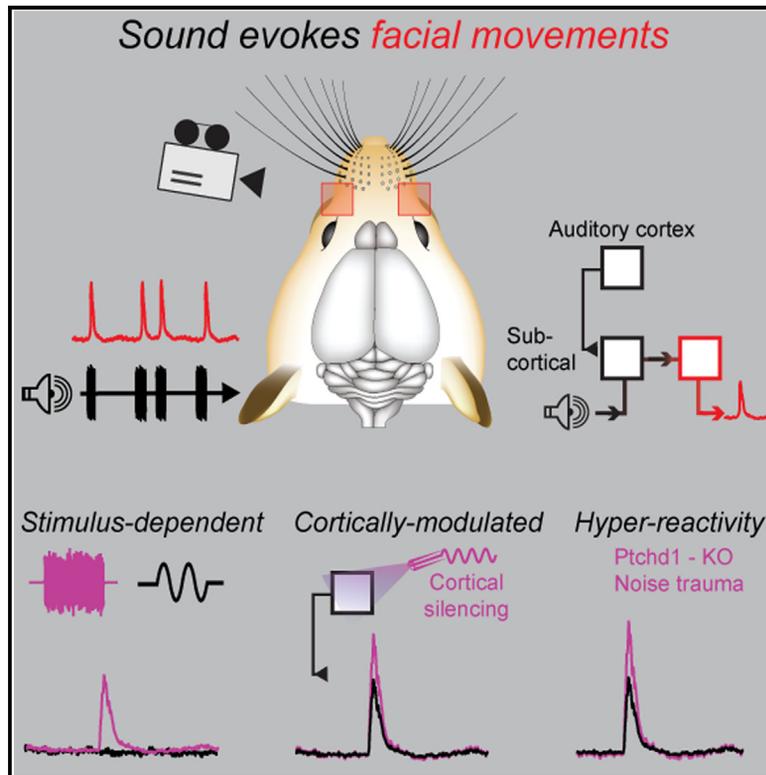


# Current Biology

## Sound elicits stereotyped facial movements that provide a sensitive index of hearing abilities in mice

### Graphical abstract



### Authors

Kameron K. Clayton, Kamryn S. Stecyk, Anna A. Guo, Anna R. Chambers, Ke Chen, Kenneth E. Hancock, Daniel B. Polley

### Correspondence

kameron\_clayton@meei.harvard.edu

### In brief

Clayton et al. show that video recordings of the face provide a window into complex sound processing in mice with normal hearing, with sensorineural hearing loss, and with mutations in autism risk genes. Facial twitches rapidly synchronize to broadband sounds across audible levels, indicating that many auditory stimuli are not strictly unisensory.

### Highlights

- A wide range of broadband sounds elicit rapid and stereotyped facial twitches
- Complex sounds like speech phonemes in noise can be decoded from facial twitches
- Sound-evoked movements are subcortically mediated but corticofugally modulated
- Facial movements are hypersensitive to sound in autism gene KO and acoustic trauma

Article

# Sound elicits stereotyped facial movements that provide a sensitive index of hearing abilities in mice

Kameron K. Clayton,<sup>1,2,4,\*</sup> Kamryn S. Stecyk,<sup>1</sup> Anna A. Guo,<sup>1</sup> Anna R. Chambers,<sup>1,2</sup> Ke Chen,<sup>1,2,3</sup> Kenneth E. Hancock,<sup>1,2</sup> and Daniel B. Polley<sup>1,2</sup>

<sup>1</sup>Eaton-Peabody Laboratories, Massachusetts Eye and Ear, Boston, MA 02114, USA

<sup>2</sup>Department of Otolaryngology–Head and Neck Surgery, Harvard Medical School, Boston, MA 02114, USA

<sup>3</sup>Present address: McGovern Institute for Brain Research, MIT, Cambridge, MA 02139, USA

<sup>4</sup>Lead contact

\*Correspondence: [kameron\\_clayton@meei.harvard.edu](mailto:kameron_clayton@meei.harvard.edu)

<https://doi.org/10.1016/j.cub.2024.02.057>

## SUMMARY

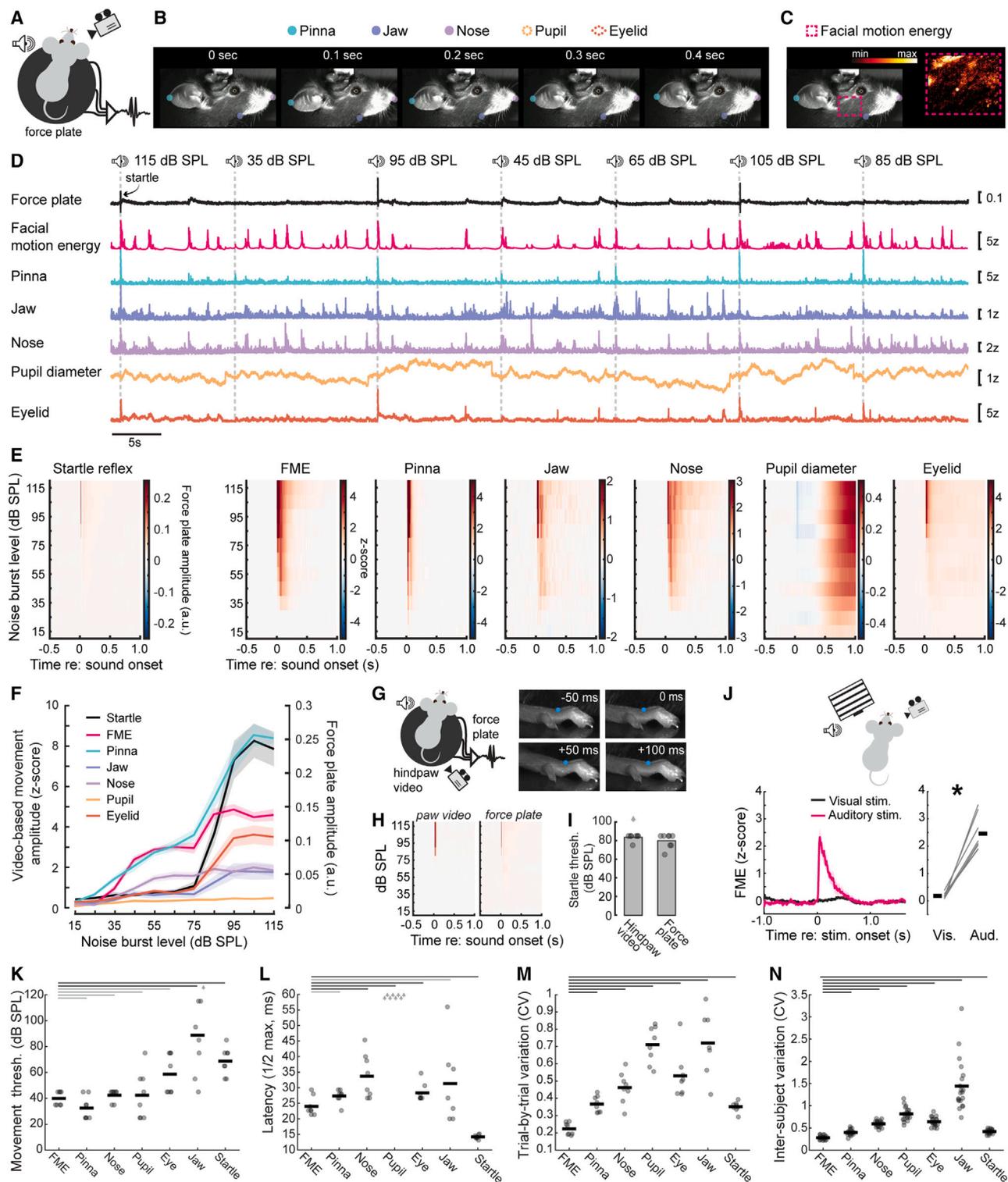
Sound elicits rapid movements of muscles in the face, ears, and eyes that protect the body from injury and trigger brain-wide internal state changes. Here, we performed quantitative facial videography from mice resting atop a piezoelectric force plate and observed that broadband sounds elicited rapid and stereotyped facial twitches. Facial motion energy (FME) adjacent to the whisker array was 30 dB more sensitive than the acoustic startle reflex and offered greater inter-trial and inter-animal reliability than sound-evoked pupil dilations or movement of other facial and body regions. FME tracked the low-frequency envelope of broadband sounds, providing a means to study behavioral discrimination of complex auditory stimuli, such as speech phonemes in noise. Approximately 25% of layer 5–6 units in the auditory cortex (ACTx) exhibited firing rate changes during facial movements. However, FME facilitation during ACTx photoinhibition indicated that sound-evoked facial movements were mediated by a midbrain pathway and modulated by descending corticofugal input. FME and auditory brainstem response (ABR) thresholds were closely aligned after noise-induced sensorineural hearing loss, yet FME growth slopes were disproportionately steep at spared frequencies, reflecting a central plasticity that matched commensurate changes in ABR wave 4. Sound-evoked facial movements were also hypersensitive in *Ptchd1* knockout mice, highlighting the use of FME for identifying sensory hyper-reactivity phenotypes after adult-onset hyperacusis and inherited deficiencies in autism risk genes. These findings present a sensitive and integrative measure of hearing while also highlighting that even low-intensity broadband sounds can elicit a complex mixture of auditory, motor, and reafferent somatosensory neural activity.

## INTRODUCTION

Less than 1 s after intense sound reaches the ear, the pupil dilates, postural muscles contract, the pinna twitches, eardrum tension increases, and molecular motors in cochlear outer hair cells are disengaged to attenuate cochlear amplification. Each of these rapid audiomotor transformations is mediated by independent neural reflex pathways that protect the inner ear from acoustic injury, initiate defensive behaviors, and orient the head and body for further analysis of the sound source. In principle, assaying these involuntary behaviors with startle reflex audiometry,<sup>1</sup> pre-pulse inhibition of startle,<sup>2,3</sup> middle ear muscle reflex audiometry,<sup>4</sup> olivocochlear reflex audiometry,<sup>5,6</sup> or pupillary dilation response audiometry<sup>7,8</sup> provides a useful middle ground to study hearing in laboratory animals, offering higher throughput than operant behaviors while providing the integrative measure of behavioral registration in awake animals that is lacking with physiological proxies for hearing, such as the auditory brainstem response.

The primary challenge with most involuntary behavioral assays is that they are mediated by different cell types and neural pathways than the central auditory neuroaxis that provides the basis for conscious sound awareness. Reflexes that protect the animal and ear from injury are often insensitive to low and moderate sound levels that are well within the audible range and generally do not reflect the involvement of neural circuits beyond the brainstem, limiting their broader use as a behavioral measure of general hearing ability.<sup>2,9,10</sup> Interestingly, in analyzing high-resolution video of the face during a sound presentation in mice, several recent studies observed un instructed movements that were synchronized to the onsets of discrete sounds.<sup>11–13</sup> Apart from remarking on their occurrence, these studies did not delve into their acoustic feature sensitivity or characterize the neural pathways that might transform sound into facial movements.

Here, we build on these reports to show that sound-evoked movements from a region of the face just caudal to the vibrissae array are 1,000 times more sensitive (30 dB) to sound than the startle reflex. We show that facial movements track the temporal



**Figure 1. Facial movements provide a sensitive and specific behavioral index of hearing in mice**

(A) Schematic depicts mouse, camera, speaker, and summed output of three piezoelectric transducers attached to the force plate. The multi-peaked startle waveform is copied from the first 115 dB SPL trial shown in (D).

(B) Video frames from a single trial depict changes in the pinna, jaw, nose, pupil, and eyelid positions determined by DeepLabCut.

(C) Motion energy calculated from a region of interest positioned caudal to the vibrissa array.

(D) Movement amplitudes for 7 consecutive trials in a representative mouse. y axis scales presented at right.

(E) Mean peristimulus force plate and facial movement responses from 8 mice.

(legend continued on next page)

envelope of sound, permitting single-trial decoding of complex sounds, such as English speech tokens presented in background noise. Single-unit recordings and optogenetic manipulations suggest that sound-evoked facial movements are mediated by a midbrain pathway but modulated by descending corticofugal projections. Finally, we show that facial movements capture both inherited and acquired auditory hypersensitivity in mice with mutations in autism risk genes and noise-induced cochlear sensorineural damage, respectively. In sum, our data suggest that a relatively simple approach of making video recordings of the face provides a new behavioral vantage point to study neural and behavioral processing of complex sounds.

## RESULTS

### Facial movements provide a sensitive index of hearing

We measured involuntary behavioral reactions to sound by placing head-fixed unanesthetized mice atop a piezoelectric force transduction plate while acquiring a high-resolution video of the face. The output of the force plate provided an index of general body movements as well as large-amplitude multiphasic events corresponding to the startle reflex (Figure 1A).<sup>14</sup> Movement of the pinna, jaw, eyelid, nose, and pupil were quantified with DeepLabCut, a video analysis method for markerless tracking of body movements based on deep neural networks (N = 8 mice; Figure 1B).<sup>15</sup> We also calculated facial motion energy (FME) from a region of interest caudal to the whisker array (Figure 1C), permitting simultaneous acquisition and comparison of sound-evoked movements across measurement modalities (Figure 1D).

We observed that broadband noise bursts elicited short-latency facial twitches and longer latency pupil dilations across a far wider range of sound levels than the acoustic startle reflex (Figure 1E). To quantify the differences in the amplitude of these movements across the full range of sound levels, Z-scored amplitudes of videographic measures were compared with changes

in force plate amplitude. Acoustic startle reflex amplitudes increased rapidly above 75 dB SPL and then saturated at 95 dB SPL (Figure 1F). Sound-evoked movement of the eyelid and jaw showed the same high-threshold saturating response as the acoustic startle response, while the pupil was less responsive overall than other facial markers. By contrast, FME, as well as the movement of the pinna and nose, grew monotonically across a far wider range of sound levels, suggesting that they may arise from a separate pathway for neural sound processing than the acoustic startle reflex.

To determine whether facial movement and startle reflex differences could be attributed to inherent differences between video and piezoelectric force plate recordings, we performed high-speed videography of the hind paw in a subset of mice but noted a close correspondence between force plate and hind paw movement amplitudes (Figure 1G). These experiments demonstrated consistently high sound-evoked thresholds for the startle reflex, whether acquired via force plate or video analysis of the hind paw (Figures 1H and 1I, statistical reporting provided in the figure legends throughout). Next, we explored whether evoked facial movements were particularly sensitive to auditory stimuli by interleaving blocks of 60 dB SPL noise bursts with high-contrast visual drifting gratings. These experiments revealed the virtual absence of visually evoked facial movements in mice with robust reactions to a moderate-intensity sound (Figure 1J). While this experiment cannot prove the absence of visually evoked facial movements, it demonstrates that an easily detectable broadband visual stimulus used widely in neuroscience experiments does not elicit time-locked facial movements, but a typical auditory broadband stimulus does. A final series of control experiments addressed whether time-locked movements to sound presentation might instead arise from a direct mechanical displacement from the sound pressure wave or a response to somatosensory cues induced by any direct movement from the sound wave. A combination of mathematical calculations, recordings in whisker-trimmed mice, and

(F) Mean  $\pm$  SEM movement amplitudes. Vertical axes on the left and right refer to the videographic movement and startle reflex, respectively.

(G) Left: as per (A), except that high-resolution video recordings (150 frames/s) are made of the hind paw. Right: blue circle indicates DeepLabCut tracking of a point on the hind paw in four frames relative to the onset of a noise burst.

(H) Mean peristimulus force plate and hind paw video responses from 8 mice. Force plate and hind paw pseudocolor scales match the startle reflex and nose videography plots presented above in (E).

(I) Startle thresholds for each mouse (circle) and sample mean. Arrow indicates one outlying value outside of the plotted range. There was no significant difference in startle reflex threshold measured via the piezoelectric force plate or hind paw videography (paired t test,  $p = 0.17$ ).

(J) Left, mean FME amplitude in response to a high-contrast drifting visual grating compared with a 60 dB SPL white noise burst (N = 8). Right, moderate-intensity noise bursts elicited significantly greater FME than visual gratings (paired t test,  $p = 3 \times 10^{-5}$ ).

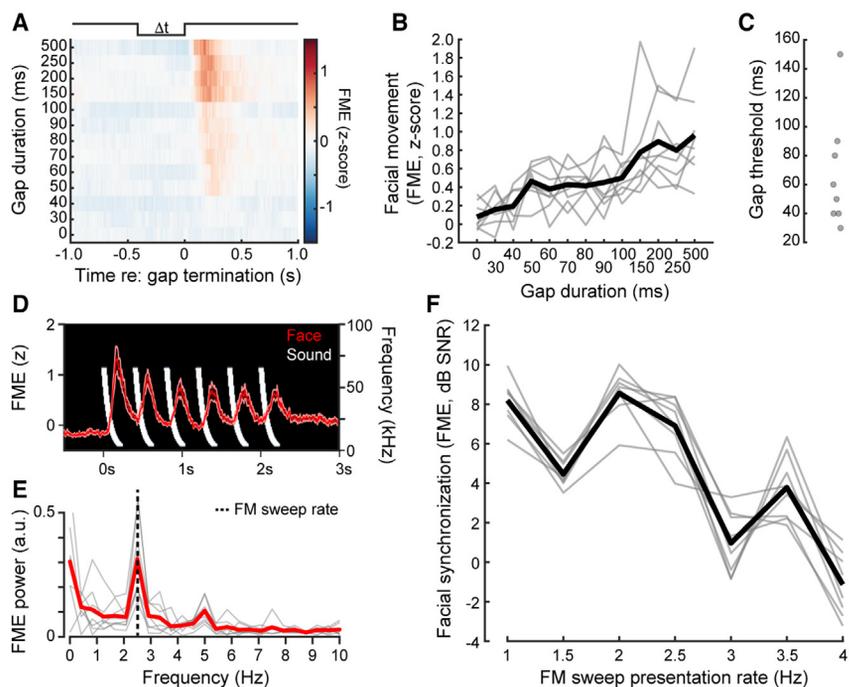
(K) Minimum sound intensity that elicited movement presented for each mouse (circle) and sample mean (horizontal bar). Threshold varied significantly across movement types (one-way repeated measures ANOVA,  $F = 13.45$ ,  $p = 2 \times 10^{-8}$ ) with post-hoc comparisons finding significant differences between FME and jaw ( $p = 0.01$ ) and startle reflex ( $p = 0.0002$ ). For all figures, black and gray horizontal bars denote significant ( $p < 0.05$ ) and non-significant differences, respectively, with FME after Holm-Bonferroni corrections for multiple comparisons.

(L) Sound-evoked movement latencies presented for each mouse (circle) and sample mean (horizontal bar). Response latency varied significantly across movement types (one-way repeated measures ANOVA,  $F = 550.67$ ,  $p = 8 \times 10^{-34}$ ) with post-hoc comparisons finding significant differences between FME and nose ( $p = 0.01$ ), pupil dilation ( $p = 2 \times 10^{-7}$ ), eyelid ( $p = 0.01$ ), and startle reflex ( $p = 0.0003$ ). Gray arrows for pupil denote that values were outside of the y axis range (mean latency = 627 ms).

(M) Trial-to-trial variability measured with the coefficient of variation presented for each mouse (circle) and sample mean (horizontal bar). Trial-to-trial variability was significantly different between movement types (one-way repeated measures ANOVA,  $F = 23.18$ ,  $p = 6 \times 10^{-11}$ ) with post-hoc comparisons finding that FME was significantly less variable than all other movement types ( $p < 0.0004$  for all comparisons).

(N) Inter-subject variability measured with the coefficient of variation across subjects for each trial (circle) and sample mean (horizontal bar). Inter-subject variability was significantly different between movement types (one-way repeated measures ANOVA,  $F = 59.12$ ,  $p = 1 \times 10^{-34}$ ) with post-hoc comparisons finding that FME was significantly less variable than all other movement types ( $p < 0.005$  for all comparisons).

See also Figure S1.



**Figure 2. Facial movements synchronize to slow changes in the sound pressure envelope**

(A) Silent gaps of varying durations were introduced in a constant background of 50 dB SPL white noise. Mean FME before, during, and after the silent gap. (B) FME significantly increased with gap duration (one-way repeated measures ANOVA,  $F = 9.43$ ,  $p = 3 \times 10^{-11}$ ). Individual mice and sample means ( $N = 8$ ) are plotted as thin gray lines and thick black lines, respectively. (C) Gap detection thresholds for each mouse. (D) Spectrogram depicts downward frequency modulated sweeps presented at 2.5 Hz with a 50% duty cycle (white) at 70 dB SPL. Mean  $\pm$  SEM FME amplitude for an example mouse (red) shows a facial twitch elicited by each of the six consecutive FM sweeps. (E) Fourier analysis of FME responses from 8 mice to the FM sweep sequence presented at 2.5 Hz yields a peak at the presentation rate (dashed vertical line). Individual mice and sample means are plotted as thin gray lines and thick red lines, respectively. (F) Facial synchronization was calculated as the power at the stimulus presentation rate relative to the noise floor. Synchronization significantly decreases across higher FM sweep presentation rates (one-way repeated measures ANOVA,  $F = 52.73$ ,  $p = 6 \times 10^{-18}$ ).

recordings on the contralateral side of the head to the sound source provided no support for this possibility (Figures S1A–S1D). Additional hearing loss experiments described below confirm that sound-evoked facial movements reflect the neural encoding of the auditory stimulus.

### FME provides a robust measure of sound-evoked facial movements

Compared with other measures of facial movement, FME appeared to offer a few advantages based on its sensitivity to low sound levels, apparent independence from the startle reflex, relative ease of measurement, and use in prior studies.<sup>11–13</sup> We quantified features of FME relative to other involuntary movements to address this point more rigorously. We found that FME thresholds were approximately 40 dB SPL and not significantly different than movement thresholds for the pinna, nose, pupil, or eyelid, but significantly more sensitive than the startle reflex or jaw movements (Figure 1K). Facial movements exhibited short latencies overall, occurring less than 50 ms following sound onset (Figure 1L). The startle reflex was the fastest of all ( $14.2 \pm 0.2$  ms), while sound-evoked pupil dilations were over an order of magnitude slower than other movements, with an average onset latency of  $627.3 \pm 22.8$  ms. Although FME, pinna, and nose movements showed comparable thresholds and latencies, FME proved the most robust measure of sound-evoked facial movements, as evidenced by significantly lower trial-by-trial variability (Figure 1M) and inter-subject variability (Figure 1N).

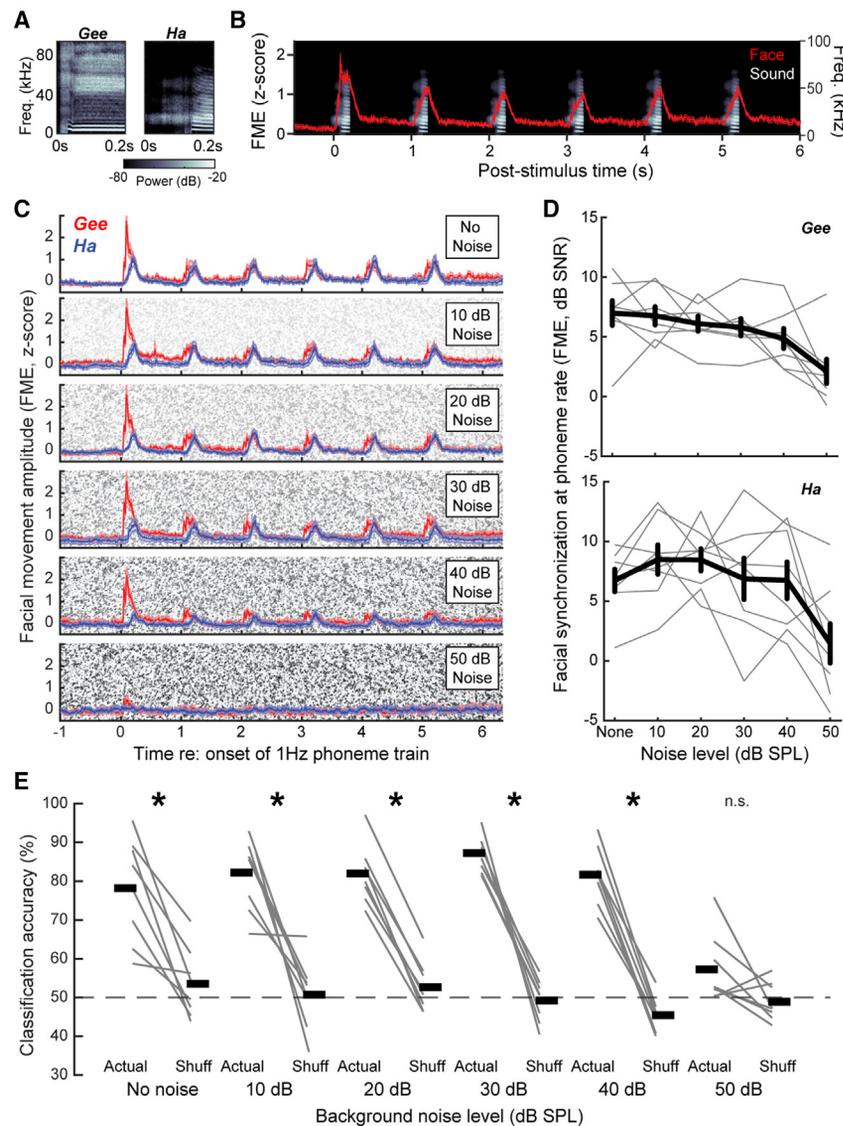
One potential concern relates to FME's specificity, as sound-evoked facial movements are intermingled with spontaneous facial movements (Figure 1D). However, calculating FME in units of Z score units accounts for the degree of spontaneous movement in a given recording session, as confirmed by the

correspondence between the growth functions of FME and the d-prime sensitivity statistic (Figures S1E and S1F). Furthermore, motion energy can be calculated from a region of interest (ROI) positioned anywhere on the body, raising the question of whether placing the ROI caudal to the whisker array is the optimal choice. We found that motion energy saturated at 45 dB SPL for a whisker pad ROI, while an ROI centered on the pinna exhibited non-monotonic growth for sound levels greater than 75 dB, indicating startle reflex involvement (Figures S1G and S1H). Because of these non-linearities and because the ROI caudal to the whisker array is less sensitive to variations in lighting conditions, we rely on motion energy measured caudal to the whisker array for all subsequent analyses.

### Facial movements track the temporal envelope of broadband sounds

Most commonly used reflexive and physiological markers of sound processing rely on discrete bursts of sounds to elicit responses of varying magnitude. However, encoding natural sounds is critically dependent on tracking fluctuations in the sound pressure envelope over time.<sup>16</sup> Hence, an involuntary behavioral readout of sound processing that indexed synchronization to the sound pressure envelope could provide insights into the neural encoding of more complex sounds not possible with other reflexive or voluntary behaviors.

To address the possibility that sound-evoked facial movements could provide this insight, we first introduced silent gaps of varying duration in continuous broadband noise. We found that facial movements were elicited at the offset of silent gaps (Figure 2A), where the magnitude of FME grew monotonically with gap duration (Figure 2B), with FME gap response thresholds of approximately 55 ms across eight mice (Figure 2C). Next, we noted that facial movements were entrained



**Figure 3. Decoding phonemes in background noise via facial movements**

(A) Spectrograms of two English speech tokens digitally resynthesized to span the mouse hearing range without distorting the spectrotemporal envelope of the source signal.

(B) Spectrogram plots six presentations of the phoneme, *Gee*, presented at 1 Hz (grayscale, right vertical axis). Mean  $\pm$  SEM FME from a representative mouse elicited by each speech token (red, left vertical axis).

(C) Mean  $\pm$  SEM FME for *Gee* and *Ha* presented at 70 dB SPL without background noise and five levels of increasingly intense background noise (N = 8).

(D) FME synchronization to the speech token presentation rate decreased significantly, but equivalently, for both phonemes across increasing levels of background noise (2-way repeated measures ANOVA, main effect for noise level [F = 12.36, p = 2  $\times$  10<sup>-8</sup>], main effect for phoneme [F = 0.78, p = 0.39]).

(E) Single-trial speech token classification accuracy with actual and shuffled (shuff) assignment of stimulus identity. Chance classification = 50%. Classification accuracy was significantly greater for actual than shuffled stimulus label assignments for all background noise levels (paired t tests, p < 0.002 for all), except at the highest noise level (50 dB SPL, p = 0.07). Thin lines and thick horizontal bars denote individual mice and sample means, respectively. See also Figure S2.

to the repetition rate of broadband sounds, such as FM sweeps (Figure 2D). We quantified facial synchronization to the sound pressure envelope by performing a Fourier analysis on FME amplitude and calculating the power at the stimulus repetition rate relative to the noise floor (Figure 2E). We found that FME became less synchronized as the presentation rate of FM sweeps increased but tracked the sound envelope for modulation rates up to 3 Hz (Figure 2F). As the dominant frequencies of facial pad movements during whisking fall in the theta range (4–8 Hz) or higher,<sup>17</sup> the 3 Hz low-pass cutoff on sound synchronization is not likely attributable to a motor-related limitation but rather an upstream process in the audiomotor transformation.

### Facial movements index speech processing in background noise

In the hearing sciences, normal response thresholds in silence often belie profoundly abnormal encoding of spectrotemporally complex communication sounds or sound in background

noise.<sup>18–20</sup> Understanding the broader impact of hearing loss or hearing restoration interventions on behavioral registration of communication sounds remains a high priority for animal models of human hearing disorders. Therefore, as a next step, we presented English syllables resynthesized to match the mouse hearing range (Figures 3A and 3B) and quantified facial movement synchronization to these repeating speech tokens in increasing levels of background noise (Figure 3C). We found that the synchronization to the speech presentation rate decreased with background noise level but remained robust up to 50 dB SPL (20 dB SNR) background noise (Figure 3D).

Measuring whether a speech stimulus can be detected in noise is less relevant than whether it can be discriminated in noise. *Gee* and *Ha* clearly differed in spectrum and voice onset timing (Figure 3A), such that the phase of facial synchronization to each syllable was slightly offset (see timing of blue versus red movements, Figure 3C). To determine whether facial registration of timing cues provided a basis for decoding speech token identity in varying levels of background noise, we performed a principal-component analysis to decompose the movement vectors into a lower dimensional space and then trained a support vector machine to classify held-out trials as either *Gee* or *Ha*. We found approximately 80% single-trial speech token classification accuracy, which was significantly higher than chance for all noise levels up to 50 dB SPL (Figure 3E). To establish the generality

of these observations, we performed additional speech decoding experiments with syllable pairs synthesized from vocoded noise (Figure S2A), syllable pairs distinguished primarily by spectral cues (Figure S2B), and syllable pairs primarily distinguished by voice onset timing (Figure S2C). These experiments demonstrated that facial movement features conserved across individual mice arose from a sensitivity to acoustic cues that also influence human speech intelligibility (Figure S2D).

### Facial movements are insensitive to pure tones

Facial videography offers several advantages over traditional indices of hearing function, such as the auditory brainstem response (ABR), in that it can be measured in unanesthetized animals and provides a direct behavioral readout for the encoding of spectrotemporally complex sounds. The principal use of ABR measurements is to provide a non-invasive physiological assay of cochlear function. For example, 2 h of exposure to a 103 dB SPL 16–32 kHz noise band damages outer hair cells and cochlear afferent synapses in the high-frequency base (Figure 4A). The sensorineural damage from this acoustic trauma protocol can be directly visualized with post-mortem cochlear histology<sup>21</sup> and indirectly indexed via an irreversible elevation of ABR thresholds evoked by tone bursts 16 kHz and higher (Figure 4B). To determine whether sound-evoked facial movements can also provide an index of cochlear sensorineural hearing loss (SNHL), we measured FME evoked by pure tones (Figure 4C) but discovered that facial movements were far less responsive to tones than broadband noise, with average response thresholds as high as 90 dB SPL (Figure 4D).

### ABR and facial movements provide congruent measures of hearing loss and excess central gain after acoustic trauma

To address whether the stimuli used for ABR and facial videography measures could be adapted to support direct comparison, we elicited the ABR with octave-band noise (OBN) centered on 8 or 32 kHz and observed low-threshold responses with the expected multi-peaked waveform (Figure 4E). 2 weeks after high-frequency noise exposure, we found that 8 kHz OBN responses were unchanged, while 32 kHz OBN thresholds were elevated by approximately 40 dB, closely matching the threshold shift observed with pure tone stimuli (Figure 4F).

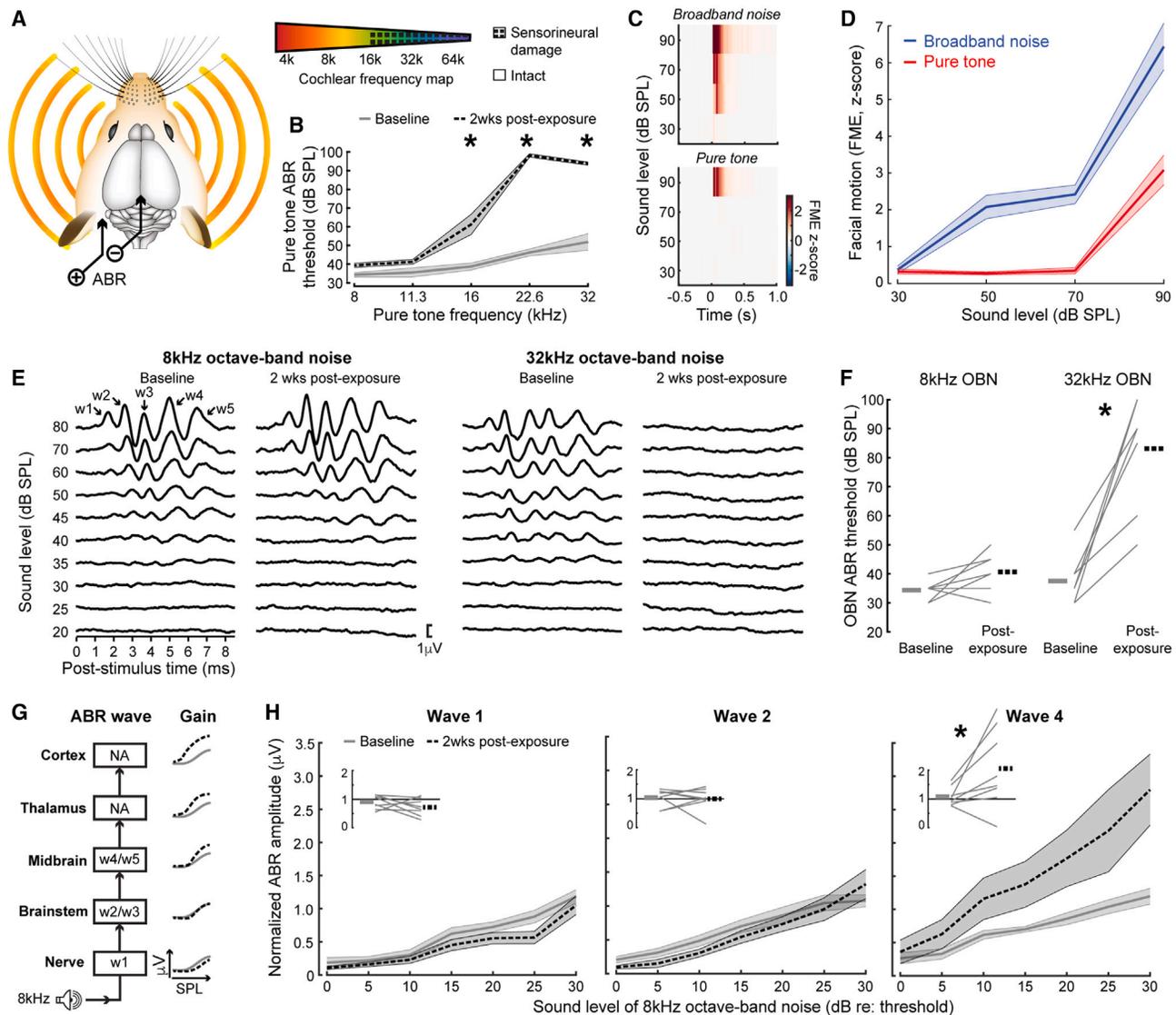
Exposure to intense noise has two types of effects on sound-evoked neural activity: first, neural responses to high-stimulus frequencies within the range of cochlear damage are reduced due to degeneration of cochlear sensory cells and primary afferent nerve endings<sup>22</sup>; second, single-unit responses to lower stimulus frequencies bordering the cochlear lesion are unaffected at early stages of auditory processing but are enhanced at higher stages of the central auditory pathway due to the increased expression of compensatory plasticity processes that are collectively described as excess central gain (Figure 4G).<sup>18,21,23–26</sup> To determine whether excess central gain becomes more prevalent at successive stages of neural processing with gross neuroelectric recordings, we analyzed the amplitude of each ABR wave before and after noise exposure. Each wave of the ABR is generated by the initial volley of synchronized

action potentials at successive stages of central auditory processing.<sup>27,28</sup> Thus, wave 1 is generated by spiral ganglion neurons, waves 2 and 3 by the cochlear nucleus and superior olivary complex, and waves 4 and 5 are generated within the auditory midbrain.<sup>29–32</sup> The 8 kHz OBN reliably elicited waves 1, 2, and 4 across our subjects, affording us the opportunity to quantify the amplitude growth for each wave across a fixed range of sound levels before and after noise exposure. As predicted, we observed that normalized input-output functions for the low-frequency stimulus were slightly attenuated for wave 1, unchanged from baseline for wave 2, but were enhanced above baseline for wave 4, consistent with the emergence of central compensatory plasticity mechanisms at higher stages of central auditory processing (Figure 4H).

We then performed facial videography in noise-exposed mice to determine if FME captured the same combination of high-frequency threshold shift and low-frequency response potentiation. We found that 8 and 32 kHz OBN stimuli elicited strong facial movements (Figure 5A) that increased over a wide range of sound levels (Figure 5B) and remained stable over a 2-week measurement period following a sham exposure to moderate sound levels that do not damage the cochlea (Figure 5C). Noise-induced high-frequency hearing loss was associated with a 40–50 dB increase in 32 kHz OBN thresholds without affecting 8 kHz OBN thresholds, matching the standard ABR-based approach for measuring noise-induced hearing loss (Figure 5D).

We hypothesized that if sound-evoked facial movements were mediated by a brainstem pathway we would—similar to the early waves of the ABR—observe an invariant 8-kHz-evoked response amplitude over the measurement period and stably depressed 32-kHz-evoked responses (Figure 5E). Conversely, if sound-evoked reflexes were mediated by auditory midbrain or forebrain nuclei, we would expect that 8-kHz-evoked facial movements would grow to exceed baseline levels within days (forebrain) or weeks (midbrain) following high-frequency SNHL, whereas 32-kHz-evoked movements would be initially reduced before staging a partial recovery to baseline levels.<sup>18,25,33</sup>

Sound-evoked facial movements exhibited a pattern of loss and recovery after SNHL most consistent with a neural pathway including midbrain or forebrain nuclei (Figure 5F). We found that FME amplitudes elicited by the spared 8 kHz noise band at near-threshold intensities were initially stable after noise exposure but then significantly exceeded baseline levels 2 weeks following noise exposure (Figure 5G, top). Similarly, facial movements evoked by the 32 kHz noise band were nearly eliminated hours after noise exposure but partially recovered 2 weeks after noise exposure (Figure 5G, bottom). Comparing changes in the 8 kHz ABR and FME amplitudes in the same mice 2 weeks after noise-induced SNHL, we found that individual differences in FME enhancement were highly correlated with ABR wave 4 potentiation, yet no association was observed with changes in waves 1 or 2 (Figure 5H). Finally, the combination of enhanced facial movements at 8 kHz and suppressed responses at 32 kHz observed after SNHL was not a consequence of repeated testing over the 2-week period, as facial movement amplitudes at both test frequencies were comparatively stable in sham-exposed mice (Figure 5I).



**Figure 4. Noise-induced SNHL causes a combination of high-frequency threshold shift and excess low-frequency gain in the ABR**

(A) 2 h of exposure to 16–32 kHz octave-band noise (OBN) at 103 dB SPL causes sensorineural damage in the high-frequency base of the cochlea, as determined with pinna-vertex ABR measurements.

(B) Tone-evoked ABR measurements reveal significant threshold elevation for test frequencies above 11.3 kHz measured before vs 2 weeks after noise exposure (2-way repeated measures ANOVA,  $N = 9$ , main effect for group,  $[F = 146.95, p = 9 \times 10^{-9}]$ ; group  $\times$  frequency interaction,  $[F = 39.73, p = 9 \times 10^{-16}]$ ). Asterisks denote post-hoc pairwise comparisons ( $p < 0.05$  for all).

(C) Mean peristimulus FME responses from 8 mice for pure tone and broadband noise stimuli.

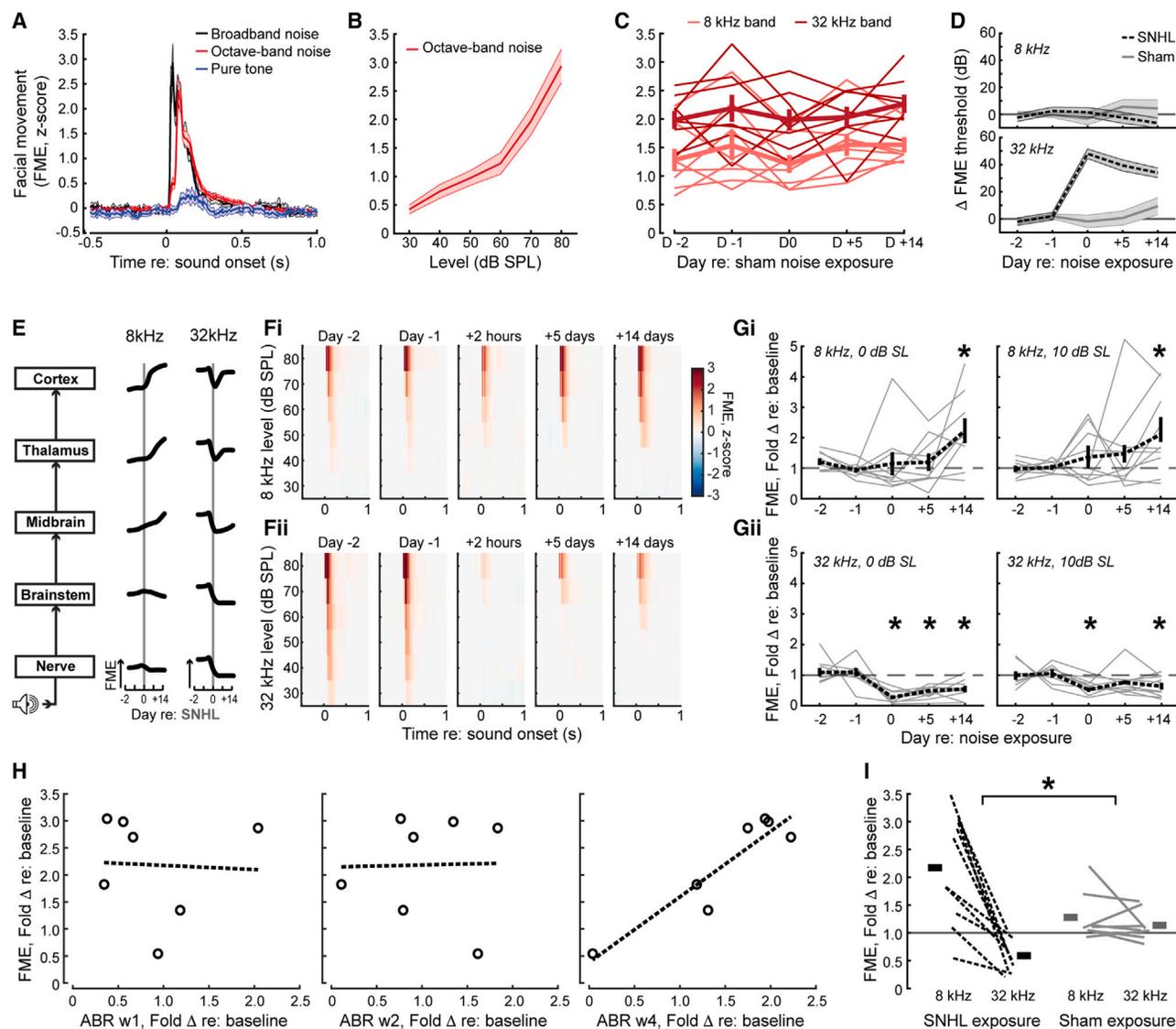
(D) Mean  $\pm$  SEM FME amplitudes over sound levels. Broadband noise evokes significantly larger facial movements than pure tones (2-way repeated measures ANOVA,  $N = 8$ , main effect for stimulus type,  $[F = 54.95, p = 4 \times 10^{-6}]$ ; group  $\times$  frequency interaction,  $[F = 10.87, p = 3 \times 10^{-5}]$ ).

(E) OBN centered at 8 and 32 kHz elicits a robust multi-peaked ABR before acoustic trauma. 2 weeks after noise-induced high-frequency SNHL, the ABR response to the 32 kHz noise band is virtually absent at sound levels up to 80 dB SPL, whereas responses to the 8 kHz noise band appear unaffected or slightly larger than baseline measurements. Arrows indicate the appearance of ABR waves (w) 1–5.

(F) OBN ABR thresholds after noise exposure (two-way repeated measures ANOVA,  $N = 8$ ; main effect for frequency  $[F = 38.26, p = 2 \times 10^{-5}]$ ; main effect for timepoint  $[F = 46.59, p = 8 \times 10^{-6}]$ , frequency  $\times$  time point  $[F = 38.259, p = 1 \times 10^{-4}]$ ).

(G) Schematic illustrating the neural generators of ABR waves 1–5 and the expected transition from slight attenuation in the 8 kHz OBN level  $\times$  amplitude input-output function for early waves to excess central gain measured in later waves. NA indicates that the neural generators of the ABR do not include central auditory structures above the midbrain.

(H) Mean  $\pm$  SEM 8 kHz OBN-evoked normalized wave amplitude growth functions plotted relative to a threshold. Inset: 8 kHz OBN-evoked normalized wave amplitudes were averaged within a 30–45 dB range above the threshold. Solid black line represents no change, gray lines represent individual subjects, and thick gray and dashed black lines represent baseline and 2 weeks post-exposure, respectively. Asterisks denote  $p < 0.05$ .



**Figure 5. Changes in sound-evoked facial movement after noise-induced SNHL parallel modifications in late ABR waves**

(A) Mean  $\pm$  SEM facial movements evoked by broadband noise, octave-band noise (OBN), and pure tones at 70 dB SPL. OBN and pure tone responses are averaged across 8 and 32 kHz.

(B) OBN responses are elicited at low sound levels and grow monotonically with sound level (one-way repeated measures ANOVA [ $F=189.9$ ,  $p=6 \times 10^{-66}$ ,  $N=20$ ]).

(C) Sound-evoked facial movements elicited by OBN centered at 8 and 32 kHz are stable over a 17-day (D) measurement period spanning sham noise exposure (2-way repeated measures ANOVA,  $N=8$ , main effect for day, [ $F=1.2$ ,  $p=0.33$ ]; main effect for frequency, [ $F=40.1$ ,  $p=4 \times 10^{-4}$ ]).

(D) Top: thresholds for sound-evoked facial movements with the 8 kHz OBN are unchanged over time after noise exposure and do not differ between SNHL and sham groups (2-way repeated measures ANOVA,  $N=8/10$  sham/SNHL, main effect for day [ $F=0.14$ ,  $p=0.93$ ]; main effect for group [ $F=0.44$ ,  $p=0.52$ ]). Bottom: Thresholds for sound-evoked facial movements with the 32 kHz OBN are elevated after SNHL but not sham noise exposure (2-way repeated measures ANOVA,  $N=18$ , main effect for day, [ $F=32.87$ ,  $p=1 \times 10^{-11}$ ]; main effect for group, [ $F=39$ ,  $p=1 \times 10^{-5}$ ]).

(E) Schematic illustrating hypothetical changes in sound-evoked facial movement amplitudes over the same 17-day period before and after an SNHL-inducing noise exposure (vertical gray line). The cartoon model assumes that central gain is progressively enhanced at successive stages of the central pathway, promoting faster and more complete recovery of the high-frequency response in the damaged region of the cochlea and hyper-responsiveness to the low-frequency noise band.

(F) Mean peristimulus FME responses from 8 mice over a 17-day period spanning noise-induced hearing loss for an 8 and 32 kHz OBN (top and bottom, respectively).

(G) Actual data after SNHL are compared against the cartoon model shown in (E) by plotting the fold change in facial movement amplitudes relative to baseline (mean of D-1 and D-2) for sound levels at threshold or 10 dB above threshold (i.e., 0 dB and 10 dB sensation level, SL). Data from individual mice ( $N=10$ ) and group mean are shown as thin gray and thick dashed lines, respectively. Asterisks denote that the change in sound-evoked facial movements are either significantly elevated relative to baseline (top) or significantly suppressed relative to baseline (bottom), as assessed with one-sample t tests relative to a population mean of 1.0 ( $p < 0.02$  for all significant time points).

(legend continued on next page)

### Suppressing auditory cortex neural activity enhances sound-evoked facial movements

To study the correspondence between facial movements and neural activity dynamics in brain regions downstream of the auditory midbrain, we performed extracellular single-unit recordings from the primary auditory cortex (A1) with 64-channel laminar probes and identified laminar boundaries based on current source density analysis (Figure 6A). Isolated regular spiking (RS) single units increased their activity shortly after the presentation of broadband sounds, making it difficult to isolate the relative contribution of sound-evoked sensory inputs from movement-related activation when they overlap in time. For this reason, we identified spontaneous FME transients that occurred during the silent inter-trial interval (Figure 6B). In L2/3 and L4, only 11% of A1 RS units exhibited significant spike rate modulation during spontaneous facial movements, whereas 25% of units in L5 and L6 were significantly modulated (Figures 6C and 6D). We and others have previously reported that self-initiated movements with or without coincident sound stimulation modulate the spike rates of layer 5 and layer 6 auditory cortex (ACtx) units more than upper layer units.<sup>34–36</sup> Here, we also observed that spike rate modulation during spontaneous facial twitches was greater in deeper layer units (Figure 6E). Interestingly, motor-preparatory activity in ACtx begins hundreds of milliseconds before the onset of comparatively complex, goal-directed movements that involve licking or level pressing,<sup>35–37</sup> but appeared coincident with the onset of spontaneous facial twitches.

Because deep-layer ACtx neurons respond both to sound and FME transients, they could be an obligatory relay for converting the sensory representation into a motor signal. On the other hand, sound-evoked first spike latencies are on the order of 10–25 ms in the deep-layer ACtx neurons,<sup>38</sup> which overlaps with the onset of sound-evoked FME responses ( $24 \pm 1.1$  ms; Figure 2B). By contrast, sound-evoked first spike latencies in the inferior colliculus, an auditory midbrain structure, are on the order of 5–10 ms, thus occurring well ahead of sound-evoked movements.<sup>39</sup> Inferior colliculus neurons also generate wave 4 of the ABR, which showed a close correspondence with changes in sound-evoked facial movements after noise exposure (Figure 6H) and also receive massive corticofugal feedback from the ACtx. Collectively, these pieces of evidence suggest an alternative model where ACtx is not a mediator of sound-evoked facial movements but instead modulates the sensorimotor transformation via its descending projections to subcortical nuclei.

To further test the role of ACtx as a mediator or modulator of sound-evoked facial movements, we optogenetically inhibited ACtx RS unit spiking. If ACtx mediated the response, sound-evoked facial movements would be eliminated in photoinhibition trials (Figure 6F, left). If ACtx played a modulatory role, sound-

evoked facial movements would be attenuated or enhanced during cortical photoinhibition (Figure 6F, right). We first confirmed that optogenetic activation of parvalbumin-expressing (PV) GABAergic interneurons effectively suppressed spiking in ACtx RS units (Figure 6G) and, in the absence of sound, did not elicit facial movements (Figure S3). Next, we bilaterally activated PV neurons throughout the peri-stimulus period and interleaved laser-on trials with laser-off trials (Figure 6H). We found that FME was significantly enhanced during ACtx photoinhibition, supporting a conceptual model where the ACtx is a modulator rather than a mediator of sound-evoked facial movements (Figure 6I).

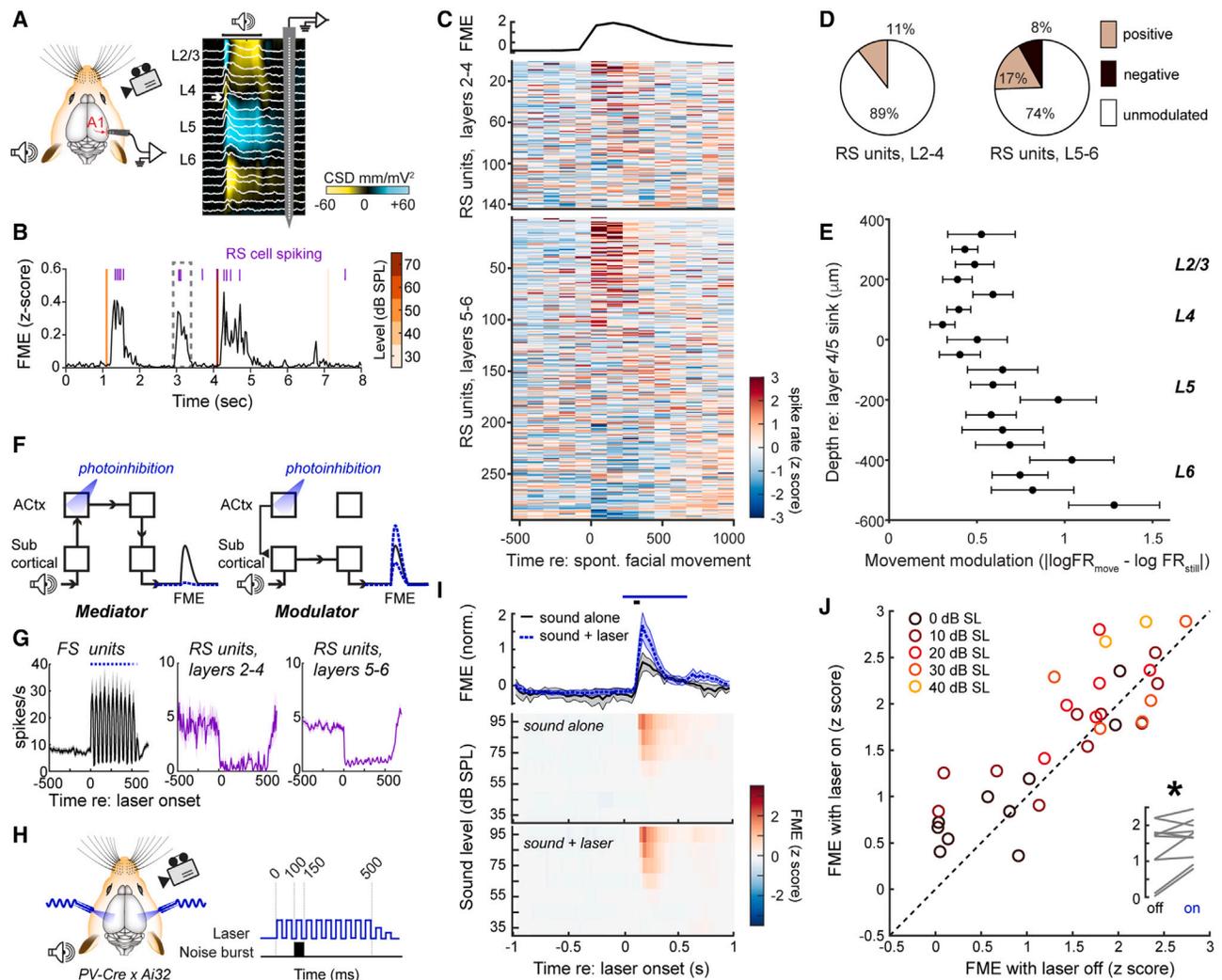
### Sound-evoked facial movements capture an auditory sensory hyper-responsivity phenotype in mice with an autism risk gene mutation

Sensory overload is a cardinal feature of autism spectrum disorder (ASD), particularly in the auditory modality, where moderate-intensity sounds are overwhelmingly loud and distressing.<sup>40,41</sup> The sensory overload ASD phenotype can be challenging to model in laboratory animals because the broader cognitive and motor impairment related to mutations in ASD risk genes can interfere with acquiring and performing the procedural demands of an operant behavioral task. Here, we focused on mice with deletion of *Ptchd1*, an ASD risk gene expressed in the thalamic reticular nucleus during prenatal development.<sup>42</sup> Male mice with *Ptchd1* deletion exhibit attention deficits, gross hyperactivity, and poor distractor suppression during a visual detection task, but an auditory hyper-reactivity phenotype has not been reported.<sup>43</sup>

We first confirmed that ABR thresholds were equivalent in *Ptchd1* KO mice and wild-type controls, demonstrating that any differences in auditory responses were unlikely attributable to the differences in early auditory processing. Facial videography revealed that the baseline pupil diameter was abnormally large in *Ptchd1* KO mice, consistent with the descriptions in several other strains of mice with ASD risk gene mutations (Figures 7B and 7C).<sup>44</sup> Sound-evoked pupil dilation growth functions were comparable between KO and WT mice after baseline pupil differences were factored out (two-way repeated measures ANOVA,  $N = 11/14$  for WT/KO; main effect for genotype,  $F = 3.4$ ,  $p = 0.08$ , data not shown). In contrast, spontaneous facial movements during the inter-trial interval were comparable between WT and KO mice (Figure 7D, two-sample *t* test,  $N = 11/15$ ,  $p = 0.25$ ), but sound-evoked FME was significantly steeper across a range of moderate sound intensities in KO mice compared with WT controls (Figure 7E). These findings demonstrate that sound-evoked facial movements can capture a heritable hyper-responsivity phenotype related to genetic mutations as well as acquired hyper-responsivity caused by noise-induced SNHL.

(H) Changes in ABR amplitude and sound-evoked facial movements elicited by the 8 kHz noise band after noise-induced high-frequency SNHL are calculated for a subset of mice with data from both measurement types ( $N = 7$ ). Dashed line presents the linear fit of the data. Facial movement changes reflect the average of 0 and 10 dB SL sound levels. Increased 8-kHz-evoked facial movements are not correlated with changes in ABR wave 1 (Pearson  $r = -0.05$ ,  $p = 0.92$ ) or ABR wave 2 ( $r = 0.02$ ,  $p = 0.96$ ) but are significantly correlated with changes in ABR wave 4 ( $r = 0.92$ ,  $p = 0.003$ ).

(I) The pattern of increased responses at 8 kHz and decreased responses at 32 kHz observed 2 weeks after SNHL is not observed following sham exposure (two-way mixed model ANOVA, main effect for group [ $F = 0.65$ ,  $p = 0.42$ ]; main effect for frequency [ $F = 22.45$ ,  $p = 3 \times 10^{-4}$ ]; Group  $\times$  Frequency interaction [ $F = 15.65$ ,  $p = 0.001$ ]). Thin and thick lines represent data from individual mice and group means from the SNHL group ( $N = 10$ ) and sham exposure group ( $N = 8$ ), respectively.



**Figure 6. Suppressing auditory cortex activity facilitates sound-evoked facial movements**

(A) Extracellular recordings were made from all layers of the primary auditory cortex (A1) with a 64-channel linear probe during contralateral sound presentation and facial videography. Electrophysiological responses are filtered offline to separate spiking activity (white trace) and the current source density (CSD). White arrow in the CSD trace identifies the early current sink in layer (L) 4 elicited by a 70 dB SPL 50 ms white noise burst that is used to assign units to layers.

(B) FME is increased following sound presentation (orange line), but facial twitches also occur spontaneously (dashed box). Action potentials (purple) from a single regular spiking (RS) unit are evoked by the combination of sound and movement but also during spontaneous facial movements.

(C) Neurograms present the Z-scored firing rates before and after bouts of spontaneous facial movements from 438 single RS units grouped into superficial (L2/3 and L4) and deep (L5 and L6) layers of the cortical column. Units are sorted by their mean activity. The line plot presents the mean FME over the same period.

(D) Pie charts represent whether and how single-unit firing rates were modulated by spontaneous facial movements.

(E) Mean  $\pm$  SEM absolute value of firing rate changes during spontaneous facial movements along the cortical column. Spike rate modulation is significantly elevated with increasing depth in the cortical column (one-way ANOVA,  $n = 438$ , main effect of depth [ $F = 2.09$ ,  $p = 0.006$ ]).

(F) Left: schematic illustrates that optogenetic suppression of auditory cortex (ACtx) spiking could eliminate sound-evoked facial movements if it were an obligatory sensorimotor relay. Right: alternatively, optogenetic suppression of ACtx spiking could amplify or attenuate sound-evoked facial movements if it modulated a subcortical sensorimotor relay.

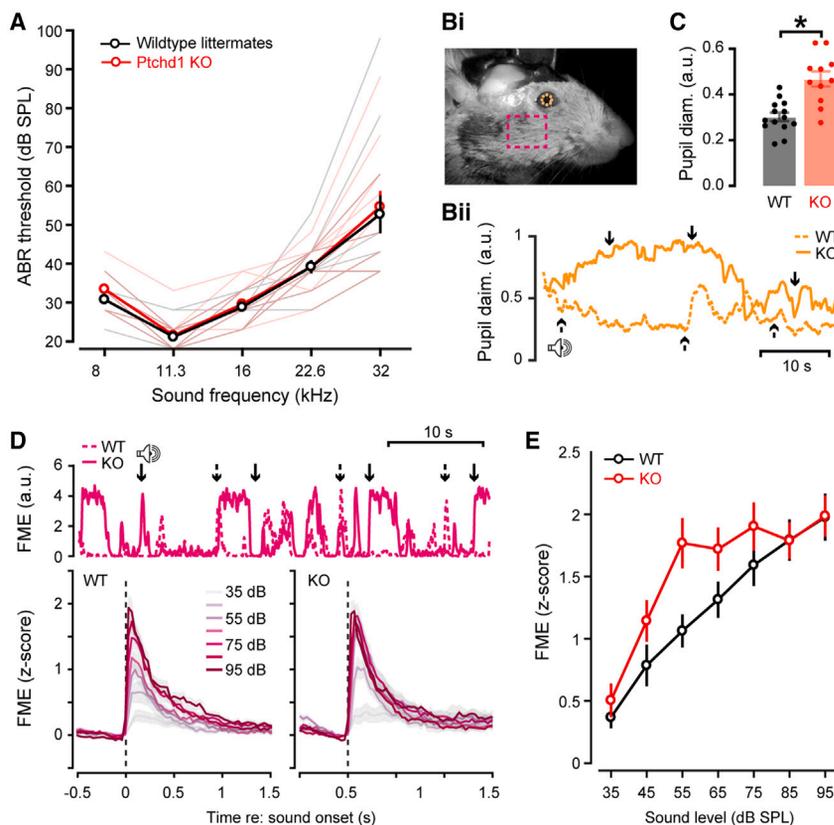
(G) Mean  $\pm$  SEM changes in spike rate in response to 20 Hz activation of parvalbumin-expressing (PV) GABAergic interneurons in fast-spiking putative PV units ( $n = 99$ ), superficial RS units ( $n = 33$ ), and deep-layer RS units ( $n = 373$ ) from  $N = 4$  mice.

(H) Experimental paradigm to test the two hypothetical scenarios described in (F). Bilateral activation of PV interneurons was interleaved with sound-only trials.

(I) Sound-evoked FME on interleaved laser-on and laser-off trials. Top: mean  $\pm$  SEM FME. Thick black and blue lines depict the relative timing of sound and laser timing, respectively. Bottom: mean FME over a range of broadband noise levels ( $N = 10$ ).

(J) Scatterplot presents mean sound-evoked FME during laser-on/photoinhibition versus laser-off trials. Each symbol represents the trial-averaged mean from a single animal color coded according to the level of the sound relative to a threshold. Data points above the line of unity (dashed diagonal) are enhanced during ACtx suppression. Inset: mean FME within 10 dB SPL above threshold is plotted for each mouse during laser-off and -on trials. Asterisk denotes a significant difference (paired  $t$  test,  $p = 0.005$ ).

See also [Figure S3](#).



**Figure 7. Hyper-responsive sound-evoked facial movements in mice with an autism risk gene mutation**

(A) ABR thresholds are not significantly different in mice with *Ptchd1* deletion (KO) and wild-type littermate controls (two-way repeated measures ANOVA, main effect for frequency [ $F = 131, p < 0.001$ ], main effect for genotype [ $F = 0.31, p = 0.58$ ], frequency  $\times$  genotype interaction [ $F = 0.74, p = 0.60$ ]). (B) (i) Placement of pupil markers and ROI for FME calculation. (ii) Representative traces of pupil diameter changes in a wild-type (WT) and KO mouse. Solid and dashed arrows denote the timing of noise bursts in the corresponding recording session. (C) Baseline pupil diameter in KO mice is significantly larger than that in WT controls (two-sample  $t$  test, [ $t = -4.64, p < 0.001$ ]). Bars and error bars represent mean  $\pm$  SEM. Each data point represents an individual mouse ( $N = 14/11, WT/KO$ ). (D) Top: representative FME traces in a wild-type (WT) and KO mouse. Solid and dashed arrows denote the timing of noise bursts in the corresponding recording session. Bottom: mean  $\pm$  SEM FME for noise bursts of increasing intensity in WT (left) and KO (right) groups ( $N = 15/11, WT/KO$ ). (E) Sound-evoked facial movements grow significantly more steeply across sound levels in KO mice compared with WT controls (two-way mixed design ANOVA, main effect for sound level [ $F = 43.46, p = 0.001$ ], main effect for genotype [ $F = 2.29, p = 0.14$ ], level  $\times$  genotype interaction [ $F = 2.67, p = 0.02$ ]).

## DISCUSSION

We used quantitative videography to characterize sound-evoked movements of the face. We found that sound elicited small movements of a region of the face just caudal to the vibrissae array that was 30–40 dB more sensitive than the acoustic startle reflex and less variable than other points of measurement on the face (Figure 1). FME faithfully encoded the low-frequency envelope of broadband stimuli up to 3 Hz (Figure 2), supporting accurate single-trial decoding of speech syllable identity in varying levels of background noise (Figure 3). We found that low-threshold sound-evoked FME was only elicited by sounds with spectral bandwidth greater than one octave (Figure 4) and showed no signs of habituation across repeated measurements over multiple weeks (Figure 5). After noise-induced damage to the high-frequency cochlear base, FME was attenuated for a high-frequency stimulus, but low-frequency sounds targeting the spared region of the cochlea elicited hyper-responsive facial movements that were strongly correlated with the potentiation of ABR wave 4, generated by the auditory midbrain. Firing rates of some deep-layer A1 RS units were entrained by spontaneous facial movements, yet optogenetic A1 RS unit silencing increased sound-evoked FME, indicating that the ACTx is a modulator—not a mediator—of sound-evoked facial movements (Figure 6). Finally, we report that FME is hypersensitive at intermediate sound intensities in the *Ptchd1* KO mouse, consistent with an ASD auditory hypersensitivity phenotype associated with ASD (Figure 7).

## Facial videography is a sensitive readout of behavioral sound registration

Hearing, the conscious awareness of sound, is a psychological construct that can only be directly assessed through behavioral measurements. Operant assays for behavioral reporting of conscious sound awareness represent the gold standard for hearing assessments. However, training animals to perform operant tasks can take weeks, which makes measurements of complex sound processing difficult in most animal models and prohibitive for models of disordered sound perception that also have generalized motor, vestibular, or cognitive impairments.<sup>45–48</sup> Alternatively, physiological proxies for hearing such as the ABR can be measured in minutes, but, as gross electrical potentials generated in the brainstem of anesthetized animals, these measures are not interchangeable with hearing assessments. For example, selective lesions of the cochlear afferent pathway render the ABR absent or grossly abnormal, yet auditory detection thresholds measured with operant behaviors are unaffected.<sup>23,49</sup>

While sound-evoked facial movements provide a high-fidelity readout of sound registration by the central nervous system, the approach also has limitations. First, the correspondence between facial movement amplitude and the behavioral report of sound detection is unknown. Perceptual awareness in operant tasks is closely linked to cortical registration of the sensory stimulus.<sup>48,50–52</sup> While ACTx is required for many forms of operant detection and discrimination,<sup>52,53</sup> silencing excitatory neurons in ACTx did not abolish or even attenuate facial movements. We also noted that sound-evoked facial movements were

more easily disrupted by background noise than operant behavior. Sound processing in noise is generally robust down to SNRs well below zero when studied with operant tasks but was at chance at 20 dB SNR based on facial movements.<sup>45,46,50,51</sup> In a recent study, we observed that FME was reduced after Pavlovian discriminative threat conditioning, but the change was not specific to the sound associated with aversive stimulation, highlighting another difference between FME and classic measures of associative learning such as freezing or increased pupil dilation.<sup>13</sup>

Auditory research relies heavily on narrowband sounds, yet sound-evoked facial movements were largely insensitive to these stimuli. Despite this broadband sound preference, we demonstrated that frequency-specific measurements using band-limited noise were possible and could reveal bi-directional changes following cochlear damage similar to operant behavior and cortical plasticity.<sup>21</sup> Additionally, while sound-evoked FME can be measured in freely moving animals,<sup>11</sup> these measurements require custom-built head-mounted cameras or multi-camera arrays coupled with machine vision approaches. These observations underscore that while FME is relatively easy to measure and offers advantages over other behavioral and electrophysiological hearing indices, it also has drawbacks and should not be equated with a measure of conscious sound awareness.

### The neural basis of sound-evoked facial movements

We found that optogenetic inactivation of ACTx paradoxically increased sound-evoked facial movements. This result suggests that the ACTx may modulate subcortical pathways for sound-evoked facial movements via the extensive network of descending corticofugal projections.<sup>38,54–57</sup> Given the strong correspondence between the potentiation of ABR wave 4 and FME at spared frequencies, our data suggest that sound-evoked facial movements are routed through the inferior colliculus, though the definitive evidence for this claim would require a series of lesions or inactivation experiments in various subcortical centers for sound processing. An auditory midbrain pathway is consistent with reports that electrical or optogenetic stimulation of the inferior colliculus induces immediate motor responses,<sup>55,58</sup> including vibrissae movement in anesthetized rats,<sup>59</sup> while optogenetic activation of the auditory thalamus is not known to induce motor activity.<sup>21,60</sup> Inferior colliculus neurons can be excited or suppressed by ACTx activation,<sup>61,62</sup> where the direction of corticofugal modulation of spike rate can reflect the recruitment of recurrent inhibition in intracollicular circuits,<sup>63</sup> and ACTx stimulation parameters.<sup>61,64</sup> Conversely, silencing ACTx increases spontaneous and sound-evoked activity in the inferior colliculus.<sup>65</sup> An increase in sound-evoked FME responses during cortical silencing is therefore consistent with a release of tonic inhibition imposed by corticofugal recruitment of collicular inhibitory circuits.

While afferent signals eliciting sound-evoked facial movements may be routed through the auditory midbrain, the selectivity for broadband sounds points to the essential involvement of brain areas outside the classical auditory pathway, where neural responses to low-intensity pure tones are observed from the cochlear nucleus to the ACTx.<sup>66–68</sup> Selectivity for broadband sounds, but not pure tones, has been found in a reticular-limbic

auditory pathway, though auditory response thresholds from neurons in these brain regions are higher than FME thresholds reported here.<sup>69</sup> Inferior colliculus neurons project primarily not only to the auditory thalamus but also to the dorsal medial periaqueductal gray and the intermediate and deep layers of the superior colliculus.<sup>55</sup> The dorsal periaqueductal gray is known to be involved in the generation of defensive behaviors,<sup>70,71</sup> but defensive behaviors elicited by sound should rapidly adapt within one to two stimulus presentations,<sup>72</sup> inconsistent with the indefatigable sound-evoked facial movements we observed. In contrast, the intermediate and ventral layers of the superior colliculus are known to respond robustly to sound, with receptive fields that are preferential toward high-frequency sounds with large spectral bandwidth.<sup>73,74</sup> These high-frequency, large spectral bandwidth receptive fields are congruent with our finding that baseline FME responses were greater for an OBN centered at 32 kHz compared with 8 kHz. Anatomical tracing in mice has identified direct projections from superior colliculus to facial nucleus motoneurons,<sup>75</sup> and optogenetic activation of the SC elicits attempted head movements in head-fixed mice.<sup>76</sup> It remains possible that sound-evoked facial movements could be routed through a brainstem pathway, as it is clear that audiomotor pathways have not been exhaustively characterized. For instance, a novel multi-synaptic pathway from the lateral lemniscus in the auditory brainstem to the anterior lateral motor cortex has recently been described.<sup>77</sup> Further work will be required to elucidate the precise pathway through which auditory signals evoke facial movements.

### Broader consequences of sound-evoked facial movements for neuroscience research

The findings reported here and in earlier descriptions of sound-evoked facial movements have important implications for the design and interpretation of multisensory integration experiments. A wide range of moderate-intensity “purely” auditory stimuli will in fact elicit a complex mixture of auditory, motor, and reafferent somatosensory activity within 30 ms following sound onset. Disentangling the relative weighting of sensorimotor interactions arising from peripheral reafferent versus central pathways will be challenging. For example, a recent study demonstrated that auditory influences on visual cortex response properties could mostly be accounted for by global, low-dimensional activity patterns arising from sound-evoked facial movements.<sup>12</sup> “Unisensory” experiments must also reckon with the influence of motor corollary and peripheral sensory reafferent inputs. We reported here that the spiking rates of a substantial minority of RS units in layers 5 and 6 of A1 were significantly modulated shortly after spontaneous facial movements, supporting earlier mesoscale recordings that described widespread activity across the dorsal cortex related to facial movements.<sup>78</sup> Clayton et al.,<sup>35</sup> Audette et al.,<sup>36</sup> and Rummell et al.<sup>37</sup> have shown that planned movements of the face or forelimb modulate the spiking of layers 5 and 6 neurons hundreds of milliseconds prior to movement and sound onset, highlighting the inextricable link between movement and sound and underscoring the difficulty of studying one in the absence of the other, at least in awake subjects. However, our work shows that facial movements are not elicited by all sounds but rather are only elicited by broadband stimuli. Thus, by carefully controlling the stimulus (i.e., using

narrowband sounds that are <90 dB SPL), potential confounds caused by sound-evoked facial movements can be mitigated, or even avoided altogether.

### Sound-evoked facial movements as a tool for active sensing

Most reflex pathways have a clear adaptive value to the animal. For example, the acoustic startle reflex has been suggested not only to protect the body from immediate harm<sup>79</sup> but also to rapidly terminate ongoing sequences of motor activity and prepare escape behaviors.<sup>80</sup> The adaptive value of sound-evoked facial movements is not entirely clear, though one can speculate that a synchronized low-dimensional volley of bottom-up auditory, somatosensory, and motor activity could induce a global neural and behavioral state shift to active sensing and orienting.<sup>76,81</sup> Facial movement amplitudes increased with spectral bandwidth and sound level, providing an analog representation of environmental sound sources that could more effectively recruit multisensory brainstem or reticular-limbic networks for threat avoidance than an internal representation based only on acoustic features.<sup>69,82–84</sup> Future work using head-mounted cameras in freely moving animals<sup>11,85,86</sup> could help to place facial movements in the context of orienting and threat avoidance behaviors. New findings confirm that emotionally evocative sounds also elicit facial movements in humans that increase with self-reported valence, though these movements occur at longer latency and are less stereotyped.<sup>87</sup> Assessing the generality of these findings to other species with various histories of sensory experience will reveal whether sound-evoked facial movements represent a broadly useful framework for studying hearing and internal state regulation or are more of a special case observed only in laboratory mice.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**
  - Animal subjects
- **METHOD DETAILS**
  - Surgical preparation
  - Optical access to auditory cortex
  - Videography
  - Stimulus generation and presentation
  - High-frequency noise exposure
  - Cochlear function testing
  - Cortical electrophysiology
  - Optogenetic stimulation
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - Video processing and analysis
  - Acoustic startle reflex measurement
  - Electrophysiology acquisition and online analysis

- Single-unit identification and analysis
- Statistical analysis

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cub.2024.02.057>.

### ACKNOWLEDGMENTS

We thank Dr. Michael Halassa for the gift of the *Ptchd1*-KO mouse. We thank Ashvini Melkote for assistance with DeepLabCut; Dr. Meenakshi Asokan for providing videography analysis code and performing initial pilot experiments; Dr. Matthew McGill for assistance with cochlear function measurements; Dr. John Rosowski for discussions regarding acoustics; and Divya Narayanan, Yurika Watanabe, and Jennifer Zhu for providing surgical support. This work was supported by the Nancy Lurie Marks Family Foundation (to D.B.P.) and NIH grant DC009836 (to D.B.P.).

### AUTHOR CONTRIBUTIONS

Conceptualization, K.K.C., K.C., and D.B.P.; methodology, K.K.C., A.A.G., K.C., and K.E.H.; investigation, K.K.C., K.S.S., A.R.C., and K.C.; software, K.K.C., A.A.G., A.R.C., K.C., and K.E.H.; formal analysis, K.K.C., A.R.C., K.C., and A.A.G.; data curation, K.K.C., A.R.C., K.C., A.A.G., and K.S.S.; visualization, K.K.C., A.R.C., K.C., and D.B.P.; writing – original draft, D.B.P. and K.K.C.; writing – review & editing, K.K.C. and D.B.P.; resources, D.B.P.; supervision, D.B.P.; funding acquisition, D.B.P.

### DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 7, 2023

Revised: January 2, 2024

Accepted: February 23, 2024

Published: March 15, 2024

### REFERENCES

1. Davis, M., Gendelman, D.S., Tischler, M.D., and Gendelman, P.M. (1982). A Primary acoustic startle circuit: Lesion and stimulation studies. *J. Neurosci.* **2**, 791–805.
2. Lauer, A.M., Behrens, D., and Klump, G. (2017). Acoustic startle modification as a tool for evaluating auditory function of the mouse: Progress, pitfalls, and potential. *Neurosci. Biobehav. Rev.* **77**, 194–208.
3. Willott, J.F., Tanner, L., O'Steen, J., Johnson, K.R., Bogue, M.A., and Gagnon, L. (2003). Acoustic startle and prepulse inhibition in 40 inbred strains of mice. *Behav. Neurosci.* **117**, 716–727.
4. Valero, M.D., Hancock, K.E., and Liberman, M.C. (2016). The middle ear muscle reflex in the diagnosis of cochlear neuropathy. *Hear. Res.* **332**, 29–38.
5. Chambers, A.R., Hancock, K.E., Maison, S.F., Liberman, M.C., and Polley, D.B. (2012). Sound-evoked olivocochlear activation in unanesthetized mice. *J. Assoc. Res. Otolaryngol.* **13**, 209–217.
6. Guitton, M.J., Avan, P., Puel, J.L., and Bonfils, P. (2004). Medial olivocochlear efferent activity in awake guinea pigs. *NeuroReport* **15**, 1379–1382.
7. Montes-Lourido, P., Kar, M., Kumbam, I., and Sadagopan, S. (2021). Pupillometry as a reliable metric of auditory detection and discrimination across diverse stimulus paradigms in animal models. *Sci. Rep.* **11**, 3108.
8. Bala, A.D.S., and Takahashi, T.T. (2000). Pupillary dilation response as an indicator of auditory discrimination in the barn owl. *J. Comp. Physiol. A* **186**, 425–434.
9. Plappert, C.F., and Pilz, P.K.D. (2005). Long-term habituation of the startle response in mice evoked by acoustic and tactile stimuli. *Behav. Brain Res.* **162**, 307–310.

10. Jero, J., Coling, D.E., and Lalwani, A.K. (2001). The use of Preyer's reflex in evaluation of hearing in mice. *Acta Otolaryngol.* *121*, 585–589.
11. Meyer, A.F., Poort, J., O'Keefe, J., Sahani, M., and Linden, J.F. (2018). A Head-Mounted Camera System Integrates Detailed Behavioral Monitoring with Multichannel Electrophysiology in Freely Moving Mice. *Neuron* *100*, 46–60.e7.
12. Bimbard, C., Sit, T.P.H., Lebedeva, A., Reddy, C.B., Harris, K.D., and Carandini, M. (2023). Behavioral origin of sound-evoked activity in mouse visual cortex. *Nat. Neurosci.* *26*, 251–258.
13. Asokan, M.M., Watanabe, Y., Kimchi, E.Y., and Polley, D.B. (2023). Potentiation of cholinergic and corticofugal inputs to the lateral amygdala in threat learning. *Cell Rep.* *42*, 113167.
14. Grimsley, C.A., Longenecker, R.J., Rosen, M.J., Young, J.W., Grimsley, J.M., and Galazyuk, A.V. (2015). An improved approach to separating startle data from noise. *J. Neurosci. Methods* *253*, 206–217.
15. Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., and Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* *21*, 1281–1289.
16. Shannon, R.V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech Recognition with Primarily Temporal Cues. *Science* *270*, 303–304.
17. Carvell, G.E., and Simons, D.J. (1995). Task- and subject-related differences in sensorimotor behavior during active touch. *Somatosens. Mot. Res.* *12*, 1–9.
18. Chambers, A.R., Resnik, J., Yuan, Y., Whitton, J.P., Edge, A.S., Liberman, M.C., and Polley, D.B. (2016). Central Gain Restores Auditory Processing following Near-Complete Cochlear Denervation. *Neuron* *89*, 867–879.
19. Parthasarathy, A., Hancock, K.E., Bennett, K., Degruittola, V., and Polley, D.B. (2020). Bottom-up and top-down neural signatures of disordered multi-talker speech perception in adults with normal hearing. *eLife* *9*, 1–22.
20. Sabesan, S., Fragner, A., Bench, C., Drakopoulos, F., and Lesica, N.A. (2023). Large-scale electrophysiology and deep learning reveal distorted neural signal dynamics after hearing loss. *eLife* *12*, 1–32.
21. McGill, M., Hight, A.E., Watanabe, Y.L., Parthasarathy, A., Cai, D., Clayton, K., Hancock, K.E., Takesian, A., Kujawa, S.G., and Polley, D.B. (2022). Neural signatures of auditory hypersensitivity following acoustic trauma. *eLife* *11*, 1–33.
22. Kujawa, S.G., and Liberman, M.C. (2009). Adding insult to injury: Cochlear nerve degeneration after “temporary” noise-induced hearing loss. *J. Neurosci.* *29*, 14077–14085.
23. Chambers, A.R., Salazar, J.J., and Polley, D.B. (2016). Persistent thalamic sound processing despite profound cochlear denervation. *Front. Neural Circuits* *10*, 72.
24. Heinz, M.G., and Young, E.D. (2004). Response Growth with Sound Level in Auditory-Nerve Fibers after Noise-Induced Hearing Loss. *J. Neurophysiol.* *91*, 784–795.
25. Cai, S., Ma, W.L.D., and Young, E.D. (2009). Encoding intensity in ventral cochlear nucleus following acoustic trauma: Implications for loudness recruitment. *J. Assoc. Res. Otolaryngol.* *10*, 5–22.
26. Schrode, K.M., Muniak, M.A., Kim, Y.H., and Lauer, A.M. (2018). Central compensation in auditory brainstem after damaging noise exposure. *eNeuro* *5*, ENEURO.0250-18.2018.
27. Dau, T., Wegner, O., Mellert, V., and Kollmeier, B. (2000). Auditory brainstem responses with optimized chirp signals compensating basilar-membrane dispersion. *J. Acoust. Soc. Am.* *107*, 1530–1540.
28. Melcher, J.R., and Kiang, N.Y.S. (1996). Generators of the brainstem auditory evoked potential in cat. III: Identified cell populations. *Hear. Res.* *93*, 52–71.
29. Melcher, J.R., Guinan, J.J., Knudson, I.M., and Kiang, N.Y.S. (1996). Generators of the brainstem auditory evoked potential in cat. II. Correlating lesion sites with waveform changes. *Hear. Res.* *93*, 28–51.
30. Land, R., Burghard, A., and Kral, A. (2016). The contribution of inferior colliculus activity to the auditory brainstem response (ABR) in mice. *Hear. Res.* *341*, 109–118.
31. Schaette, R., and McAlpine, D. (2011). Tinnitus with a Normal Audiogram: Physiological Evidence for Hidden Hearing Loss and Computational Model. *J. Neurosci.* *31*, 13452–13457.
32. Møller, A.R. (1998). Neural generators of the brainstem auditory evoked potentials. *Semin. Hear.* *19*, 11–27.
33. Qiu, C.X., Salvi, R., Ding, D., and Burkard, R. (2000). Inner hair cell loss leads to enhanced response amplitudes in auditory cortex of unanesthetized chinchillas: Evidence for increased system gain. *Hear. Res.* *139*, 153–171.
34. Rummell, B.P., Klee, J.L., and Sigurdsson, T. (2016). Attenuation of Responses to Self-Generated Sounds in Auditory Cortical Neurons. *J. Neurosci.* *36*, 12010–12026.
35. Clayton, K.K., Williamson, R.S., Hancock, K.E., Tasaka, G.I., Mizrahi, A., Hackett, T.A., and Polley, D.B. (2021). Auditory corticothalamic neurons are recruited by motor preparatory inputs. *Curr. Biol.* *31*, 310–321.e5.
36. Audette, N.J., Zhou, W.X., La Chioma, A., and Schneider, D.M. (2022). Precise movement-based predictions in the mouse auditory cortex. *Curr. Biol.* *32*, 4925–4940.e6.
37. Rummell, B.P., Bikas, S., Babl, S.S., Gogos, J.A., and Sigurdsson, T. (2023). Altered corollary discharge signaling in the auditory cortex of a mouse model of schizophrenia predisposition. *Nat. Commun.* *14*, 7388.
38. Williamson, R.S., and Polley, D.B. (2019). Parallel pathways for sound processing and functional connectivity among layer 5 and 6 auditory corticofugal neurons. *eLife* *8*, e42974.
39. Egorova, M.A., Akimov, A.G., Khorunzhii, G.D., and Ehret, G. (2020). Frequency response areas of neurons in the mouse inferior colliculus. III. Time-domain responses: Constancy, dynamics, and precision in relation to spectral resolution, and perception in the time domain. *PLoS One* *15*, e0240853.
40. Williams, Z.J., He, J.L., Cascio, C.J., and Woynarowski, T.G. (2021). A review of decreased sound tolerance in autism: Definitions, phenomenology, and potential mechanisms. *Neurosci. Biobehav. Rev.* *121*, 1–17.
41. Tavassoli, T., Miller, L.J., Schoen, S.A., Jo Brout, J., Sullivan, J., and Baron-Cohen, S. (2018). Sensory reactivity, empathizing and systemizing in autism spectrum conditions and sensory processing disorder. *Dev. Cogn. Neurosci.* *29*, 72–77.
42. Wells, M.F., Wimmer, R.D., Schmitt, L.I., Feng, G., and Halassa, M.M. (2016). Thalamic reticular impairment underlies attention deficit in *Ptchd1*(Y/–) mice. *Nature* *532*, 58–63.
43. Nakajima, M., Schmitt, L.I., Feng, G., and Halassa, M.M. (2019). Combinatorial Targeting of Distributed Forebrain Networks Reverses Noise Hypersensitivity in a Model of Autism Spectrum Disorder. *Neuron* *104*, 488–500.e11.
44. Artoni, P., Piffer, A., Vinci, V., LeBlanc, J., Nelson, C.A., Hensch, T.K., and Fagioli, M. (2020). Deep learning of spontaneous arousal fluctuations detects early cholinergic defects across neurodevelopmental mouse models and patients. *Proc. Natl. Acad. Sci. USA* *117*, 23298–23303.
45. Whitton, J.P., Hancock, K.E., and Polley, D.B. (2014). Immersive audiomotor game play enhances neural and perceptual salience of weak signals in noise. *Proc. Natl. Acad. Sci. USA* *111*, E2606–E2615.
46. Burke, K., Screven, L.A., Kobrina, A., Charlton, P.E., Schrode, K., Villavisanis, D.F., Dent, M.L., and Lauer, A.M. (2022). Effects of Noise Exposure and Aging on Behavioral Tone Detection in Quiet and Noise by Mice. *eNeuro* *9*, 1–15.
47. Takahashi, N., Oertner, T.G., Hegemann, P., and Larkum, M.E. (2016). Active cortical dendrites modulate perception. *Science* *354*, 1587–1590.
48. Francis, N.A., Winkowski, D.E., Sheikhattar, A., Armengol, K., Babadi, B., and Kanold, P.O. (2018). Small Networks Encode Decision-Making in Primary Auditory Cortex. *Neuron* *97*, 885–897.e6.

49. Lobarinas, E., Spankovich, C., and Le Prell, C.G. (2017). Evidence of "hidden hearing loss" following noise exposures that produce robust TTS and ABR wave-I amplitude reductions. *Hear. Res.* **349**, 155–163.
50. Resnik, J., and Polley, D.B. (2021). Cochlear neural degeneration disrupts hearing in background noise by increasing auditory cortex internal noise. *Neuron* **109**, 984–996.e4.
51. Clayton, K.K., Asokan, M.M., Watanabe, Y., Hancock, K.E., and Polley, D.B. (2021). Behavioral Approaches to Study Top-Down Influences on Active Listening. *Front. Neurosci.* **15**, 666627.
52. Ceballos, S., Piwkowska, Z., Bourg, J., Daret, A., and Bathellier, B. (2019). Targeted Cortical Manipulation of Auditory Perception. *Neuron* **104**, 1168–1179.e5.
53. Lakunina, A.A., Menashe, N., and Jaramillo, S. (2022). Contributions of Distinct Auditory Cortical Inhibitory Neuron Types to the Detection of Sounds in Background Noise. *eNeuro* **9**, ENEURO.0264-21.2021.
54. Bajo, V.M., Nodal, F.R., Bizley, J.K., Moore, D.R., and King, A.J. (2007). The ferret auditory cortex: Descending projections to the inferior colliculus. *Cereb. Cortex* **17**, 475–491.
55. Xiong, X.R., Liang, F., Zingg, B., Ji, X.Y., Ibrahim, L.A., Tao, H.W., and Zhang, L.I. (2015). Auditory cortex controls sound-driven innate defense behaviour through corticofugal projections to inferior colliculus. *Nat. Commun.* **6**, 7224.
56. Winer, J.A., Larue, D.T., Diehl, J.J., and Hefti, B.J. (1998). Auditory cortical projections to the cat inferior colliculus. *J. Comp. Neurol.* **400**, 147–174.
57. Yan, W., and Suga, N. (1998). Corticofugal modulation of the midbrain frequency map in the bat auditory system. *Nat. Neurosci.* **1**, 54–58.
58. Melo, L.L., and Brandão, M.L. (1995). Role of 5-HT<sub>1A</sub> and 5-HT<sub>2</sub> receptors in the aversion induced by electrical stimulation of inferior colliculus. *Pharmacol. Biochem. Behav.* **51**, 317–321.
59. Melo-Thomas, L., and Thomas, U. (2015). Deep brain stimulation of the inferior colliculus: A possible animal model to study paradoxical kinesia observed in some parkinsonian patients? *Behav. Brain Res.* **279**, 1–8.
60. Schneider, D.M., Nelson, A., and Mooney, R. (2014). A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature* **513**, 189–194.
61. Vila, C.H., Williamson, R.S., Hancock, K.E., and Polley, D.B. (2019). Optimizing optogenetic stimulation protocols in auditory corticofugal neurons based on closed-loop spike feedback. *J. Neural Eng.* **16**, 066023.
62. Syka, J., and Popelář, J. (1984). Inferior colliculus in the rat: Neuronal responses to stimulation of the auditory cortex. *Neurosci. Lett.* **51**, 235–240.
63. Oberle, H.M., Ford, A.N., Czarny, J.E., Rogalla, M.M., and Apostolides, P.F. (2023). Recurrent circuits amplify corticofugal signals and drive feed-forward inhibition in the inferior colliculus. *J. Neurosci.* **43**, 5642–5655.
64. Qi, J., Zhang, Z., He, N., Liu, X., Zhang, C., and Yan, J. (2020). Cortical Stimulation Induces Excitatory Postsynaptic Potentials of Inferior Colliculus Neurons in a Frequency-Specific Manner. *Front. Neural Circuits* **14**, 591986.
65. Nwabueze-Ogbo, F.C., Popelář, J., and Syka, J. (2002). Changes in the acoustically evoked activity in the inferior colliculus of the rat after functional ablation of the auditory cortex. *Physiol. Res.* **51**, S95–S104.
66. Ehret, G., and Romand, R. (1992). Development of tone response thresholds, latencies and tuning in the mouse inferior colliculus. *Brain Res. Dev. Brain Res.* **67**, 317–326.
67. Guo, W., Chambers, A.R., Darrow, K.N., Hancock, K.E., Shinn-Cunningham, B.G., and Polley, D.B. (2012). Robustness of Cortical Topography across Fields, Laminae, Anesthetic States, and Neurophysiological Signal Types. *J. Neurosci.* **32**, 9159–9172.
68. Young, E.D., and Brownell, W.E. (1976). Responses to tones and noise of single cells in dorsal cochlear nucleus of unanesthetized cats. *J. Neurophysiol.* **39**, 282–300.
69. Zhang, G.W., Sun, W.J., Zingg, B., Shen, L., He, J., Xiong, Y., Tao, H.W., and Zhang, L.I. (2018). A Non-canonical Reticular-Limbic Central Auditory Pathway via Medial Septum Contributes to Fear Conditioning. *Neuron* **97**, 406–417.e4.
70. Kim, E.J., Horowitz, O., Pellman, B.A., Tan, L.M., Li, Q., Richter-Levin, G., and Kim, J.J. (2013). Dorsal periaqueductal gray-amygdala pathway conveys both innate and learned fear responses in rats. *Proc. Natl. Acad. Sci. USA* **110**, 14795–14800.
71. Vianna, D.M.L., and Brandão, M.L. (2003). Anatomical connections of the periaqueductal gray: Specific neural substrates for different kinds of fear. *Braz. J. Med. Biol. Res.* **36**, 557–566.
72. Li, Z., Wei, J.X., Zhang, G.W., Huang, J.J., Zingg, B., Wang, X., Tao, H.W., and Zhang, L.I. (2021). Corticostriatal control of defense behavior in mice induced by auditory looming cues. *Nat. Commun.* **12**, 1040.
73. Drager, U.C., and Hubel, D.H. (1975). Physiology of visual cells in mouse superior colliculus and correlation with somatosensory and auditory input. *Nature* **253**, 203–204.
74. Ito, S., Si, Y., Feldheim, D.A., and Litke, A.M. (2020). Spectral cues are necessary to encode azimuthal auditory space in the mouse superior colliculus. *Nat. Commun.* **11**, 1087.
75. Benavidez, N.L., Bienkowski, M.S., Zhu, M., Garcia, L.H., Fayzullina, M., Gao, L., Bowman, I., Gou, L., Khanjani, N., Cotter, K.R., et al. (2021). Organization of the inputs and outputs of the mouse superior colliculus. *Nat. Commun.* **12**, 4004.
76. Zahler, S.H., Taylor, D.E., Wong, J.Y., Adams, J.M., and Feinberg, E.H. (2021). Superior colliculus drives stimulus-evoked directionally biased saccades and attempted head movements in head-fixed mice. *eLife* **10**, 1–25.
77. Inagaki, H.K., Chen, S., Ridder, M.C., Sah, P., Li, N., Yang, Z., Hasanbegovic, H., Gao, Z., Gerfen, C.R., and Svoboda, K. (2022). A midbrain-thalamus-cortex circuit reorganizes cortical dynamics to initiate movement. *Cell* **185**, 1065–1081.e23.
78. Salkoff, D.B., Zagha, E., McCarthy, E., and McCormick, D.A. (2020). Movement and Performance Explain Widespread Cortical Activity in a Visual Detection Task. *Cereb. Cortex* **30**, 421–437.
79. Yeomans, J.S., Li, L., Scott, B.W., and Frankland, P.W. (2002). Tactile, acoustic and vestibular systems sum to elicit the startle reflex. *Neurosci. Biobehav. Rev.* **26**, 1–11.
80. Koch, M. (1999). The neurobiology of startle. *Prog. Neurobiol.* **59**, 107–128.
81. Zhou, J., Hormigo, S., Busel, N., and Castro-Alamancos, M.A. (2023). The Orienting Reflex Reveals Behavioral States Set by Demanding Contexts: Role of the Superior Colliculus. *J. Neurosci.* **43**, 1778–1796.
82. Elbaz, M., Callado Perez, A., Demers, M., Zhao, S., Foo, C., Kleinfeld, D., and Deschenes, M. (2022). A vibrissa pathway that activates the limbic system. *eLife* **11**, 1–18.
83. Cohen, J.D., and Castro-Alamancos, M.A. (2007). Early sensory pathways for detection of fearful conditioned stimuli: Tectal and thalamic relays. *J. Neurosci.* **27**, 7762–7776.
84. Xiao, C., Wei, J., Zhang, G.W., Tao, C., Huang, J.J., Shen, L., Wickersham, I.R., Tao, H.W., and Zhang, L.I. (2023). Glutamatergic and GABAergic neurons in pontine central gray mediate opposing valence-specific behaviors through a global network. *Neuron* **111**, 1486–1503.e7.
85. Sattler, N.J., and Wehr, M. (2021). A Head-Mounted Multi-Camera System for Electrophysiology and Behavior in Freely-Moving Mice. *Front. Neurosci.* **14**, 592417.
86. Parker, P.R.L., Abe, E.T.T., Leonard, E.S.P., Martins, D.M., and Niell, C.M. (2022). Joint coding of visual input and eye/head position in V1 of freely moving mice. *Neuron* **110**, 3897–3906.e5.
87. Smith, S.S., Jahn, K.N., Sugai, J.A., Hancock, K.E., and Polley, D.B. (2024). The human pupil and face encode sound affect and provide objective signatures of tinnitus and auditory hypersensitivity disorders. Preprint at bioRxiv.
88. Johnson, K.R., Tian, C., Gagnon, L.H., Jiang, H., Ding, D., and Salvi, R. (2017). Effects of Cdh23 single nucleotide substitutions on age-related

- hearing loss in C57BL/6 and 129S1/Sv mice and comparisons with congenic strains. *Sci. Rep.* **7**, 44450.
89. Kawahara, H., and Morise, M. (2011). Technical foundations of TANDEM-STRAIGHT, a speech analysis, modification and synthesis framework. *Sadhana* **36**, 713–727.
90. Guo, W., Clause, A.R., Barth-Marion, A., and Polley, D.B. (2017). A cortico-thalamic circuit for dynamic switching between feature detection and discrimination. *Neuron* **95**, 180–194.e5.
91. Li, N., Chen, S., Guo, Z.V., Chen, H., Huo, Y., Inagaki, H.K., Chen, G., Davis, C., Hansel, D., Guo, C., and Svoboda, K. (2019). Spatiotemporal constraints on optogenetic inactivation in cortical circuits. *eLife* **8**, 1–31.
92. Vincis, R., Chen, K., Czarnecki, L., Chen, J., and Fontanini, A. (2020). Dynamic representation of taste-related decisions in the gustatory insular cortex of mice. *Curr. Biol.* **30**, 1834–1844.e5.
93. Robert, B., Kimchi, E.Y., Watanabe, Y., Chakoma, T., Jing, M., Li, Y., and Polley, D.B. (2021). A functional topography within the cholinergic basal forebrain for encoding sensory cues and behavioral reinforcement outcomes. *eLife* **10**, 1–28.
94. Pachitariu, M., Sridhar, S., and Stringer, C. (2023). Solving the spike sorting problem with Kilosort. Preprint at bioRxiv.
95. Schmitzer-Torbert, N., Jackson, J., Henze, D., Harris, K., and Redish, A.D. (2005). Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience* **137**, 1–11.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Bacterial and virus strains</b>		
AAV5-Ef1a-DIO hChR2(E123T/T159C)-EYFP	Addgene	35509
<b>Deposited data</b>		
Source data	This paper; Mendeley	<a href="https://doi.org/10.17632/z6ynns9pdc.1">https://doi.org/10.17632/z6ynns9pdc.1</a>
<b>Experimental models: Organisms/strains</b>		
Mouse: C57BL/6J	The Jackson Laboratory	RRID: IMSR_JAX:000664
Mouse: B6N(Cg)-Cdh23tm2.1Kjn/Kjn	The Jackson Laboratory	RRID: IMSR_JAX:018399
Mouse: B6.129P2-Pvalbtm1(cre)Arbr/J	The Jackson Laboratory	RRID: IMSR_JAX:017320
Mouse: B6.Cg-Gt(ROSA)26Sortm32(CAG-COP4*H134R/EYFP)Hze/J	The Jackson Laboratory	RRID: IMSR_JAX:024109
Mouse: Ptchd1-KO	Gift from Michael Halassa	MGI:5792693
<b>Software and algorithms</b>		
Labview 2016	National Instruments	<a href="https://www.ni.com/en-us/shop/labview.html">https://www.ni.com/en-us/shop/labview.html</a>
Synapse	Tucker-Davis Technologies	<a href="http://www.tdt.com/component/synapse-software/">http://www.tdt.com/component/synapse-software/</a>
Kilosort 2.0	GitHub	<a href="https://github.com/MouseLand/Kilosort">https://github.com/MouseLand/Kilosort</a>
DeepLabCut	GitHub	<a href="https://github.com/AlexEMG/DeepLabCut">https://github.com/AlexEMG/DeepLabCut</a>
Python 3.8	Python	<a href="https://www.python.org/downloads/release/python-3817/">https://www.python.org/downloads/release/python-3817/</a>
Matlab 2020a	Mathworks	<a href="https://www.mathworks.com/products/matlab.html">https://www.mathworks.com/products/matlab.html</a>
Original code for data analysis	GitHub	<a href="https://github.com/Plastic-Polley-Lab/Clayton_et_al_2024_Current_Biology">https://github.com/Plastic-Polley-Lab/Clayton_et_al_2024_Current_Biology</a>
<b>Other</b>		
PXI controller	National Instruments	PXIe-8840
Siicone recording probes	Cambridge Neurotech	H3
Neurodigitizer and preamplifier	Tucker-Davis Technologies	PZ5
Data processor and real-time controller	Tucker-Davis Technologies	RZ2
Data streamer	Tucker-Davis Technologies	RS4
CMOS camera	Teledyne Dalsa	M2020
Lens	Tamron	032938
Long pass filter	MidWest Optical	LP830
Tweeter	Scan-speak	D3004/66000
Diode laser (473 nm)	Omnicon	LuxX-473
Titanium headplate	iMaterialise	Custom

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests should be directed to and will be fulfilled by the lead contact, Kameron K. Clayton ([kameron\\_clayton@meei.harvard.edu](mailto:kameron_clayton@meei.harvard.edu)).

#### Materials availability

This study did not generate new reagents.

#### Data and code availability

- Data associated with this article were deposited in Mendeley Data. DOIs are listed in the [key resources table](#).
- Original code is publicly available on GitHub and is publicly available as of the date of publication. DOIs are listed in the [key resources table](#).
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

### Animal subjects

We used 94 adult male and female mice aged 8–12 weeks. Mice were maintained on a 12/12 hr light/dark cycle and all experiments performed during the dark cycle, with food and water available *ad libitum* in the home cage. Mice were housed individually following major survival surgery. All procedures were approved by the Massachusetts Eye and Ear Animal Care and Use Committee and followed the guidelines established by the National Institute of Health for the care and use of laboratory animals.

For simultaneous videography and acoustic startle testing, 8 C57 x Cdh23 mice were used. C57 mice were crossed with homozygous Cdh23 mice to prevent the precocious high frequency hearing loss typically found in C57 mice, which can begin as early as 12 weeks post-natal.<sup>88</sup> For videography experiments with temporally modulated stimuli, an additional 16 C57BL6 x Cdh23 mice were used, 8 of which were also used for visual stimulation and whisker trimming experiments. For noise exposure experiments, we used a total of 20 C57BL6 x Cdh23 (10 exposed/8 sham). Auditory brainstem responses were collected in a subset of 9 noise-exposed mice, with one mouse excluded from further analysis because of the absence of wave IV in pre-exposure measurements. Single-unit electrophysiology with simultaneous videography was performed in 5 C57BL6 x Cdh23 mice. Experiments to verify cortical silencing via PV-mediated inhibition were performed in 4 PV-Cre mice with unilateral channelrhodopsin expression in right ACTx. Videography with optogenetic stimulation was performed in 7 PV-Cre x Ai32 mice and 3 PV-Cre mice with injected with virus to express channelrhodopsin bilaterally in ACTx. Experiments characterizing sound-evoked facial movements in the Ptchd1-KO mouse were performed in 11 KO mice and 15 C57 littermate controls.

## METHOD DETAILS

### Surgical preparation

Mice were anesthetized with isoflurane in oxygen (5% induction, 1.5–2% maintenance). Body temperature was maintained at 36.6°C with a homeothermic blanket system (FHC). Lubricating ointment was placed on the eyes. Lidocaine hydrochloride (0.1 mL) was administered subcutaneously to numb the scalp. For analgesia, Buprenex (0.05 mg/kg) and meloxicam (0.1 mg/kg) were administered subcutaneously at the beginning of the procedure and again 24 and 48 h from the initial dosing. Following surgery, the mice were transferred to a heated recovery chamber.

Following repeated serial applications of Betadine and 70% ethanol, the skin overlying the dorsal cranium was retracted and the periosteum was removed. Etchant (C&B metabond) and 70% ethanol was applied to prepare the exposed skull surface. For mice not undergoing optogenetics experiments, a custom titanium headplate (iMaterialise) was affixed to the skull with dental cement (C&B metabond).

For whisker trimming experiments, mice were briefly anesthetized with isoflurane prior to videography recording and the whiskers were removed bilaterally using surgical scissors. Care was taken to trim the whiskers as close as possible to the epidermis and microvibrissae were also removed.

### Optical access to auditory cortex

To prepare mice for bilateral optogenetic stimulation, the skull overlying each auditory cortex was made optically transparent. First, a layer of super glue (Krazy-Glue) was applied and allowed to dry, followed by layers of clear cement (C&B metabond) and nail polish (L.A. Colors). While the nail polish dried, black plastic casings (Freelin-Wade) were affixed around the transparent portion of the skull to allow for interfacing with optic fiber patch cables. Once the casings were affixed, a custom titanium headplate was affixed to the dorsal surface of the skull with cement.

For PV-Cre mice only, channelrhodopsin was expressed in PV neurons by injecting AAV5-EF1a-DIO-ChR2 into ACTx. Briefly, two small burr holes were made in the skull overlying ACTx in each hemisphere using a 31-gauge needle, 1.5–2.25 mm rostral to the lambdoid suture. The viral solution was backfilled into a pulled glass pipette (Drummond, Wiretrol II) and a precision injection system (Drummond, Nanoject III) was used to deliver 200 nL of virus per injection site at a rate of 9 nL/min, 0.3 mm below the pial surface. The pipette was left to rest for at least 10 min following the end of the injection and burr holes were filled with KWIK-SIL (WPI). Injections took place immediately prior to the skull clearing procedure.

### Videography

A CMOS camera (Teledyne Dalsa, model M2020) equipped with a lens (Tamron 032938) and a long-pass filter (MidWest Optical, LP830, 25.5nm cutoff) was positioned approximately 25 cm to the right of the animal's face, and illuminated by an array of infrared LEDs (Vishay, 850 nm). A white LED (Thorlabs, MBB1F1) was used to provide ambient illumination sufficient to keep the pupil diameter in an intermediate range. Video recordings of the face and foot were acquired with a 512 x 512 pixel resolution at 150 Hz (for data shown in [Figures 1, 2, 3, 4, and 5](#)) or 30 Hz (for data shown in [Figures 6 and 7](#)).

Auditory stimuli were generated with a 24-bit digital-to-analog converter (National Instruments model PXI-4461), amplified (Samson, 120a Power Amplifier), and presented via a tweeter (ScanSpeak) positioned 25 cm from the left ear. Speaker output calibrated with a 1/4" prepolarized microphone (PCB Electronics). Visual stimuli were presented on a 5" display with 800 x 480 resolution

(Adafruit TFP401). Mice were head-fixed atop an acrylic plate resting on three piezoelectric force transducers (PCB Piezotronics) coupled to a summation amplifier, allowing measurement of the downward force caused by skeletal muscle contraction. Mice were allowed to acclimate for 15 min before recordings began

### Stimulus generation and presentation

Each stimulus was repeated in 20 trials in a pseudo-random order separated by an intertrial interval duration selected at random from a truncated exponential distribution to produce a flat hazard function. Broadband noise bursts (50 ms duration, 5 ms raised cosine onset and offset ramps) were presented between 15 to 115 dB SPL in 10 dB steps with a 10-20s intertrial interval. Visual stimuli were full field, full contrast sinusoidal gratings (500 ms duration, 0.4 cycles/degree spatial frequency, 2 Hz velocity) presented to the eye contralateral to the camera at multiple orientations (0°, 45°, 90°, 180°) with a 7 – 11 s inter-trial interval. Spatial tuning of sound-evoked facial movements was tested by presenting a 60 dB SPL, 50 ms noise burst at 90° and -90° in azimuth relative to the mouse's head with a 7 to 11 s inter-trial interval. For gaps in noise, silent gaps (30 - 500 ms in duration, 0.1ms onset and offset ramps) were inserted in 50 dB SPL continuous broadband noise with a 14 - 19 s inter-trial interval. For FM sweeps, sequences of 6 sweeps (200 ms duration, ±20 octaves/s between 4-64 kHz, 5 ms raised cosine ramps) were presented at 1 - 4 Hz in 0.5 Hz steps with a 7 - 11 s inter-trial interval. Speech stimuli were 200 ms tokens produced by an adult female speaker resynthesized to be four octaves higher than the original source material with the TANDEM-STRAIGHT vocoder,<sup>89</sup> as originally synthesized by Chambers et al.<sup>18</sup> Speech tokens (70 dB SPL) were presented in trains of 6 tokens at 1 Hz in continuous broadband noise (10 - 50 dB SPL in 10 dB increments) or in silence, with a 7-11s intertrial interval. To contrast tone and broadband noise responses, tone bursts (50 ms duration with 5 ms raised cosine ramps at 8 and 32 kHz or broadband noise bursts were presented from 30 - 90 dB SPL in 20 dB increments with a 7 to 11 s intertrial interval. All stimuli were normalized for total energy. Octave-band noise (OBN) bursts were made by applying a 4<sup>th</sup> order Butterworth filter centered at 8 or 32 kHz to broadband noise. OBN noise bursts (50 ms duration with 5 ms raised cosine ramps) were presented from 30 - 100 dB SPL in 10 dB increments with a 7 - 11 s inter-trial interval.

To understand if sound stimuli might be driving direct whisker motion, we used the spherical wave equation to estimate the air particle displacement at the position of the mouse's face relative to the speaker. The spherical wave equation was used as follows:

$$x = \left| \frac{v}{j\omega} \right|$$

where  $x$  is the absolute particle displacement,  $\omega$  is the frequency in radians and  $v$  is the particle velocity given by:

$$v = \frac{P}{Z}$$

where  $P$  is the pressure in pascals and  $Z$  is the acoustic impedance which for a spherical wave is calculated as:

$$Z = \frac{\rho c}{1 + \frac{1}{j\omega r * \frac{1}{c}}}$$

where  $c$  is the speed of sound (345 m/s),  $\rho$  is the density of air (1.12 kg/m<sup>3</sup>), and  $r$  is the distance between the sound source and the mouse's head (25 cm).

### High-frequency noise exposure

To induce high-frequency SNHL, OBN at 16 - 32 kHz was presented at 103 dB SPL for 2 h. Exposure stimulus was delivered via a tweeter fixated inside a custom-made exposure chamber (51 × 51 × 51 cm). The interior walls of the acoustic enclosure joined at irregular, non-right angles to minimize standing waves. Additionally, to further diffuse the high-frequency sound field, irregular surface depths were achieved on three of the interior walls by attaching stackable ABS plastic blocks (LEGO). Prior to exposure, mice were placed, unrestrained, in an independent wire-mesh chamber (15 × 15 × 10 cm). This chamber was placed at the center of a continuously rotating plate, ensuring mice were exposed to a relatively uniform sound field. Sham-exposed mice underwent the same procedure except that the exposure noise was presented at an innocuous level (70 dB SPL). All sham and noise exposures were performed at the same time of day.

### Cochlear function testing

Animals were anesthetized with ketamine (120 mg/kg) and xylazine (12 mg/kg), placed on a homeothermic heating blanket during testing, and administered half the initial ketamine dose as a booster when required. Acoustic stimuli were presented via in-ear acoustic assemblies consisting of two miniature dynamic earphones (CUI CDMG15008-03A) and an electret condenser microphone (Knowles FG-23339-PO7) coupled to a probe tube. Stimuli were calibrated in the ear canal in each mouse before recording. The ABR was measured with tone bursts (5 ms with 0.5 ms raised cosine ramps at 8, 11.3, 16, 22.6, and 32 kHz delivered at 26.99 Hz from 20-100 dB SPL in 5 dB increments) and the same OBN bursts described above delivered at 10.02 Hz. Recordings were made with one transdermal electrode behind the right pinna and one electrode attached to a silver wire (A-M Systems) placed on the surface of the brain at vertex through a burr hole and cemented in place during the initial headplate surgery. Threshold was

determined as the lowest intensity which elicited a repeatable waveform. Positive and negative peaks of each ABR wave were quantified as the peak to trough amplitude of each wave, subtracted by the peak to trough amplitude of the pre-stimulus baseline to correct for the measurement noise floor. ABR testing was performed 1 week before noise or sham exposure and again two weeks after exposure, following the final behavioral measurement.

### Cortical electrophysiology

#### **Preparation for acute insertion of high-density probes in awake, head-fixed mice**

A ground wire was placed over the left occipital cortex through a small burr hole during the initial head plate attachment surgery. On the day of recording, the mouse was briefly anesthetized with isoflurane in oxygen (5% induction, 2% maintenance) and a scalpel was used to make a small (1 × 1 mm) craniotomy over right auditory cortex, centered on the temporal ridge between 1.5 and 2.5 mm anterior from the lambdoid suture. A well was constructed around the craniotomy using UV-cured composite (Flow-It ALC) was filled with lubricating ointment (Paralube Vet Ointment), after which isoflurane was discontinued and the mouse was moved to a body cradle where the animal's head was immobilized by attaching head plate to a fixation post in a dimly lit double walled acoustic chamber. Recordings began 30 min after the cessation of isoflurane to allow for full recovery from anesthesia. Following each experiment, the recording chamber was flushed with saline, lubricating ointment was reapplied, and capped with UV-cured composite.

#### **Extracellular recordings**

A single shank 64-channel silicon probe (Cambridge Neurotech: H3, 20 μm between contacts) was positioned perpendicular to the cortical surface using a micromanipulator (Narishige) and inserted using a hydraulic micromanipulator (Narishige). The probe was advanced at 100 μm/s until the probe tip was approximately 1.3–1.4 mm below the cortical surface. The probe was allowed to settle for 15 min before recordings began. At the beginning of each recording, noise bursts (50 ms duration, 70 dB SPL, 50 repetitions) were presented to confirm the depth of the probe relative to auditory cortex using the current source density.<sup>35,90</sup> Additionally, frequency-receptive fields were derived to confirm the placement of the probe within the primary auditory cortex, using the tonotopic reversal which marks the rostral border of A1.<sup>67</sup> We simultaneously recorded facial videography and laminar electrophysiology as mice were presented with frequency-modulated sweeps (4–64 kHz, 80 and -80 oct/s, 50 ms duration, 30–70 dB SPL in 10 dB increments, 60–80 repetitions) from a free-field speaker placed approximately 10 cm from the left ear.

### Optogenetic stimulation

Neurons transduced with channelrhodopsin were activated through the optically cleared skull with light (10 mW at fiber tip, 500 ms, 20 Hz pulse rate, 25 ms pulse width, terminating with 100 ms linearly decreasing ramp) that have been shown to trigger robust inhibition of pyramidal neuron spiking without rebound excitation.<sup>91</sup> Blue light was delivered by two 473 nm diode lasers (Omicron LuxX) via optic fibers terminating in ferrules (Doric, 0.2 mm diameter, 0.22 NA), which were fit with mating sleeves to create a snug interface with the skull-mounted plastic casings. Laser onset preceded sound onset by 100 ms. Laser stimulation occurred in half of trials, which consisted of 50 ms noise bursts presented from 35 to 95 dB SPL in 10 dB increments, with a 6–10 s inter-trial interval. To control for the potential effects of ACTx PV cell activation alone on facial movements, experiments were performed in an additional cohort of mice where laser stimulation was delivered without sound presentation. For single-unit electrophysiology experiments, the optic fiber was placed directly above the exposed surface of the ACTx and laser-evoked responses were recorded without sound stimulation (identical laser stimulation parameters as for behavior, 20 trials per stimulus).

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Video processing and analysis

#### **Facial movement energy**

A region of interest (ROI) was manually drawn on the rostral cheek, just caudal to the vibrissae array. FME was defined as the sum of the absolute difference in intensity between consecutive frames for each pixel within the ROI.<sup>13,92</sup> FME was z-scored with respect to the mean and standard deviation of the session.

#### **Markerless behavior tracking**

DeepLabCut was used to track the nose, ear, jaw, hind paw, eyelid diameter, and pupil diameter. The anterior tip of the nose, posterior edge of the pinna, jaw, and the metatarsal joint of the hindfoot were labeled with single markers. The eyelid and the pupil were labeled with 8 markers that spanned the four cardinal and four intracardinal compass points. Separate DeepLabCut models were used for points on the face and the foot. For the face model, 300 frames from 30 mice were used to train the model (model originally used in Robert et al.<sup>93</sup>). For the foot model, 40 frames from 6 mice were used. For both models, a ResNet-101 based neural network was trained on 95% of the labeled frames for 1,030,000 training iterations, using the default parameters in DeepLabCut.

Any time point for which model tracking likelihood dropped below 90% was discarded and interpolation was performed between neighboring frames. For single-point tracking, movement amplitudes were taken as the square root of the sum of the squared x and y velocities. For eyelid and pupil diameter, tracked points were used to fit an ellipse using a least-squares criterion and calculating the long axis diameter. Because the eyelid was widest at rest and narrowed with sound stimulation, we took the absolute value of the eyelid diameter so that movement would be positive going. All movement traces were z-scored with respect to their mean and standard deviations for the session.

### Quantification of sound-evoked movements

Sound-evoked response amplitude for all movements were taken as the mean of the 5 frames surrounding the peak response within 1 s of stimulus onset. To determine the threshold of sound-evoked movement, we performed a paired t-test comparing the mean of 1 s pre-stimulus and the post-stimulus response across trials. Threshold was defined as the lowest intensity at which reliably elicited a significant response; all levels above threshold were required to have a significant response. Latency was defined as time to the half-maximum of the response. Intertrial variance was defined as the coefficient of variation, the standard deviation of trial-by-trial sound-evoked responses divided by the mean response for each subject individually. As computing the coefficient of variation for intersubject variance across subjects would give a single value for each measured effector, we treated each trial as an individual measurement and computed the standard deviation of the sound-evoked response divided by the mean response across subjects for each trial separately. The signal detection metric  $d'$  was computed to compare the spontaneous rate of z-scored FME during the intertrial interval (0-1 s prestimulus and 1 s from stimulus onset to the end of the trial, analogous to estimating the false alarm rate) to the rate of z-scored FME during and immediately following sound stimulation (0 to 500 ms from sound onset, analogous to the hit rate) as follows:

$$d' = \frac{2 (\mu FME_{evoked} - \mu FME_{spontaneous})}{(\sigma FME_{evoked} + \sigma FME_{spontaneous})}$$

where the mean and standard deviation of FME rates are taken across trials and computed separately for each intensity tested.

To quantify the rhythmic entrainment of facial movements, we used the fast Fourier transform to compute the power spectral density (PSD) within the stimulus period, which varied in duration depending on the repetition rate from 1.5 to 6 s. The dB SNR metric was computed as 10 times the common logarithm of the PSD amplitude at the stimulus repetition rate divided by the average PSD amplitude at neighboring frequencies. To decode syllable identity trial by trial, a support vector machine with a linear kernel was trained for each mouse separately using time series responses to each syllable in quiet. All data was first transformed using principal components analysis and only the principal components which cumulatively explained 80% of the variance were included to prevent overfitting. We used 10-fold cross-validation to train and test the classifier and repeated this process for 1000 iterations using different random train/test splits. To estimate the classification noise floor, we shuffled the trial-by-trial syllable labels and repeated the training and testing process. Models were fit using the Matlab function 'fitcsvm'. The same training procedure was applied for cross-mouse decoding, but the support vector machine was then tested on data from each other mouse. Cross-mouse decoding was quantified as the average performance of each mouse tested on each other mouse's decoder. For FME response tracking over time, fold change was calculated as the response for each day divided by the baseline response, which included the two pre-exposure days, for 0 and 10 dB re: threshold.

### Acoustic startle reflex measurement

The acoustic startle reflex is highly stereotyped triphasic signal which corresponds to the contraction of skeletal muscles.<sup>14</sup> We quantified the startle reflex amplitude as the peak to trough of the response in the 1 s following stimulus onset and computed measures of threshold, latency and variance as we did for facial videography.

### Electrophysiology acquisition and online analysis

Raw neural signals were acquired at 24.4 kHz and digitized at 32-bit resolution (PZ5 Neurodigitizer, RZ2 BioAmp Processor, RS4 Data Streamer; Tucker-Davis Technologies). For online data visualization, the common mode, channel-averaged, signal was subtracted from the raw signals for all channels to eliminate artifacts shared across all channels. To examine multiunit activity, signals were band-pass filtered (300-3000 Hz, second-order Butterworth filters) and the threshold for spike detection was set as a negative deflection greater than 3.5 standard deviations above the background hash. Following notch filtering at 60 Hz and downsampling to 1 kHz, the CSD was computed as the second spatial derivative of the local field potential. Signals were spatially smoothed across channels using a 5-point Hanning window to mitigate potential artifacts induced by impedance mismatches between neighboring channels. The layer 4/5 boundary (0.5 mm from the pial surface) was identified using the noise-evoked CSD.<sup>35,90</sup>

### Single-unit identification and analysis

Kilosort 2.0 was used to sort spikes into single-unit clusters.<sup>94</sup> Single-unit classification was based on the presence of a well-defined refractory period in the interspike interval histogram and an isolation distance (>10) which indicated that the cluster was well separated from noise.<sup>95</sup> Units were classified as RS if the peak-to-trough delay of its mean spike waveform was greater than 0.6 ms.

### Statistical analysis

All statistical analysis was performed in Matlab. We used the standard  $p < 0.05$  threshold for assigning statistical significance. Multiple post-hoc comparisons were corrected with the Holm-Bonferroni correction.