

Integrating Visual Information into the Auditory Cortex Promotes Sound Discrimination through Choice-Related Multisensory Integration

Song Chang,^{1*} Jinghong Xu,^{1*} Mengyao Zheng,¹ Les Keniston,² Xiaoming Zhou,¹ Jiping Zhang,¹ and Liping Yu¹

¹Key Laboratory of Brain Functional Genomics (Ministry of Education and Shanghai), East China Normal University, Shanghai 200062, China, and ²Department of Biomedical Sciences, Kentucky College of Osteopathic Medicine, University of Pikeville, Pikeville, Kentucky 41501

An increasing number of studies have shown that cross-modal interaction can occur in early sensory cortices. Yet, how neurons in sensory cortices integrate multisensory cues in perceptual tasks and to what extent this influences behavior is largely unclear. To investigate, we examined visual modulation of auditory responses in the primary auditory cortex (A1) in a two-alternative forced-choice task. During the task, male rats were required to make a behavioral choice based on the pure tone frequency (low vs high) of the self-triggered stimulus to get a water reward. The result showed that the presence of a noninformative visual cue did not uniformly influence auditory response, with frequently enhancing just one of them. Closely correlated with behavioral choice, the visual cue mainly enhanced responsiveness to the auditory cue indicating a movement direction contralateral to A1 being recorded. Operating in this fashion provided A1 neurons a superior capability to discriminate sound during multisensory trials. Concomitantly, behavioral data and decoding analysis revealed that visual cue presence could speed the process of sound discrimination. We also observed this differential multisensory integration effect in well-trained rats when tested with passive stimulation and under anesthesia, albeit to a much lesser extent. We did not see this differentially integrative effect while recording in A1 in another similar group of rats performing a free-choice task. These data suggest that auditory cortex can engage in meaningful audiovisual processing, and perceptual learning can modify its multisensory integration mechanism to meet task requirements.

Key words: auditory cortex; behavioral task; cross-modal interaction; multisensory integration; task engagement; visual

Significance Statement

In the natural environment, visual stimuli are frequently accompanied by auditory cues. Although multisensory integration has traditionally been seen as a feature of associational cortices, recent studies have shown that cross-modal inputs can also influence neuronal activity in primary sensory cortices. However, exactly how neurons in sensory cortices integrate multisensory cues to guide behavioral choice is still unclear. Here, we describe a novel model of multisensory integration used by A1 neurons to shape auditory representations when rats performed a cue-guided task. We found that a task-irrelevant visual cue could specifically enhance the response of neurons in sound guiding to the contralateral choice. This differentially integrative model facilitated sound discrimination and behavioral choice. This result indicates that task engagement can modulate multisensory integration.

Introduction

Sensory cortices were long thought to be exclusively dedicated to a single sensory modality. However, an increasing number of studies show that early sensory cortices are involved in processing cross-modal interaction (Ghazanfar and Schroeder, 2006; Driver and Noesselt, 2008; Vasconcelos et al., 2011). Electrophysiological studies have demonstrated that neuronal activities in auditory cortex can be modulated by visual (Ghazanfar et al., 2005; Bizley et al., 2007; Kayser et al., 2008; Perrotin et al., 2015; Atilgan et al., 2018) and somatosensory stimulation (Fu et al., 2003; Meredith and Allman, 2015). Intracellular recordings in mice have shown that many neurons in sensory cortices are unisensory in spiking

Received Apr. 22, 2022; revised Aug. 22, 2022; accepted Sep. 14, 2022.

Author contributions: J.X. and L.Y. designed research; S.C. and M.Z. performed research; L.Y. contributed unpublished reagents/analytic tools; S.C., M.Z., J.X., X.Z., J.Z., and L.Y. analyzed data; J.X., L.K., and L.Y. wrote the paper.

This work was supported by Technology Innovation 2030 Major Projects on Brain Science and Brain-like Computing of the Ministry of Science and Technology of China Grant 2021ZD0202600, National Natural Science Foundation of China Grant 31970925, and Shanghai Natural Science Foundation Grant 20ZR1417800. We thank Qin Nan for technical assistance in preparing the paper.

The authors declare no competing financial interests.

*S.C. and J.X. have contributed equally to this work.

Correspondence should be addressed to Liping Yu at lpyu@bio.ecnu.edu.cn.

<https://doi.org/10.1523/JNEUROSCI.0793-22.2022>

Copyright © 2022 the authors

behavior but multisensory in intracellular signals (Olcese et al., 2013; Ibrahim et al., 2016). It is an emerging notion that multisensory association learning could increase the prevalence of cross-modally responsive neurons in sensory cortices (Xu et al., 2014; Vincis and Fontanini, 2016; Knöpfel et al., 2019; Han et al., 2021). In addition, anatomic studies have disclosed reciprocal nerve projections among the somatosensory, auditory, and visual cortices in monkeys, ferrets, and rats (Falchier et al., 2002; Cappe and Barone, 2005; Bizley et al., 2007; Campi et al., 2010; Stehberg et al., 2014).

Despite these observations, essential questions remain on the extent, origin, modulatory effect, and function of inputs from other sensory areas into regions of the brain concerned primarily with modality-specific processing. Also, many previous studies examining neuronal multisensory integration in sensory cortices were conducted either in anesthetized or awake animals in a passive state (Kayser et al., 2008; Meijer et al., 2017; Atilgan et al., 2018; Deneux et al., 2019). Indeed, relatively few studies have investigated cross-modal interaction in sensory cortices of awake, freely moving animals, and fewer in rodents specifically (Hirokawa et al., 2011; Raposo et al., 2014). In addition, the majority of studies examining cross-modal interaction at the neuronal level were focused on the influence of physical properties, spatial location, and the timing of stimuli on the combination of multisensory cues (Kayser et al., 2008; Meijer et al., 2017; Xu et al., 2018; Deneux et al., 2019; Wallace et al., 2020). To date, we know little about how sensory cortices process cross-modal interaction in freely moving animals during task engagement. It also remains an open question about how cue discrimination, behavioral choice, and perceptual learning influence the multisensory process in sensory cortices. A growing number of studies report that primary sensory areas do not just extract and encode stimulus physical attributes but can reflect factors such as motor activity, arousal, learning, and so on (Brosch et al., 2005; Niell and Stryker, 2010; Bizley et al., 2013; Zhou et al., 2014; Kuchibhotla et al., 2017; Huang et al., 2019). For instance, previous studies have shown that neurons in the early sensory cortex, analogous to those in the frontal cortex, can adapt rapidly to contextual demands when animals engage in cue discrimination tasks (Bagur et al., 2018; Elgueda et al., 2019). Thus, we are inspired to design the corresponding behavioral tasks to explore the cross-modal interaction in sensory cortices.

In the present study, we scrutinized the impact of visual stimulus on auditory responses by recording single units from the A1 of rats performing a sound frequency discrimination task. During the course of the task, rats were required to discriminate stimuli based on auditory signal alone (two pure tones of different frequencies sometimes paired with an invariant uninformative visual cue) and then make a behavioral choice (left or right). The result showed that A1 neurons, in the presence of a visual cue, frequently developed an enhanced response that facilitated sound discrimination. This differentially integrative processing model was not seen in a free-choice task. As a result, we conclude that the neural activities forwarded from A1 to subsequent sensory processing stages not only contain information about acoustical events but also can reflect aspects of the visual environment and task contexts.

Material and Methods

Rat subjects. Animal procedures were approved by the Local Ethical Review Committee of East China Normal University and conducted in accordance with the *Guide for the Care and Use of*

Laboratory Animals of East China Normal University. Twenty-one adult male Sprague Dawley rats, provided by the Shanghai Laboratory Animal Center, were used for the experiments. At the start of behavioral training, these rats were 250–300 g each and were 4–6 months old. Each was housed as one animal per cage under constant temperature ($23 \pm 1^\circ\text{C}$) with a normal diurnal light cycle. All rats had access to food *ad libitum*. Water was restricted only on experimental days up to the behavioral session. Rats are usually trained 6 d per week, in one 50–80 min session held at approximately the same time of day. Body weight was carefully monitored and kept above 80% of the age-matched control rats undergoing no behavioral training.

Behavioral task. Rats were required to perform a cue-guided two-alternative forced-choice task slightly modified from other published protocols (Jaramillo and Zador, 2011; Zheng et al., 2021). Automated training was controlled using a custom-built program running on MATLAB 2015b (MathWorks). The training was conducted in an open-topped custom-built operant chamber made of opaque plastic (size, $50 \times 30 \times 40$ cm, length \times width \times height) in a sound-insulated double-walled room with its inside walls and ceiling covered by 3 inches of sound-absorbing foam to reduce acoustic reflections. Three snout ports, each monitored by a photoelectric switch, are located on one sidewall of the operant chamber (Fig. 1A). The signals from the photoelectric switches were first fed to an analog-digital multifunction card and digitized (DAQ NI USB-6363, National Instruments) and sent via USB to a PC running the training program.

Rats initiated a trial by poking their nose into the center port. Following a short variable delay (500–700 ms), a stimulus (one of four options, two auditory and two auditory visual) was presented randomly. After the presentation of this cue, rats could immediately initiate their behavioral choice, moving to the left or right port (Fig. 1A). If rats made a correct choice (hit trial), they could obtain a water reward, and a new trial could immediately follow. If rats made a wrong choice or made no behavioral choice within 3 s after cue onset, a punishment time-out of 5–6 s was applied.

The auditory cue was delivered via a speaker (FS Audio), using a 300-ms-long 3 kHz (low) or 10 kHz (high) pure tone with a 25 ms ramp decay given at 60 dB sound pressure level (SPL) against an ambient background of 30–35 dB SPL. SPLs were measured at the position of the central port (the starting position). The visual cue was a 300-ms-long flash of white light given at $5 \sim 7 \text{ cd/m}^2$ intensity, delivered by a light-emitting diode. The auditory-visual cue (multisensory cue) was the simultaneous presentation of both auditory and visual cues. Rats were trained to the competency of >80% correct in three consecutive sessions before surgical implantation of recording electrodes.

Assembly of tetrodes. Formvar-insulated Nichrome wire (bare diameter, 17.78 μm , A-M Systems) was twisted in groups of four as tetrodes (impedance, 0.5–0.8 M Ω at 1 kHz). Two 20 cm-long wires were folded in half over a horizontal bar for twisting. The ends were clamped together and manually twisted clockwise. Finally, their insulation coating was fused with a heat gun at the desired twist level and cut in the middle to produce two tetrodes. To reinforce each tetrode longitudinally, each tetrode was then inserted into polyimide tubing (inner diameter, 0.045 inches; wall, 0.005 inches; A-M Systems) and fixed in place by cyanoacrylate glue. An array of 2×4 tetrodes were then assembled using an inter-tetrode gap of 0.4–0.5 mm. After assembly, the insulation coating of each wire was gently removed at the tip and

soldered to a connector pin. The reference electrode used was a tip-exposed Ni-Chrome wire, diameter 50.8 μm (A-M Systems), and a ground electrode was a piece of copper wire with a diameter of 0.1 mm. Each of these was also soldered to a corresponding pin on the connector. Tetrodes and connectors were then carefully arranged and cemented with silicone gel. The tetrodes were trimmed to an appropriate length immediately before implantation.

Electrode implantation. The animal received subcutaneous injections of atropine sulfate (0.01 mg/kg) before surgery to reduce bronchial secretions. Immediately before surgery, the animal was anesthetized with an intraperitoneal injection of sodium pentobarbital (40–50 mg/kg) and then fixed on a stereotaxic apparatus (RWD). The skin was incised and the temporal muscle recessed. Then, a craniotomy and a durotomy were performed. The tetrode array was then implanted in the primary auditory cortex (3.5 to –5.0 mm posterior to bregma and 6.4 mm lateral to the midline) by slowly advancing a micromanipulator (RWD). Tissue gel (3M) was then used to seal the craniotomy. The tetrode array was next secured to the skull with stainless steel screws and dental acrylic. After surgery, the animals were given a 4 d prophylactic course of antibiotics (Baytril, 5 mg/kg, body weight; Bayer). They had a recovery period of at least 7 d (usually 9–12 d with free access to food and water).

Neural recordings in task engagement. When fully recovered from the surgery, animals resumed performing the same behavioral task. Recording sessions began after the behavioral performance of the animal recovered to the level attained before surgery (typically 2–3 d). Wideband neural signals (250–6000 Hz) were recorded using a head-stage amplifier (RHD2132, IntanTech). Amplified ($\times 20$) and digitized (at 20 kHz) neural signals were combined with trace signals representing both the stimuli and session performance information and sent to a USB interface board (RHD2000, Intan Technologies), and then to a PC for online observation and data storage.

Neural recordings in anesthesia. Rats were anesthetized with an intraperitoneal injection of sodium pentobarbital (40 mg/kg, body weight), and the anesthesia was maintained throughout the experiment by continuous intraperitoneal infusion of sodium pentobarbital (0.004–0.008 g/Kg/h) via an automatic microinfusion pump (WZ-50C6, Smiths Medical). Body temperature was maintained at 37.5°C using a heating blanket. We placed the anesthetized rats in the behavioral training chamber and put the head into the cue port to imitate the cue trigger as they did in task engagement. The stimuli (auditory, visual, and multisensory) used were the same as task engagement and presented randomly. We recorded 40–60 trials of response for each cue condition.

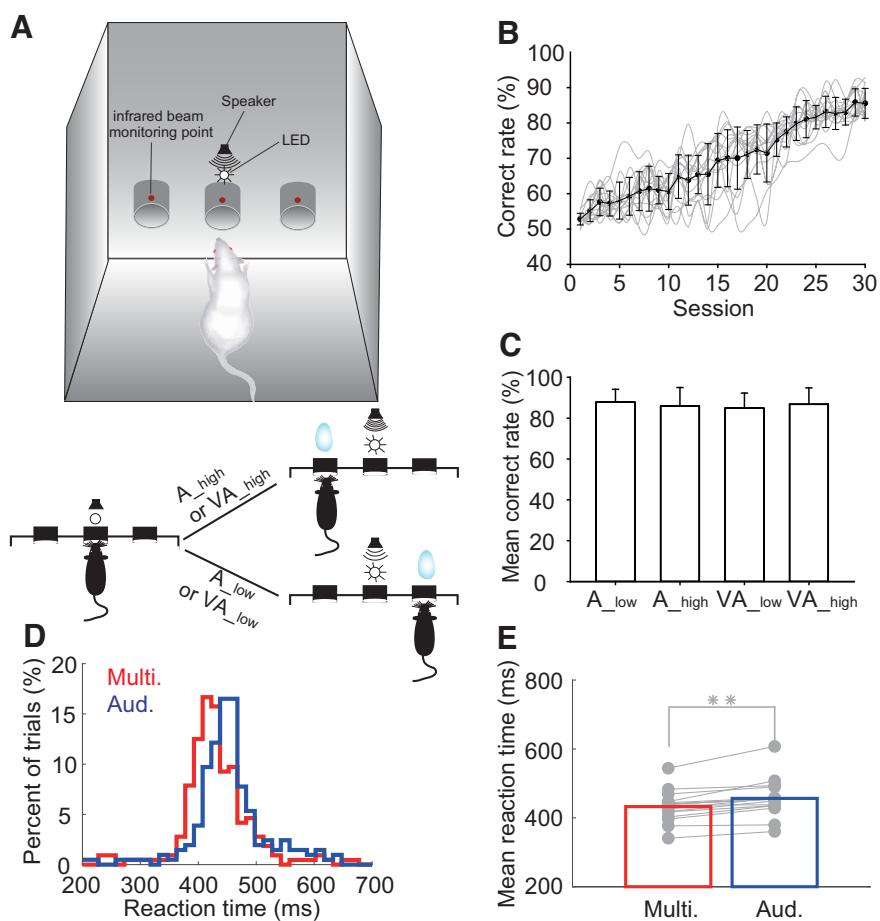


Figure 1. Stimulus discrimination task and behavioral performance. **A**, A schematic of the behavioral paradigm used. An individual trial started when trained rats placed their nose in the central port. Next, a lone auditory stimulus or a combined auditory and visual stimulus was presented via a centrally positioned light-emitting diode (LED) and speaker to cue the location for a water reward. When the stimulus was either a 10 kHz pure tone (A_{high}) or a combination of a 10 kHz pure tone and a flash of light (VA_{high}), the animal would be rewarded at the left port. If the stimulus given was a 3 kHz pure tone alone (A_{low}) or the same tone paired with a flash of light (VA_{low}), the animal would receive the reward in the right port. Trials of different stimuli (A_{low} , A_{high} , VA_{low} , VA_{high}) were presented in a randomized order. **B**, Performances of all rats ($n = 15$) while learning this two-alternative forced-choice task. Each gray line represents data from one rat. The black line represents the overall mean. Error bar indicates SD. **C**, The mean performance across all rats for each stimulus condition. Error bar indicates SEM. **D**, The distribution of reaction times for both auditory (blue) and multisensory (red) trials performed by a well-trained rat. **E**, A comparison of mean reaction times between auditory and multisensory trials across all rats. ** $p < 0.001$.

Histology. After the last data recording session, the final tip position of the recording electrode was marked with a small DC lesion ($-30 \mu\text{A}$ for 15 s). Afterward, rats were deeply anesthetized with sodium pentobarbital (100 mg/kg) and perfused transcardially with saline for several minutes, followed immediately by PBS with 4% paraformaldehyde (PFA). Their brains were carefully removed and stored in the 4% PFA solution overnight. After cryoprotection in PBS with 20% sucrose solution for at least 3 d, the fixed brain tissue was sectioned in the coronal plane on a freezing microtome (Leica) at a slice thickness of 50 μm . Sections containing the auditory cortex were stained with methyl violet to verify the lesion sites and/or the trace of electrode insertion to be located within the primary auditory cortex.

Data analysis. The reaction time was defined as the temporal gap between the cue onset and the time that an animal withdrew its nose from the infrared beam in the cue port. The correct performance rate was defined by the following:

$$\text{Correct performance rate(\%)} = 100 * \frac{\text{hit trials}}{\text{total number of trials}}$$

Raw neural signals were recorded and stored for off-line analysis. Spike sorting was later performed using Spike 2 software (Cambridge Electronic Design, version 8). Recorded raw neural signals were first bandpass filtered in 300–6000 Hz to remove field potentials. A threshold criterion of no less than threefold SDs above background noise was then used for identifying spike peaks. The detected spike waveforms were then clustered by principal component analysis and a template-matching algorithm. Waveforms with interspike intervals of <2.0 ms were excluded. Relative spike timing data for a single unit were then obtained for different trials of different cued conditions and used to construct raster plots and peristimulus time histograms (PSTHs) using custom MATLAB scripts. Only neurons for which the overall mean firing rate within the session was at least 2 Hz were included for analysis. As generally observed, behavioral and neuronal results were similar across all relevant animals for a particular testing paradigm. Thus, the data across sessions were combined to study population effects.

To render PSTHs, all spike trains were first binned at 10 ms and convolved with a Gaussian smoothing kernel ($\delta = 100$ ms) to minimize the impact of random spike-time jitter at the borders between bins. The mean spontaneous firing rate was calculated from a 500 ms window immediately preceding stimulus onset. Cue-evoked responses were quantified as mean firing rates found in a 400 ms window after cue onset minus the mean spontaneous firing rate.

We quantified the cue selectivity between two different cue conditions (e.g., low-tone trials vs high-tone trials) by using an analysis based on a receiver operating characteristic (ROC; Britten et al., 1992). First, we set 12 threshold levels of activity covering the range of firing rates obtained in cue_A and cue_B trials. Following that, a ROC curve is generated for each threshold criterion by plotting the proportion of cue_A trials on which the response exceeded the criterion against the proportion of cue_B trials on which the response exceeded the criterion. The value of cue selectivity is defined as $2 * [(area \ under \ the \ ROC \ curve) - 0.5]$. Therefore, a value of zero indicates no difference in the distribution of responses between cue_A and cue_B. A value of 1/−1 represents the highest selectivity; that is, responses triggered by cue_A were always higher or lower than those evoked by cue_B.

We ran a permutation test to test the significance of each cue selectivity value. This was accomplished by randomly distributing all trials from a neuron into two groups, independent of the actual cue conditions. These groups were nominally called cue_A trials and cue_B trials and contained the same number of trials as the experimentally obtained groups. The cue selectivity value was then calculated from the redistributed data, and the procedure was repeated 5000 times, thereby giving a distribution of values from which to calculate the probability of the result we obtained. When our actual value was found in the top 5%, it was defined as significant (i.e., $p < 0.05$).

Support Vector Machine decoding. To compare the population decoding accuracy of auditory neurons for tone frequencies with or without the presence of visual cue, we trained Support Vector Machine (SVM) classifier with a linear kernel to predict frequency selectivity (itcsvm in MATLAB). Briefly, the spike counts for each neuron in correct trials were grouped by

triggered cues and binned into a 100 ms window moving at 10 ms resolution. Only neurons with >40 trials for each cue were included to minimize overfitting. All these neurons were pooled as a pseudo population. Responses of population neurons were arranged in an $M \times N \times T$ matrix, where M is the number of trials, N is the number of neurons, and T is the number of bins. For each resample run, 40 trials were randomly selected for each cue from each neuron. For each cross-validation, we randomly sampled 90% of trials as the training set and the remaining 10% of trials as the test set. The training set was used to compute the linear hyperplane that optimally separated the population response vectors corresponding to high-tone versus low-tone trials. Performance was calculated as the fraction of correct classifications of the test trials. The classification accuracy was estimated over the 10-fold cross-validation procedures. We repeated the resampling process 100 times and computed the mean and SD of decoding accuracy across the 100 resampling iterations. Decoders were trained and tested independently for each bin.

Experimental design and statistical analyses. All statistical analyses were conducted in MATLAB with statistical significance assigned for findings attaining a p value of < 0.05 . To determine whether an A1 neuron was responsive to sensory stimulus, we computed the baseline response during the 0.5 s before stimulus onset and the evoked firing rate within the time window of 0.4 s after the stimulus onset for each trial. Those A1 neurons in which the evoked responses were larger than 2 spikes/s and were significantly higher than the baseline response ($p < 0.05$, Wilcoxon signed-rank test) were used in the analysis. All behavioral data (e.g., mean reaction time differences between auditory and multisensory trials) and multisensory integrative index between different tone conditions were compared using the Wilcoxon signed-rank test or paired t test. We performed the chi-square test to analyze the difference in proportions of neurons showing tone-frequency selectivity. Permutation tests were used to measure the significance of the multisensory enhancement and the strength of sound frequency preference. The data were not collected in a blinded fashion. Unless stated otherwise, all data group results are presented as mean \pm SD.

Results

Unisensory and multisensory decision-making behavior in freely moving rats

We trained rats ($n = 15$) to perform a well-established cue-guided two-alternative, forced-choice task (Fig. 1A). Rats were required to make a specific behavioral choice based on whether the auditory stimulus or the auditory component of a multisensory cue was a lower (3 kHz) or higher frequency (10 kHz) pure tone. A trial was initiated when the rat poked its nose into the central port in a line of three ports on one wall of the training box (Fig. 1A). After the waiting period of 500–700 ms, a cue, randomly chosen from a group of four cues (3 kHz pure tone, A_{low} ; 10 kHz pure tone, A_{high} ; 3 kHz pure tone plus flash of light, VA_{low} ; 10 kHz pure tone plus flash of light, VA_{high}), was presented in front of the central port. Based on the cue, rats were required to choose a port (left or right) to obtain a water reward within 3 s. If the stimulus was A_{high} or VA_{high} , the rat should move to the left port for harvesting the water reward (Fig. 1A). The other two cues indicated rats should move to the right port for a reward. After rats reached the criteria of 80% correct in three consecutive sessions, they were deemed well trained. On

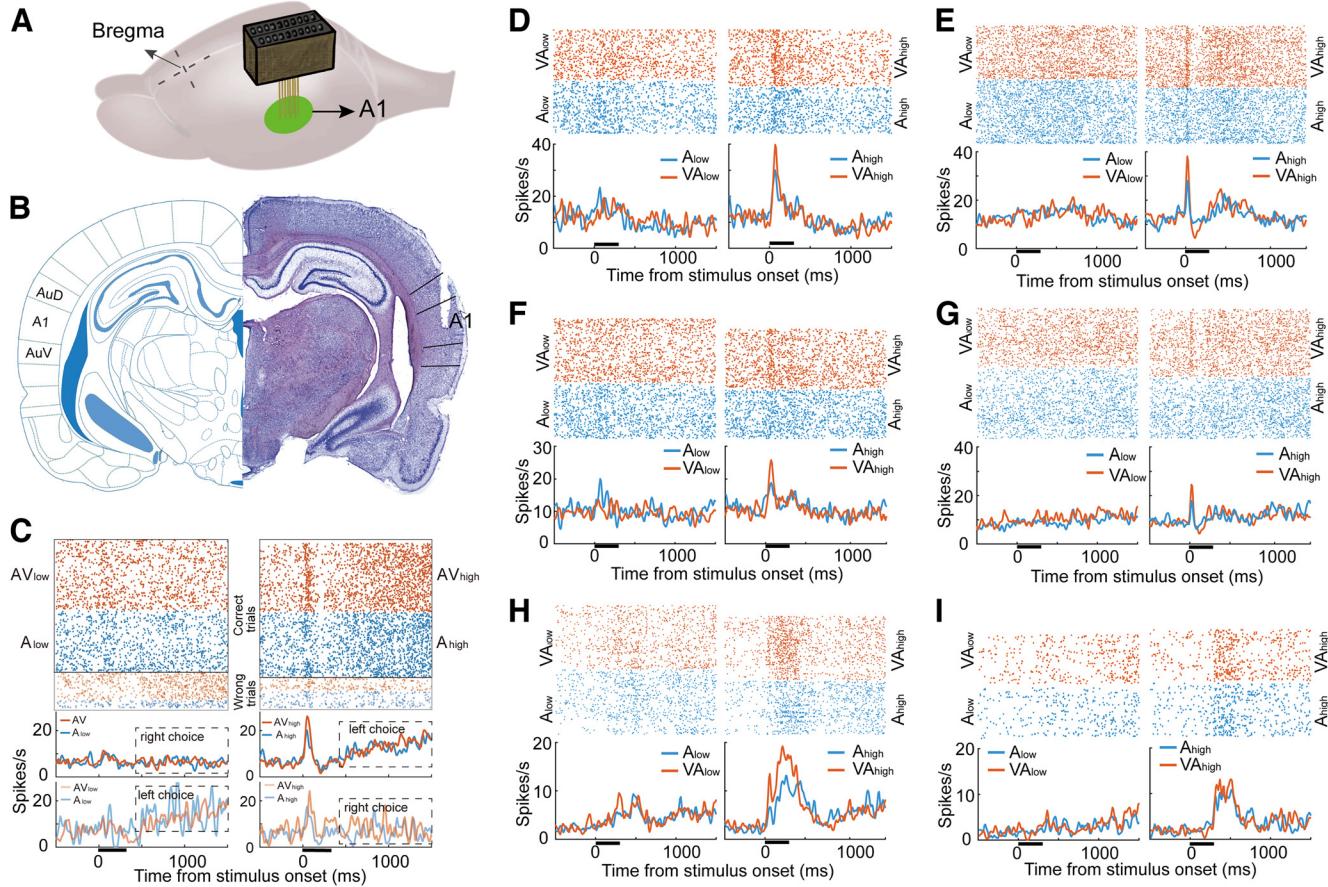


Figure 2. Neuronal activity during task performance and histologic verification of recording sites. **A**, The target area, the primary auditory cortex (A1, green circle), was recorded by multi-channel electrodes. **B**, The histologic section demonstrates the recording site verified within A1. **C**, Rasters (top rows) and PSTHs (bottom traces) in both correct (dark color) and error (light color) trials of each given cue condition. Note that the sensory signal preceded the behavioral choice signal denoted by a dashed rectangle. In PSTHs, mean spike counts across trials of each cue were computed in 10 ms time windows and smoothed with a Gaussian kernel ($\sigma = 100$ ms). Responses were aligned to the initial cue presentation. In multisensory trials, visual and auditory stimuli were presented simultaneously. Black bars indicate the stimulus onset and duration. **D–I**, Shows six more example neurons (correct trials only).

average, the behavioral performance stabilized at $86 \pm 7.4\%$ (Fig. 1B). There was no difference in behavioral performance between auditory and multisensory-cued trials (Fig. 1C). Despite this, the presence of a visual cue sped up the process of cue discrimination. The reaction time, defined as the temporal gap between the cue onset and the moment when the animal withdrew its nose from the infrared beam monitoring point in the central port (Fig. 1A), was compared between auditory and multisensory trials (Fig. 1D,E). Note that rats responded more quickly in multisensory trials (mean reaction time across rats, 432 ± 46 ms in multisensory conditions; 456 ± 54 ms in auditory conditions; $p < 0.001$, Wilcoxon signed-rank test; Fig. 1E).

Frequency discrimination of A1 neurons in task engagement

To investigate the visual modulation of the response of A1 neurons in the cue discrimination task, we used tetrode recordings to characterize the activity of A1 neurons (Fig. 2A,B). We measured each the characteristic frequency (CF) of each neuron directly after successful tetrode implantation. We did not see any frequency biasing of the CF sampling distribution throughout experimentation. On average, rats performed 254 ± 49 trials in a daily session. We recorded a total of 721 A1 neurons (48 ± 24 neurons per animal) when well-trained rats performed the task. Among them, 68% (489/721) of neurons elicited apparent cue-evoked responses. However, 14% (70/489) of them showed cue-evoked inhibitory responses and were precluded from further

analysis. Thus, our following analysis was focused on the remaining 419 neurons, 224 neurons from the right auditory cortex and 195 from the left auditory cortex.

The majority of neurons (72%, 302/419) showed a single-peaked response during the initial period of cue presentation (Fig. 2C–G), and 19% (79/419) of neurons showed two separated responses, with a quick onset response and a late offset response (Fig. 2E). A small subset of neurons elicited a sustained response (Fig. 2H). Fewer neurons appeared to show an offset response (Fig. 2I). In addition, some neurons (9%, 38/419) showed both cue-evoked and behavioral-choice-related responses (Fig. 2C). However, these two distinct responses could be easily separated as the behavioral-choice-related response occurred much later (frequently 400 ms later following the cue onset).

Because the vast majority of neurons showed their main cue-evoked response in the initial period of cue presentation, we focused our analysis on responses obtained at the window of 0–150 ms after cue onset. We used an ROC analysis to generate an index of auditory preference that measures how strongly the firing rate of a neuron for low-tone trials diverged from the firing rate for high-tone trials. In the same way, an index of multisensory preference was defined. We found that 47% (199/419) of neurons showed a significant frequency preference in task engagement (permutation test, $p < 0.05$; Fig. 2C–F, individual cases; Fig. 3A, population data). A subset of neurons (142/419)

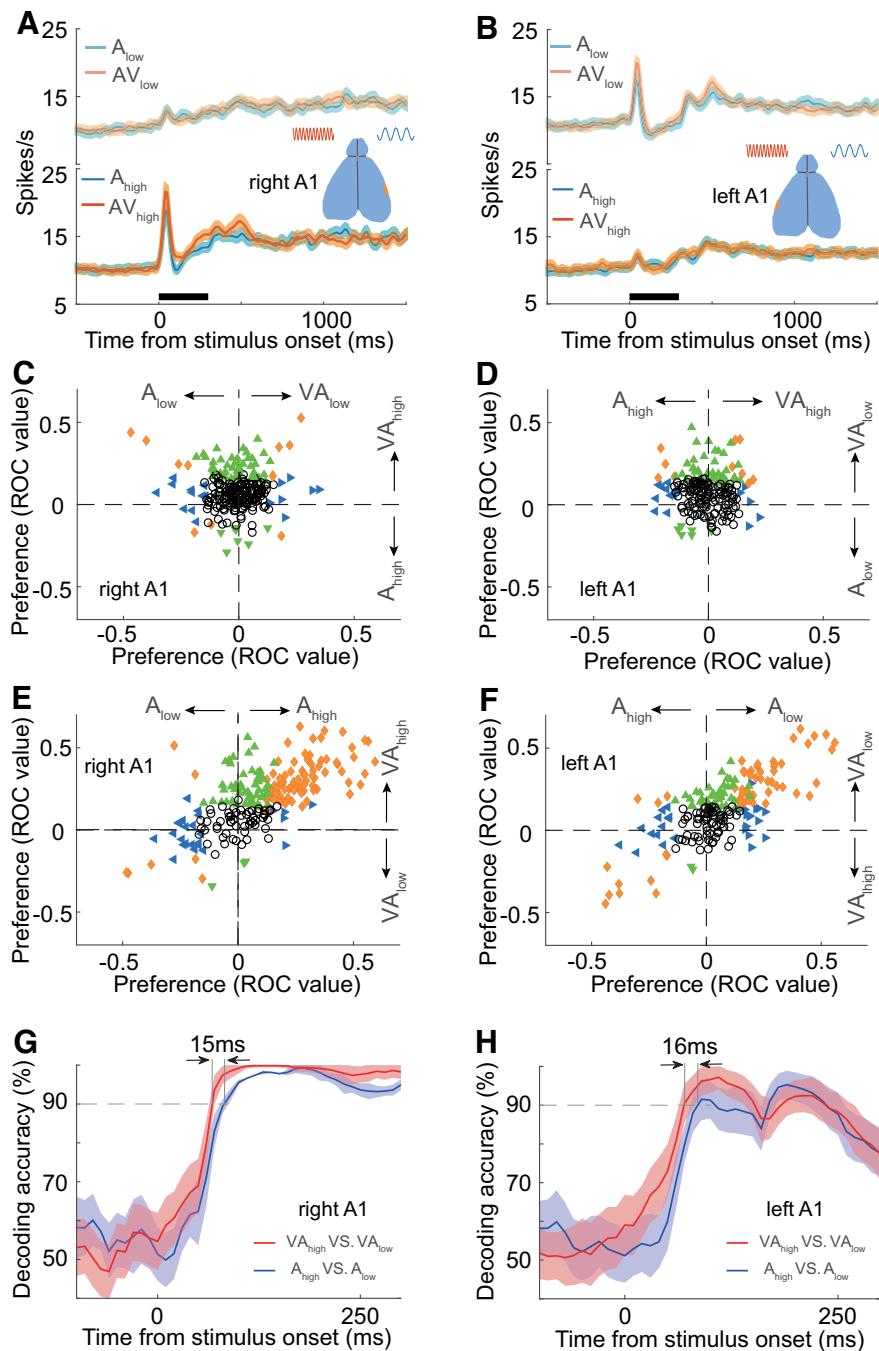


Figure 3. Distributions of neuronal cue preferences and mean responses in task engagement. **A, B**, PSTHs show mean responses across populations for different stimulus trials (correct only) for neurons recorded in the right (as denoted by the brain schematic, **A**) and left (**B**) A1. Note that neurons preferred the tone sound denoting the contralateral choice being recorded, and the visual cue specifically enhanced this preferred response only. Shaded areas indicate SEM. **C, D**, Modality preference (auditory vs multisensory) was compared between low- (abscissa) and high-tone (ordinate) conditions for neurons recorded in the right (**C**) and left (**D**) A1. Each point shows values for a single neuron. Open circles indicate modality preference in neither low- nor high-tone condition was significant ($p < 0.05$, permutation test, 5000 iterations), triangles indicate modality preference in either low- (blue) or high-tone (green) condition was significant, and red diamonds indicate modality preference in both low- and high-tone conditions was significant. Dashed lines represent zero ROC values. **E, F**, Similarly, auditory and multisensory preferences were compared for neurons recorded in the right (**E**) and left (**F**) A1. **G, H**, Decoding accuracy of population neurons (cross-validation accuracy based on SVM) in task engagement for tone discrimination both in auditory conditions (blue) and in multisensory conditions (red). The dashed line represents 90% of decoding accuracy. Shaded areas indicate SD. Note that the visual cue hastened tone discrimination regardless of whether the neuron was recorded on the right (**G**) or the left (**H**) A1.

appeared to show responsiveness only to one of two tones in task engagement. We failed to see such extreme frequency selectivity in the passive and anesthetized conditions (see below), suggesting that frequency preference is related to the perceptual task. Figure 2, *E* and *G*, illustrates such two exemplar cases. In addition, we found a significant difference in frequency preference between left A1 neurons versus right A1 neurons. Most right A1 neurons preferred the high tone (the number of neurons preferring high tone, 88; preferring low tone, 26; Fig. 3*A*). For example, all neurons shown in Figure 2 were recorded from the right auditory cortex, indicating that all of them preferred high-tone sound.

In contrast, left A1 neurons showed the reversal (the number of neurons preferring high tone, 25; preferring low tone, 60). This result was in line with a previous study conducted in the mouse auditory cortex that a population tone frequency preference could be induced in a tone discrimination task (Xin et al., 2019). Considering that low and high tones denoted the right and left behavioral choices, our result indicates that A1 neurons selectively preferred the tone associated with contralateral behavioral choice in task engagement. We believe that such frequency-preference biasing could enable the learned transformation of sound into action as sensorimotor association learning could create direct corticocortical pathways between sensory and motor cortices (Makino et al., 2016). Similar biasing was also reported in a previous study that auditory discrimination learning preferentially potentiates corticostriatal synapses from neurons representing either high or low frequencies associated with contralateral choices (Xiong et al., 2015).

Visual modulation of the response of A1 neurons in task engagement

Visual modulation of auditory cortical responses has been previously characterized in anesthetized and passively listening animals but not while engaged in perceptual tasks. We examined the influence of the visual cue on auditory responses recorded in task engagement. As shown in exemplar cases (Fig. 2*C–I*), the simultaneous presentation of a noninformative visual cue could yet significantly modulate the responsiveness to auditory cues (permutation test, $p < 0.05$). Multisensory enhancement appeared to be the favored processing

strategy in use. Surprisingly, we found that the visual cue did not uniformly influence auditory responses and frequently just facilitated one of them (Fig. 2C–I). After peer examination, we found that for neurons showing tone preference and multisensory enhancement, the visual cue mostly enhanced the preferred auditory response and left the nonpreferred response with no significant change (Fig. 2C–E). Someone might attribute this differential response to the failure of the neurons respond to the nonpreferred sound, as shown in exemplar cases (Fig. 2E,G). However, it is not the case. For instance, like the neurons in Figure 2, D and F, show, many neurons exhibited responsiveness to both tone stimuli, but still, only the preferred was boosted by the visual stimulus. Also, we found that one-quarter of neurons (25%, 124/419) showed no significant tone preference in auditory-alone trials but exhibited preference in multisensory trials because of this multisensory enhancement biasing (Fig. 2F). This process of integration further increased frequency selectivity of neurons. The intriguing question to us was then what factor(s) govern this differentially integrative model?

Previous studies showed that perceptual discrimination and behavioral choice influence cue encoding (Fritz et al., 2003; Otazu et al., 2009; David et al., 2012; Niwa et al., 2012; Schneider et al., 2014). Our data shown above support this conclusion. We then asked whether this model of multisensory integration is also behavioral-choice dependent. Indeed, we found that visual cue specifically boosted the response to the tone associated with the contralateral choice (Fig. 3A–D), indicating that behavioral choices influence the strategy of neurons of multisensory integration. For neurons recorded from right auditory cortex ($n = 224$), visual cue primarily enhanced the response to the high tone associated with the left choice (Fig. 3A,C). Twenty-five percent (56/224) of these neurons showed the high-tone response significantly enhanced by a visual cue, and in rare cases (4%, 10/224), they showed visually induced inhibition (Fig. 3C). When the sound was a low tone, the low proportion (13%, 28/224) of neurons showed auditory responses significantly influenced by visual cue, and in this, cross-modal inhibition was even a little bit more pronounced (inhibited, 8%, 17/224; enhanced, 5%, 11/224). This choice-related multisensory integration was the same regardless of whether the neuron showed a tone preference or not. We also used ROC-based analysis to generate an index of multisensory integration (MI) that measures how strongly the response of multisensory trials diverges from that in auditory trials. Mean MI in the high-tone condition was significantly higher than in the low-tone condition (high, 0.09 ± 0.12 ; low, -0.01 ± 0.11 ; Wilcoxon signed-rank test, $p < 0.001$). Thus, on average, the $V_{A_{low}}$ response was very close to the A_{low} response (Fig. 3A), but $V_{A_{high}}$ response was a bit greater than the A_{high} response as more neurons showed multisensory enhancement.

In contrast, neurons recorded from left A1 biased visually induced multisensory enhancement primarily to low-frequency sound (low tone, enhanced, 26%, 50/195; inhibited, 3%, 6/195; high tone, enhanced, 6%, 11/195; inhibited, 7%, 14/195). Mean MI in the low-tone condition was significantly higher than in the high-tone condition (low, 0.08 ± 0.12 ; high, -0.01 ± 0.09 ; paired t test, $p < 0.001$). This result indicated that the visual cue, albeit uninformative, could facilitate perceptual discrimination, which was consistent with previous studies conducted in humans (Thorne and Debener, 2008; McDonald et al., 2013; Feng et al., 2014). Again, such choice-related multisensory integration appeared to strengthen frequency selectivity of neurons (Fig. 3E,F).

The multichannel recording method enabled us to sample a large number of neurons simultaneously, which in turn allowed us to decode sensory information on a single-trial basis from population activity in auditory cortex. We used cross-validated linear classifiers based on SVMs to decode stimulus information from the population activity by training the classifier to discriminate pairs of tone frequencies (see above, Materials and Methods). This population decoding analysis demonstrated that although the decoding accuracy could arrive at a nearly equal height in both auditory and multisensory conditions, the presence of a visual cue sped up the decoding process and facilitated the decoding accuracy of neurons to reach a high level earlier (Fig. 3G,H). This decoding result is in line with the behavioral data. It is almost 15 ms earlier in multisensory condition than in auditory-alone condition that the population neurons achieved a decoding accuracy of 90%. Summarily, these results indicate that multisensory integration can serve the task goal and aid in successful task completion.

Choice-related cross-modal interaction is independent of visual response

There is the possibility that the choice-related differential cross-modal interaction we observed might be just because of the difference in visual response between left and right choices. To test this, we introduced an individual visual cue into the task after rats performed the auditory-cue-dependent choice task. Without any training, rats ($n = 8$) could make the behavioral choice in visual-only trials. Rats could get water rewards in both left- and right-side water ports when the triggered cue was visual only. The behavioral data showed that adding visual trials decreased the behavioral performance of the rats in multisensory trials a bit (correct rate, auditory-only trials, $86.4 \pm 4.9\%$; multisensory trials, $80.8 \pm 4.0\%$; two-tailed p value = 0.0213), indicating that a noise effect was induced on decision-making in the multisensory condition. We examined a total of 324 auditory neurons recorded in this task. In task engagement, neurons recorded in the left and right A1 exhibited the same stimulus tone preference denoting the contralateral choice (Fig. 3). To examine all A1 neurons together, left A1 neurons were appropriately adjusted (the trial identity was flipped for low- vs high- tone and right vs left choice, making the orientation of cuing and choice identical) and placed in with neurons recorded in right A1. The result showed that the vast majority of neurons (92%, 298/324) failed to show visual responses in task engagement. Such a case is shown in Figure 4A, in which neurons showed no visual response but exhibited choice-related multisensory enhancement. For those ($n = 26$) that responded to visual cue, most of them (23/26) exhibited similar visual responses between left- and right-choice trials (Fig. 4B). Figure 4, C and D, shows the mean responses across populations in which we failed to see the difference in mean responses in left and right visual trials. Thus, our result indicates that choice-related differential multisensory integration should not be attributable to differences in visual response.

Also, we notice that the noise effect caused by the addition of visual trials also appeared to decrease the incidence of neurons showing visually induced enhancement of high-tone response (Fig. 4E), with 16% (52/324) versus 25% (106/419; with vs without visual trials; χ^2 test, $p = 0.0022$). Despite that, still, more neurons showed frequency preference in multisensory trials (57%, 186/324; auditory trials, 43%, 140/324; χ^2 , $p = 0.0003$, Fig. 4F).

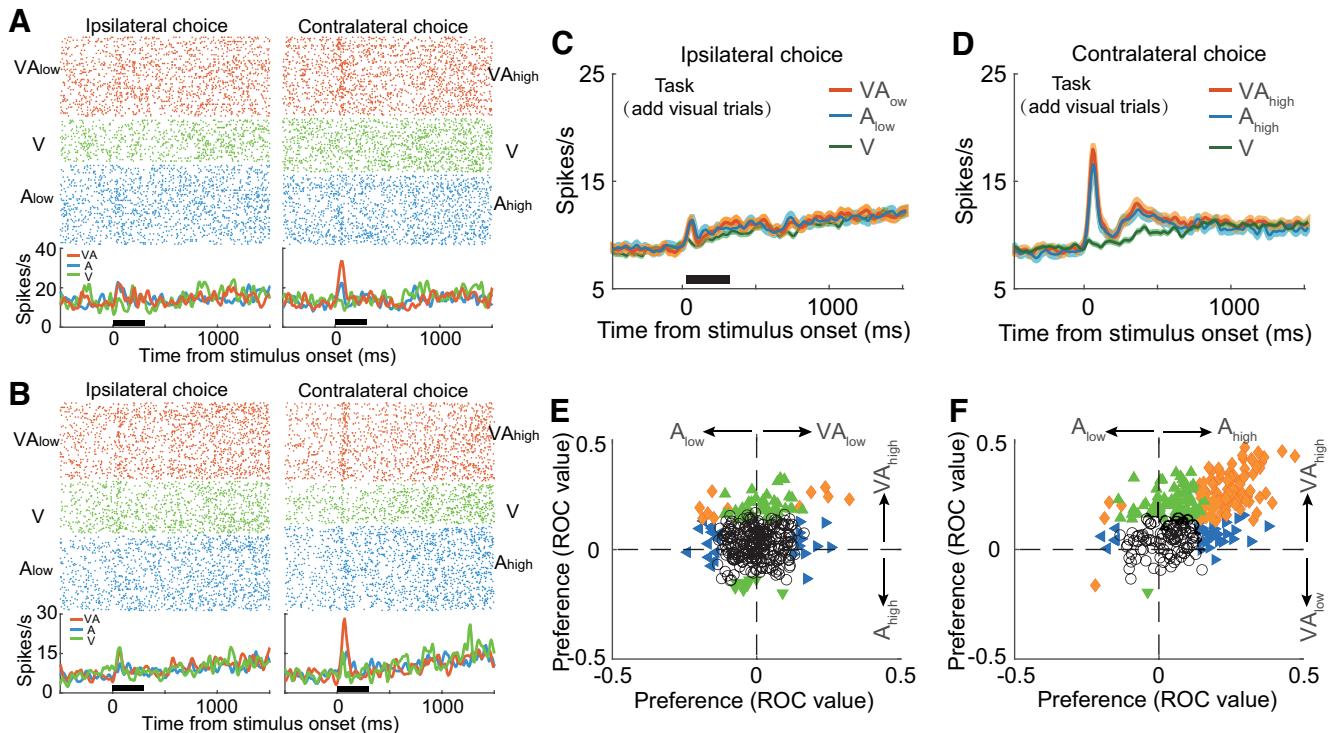


Figure 4. Neural responses, cue preferences, and multisensory integration when visual trials were added to the task. **A, B**, Rasters and PSTHs show auditory (blue), visual (green), and multisensory (orange) responses of two exemplar neurons. Left, Responses in ipsilateral choice trials. Right, Responses in contralateral choice trials. **C, D**, The mean PSTHs of different cued trials across neurons. **E**, Modality preference of neurons (auditory vs multisensory). **F**, Frequency preference of neurons (high vs low). The conventions used are the same as in Figures 2 and 3.

Visual modulation of response of A1 neurons in passive stimulation

In task engagement, rats were required to make behavioral choices for water rewards. We then asked whether the differentially integrative model could be seen only during task engagement. To answer this question, we recorded responses of A1 neurons in well-trained rats in the passive condition, conducted immediately after a task session on the same day or the following 2 d. Rats were placed into a box (Fig. 5A) and could sense the same stimuli as those they triggered in task engagement but had no task requirement. We found that the proportion of neurons showing visually induced enhancement of high-tone response declined to 11% (45/399; Fig. 5D), the number of which is 25% (106/419) in task engagement (χ^2 test, $p < 0.0001$). Despite that, this proportion is still higher than those showing multisensory inhibition (5%, 21/399; χ^2 test, $p = 0.002$). In low-tone trials, the ratio of neurons showing multisensory enhancement versus multisensory inhibition is nearly the same (enhanced, 4.8%, 19/399; inhibited, 5.5%, 22/399). The mean MI across populations for high-tone and low-tone conditions was 0.038 ± 0.113 and 0.001 ± 0.099 , respectively (Wilcoxon signed-rank test, $p < 0.001$). These data indicate that the differentially integrative model is still present but weakened in passive stimulation (Fig. 5D,F,G). As shown in Figure 5B, some neurons showed multisensory enhancement to VA_{high} cuing the contralateral movement in task engagement but failed to show the same in passive stimulations. Like the neuron shown in Figure 5C, some neurons exhibited multisensory enhancement to both preferred and nonpreferred auditory responses in the passive stimulation, which was rare in task engagement. Therefore, the differential multisensory integration in these auditory cortical neurons mainly arose from task-related modulation.

Consistently, during passive stimulations, the proportion of neurons exhibiting frequency preference (Fig. 5E) was lower than we saw in task engagement in both auditory and multisensory trials (auditory, passive vs task, 35% vs 47%; χ^2 test, $p = 0.0002$; multisensory, passive vs task, 46% vs 57%; χ^2 test, $p = 0.0008$). We believe the task-dependent enhancement of tone selectivity well served an increased discrimination demand. Still, more neurons favored the high tone in passive conditions (high vs low, 99 vs 40 for auditory trials; 149 vs 33 for multisensory trials). These exhibited residual effects (differentially integrative fashion and tone selectivity) suggest that perceptual training itself might leave a plastic trace in the auditory cortex. But another possibility is that the alert brain in passive stimulation processed sensory information in a fashion similar in task engagement despite no behavioral choice.

Visual modulation of auditory cortical responses in anesthesia

To further probe whether multisensory perceptual learning induced a plastic change in A1, we examined visual modulation of auditory cortical response when well-trained rats were under anesthesia. The vast majority of A1 neurons recorded in anesthesia failed to show multisensory enhancement (Fig. 6A, example, E,F, mean PSTHs). However, consistent with task engagement, more neurons showed multisensory enhancement in A_{high} trials (enhanced, 9%, 46/501; inhibited, 3%, 16/501; χ^2 test, $p = 0.0008$). Figure 6B provides such a case. In A_{low} trials, the ratio of neurons showing multisensory enhancement versus inhibition was identical (enhanced, 6%, 23/501; inhibited, 8%, 34/501; χ^2 test, $p = 0.13$). This result reveals that the differential multisensory integration is still present during anesthesia, suggesting the perceptual training of associating multisensory cues with the target produced a plastic change in A1. Compared with the data in

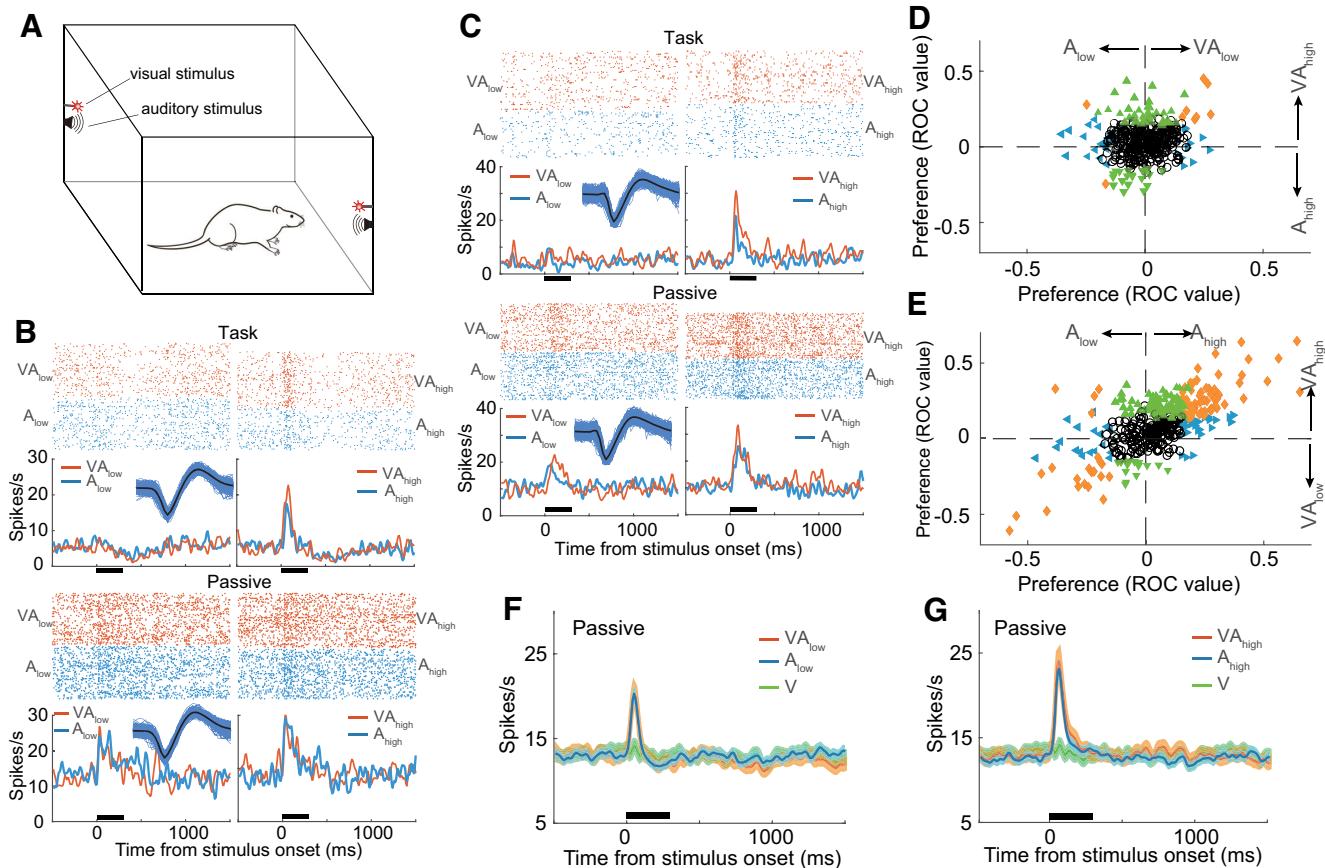


Figure 5. Visual modulation of A1 auditory response in passive stimulation. **A**, The schematic diagram of passive stimulation. Rats could freely move in the box (size, 40×40 cm) in which tone sounds, a flash of light, or a combination of the two could randomly elicit from two corners. **B**, Comparing responses of a typical neuron recorded in task engagement (top) and in passive stimulation (bottom). **C**, Comparison of responses of another neuron. **D**, Modality preference of neurons (auditory vs multisensory). **E**, Frequency preference of neurons (high vs low). **F**, **G**, The mean PSTHs of different cued trials across neurons. The conventions used are the same as in Figures 2 and 3.

task engagement, a lower proportion of neurons displayed frequency preference in both auditory (34%, 172/501) and multisensory (30%, 151/501) trials when rats were anesthetized, and cross-modal interaction could not increase the incidence of neurons favoring the high tone (auditory condition, 24%, 120/501; multisensory condition, 20%, 101/501). Again, this result suggests that the process of cross-modal interaction could be modified to meet task requirements in task engagement.

Visual modulation of response of A1 neurons in a free-choice task

To determine whether the behavioral choice itself could induce the differentially integrative model, we recorded the response of A1 neurons in a no-cue-discrimination free-choice task, where rats ($n = 6$) poked the head into the cue port to trigger a cue (A_{low} , A_{high} , V , VA_{low} or VA_{high}) and then could get water reward at the water port of either side. We gathered the responses of 240 A1 neurons when rats performed this free-choice task. For consistency with the earlier analyses, responses of neurons in high-tone trials moving to the left port were compared with responses in low-tone trials moving to the right port, which was done for multisensory comparison. The result showed that unlike those recorded in the cue-discrimination task, neurons recorded in the free-choice task did not show evidence of possessing a differentially integrative model (Fig. 7A, B, examples, C,D, populations). For the majority of neurons (69%, 165/240), their multisensory response is similar to the

corresponding auditory response regardless of tone frequencies ($p > 0.05$, permutation test; Fig. 7A–E). For those neurons with the auditory response(s) influenced by visual cue, they showed induced inhibitory or facilitatory effects that appear similar in each tone condition (low tone, enhanced, 11%, 27/240; inhibited, 10%, 25/240, $p = 0.77$, χ^2 test; high tone, enhanced, 9%, 21/240; inhibited, 11%, 27/240, $p = 0.31$, χ^2 test; Fig. 7E). Thus, across the populations, the mean multisensory response is equal to the auditory (Fig. 7C,D). The mean MIs for both tone conditions were similar (low tone vs high tone, 0.001 ± 0.184 vs -0.001 ± 0.183 ; paired t test, $p = 0.868$). This result indicates that a consequence-free behavioral choice itself fails to induce the differentially integrative model seen earlier, and associative learning between cues and behavioral choices might be necessary to do so.

In addition, we did not see neurons display tone biasing as strongly as their counterparts did in the cue-discrimination task (auditory, preferring A_{low} , 15%, 37/240; preferring A_{high} , 22%, 52/240, χ^2 test, $p = 0.08$; multisensory, preferring VA_{low} , 16%, 38/240; preferring VA_{high} , 18%, 44/240, χ^2 test, $p = 0.47$; Fig. 6F). This result, combined with those given above, reveals that the differential auditory-visual interaction in auditory cortex likely reflects the requirement of the given task. When stimulus discrimination is not required, the cross-modal interaction exhibits no choice-related difference. When demanded by a task, an appropriate interaction is developed to serve sensory discrimination.

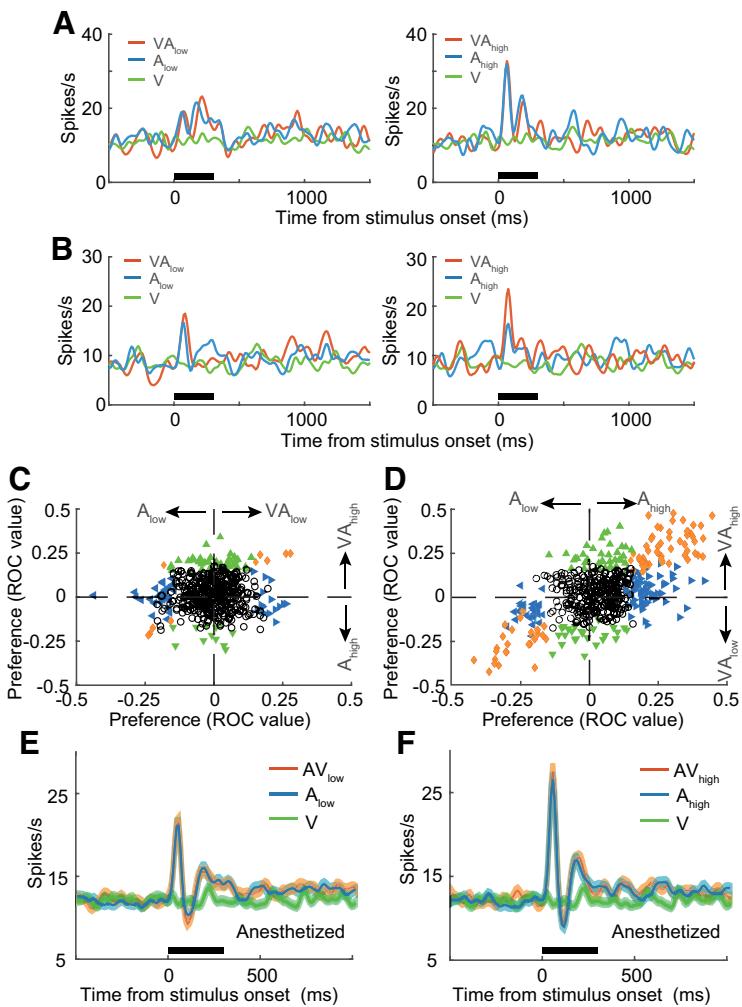


Figure 6. Visual modulation of A1 auditory response in anesthetized rats being well trained to perform the task. **A**, **B**, Rasters and PSTHs showed responses of two neurons. **C**, Modality preference of neurons (auditory vs multisensory). **D**, Frequency preference of neurons (high vs low). **E**, **F**, The mean PSTHs of different cued trials across neurons. The conventions used are the same as in Figures 2 and 3.

Discussion

As mentioned in the introduction, several previous studies have reported that sensory cortices process more than one sensory modality, challenging the long-held view that they process only one. However, it is still unclear how cross-modal interaction occurs in these areas and whether they contribute to perceptual judgments. This study investigated the visual modulation of auditory cortical activities of neurons when rats performed a two-alternative forced-choice task. We found that a choice-related differentially integrative model provides a good description of how auditory cortical neurons integrate a visual cue into an auditory discrimination task to guide behavioral choice. When rats performed the task, we observed that a task-irrelevant visual cue typically could enhance the response of auditory cortical neurons specifically to the tone sound guiding the contralateral choice. This differentially integrative model facilitated sound discrimination and shortened reaction time. We believe this new integrative model provides a unique explanation for how cross-modal interaction helps facilitate sensory encoding and perceptual discrimination, which extends our understanding of multisensory integration in the brain.

Cross-modal influences on auditory cortex have been reported by a host of imaging studies that provided strong

evidence that auditory activity can be modulated by visual and/or somatosensory stimulation (Calvert et al., 1997; Foxe et al., 2002; Kayser et al., 2005; Schürmann et al., 2006). For instance, an fMRI study found that checkerboards activated primary auditory cortex and noise bursts activated primary visual cortex, and when presented together, these stimuli shortened the latency of the hemodynamic BOLD response in each area, indicating multisensory facilitation (Martuzzi et al., 2007). Also, a host of electrophysiological studies have shown that cross-modal responses are present in auditory cortices (Bizley et al., 2007; Bizley and King, 2008; Meredith et al., 2020; Lohse et al., 2021; Opoku-Baah et al., 2021). For instance, visual stimuli and touch are found to modulate the firing of neurons in auditory cortex (Kayser et al., 2008). When the luminance of a visual stimulus is temporally coherent with the amplitude fluctuations of one sound, the representation of that sound is enhanced in the auditory cortex (Atilgan et al., 2018). However, most of these electrophysiological studies regarding cross-modal representation in sensory cortices were done in anesthetized or alert animals in a passive state (Driver and Noesselt, 2008; Xu et al., 2014; Meijer et al., 2017), leaving the issue of cross-modal interaction in early cortical sensory areas unclear. The present study shows that task engagement can combine more neurons into cross-modal interaction in a region not long ago considered pristine for a single sense. This finding is consistent with claims that presumptively unimodal areas of the neocortex are fundamentally multisensory (Ghazanfar and Schroeder, 2006).

Most previous studies showed that stimulus properties such as spatial location, intensity, direction of motion, and temporal relationships between different stimuli influence the

process of multisensory integration and the physiological salience of external events (Stein et al., 2014; Wallace et al., 2020). Our present study shows that auditory cortical neurons apply a differentially integrative model to process auditory-visual interaction to support the task goal of tone-frequency discrimination, despite all cue combinations being spatiotemporally congruent. In this experiment, this novel integrative model helps discriminate multisensory cues. We speculate that this model might be a generalized integrative principle in perceptual tasks as many tasks involve multisensory cue discrimination. To our knowledge, no research articles previously showed a similar integrative model at the neuronal level. However, our result is in line with more recent experiments in awake preparations that contend that the perceptions of objects (or events) are influenced by many factors such as corresponding values, reward, memory, behavioral state, and decision-making (Miller and Cohen, 2001; Gourley and Taylor, 2016; Makino et al., 2016; Gold and Stocker, 2017; Han et al., 2021).

Our data showed that after rats had learned a sound frequency discrimination task well, A1 neurons exhibited a strong frequency preference in task engagement, consistent with previous studies (Znamenskiy and Zador, 2013; Xiong et al., 2015;

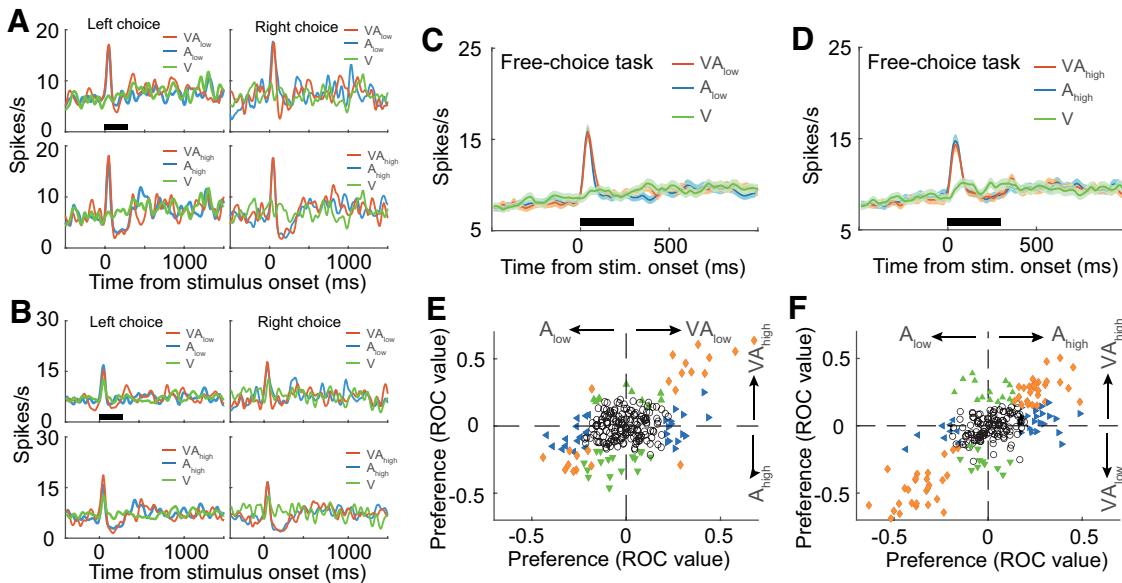


Figure 7. Activities and cue preferences of auditory cortical neurons in a free-choice task. **A, B**, PSTHs show responses of two exemplar neurons of different cue trials for different behavioral choices (left, contralateral choice; right, ipsilateral choice). **C, D**, The mean PSTHs of different cue trials across neurons. **E**, Modality preference of neurons (auditory vs multisensory). **F**, Frequency preference of neurons (high vs low). The conventions used are the same as in Figures 2 and 3.

Xin et al., 2019). Auditory cortex in rodents receives several top-down projections from association cortices such as posterior parietal cortex (PPC; Zingg et al., 2014; Zhong et al., 2019), orbitofrontal cortex (Winkowski et al., 2018; Sharma and Bandyopadhyay, 2020), and medial prefrontal cortex (Gao et al., 2022). These connected association cortices can both process multiple sensory inputs (Raposo et al., 2014; Sharma and Bandyopadhyay, 2020; Zheng et al., 2021) and give strong decision-related responses (Raposo et al., 2014; Lyamzin and Benucci, 2019; Reinert et al., 2021). Top-down projections from high-order cortices can powerfully influence sensory processing in sensory cortices. In mice, neurons of the cingulate region of the frontal cortex enhanced responses of visual cortical neurons and improved visual discrimination (Zhang et al., 2014). Similarly, PPC projections to auditory cortex were shown to play a critical role in the auditory categorical decisions made on new sensory stimuli (Zhong et al., 2019). In addition, associative learning enhances the relative impact of top-down processing in the visual cortex (Makino and Komiyama, 2015). In the present study, rats were required to perform an auditory detection task and make a behavioral choice. We believe that top-down projections similarly regulate the superior sound selectivity of A1 neurons and the choice-related cross-modal interaction that we have observed here.

As is well known, the cortex projects to the dorsal striatum topographically to regulate behavior (Friedman et al., 2015; Hintiryan et al., 2016), and each part of the striatum precisely mirrors activity in topographically associated cortical regions (Peters et al., 2021). Learning has the effect of potentiating corticostratial synapses from A1 neurons representing the frequencies denoting the contralateral choice (Xiong et al., 2015). We believe that the signal of tone selectivity observed in A1 could be being forwarded to the striatum to trigger a specific choice, which might explain our finding that most neurons preferred the sound indicating the contralateral choice. Also, it helps explain why nonspecific activation of the auditory cortex resulted in predominantly contralateral biases after associative learning (Znamenskiy and Zador, 2013).

Compared with passive listening, cue-evoked responses in task engagement were attenuated in amplitude, which is consistent with previous studies (Otazu et al., 2009; Schneider et al., 2014; Zhou et al., 2014; Kuchibhotla et al., 2017). The possible underlying mechanism might be attributed to strengthened motor-related activation of auditory cortical inhibitory neurons (Schneider et al., 2014; Zhou et al., 2014; Kuchibhotla et al., 2017) or a decreased signal-to-noise ratio in the thalamocortical pathway (Otazu et al., 2009).

In normal anesthetized rats, visual cue frequently induces an inhibitory effect on auditory cortex, and no sound-frequency-dependent auditory-visual interaction is observed (Wallace et al., 2004; Xu et al., 2014; Han et al., 2021). In the present study, we found that a subset of neurons still kept a differentially integrative model in anesthesia, suggesting that perceptual training leaves a plastic change in the auditory cortex. Our result here is in line with an increasing number of studies showing that multisensory perceptual learning can effectively drive plastic change in both sensory and association cortices (Shams and Seitz, 2008; Proulx et al., 2014), as well as in subcortical areas (Yu et al., 2010; Xu et al., 2015). Previous studies have shown that multisensory associative learning can enhance cross-modal representation in sensory cortices (Vincis and Fontanini, 2016; Knöpfel et al., 2019; Han et al., 2021). In the present study, although rats experienced a lot of audiovisual stimuli in task engagement, we did not see an increase in visual representation in auditory cortex. The reason, we believe, is that the visual cue was not a necessary factor for rats to perform this task. As an efficiency, rats could just focus their attention on the auditory modality. An intriguing question remaining for us is whether auditory cortex can deeply involve visual presentation and discrimination if visual cues (like the auditory here) become necessary for performing multisensory cue discrimination tasks. This idea will be explored by our group in future studies.

In this study, although irrelevant to this task, visual cue hastened the process of frequency discrimination. This result is consistent with previous studies that show task-irrelevant visual information can enhance perceptual outcomes on a variety of

low-level auditory tasks, including, but not limited to, auditory detection (Lovelace et al., 2003), loudness perception (Odgaard et al., 2004); spatial localization (Bolognini et al., 2007), and frequency discrimination (Thorne and Debener, 2008). For example, an auxiliary light could enhance the accuracy of localizing a near-threshold auditory target (Bolognini et al., 2007). Salient sounds, regardless of task relevance, can activate the human visual cortex automatically (McDonald et al., 2013) and improve the discriminative processing of colocalized visual events (Feng et al., 2014). In the sound discriminative task used here, we believe the combined auditory-visual cue induced more attention, which might partly explain the enhancement of perceptual discrimination. Attentional mechanisms are believed to compute approximate solutions to the multisensory binding problem in naturalistic environments when complex signals arise from myriad causes (Noppeney, 2021).

References

- Atilgan H, Town SM, Wood KC, Jones GP, Maddox RK, Lee AKC, Bizley JK (2018) Integration of visual information in auditory cortex promotes auditory scene analysis through multisensory binding. *Neuron* 97:640–655.e4.
- Bagur S, Averseng M, Elgueda D, David S, Fritz J, Yin P, Shamma S, Boubenec Y, Ostoicic S (2018) Go/no-go task engagement enhances population representation of target stimuli in primary auditory cortex. *Nat Commun* 9:2529.
- Bizley JK, King AJ (2008) Visual-auditory spatial processing in auditory cortical neurons. *Brain Res* 1242:24–36.
- Bizley JK, Nodal FR, Bajo VM, Nelken I, King AJ (2007) Physiological and anatomical evidence for multisensory interactions in auditory cortex. *Cereb Cortex* 17:2172–2189.
- Bizley JK, Walker KM, Nodal FR, King AJ, Schnupp JW (2013) Auditory cortex represents both pitch judgments and the corresponding acoustic cues. *Curr Biol* 23:620–625.
- Bolognini N, Leo F, Passamonti C, Stein BE, Lådavas E (2007) Multisensory-mediated auditory localization. *Perception* 36:1477–1485.
- Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci* 12:4745–4765.
- Brosch M, Selezneva E, Scheich H (2005) Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *J Neurosci* 25:6797–6806.
- Calvert GA, Bullmore ET, Brammer MJ, Campbell R, Williams SC, McGuire PK, Woodruff PW, Iversen SD, David AS (1997) Activation of auditory cortex during silent lipreading. *Science* 276:593–596.
- Campi KL, Bales KL, Grunewald R, Krubitzer L (2010) Connections of auditory and visual cortex in the prairie vole (*Microtus ochrogaster*): evidence for multisensory processing in primary sensory areas. *Cereb Cortex* 20:89–108.
- Cappe C, Barone P (2005) Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *Eur J Neurosci* 22:2886–2902.
- David SV, Fritz JB, Shamma SA (2012) Task reward structure shapes rapid receptive field plasticity in auditory cortex. *Proc Natl Acad Sci U S A* 109:2144–2149.
- Deneux T, Harrell ER, Kempf A, Ceballo S, Filipchuk A, Bathellier B (2019) Context-dependent signaling of coincident auditory and visual events in primary visual cortex. *Elife* 8:e44006.
- Driver J, Noesselt T (2008) Multisensory interplay reveals crossmodal influences on “sensory-specific” brain regions, neural responses, and judgments. *Neuron* 57:11–23.
- Elgueda D, Duque D, Radtke-Schuller S, Yin P, David SV, Shamma SA, Fritz JB (2019) State-dependent encoding of sound and behavioral meaning in a tertiary region of the ferret auditory cortex. *Nat Neurosci* 22:447–459.
- Falchier A, Clavagnier S, Barone P, Kennedy H (2002) Anatomical evidence of multimodal integration in primate striate cortex. *J Neurosci* 22:5749–5759.
- Feng W, Störmer VS, Martinez A, McDonald JJ, Hillyard SA (2014) Sounds activate visual cortex and improve visual discrimination. *J Neurosci* 34:9817–9824.
- Foxe JJ, Wylie GR, Martinez A, Schroeder CE, Javitt DC, Guilfoyle D, Ritter W, Murray MM (2002) Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J Neurophysiol* 88:540–543.
- Friedman A, Homma D, Gibb LG, Amemori K, Rubin SJ, Hood AS, Riad MH, Graybiel AM (2015) A corticostratial path targeting striosomes controls decision-making under conflict. *Cell* 161:1320–1333.
- Fritz J, Shamma S, Elhilali M, Klein D (2003) Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat Neurosci* 6:1216–1223.
- Fu KM, Johnston TA, Shah AS, Arnold L, Smiley J, Hackett TA, Garraghty PE, Schroeder CE (2003) Auditory cortical neurons respond to somatosensory stimulation. *J Neurosci* 23:7510–7515.
- Gao L, et al. (2022) Single-neuron projectome of mouse prefrontal cortex. *Nat Neurosci* 25:515–529.
- Ghazanfar AA, Schroeder CE (2006) Is neocortex essentially multisensory? *Trends Cogn Sci* 10:278–285.
- Ghazanfar AA, Maier JX, Hoffman KL, Logothetis NK (2005) Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J Neurosci* 25:5004–5012.
- Gold JI, Stocker AA (2017) Visual decision-making in an uncertain and dynamic world. *Annu Rev Vis Sci* 3:227–250.
- Gourley SL, Taylor JR (2016) Going and stopping: dichotomies in behavioral control by the prefrontal cortex. *Nat Neurosci* 19:656–664.
- Han X, Xu J, Chang S, Keniston L, Yu L (2021) Multisensory-guided associative learning enhances multisensory representation in primary auditory cortex. *Cereb Cortex* 32:1040–1054.
- Hintiryan H, Foster NN, Bowman I, Bay M, Song MY, Gou L, Yamashita S, Bienkowski MS, Zingg B, Zhu M, Yang XW, Shih JC, Toga AW, Dong HW (2016) The mouse cortico-striatal projectome. *Nat Neurosci* 19:1100–1114.
- Hirokawa J, Sadakane O, Sakata S, Bosch M, Sakurai Y, Yamamori T (2011) Multisensory information facilitates reaction speed by enlarging activity difference between superior colliculus hemispheres in rats. *PLoS One* 6:e25283.
- Huang Y, Heil P, Brosch M (2019) Associations between sounds and actions in early auditory cortex of nonhuman primates. *Elife* 8:e43281.
- Ibrahim LA, Mesik L, Ji XY, Fang Q, Li HF, Li YT, Zingg B, Zhang LI, Tao HW (2016) Cross-modality sharpening of visual cortical processing through layer-1-mediated inhibition and disinhibition. *Neuron* 89:1031–1045.
- Jaramillo S, Zador AM (2011) The auditory cortex mediates the perceptual effects of acoustic temporal expectation. *Nat Neurosci* 14:246–251.
- Kayser C, Petkov CI, Augath M, Logothetis NK (2005) Integration of touch and sound in auditory cortex. *Neuron* 48:373–384.
- Kayser C, Petkov CI, Logothetis NK (2008) Visual modulation of neurons in auditory cortex. *Cereb Cortex* 18:1560–1574.
- Knöpfel T, Sweeney Y, Radulescu CI, Zabouri N, Doostdar N, Clopath C, Barnes SJ (2019) Audio-visual experience strengthens multisensory assemblies in adult mouse visual cortex. *Nat Commun* 10:5684.
- Kuchibhotla KV, Gill JV, Lindsay GW, Papadoyannis ES, Field RE, Sten TA, Miller KD, Froemke RC (2017) Parallel processing by cortical inhibition enables context-dependent behavior. *Nat Neurosci* 20:62–71.
- Lohse M, Dahmen JC, Bajo VM, King AJ (2021) Subcortical circuits mediate communication between primary sensory cortical areas in mice. *Nat Commun* 12:3916.
- Lovelace CT, Stein BE, Wallace MT (2003) An irrelevant light enhances auditory detection in humans: a psychophysical analysis of multisensory integration in stimulus detection. *Brain Res Cogn Brain Res* 17:447–453.
- Lyamzin D, Benucci A (2019) The mouse posterior parietal cortex: anatomy and functions. *Neurosci Res* 140:14–22.
- Makino H, Komiyama T (2015) Learning enhances the relative impact of top-down processing in the visual cortex. *Nat Neurosci* 18:1116–1122.
- Makino H, Hwang EJ, Hedrick NG, Komiyama T (2016) Circuit mechanisms of sensorimotor learning. *Neuron* 92:705–721.
- Martuzzi R, Murray MM, Michel CM, Thiran JP, Maeder PP, Clarke S, Meuli RA (2007) Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cereb Cortex* 17:1672–1679.

- McDonald JJ, Störmer VS, Martinez A, Feng W, Hillyard SA (2013) Salient sounds activate human visual cortex automatically. *J Neurosci* 33:9194–9201.
- Meijer GT, Montijn JS, Pennartz CMA, Lansink CS (2017) Audiovisual modulation in mouse primary visual cortex depends on cross-modal stimulus configuration and congruity. *J Neurosci* 37:8783–8796.
- Meredith MA, Allman BL (2015) Single-unit analysis of somatosensory processing in the core auditory cortex of hearing ferrets. *Eur J Neurosci* 41:686–698.
- Meredith MA, Keniston LP, Prickett EH, Bajwa M, Cojanu A, Clemo HR, Allman BL (2020) What is a multisensory cortex? A laminar, connectional, and functional study of a ferret temporal cortical multisensory area. *J Comp Neurol* 528:1864–1882.
- Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24:167–202.
- Niell CM, Stryker MP (2010) Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron* 65:472–479.
- Niwa M, Johnson JS, O'Connor KN, Sutter ML (2012) Activity related to perceptual judgment and action in primary auditory cortex. *J Neurosci* 32:3193–3210.
- Noppeney U (2021) Perceptual inference, learning, and attention in a multisensory world. *Annu Rev Neurosci* 44:449–473.
- Odgaard EC, Arieh Y, Marks LE (2004) Brighter noise: sensory enhancement of perceived loudness by concurrent visual stimulation. *Cogn Affect Behav Neurosci* 4:127–132.
- Olcese U, Iurilli G, Medini P (2013) Cellular and synaptic architecture of multisensory integration in the mouse neocortex. *Neuron* 79:579–593.
- Opoku-Baah C, Schoenhaut AM, Vassall SG, Tovar DA, Ramachandran R, Wallace MT (2021) Visual influences on auditory behavioral, neural, and perceptual processes: a review. *J Assoc Res Otolaryngol* 22:365–386.
- Otazu GH, Tai LH, Yang Y, Zador AM (2009) Engaging in an auditory task suppresses responses in auditory cortex. *Nat Neurosci* 12:646–654.
- Perrodin C, Kayser C, Logothetis NK, Petkov CI (2015) Natural asynchronies in audiovisual communication signals regulate neuronal multisensory interactions in voice-sensitive cortex. *Proc Natl Acad Sci U S A* 112:273–278.
- Peters AJ, Fabre JM, Steinmetz NA, Harris KD, Carandini M (2021) Striatal activity topographically reflects cortical activity. *Nature* 591:420–425.
- Proulx MJ, Brown DJ, Pasqualotto A, Meijer P (2014) Multisensory perceptual learning and sensory substitution. *Neurosci Biobehav Rev* 41:16–25.
- Raposo D, Kaufman MT, Churchland AK (2014) A category-free neural population supports evolving demands during decision-making. *Nat Neurosci* 17:1784–1792.
- Reinert S, Hübener M, Bonhoeffer T, Goltstein PM (2021) Mouse prefrontal cortex represents learned rules for categorization. *Nature* 593:411–417.
- Schneider DM, Nelson A, Mooney R (2014) A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature* 513:189–194.
- Schürmann M, Caetano G, Hlushchuk Y, Jousmaki V, Hari R (2006) Touch activates human auditory cortex. *Neuroimage* 30:1325–1331.
- Shams L, Seitz AR (2008) Benefits of multisensory learning. *Trends Cogn Sci* 12:411–417.
- Sharma S, Bandyopadhyay S (2020) Differential rapid plasticity in auditory and visual responses in the primarily multisensory orbitofrontal cortex. *eNeuro* 7:ENEURO.0061-20.2020.
- Stehberg J, Dang PT, Frostig RD (2014) Unimodal primary sensory cortices are directly connected by long-range horizontal projections in the rat sensory cortex. *Front Neuroanat* 8:93.
- Stein BE, Stanford TR, Rowland BA (2014) Development of multisensory integration from the perspective of the individual neuron. *Nat Rev Neurosci* 15:520–535.
- Thorne JD, Debener S (2008) Irrelevant visual stimuli improve auditory task performance. *Neuroreport* 19:553–557.
- Vasconcelos N, Pantoja J, Belchior H, Caixeta FV, Faber J, Freire MA, Cota VR, Anibal de Macedo E, Laplagne DA, Gomes HM, Ribeiro S (2011) Cross-modal responses in the primary visual cortex encode complex objects and correlate with tactile discrimination. *Proc Natl Acad Sci U S A* 108:15408–15413.
- Vincis R, Fontanini A (2016) Associative learning changes cross-modal representations in the gustatory cortex. *Elife* 5:e16420.
- Wallace MT, Ramachandran R, Stein BE (2004) A revised view of sensory cortical parcellation. *Proc Natl Acad Sci U S A* 101:2167–2172.
- Wallace MT, Woynaroski TG, Stevenson RA (2020) Multisensory integration as a window into orderly and disrupted cognition and communication. *Annu Rev Psychol* 71:193–219. Vol 71
- Winkowski DE, Nagode DA, Donaldson KJ, Yin P, Shamma SA, Fritz JB, Kanold PO (2018) Orbitofrontal cortex neurons respond to sound and activate primary auditory cortex neurons. *Cereb Cortex* 28:868–879.
- Xin Y, Zhong L, Zhang Y, Zhou T, Pan J, Xu NL (2019) Sensory-to-category transformation via dynamic reorganization of ensemble structures in mouse auditory cortex. *Neuron* 103:909–921.e6.
- Xiong Q, Znamenskiy P, Zador AM (2015) Selective corticostriatal plasticity during acquisition of an auditory discrimination task. *Nature* 521:348–351.
- Xu J, Sun X, Zhou X, Zhang J, Yu L (2014) The cortical distribution of multisensory neurons was modulated by multisensory experience. *Neuroscience* 272:1–9.
- Xu J, Yu L, Stanford TR, Rowland BA, Stein BE (2015) What does a neuron learn from multisensory experience? *J Neurophysiol* 113:883–889.
- Xu J, Bi T, Wu J, Meng F, Wang K, Hu J, Han X, Zhang J, Zhou X, Keniston L, Yu L (2018) Spatial receptive field shift by preceding cross-modal stimulation in the cat superior colliculus. *J Physiol* 596:5033–5050.
- Yu L, Rowland BA, Stein BE (2010) Initiating the development of multisensory integration by manipulating sensory experience. *J Neurosci* 30:4904–4913.
- Zhang S, Xu M, Kamigaki T, Hoang Do JP, Chang WC, Jenvay S, Miyamichi K, Luo L, Dan Y (2014) Selective attention. Long-range and local circuits for top-down modulation of visual cortex processing. *Science* 345:660–665.
- Zheng M, Xu J, Keniston L, Wu J, Chang S, Yu L (2021) Choice-dependent cross-modal interaction in the medial prefrontal cortex of rats. *Mol Brain* 14:13.
- Zhong L, Zhang Y, Duan CA, Deng J, Pan J, Xu NL (2019) Causal contributions of parietal cortex to perceptual decision-making during stimulus categorization. *Nat Neurosci* 22:963–973.
- Zhou M, Liang F, Xiong XR, Li L, Li H, Xiao Z, Tao HW, Zhang LI (2014) Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex. *Nat Neurosci* 17:841–850.
- Zingg B, Hintiryan H, Gou L, Song MY, Bay M, Bienkowski MS, Foster NN, Yamashita S, Bowman I, Toga AW, Dong HW (2014) Neural networks of the mouse neocortex. *Cell* 156:1096–1111.
- Znamenskiy P, Zador AM (2013) Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature* 497:482–485.