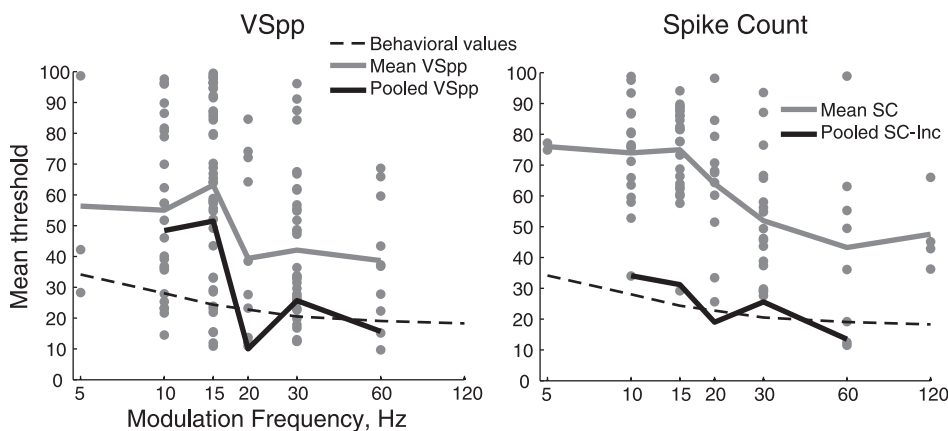


Fig. 9. Comparison of behavioral thresholds and neurometric thresholds. Behavioral values are replotted from O'Connor et al. (2011) and are the same in both panels (threshold is $d' \geq 1.0$). The thresholds measured in O'Connor et al. (2000) under a slightly different paradigm were calculated with an ROC area of 0.75 and were similar to those found in O'Connor et al. (2011). Gray points indicate thresholds of individual depth sensitivity functions. Pooling is done with the across-cell method. For VS_{PP} (left), mean of 1,000 random draws of 25 cells from a pool of all cells. For SC-Inc (right), mean of 1,000 random draws of 25 cells from a pool of all cells with increasing spike count functions.

DISCUSSION

Comparing neurons with behavior. A major focus of this study was to determine neurometric thresholds for amplitude modulation detection in macaque A1 with rate and temporal measures and to compare those thresholds to the known detection ability of macaques. Previous psychophysical studies in rhesus and other macaques have found that behavioral sensitivity to AM noise plateaus between modulation frequencies of ~10 and 120 Hz (Moody 1994; O'Connor et al. 2000, 2011). Under conditions similar to those used in this study, this plateau corresponds to a threshold of ~18–25% modulation depth. In comparison, for single units we find that the mean threshold for SC is above 60% modulation depth and for VS_{PP} thresholds are around 50%. As a result, the “average” A1 neuron is not sensitive enough to account for known behavioral performance. On the other hand, we find 20 depth sensitivity functions with estimated thresholds below 20%, 3 for SC (minimum 11.5%) and 17 for VS_{PP} (minimum 9.7%), so it is clear that while the average neuron is not sensitive enough to account for behavioral performance, the best neurons are more sensitive than the animal, particularly for VS_{PP} measures.

However, the idea that the animal would rely on an average sensory neuron—or even the best sensory neuron (often known as the lower-envelope principle)—to effect a behavioral decision makes little sense in light of the general properties of cortical architecture. The large number of synaptic inputs impinging on each cortical neuron instead suggests another strategy for sensory decision-making—that of pooling the responses of converging neurons and using an aggregate signal for performing such tasks as AM detection. We found that both multiunit activity and simple two-trial pooling within cells were able to reduce thresholds on a cell-by-cell basis for both SC and vector strength measures.

The cell-by-cell improvement we found, however, often did not result in a change in mean threshold because of a recruitment effect. We observed such an effect under two different comparisons, when comparing thresholds of pooled/multiunit responses to those of single-cell responses and also when comparing thresholds of recordings made at the BMF to recordings made away from the BMF. In both cases, the introduction of high thresholds from cells that previously did not reach threshold (hence cells not included in the previous mean threshold calculations because they were not initially classified as AM sensitive) counteracts the improvement seen for individual cells, resulting in approximately equal mean thresholds. Such an effect is unlikely to be limited to the domain of AM sensitivity and may be an important factor to account for in any studies that do not isolate single units.

For within-cell pooling, we avoided a recruitment effect by restricting our pooling to cells that reached threshold without pooling and found improvements in sensitivity up to the maximum number of trials (25) we could pool. Cells generally were more sensitive with temporal measures than with SC measures, with temporal thresholds dropping to depths of ~20% and SC thresholds dropping to ~25% modulation depth, values that lie in the range of macaque behavioral thresholds. However, these calculations are hampered by the fact that the greater the number of trials we pooled, the fewer “pooled trials” were available for our neurometric analysis, which would increase the uncertainty of the threshold estimate.

Our across-cell pooling method, which we believe would be the model of convergence more easily implemented biologically, does not suffer from a recruitment effect because each simulated pooling results from a random draw taken across the population of recordings, including both sensitive and nonsensitive (not reaching threshold) cells—as the number of cells pooled increases, there is no paired correspondence with previous pools as there is for within-cell and multiunit pooling. With across-cell pooling we found that our mean temporal thresholds drop below modulation depths of 30% for 25 and 50 cells pooled, which suggests that, despite any variability in response phase that might be found for phase-locking neurons in A1, indiscriminate pooling can still result in a phase-locked signal nearly sensitive enough to explain the animal's performance (even at 60 Hz). The situation is slightly more complicated for SC measures because a good number of our cells exhibit decreasing, rather than increasing, SC as a function of modulation depth. If these two types of responses are allowed to be pooled together, the resulting thresholds remain higher than temporally based thresholds up through at least 50 pooled cells. However, if increasing functions are segregated from decreasing functions, SC measures on pools of increasing SC cells begin to outperform temporal measures by the time 20 cells are pooled and eventually reach thresholds below modulation depths of 20%, which is in line with the animal's ability to detect AM.

When we look across modulation frequencies, the lower-envelope principle and pooling seem plausible candidates for decoding AM information with SC measures. Pooling of ~25 cells (with increasing functions only) seems sufficient to account for behavior. For VS_{PP}, the lower-envelope principle does not appear to be a good hypothesis, for many individual cells have thresholds below the behavioral threshold. Pooling (again 25 cells) seems sufficient to account for the behavioral results at higher modulation frequencies (20 Hz and above) but not at low modulation frequencies. That pooled spike trains analyzed with VS_{PP} may be sufficient to account for behavioral performance for higher modulation frequencies and not lower ones is unusual because most of the literature has suggested that temporal codes are more effective at low modulation frequencies and phase locking to higher modulation frequencies diminishes at higher levels of the auditory system (Creutzfeldt et al. 1980; Joris et al. 2004; Krishna and Semple 2000; Rhode and Greenberg 1994; Schreiner and Urbas 1988). One possible explanation for this result might lie in a recent model that proposes that phase-locked AM responses are driven by the amplitude of the stimulus envelope rising (or falling) through the cell's preferred intensity (Malone et al. 2007; see also Heil 2003; Heil and Irvine 1998; Krebs et al. 2008; Zheng and Escabi 2008; Zhou and Wang 2010 for neuronal sensitivity to envelopes). The protracted envelope changes that occur for low modulation frequencies may result in more temporally smeared responses than those arising from faster-changing envelopes, and VS_{PP}, which measures how tightly spikes are clustered to a particular phase of the stimulus cycle, would naturally be reduced for more spread-out responses.

Pooling of auditory responses. In our method of across-cell pooling, the actual spike trains from individual trials were pooled together in a literal fashion. If one imagines that all of the pooled cells were to synapse (with identical weights) onto

the same dendritic compartment of a target neuron, the pooled spike train that we create would represent the input to that dendritic compartment. From this point of view, our pooling method is an empirical first-pass attempt to estimate the sensitivity that could be obtained in a downstream neuron by convergent inputs. While this approach has rarely been taken, in a study of the encoding of marmoset twitter calls in ferret A1, Walker et al. (2008) used an aggregate signal similar to our across-cell pooling. Although the natural twitter calls do not have precise regular periodic AM, the envelope modulates roughly periodically at $\sim 7\text{--}9$ Hz (Wang et al. 1995; Wang and Kadia 2001). Walker et al. investigated how well the neurons could distinguish natural twitters from manipulated twitters and found that at the single-unit level temporal coding more accurately reflected human psychophysical thresholds than rate coding. They used a principal component coding algorithm that did not necessarily measure phase locking in the same way that vector strength does, but the examples and the aggregate activity shown suggest that phase locking contributes to the temporal coding. As was the case with our study when the across-cell pooling method was used, temporal-based neurometric thresholds improved (rate codes were not tested with this kind of pooling).

The mechanics of our pooling method differ somewhat from most previous pooling studies, which have tended to use simulated responses based on draws from an approximation of the observed distribution rather than pooling actual responses (e.g., Rosen et al. 2010; Shadlen et al. 1996; but see Schneider and Woolley 2010). For SC methods (e.g., Shadlen et al. 1996) there should be little difference between using simulated responses and actual responses, but our method allowed us to perform temporal analysis on pooled raw data. In comparison, Rosen et al. (2010), who investigated AM tone responses in gerbils, obtained their temporal neurometric measures on mean vector strength values drawn from an estimated distribution. This difference is crucial because when precalculated vector strength values are pooled any phase difference between pooled responses has already been removed. In our pooling method, spike phase information is preserved up to the point of the final phase-locking calculation, just as it would be in an in vivo pooling, and differences in response phase will be reflected in the resulting VS_{PP} values. We find that for the across-cell condition pooling resulted in improvement of thresholds—even at 60 Hz—a result we would not expect to see were the phase relationships between different cells random. At the same time, we note that our pooled VS_{PP} thresholds are not as sensitive in the across-cell case as the within-cell case, which suggests that inconsistent phase relationships do contribute to the difference. Still, even indiscriminate pooling results in a notable increase in temporal information, which is consistent with the finding that phase-locked AM responses in A1 generally synchronize to the same phase in the stimulus cycle (Bendor and Wang 2008; Yin et al. 2011).

One important issue to keep in mind when pooling is the possible response correlations between neurons. In the case of SC, across-cell correlation has been noted to reduce the effectiveness of pooling (Alves-Pinto et al. 2010; Zohary et al. 1994; for a counterexample, see Romo et al. 2003). Since our pooling methods (except in the case of the multiunit considered as a pool) combine trials that were collected at different times, any correlation in SC that might have been present in an in vivo

pooling operation will be missing in our simulation, and our SC threshold estimates may be too low (i.e., too sensitive) to the extent that we have failed to capture existing correlations. However, interinput correlation only reduces the effectiveness of pooling when the inputs have identical tuning curves (Abbott and Dayan 1999), a condition that is only true for our within-cell pooling method—our across-cell method encompasses cells with diverse tuning properties and should be relatively robust to SC correlations.

For measures of phase locking, such as VS_{PP} , the influence of across-cell correlation is more complicated. Whereas for SC it is clear that high intercell correlation decreases the benefits of pooling, for temporal-based measures it is not clear that higher intercell correlation will always result in decreasing the benefits of pooling, nor is it clear that lower intercell correlation will benefit pooling (Elhilali et al. 2009; Walker et al. 2008). The ideal method to determine the effect of potential correlation issues for pooling with temporal-based measures would be multiunit array recording (e.g., Bizley et al. 2010 in ferrets). Unfortunately, such arrays have yet to be employed successfully in the macaque model, and direct correlation analysis has to date proven elusive.

Implications of pooling for rate and temporal coding. The results of this study have interesting implications for the usefulness of rate versus temporal coding of temporal envelope modulation. Previous studies with AM (e.g., Malone et al. 2010; Nelson and Carney 2007) and communication vocalizations with strong AM (e.g., Huetz et al. 2009; Schnupp et al. 2006; Walker et al. 2008) indicate that temporal codes (including VS) represent the sounds better than rate codes. This could mean that temporal codes are preferentially used to encode AM, but our results suggest that a preference for rate or temporal codes should depend on how the brain integrates information across the neuronal population. For example, if the brain were able to select only the most sensitive A1 units, a rate code would fairly accurately predict performance but a phase locking code would not, suggesting performance better than that obtained psychophysically. The inability to perform better than the best phase-locking units from A1 suggests that the brain is unable to simply isolate those units.

An alternate population decoding strategy is neuronal pooling. Indiscriminate pooling (pooling without respect to coding efficiency, best modulation frequency, etc.) would be the easiest form of pooling to implement. For truly indiscriminate pooling, phase locking measures outperform rate-based measures through pools of at least 50 neurons. However, our pooling simulations find that thresholds based on SC are lower and more accurately reflect behavior across modulation frequency than thresholds based on phase locking measures—but require one restriction, that only cells with increasing depth sensitivity functions are included in the pool, because at the pooling level excited and suppressed responses counteract (e.g., Oshurkova et al. 2008). This sort of segregation seems to represent an intermediate level of specificity in convergence and might be biologically plausible in several ways. Phase locking strength can vary with laminar depth (Wallace et al. 2011), so it is not unlikely that other AM response properties might as well. Because we do not have information about the laminar location of our recorded neurons it is possible that cells with increasing and decreasing depth functions differ in their laminar location and have different synaptic targets. Hebbian

mechanisms or other physiological factors (decreasing cells could be different cell types, for instance, local interneurons) could also allow these two classes of cells to be segregated onto separate targets.

For within-cell pooling, which is analogous to pooling from neurons with identical properties, phase locking measures appear to give even lower detection thresholds, which would allow smaller pool sizes to account for behavior. This result is consistent with the observation that phase locking is better for multiunits than single units (Oshurkova et al. 2008), and if such a strategy were to be implemented the selection of cells with similar properties might well be based on columnar tuning. It would be interesting to see whether pooling can improve detection thresholds found with other temporally based codes, such as those based on overall spike timing (Furukawa and Middlebrooks 2002; Kajikawa et al. 2008; Malone et al. 2007; Wang et al. 2007) and interspike interval (ISI) distributions (Imaizumi et al. 2010), and to what extent pooling results using alternate codes depend on whether neurons with similar properties are sampled.

Comparison to previous studies. The neurometric thresholds that we found for single cells were somewhat higher (less sensitive) than those found in previous studies of AM depth sensitivity in midbrain and in cortex, where AM tones rather than AM noise were used. AM noise has often been used in psychophysical studies because spectral and temporal cues in AM tones can be confounded, a problem that is alleviated with a noise carrier. Therefore, in addition to determining modulation encoding for tone carriers, it is important to determine the relationship of neuronal to behavioral sensitivity for noise carriers.

In the IC, Nelson and Carney (2007) recorded from awake rabbits while presenting 2-s AM tones (further broken down into nine 500-ms segments, discarding onset) with the carrier at the cell's best frequency. When using a signal detection method as a neurometric, they found temporal-based thresholds to be lower than rate-based thresholds—the median of single-cell rate thresholds was ~30% modulation depth, and synchrony thresholds were somewhat lower with a median below ~20% modulation depth. Krishna and Semple (2000) recorded from IC of anesthetized gerbils; they also report that the median cutoff depth for significant phase locking (using the Rayleigh statistic) is ~20% and do not report rate-based thresholds. The better ability to phase lock to lower depths compared with our findings (particularly for temporal measures) is likely to be a consequence of recording in IC, where phase locking is stronger than in cortex, but may reflect species differences in AM sensitivity as well.

In auditory cortex AM sensitivity has been found to be worse. Eggermont (1994) found that most units did not phase lock to 25% or lower modulation depth with AM noise. This is consistent with our results. Malone et al. (2010) recorded from A1, R, and three belt areas in macaque while presenting very long (10 s) AM tones and focusing on lower modulation frequencies, making only 11 recordings with modulation frequencies above 20 Hz. For spike count they found that 41% of their cells could detect modulation (using a comparison to the unmodulated tone) with a median threshold of 50% modulation depth. While these spike count data were similar to ours, for VS they found a remarkably high percentage of single cells that were capable of detecting modulation (99%, compared with

51% for the data we report here) with a median threshold of 20% modulation depth. Several factors may contribute to the difference between their result and ours. Their long stimulus duration provides 25 times more cycles per stimulus than ours (10 s compared with 400 ms in this study) and so approximates our within-cell pooling condition. When we pooled 25 trials within cell, our mean VS_{pp} thresholds were also ~20% modulation depth and we found ~75% of pools to reach threshold. In addition, their VS detection thresholds were based on a significant Rayleigh statistic, testing for nonuniformity in the cycle histogram instead of a comparison to an unmodulated tone, and did not exclude onset responses. We have previously shown that the Rayleigh statistic is prone to producing false alarms at low modulation frequencies (Yin et al. 2011). The presence of onset responses may also result in spurious nonuniformity in the cycle histogram, particularly for low modulation frequencies (long cycles) and cells with low sustained responses.

Electrical stimulation studies using cochlear implants in guinea pigs (Middlebrooks 2008b) find more sensitive rate and temporal multiunit responses than our study (the median for both is slightly greater than 10%). This cannot be accounted for by their using essentially four-unit within-cell pooling (see Fig. 8A). One possibility is that electrical stimulation is more effective at entraining neural responses. The electrical stimulus consists of pulse trains with a constant interpulse interval, and the amplitude of each pulse is adjusted to create the AM envelope. Because A1 neurons are most sensitive for carriers with lower interpulse intervals (Middlebrooks 2008b), the pulsing carrier might aid in phase locking to the envelope. Middlebrooks (2008a) also shows that many units have nonmonotonic depth functions and there is a tendency to see more nonmonotonic functions at higher modulation frequencies, while we find no strongly nonmonotonic depth functions for VS_{pp} and only 3 of 55 for increasing SC, consistent with Liang et al. (2002). This suggests that the high proportion of nonmonotonic depth functions found by Middlebrooks may be a factor of electrical stimulation and that nonmonotonic depth functions for AM are not common for acoustical stimulation.

Properties of pooling and temporal codes in other sensory systems. The kinds of analyses performed here have also been used in other sensory systems, so some general conclusions about the function of sensory cortex can be discussed. For one, neural pooling appears to be a general strategy used in decoding a wide variety of sensory information in cortex (e.g., vision: Shadlen et al. 1996; somatosensory: Panzeri et al. 2003; olfactory: Kazama and Wilson 2009) as well as retina (Pahlberg and Sampath 2011) and invertebrate nervous systems (Warrant 2008), particularly in the context of improving the signal-to-noise ratio beyond that found in single neurons. Pooling of neural data has been successful in accounting for behavioral performance at various levels of the mammalian nervous system (Gold et al. 2010; Swanson et al. 2008) including primary auditory cortex (cat: Qin et al. 2009; gerbil: Sarro et al. 2011).

Three different pooling strategies have been commonly used: pooling neuronal responses; pooling or averaging a property that is calculated for each neuron (such as using average threshold); and relying on the most sensitive neurons.

When neuronal responses are pooled as the first step of a pooling analysis, the number of neurons required to account for

behavior differs across studies, depending on several experimental details. The duration of response that is used is particularly important because it can account for a lot of observed differences (Cohen and Newsome 2009; Cook and Maunsell 2002)—in general, increasing the duration of responses analyzed is similar to increasing the number of neurons in the pool. Whether pooling is done with spatially proximate neurons or on a global basis (Panzeri et al. 2003; Reich et al. 2001) can also affect the efficacy of pooling. In this study we focused on whether neurons are selected for pooling on the basis of similar response properties for optimal stimuli or indiscriminately. Typically, restricting the pool to neurons that are tuned to the parameter of interest allows a smaller number of neurons to account for behavior, but asymptotic performance at increased numbers of neurons leads to thresholds that outperform behavior (Cohen and Newsome 2009; Palmer et al. 2007). Some studies that are less restrictive in choosing which neuronal responses to use require on the order of 25–100 neurons, but asymptotic thresholds approach behavioral thresholds (Cohen and Newsome 2009), similar to our observations.

Another pooling method is to measure each neuron's threshold separately and then compute the average of these thresholds (Celebrini and Newsome 1994; Hernandez et al. 2000; Heuer and Britten 2004; Uka and DeAngelis 2003). Average single neural thresholds based on firing rate vary widely. Some studies find average thresholds better than or similar to behavior (Adibi and Arabzadeh 2011; Britten et al. 1992; Uka and DeAngelis 2003), but most studies (including our own, Fig. 4, A and B) find average thresholds to be worse (Cohen and Newsome 2009; Cook and Maunsell 2002; Liu et al. 2010; Matsumora et al. 2008; Prince et al. 2000).

Relying on only the most sensitive neurons (the “lower-envelope” model) has also been claimed as a means to support behavior (Liu and Newsome 2005; Osborne et al. 2004; Parker and Newsome 1998; Prince et al. 2000; Purushothaman and Bradley 2005; Vogels and Orban 1990). Our data support the possibility of such a coding scheme with firing rate, but for temporal coding we find the most sensitive neurons are superior to behavioral thresholds.

Given the dense interconnections between cortical neurons, the idea that pooling strategies might be widely utilized in decoding and responding to sensory stimuli seems natural, if not practically necessary. The results here provide further evidence that sensory-based behavioral thresholds can be predicted from the pooling of responses in sensory cortex, and further bolster the idea that pooling mechanisms should be a useful tool when trying to understand sensory-based decisions in general.

Like pooling strategies, temporal codes can be found in sensory cortex across several modalities. In addition to phase locking codes, general temporal coding (e.g., pattern recognition: Hopfield 1995; information theory: Kajikawa and Hackett 2005), which asks whether any pattern could be used as a code rather than requiring that the temporal pattern of neural activity track the temporal pattern of the stimulus, can also be used to predict the stimulus based on the firing pattern of the neuron. In the visual system, both strategies have been used—for example, a neurometric analysis shows that phase locking can detect coherent motion (Masse and Cook 2008), and information theoretical analysis on spike trains suggests that as few as 5–10 neurons can account for behavior (Ghose and Harrison

2009). In the auditory and somatosensory systems phase locking to the stimulus is most commonly studied. In the somatosensory system, phase locking has been shown to encode stimulus flutter frequency well at levels up to primary somatosensory cortex (e.g., Mountcastle et al. 1969, 1990; Recanzone et al. 1992) but less well in higher somatosensory areas (Salinas et al. 2000). These phase-locked codes tend to be more sensitive than firing rate codes. In primary somatosensory cortex, individual neuronal thresholds in discriminating two flutter frequencies were more sensitive (in fact better than the animal's performance) with phase locking measures, whereas rate-based coding for individual neurons was less sensitive but closer to behavioral thresholds (Hernandez et al. 2000). This result is conceptually similar to our findings that a higher percentage of neurons reach threshold with VS_{PP} than rate (Fig. 3) and that single-unit thresholds are more sensitive for VS_{PP} than rate (Fig. 4). One difference is that Hernandez et al. (2000) found very few neurons that represented stimulus frequency with both rate and temporal codes, while in our detection task many do. Altogether, these results suggest that the visual, somatosensory, and auditory systems might share common properties that allow phase locking or other temporal codes to encode stimulus information.

ACKNOWLEDGMENTS

Present address of P. Yin: Neural Systems Laboratory, Institute for Systems Research, University of Maryland, College Park, MD 20742.

GRANTS

This work was funded by National Institute on Deafness and Other Communication Disorders Grants DC-02514 and T32 DC-008072.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: J.S.J. and P.Y. analyzed data; J.S.J., K.N.O., and M.L.S. interpreted results of experiments; J.S.J. prepared figures; J.S.J. and M.L.S. drafted manuscript; J.S.J., K.N.O., and M.L.S. edited and revised manuscript; J.S.J., P.Y., K.N.O., and M.L.S. approved final version of manuscript; P.Y., K.N.O., and M.L.S. conception and design of research; P.Y. and K.N.O. performed experiments.

REFERENCES

- Abbott LF, Dayan P.** The effect of correlated variability on the accuracy of a population code. *Neural Comput* 11: 91–101, 1999.
- Adibi M, Arabzadeh E.** A comparison of neuronal and behavioral detection and discrimination performances in rat whisker system. *J Neurophysiol* 105: 356–365, 2011.
- Alves-Pinto A, Baudoux S, Palmer AR, Sumner CJ.** Forward masking estimated by signal detection theory analysis of neuronal responses in primary auditory cortex. *J Assoc Res Otolaryngol* 11: 477–494, 2010.
- Attias H, Schreiner CE.** Coding of naturalistic stimuli by auditory midbrain neurons. *Adv Neur Info Proc Syst* 10: 103–109, 1998.
- Auerbach SH, Allard T, Naeser M, Alexander MP, Albert ML.** Pure word deafness. Analysis of a case with bilateral lesions and a defect at the prephonemic level. *Brain* 105: 271–300, 1982.
- Bendor D, Wang X.** Neural response properties of primary, rostral, and rostrotemporal core fields in the auditory cortex of marmoset monkeys. *J Neurophysiol* 100: 888–906, 2008.
- Bieser A, Müller-Preuss P.** Auditory responsive cortex in the squirrel monkey: neural responses to amplitude-modulated sounds. *Exp Brain Res* 108: 273–284, 1996.

- Bizley JK, Walker KMM, King AJ, Schnupp JWH.** Neural ensemble codes for stimulus periodicity in auditory cortex. *J Neurosci* 30: 5078–5091, 2010.
- Bregman AS.** *Auditory Scene Analysis*. Cambridge, MA: MIT Press, 1990.
- Britten KH, Shadlen MN, Newsome WT, Movshon JA.** The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci* 12: 4745–4765, 1992.
- Burns EM, Viemeister NF.** Nonspectral pitch. *J Acoust Soc Am* 58: 863–869, 1976.
- Burns EM, Viemeister NF.** Played-again SAM: further observations on the pitch of amplitude-modulated noise. *J Acoust Soc Am* 70: 1655–1660, 1981.
- Carmon A, Nachshon I.** Effect of unilateral brain damage on perception of temporal order. *Cortex* 7: 411–418, 1971.
- Celebrini S, Newsome WT.** Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. *J Neurosci* 14: 4109–4124, 1994.
- Chandrasekaran C, Tureson HK, Brown CH, Ghazanfar AA.** The influence of natural scene dynamics on auditory cortical activity. *J Neurosci* 30: 13919–13931, 2010.
- Cohen MR, Newsome WT.** Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. *J Neurosci* 29: 6635–6648, 2009.
- Cook EP, Maunsell JH.** Dynamics of neuronal responses in macaque MT and VIP during motion detection. *Nat Neurosci* 5: 985–994, 2002.
- Creutzfeldt O, Hellweg FC, Schreiner C.** Thalamic transformation of responses to complex auditory stimuli. *Exp Brain Res* 39: 87–104, 1980.
- Delgutte B, Hammond BM, Cariani PA.** Neural coding of the temporal envelope of speech: relation to modulation transfer functions. In: *Psychophysical and Physiological Advances in Hearing*, edited by Palmer AR, Rees A, Summerfield AQ, Meddis R. London: Whurr, 1998, p. 595–603.
- DiMattina C, Wang X.** Virtual vocalization stimuli for investigating neural representations of species-specific vocalizations. *J Neurophysiol* 95: 1244–1262, 2006.
- Drullman R, Festen JM, Plomp R.** Effect of temporal envelope smearing on speech reception. *J Acoust Soc Am* 95: 1053–1064, 1994.
- Efron R, Yund EW, Nichols D, Crandall PH.** An ear asymmetry for gap detection following anterior temporal lobectomy. *Neuropsychologia* 23: 43–50, 1985.
- Eggermont J.** Temporal modulation transfer functions for AM and FM stimuli in cat auditory cortex. Effects of carrier type, modulating waveform and intensity. *Hear Res* 74: 51–66, 1994.
- Elhilali M, Ma L, Micheyl C, Oxenham AJ, Shamma SA.** Temporal coherence in the perceptual organization and cortical representation of auditory scenes. *Neuron* 29: 317–329, 2009.
- Fishman YI, Steinschneider M.** Neural correlates of auditory scene analysis based on inharmonicity in monkey primary auditory cortex. *J Neurosci* 30: 12480–12494, 2010.
- Fishman YI, Volkov IO, Noh MD, Garell PC, Bakken H, Arezzo JC, Howard MA, Steinschneider M.** Consonance and dissonance of musical chords: neural correlates in auditory cortex of monkeys and humans. *J Neurophysiol* 86: 2761–2788, 2001.
- Fitch RH, Tallal P, Brown CP, Galaburda AM, Rosen GD.** Induced microgyria and auditory temporal processing in rats: a model for language impairment? *Cereb Cortex* 4: 260–270, 1994.
- Furukawa S, Middlebrooks JC.** Cortical representation of auditory space: information-bearing features of spike patterns. *J Neurophysiol* 87: 1749–1762, 2002.
- Ghose GM, Harrison IT.** Temporal precision of neuronal information in a rapid perceptual judgment. *J Neurophysiol* 101: 1480–1493, 2009.
- Gold JL, Law C, Connolly P, Benuer S.** Relationships between the threshold and slope of psychometric and neurometric functions during perceptual learning: implications for neuronal pooling. *J Neurophysiol* 103: 140–154, 2010.
- Goldwyn JH, Shea-Brown E, Rubinstein JT.** Encoding and decoding amplitude-modulated cochlear implant stimuli—a point process analysis. *J Comput Neurosci* 28: 405–424, 2010.
- Green DM, Swets JA.** *Signal Detection Theory and Psychophysics*. New York: Wiley, 1966.
- Grimault N, Bacon SP, Micheyl C.** Auditory stream segregation on the basis of amplitude-modulation rate. *J Acoust Soc Am* 111: 1340–1348, 2002.
- Hand DJ, Till RJ.** A simple generalisation of the area under the ROC curve for multiple class classification problems. *Machine Learn* 45: 171–186, 2001.
- Hanley JA, McNeil BJ.** The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143: 29–36, 1982.
- Heil P.** Coding of temporal onset envelope in the auditory system. *Speech Communication* 41: 123–134, 2003.
- Heil P, Irvine DRF.** The posterior field P of cat auditory cortex: coding of envelope transients. *Cereb Cortex* 8: 125–141, 1998.
- Heilman KM, Hammer LC, Wilder BJ.** An audiometric defect in temporal lobe dysfunction. *Neurology* 23: 384–386, 1973.
- Helmholtz H.** *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. New York: Dover, 1954.
- Hernandez A, Zainos A, Romo R.** Neuronal correlates of sensory discrimination in the somatosensory cortex. *Proc Natl Acad Sci USA* 97: 6191–6196, 2000.
- Hopfield JJ.** Pattern recognition computation using action potential timing for stimulus representation. *Nature* 376: 33–36, 1995.
- Heuer HW, Britten KH.** Optic flow signals in extrastriate area MST: comparison of perceptual and neuronal sensitivity. *J Neurophysiol* 91: 1314–1326, 2004.
- Huetz C, Philibert B, Edeline JM.** Spike-timing code for discriminating conspecific vocalizations in the thalamocortical system of anesthetized and awake guinea pigs. *J Neurosci* 29: 334–350, 2009.
- Imaizumi K, Priebe NJ, Sharpee TO, Cheung SW, Schreiner CE.** Encoding of temporal information by timing, rate, and place in cat auditory cortex. *PLoS One* 5: e11531, 2010.
- Ison JR, Bowen GP.** Scopolamine reduces sensitivity to auditory gaps in the rat, suggesting a cholinergic contribution to temporal acuity. *Hear Res* 145: 169–176, 2000.
- Ison JR, O'Connor K, Bowen GP, Bocirnea A.** Temporal resolution of gaps in noise by the rat is lost with functional decortication. *Behav Neurosci* 105: 33–40, 1991.
- Joris PX, Schreiner CE, Rees A.** Neural processing of amplitude-modulated sounds. *Physiol Rev* 84: 541–577, 2004.
- Kajikawa Y, Hackett TA.** Entropy analysis of neuronal spike train synchrony. *J Neurosci Methods* 149: 90–93, 2005.
- Kajikawa Y, de la Mothe LA, Blumell S, Sterbing-D'Angelo SJ, D'Angelo W, Camalier CR, Hackett TA.** Coding of FM sweep trains and twitter calls in area CM of marmoset auditory cortex. *Hear Res* 239: 107–125, 2008.
- Kazama H, Wilson RI.** Origins of correlated activity in an olfactory circuit. *Nat Neurosci* 12: 1136–1144, 2009.
- Kelly JB, Rooney BJ, Phillips DP.** Effects of bilateral auditory cortical lesions on gap-detection thresholds in the ferret (*Mustela putorius*). *Behav Neurosci* 110: 542–550, 1996.
- Krebs B, Lesica NA, Grothe B.** The representation of amplitude modulations in the mammalian auditory midbrain. *J Neurophysiol* 100: 1602–1609, 2008.
- Krishna BS, Semple MN.** Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. *J Neurophysiol* 84: 255–273, 2000.
- Lackner JR, Teuber HL.** Alterations in auditory fusion thresholds after cerebral injury in man. *Neuropsychologia* 11: 409–415, 1973.
- Liang L, Lu T, Wang X.** Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates. *J Neurophysiol* 87: 2237–2261, 2002.
- Liu J, Newsome WT.** Correlation between speed perception and neural activity in the middle temporal visual area. *J Neurosci* 25: 711–722, 2005.
- Liu S, Yakusheva T, DeAngelis GC, Angelaki DE.** Direction discrimination thresholds of vestibular and cerebellar nuclei neurons. *J Neurosci* 30: 439–448, 2010.
- Lorenzi C, Micheyl C, Berthommier B.** Neuronal correlates of perceptual amplitude-modulation detection. *Hear Res* 90: 219–227, 1995.
- Lu T, Liang L, Wang X.** Temporal and rate representations on time-varying signals in the auditory cortex of awake primates. *Nat Neurosci* 4: 1131–1138, 2001.
- Lu T, Wang X.** Information content of auditory cortical responses to time-varying acoustic stimuli. *J Neurophysiol* 91: 301–313, 2004.
- Malone BJ, Scott BH, Semple MN.** Temporal codes for amplitude contrast in auditory cortex. *J Neurosci* 30: 767–784, 2010.
- Malone BJ, Scott BH, Semple MN.** Dynamic amplitude coding in the auditory cortex of awake rhesus macaques. *J Neurophysiol* 98: 1451–1474, 2007.
- Masse NY, Cook EP.** The effect of middle temporal spike phase on sensory encoding and correlates with behavior during a motion-detection task. *J Neurosci* 28: 1343–1355, 2008.
- Matsumora T, Koida K, Komatsu H.** Relationship between color discrimination and neural responses in the inferior temporal cortex of the monkey. *J Neurophysiol* 100: 3361–3374, 2008.

- Middlebrooks JC.** Auditory cortex phase locking to amplitude-modulated cochlear implant pulse trains. *J Neurophysiol* 100: 76–91, 2008a.
- Middlebrooks JC.** Cochlear-implant high pulse rate and narrow electrode configuration impair transmission of temporal information to the auditory cortex. *J Neurophysiol* 100: 92–107, 2008b.
- Moody DB.** Detection and discrimination of amplitude-modulated signals by macaque monkeys. *J Acoust Soc Am* 95: 3499–3510, 1994.
- Mountcastle VB, Steinmetz MA, Romo R.** Frequency discrimination in the sense of flutter: psychophysical measurements correlated with postcentral events in behaving monkeys. *J Neurosci* 10: 3032–3044, 1990.
- Mountcastle VB, Talbot WH, Sakata H, Hyvärinen J.** Cortical neuronal mechanisms in flutter-vibration studied in unanesthetized monkeys: neuronal periodicity and frequency discrimination. *J Neurophysiol* 32: 452–484, 1969.
- Müller-Preuss P, Flachskamm C, Bieser A.** Neural encoding of amplitude modulation within the auditory midbrain of squirrel monkeys. *Hear Res* 80: 197–208, 1994.
- Nagarajan SS, Cheung SW, Bedenbaugh P, Beitel RE, Schreiner CE, Merzenich MM.** Representation of spectral and temporal envelope of twitter vocalizations in common marmoset primary auditory cortex. *J Neurophysiol* 87: 1723–1737, 2002.
- Narayan R, Graña G, Sen K.** Distinct time scales in cortical discrimination of natural sounds in songbirds. *J Neurophysiol* 96: 252–258, 2006.
- Nelken I, Rotman Y, Bar Yosef O.** Responses of auditory-cortex neurons to structural features of natural sounds. *Nature* 397: 154–157, 1999.
- Nelson PC, Carney LH.** Neural rate and timing cues for detection and discrimination of amplitude-modulated tones in the awake rabbit inferior colliculus. *J Neurophysiol* 97: 522–539, 2007.
- O'Connor KN, Barruel P, Sutter ML.** Global processing of spectrally complex sounds in macaques (*Macaca mulatta*) and humans. *J Comp Physiol A* 186: 903–912, 2000.
- O'Connor KN, Johnson JS, Niwa MN, Noriega NC, Marshall EA, Sutter ML.** Amplitude modulation detection as a function of modulation frequency and stimulus duration: comparisons between macaques and humans. *Hear Res* 277: 37–43, 2011.
- O'Connor KN, Petkov CI, Sutter ML.** Adaptive stimulus optimization for auditory cortical neurons. *J Neurophysiol* 94: 4051–4067, 2005.
- Olsen WO, Noffsinger D, Kurdziel S.** Speech discrimination in quiet and in white noise by patients with peripheral and central lesions. *Acta Otolaryngol (Stockh)* 80: 375–382, 1975.
- Osborne LC, Bialek W, Lisberger SG.** Time course of information about motion detection in visual area MT of macaque monkeys. *J Neurosci* 24: 3210–3222, 2004.
- Oshurkova E, Scheich H, Brosch M.** Click train encoding in primary and non-primary auditory cortex of anesthetized macaque monkeys. *Neuroscience* 153: 1289–1299, 2008.
- Pahlberg J, Sampath AP.** Visual threshold is set by linear and nonlinear mechanisms in the retina that mitigate noise: how neural circuits in the retina improve the signal-to-noise ratio of the single-photon response. *Bioessays* 33: 438–447, 2011.
- Palmer C, Cheng SY, Seidemann E.** Linking neuronal and behavioral performance in a reaction-time visual detection task. *J Neurosci* 27: 8122–8137, 2007.
- Panzeri S, Petroni F, Petersen RS, Diamond ME.** Decoding neuronal population activity in rat somatosensory cortex: role of columnar organization. *Cereb Cortex* 13: 45–52, 2003.
- Parker AJ, Newsome WT.** Sense and the single neuron: probing the physiology of perception. *Annu Rev Neurosci* 21: 227–277, 1998.
- Phillips DP, Farmer ME.** Acquired word deafness, and the temporal grain of sound representation in the primary auditory cortex. *Behav Brain Res* 40: 85–94, 1990.
- Prince SJ, Pointon AD, Cumming BG, Parker AJ.** The precision of single neuron responses in cortical area V1 during stereoscopic depth judgments. *J Neurosci* 20: 3387–3400, 2000.
- Purushothaman G, Bradley DC.** Neural population code for fine perceptual decisions in area MT. *Nat Neurosci* 8: 99–106, 2005.
- Qin L, Liu Y, Wang J, Li S, Sato Y.** Neural and behavioral discrimination of sound duration by cats. *J Neurosci* 29: 15650–15659, 2009.
- Recanzone GH, Merzenich MM, Schreiner CE.** Changes in the distributed temporal response properties of SI cortical neurons reflect improvements in performance on a temporally based tactile discrimination task. *J Neurophysiol* 67: 1071–1091, 1992.
- Reich DS, Mechler F, Victor JD.** Independent and redundant information in nearby cortical neurons. *Science* 294: 2566–2568, 2001.
- Rhode WS, Greenberg S.** Encoding of amplitude modulation in the cochlear nucleus of the cat. *J Neurophysiol* 71: 1797–1825, 1994.
- Romo R, Hernández A, Zainos A, Salinas E.** Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. *Neuron* 38: 649–657, 2003.
- Rosen MJ, Semple MN, Sanes DH.** Exploiting development to evaluate auditory encoding of amplitude modulation. *J Neurosci* 30: 15509–15520, 2010.
- Salinas E, Hernandez A, Zainos A, Romo R.** Periodicity and firing rate as candidate neural codes for the frequency of vibrotactile stimuli. *J Neurosci* 20: 5503–5515, 2000.
- Sarro EC, Rosen MJ, Sanes DH.** Taking advantage of behavioral changes during development and training to assess sensory coding mechanisms. *Ann NY Acad Sci* 1225: 142–154, 2011.
- Schneider DM, Woolley SMN.** Discrimination of communication vocalizations by single neurons and groups of neurons in the auditory midbrain. *J Neurophysiol* 103: 3248–3265, 2010.
- Schnupp JWH, Hall TM, Kokelaar RF, Ahmed B.** Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. *J Neurosci* 26: 4785–4795, 2006.
- Schreiner CE, Urbas JV.** Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields. *Hear Res* 32: 49–64, 1988.
- Shadlen MN, Britten KH, Newsome WT, Movshon JA.** A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J Neurosci* 16: 1486–1510, 1996.
- Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M.** Speech recognition with primarily temporal cues. *Science* 270: 303–304, 1995.
- Singh NC, Theunissen FE.** Modulation spectra of natural sounds and ethological theories of auditory processing. *J Acoust Soc Am* 114: 3394–3411, 2003.
- Steinschneider M, Fishman YI, Arezzo JC.** Representation of the voice onset time (VOT) speech parameter in population responses within primary auditory cortex of the awake monkey. *J Acoust Soc Am* 114: 307–321, 2003.
- Swanson WH, Pan F, Lee BB.** Chromatic temporal integration and retinal eccentricity: psychophysics, neurometric analysis and cortical pooling. *Vision Res* 48: 2657–2662, 2008.
- Uka T, DeAngelis GC.** Contribution of middle temporal area to coarse depth discrimination: comparison of neuronal and psychophysical sensitivity. *J Neurosci* 23: 3515–3530, 2003.
- Vogels R, Orban GA.** How well do response changes of striate neurons signal differences in orientation: a study in the discriminating monkey. *J Neurosci* 10: 3543–3558, 1990.
- Walker KMM, Ahmed B, Schnupp JWH.** Linking cortical spike pattern codes to auditory perception. *J Cogn Neurosci* 20: 135–152, 2008.
- Wallace MN, Coomber B, Sumner CJ, Grimsley JMS, Shackleton TM, Palmer AR.** Location of cells giving phase-locked responses to pure tones in the primary auditory cortex. *Hear Res* 274: 142–151, 2011.
- Wang L, Narayan R, Graña G, Shamir M, Sen K.** Cortical discrimination of complex natural stimuli: can single neurons match behavior? *J Neurosci* 27: 582–589, 2007.
- Wang X, Kadia SC.** Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. *J Neurophysiol* 86: 2616–2620, 2001.
- Wang X, Merzenich MM, Beitel R, Schreiner CE.** Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J Neurophysiol* 74: 2685–2706, 1995.
- Warrant EJ.** Seeing in the dark: vision and visual behavior in nocturnal bees and wasps. *J Exp Biol* 211: 1737–1746, 2008.
- Yin P, Johnson JS, O'Connor KN, Sutter ML.** Coding of amplitude modulation in primary auditory cortex. *J Neurophysiol* 105: 582–600, 2011.
- Yost WA.** Auditory image perception and analysis: the basis for hearing. *Hear Res* 56: 8–18, 1991.
- Young ED.** Neural representation of spectral and temporal information in speech. *Philos Trans R Soc B Biol Sci* 363: 923–945, 2008.
- Zheng Y, Escabi MA.** Distinct roles for onset and sustained activity in the neuronal code for temporal periodicity and acoustic envelope shape. *J Neurosci* 28: 14230–14244, 2008.
- Zhou Y, Wang X.** Cortical processing of dynamic sound envelope transitions. *J Neurosci* 30: 16741–16754, 2010.
- Zohary E, Shadlen MN, Newsome WT.** Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* 370: 140–143, 1994.