University of Oxford

Medical Sciences Division

The Influence of Whisker Stimulation and Active Whisking on Auditory Cortical Processing and Perception

Paul Zimmer-Harwood

Thesis submitted for the Degree of Doctor of Philosophy

April 2025

Abstract

Sensory processing is profoundly influenced by its behavioural context, and the primary sensory cortices are recognised as dynamic integrators of multisensory and motor-related inputs. In rodents, whiskers are critical for near-field exploration, and mounting evidence demonstrates that somatosensory inputs can modulate auditory processing even at early subcortical stages. However, how these interactions impact perception and how behavioural relevance and motor activity reshape cross-modal influences in the auditory cortex is not yet known.

In this thesis, I explore the impact of passive whisker stimulation and active whisking on auditory cortical processing and perception in mice, and ask how behavioural demands and stimulus relevance dynamically tune tactile–auditory interactions. Firstly, using a combination of two-photon calcium imaging and near-threshold behavioural tasks, I confirm that whisker stimulation consistently suppresses sound-evoked activity in the auditory cortex during goal-directed behavior. Despite this suppression, detection performance remains stable across both tone and broadband noise stimuli, while frequency discrimination improves during simultaneous whisker stimulation, suggesting that divisive suppression sharpens spectral representations without impairing detectability.

Building on these results, I examine how the behavioural relevance of whisker input modulates cross-modal suppression. Mice trained to detect whisker stimulation exhibit reduced somatosensory-auditory suppression compared to those detecting sounds, contrary to the hypothesis that reward association would amplify tactile gating. Instead, whisker-driven suppression was strongest when whisker input was behaviourally irrelevant, suggesting instead, that suppression at the thalamocortical level reflects learned prioritisation of sensory modalities rather than fixed circuit properties.

Finally, I demonstrate that active whisking itself modulates auditory cortical responses in a context-dependent manner. A subset of neurons responds robustly to spontaneous whisking—accompanied by pupil dilation and brain state changes—but not to sound-triggered whisking of similar kinematics, revealing that brain state transitions, rather than motor output per se, might shape auditory cortical activity. Whisking-sensitive neurons encode spontaneous and stimulus-evoked movements differently, highlighting the flexible integration of motor, arousal, and sensory signals in the auditory cortex.

Together, these studies establish that subcortical tactile-auditory suppression is not static but dynamically reweighted by behavioural context, behavioural state and task demands.

Acknowledgements

The experience of being a PhD student depends, more than anything, on one's supervisors. While many other PhD students in my social circle struggled with unavailable or unsupportive supervisors - and sometimes questioned their decision to pursue a PhD as a result - I could not feel more differently. I am deeply grateful that Johannes and Andy chose me as their student.

Through the beginning of the COVID-19 pandemic, when uncertainty led me to return to Germany, and during other periods when opportunities in my life temporarily took priority over my studies, I always knew you had my back. You provided the best guidance I could have imagined, were always available when I had questions, and took time and care to ensure I was on the right track. I never felt that I had to hide mistakes; you would come back to the lab late at night when I had an issue, without flinching or showing the slightest sign of annoyance.

I believe your supervision went far beyond what anyone could reasonably expect. To me, you are role models as researchers, and you have shown me what genuine supervision and support can look like. I thoroughly enjoyed my PhD experience, and that would not have been possible without you.

I want to thank my parents and my brother, who have always trusted and supported my decisions. From an early age, you nurtured my interest in science and raised me to find joy and wonder wherever I go. Even though we have missed each other throughout the many years I have lived far away since I finished school, I hope you know that it is the warmth and safety I had while growing up that gave me the confidence to take on new challenges and to trust that I would always be fine.

To my friends - or replacement family - in Oxford, to Khadija, Antonis and Lisa: all your "distractions", my favourite memories from Oxford, will be the bits I remember most. It's not the full weeks in the lab,

but one hour going for a run in port meadow at sunrise, or thirty minutes with a surprise moussaka, that will stay with me.

To everyone else across the channel, in Zurich, Bielefeld, Berlin, Amsterdam - you know who you are - thanks for bearing with me. I can't wait to spend more time with you all again and thank you for "Phd Heaven", Cologne.

Thank you to the 1Day Crew for letting me finish this thing.

Thank you to all the mice who have given their lives to science.

And finally, to everyone in the lab: thank you for welcoming me, for all the help at the microscope and the whiteboard, for sharing your tricks, your data, your code, your snacks. It was a pleasure to work with you.

Table of Contents

I General Introduction

1.1 Broad Context of Multisensory Processing

Multisensory integration refers to the process by which the brain combines information from different sensory modalities - such as vision, hearing, and touch - to create a more robust or efficient perceptual experience than would be possible from a single modality alone (Ernst & Bülthoff, 2004; Ghazanfar & Schroeder, 2006). This integration often reduces ambiguities, speeds up reaction times, and improves overall perceptual accuracy, particularly in complex or noisy environments (Diederich & Colonius, 2004; Ernst & Bülthoff, 2004; Stein & Stanford, 2008). Traditionally, research has emphasized audio-visual interactions, demonstrating how vision can refine auditory cues, for instance, by helping to localise sound sources or by enhancing speech understanding through lip-reading (Opoku-Baah et al., 2021; Schroeder et al., 2008). Although early work recognized the importance of audio-tactile integration for near-field events (Moore & King, 1999; Von Bekesy, 1959), contemporary research continues to highlight how critical these interactions become when objects approach the head or body (Lohse et al., 2022; Serino, 2019). Unlike vision, which may be limited in low-light conditions or obstructed environments, touch and audition can together provide important information about the physical properties and location of objects we actively manipulate or come into close contact with (Occelli et al., 2011; Soto-Faraco & Deco, 2009). Indeed, many everyday tasks - handling tools, playing musical instruments, or even simply rubbing our hands, produce correlated tactile and auditory feedback. Understanding how the brain merges these signals is essential for comprehending real-world behaviours, particularly in species that rely on vibrissae (whiskers) or other tactile appendages to explore their immediate surroundings (Godenzini et al., 2021; Sofroniew & Svoboda, 2015).

1.2 Audio-Tactile Cues in Peripersonal Space

When a sound source lies within roughly one acoustic wavelength of the listener, the physics of the sound field changes dramatically compared with the far field. Whereas in the far field the sound level usually follows an inverse-square law drop-off with distance in which listeners can depend on stable binaural cues such as interaural time and level differences, near-field propagation is dominated by complex relationships between distance, sound-pressure level and frequency content (Brungart & Rabinowitz, 1999; Kolarik et al., 2016; Moore & King, 1999). Small changes in distance can alter reverberant-to-direct energy ratios and even flip the sign of interaural level differences, providing salient information that an object is looming towards the head and may warrant defensive or reaching responses (M. S. A. Graziano & Cooke, 2006; Seifritz et al., 2002).

This acoustic near-field region largely overlaps with peripersonal space, the multisensory zone surrounding the body within which tactile, visual and auditory inputs are bound together to guide immediate actions (Serino, 2019). Neurophysiological work in non-human primates shows that fronto-parietal neurons represent sounds differently when they arise within reachable distance (M. S. Graziano et al., 1999; Serino, 2019), and human behavioural studies demonstrate tighter coupling between touch and audition for sounds delivered close to the head (Bruns et al., 2011; Teneggi et al., 2013).

Rodents offer a complementary perspective. Mice and rats explore nearby objects primarily with their whiskers, using them for texture discrimination, obstacle avoidance and social investigation (Sofroniew & Svoboda, 2015). Recent work shows that whisker-related somatosensory signals powerfully modulate auditory processing from the inferior colliculus (IC) to the auditory cortex, particularly suppressing near-field tone responses in the lemniscal pathway while sometimes facilitating

activity in non-lemniscal nuclei (Lohse et al., 2021; Rao et al., 2014). These findings suggest a dedicated subcortical mechanism for prioritising tactile information from objects that impinge on the whisker array (Lohse et al., 2022).

Recent evidence reveals a further dimension to this multisensory scene: when whiskers strike a surface they produce broadband sounds (10–60 kHz) that rise well above background noise and fall squarely within the mouse hearing range. Extracellular recordings show that such whisker-generated sounds drive vigorous spiking in the IC and in auditory-cortex neurons even after tactile input is abolished by cutting the infra-orbital nerve. At the behavioural level, head-fixed mice can learn to identify objects purely from these self-generated acoustic signatures: performance collapses when the sounds are masked by white noise (Efron et al., 2025). Thus, active whisking supplies its own near-field acoustics, furnishing rodents with an additional auditory channel that complements touch and may enhance object recognition in dim light or cluttered environments.

Taken together, peripersonal space in rodents is enriched not only by exafferent auditory cues from external sources but also by the reafferent sounds produced by the animals' own sensory organs. To understand the mechanisms behind audio-tactile integration, we next provide an overview of the underlying sensory pathway.

1.3 The Auditory Hierarchy and Cortical–Subcortical Auditory-Tactile Interactions

In mammals, the cochlear nucleus (CN) is the first central station for sound-frequency organisation, receiving afferents directly from the auditory nerve (Watson et al., 2011; Winer &

Schreiner, 2005). Traditionally viewed as a purely acoustic relay, the CN (dorsal and ventral divisions) can already integrate non-auditory inputs, such as trigeminal signals tied to self-generated sounds (Ansorge et al., 2021; Shore & Zhou, 2006; Singla et al., 2017; Wu et al., 2015) and may help distinguish external stimuli from internally generated noise, suggesting that contextual filtering emerges at the earliest central stage.

After traversing other brainstem nuclei, such as those comprising the superior olivary complex, ascending auditory signals converge in the inferior colliculus (IC) of the midbrain (Gruters & Groh, 2012; Lesicko et al., 2016). The IC receives multisensory and motor-related inputs, especially within its shell or lateral cortex (Bajo & King, 2012; Wu et al., 2015).

Whisker-related activity reaching the IC shell may influence the timing and coordination of ascending auditory responses, so that these responses occur in synchrony with whisker contact or body movements (Lesicko et al., 2016; Lohse et al., 2021). By integrating tactile cues at this relatively early stage, the auditory system may be able to incorporate movement context into sound processing before signals proceed towards the thalamus.

The medial geniculate body (MGB) is the principal auditory relay to the cortex, yet recent research shows it functions as a dynamic, context-sensitive filter rather than merely a passive conduit (Bordi & LeDoux, 1994; Guillery et al., 2001; Kimura & Imbe, 2018; Lohse et al., 2021; Suga et al., 2002; Tang et al., 2012). Its ventral division (MGBv) relays tonotopically organized acoustic signals to primary auditory cortex (A1), whereas the dorsal (MGBd) and medial (MGBm) divisions also receive multisensory inputs—potentially linked to limbic or higher cognitive functions (Bartlett, 2013; Lohse et al., 2021). Ongoing corticofugal feedback can re-tune MGB responses according to stimulus salience, motor state, or attentional goals (Kimura & Imbe, 2018). Moreover, lemniscal subdivisions like MGBv

may exhibit suppressive influences from tactile inputs, while non-lemniscal divisions (e.g., MGBm/d) can facilitate auditory processing or integrate with limbic networks (Lohse et al 2021). This duality may allow the MGB to regulate the signal-to-noise ratio of auditory information reaching the cortex, effectively determining whether whisker-driven cues overshadow or complement coincident sounds.

Projections from the MGB to A1 mark the onset of higher-level auditory processing; however, the signals arriving in A1 already incorporate subcortical computations (Lohse et al., 2024) and non-auditory modulations from earlier stages. Far from a unidirectional pathway, corticofugal loops through layers 5 and 6 neurons maintain a continuous dialogue with subcortical stations (Bajo & King, 2012; Mo & Murray Sherman, 2019; Suga et al., 2000), potentially shaping coding precision and highlighting salient auditory features. This interplay between cortex and subcortex is evident in the descending projections from primary and higher auditory fields, which target the IC, the MGB, and even the cochlear nucleus (Bajo & King, 2012; Tang et al., 2012). These feedback mechanisms can re-tune neuronal frequency selectivity, adjust temporal response profiles, prioritize task-relevant stimuli, and provide a substrate for learning-dependent plasticity. At the same time, subcortical stations like the MGB (McGinley et al., 2015; Williamson et al., 2015), (Yang et al., 2020)are suppressed during locomotion and may suppress predictable, self-generated sounds while enhancing faint or near-threshold cues (Godenzini et al., 2021; Rao et al., 2014). The precise origin of these suppressive effects on subcortical auditory responses remains uncertain; however, they may arise from somatosensory signals activated during locomotion or descending projections originating in motor cortical regions (Lohse et al., 2022; Olthof et al., 2019)

Taken together, from the CN onward, ascending and descending pathways form a multilayered, context-aware filtering system that enables rodents and other animals to integrate whisker-based somatosensory information with audition. This interconnected architecture may improve detection of

near-field stimuli and elevate the salience of tactile inputs when whisker contact is especially pertinent. These subcortical–cortical interactions are therefore essential for appreciating audio–tactile convergence and its influence on behaviour, even before sensory signals reach higher-level perceptual processing.

1.4 Whisker-Driven Suppression, the S1–IC–MGB–A1 Loop and co-existence of facilitative interactions

This thesis is inspired by the work of Lohse et al. (2021) who demonstrated the origin of suppression of sound-evoked responses in lemniscal auditory nuclei, most notably the ventral division of the medial geniculate body (MGBv) and primary auditory cortex (A1), when whiskers are stimulated. In mice, whisker input exerts a divisive gain on neuronal sound frequency–response curves, scaling down tone-evoked amplitudes, especially near each neuron's best frequency (Lohse et al., 2021). Similar phenomena appear in ferrets (Meredith & Allman, 2015), rats (Rao et al., 2014), and cats (Khorevin & Shelest, 1998; Rao et al., 2014), suggesting a broadly conserved principle of tactile-driven downscaling in the auditory pathway.

Mechanistically, whisker signals activate GABAergic circuitry in the inferior colliculus (IC), which in turn suppresses MGBv relay cells projecting to A1 (Kimura & Imbe, 2018; Lohse et al., 2021) (Figure 1). Such gating could prioritize tactile cues during close-range exploration, minimizing auditory "distractions" at the very moment whiskers contact an object (Rao et al., 2014). However, the behavioural repercussions of this suppression—especially for tasks that require balancing whisker-based and auditory cues—remain a mystery.

Lohse et al. 2021 identified a descending somatosensory–auditory circuit that partly explains these suppressive effects. In this pathway, layer 5 neurons in the primary somatosensory cortex (S1) activate GABAergic neurons in the lateral cortex of the inferior colliculus (IC). These neurons, in turn, suppress sound-evoked activity in the medial geniculate body (MGB), likely through inhibitory projections or reduced excitatory drive within the IC. This *corticocolliculothalamic* pathway thus enables S1 to impose a whisker-driven suppression on ascending auditory signals, such that inputs reaching primary auditory cortex (A1) have already undergone cross-modal modulation at the thalamic level.

In fact, not all audio-tactile interactions at this level are inhibitory. Concurrently, an alternative direct projection from S1 to the medial sector of the MGB can facilitate auditory responses destined for non-A1 pathways, such as limbic or higher-order sensory targets, allowing facilitation and suppression to coexist, perhaps according to the salience of whisker contact or the behavioural context (Lohse et al., 2021) .

Building upon the discussion of whisker-generated auditory signals in Section 1.2, recent findings by Efron et al. (2025) demonstrate that self-generated broadband acoustic stimuli arising specifically from whisker-object interactions robustly activate neurons in the inferior colliculus shell (ICs) and subsequently drive spiking activity in auditory cortical units, even without direct trigeminal input, while externally presented broadband sounds are subject to tactile-driven suppression when whiskers were stimulated externally (Lohse et al., 2021). Thus, while externally generated auditory signals may be selectively suppressed by descending tactile pathways to minimise distraction, self-generated broadband auditory signals from whisker contacts appear to be processed differently—potentially amplified or selectively preserved due to their behavioural relevance in conveying immediate spatial or textural information.

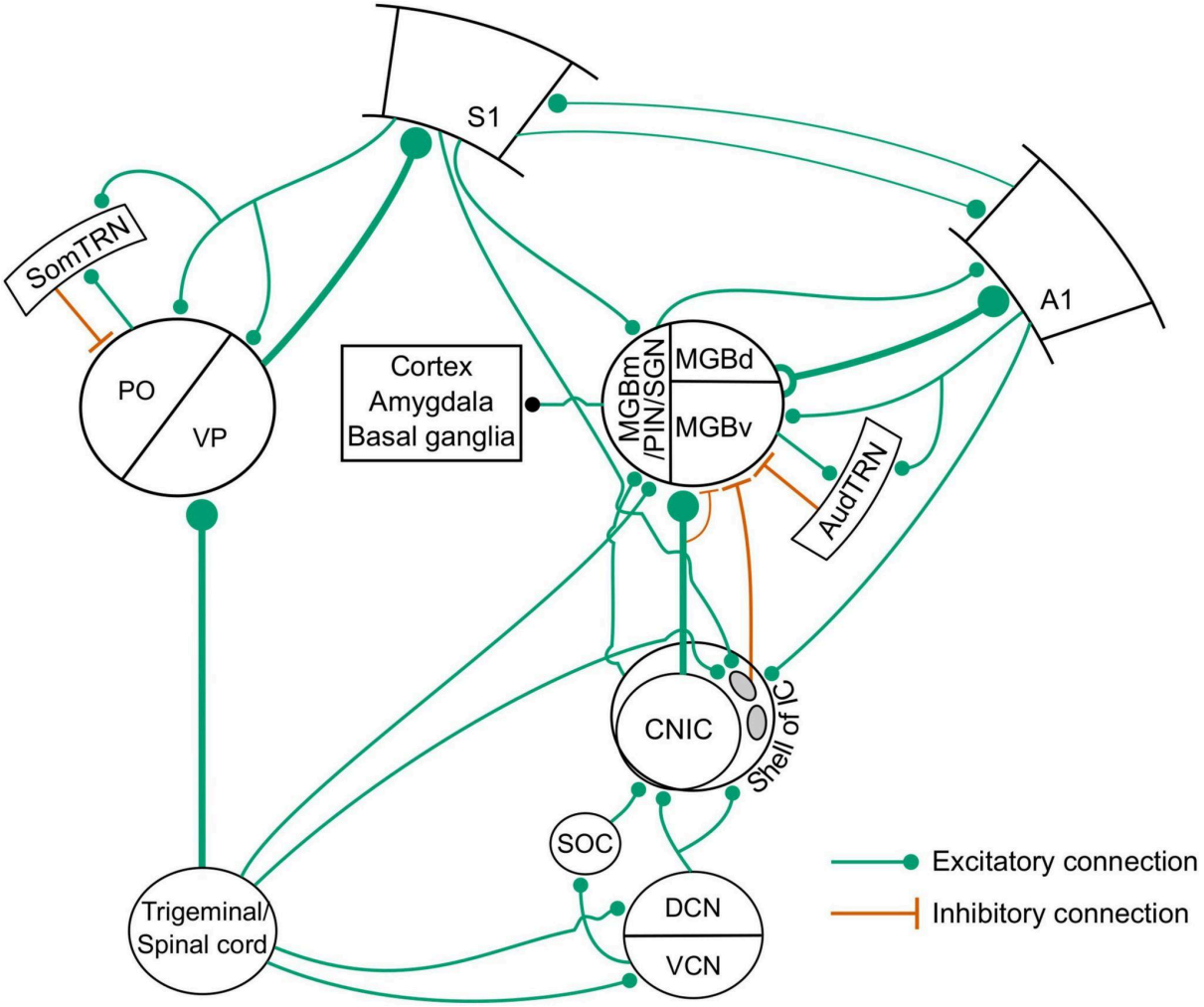

Figure 1. (From Lohse et al. 2022) Neural circuitry underlying somatosensory-auditory interactions. Schematic showing excitatory (green) and inhibitory (orange) connections between somatosensory and auditory structures. Key structures include the cochlear nuclei (DCN, dorsal cochlear nucleus; VCN, ventral cochlear nucleus), superior olivary complex (SOC), central nucleus (CNIC) and shell region of the inferior colliculus (Shell of IC), medial geniculate body subdivisions (ventral: MGBv; dorsal: MGBd; medial: MGBm/PIN (posterior intralaminar nucleus)/SGN (suprageniculate nucleus)), primary auditory cortex (A1), primary somatosensory cortex (S1), posterior medial nucleus (PO), ventroposterior nucleus (VP), and thalamic reticular nucleus sectors (AudTRN, auditory sector; SomTRN, somatosensory sector). Additional inputs arise from the trigeminal and spinal cord nuclei, as well as higher-level cortical, amygdalar, and basal ganglia regions.

Findings in primates indicate that some subdivisions of auditory cortex and possibly higher association cortices show enhanced or phase-reset responses to somatosensory input (Fu et al., 2003; Kayser et al., 2005; Lakatos et al., 2007). These facilitative patterns may route whisker-driven signals into circuits governing rapid orienting, valence appraisal, or general multisensory object representation.

The coexistence of suppressive gating in lemniscal pathways and facilitative interactions in non-lemniscal or higher-order circuits underscores the dynamic complexity of multisensory processing. On the one hand, whisker stimulation may thus reduce irrelevant or lower-priority auditory cues at primary thalamocortical relays; on the other, it can amplify signals in specialised or limbic regions, possibly shaping emotional responses, alertness, or more elaborate perceptual judgments.

1.5 Behavioural Findings: Enhanced Detection, Discrimination, and Reaction Times

From a functional perspective, integrating signals across multiple sensory modalities often reduces perceptual ambiguities and optimises behaviour (Bizley, Maddox, et al., 2016; Stein & Stanford, 2008). In humans, audio–tactile integration has been shown to lower detection thresholds, speed up reaction times, and improve discrimination in difficult or noisy conditions (Godenzini et al., 2021; Ro et al., 2009; Rowland et al., 2007; Schürmann et al., 2004; Sperdin et al., 2009). Vibrotactile stimulation, for instance, can shift perceived loudness of a faint tone (Gillmeister & Eimer, 2007), or amplify pitch/intensity discrimination if the frequency ranges align (Ro et al., 2009; Wilson et al., 2010).

Rodent studies, while fewer, suggest that whisker stimulation can similarly facilitate auditory tasks under certain paradigms, for example, by redirecting attentional resources (Godenzini et al., 2021; Rao et

al., 2014). Yet in other contexts, whisker cues may partially suppress or reweigh how animals process dual-modality inputs. These discrepancies suggest that the net outcome of audio–tactile convergence may depend on factors such as task demands, stimulus salience, and current behavioural goals (McGinley et al., 2015). Pinpointing exactly when whisker-based inputs provide net performance gains mostly remains an open question, particularly for rodents relying heavily on near-field tactile cues.

1.6 Active Whisking and Sensorimotor Gating

So far, we have seen that whisker-based inputs can modulate the auditory pathway at multiple levels, from the IC to the MGB and A1. Most of this evidence, however, arises from externally applied whisker stimuli, where the experimenter deflects the whiskers in non-behaving animals. In the wild, rodents primarily use active whisking - voluntary, self-generated vibrissa movements - for near-field exploration (Carvell & Simons, 1990; Sofroniew & Svoboda, 2015). Unlike passive deflections, self-initiated whisker sweeps introduce motor-related signals (such as corollary discharges or efference copies) that can modulate sensory processing before touch signals even arrive (Schneider et al., 2014; Sreenivasan et al., 2016). Below, we touch on three critical aspects of active whisking: (1) the neurophysiological mechanisms controlling and gating whisker movements, (2) the temporal dynamics of whisking and how they implement predictive coding, and (3) how rodents adjust whisking strategies based on behaviour and context—to illustrate how these factors might shape audio–tactile interactions.

1.6.1 Neurophysiological Mechanisms: Basal Ganglia–Thalamocortical Loops

While whisker movements can be generated by brainstem pattern generators, higher motor control of whisking involves basal ganglia–thalamocortical loops (Kleinfeld & Deschênes, 2011;

Sreenivasan et al., 2016). The basal ganglia (BG)—long associated with action selection—receive cortical inputs and, via direct and indirect pathways, regulate thalamic drive to the motor cortex. This circuit helps facilitate desired whisker movements and suppress competing ones (Kleinfeld & Deschênes, 2011). Through inhibitory connections to midbrain or thalamic nuclei, the BG can also gate somatosensory feedback during whisking, suppressing predictable reafferent signals so that any unexpected tactile contact stands out (Crapse & Sommer, 2008). Notably, facial motor areas (sometimes termed "whisker motor cortex") project both to the facial nucleus (driving whisker muscles) and to sensory areas (S1, S2), generating corollary discharge signals (Sreenivasan et al., 2016). These signals can prime somatosensory circuits to expect whisker contact, effectively preventing self-induced rustling noises or minor tactile feedback from interfering with more salient cues (Singla et al., 2017).

In turn, descending projections from somatosensory cortex to the IC and MGB could be speculated to interact with BG-thalamocortical loops, thereby allowing whisker movements to modulate not just tactile but also auditory pathways, though direct evidence for BG gating of this circuit is missing. For instance, while the BG disinhibits whisker movements, inhibitory neurons in the IC shell could receive whisker-driven inputs in parallel that dampen concurrent sound responses.

1.6.2 Temporal Dynamics of Predictive Coding in Whisking

Rodents whisk rhythmically at frequencies typically ranging from 4–12 Hz (Kleinfeld & Deschênes, 2011). This oscillatory behaviour provides a temporal framework for predictive coding, whereby the brain aligns motor output (whisker sweeps) with anticipated tactile or auditory feedback. Much like the coupling of sniffing and olfactory sampling, the rhythmic nature of whisking can phase-lock the processing of relevant stimuli (Carvell & Simons, 1990). During the protraction phase—when whiskers move forward to contact objects—somatosensory cortex exhibits heightened

responsiveness to unexpected touches (Crochet & Petersen, 2006). Meanwhile, inhibitory mechanisms can dampen neural responses to the expected self-generated signals, preventing them from overwhelming the system. This can similarly be observed in the visual and auditory systems (S.-H. Lee et al., 2012; Singla et al., 2017).

This phase dependence extends to crossmodal interactions. For example, if rodents produce whisker-generated noise or if whisking coincides with a tone, the auditory cortex may temporarily downregulate its response (Schneider et al., 2014). In parallel, the trigeminal pathway conveying whisker afferents can receive efference copy signals that mark the movement's timing—enabling suppressive gating of self-generated sounds at precisely the whisking phase when they are most likely (Crapse & Sommer, 2008). When the infra-orbital nerve is severed, the tactile branch of the S1 → IC → MGB suppressive loop is silent and IC and AC fire vigorously to the broadband 'scratching' sounds of active whisking (Efron et al. 2025). Importantly, responses are even larger than to matched, passively replayed sounds, implying that a facilitatory efference-copy signal combines with the acoustic input whenever the animal generates the movement itself.

In sum, the whisking cycle may help the rodent's brain distinguish predictable self-driven inputs from novel external events, thus optimising both tactile and auditory processing through temporally-coordinated gating.

1.6.3 Whisking Strategies and Behavioural Context

Rodents adjust their whisking patterns and amplitude based on context. For exploratory whisking, they often sweep whiskers widely and synchronously across space at moderate frequencies, aiming to detect obstacles or surfaces (Carvell & Simons, 1990). Once an object is located, they may shift to "foveal" whisking—smaller, high-frequency movements of specific whiskers—enabling finer texture

discrimination or object contour mapping (Kleinfeld & Deschênes, 2011). During social interactions, rodents whisk each other's faces or bodies, potentially temporarily coordinating whisker movements with ultrasonic vocalizations in a way that avoids cross-sensory interference.

Each of these behaviours entails a different sensorimotor gating profile. For instance, exploratory whisking relies on gating out the constant reafferent signals of wide sweeps while remaining highly sensitive to unanticipated contacts. During foveal whisking, the system might boost relevant touch signals to discern fine tactile details, which can lead to a more pronounced suppression of distracting auditory inputs. Even subtle differences in whisking amplitude or frequency could modulate the interplay of whisker and auditory signals in subcortical loops, potentially altering how sound-evoked responses are scaled in the MGB and A1. Thus, the sensorimotor gating that emerges in active whisking may not be monolithic—it could instead vary according to task goals (e.g., broad exploration vs. detailed palpation) and behavioural states (locomotion, arousal, or social context).

1.6.4 Relevance for Audio–Tactile Convergence

Active whisking reveals an intricate layer of contextual modulation that goes beyond simply "touch suppresses sound". The circuits engaged by whisker movements can exert both inhibitory and facilitatory effects on concurrent auditory processing, depending not only on auditory processing stage but also potentially on motor output, whisking phase, and behavioural salience. This dynamic gating is especially relevant to near-field exploration, where rodents rely on whiskers for spatial navigation and object discrimination while also needing to remain vigilant to local sounds . The dual-mode picture that emerges from the new data is therefore one of dynamic balance: narrow-band tonal input is divisively scaled to minimise distraction, whereas unpredictable broadband whisking sounds are transmitted, and sometimes even facilitated, because they convey salient information e.g. about nearby surfaces. In doing

so, rodents could achieve an efficient multisensory perception that is highly tuned to the demands of real-world exploration.

1.7 Auditory influences on somatosensory and visual cortices

While somatosensory signals can suppress auditory processing, the reverse influence, auditory input modulating tactile or visual representations, demonstrates that crossmodal communication is bidirectional and widely distributed across cortical hierarchies. Two-photon imaging in mice demonstrates that auditory stimulation profoundly reshapes tactile coding in the somatosensory cortex. In Zhang et al. (2020), concurrent sounds altered whisker-evoked responses in both S1 and S2, most prominently producing frequency-dependent inhibition in S2 neurons (Zhang et al., 2020). This inhibitory effect was strongest for low-frequency whisker deflections and persisted after controlling for locomotor activity, suggesting an intrinsic cortical mechanism. The same study identified a subset of S2 neurons selectively responsive to sound whose activity was reciprocally suppressed by high-frequency whisker inputs, implying a mutually inhibitory local circuit that dynamically gates between touch and sound depending on spectral content. Godenzini et al. (2021) showed that auditory inputs can enhance rather than suppress tactile encoding, particularly within the forepaw representation of S1. There, auditory stimulation increased the reliability and precision of whisker-evoked spike timing and improved tactile discrimination behavior, suggesting that crossmodal influences can be facilitatory for sensory congruent or behaviorally relevant stimuli (Godenzini et al., 2021).

Kato and Bruno (2025) systematically tested how robust and functionally meaningful auditory influences on S1 actually are. Using in vivo calcium imaging and behavioural paradigms, they found that while brief noise bursts elicit small, consistent responses in a minority of barrel cortex neurons, these signals convey no information about sound identity and remain stable across wakefulness, locomotion,

and reward learning. Passive or reinforced audio–tactile pairing did not induce plastic changes, and concurrent sound produced uniform, nonspecific suppression of whisker responses rather than feature-selective integration. Thus, auditory inputs to S1 appear largely diffuse, excitatory, and noninformative, modulating cortical excitability rather than encoding sensory content. This stands in sharp contrast to the S1→A1 pathway described by Lohse et al. (2021), where structured inhibitory projections from somatosensory to auditory circuits implement a precise, state-dependent gating of ascending sound signals. The asymmetry between these directions—strong, functionally specific suppression from S1 to A1 versus weak, nonspecific modulation from A1 to S1—suggests that multisensory communication among primary cortices is not reciprocal in effect but instead weighted to prioritize tactile information during near-field exploration (D. D. Kato & Bruno, 2025).

In the visual cortex, Bimbard et al. (2023) demonstrated that sound-evoked responses in V1, long thought to represent true multisensory integration, largely arise from sound-induced body movements and postural adjustments. When movement was precisely tracked and regressed out, most "auditory" activity disappeared, indicating that some cross-sensory correlations in the cortex may reflect behavioral coupling rather than direct sensory convergence (Bimbard et al., 2023).

Together, these studies show that crossmodal influences can flow in both directions but differ in their mechanisms and functional roles. The S1→A1 pathway described by Lohse et al. (2021) acts through long-range inhibitory circuits that suppress auditory relays, whereas A1→S1 communication can be excitatory as well, locally implemented, but also often be modulated by behavioral state. The differences may reflect the priorities of each modality: Tactile cues may dominate when whiskers are in contact with nearby objects, but auditory inputs can augment somatosensory perception when external sounds predict or coincide with tactile events. These reciprocal yet context-dependent auditory influences in S1 show how cortical networks achieve flexible multisensory weighting. In contrast, this thesis focuses on

the reciprocal direction of interaction, investigating how tactile signals are integrated into auditory cortical processing under different behavioural and contextual conditions, building on the passive stimulation paradigm described by Lohse et al.

1.8 Thesis Overview and Objectives

In the following three experimental chapters, we aim to answer the overarching research question of this thesis: "How does whisker-based somatosensory input interact with auditory cortical processing under varying behavioural contexts?".

Chapter II investigates perceptual consequences of whisker stimulation on task performance, in a near-threshold sound detection and frequency discrimination task.

Chapter III explores how changes in behavioural context and reward contingencies influence multisensory integration in the auditory cortex, particularly examining the modulation of auditory cortical responses when mice switch between auditory and tactile detection tasks.

Chapter IV examines how auditory cortical neurons encode spontaneous versus sound-triggered whisking. By analysing whisker movements, pupil dilation, and locomotion alongside two-photon calcium imaging in the auditory cortex, this chapter addresses whether and how internal states and sensory contexts modulate auditory processing through active whisking.

Finally, the General Discussion (Chapter VI) synthesises these findings, placing them in the broader context of multisensory and sensorimotor integration, and proposes directions for future research.

II Perceptual Implications of Somatosensory-Auditory Suppression

2.1 Introduction

The capacity of the brain to seamlessly integrate different sensory inputs is particularly evident when considering how somatosensory and auditory signals interact in rodents that heavily rely on whiskers (vibrissae) for close-range exploration (Sofroniew & Svoboda, 2015). Recent work has shown that stimulation of the whiskers can suppress sound-evoked responses within key auditory relays (Lohse et al., 2021; Rao et al., 2014). However, the functional consequences of this suppression, especially for fundamental perceptual tasks such as sound detection and frequency discrimination, remain largely unexplored.

Studies in multiple species (e.g., ferrets, rats, mice) have documented how whisker-driven inputs can attenuate sound responses via midbrain or thalamic circuits, in turn leading to reductions in sound-evoked activity at both subcortical and cortical levels (Kimura & Imbe, 2018; Lohse et al., 2021; Meredith & Allman, 2015). A substantial body of evidence indicates that the divisive suppression revealed by Lohse et al. (2021), with the strongest suppressive effects observed at the sound frequencies to which the neurons responded most strongly, may help animals cope with internally generated sound or filter out irrelevant stimuli while whisking (Rao et al., 2014). Yet, it is by no means self-evident that flattening the tuning curve translates to poorer perceptual performance. Indeed, theoretical and experimental work suggests that effectively reducing the overlap of adjacent tuning curves by reducing their amplitude in reference to perceptual thresholds can improve certain aspects of sound discrimination, even if overall firing amplitudes decrease (Atiani et al., 2009; Rabinowitz et al., 2013).

Some studies imply that rodents can modulate or gate multisensory responses, particularly when certain tactile cues are less behaviourally relevant (Iurilli et al., 2012; Rao et al., 2014). This raises the possibility that long-term training might diminish the robust whisker-driven suppression observed in acute or passive paradigms (Lohse et al., 2021). Moreover, near-threshold detection could be modulated differently from tasks that rely on precise frequency distinctions (Guo et al. 2017). The interplay between auditory and tactile cues may also vary with the breadth of auditory recruitment: pure tones recruit narrower populations in A1, whereas broadband noise or complex sounds engage a wider range of frequency channels (Christensen et al. 2019). Hence, the same "suppressive circuit" could manifest different behavioural outcomes depending on stimulus, training history, or reward context.

A study by Christensen et al. (2019) illustrates this principle in the context of background noise. Their experiments in mouse auditory cortex demonstrated that adding background noise suppressed cortical tuning curves, increasing frequency selectivity and thereby improving the animals' ability to discriminate closely-spaced tones. Similarly, artificially suppressing cortex by optogenetic activation of inhibitory interneurons yielded a comparable improvement in frequency discrimination. This highlights a fundamental insight: firing-rate suppression does not necessarily impede perceptual acuity; under certain conditions, it can refine tuning and facilitate better performance on tasks requiring fine frequency resolution.

Building on these findings, our goal in this chapter is to assess how whisker-evoked suppression of auditory responses influences rodent behaviour in two quintessential tasks: (1) near-threshold sound detection, and (2) frequency discrimination. The logic is as follows:

- In detection tasks, whisker stimulation might impair noticing faint sounds by reducing their effective strength through divisive scaling. Conversely, partial gating could help animals ignore

predictable, self-generated whisker signals (noise), preventing these signals from overshadowing faint but behaviourally relevant auditory cues. Whether this gating ultimately helps or hinders detection likely depends on how whisker input is tied to reward—if whisker stimulation is never rewarded, animals can learn to disregard it, thus minimizing any interference with detecting near-threshold sounds..

- In a frequency discrimination task, we probe whether whisker stimulation impairs or improves spectral resolution. If whisker input narrows neuronal tuning or diminishes overlap between frequency channels, as suggested by Christensen et al. (2019), we might see enhanced discriminability for closely-spaced tones.

We measure detection and discrimination performance with and without simultaneous whisker stimulation while monitoring neural responses via two-photon calcium imaging in primary auditory cortex (A1). We ask whether whisker-induced suppression of auditory activity elevates detection thresholds or instead facilitates detectability, and whether it sharpens frequency representations enough to enhance discrimination. By comparing detection versus discrimination, we aim to delineate when and why suppression acts as a helpful filter rather than a handicap—extending the core ideas from Christensen et al. (2019) to a whisker-driven, context-dependent scenario in mice.

2.2 Results

We trained water-regulated mice on a sound detection task in which they were rewarded with a drop of water for licking in response to a sound. During the initial training phase, mice encountered three trial types: 'Sound-alone' trials (12 kHz pure tone), 'whisker-alone' trials (approximately 1 mm vertical displacement of the whiskers) and catch trials (no stimulus, Figure 2). These were presented randomly interleaved and with a probability of 45%, 45% and 10%, respectively. Mice were rewarded only for licking during sound-alone trials but not for licking during whisker-alone or catch trials. Over the course of the training, we established each animal's detection threshold by progressively lowering the sound intensity until the detection rate was reduced to approximately 50% (Figure 2). At this stage, a fourth trial type ('multi') was introduced, in which sound presentation and whisker stimulation occurred simultaneously. Licking during multi-trials was rewarded but in order to avoid mice learning to associate whisker stimulation with the availability of a reward, multi-trials occurred with a low probability (test phase probabilities: multi: 10%, sound-alone: 40%, whisker-alone: 40%, catch: 10%). During the test phase, sounds were presented at three different intensities near the threshold (e.g. -5dB, 0dB, +5dB relative to threshold) plus a fourth easily detectable sound level. Behavioural testing was combined with two-photon calcium imaging of cellular activity in the auditory cortex.

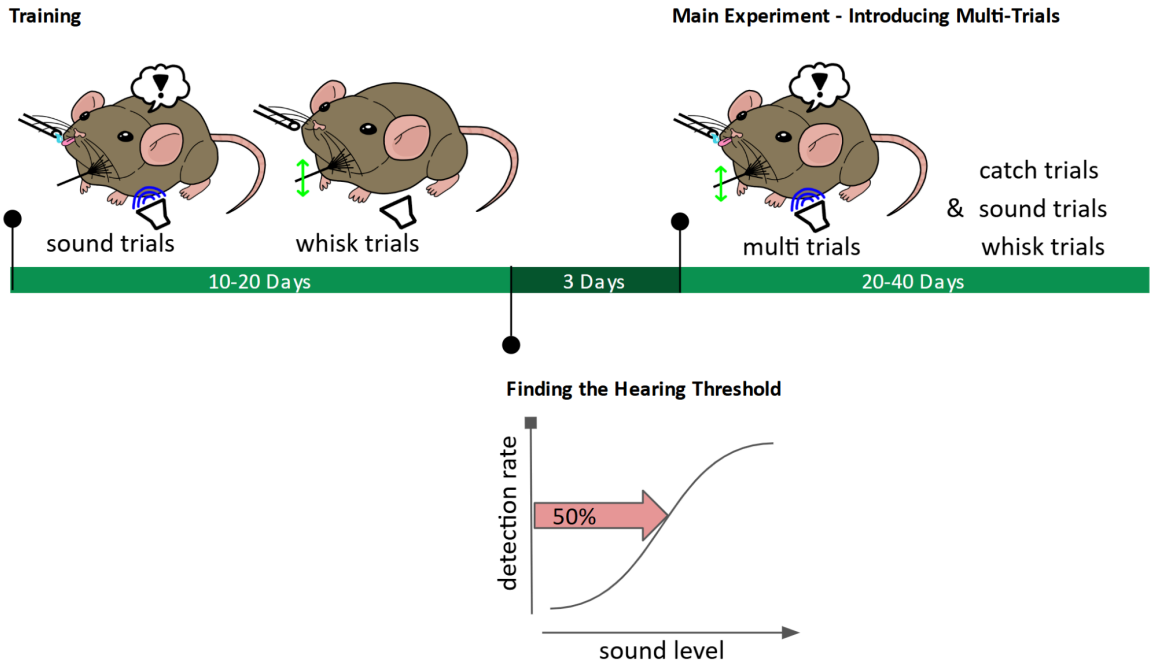

Figure 2. Training timeline and experimental procedure for the pure tone experiment. The training process consisted of multiple phases: (1) An initial training period (10–20 days) in which mice were exposed to sound-alone trials and whisker-alone trials to establish a behavioural association between stimuli and water reward. (2) A brief 3-day intermediate phase followed, during which the detection threshold was determined by progressively lowering the sound pressure level until mice detected tones at a 50% detection rate (bottom panel). (3) The main experimental phase (20–40 days) introduced multi-trials, where sound and whisker stimuli were presented simultaneously. Trial types included sound-alone, whisker-alone, multi-trials, and catch trials, which were randomly interleaved. Multi-trials were introduced at a low probability of 10% of trials to prevent mice from associating whisker stimulation with rewards.

2.2.1 Sound Detection Task Learning

Mice didn't typically complete the training phase after a particular time point, but showed variability in their learning rates across the cohort (data from 9 mice included). To assess learning progress, we tracked each mouse's detection sensitivity (d-prime) for sound-alone trials over time (Figure 3, A). Each mouse's performance was plotted as a moving average of d-prime, with the criterion for successful task learning set at a d-prime of 2, indicating reliable discrimination between target and non-target stimuli.

The left panel of Figure 3 shows that individual mice differed markedly in how quickly they reached the d–prime criterion. Some mice, such as GTMK54_2R and GTMK54_2L, achieved the threshold within just a few days of training, whereas others, like GTMK56_1b, required longer to reach the same performance level. This variability reflects differences in learning rates and task engagement across the cohort. To quantify the overall learning progress across the cohort, we calculated the cumulative distribution function (CDF) of the day each mouse reached the d-prime threshold of 2 (Figure 3, B). The CDF shows that approximately 50% of the mice reached the criterion within the first 5 days of training, with most mice achieving it by day 17. The logistic fit to the CDF further illustrates a steep initial rise in the proportion of mice reaching the threshold, followed by a plateau as the remaining slower learners reached the target performance level.

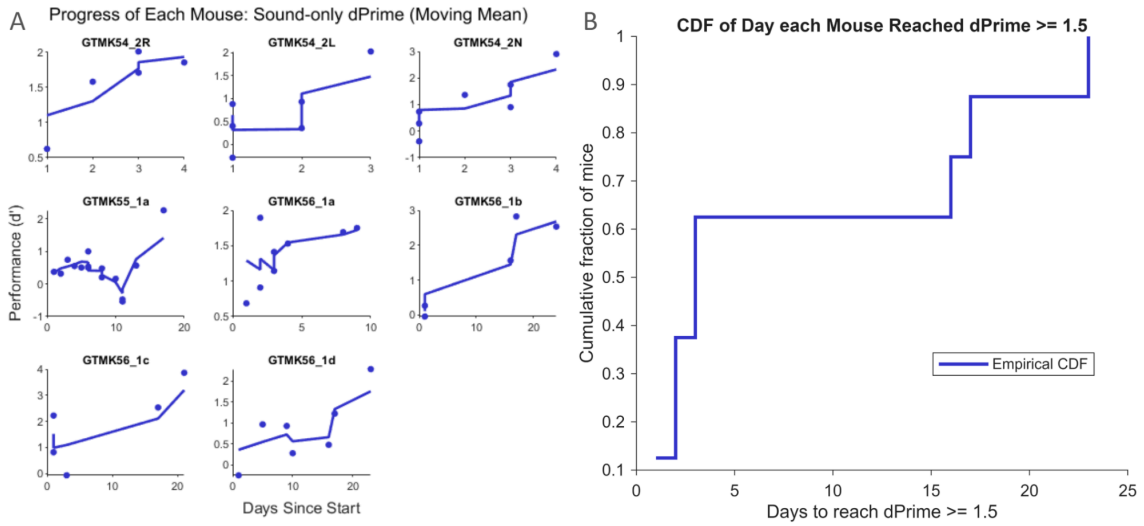

Figure 3. Sound Detection Task Learning. (A) Detection performance (d-prime) as a function of training day for individual mice. The line represents the moving average of 3 days. During some training sessions, "free" rewards were given to encourage the animals to lick. These sessions were not included in the learning progress plots, resulting in large gaps between consecutive data points for some mice. **(B)** Cumulative distribution function (CDF) showing the proportion of mice reaching the d-prime criterion of 2.0 as a function of time.

2.2.2 Effects of Somatosensory-Auditory Stimulation on Task Performance

In order to establish whether stimulating their whiskers affects the animals' ability to detect sounds, we compared their performance (d-Prime) in sound-alone and multi-trials (Figure 4), which had a d-Prime of 1.33 and 1.38, respectively ($p=0.459$, Wilcoxon sign-rank test). No mouse showed significantly higher sensitivity in multi-trials than sound-alone trials (Figure 4, C). There was also no significant difference between conditions when the data from all mice were combined (Figure 4, lA). Given that extreme values (hit and false-alarm rates approaching 0 or 1) can heavily skew the associated d-prime values (Stanislaw & Todorov, 1999), and in order to present a comprehensive account of the

control conditions for catch trials (randomly interleaved trials with no stimulus presentation) and whisker stimulation, we further scrutinised the response rates for each stimulus condition, as depicted in Fig 5. The mean response rates were 0.540 for tone-trials and 0.538 for multi-trials, respectively ($p=0.682$, Wilcoxon sign-rank test), confirming that there is no significant difference in the detection performance between the conditions. There were also no significant differences between the conditions for individual mice (Figure 5, C).

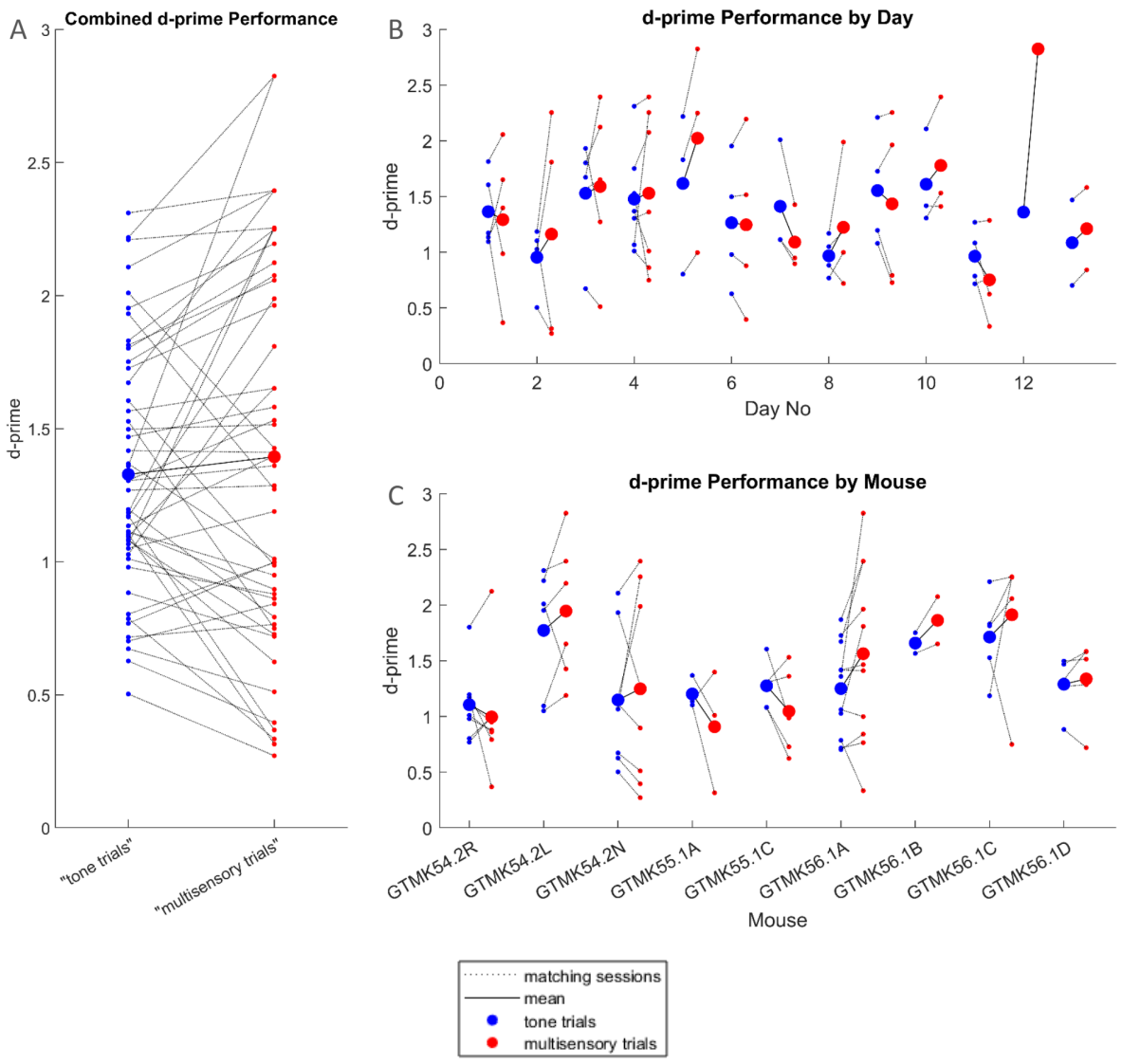

Figure 4. Comparison of detection performance (d-prime) between sound-alone and multi trials. (A) d-prime of all sessions of all mice for sound-alone (blue) and multi trials (red). The two large dots indicate the mean d-prime across the two conditions. Dashed lines connect matching sessions. **(B)** d-prime in sound-alone and multi trials for different days. Small dots indicate individual mice, large dots indicate the means across mice. Day 1 was defined as the first day on which a mouse was introduced to multi-trials. **(C)** d-prime on sound-alone and multi trials for different mice. Small dots indicate individual sessions, large dots indicate the means across sessions.

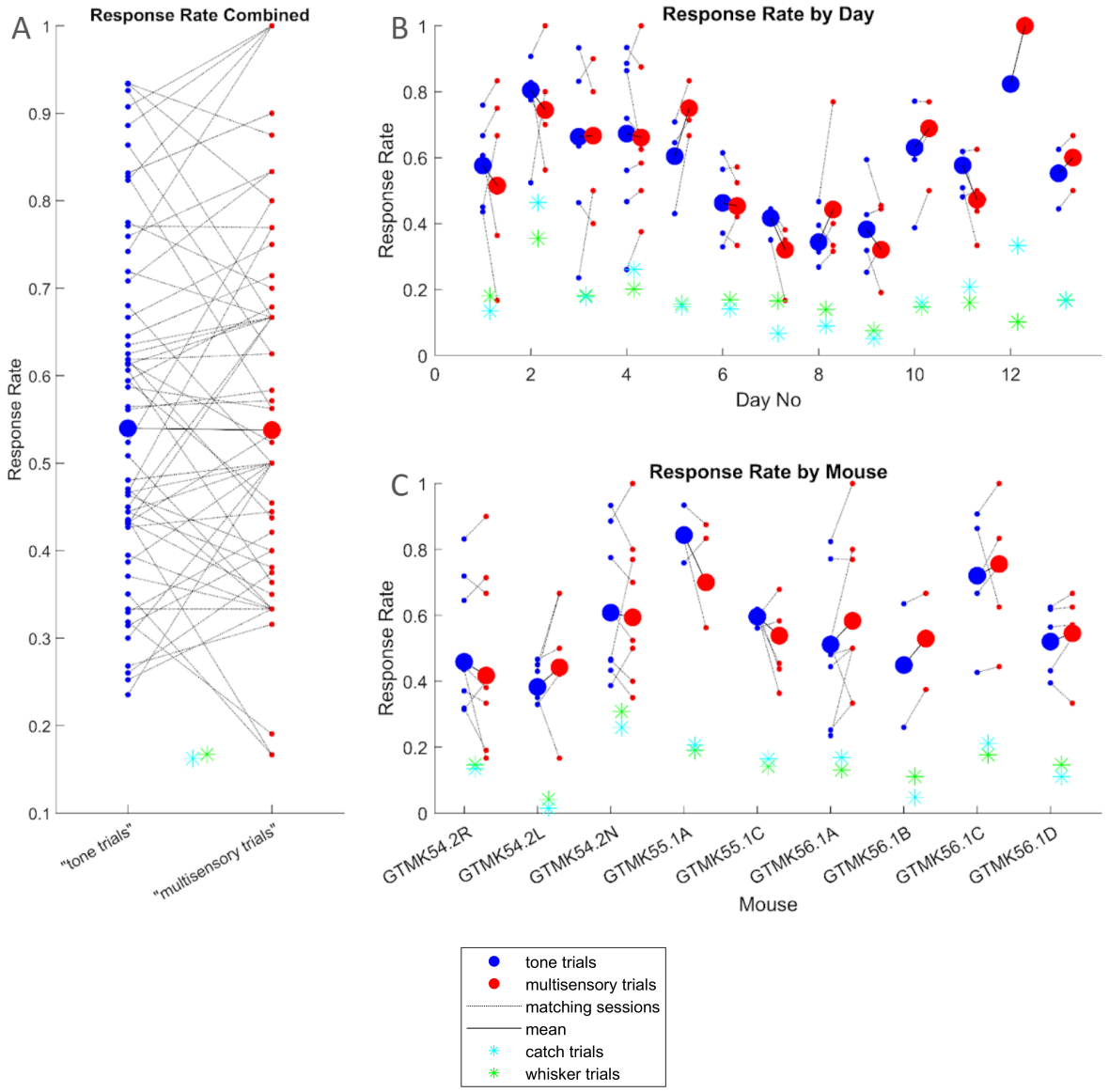

Figure 5. Response rates across sound-alone, multi, whisker-alone, and catch trials by session, day, and mouse. (A) Response rates of all sessions of all mice for sound-alone (blue) and multi (red) trials. Dashed lines connect matching sessions. The two large blue and red dots indicate the mean response rates across the two conditions. The large green and cyan asterisks indicate the mean response rates for whisker-alone and catch trials, respectively. **(B)** Response rates on sound-alone (blue) and multi (red) trials for different days. Small dots indicate individual mice, large dots indicate the means across mice. The large green and cyan asterisks indicate the mean response rates for whisker-alone and catch trials, respectively. Day 1 was defined as the first day on which a mouse was introduced to multi-trials. **(C)** Response rates on

sound-alone (blue) and multi (red) trials for different mice. Small dots indicate individual sessions, large dots indicate the means across sessions. The large green and cyan asterisks indicate the mean response rates for whisker-alone and catch trials, respectively.

Analysis of the response rates for the control conditions revealed that false alarms (catch trials) and responses to whisker-alone stimulation (a control condition that allowed us to evaluate whether mice began to perceive the whisker-stimulation component of the multi-sensory trials as a target after introduction of the whisker-alone trials) were generally low, with responses occurring in 16.2% of all catch-trials and 16.7% of all whisker-alone trials (Fig 5, A) in these sessions. They remained low over the course of the experiment indicating that no association between the reward and whisker stimulation was formed as a consequence of receiving rewards in multi-trials (Fig 5, C). Furthermore, most mice had very similar average response rates for catch trials and whisker-alone stimulation, indicating that mice didn't selectively lick for whisker-alone trials.

Together this suggests that whisker stimulation has very little, if any, impact on sound detectability. Therefore, we next assessed whether whisker stimulation affected the timing of the mice' responses and found that they responded faster in multi-trials than in sound-alone trials. All mice showed nominally faster response times and there was a statistically significant reduction in median latency from 318 ms (+) to 302 ms (+) across all trials ($p=0.0012$ Mann-Whitney-Wilcoxon non-parametric test, n= 4206 sound-alone trials, 475 multi trials) overall (Figure 6).

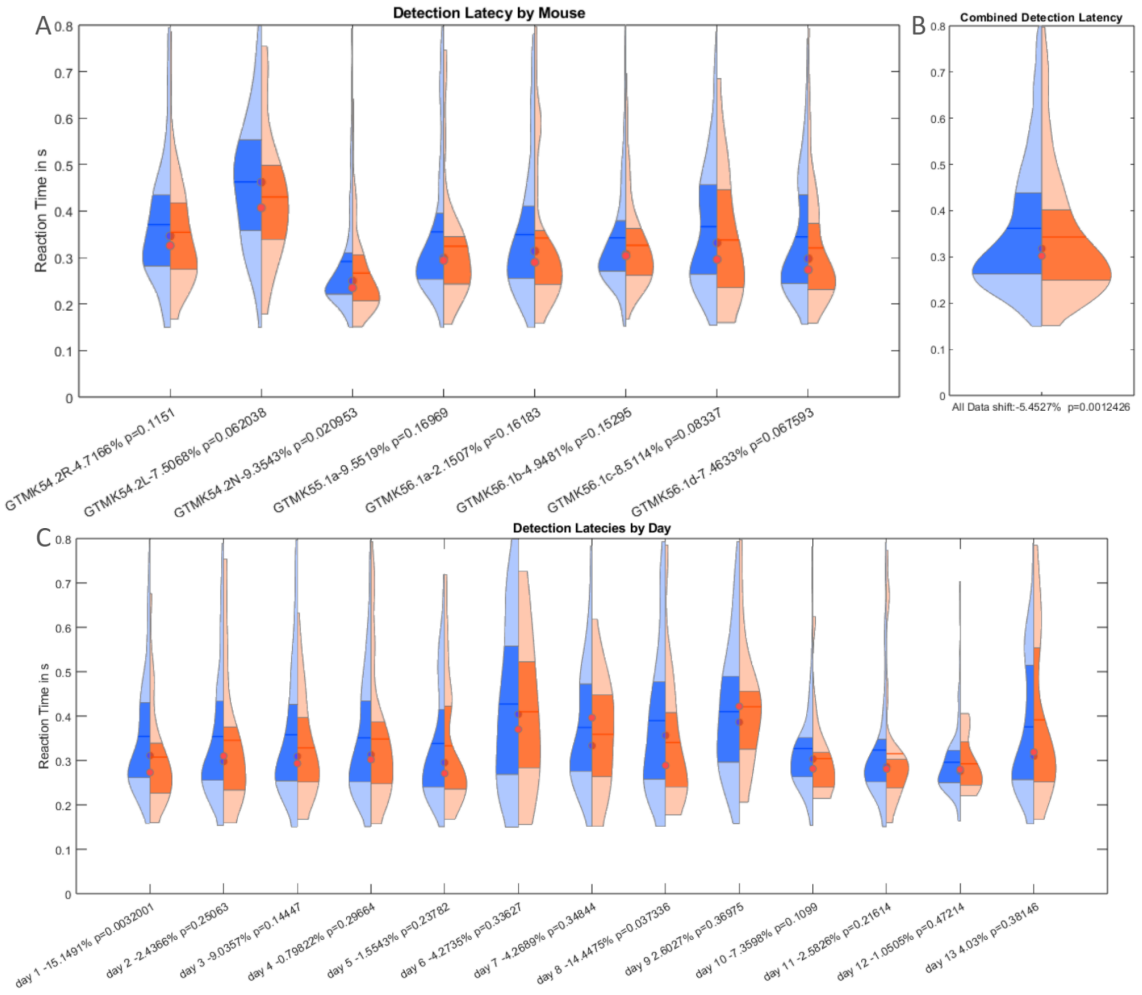

Figure 6: Response latencies across sound-alone and multi trials. (A) Response latencies in sound-alone (blue) and multi trials (red) for different mice. **(B)** Response latencies in sound-alone (blue) and multi trials (red) for all mice and all sessions. **(C)** Response latencies in sound-alone (blue) and multi trials (red) across days. Day 1 was defined as the first day on which a mouse was introduced to multi-trials. Dark and light shaded areas in violet plots indicate the distribution of detection timings while the dots and lines indicate the median and the mean, respectively.

We had predicted that the suppression of sound-evoked neural activity caused by whisker stimulation would make it harder for mice to detect that sound. That whisker stimulation did not lead to lower sensitivity and even had a beneficial effect on response speed is clearly inconsistent with this

prediction. However, some aspects of our methodology may have precluded the somatosensory stimulation from having the impact on sound response that we had expected. First, our previous work showed that suppression of tone-evoked neural activity is strongest near the peak of a neuron's tuning curve and that suppression tends to be stronger for broadband noise (Lohse et al., 2021). Here, we presented pure tones of a single sound frequency. Second, mice were exposed to whisker stimulation (whisker-alone trials) from the beginning of the training phase and this likely taught them, over weeks of training, to ignore this stimulation. Consequently, when multi-trials were eventually introduced, whisker stimulation may have lost much of its saliency, its ability to capture processing resources and potentially even its capacity to suppress sound-evoked activity. Nevertheless, by comparing the neural activity of sound-driven neurons in sound-alone and multi trials we could ascertain that whisker stimulation does significantly suppress sound-evoked activity in mice even after weeks of training with exposure to whisker-alone trials (12% suppression, $p<0.001$, Figure 7). However, the degree of suppression was moderate compared to a >20% difference reported previously in passively listening mice (Lohse et al., 2021).

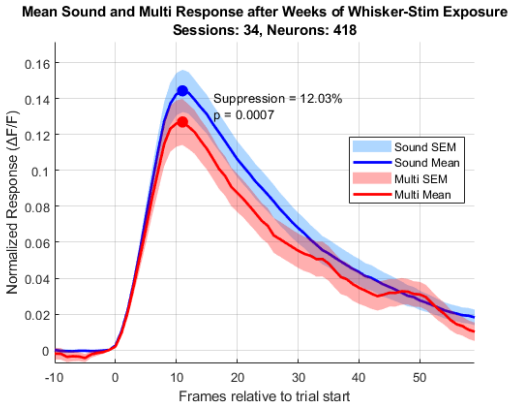

Figure 7. Mean sound-evoked and multi-trial responses in miss trials after weeks of exposure to whisker-alone trials. Normalised calcium responses (ΔF/F) of sound-driven neurons in sound-alone (blue) and multi trials (red) from the last week of experiments. These data were recorded after mice had undergone extensive training, including weeks of exposure to whisker-alone trials during task learning, threshold determination, and at least one week of multi trial sessions.

Therefore, we trained additional mice on a sound detection task with a modified methodology. The main changes were to present broadband noise bursts instead of pure tones, as broadband noise was expected to recruit a larger proportion of auditory neurons and because Lohse et al. (2021a) had shown that suppression during noise was greater than during pure tones. Whisker-alone trials were introduced only at the same stage as multi-trials, i.e. after the training phase (Figure 8), rather than from the beginning of training as we had done previously, in order to prevent animals from learning to ignore them through prolonged exposure. Additionally, we increased the proportion of multi-trials to 30% (test phase probabilities: 30% multi trials, 30% sound trials, 30% whisk trials, 10% catch trials), as earlier experiments had indicated that mice did not learn to treat multi-trials differently from sound-alone trials, and the higher proportion ensured sufficient sampling. We also replaced the brush stimulator with a fine steel mesh grid, since mice were often able to pull their whiskers free of the brush; the mesh, with ~0.7

mm square holes, provided more secure engagement while delivering the same displacement profile. Finally, we restricted stimulation to a single near-threshold sound level determined by an automated procedure at the beginning of each session, which maximized the number of multi-trials available per session for robust statistical analysis. This refinement addressed a major limitation of the original design, in which the small 10% proportion of multi-trials was further split across several sound levels, leaving too few usable trials. Finally, the ISI was now based on the last lick and not the last stimulus to further discourage anticipatory licking. When we, again, compared neural responses in sound-alone and multi-trials we found that somatosensory suppression of sound evoked activity was stronger in the mice trained with the modified paradigm than with the original one, suggesting that introducing whisker-alone trials later (or the use of broadband noise instead of pure tones) increased the capacity of the somatosensory stimulus to suppress sound responses, possibly through decreasing the mice having learned to ignore it.

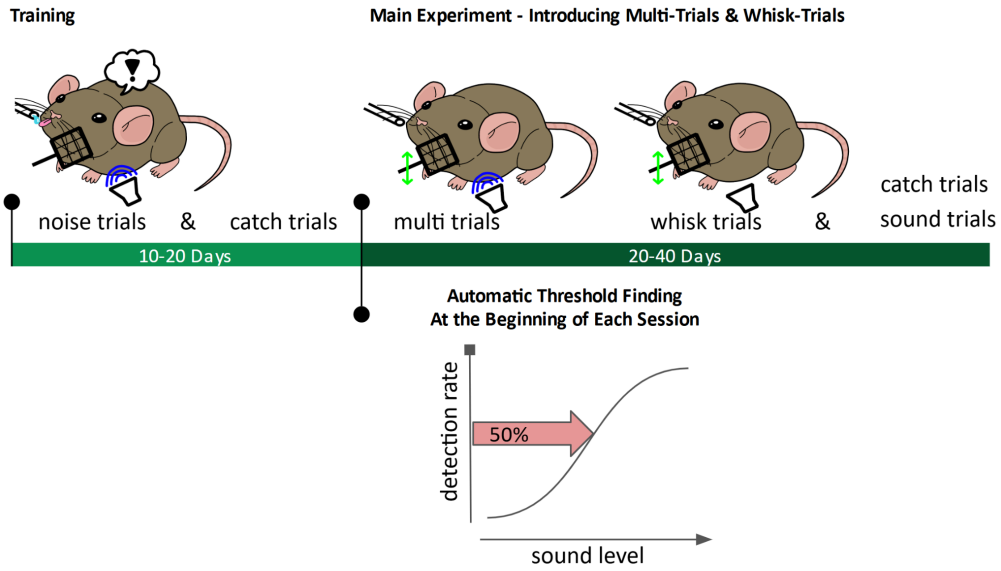

Figure 8: Training timeline and experimental procedure for the broadband-noise detection control experiment. (1) During the initial training phase (10–20 days), mice were exposed to

sound-alone trials and catch trials to establish a behavioural association between auditory stimuli and water rewards. (2) The main experimental phase (20–40 days) introduced multi trials (noise + whisker stimulation and whisker-alone trials alongside the previous sound-alone and catch trials. Unlike the pure tone experiment, whisker-alone trials were introduced simultaneously with multi trials to assess the immediate impact of somatosensory stimuli without prolonged prior exposure. To ensure consistency, automatic threshold detection was conducted at the beginning of each session by presenting a range of noise levels and fitting a logistic function to the detection rates. A more dynamic inter-stimulus interval (ISI) based on the last lick (instead of the last trial) was implemented, ranging from 4.5 to 9 seconds, to minimize random hits. The training script automatically terminated once 30 multi-trials were completed, ensuring a balanced trial distribution across sessions.

However, despite stronger neural suppression, whisker stimulation did not cause an impairment in detection sensitivity (Figure 9). Both d-prime performance and response rates revealed no significant differences between broadband noise and multi-trials. Across all sessions, the mean d-prime values for noise-alone and multi-trials were nearly identical (0.771 vs 0.768, $p=0.995$, Wilcoxon signed-rank test), indicating that whisker stimulation did not reduce the animals' ability to detect auditory stimuli. These results were consistent across individual days and both mice, with low variability in detection sensitivity. Notably, catch trial response rates remained low, confirming that mice were not exploiting unintentional patterns in trial timing in the task structure.

Response speed, on the other hand, was dramatically impacted by the addition of whisker stimulation. Reaction times were substantially longer in multi trials compared to sound-alone trials in both mice. Mouse GTMK64.4a exhibited a median reaction time increase from 234ms for noise-alone trials to 283ms for multi-trials, while for mouse GTLA24.1C, reaction times increased from 247ms to

332ms. When combined, the median reaction time for multi-trials was 307ms, compared to 240ms for noise-alone trials—a statistically significant 20% increase ($p<0.001$, Wilcoxon rank-sum test).

Interestingly, the impact of whisker stimulation on response latency was reversed compared to the pure-tone experiment, where whisker stimulation slightly improved reaction times. Here, whisker stimulation significantly delayed the response to the sound. This difference likely arises from the fact that in the new paradigm tactile stimulation was novel when multi-trials were introduced as whisker-alone trials were only introduced at the same stage as multi-trials whereas they were included from the start in the original experiment. Due to its novelty the tactile stimulation may have distracted the mice and delayed their response whereas, in the original experiment, mice had had ample opportunity to adapt to the whisker stimulus and actually ought to have learned to ignore it as it was not rewarded. Together, these findings highlight the nuanced role of somatosensory inputs in modulating auditory detection, demonstrating that tactile stimulation can delay responses without impairing sensitivity and that the procedural details of the learning paradigm and stimulus selection are important considerations potentially profoundly influencing the results.

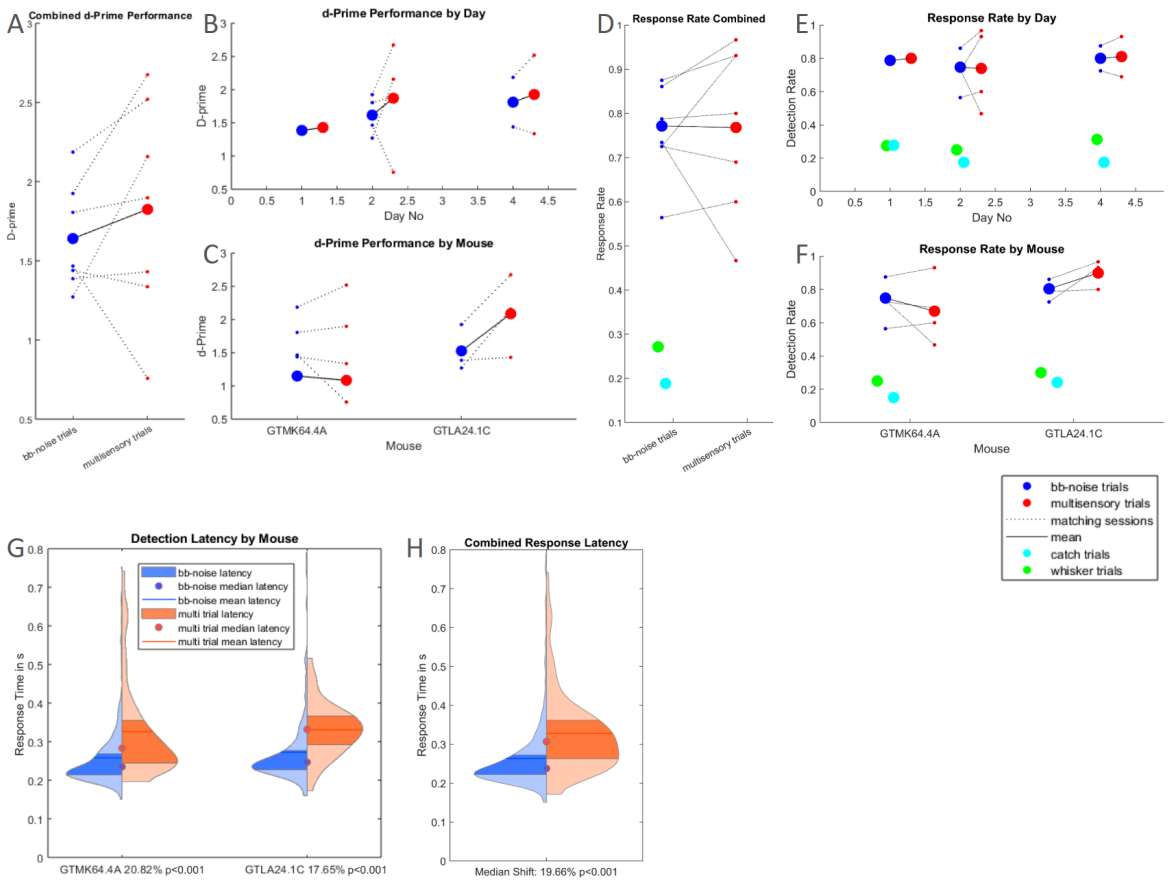

Figure 9. Detection performance and latencies in mice trained with broadband noise bursts and a modified behavioural paradigm. (A) Detection performance (d-prime) for sound-alone (blue) and multi trials (red). The large dots indicate means. Dashed lines connect matching sessions. **(B)** Detection performance (d-prime) for sound-alone (blue) and multi trials (red) by day. Day 1 was defined as the first day on which a mouse was introduced to multi-trials. **(C)** Detection performance (d-prime) for sound-alone (blue) and multi trials (red) for different mice. **(D)** Response rates of all sessions of all mice for sound-alone (blue) and multi (red) trials. Dashed lines connect matching sessions. The two large blue and red dots indicate the mean response rates across the two conditions. The large green and cyan dots indicate the mean response rates for whisker-alone and catch trials, respectively. **(E)** Response rates on sound-alone (blue) and multi (red) trials for different days. Small dots indicate individual mice, large dots indicate the means across mice. The large green and cyan dots indicate the mean response rates for

whisker-alone and catch trials, respectively. Day 1 was defined as the first day on which a mouse was introduced to multi-trials. **(F)** Response rates on sound-alone (blue) and multi (red) trials for different mice. Small dots indicate individual sessions, large dots indicate the means across sessions. The large green and cyan dots indicate the mean response rates for whisker-alone and catch trials, respectively. **(G)** Response latencies in sound-alone (blue) and multi -trials (red) for different mice. **(H)** Response latencies in sound-alone (blue) and multi trials (red) for all mice and all sessions.

2.2.3 Audio-Tactile Integration and Frequency Discrimination

Having assessed how tactile stimulation affects sound detection, we sought to establish whether it affects discriminability. To this end, we trained 10 mice on a go no-go frequency discrimination task. Mice were rewarded for licking after a 7 kHz target pure tone ('target-alone' trial) and received a 4s time-out for licking after distractor tones ('distractor-alone') (Figure 10). The distractor tone frequency was initially set to 12 kHz and as discrimination performance improved, that is whenever the d-prime for a given session exceeded 2, the distractor frequency was adjusted downwards. Poor performance (d-prime < 0.75) meant that the distractor frequency was adjusted upwards for the next session. The training period ended when the performance stopped improving and had stabilized. Mice took between 10 to 19 days to transition from training to test stage. At this point multi-trials, i.e. trials in which the sound presentation coincided with stimulation of the whiskers were introduced. 'Target-multi' and 'distractor-multi' trials were randomly interleaved with 'target-alone' and 'distractor-alone' trials.

In order to ascertain whether whisker stimulation affected discriminability, we then derived d-prime scores separately for multi trials and sound-alone trials. This revealed that performance was significantly better for sounds that were presented together with stimulation of the whiskers (Figure 11,

multi d-prime=1.65, sound only d-prime= 1.42, p=0.019, Wilcoxon signed rank test). The improvement in d-prime is the result of higher hit rates in combination with lower false alarm rates in multi-trials than in sound-alone trials (Figure 12). In contrast to the detection tasks, whisker stimulation had no impact on response timing (median sound-alone vs multi trial latency = 205 ms vs 206 ms, variance 14.8 ms and 14.6ms, p = 0.562 Wilcoxon rank-sum test, Figure 13) but, overall, response latencies in this tone discrimination were faster than in the tone detection task, possibly due to presentation level differences. Similar to the lower average d-prime performance, a tighter grip on the performance threshold through the automated script could have resulted in the sound pressure level being much lower in relation to the detectability threshold then in the previous (pure tone) cohort.

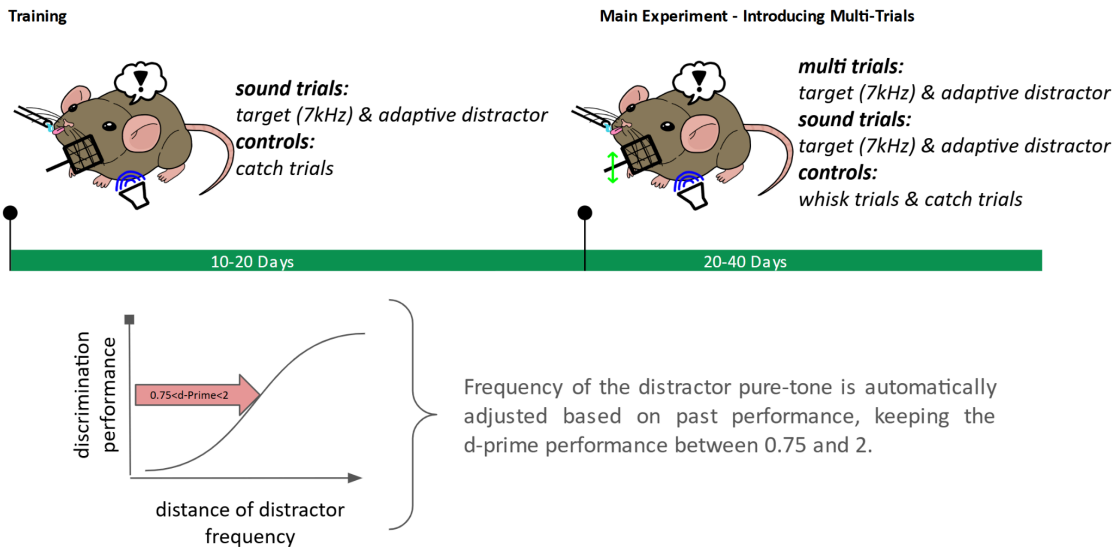

Figure 10: Training Procedure for the Frequency Discrimination Task. The training and experimental timeline for the go/no-go frequency discrimination task. During the training phase (10–20 days), mice were exposed to sound trials, consisting of a 7 kHz target pure tone and an adaptive distractor tone, alongside catch trials as control conditions. Performance was continuously evaluated, with the distractor tone frequency automatically adjusted based on past performance: frequencies moved closer to the target for high d-prime performance (d-prime > 2) and farther away for low performance (d-prime < 0.75), maintaining discrimination difficulty within an optimal range (0.75 ≤ d-prime ≤ 2). In the main experiment (20–40 days), multi-trials were introduced, where whisker stimulation coincided with the sound presentation. The experimental session consisted of target-alone, distractor-alone, target-multi, distractor-multi, and control (whisk and catch) trials.

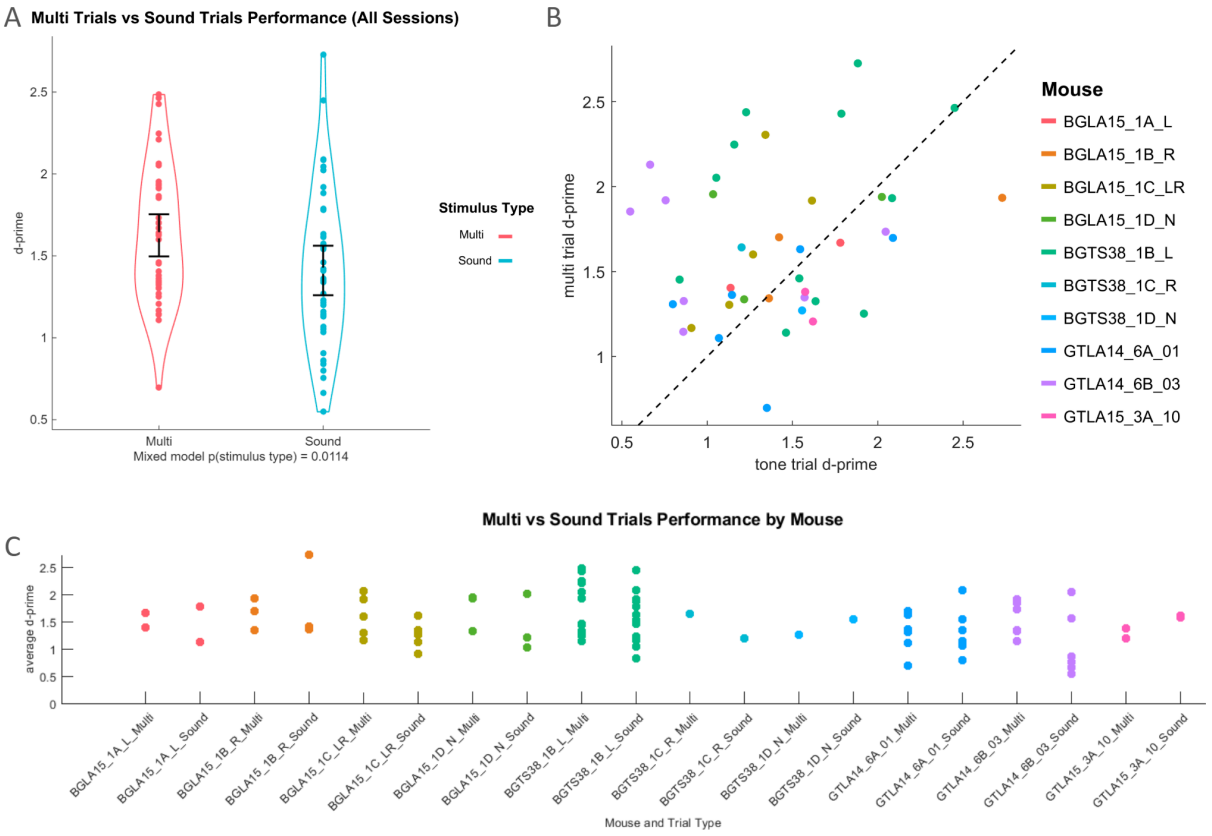

Figure 11. Performance in a go/no-go frequency discrimination task with and without whisker stimulation. (A) Discrimination performance (d-prime) in multi (red) and sound-alone (blue) trials for all mice and all sessions. Each dot indicates the data from one session. The black bars indicate the 95% confidence intervals. A linear mixed-effects model with random intercepts for mice revealed significantly higher d′ during Multi trials ($F(1, 82) = 6.70$, $p = 0.011$). Mean ± SE estimated d′: Sound = 1.40 ± 0.08; Multi = 1.61 ± 0.08; Δ = +0.22 (95% CI 0.05–0.38). (B) Discrimination performance (d-prime) in multi versus sound-alone trials, with each dot representing a session and being colour-coded for the identity of the mouse. The diagonal line represents equal performance between the two conditions. The clustering of data points above the diagonal indicates better performance in multi trials. (C) Discrimination performance (dprime) in multi and sound-alone trials separately for each mouse and session.

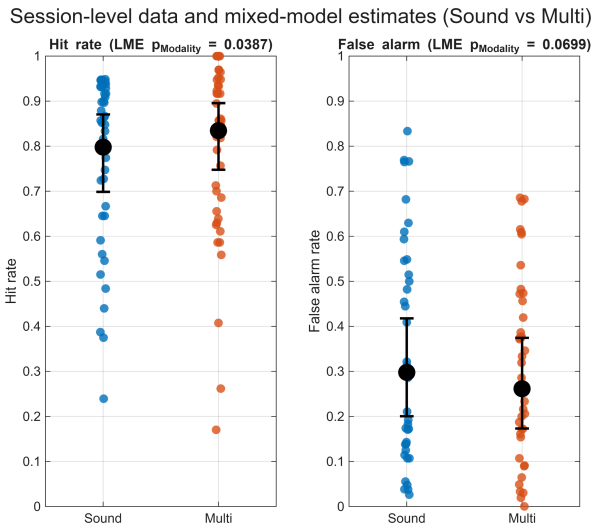

Figure 12. Session-level hit and false-alarm rates in the frequency-discrimination task with and without whisker stimulation. Each dot shows the hit rate (left) or false-alarm rate (right) for a single session in sound-alone (blue) and multi (orange) trials. Large black circles and error bars indicate the fixed-effect estimates (mean ± 95% CI on the probability scale) from logistic mixed-effects models with Modality (Sound vs Multi) as a fixed factor and random intercepts for Mouse and Session. Modality had a significant positive effect on hit rate ($F(1,82) = 4.42$, $p = 0.0387$) and a trend towards lower false-alarm rates in multi trials ($F(1,82) = 3.37$, $p = 0.0699$)), consistent with improved discrimination during multi trials.

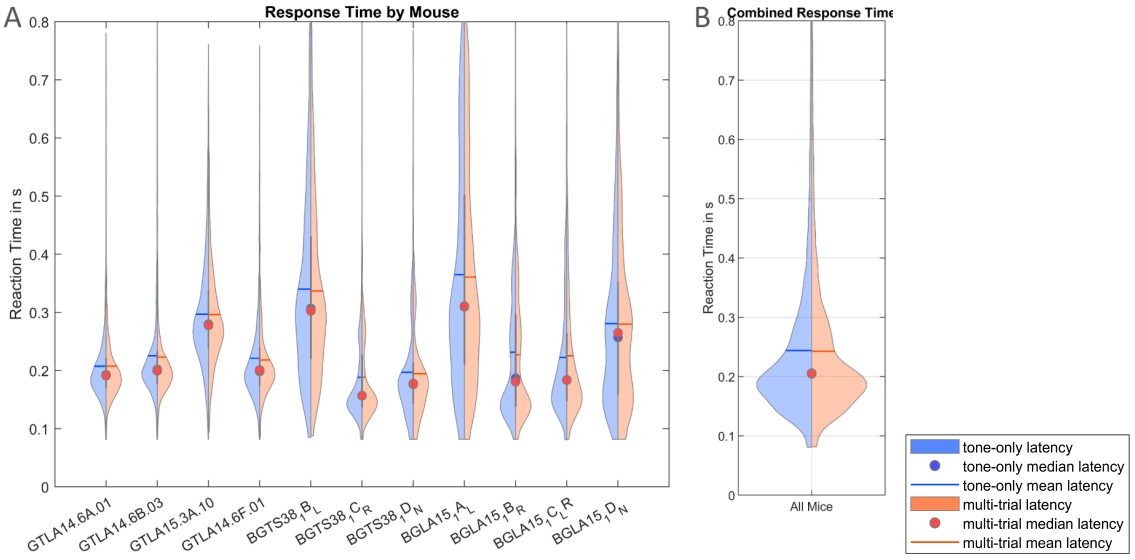

Figure 13. Response Latency in the Frequency Discrimination Task. (A) Response latencies in sound-alone (blue) and multi (red) trials separately for each mouse **(B)** and all mice combined . Dots indicate the medians and lines the means. Response latencies were remarkably consistent between sound-alone and multi trials, with no significant differences between the conditions for any of the mice or overall.

2.3 Discussion

Across three complementary behavioural paradigms we set out to investigate the perceptual implications of somatosensory-auditory suppression.

Whisker stimulation neither improved nor impaired the probability of detecting faint sounds, irrespective of whether the acoustic cue was a pure tone (Experiment 1) or a broadband noise burst (Experiment 2). Somatosensory inputs did, however, have a consistent effect on when the mice responded. In the original tone detection paradigm, where whisker stimulation had become

behaviourally irrelevant through weeks of exposure, response latencies shortened by ~5 % in the multi trials. In the modified noise paradigm, where the tactile stimulus was only introduced in the test phase, latencies lengthened by ~20 %. By contrast, in a go/no-go task that demanded fine spectral resolution, concurrent whisker stimulation improved d′ by ~16 %, driven by both higher hit rates and lower false-alarm rates, without any cost in response speed. Two-photon imaging confirmed that whisker input continued to divisively scale sound-evoked responses in core auditory cortex (12 % mean suppression) even after weeks of training, although the magnitude was smaller than that reported by Lohse et al. (2021) in passively-listening animals.

Collectively, these results reveal that the auditory perceptual consequences of concurrent somatosensory inputs and even the somatosensory-auditory suppression observed in the auditory cortex depend on task demands, stimulus context, and the animal's learning history.

2.3.1 Reconciling Neural Suppression with Preserved — or Even Improved — Behaviour

At first glance, it seems paradoxical that suppressing cortical activity can leave perception intact or even sharpen it. Yet across sensory systems a common motif has emerged: Suppression can lower absolute firing while redistributing dynamic range so that information about behaviourally critical dimensions is enhanced. Our data fit into that framework but must be interpreted in light of important training differences between the tasks.

Divisive gain control is most widely studied in vision. Classic contrast-adaptation experiments in V1 show that prolonged exposure to a high-contrast adaptor reduces neuronal responsiveness through a combination of synaptic depression and inhibitory normalization: excitatory drive becomes scaled down

relative to the population's overall activity, effectively compressing the gain of neurons most tuned to the adaptor. As a result, their firing rates to subsequent stimuli are suppressed, and the peak of their tuning curves shifts away from the adaptor (Dragoi et al., 2000, 2002). By reducing the overlap between neighbouring orientation channels the population's Fisher information rises, so orientation discrimination improves even though contrast-detection thresholds remain stable (Kohn, 2007). The same principle appears in multisensory settings. In mouse V1, a brief sound drives a layer-1 inhibitory circuit that divisively scales visual firing yet sharpens orientation tuning, boosting single-trial decoding accuracy—particularly for low-contrast gratings (Ibrahim et al., 2016). In macaque A1, adding the image component of a natural movie trims peak firing but quells trial-to-trial variability so effectively that mutual information about the soundtrack increases by ~30 % (Kayser et al., 2010). Co-located flashes suppress sound-driven activity but increase the information conveyed about stimulus location in ferret auditory cortex (Bizley & King, 2009) and improve sound-localisation behaviour (Hammond-Kenny et al., 2017). Furthermore, in humans a vibrotactile cue has been shown to speed up auditory-motion judgements without altering detectability (Soto-Faraco et al., 2004).

More recent large-scale recordings reveal that cross-modal modulation is neither uniform nor invariably divisive. In awake mice, only ~20 % of A1 neurons are influenced by a flash, and the sign of the effect depends on cell type and layer; nevertheless, the net population outcome is a *variance quench* reminiscent of contrast adaptation in V1 (Bigelow et al., 2022). Conversely, concurrent broadband noise in mouse V1 can elicit an *additive* gain: contrast-response functions shift upward in parallel, allowing a decoder to read out orientation ~15–20 ms sooner (Bigelow et al., 2022; Williams et al., 2023). Task rules can flip the sign altogether—an LED selectively adds spikes to the representation of a rewarded tone in rat A1 but leaves the opposite choice channel untouched, and the effect vanishes when the task is passive

or choices are unconstrained (Chang et al., 2022). Hence, it appears that multisensory circuits deploy divisive, additive and choice-gated mechanisms as the behavioural context demands.

Our three paradigms extend these ideas to somatosensory–auditory interactions in whisker-using mice. However, as differences between the training paradigms may lead to perceptual differences and may impact behavioural strategies, we must also consider these.

Table 1. Overview of training variables and performance consequences

Training variable	Tone-detection	Noise-detection	Frequency-discrimination
Whisker (or multi-) trials first presented	**Day 1** of training	**Test phase** only	**Test phase** only (after discrimination had stabilised)
Proportion of multi-trials / whisk-only trials	10 % / 40 %	30 % / 30 %	≈ 40 % (target-multi + distractor-multi) / 10 %
ISI rule	Random ISI timed from **last stimulus** (licks **do not** reset) → *no penalty for guessing*	Random ISI timed from **last lick** (each lick **resets** the timer) → *penalty for guessing*	Same "lick-reset" rule as noise-detection
Sound-alone median latency	**318 ms**	**238 ms**	**206 ms**
Multi-trial median latency	**302 ms**	**307 ms**	**205 ms**
Whisker effect on latency	**−16 ms** (≈ 5 % faster)	**+69 ms** (≈ 29 % slower)	**−1 ms** (not significant)
Whisker effect on performance	**not significant**	**not significant**	**improved d-prime +0.23** (≈ +16 %; $p < 0.05$)
Likely motivational / attentional consequence	Whisker cue long familiar → treated as background; more random licks and slightly worse overall performance, but when whisker coincides with tone, the responses may be more deliberate or enjoy an additive perceptive boost → **multi trials faster**	Whisker cue novel & salient → distracting; lick-reset penalty encourages deliberate licking; broadband noise already drives a large neural population, so redundant tactile input offers no extra speed benefit → **multi trials slower**	Whisker cue not informative for discrimination; high amount of multi-trials → multisensory input neither distracts nor strategically helps to decide trial outcome, **latencies stable**

Table 1 indicates that the three “whisker + sound” paradigms were not separated solely by the acoustic stimulus but by how—and when—the tactile channel entered the animal’s training history. Those procedural details provide a parsimonious account of the otherwise puzzling pattern in both latency and d′.

In the tone detection experiment, whisker stimulation was present from day 1. Mice therefore had become familiar with the whisker stimulus and potentially learned to ignore it by the time multi trials were introduced. When the familiar tactile input happened to coincide with a tone, however, it may have supplied a mild additive effect on thresholds or encouraged a more deliberate lick, trimming the median reaction time by ~5 % (318 → 302 ms) without changing sensitivity.

In the noise detection experiment, the tactile cue was introduced only during the test phase and, crucially, every premature lick reset the inter-trial timer. The opportunity cost of guessing rose sharply, prompting a more deliberate strategy, as reflected in shorter and narrower tails in the lick latency distributions and better overall performance: sound-alone responses were already faster than in the tone task (at 238 ms) but adding the novel and hence distracting whisker event may have incurred the delay (to 307 ms). As an additional explanation, the better training might already have yielded median reaction speeds near a perceptual limit or broadband noise may already recruit a wider neural population than pure tones, and hence the redundant somatosensory input brought no further advantage, leaving d′ unchanged.

In the frequency discrimination experiment, whisker input was equally paired with targets and distractors, so it carried no category information. Nevertheless, coupling the inter-trial interval to the last lick prevented guessing and improved task learning, as visible by much narrower reaction time tails. Latencies were near-identical for sound-alone and multi trials. What changed was discriminability:

Whisker stimulation improved d′ by ~16 % (Δd′ ≈ +0.23), possibly because divisive suppression narrowed spectral tuning and reduced overlap between frequency channels. The absence of a latency cost suggests that the sharpening benefit outweighed any momentary distraction perhaps due to the high rate of multi-trials compared to the sound detection experiments.

With those caveats in mind, the data suggests that somatosensory-evoked suppression operates as a context-sensitive filter: In near-threshold detection, divisive scaling leaves sensitivity intact and, once the cue is familiar, can perhaps even expedite decisions. In frequency discrimination the same scaling decreases channel overlap, improving discriminability. When the cue is novel or costly (noise detection with lick-reset ISI) attentional capture and strategic differences may outweigh sensory gain changes, delaying responses while sparing d′. That said, it could be argued that delayed detection is impaired detection and would result in lower detection rates and d-prime if the response window were sufficiently short.

Thus, while our findings are consistent with divisive-normalisation theory, they are not conclusive. They reinforce a broader literature showing that multisensory circuits flexibly mix divisive suppression, additive boosts and task-dependent routing to optimize exactly those dimensions of the neural code that matter for current behaviour—and they highlight how subtle shifts in training regimen can tip the scales. In the next chapter we will hence look into whether tactile-auditory suppression does or does not imply a fixed gain rule or can be adapted through stimulus relevance.

2.3.2 Limitations and Future Directions

As mentioned, Our “tone-detection”, “noise-detection” and “frequency-discrimination” cohorts differed not only in the acoustic cue but also in when and how often whisker and multi-sensory trials

were presented. Those procedural differences were deliberate, yet they mean that the three data sets are not strictly matched for prior sensory exposure, reward contingency or novelty of the tactile cue. Because learning can reshape cross-modal gain, some of the task-dependent effects we attribute to neural mechanisms could instead reflect history-dependent shifts in attention or expectation. A future factorial design in which stimulus statistics and reward contingencies are crossed systematically would allow a cleaner dissociation of these influences.

Throughout this chapter we have focused on the suppression measured in the core fields and have treated those signals as the neural substrate most tightly linked to the animals' detection and discrimination performance. That assumption is reasonable for near-threshold tasks because transient inactivation of the auditory cortex degrades both tone detection and fine frequency judgments in mice (Ceballo et al., 2019; H. K. Kato et al., 2015; T. Y. Lee et al., 2024; O'Sullivan et al., 2019) even though lesion studies suggest that mice can perform simple auditory tasks without auditory cortex (Ceballo et al., 2019; T. Y. Lee et al., 2024; O'Sullivan et al., 2019). Nevertheless, it is important to recognise that multisensory choices could be sculpted downstream, in the mouse equivalents of the primate belt. Higher-order auditory cortical fields receive dense non-lemniscal thalamic input, and harbour the highest fraction of neurons modulated by whisker or visual stimulation (Budinger et al., 2006), properties that closely parallel the caudomedial/caudolateral belt where supra-additive audio-tactile responses were first described in macaques (Kayser et al., 2005). Their broader receptive fields and heavier non-acoustic drive may make them even better suited than A1 for integrating touch with sound. Future experiments that extend two-photon or Neuropixels recordings into DP/A2, or that transiently silence these belt-like regions while preserving A1, will therefore be essential for pinning down the circuit locus of the perceptual effects reported here.

Evidence from anaesthetised macaques illustrates the point: high-resolution fMRI revealed supra-additive responses to concurrent tactile and broadband-noise stimulation in the caudal auditory belt, posterior and lateral to A1. These interactions obeyed both temporal-coincidence and inverse-effectiveness rules that are hallmarks of behaviourally useful multisensory integration (Kayser et al., 2005). While the primate auditory cortex comprises more fields than the mouse, rodent tracing and imaging likewise show that higher-order regions (e.g. dorsal posterior/temporal association cortex, the caudal A2 strip) receive stronger somatosensory and motor-related drive than A1. If those regions sit closer to the decision stage, correlations between A1 activity and behaviour could underestimate—or occasionally miss—the causal contribution of cross-modal signals encoded later in the hierarchy.

Two pragmatic factors kept our recordings in A1/AAF: (i) the need for optical access through a chronic cranial window, and (ii) the goal of comparing our data directly with earlier studies of whisker-evoked suppression that were also confined to primary fields. Nonetheless, future work should extend calcium imaging or Neuropixels sampling to other auditory fields and temporal association cortex—ideally coupled with pathway-specific perturbations—to test whether the task-dependent benefits and costs we report sharpen or even reverse at those higher tiers. Such experiments would also clarify whether the supra-additive gain seen in monkeys emerges in mice under more naturalistic, freely moving conditions.

The tone-detection paradigm purposely limited multi-trials to 10 % of the session in order to prevent reward-whisker associations. Although this achieved the intended behavioural control, it left relatively few multi-trials per sound level (often < 10), reducing statistical power and widening confidence intervals on both neural and behavioural metrics. A parallel issue arose from animal attrition: only 3 of 5 mice learned the noise-detection task (one of which only learned the task near the end of experiments and didn't yield sessions within our inclusion criteria) to criterion, and just 8 of 21

completed the discrimination task. While the effect sizes reported are robust to leave-one-out checks, larger cohorts will be needed to benchmark inter-animal variability and to permit sex-specific analyses.

All behavioural testing was done under head fixation so that two-photon imaging could be performed simultaneously. While necessary for cellular-level recordings, head fixation restricts the animal's natural posture. The enforced posture also required the whisker stimulator to deflect the whiskers within a vertical movement that differed from those observed during free-moving whisking. Consequently, the degree of somatosensory-auditory interaction could differ in freely behaving animals: We used a brush (tone task) or fine mesh (noise & discrimination tasks) that imposed a 1 mm, 20 Hz sinusoid over three cycles. Such global deflections activate many whiskers in near synchrony and may not mimic the spatiotemporal patterns generated during natural object contact. Because cross-modal effects in auditory thalamus scale with the number of simultaneously stimulated whiskers (Rao et al., 2014), our stimulator could overestimate—or underestimate—suppression relative to real-world whisking. High-speed videography in freely moving mice would allow a closer match to ethological conditions.

Finally, our imaging experiments are observational: we correlate the magnitude of whisker-induced suppression with behaviour but do not manipulate the circuit. Optogenetic suppression or activation of the putative S2 → IC → A1 pathway during behaviour, together with cell-type specific imaging or holographic stimulation, will be required to establish causality and to pinpoint which node(s) determine the task-dependent sign of the perceptual effect.

2.4 Methods

2.4.1 Animals

The local ethical review committee approved all experimental procedures, which were also licensed by the UK Home Office. The experimenter obtained a personal license (PIL A, B, & C) for animal testing and underwent specific training for the required tasks before conducting any experiments. Eight mice out of 10 expressing the fluorescent calcium indicator GCaMP6f in excitatory cortical neurons (T.-W. Chen et al., 2013) successfully learned the tone detection paradigm and three out of five mice the broadband noise detection task, of which one did not provide usable imaging data. For the frequency discrimination task, 8 out of 21 mice were included in the data-set. Here, two litters of mice from a separate genetic background (C57BL/6NTac.Cdh23$^{753A>G}$ (MRC Harwell, UK)) proved incapable of learning the task. All other mice were bred by crossing the floxed Ai95 (RCL-GCaMP6f)-D line with the CaMKIIalpha-Cre T29-1 line (Jackson Laboratories; stock numbers 024105 and 005359). All mice were between 7 and 20 weeks old at the time of data collection. They were kept on a 12-hour light/dark cycle and housed at temperatures between 20 and 24 °C with relative humidity levels of 45–65%. All experiments were conducted in sound-attenuated chambers.

2.4.2 Surgery

For all surgical procedures, mice were premedicated with intraperitoneal injections of dexamethasone (Dexadreson, 4 mg), atropine (Atrocare, 1 mg) and carprofen (Rimadyl, 0.15 mg) before being anaesthetised with isoflurane (1.5-2%) and administered with buprenorphine (Vetergesic, 1 ml/kg) postoperatively. Mice were then placed in a stereotaxic frame (Model 900LS, David Kopf Instruments)

and their body temperature was kept constant at 37°C by the use of a heating mat and a DC temperature controller in conjunction with a temperature probe (FHC). Head-bar implantation for mice that only underwent behavioural assessment without imaging involved removing a flap of skin of approximately 1 cm diameter over the centre of the skull and using dental acrylic (Super-Bond C&B, Sun Medical) to embed a custom steel post and to seal the exposed skull. Imaging window implantation involved the removal of slightly more skin on the right side of the skull as well as the partial removal of the right temporal muscle to make room for a circular 4 mm craniotomy covering the auditory cortex. A 4-mm diameter glass coverslip that had been glued to a ~1 mm tall steel cylinder with 0.5 mm wall thickness was inserted into this craniotomy. The cylinder allowed us to press the glass window gently onto the brain (in order to minimise brain movement during experiments) and was then glued (Pattex Ultra Gel, Henkel) to the edges of the skull. Mice were allowed to recover for at least one week before the first recording session.

2.4.3 Auditory Cortex Definition through Wide Field Imaging

In order to localise the different auditory cortical fields (A1, AAF, A2) so A1 and AAF could be targeted for 2P imaging, functional maps were attained through wide field imaging using a brief stimulation protocol and rapid analysis pipeline. Images were captured in a sound- and light-shielded chamber using a TL2x-SAP objective (Thorlabs), a 340M-GE CCD camera (Thorlabs), Thorcam 3.3 software, and a 470 nm LED (Thorlabs). The recording resolution was decreased to 128x96 pixels at an acquisition rate of 10 Hz controlled via a NI-DAQ I/O device (USB-6501, National Instruments). Sinusoidally amplitude-modulated (SAM) tones of 500 ms duration and 10 Hz frequency were used. These tones were produced at two sound pressure levels and four frequencies based on the mice's

auditory sensitivity, generated using a custom LabView script, and broadcasted through a calibrated speaker. During imaging sessions, mice were placed under the objective in a tube, with only their heads visible. SAM tones were presented sequentially with a three-second gap and repeated ten times. Data were analysed promptly, with additional repeats for enhanced signal-to-noise ratio. Slow drifts in the data were removed, and fluorescence normalisation was performed. Responses were examined by computing baseline and standard deviations prior to stimulus onset. Subsequently, the response mean was calculated in different time bins following the presentation and Z-scored relative to the pre-stimulus baseline.

2.4.4 2-Photon Calcium Imaging

Calcium imaging was performed using a two-photon laser-scanning microscope (B-Scope, Thorlabs) equipped with a 16x/0.80W water immersion objective (CFI75 LWD 16X W, Nikon). Excitation light at 930 nm was produced by a pulse-modulated laser (Mai-Tai eHP, Spectra-Physics) with a 70 fs pulse width and 80 MHz repetition rate. The laser beam was steered through a galvanometric scan mirror (for the Y-axis) and a resonant scanner (for the X-axis, bidirectional), enabling image acquisition at approximately 30 frames per second over a 512 x 512-pixel field of view. Emitted photons from the fluorescent neurons were filtered through a 525 nm bandpass filter (50 nm window) and detected by GaAsP photomultiplier tubes (Hamamatsu). The ScanImage toolbox for MATLAB was used for microscope control. To synchronise imaging frame numbers with behavioural data, a digital I/O device (USB-6501, National Instruments) was used, as two separate computers managed the imaging and behavioural data acquisition systems.

2.4.5 Image Processing

The raw ".tif" (Tagged Image File Format) imaging data were first processed using the suite2p toolbox (Pachitariu et al., 2017). The processing workflow included four main stages: 1) Registration of images to correct for any movement-induced drift; 2) Detection of Regions of Interest (ROIs) using clustering methods in a low-dimensional space; 3) Testing and labelling of ROIs based on criteria such as shape or correlation; 4) After this initial processing the imaging data were further neuropil corrected using a coefficient of 0.7 (Kerlin et al., 2010) and calcium $\Delta F/F_0$ traces were obtained by using the median over the entire fluorescence trace (excluding fluorescence values below the 10th and above the 70th percentile) as the F_0.

2.4.6 Behavioural Training, set-ups and stimuli

All auditory stimuli were generated at a 192 kHz sampling rate using a 16-bit sound card (Startech) and delivered via an ultrasonic speaker (Avisoft Vifa) positioned 23 cm in front of the mouse and slightly off to its left. Stimuli were linearly calibrated using a precision microphone (G.R.A.S. 40BE ¼" microphone, 26AK Preamplifier) to ensure consistent sound pressure levels (SPL) at the position of the mouse's head.

Somatosensory stimuli were applied to the left whiskers using a custom-built piezoelectric whisker stimulator (Lohse et al., 2021). Two stimulator designs were employed. (1) A brush-like stimulator, consisting of a rigid bar attached to a piezoelectric actuator, imparted a 1 mm vertical sinusoidal displacement at 20 Hz (corresponding to a peak velocity of ~0.13 m/s). Stimulation consisted of three consecutive cycles lasting 150 ms in total. Contact occurred as the brush pressed upward against the

whisker pad, deflecting initially about 15 of the whiskers simultaneously in a smooth, oscillatory manner. However, because the mice were often able to rapidly pull their whiskers free of the brush, (2) a fine wire mesh stimulator was introduced to increase whisker engagement. The mesh, also driven by an identical piezoelectric actuator, contained approximately square holes with a side length of ~0.7 mm, in which about 8+ individual whiskers were transiently trapped during deflection. This produced a more secure contact than the brush, engaging whiskers at multiple points while still delivering the same overall displacement and velocity profile. The specific stimulator type used depended on the experiment (details below).

Licking was monitored using a custom lickometer, involving a conductive foil sheet covering the bottom half of the acrylic tube in which the mouse's body was positioned and a metallic lick spout wired into a circuit that completed upon tongue contact. This setup enabled high-resolution lick detection (~1 ms accuracy) and precise synchronisation with stimulus presentation and neural activity imaging.

All stimuli, behavioural responses, and reward delivery were coordinated via a MATLAB-controlled system (release R2015b; MathWorks, MA, USA) with Psychophysics Toolbox extensions. The system interfaced with a data acquisition device (USB-6008; National Instruments, TX, USA).

To ensure sufficient motivation for task engagement, water access was restricted to 1ml/day for 2–3 days prior to training, reducing body weight to approximately 85% of baseline. During both training and testing phases, mice were head-fixed under a two-photon microscope, enabling simultaneous recording of neuronal activity.

2.4.6.1 Pure Tone Detection

Auditory stimuli consisted of 12 kHz pure tones, each lasting 50 ms. The sound pressure levels were manually adjusted according to the animals' detection performance. Once a detection threshold was established (defined as the sound pressure level at which the mouse detected tones 50% of the time), tones were presented at four different levels: three near the threshold and one at a louder, easily detectable level. The brush-like whisker stimulator was used to deliver somatosensory stimuli, imparting a 1 mm vertical sinusoidal displacement at 20 Hz over three cycles lasting 150 ms (as discussed in more detail above).

Mice were trained to associate the auditory stimulus with a water reward by licking a spout. Mice received a drop of water for licking the spout within 1.5 seconds of sound onset during sound-alone and multi-trials. Only trials in which a lick occurred within 0.85 seconds of sound onset were counted as hits to accurately assess detection performance. No rewards were provided for licking during whisker-alone or catch trials.

Auditory and somatosensory stimuli were presented in a randomly interleaved manner, with trial types distributed as follows: 40% sound-alone trials, 40% whisker-alone trials, 10% multi-trials (simultaneous sound and whisker stimulation), and 10% catch trials (no stimulus). The inter-stimulus interval (ISI) was randomised between 4 and 7 seconds, based on the timing of the previous stimulus presentation.

During early training phases, mice occasionally received manual water rewards without licking, particularly during sound-alone trials. This was done to encourage initial engagement with the lick-spout and to help establish an association between the auditory stimulus and the reward. Sessions involving

manual rewards were excluded from d-prime calculations to avoid skewing measures of detection sensitivity.

Once mice reached a criterion d-prime value of at least 2.0, indicating reliable discrimination between sound and whisker stimuli, multi-trials were introduced at a 10% probability. This low probability was chosen to ensure that whisker stimulation did not become associated with reward availability. Catch trials were included as a control condition, monitoring for any learning of the reward schedule based on the ISI or other unintended cues.

2.4.6.2 Broadband-Noise Detection

The noise detection experiment used broadband noise bursts (1–64 kHz), each lasting 50 ms, instead of pure tones. This stimulus type was chosen to recruit a higher proportion of neurons. A fine mesh whisker stimulator replaced the brush-like structure used in the other experiment to improve contact between the whiskers and the stimulator.

To determine the appropriate sound pressure level for each session, an automated threshold-finding script was used at the beginning of each recording session of the main experimental phase. This script presented a range of sound levels near those used in the most recent training or recording session. Eight sound levels were tested, and a logistic function was fitted to the detection rates to calculate the threshold. Only a single near-threshold sound level was then used during the session.

Behavioural training incorporated a dynamic inter-stimulus interval (ISI) that was randomised between 4.5 and 9 seconds, measured from the last lick of the mouse to reduce the likelihood of random hits. Multi trials (noise + whisker stimulation) and whisker-alone trials were introduced only after mice were reliably trained, to minimise potential adaptation effects to somatosensory stimuli.

During the test phase, trial types were distributed as follows: 30% sound-alone trials, 30% whisker-alone trials, 30% multi trials, and 10% catch trials (no stimulus). The training script automatically terminated once a total of 30 multi trials had been presented to ensure a balanced trial distribution across the session.

The refined training and experimental setup consistently produced more reliable detection performance while minimizing potential confounds related to task structure and stimulus adaptation.

2.4.6.3 Frequency Discrimination

The frequency discrimination task was designed to evaluate whether tactile stimulation affects the ability of mice to distinguish between acoustic frequencies. Mice were trained on a go/no-go paradigm where they were rewarded for licking a spout after hearing a 7 kHz target pure tone ('target-alone' trials) and penalized with a 4-second timeout for licking after distractor tones ('distractor-alone' trials).

During the training phase (10–20 days), distractor tones were initially set to a frequency of 12 kHz. To dynamically adapt the task difficulty, the distractor tone frequency was adjusted based on the mice's performance. If a mouse achieved a session d-prime score greater than 2, the distractor frequency was moved closer to the target tone in subsequent sessions. Conversely, poor performance (d-prime < 0.75) resulted in the distractor tone frequency being moved further away from the target. This adaptive approach ensured that the task remained appropriately challenging while enabling the animals to maintain discrimination performance within an optimal range ($0.75 \leq$ d-prime ≤ 2). The training phase concluded once performance stabilized, typically within 13 days.

The main experimental phase (20–40 days) introduced multi-trials, in which whisker stimulation coincided with sound presentation. Trial types were distributed as follows: target-alone trials, distractor-alone trials, target-multi trials, distractor-multi trials, and control trials (catch trials and

whisker-alone trials). The inclusion of multi-trials allowed us to investigate whether whisker stimulation affected acoustic discriminability.

2.4.7 Behavioural Performance

The discriminative sensitivity index, commonly referred to as d-prime, a metric derived from signal detection theory (Stanislav & Todorov 1999), which is robust against fluctuations in response bias (Macmillan and Douglas Creelman 2004), was calculated to evaluate the ability of the mice to detect the presence of a target stimulus.

The calculation of d-prime is based on the hit rate (pHit) and the false alarm rate (pFA). The hit rate represents the proportion of trials in which the mice correctly identified the presence of a stimulus (either sound-alone, multi or whisker-alone trials), indicated through at least one lick of the reward spout, while the false alarm rate denotes the proportion of catch trials (trials without stimulus) in which the mice incorrectly indicated the detection of a stimulus. The formula for d-prime (d') is derived from the Z scores of the hit and false alarm rates, calculated as follows:

$$d' = Z(pHit) - Z(pFA)$$

In scenarios where (pHit = 1) or (pFA = 0), corrections are applied based on the numbers of target (nTarget) and distractor (nDistract) trials, respectively, to avoid infinite values. This correction is known as the "1/2N rule", which slightly adjusts perfect hit rates and zero false alarm rates towards more realistic values by adding 0.5 to the count of hits (or false alarms) and 1 to the total number of relevant trials, mitigating the issues associated with calculating Z scores for probabilities of 1 or 0.

2.4.8 Selection of Neurons Included in the Analysis

Fluorescence signals within each ROI were compared using a one-sided paired t-test to assign neurons to these categories. The 30-frame baseline period (~1 second) prior to trial onset was contrasted with the response during frames 2 to 15 post onset, and a significance threshold of $p < 0.05$ was employed. To ensure that neurons were categorized purely on the basis of their stimulus responses, rather than neural activity associated with reward expectation, collection or licking activity, only miss trials were used for this step.

2.4.9 Adjustments for Multisensory Trial Scarcity and Variability in Detection Performance

To avoid the mice learning to associate the whisker stimulus with rewards, we limited the proportion of multisensory trials to 10% of the total trial count in the first detection experiment. This intentional constraint meant that the number of multi trials available for analysis per sound level and day tended to be small. Therefore, we aggregated trials across the three sound levels of a session. To avoid ceiling and floor effects, we required the d-prime for each sound to be between 0 and 2.5 and its detection rate lower than 0.95 and higher than 0.25. Furthermore, the session needed to have at least 5 multi-trials with one of the relevant sound levels in order for the session to be used for analysis.

Another issue associated with the low number of multi-trials, was that the confidence intervals of the evoked responses tended to be higher for multi-trials than for the other conditions, which had more responses per trial-type available to calculate the mean evoked response. This in turn means that, on average, fewer neurons would pass a significance test for responding to multi-trials compared to other

trials, whose high numbers allowed for higher confidence. This does not affect the majority of results, which are based on neurons that respond significantly to sounds. To achieve meaningful comparisons between these fractions of neurons, the number of tone- and whisker-trials included for the significance testing was reduced to match the number of multi-trials that were available in a given recording session.

III Impact of Behavioural Context on Multisensory Integration

3.1 Introduction

In the previous chapter, we examined how whisker stimulation can modulate auditory perception in mice performing sound detection or frequency discrimination tasks. All paradigms had in common that the reward contingencies remained the same throughout the experiment insofar as it was always an auditory stimulus that was rewarded and the whisker stimulation that was irrelevant. We demonstrated that whisker stimulation can shorten near-threshold tone detection latencies and enhance frequency discrimination. Whisker-induced suppression at the cortical level tended to be modest compared to passive listening conditions, potentially because mice had learned to ignore the whisker stimulation. This might suggest that the circuit was able to operate with a certain degree of flexibility depending on the behavioural context and could downregulate crossmodal gating when the auditory domain is important.

As the general introduction describes in more detail, a wealth of evidence demonstrates that integration of tactile and auditory signals can arise at multiple subcortical and cortical levels (Bizley, Jones, et al., 2016): from the cochlear nuclei (Ansorge et al., 2021; Podachin et al., 1979) and inferior colliculus (Lohse et al., 2021) to the medial geniculate body (Khorevin, 1980; Kimura & Imbe, 2018) and auditory cortex (Bizley, Jones, et al., 2016; Budinger et al., 2006; Lakatos et al., 2007; Meredith & Allman, 2015). The processing stage at which different sensory modalities are integrated possibly depends on the nature and saliency of the stimuli and the behavioural state of the animal, with the existence of multiple points of integration potentially allowing for a trade-off between higher speed/efficiency (subcortical integration) and higher flexibility (cortical integration).

While whisker stimulation often suppresses auditory responses (Kimura & Imbe, 2018; Lohse et al., 2021; Meredith & Allman, 2015; Podachin et al., 1979) through mechanisms such as divisive

normalisation or competitive gating (Ohshiro et al., 2011, 2017; Yu et al., 2019), facilitative effects have also been documented, depending on the neural circuitry, sensory context, and behavioural goals (Fu et al., 2003; Kayser et al., 2005; Lakatos et al., 2007; Lohse et al., 2021). For example, a direct S1→MGB projection can enhance certain auditory responses destined for non-lemniscal or limbic pathways (Lohse et al., 2021), and primate studies reveal that specific auditory cortical subdivisions can show phase resets or increased spiking in response to tactile cues (Fu et al., 2003; Kayser et al., 2005). Thus, somatosensory and auditory signals may interact in either suppressive or facilitative ways, depending on whether tactile contact is incidental or highly salient for orienting, affective appraisal, or broader multisensory object representations.

The suppressive effects Lohse et al. (2021) demonstrated in passive listening conditions in mice and which we also observed during goal-directed behaviour in chapter I, are implemented through a subcortical S1→LCIC→MGB pathway that bypasses direct S1→A1 or classic corticothalamic connections. While S1 does send additional projections to more medial sectors of the MGB, these do not appear to counteract the dominant suppression of thalamocortical drive to A1, as auditory cortex input remains strongly suppressed by whisker stimulation. Thus, although the S1→medial MGB route may play an important role in shaping behaviourally-relevant signals elsewhere (e.g., to the amygdala), it is unlikely to underlie the apparent flexibility observed at the level of A1. Instead, this flexibility may reflect dynamic modulation within the S1→LCIC→MGB→A1 loop itself, potentially allowing the cortex to exert contextual control over how auditory information is filtered at the level of the midbrain and thalamus.

In this chapter, we investigate whether mice can switch between detecting an auditory stimulus and detecting whisker stimulation and how that switch in behavioural context and reward contingency affects stimulus processing at the level of the cortex. Among other things, we expected that imparting

behavioural relevance to the somatosensory stimulus and, in turn, making the auditory input irrelevant would maximise somatosensory-auditory suppression. Vice versa, we reasoned that teaching mice to ignore the whisker stimulation would diminish its capacity to suppress sound-evoked activity.

3.2 Results

We trained water-restricted mice to perform one of two behavioural tasks: A tone detection task, in which mice were rewarded for licking in response to an auditory stimulus, and a whisker detection task, in which rewards were contingent on licking in response to whisker stimulation (approximately 1 mm vertical whisker displacement). Across both tasks, catch trials (no stimulus), and same stimulus trials (sound-alone trials and whisker-alone trials) were presented randomly interleaved with equal probability.

To investigate not only how reward contingency influences the processing of auditory and somatosensory stimuli, but also how these stimuli are integrated, we introduced multi-sensory trials ('multi-trials'). In these trials, sound presentation and whisker stimulation occurred simultaneously. Pilot testing revealed that introducing multi-trials from the outset made it more challenging for mice to learn the task. Therefore, the probability of multi-trials was initially set to 0% and gradually increased to 30% as the mice's detection performance improved.

The performance criterion for mice to progress to the next stage of the experiment after initial training was achieving a sensitivity index d′ of at least 2.25 on the previous training day, or maintaining a d-prime of at least 1.5 for two consecutive training days. Once they progressed consistently through the stages of 0, 10, 20 and 30% probability of multi-trials, the reward contingency was inverted. Mice initially trained on the sound detection task were then rewarded for licking only during whisker-alone trials,

while mice initially trained on the whisker detection task were then rewarded for licking only during sound-alone trials (See Methods: Behavioural Training for details).

3.2.1 Task Learning and Adaptation to Switched Reward Contingencies

We trained 17 mice, 13 (see Figure 15 for individual d-prime performance curves) of which reached criterion and were then trained on the other task. Mice successfully learned both the sound and whisker detection tasks but those trained first on the sound detection task reached criterion for task-switching within 9–12 days, while those starting with whisker detection took 18–22 days, prompting us to analyse differences in performance between the tasks in more detail.

During initial training, sound-rewarded mice showed a rapid increase in d-prime values, reaching a mean of 3.5 within 10 days, while whisker-rewarded mice reached a d-prime of only about 2.0 over the same period (Figure 14, A), indicating that the sound detection task was easier for the mice to learn.

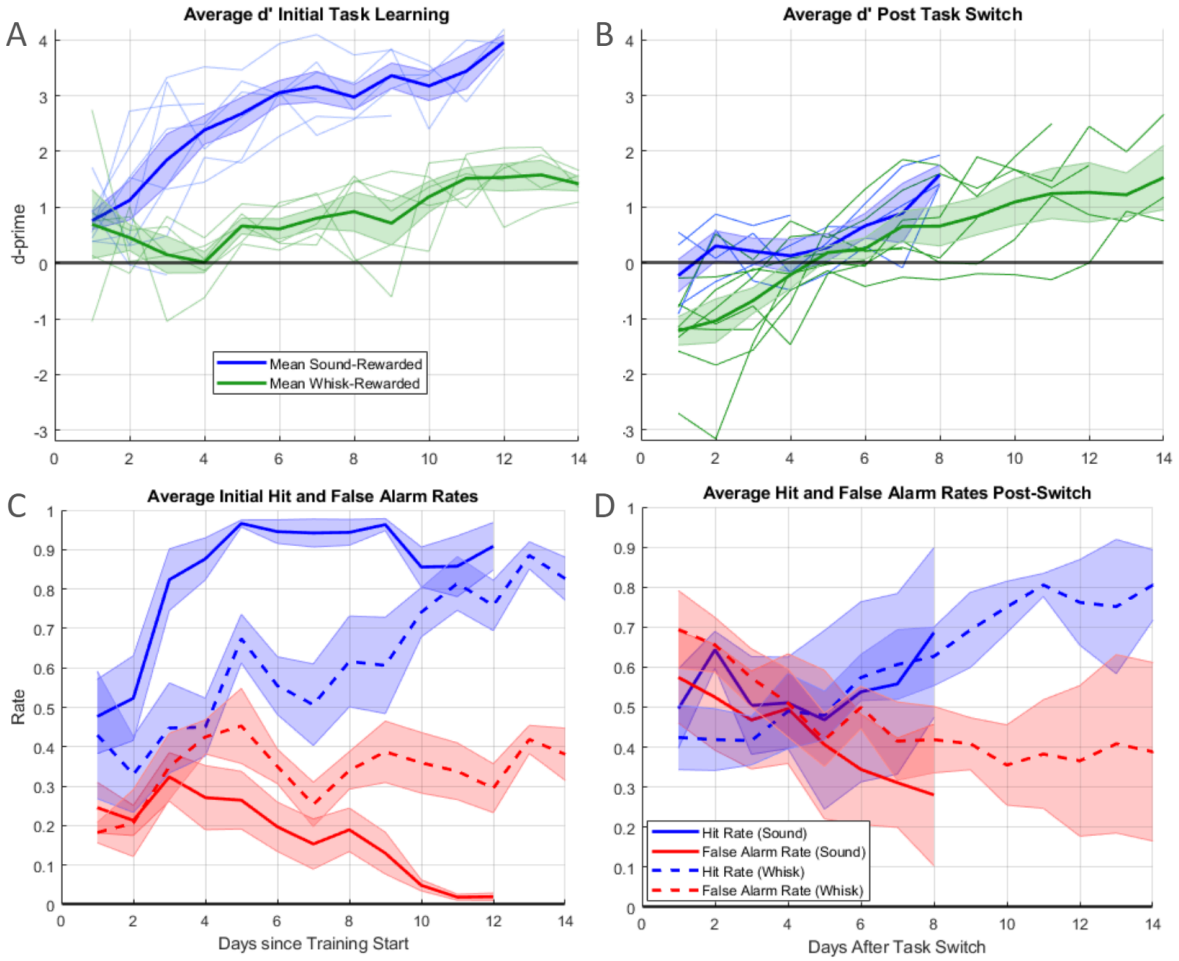

Figure 14. Learning curves for initial task learning and after switch of reward contingency. **(A-B)** Mean d-prime across mice as a function of training day for initial training (**A**) and after task switch (**B**) for sound-rewarded (blue) and whisker-rewarded (green) sessions. **(C-D)** Hit and false alarm (FA) rates during initial training (**C**) and after task switch (**D**). Shaded areas represent the standard error of the mean.

Hit and false alarm (FA) rates (Figure 14, C) provide further insight into the different learning rates. FA rates during training were based on the non-rewarded stimulus. In the sound detection task, the mean hit rate approached 1 within 5 days, suggesting that mice rapidly associated the sound with availability of reward. Further improvements in d-prime after day 5 came about primarily through a

steady decrease in the FA rate. In the whisker detection task, the mean hit rate increased much more gradually , while FA rates remained relatively unchanged and, thus, prevented d-prime values from rising to the same level as in the sound detection task. Why FA rates failed to decrease in the whisker detection task is unclear but could potentially suggest confusion or uncertainty about task rules (Fig 14, 15).

When the reward contingencies were reversed and the mice were switched from the sound detection task to the whisker detection task and vice versa, performance plummeted with d-prime values sometimes remaining negative for several days, indicating that mice still behaved more according to the old reward contingency than the new one, especially those mice that initially learned the sound detection task. After the task switch, the learning rates, as measured by d-prime, were much more similar between the two tasks, mostly because mice learned the sound detection task much more slowly than before the switch with the mean d-prime exceeding 1 only on day 6 (vs day 2 pre-switch), whereas they learned the whisker detection task no more slowly than before the switch (mean d-prime exceeded 1 on 10 day both before and after the switch, Figure 14, B). This is also reflected in the hit rates which are much more closely matched between the sound and the whisker detection tasks than before the switch. FA rates, however, diverge in a similar way as during the pre-switch learning with FA rates in the whisker detection task remaining much higher than in the sound detection task (Figure 14, D).

This pattern of results also indicates that pre-switch differences between the tasks are unlikely to be due to differences between the cohort of mice starting on the sound detection task and the cohort starting on the whisker detection task and instead suggests that mice generally find our whisker stimulation task much harder to master than the sound detection task. One exception was the mouse GTLA10_1E, which underwent a second task switch but failed to adapt successfully. Despite positive d-prime values during the original and first switch phases, this mouse maintained those values through the second switch without significant improvement (Figure 16).

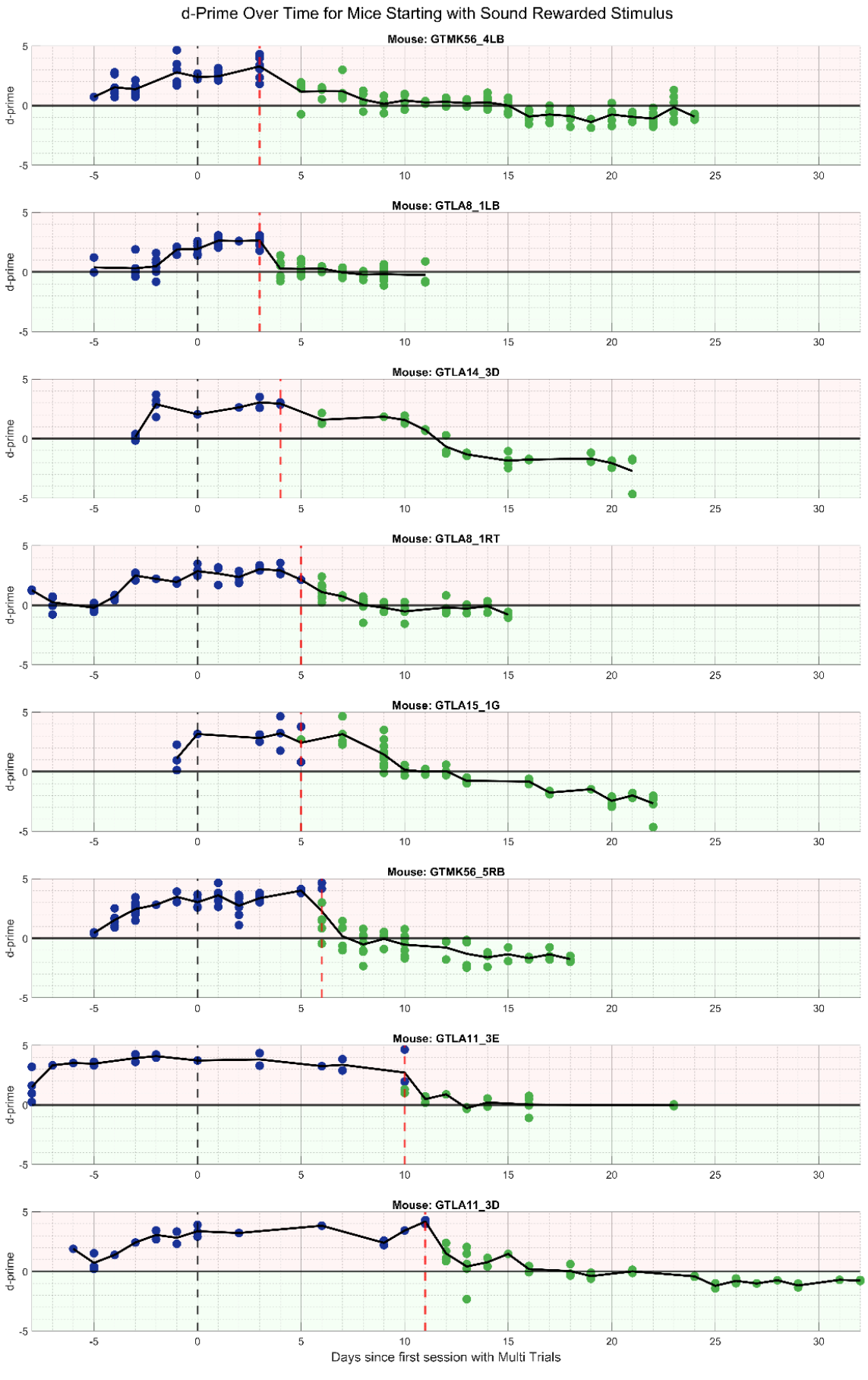
d-Prime Over Time for Mice Starting with Sound Rewarded Stimulus
Mouse: GTMK56_4LB
Mouse: GTLA8_1LB
Mouse: GTLA14_3D
Mouse: GTLA8_1RT
Mouse: GTLA15_1G
Mouse: GTMK56_5RB
Mouse: GTLA11_3E
Mouse: GTLA11_3D
d-prime
Days since first session with Multi Trials
Sound Rewarded
Whisk Rewarded
Task Switch
First Multi-Trial Day

Figure 15. Individual learning curves. d-prime as a function of time for all mice trained on the sound detection task first. Each dot indicates the d-prime of one session in the sound (blue) or the whisker (green) task and the continuous line represents the mean d-prime of all sessions from a given day. The red dashed line marks the day on which the reward contingency was switched and the black dashed line indicates the first day on which the mouse encountered multi-trials. Background shading distinguishes positive (red) d-prime values from negative (green) d-prime values. Near the beginning of training, "free" rewards were given occasionally to help build the association between stimulus and reward. Those sessions were excluded. Additionally, there were a few days where not all mice could be run, either due to debugging procedures or unavailability of the experimenter together accounting for some additional days with missing sessions.

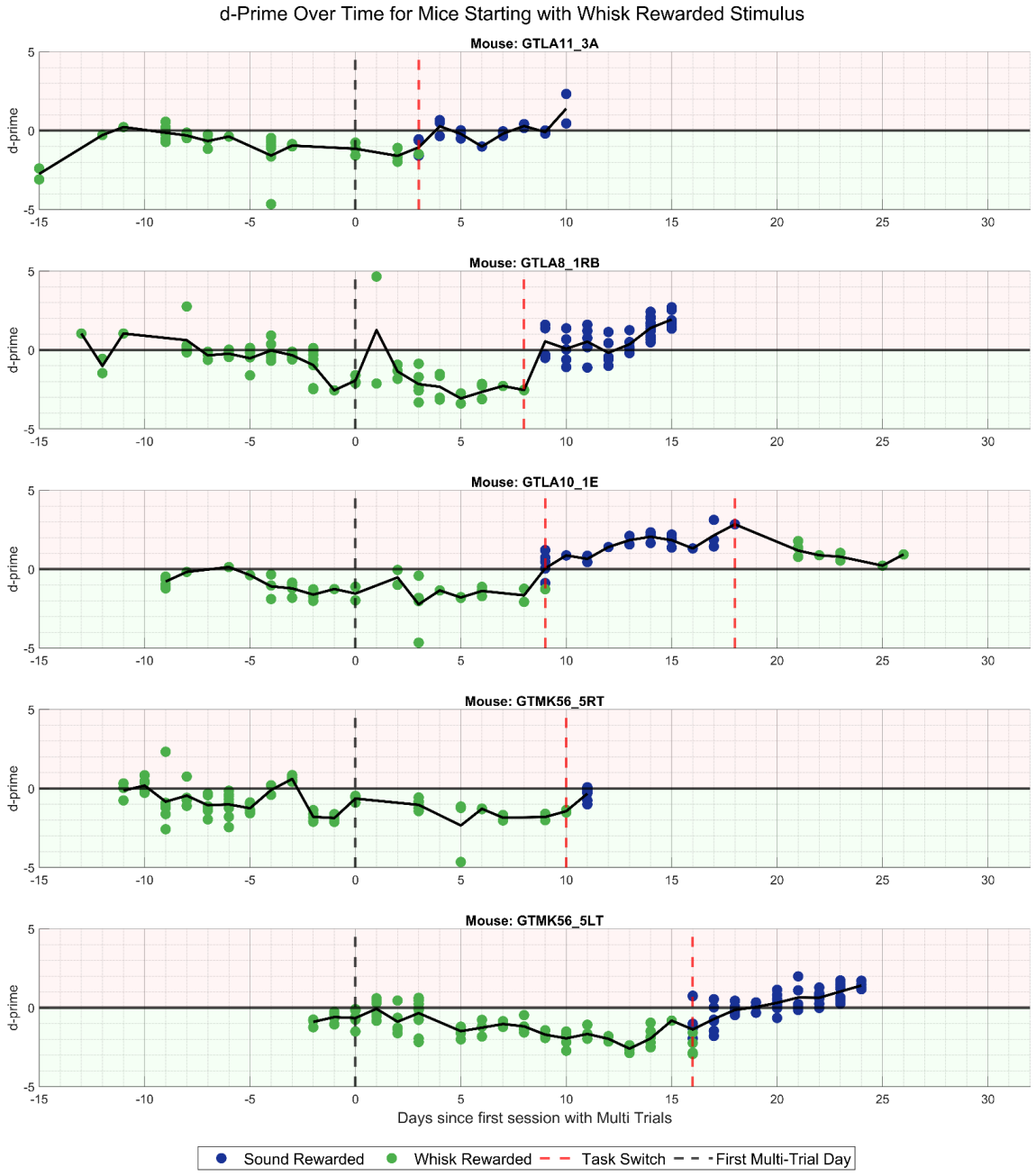

Figure 16. Individual learning curves. Same as Figure X2 for mice trained on the whisker detection task first. GTLA10_1E underwent a second task switch.

3.2.2 Evoked Responses

While the mice performed the task, we recorded the activity of excitatory neurons in the auditory cortex using two-photon calcium imaging. To find out whether the reward contingency affects cortical sound processing, we compared the sound-evoked activity of neurons in mice engaged in the sound detection task with the sound-evoked activity of neurons in mice engaged in the whisker detection task. The neural activity observed on a given trial may reflect brain activity associated with licking and other movements related to responding to the stimulus as well as, potentially, consuming the reward, rather than merely a response to the sensory stimulus that was presented on that trial. This becomes evident when we compare the response profiles recorded during trials in which the mice licked (hit trials and false alarms) (Appendix Supplementary Figure S1) with those recorded during trials in which they did not (miss trials and correct rejections, Figure 17). Therefore, we only included trials in this and subsequent analyses in which the mice did not lick (miss trials).

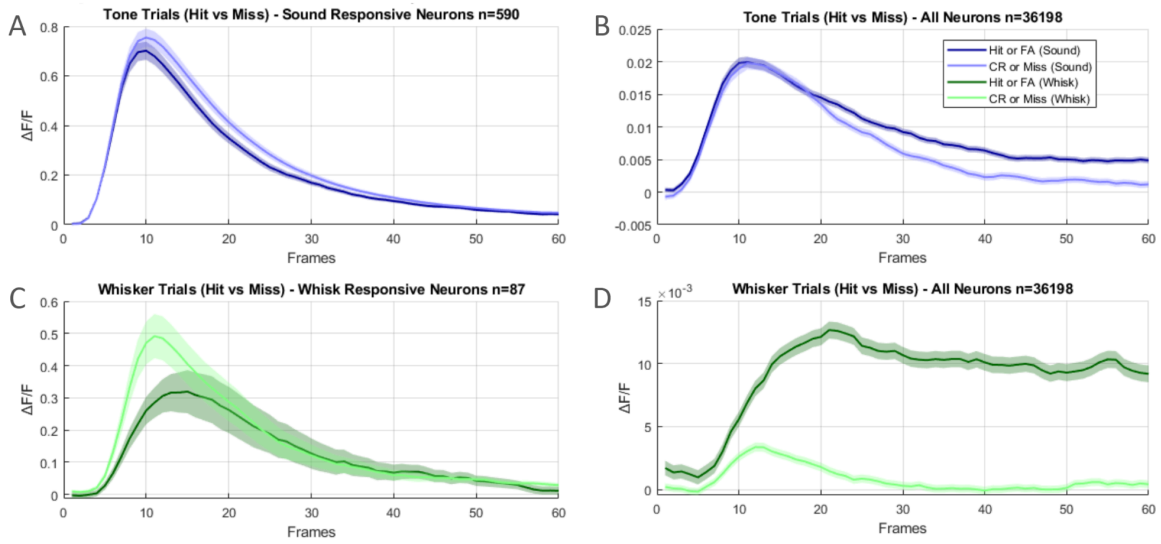

Figure 17. Averaged response profiles of neurons recorded in the auditory cortex for different trial types and outcomes. (A&B) Response profiles for different trial outcomes (hits & false alarms vs misses & correct rejections) recorded during sound-alone trials, depicted separately for sound-responsive neurons ($p = 0.051$, left, Wilcoxon sign-rank test) and all neurons ($p < 0.001$, right, Wilcoxon sign-rank test). **(C&D)** Response profiles for different trial outcomes (hits & false alarms vs misses & correct rejections) recorded during whisker-alone trials, depicted separately for whisker-responsive neurons ($p < 0.001$, left, Wilcoxon sign-rank test) and all neurons ($p < 0.001$, right, Wilcoxon sign-rank test). Shaded areas represent the standard error of the mean. All significance testing was done using a sign-rank test between the average of frames 6 to 15 from each neuron.

3.2.2.1 Sound Evoked Activity in Sound- vs Whisk-Task

The experimental design, training the same mice with one reward contingency and then the other reward contingency, was meant to enable within-mouse comparisons between conditions. Furthermore, we hoped that mice would learn to switch between reward contingencies quickly - if not within the same

day then within a few days - as that would have enabled us to make within-neuron comparisons between conditions. However, it took the mice much longer than anticipated to learn the whisker stimulation task. Furthermore, learning after the switch in reward contingency was no faster than the initial learning, either for the auditory or the somatosensory task. As a consequence of this protracted period of training, visibility deteriorated significantly in several mice, which meant that we did not make an attempt to match fields of view between conditions to make within-neuron comparisons and that, in some mice, imaging became impossible before the training had been completed. Moreover, not all mice achieved good, let alone expert, performance in both tasks. Setting a very low performance threshold would have allowed us to make a within-mouse comparison for nine mice (Figure 18, A). However, to make meaningful comparisons, we need to contrast data obtained when mice performed at expert level (d-prime > 2) in one condition with data obtained when mice performed at expert level in the other condition. Therefore, we pooled all sound-responsive neurons from all mice and all sessions in which performance was at d-prime > 2. Sound-evoked activity was significantly lower in the sound detection task than in the whisker detection task (Figure 18, B) and lower than the sound-evoked activity recorded before training. This finding is expected, as previous studies have shown that task engagement suppresses sound-elicited activity in the auditory cortex (Kuchibhotla et al., 2017; Otazu et al., 2009; Town et al., 2018) relative to passive conditions.

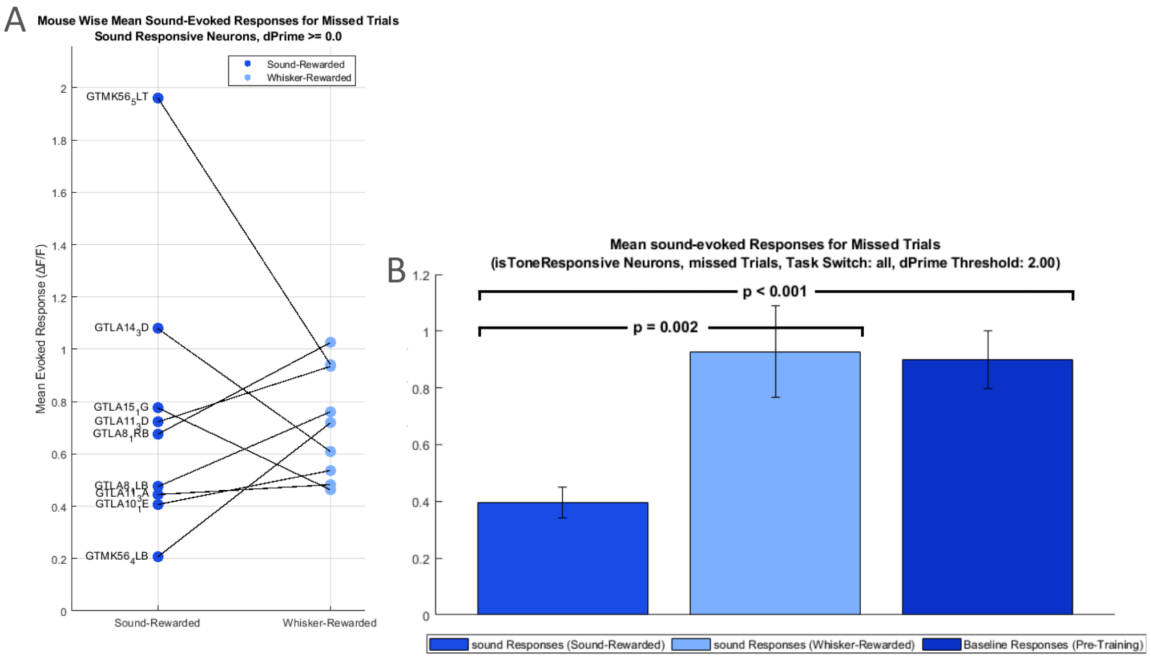

Figure 18. Sound-evoked responses in different behavioural contexts. (A) Mouse-wise comparison of mean sound-evoked activity of all sound-responsive neurons in the sound task (dark blue) and the whisker task (light blue). Only trials without licking were considered (miss trials and correct rejections). Inclusion criterion: d-prime ≥ 0. **(B)** Mean sound-evoked activity of all sound-responsive neurons recorded in three different contexts: performing the sound task, performing the whisker task, passive stimulus exposure before any training had been conducted. Inclusion criterion for behavioural performance: d-prime ≥ 2.

To further evaluate the relationship between task performance and sound-evoked activity, we exploited the anticommutative property of d-prime, where d-prime(sound, whisk) = −d-prime(whisk, sound), allowing us to use a continuous performance metric across both tasks. A negative d-prime(sound, whisk) value indicates that the mouse responded in a greater proportion of whisker-alone trials than sound-alone trials, while a positive value indicates the opposite. Note that this analysis ignores the reward contingency, i.e. whether the mouse responded correctly or not. So a mouse could have a negative d-prime(sound, whisk) in a session, indicating a preference for whisker stimulation,

even though it was not rewarded for licking in whisker-alone trials in that session, i.e. it was meant to detect sounds, and vice versa. This kind of contingency-inconsistent behaviour was typically observed after a switch in reward contingency and sometimes persisted for several days (Figure 15, 16). Regressing d-prime(sound, whisk) against sound-evoked activity (Figure 19) suggests that sound-evoked activity scales with task performance even though the regression does not quite reach statistical significance ($F = 2.83$, $p = 0.096$, F-test for the regression (equivalent to a two-sided test of non-zero slope). This finding suggests that engagement in a task involving a different sensory modality (whisker detection) may reduce suppression of auditory responses, leading to stronger tone-evoked activity during those sessions.

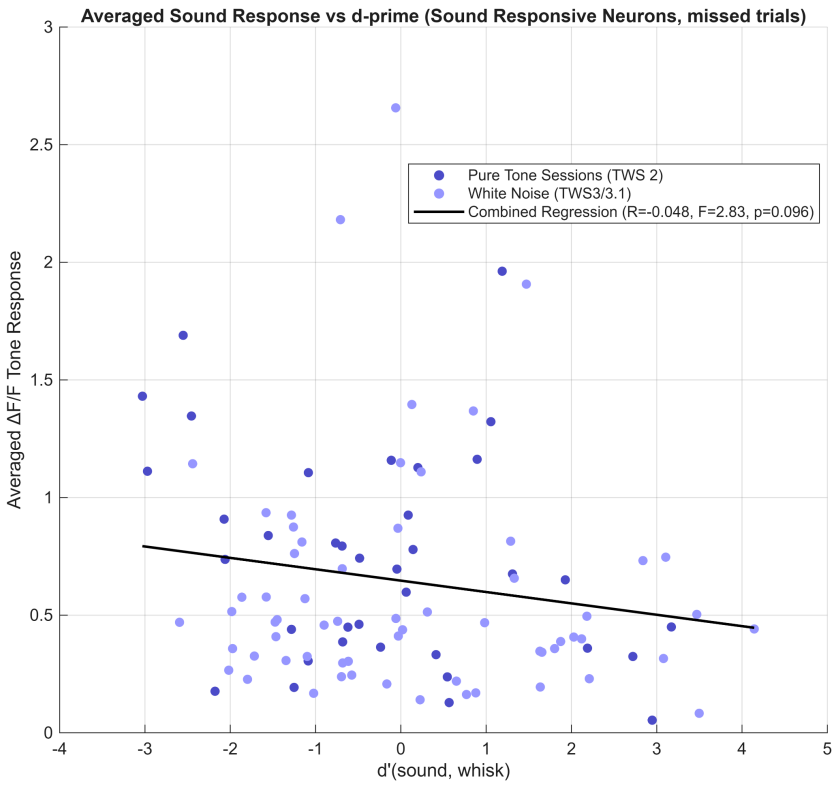

Figure 19. Relationship between sound-evoked activity and task performance. Mean sound-evoked response (ΔF/F) of all sound-responsive neurons recorded in a session versus the task performance in that session. Only trials without licking were considered (miss trials and correct rejections). The dot colour indicates whether the sound was a pure tone (dark blue) or broadband noise (light blue). The x-axis shows d′(sound, whisk), with negative values indicating that the mouse responded in a greater proportion of whisker-alone trials than sound-alone trials and vice versa.

3.2.2.2 Whisker Stimulation Evoked Activity

Neurons in the auditory cortex typically do not respond to touch. In our baseline recordings from naïve animals, only 0.3% of neurons (mean proportion across sessions, total n = 19 of 3506 neurons, range: 0% to 0.83%, SD: 0.42%) exhibited a significant response to whisker stimulation, as determined by a sign-rank comparison between pre-stimulus activity and the response window (frames 6 to 15 post-stimulus, $p < 0.01$). While such touch-sensitive neurons are rare, their responses were robust and comparable in amplitude to tone-evoked responses in sound-responsive neurons (Figure 20, A vs Figure 22, B). While we do not know whether these responses are the result of cortico-cortical connectivity or the product of subcortical crossmodal interactions and inherited from lower stages of the auditory system, such as the thalamus or inferior colliculus (Lohse et al., 2021), they demonstrate that somatosensory input can drive activity in the auditory cortex.

Given prior evidence of cross-modal plasticity in other sensory areas, such as the emergence of whisker responses in the secondary visual thalamus following reward association (Petty and Bruno, 2024), we asked whether touch sensitivity would increase if mice had learnt that whisker stimulation predicted the availability of reward. However, whisker stimulation did not, on average, evoke larger responses in mice performing the whisker task at expert level than in naive mice (Figure 22, B). In mice performing the sound detection task at expert level, touch sensitivity was reduced compared to both those performing the whisker task and naive mice. Together with the aforementioned results on sound sensitivity (Figure 18 & 19), this could suggest that engaging in a sound-detection task downregulates stimulus-evoked activity in the auditory cortex regardless of modality.

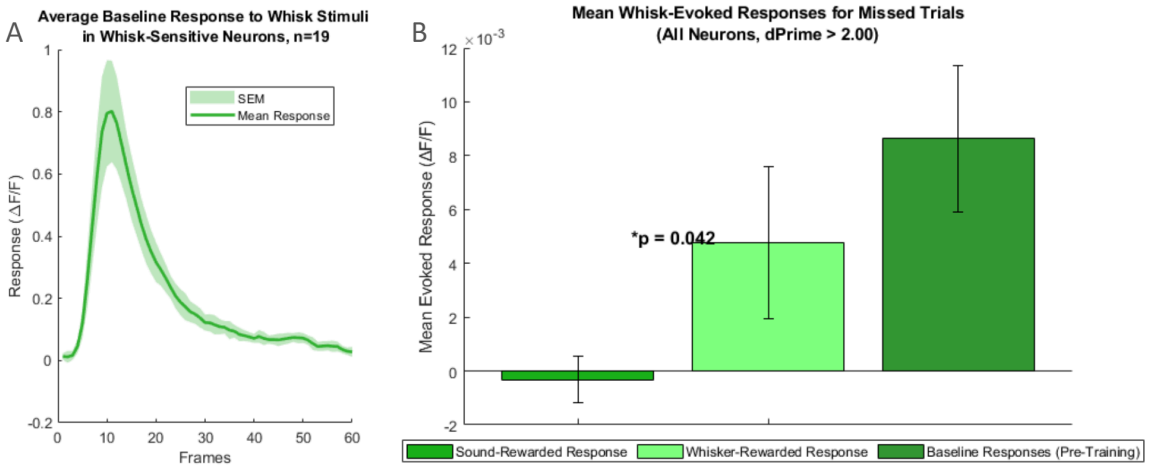

Figure 20. Touch-evoked responses in different behavioural contexts. (A) The average response profile elicited by stimulation of the whiskers for touch-sensitive neurons (n = 19) recorded in naive mice, i.e. before any training had been conducted. Shaded areas represent the standard error of the mean (SEM). **(B)** Mean touch-evoked activity of all neurons recorded in three different contexts: performing the sound task, performing the whisker task, passive stimulus exposure before any training had been conducted. Inclusion criterion for behavioural performance: d′ ≥ 2.

3.2.2.3 Suppression of Sound-Evoked Responses Correlates with Behavioural Performance in Pure Tone and Broadband Noise Mice

Finally, we asked whether the reward contingency affects the magnitude of somatosensory-auditory suppression, i.e. the reduction of a sound-evoked response that typically can be observed in auditory cortical neurons when the whiskers are stimulated at the same time compared to when the sound is presented alone (Figure 22). We observed robust suppression in naive mice in most neurons, but also facilitation in some (Figure 21). Averaged, the overall suppression was comparable in magnitude to that reported previously (Lohse et al., 2021). Assuming that suppression of sound-evoked

neural activity by somatosensory input reflects a mechanism that prioritises the processing of tactile cues from nearby objects over auditory input, we hypothesized that imparting behavioural relevance to the somatosensory stimulus and, in turn, making the auditory input irrelevant would maximize somatosensory-auditory suppression. Vice versa, we reasoned that teaching mice to ignore the whisker stimulation would diminish its capacity to suppress sound-evoked activity. Consequently, we expected somatosensory-auditory suppression to be greater in mice performing the whisker detection task than in mice performing the sound detection task. Surprisingly, the opposite was true. Considering expert performance, suppression was significantly greater when mice performed the sound detection task than when they performed the tactile detection task. In fact, on average, no suppression occurred when mice performed the tactile detection task at expert level (Figure 21). Furthermore, regressing d-prime(sound,whisk) against somatosensory-auditory suppression corroborates this result and indicates that suppression scales with task performance (Figure 23, D-F). The reward contingency, thus, affects somatosensory-auditory suppression but not in the way that we expected.

Across sessions, the absolute magnitude of somatosensory-auditory suppression increased with the size of the tone-evoked response, and this relationship just reached significance when pooling pure-tone and noise experiments (linear regression, $p = 0.048$), suggesting a gain-like modulation in which whisker input preferentially suppresses strong auditory responses (Figure 23, A-C).

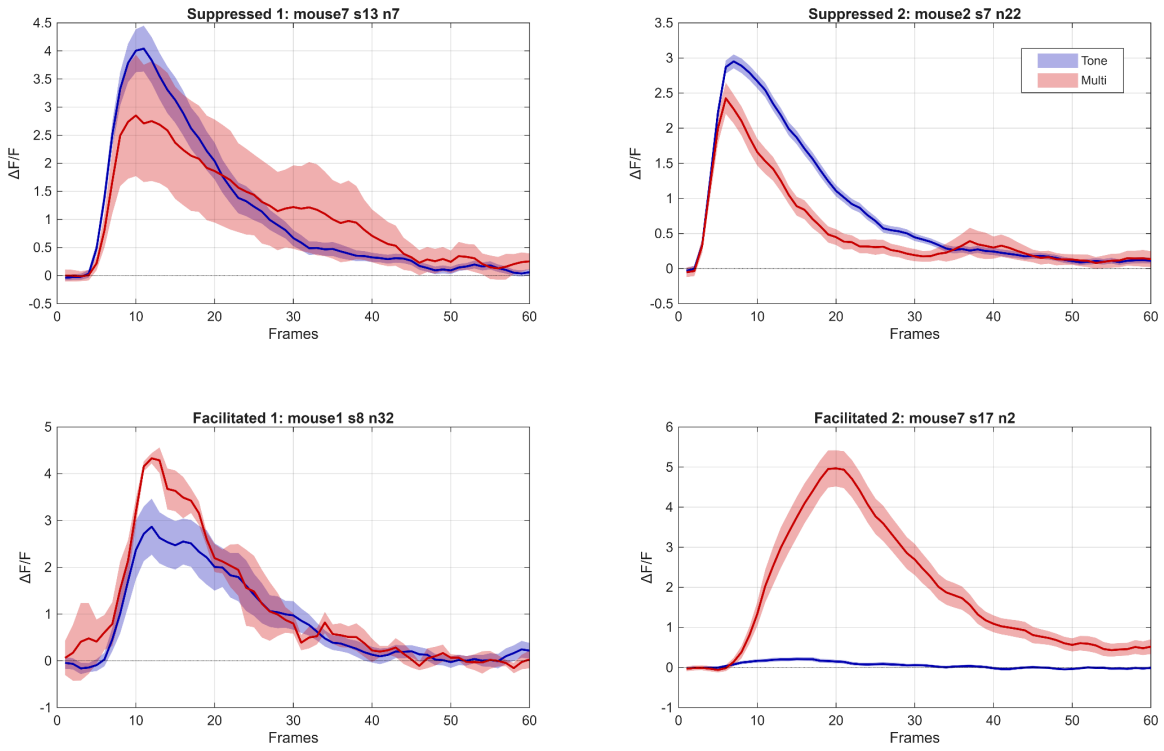

Figure 21. Single-cell examples of whisker-driven modulation. ΔF/F traces for tone-alone (blue) and tone + whisker (red) trials from two suppressed neurons (top) and two facilitated neurons (bottom).

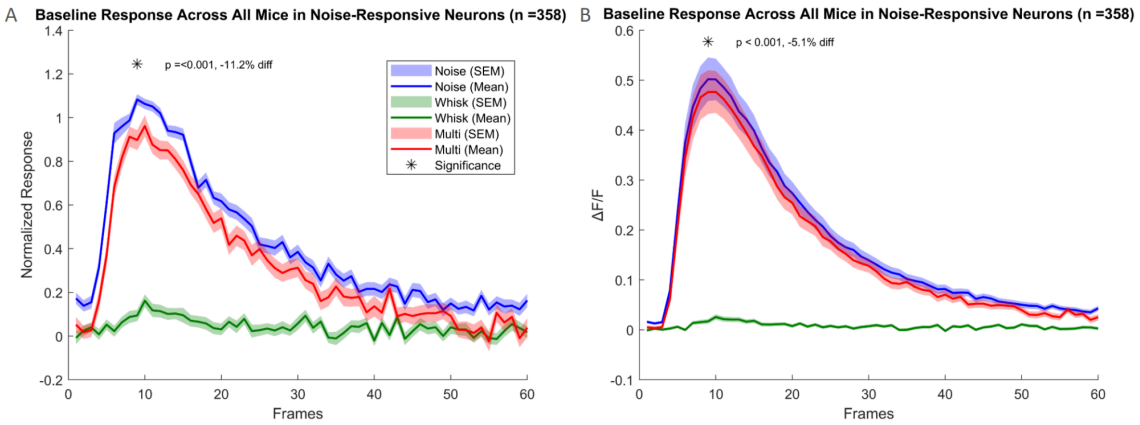

Figure 22. Baseline Responses to Noise, Whisker, and Multisensory Stimuli in Noise-Responsive Neurons Average evoked responses of 358 noise-responsive neurons to bb-noise (blue), whisker (green), and multisensory (red) stimuli during baseline sessions, prior to behavioural training and without the reward spout. The Responses in panel **A** are from neurons normalised to their own activity while panel **B** uses the unnormalised ΔF/F response. The tone response is significantly higher than the multisensory response ($p < 0.001$, -11.2% difference for normalised, $p<0.001$, -5.1% difference for the unnormalised traces). Significance tests were performed using frame 6 to 15 of the evoked responses.

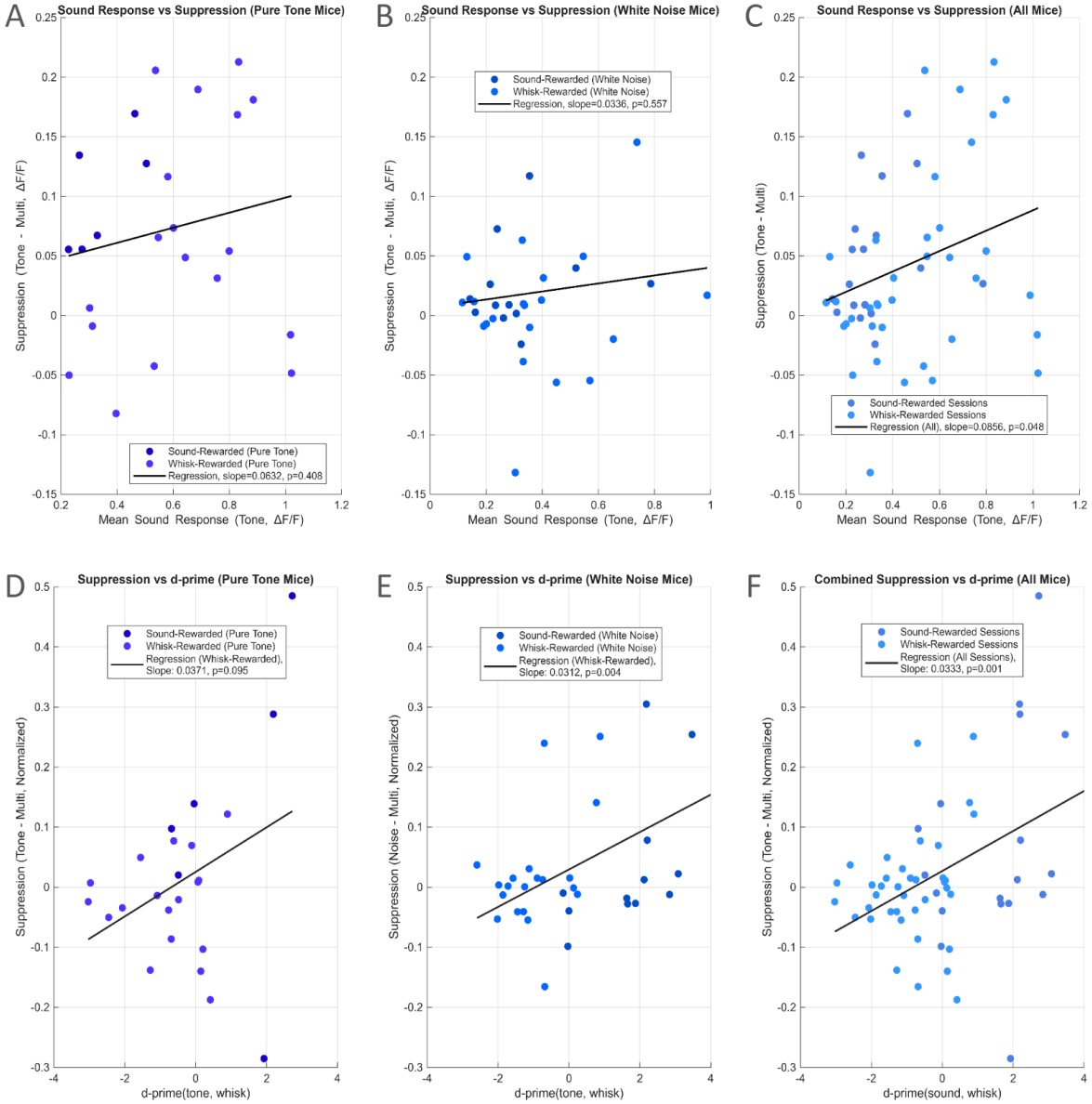
A
Sound Response vs Suppression (Pure Tone Mice)
Suppression (Tone - Multi, ΔF/F)
Mean Sound Response (Tone, ΔF/F)
Sound-Rewarded (Pure Tone)
Whisk-Rewarded (Pure Tone)
Regression, slope=0.0632, p=0.408
B
Sound Response vs Suppression (White Noise Mice)
Suppression (Tone - Multi, ΔF/F)
Mean Sound Response (Tone, ΔF/F)
Sound-Rewarded (White Noise)
Whisk-Rewarded (White Noise)
Regression, slope=0.0336, p=0.557
C
Sound Response vs Suppression (All Mice)
Suppression (Tone - Multi)
Mean Sound Response (Tone, ΔF/F)
Sound-Rewarded Sessions
Whisk-Rewarded Sessions
Regression (All), slope=0.0856, p=0.048
D
Suppression vs d-prime (Pure Tone Mice)
Suppression (Tone - Multi, Normalized)
d-prime(tone, whisk)
Sound-Rewarded (Pure Tone)
Whisk-Rewarded (Pure Tone)
Regression (Whisk-Rewarded), Slope: 0.0371, p=0.095
E
Suppression vs d-prime (White Noise Mice)
Suppression (Noise - Multi, Normalized)
d-prime(tone, whisk)
Sound-Rewarded (White Noise)
Whisk-Rewarded (White Noise)
Regression (Whisk-Rewarded), Slope: 0.0312, p=0.004
F
Combined Suppression vs d-prime (All Mice)
Suppression (Tone - Multi, Normalized)
d-prime(sound, whisk)
Sound-Rewarded Sessions
Whisk-Rewarded Sessions
Regression (All Sessions), Slope: 0.0333, p=0.001

Figure 23. Relationship between sound evoked activity, somatosensory-auditory suppression and task performance.

(A–C) Session-wise relationship between the mean sound-evoked response of sound-responsive neurons and the magnitude of somatosensory–auditory suppression. Each dot corresponds to one imaging session and shows the mean sound-alone response (x-axis; ΔF/F) plotted against the reduction in response when sound and whisker stimulation were presented together (y-axis; suppression = sound-alone − multi-trial response in ΔF/F). Panels **A** and **B** show experiments in which pure tones and broadband noise, respectively, were used as the auditory stimulus; panel **C** shows the combined data from both cohorts. Only trials without licking (misses and correct rejections) were included when computing response magnitude and suppression. Dot colour indicates whether the mouse was rewarded for responding to sound-alone trials or whisker-alone trials in that session. **(D–F)** Same sessions as in **A–C**, but suppression is plotted as a function of task performance. The y-axis shows normalised somatosensory–auditory suppression (suppression = (sound-alone − multi-trial response) / sound-alone response; negative values indicate facilitation). The x-axis shows d′(sound, whisk), where negative values indicate that the mouse responded in a greater proportion of whisker-alone than sound-alone trials and positive values indicate the opposite. Panel **D** shows pure-tone experiments, panel **E** broadband-noise experiments, and panel **F** the combined dataset. Only trials without licking were considered when calculating suppression. Dot colour again indicates reward contingency (sound- vs whisker-rewarded sessions).

In all panels, solid lines show ordinary least-squares linear regression fits across the plotted sessions; p-values report the result of an F-test for the regression (equivalent to a two-sided test of a non-zero slope).

Across panels (A-C) of Figure 23 , mice performing better in the whisker task showed less somatosensory-auditory suppression. This trend was consistent for both pure tone (TWS2) and broadband noise (TWS3/3.1) mice, indicating that task engagement and reward contingencies shape sensory integration. In the tone cohort, the relationship between suppression and d′ did not reach significance ($p = 0.095$, F-test for the regression equivalent to a two-sided test of non-zero slope) but still followed the same pattern. In the noise cohort, the correlation was significant ($p = 0.004$), confirming that whisker task proficiency (negative d-prime values) corresponded to lower auditory suppression through whisker stimulation.

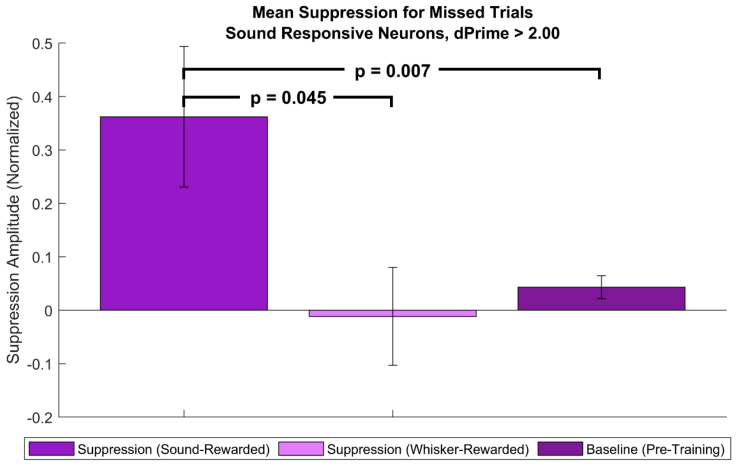

Figure 24. Somatosensory-auditory suppression in different behavioural contexts. Mean somatosensory-auditory suppression of all sound-responsive neurons recorded in two different behavioural contexts: performing the sound task and performing the whisker task. Suppression was defined as the reduction in response magnitude in sound-alone trials compared to multi-trials, normalised by the sound-alone trial response magnitude. Negative values indicate somatosensory-auditory facilitation. Only trials without licking were considered. Inclusion criterion for behavioural performance: d-prime ≥ 2.

While behavioural context modestly influenced detection latencies, the main finding was that mice generally responded faster during multisensory trials than unimodal ones, regardless of reward contingency. However, the effect was inconsistent across conditions and animals, and may reflect a combination of training history, task engagement, and individual variability (see Appendix for detailed results and statistical analyses).

3.3 Discussion

In this chapter, we investigated how reversing reward contingencies, making whisker stimulation the rewarded cue while rendering the auditory cue irrelevant (and vice versa), affects multisensory integration in the auditory cortex.

While mice learned the sound detection task quickly, learning to respond to whisker stimulation took them much longer, potentially because the whisker stimulation lacked the salience of the auditory stimulation. This is particularly surprising given that somatosensory stimulation comparable to that employed here has the capacity to modulate cortical sound processing (Lohse et al., 2021).

Contrary to the notion that rewarding a whisker stimulus would amplify its capacity for suppression, we found that whisker-induced suppression of sound responses was strongest when sound was rewarded. In the whisker-rewarded condition, somatosensory input exerted a much weaker suppressive effect on auditory activity and correspondingly, regression analysis revealed an inverse relationship between whisker-task proficiency and the magnitude of whisker-induced suppression of auditory responses. In sessions where mice excelled at detecting whisker stimulation, simultaneous auditory responses were minimally suppressed. Conversely, in sound-rewarded sessions (where whiskers were irrelevant), whisker stimulation robustly suppressed auditory responses.

Baseline recordings confirmed that excitatory responses of neurons in auditory cortex to whisker stimulation are rare. However, some A1 neurons did exhibit whisker-evoked responses, particularly under conditions of novelty or heightened attention in early stages when whisker stimulation became newly relevant. These effects subsided with further training, suggesting that attention and novelty, rather than reward association alone, drive whisker-specific responses in A1.

3.3.1 Modulation of Multisensory Integration by Behavioural Context and Reward

Contrary to our expectation, whisker input suppressed sound responses most when sound was the rewarded cue and whiskers were irrelevant, and least when whisker deflections were rewarded. Multiple factors may help explain these findings.

Sound-driven responses in the auditory cortex were strengthened when whisker stimulation itself was rewarded. Analogously, whisker-related neural activity in the somatosensory cortex could become more pronounced when the sound is rewarded. If whisker responses in S1 are indeed upregulated during sound-rewarded sessions, the subcortical loop from S1 through the lateral cortex of the IC to MGB and then on to the auditory cortex would likely experience heightened inhibitory drive. Specifically, stronger somatosensory signals arriving in the IC would further activate its inhibitory neurons, thereby inducing more robust suppression of sound-evoked responses in MGB and ultimately in the auditory cortex.

Lohse et al. (2021) identified a subcortical gate—S1 → lateral IC → MGB → auditory cortex—that mediates whisker-driven suppression of sound. Our data fit this circuit with a simple, symmetric explanation that avoids extra assumptions: When whisker stimulation is rewarded, sound-evoked firing in A1 is comparatively higher (we observe larger tone responses). By analogy, when sound is rewarded, whisker-evoked firing in S1 is expected to rise. The resulting stronger S1 output would drive the inhibitory neurons in lateral IC more forcefully, further attenuating the sound signal at the IC → MGB synapse and yielding the pronounced suppression we record in A1.

Because the gate is already tightened at this subcortical stage, additional S1 collaterals to MGB or corticofugal feedback can modulate but are unlikely to overturn the dominant effect. Direct cortical whisker inputs (Fu et al., 2003; Kayser et al., 2005) and top-down pathways may fine-tune the outcome,

yet the bulk of the task dependence can be accounted for by reward-driven shifts in S1 activity feeding into the established S1 → IC → MGB loop.

Alternatively or in addition, the observations also fit into crossmodal competition or divisive normalisation frameworks (Ohshiro et al. 2011; Yu et al. 2019), wherein signals that lack a strong incentive for integration can competitively attenuate each other. Once whisker deflections become the rewarded cue, their "competition" with the tone wanes and sound-evoked suppression relaxes. Neurons that lacked whisker-driven suppression, or even showed modest facilitation, were also correlated with a greater proportion of whisker responses, that could be explained by a distinct anatomical position or projection pattern for these units. Moreover, periods of heightened activity in the auditory cortex may saturate this competitive pool and pull attention away from the auditory task itself, further diminishing tone sensitivity.

Conversely, when whiskers are not relevant to the task, they may be processed as a competing or distracting cue, and potentially given priority by subcortical gating circuits that attenuate auditory processing. Under these conditions, whisker stimulation might overshadow the tone, yielding stronger whisker-auditory suppression. Once the whisker input acquires a positive predictive value, the gating dynamics shift: whiskers are no longer treated as "foreign" or "irrelevant," so the net suppression of sound responses relaxes. In other words, as soon as whisker stimuli become behaviourally salient, the system adapts to integrate or even facilitate that modality rather than reflexively suppressing it.

This shift in gating also demonstrates how salience is dynamically assigned based on learned associations. In the sound-rewarded context, whisker deflections, especially if novel, could function like a distractor that competes robustly with the primary (auditory) target, leading to strong suppression of tone responses. By contrast, in whisker-rewarded sessions, whisker signals are prioritized, so any

"competition" with the tone becomes reduced. Indeed, if the mouse's primary goal is to detect somatosensory input, there is no advantage in vigorously suppressing other stimuli; the crossmodal system may simply avoid amplifying the non-rewarded modality.

While novelty and task engagement likely explain the difference between baseline responses and task engagement, it does not account for the difference in whisker-evoked responses between the sound and whisker detection tasks. During task sessions, mice were equally accustomed to the experimental setup across both conditions, and our experimental design included mice that started in either task condition and later switched, ensuring that task order did not bias the results. Therefore, the observed difference — with whisker-evoked responses being significantly lower during the sound-rewarded task compared to the whisker-rewarded task — is likely driven by reward contingencies. Specifically, mice appear to have learned to ignore the whisker stimulus during the sound task, as the whisker stimulus was irrelevant for obtaining a reward in that condition. This interpretation aligns with our false alarm analysis and task learning rates, which showed that mice became better at selectively responding to task-relevant stimuli and suppressing responses to irrelevant whisk stimuli during the sound task.

In summary, while our results suggest that auditory cortical neurons do not inherently develop whisker responses through reward association, the difference in whisker-evoked activity between tasks is likely influenced by task-specific reward contingencies. The reduction of whisker responses during the sound task suggests that irrelevant stimuli are suppressed, demonstrating a flexible modulation of cortical activity based on task demands rather than a simple novelty effect.

3.3.2 Behavioural and Perceptual Implications

A notable behavioural outcome of this experiment was that mice acquired sound detection relatively quickly but struggled to master whisker detection, both initially and after the reward contingencies were switched. While learning speed was not dissimilar to what has been reported by other groups (Sachidanandam et al., 2013; Miyashita and Feldman, 2012 in rats), several studies have demonstrated that rodents can learn very subtle whisker stimulation (refs), sometimes responding to minimal whisker displacements. Consequently, our whisker stimulus, combined with head-fixation, may have been perceived as aversive and therefore may have been harder to reinforce. Additionally, the typical exploratory role of whiskers might be diminished by head fixation, making a somatosensory-based task less intuitive than an auditory-based one in this specific setup.

We also found evidence that pure tones versus broadband noise produce somewhat different magnitudes of somatosensory–auditory suppression. Because of its tonotopic organisation, broadband noise tends to recruit a broader range of frequency channels in the IC and other parts of the auditory system compared to pure tones (Barnstedt et al. 2015; Graña et al. 2017). Consequently, an expanded population of neurons may be engaged, creating a greater potential for crossmodal competition when whisker input is present. In contrast, pure tones are processed in more frequency-specific subdomains, possibly leading to relatively weaker competition or gating. Differences in stimulus complexity and predictability in multisensory integration hence should be considered (Bizley, Jones, and Town 2016), even if the precise mechanisms by which noise engages somatosensory pathways remain to be clarified.

Finally, our data support the general view that motor activity—especially licking or head movements—can mask or confound purely sensory-driven responses (Musall et al., 2019; Zagha et al., 2022). We attempted to control for this by examining only “miss” trials, nevertheless, the differences in

whisker-induced gating between tasks underscore that contextual factors (reward rules, attention, novelty) must be considered alongside motor confounds when interpreting neural data.

3.3.3 Limitations

There are several limitations inherent to the experimental design and the analysis, which should be taken into account when interpreting the findings.

3.3.3.1 Small Animal Sample Size Across Experimental conditions

Given the high number of different experimental variables, the most significant limitation of this study is the small number of mice representing each experimental condition. Our results are based on comparisons between different starting- and reward conditions (e.g., tone-rewarded versus whisker-rewarded sessions), but each subgroup contains relatively few animals. This is particularly evident in the analysis of whisker-reward starting mice, where only three mice had sufficient "high-performance" (d-prime > 2) data to be included in some of the analyses. This limitation is largely a consequence of the difficulty of training mice to perform the task switch at a satisfactory performance level. This bottleneck reflects the training burden rather than an intrinsic weakness of the paradigm. In studies that do not require a modality switch, head-fixed mice typically need only 4 sessions or less to master a tactile cue and around ≈ 9 sessions for a visual cue (Petty & Bruno, 2024). Much faster than the time our animals needed to learn their first modality, indicating that initial acquisition was slow which might be explained by the stronger stimulus amplitude and perhaps an additional auditory cue generated by the used air puff used in their experiments. What proved especially difficult in our experiment was the second learning phase, in which the rewarded modality was reversed; here many animals stalled below

criterion despite extended training, leaving us with few high-performers for the whisker-start analyses. Such attrition is not surprising: even without a reversal, passive-whisker detection in freely moving rats can take 11–21 sessions, and left–right whisker discrimination 40–70 sessions (Miyashita & Feldman, 2013). Requiring mice to unlearn an established rule and adopt the opposite one is therefore expected to lengthen training and reduce the final n. Finkel et al. trained mice to alternate between respond-to-touch / ignore-vision and respond-to-vision / ignore-touch blocks; reaching criterion (70 % correct for two consecutive days) took a median of 19 sessions (Finkel et al., 2024). Hasegawa et al. used an auditory-versus-visual Go/Nogo task in which the Go cue was reversed after initial learning. All eight mice relearned the new rule, but it still required ≈ 10 days and an initial performance collapse before recovery (Hasegawa et al., 2024).

The limited sample size restricts the generalisability of our findings. While we observed significant differences in suppression effects across conditions, these results may be driven to some degree by individual variability rather than representing a robust, generalisable trend. Future studies should aim to increase the number of animals per condition to ensure that observed effects are consistent across a larger population.

3.3.3.2 Diversity of Starting Conditions

As an extension of the first limitation above, the study's focus on training mice under different starting conditions, with some animals initially learning to detect sound stimuli and others trained to detect whisker stimuli leads to a further issue: This experimental setup was intended to remove any bias toward effects of starting conditions by ensuring that any effect could be canceled out by the inclusion of mice starting on the opposite condition. Additionally, having individual mice undergo task switches provided opportunities for within-subject analyses. However, the benefits of this design did not fully materialize, as meaningful statistical analyses often required pooling mice across conditions to achieve sufficient sample sizes.

3.3.3.3 Task Engagement and Motor-Related Activity

The neural responses recorded during task performance reflect not only sensory processing but also motor-related activity associated with licking and reward collection as well as potentially uninstructed movements (Zagha et al., 2022). In particular, the delayed peaks observed in whisker-rewarded sessions demonstrated that some of the recorded activity is likely related to reward anticipation, motor planning, motor activity and/or reward reception rather than purely somatosensory and auditory responses. While we attempted to minimise the influence of licking-related activity by excluding trials with licking during the stimulus window, it is difficult to completely disentangle sensory responses from motor-related activity in behavioural experiments. Future studies should benefit from additional controls, such as using non-rewarded trials or varying the timing of reward delivery to better isolate sensory-evoked responses.

3.3.3.4 Definition and Interpretation of Miss Trials

Using only "miss trials" for the analysis of suppression effects and evoked responses removed most confounds from motor and reward-related activity but may introduce its own problem: In sound-rewarded sessions, a missed sound trial might indicate that the mouse did not detect the sound stimulus or was disengaged. In whisker-rewarded sessions, however, a missed sound trial could also represent a case where the mouse detected the sound but chose not to respond, as the task rule required a response only to whisker stimuli. This distinction is crucial because it implies that a missed trial may not always reflect a failure of perception but rather a correct behavioural choice based on task rules.

Similarly, whisker-only trials that were not responded to in the sound-rewarded task may not reflect true misses. In this case, the mice might have perceived the whisker stimulus but chose to withhold a response because the task required them to respond only to sound stimuli. In contrast, ignored whisker

responses in the whisker-rewarded task are more likely to represent genuine missed detections, as the whisker stimulus was task-relevant and expected to elicit a response.

These discrepancies in the definition of missed trials may introduce biases in the analysis: Higher sound-evoked responses observed during missed trials in whisker-rewarded sessions could reflect neurons responding to the sound stimulus even without a behavioural response, potentially inflating sound responses in these trials. Similarly, whisker-only trials that were not responded to in sound-rewarded sessions may represent cases where the whisker stimulus was perceived but the mouse correctly withheld a response, leading to an overestimation of missed whisker trials in this context. In contrast, missed trials during their respective rewarded tasks are more likely to reflect genuine misses, where the stimulus was not consciously detected by the mouse.

These differences suggest, then, that evoked responses are likely to vary depending on whether a stimulus was consciously perceived or not. We would expect stronger neural responses when a stimulus is consciously detected, as this would involve more robust engagement of sensory pathways. Conversely, unconscious or unattended stimuli may elicit weaker responses due to reduced attention or task engagement.

It is important to note that the sound level used in our experiments was well above the threshold for detection and clearly audible to the mice. Therefore, the vigilance of the mouse during the onset of a trial may be a critical determinant of whether a response occurs. Fluctuations in attention and engagement during task performance could significantly influence the magnitude of evoked responses. These variations in vigilance likely play a key role in shaping the observed differences in neural responses across missed and detected trials, highlighting the need for careful interpretation of suppression effects within different task contexts.

3.3.3.5 Interpretation of Reaction Latency Differences

The reaction latency differences observed in our study provide some insight into how task context influences detection speed. However, these results must be interpreted with caution due to the small number of mice representing each condition. The observed trends may also be influenced by individual variability, task order effects, or other uncontrolled factors.

Additionally, the variability in reaction times across different conditions suggests that task engagement and sensory prioritization are influenced by multiple factors, including training history, reward contingencies, and stimulus characteristics. While our findings suggest that early training shapes sensory prioritization, further research is needed to fully understand the relationship between task context and reaction time.

3.3.3.6 Limited Generalisability Across Species

The cortico-colliculo-thalamo-cortical loop represents a novel circuit through which one sensory modality influences another, and hence understanding how it adapts to behavioural demands is particularly interesting. While cortico-thalamo-cortical loops provide an important route for communication between different cortical areas in multiple species (Sherman & Usrey, 2024), the generalisability of these findings may be limited by the unique characteristics of the rodent whisker system.

The whisker system, critical for rodent-specific behaviours like burrowing and social interaction, is highly specialised and has less or no relevance in other species. Mice, as small prey animals, display distinct "flight" behaviours in response to nearby stimuli, which may differ from larger animals like rats. These unique traits complicate extrapolation of our findings to multisensory systems more generally.

Future research should consider species-specific differences to better understand general principles of multisensory integration.

3.4 Methods

3.4.1 Animals and Experimental Setup

All experiments were approved by the Committee on Animal Care and Ethical Review at the University of Oxford and were licensed by the UK Home Office (Animal Scientific Procedures Act, 1986, amended in 2012). We used double-transgenic mice expressing the genetically encoded calcium indicator (GECI) GCaMP6f (T.-W. Chen et al., 2013) in CaMKIIα+ neurons (GTMK, see table) as well as triple-transgenic mice (GTLA) in which the Cdh23ahl allele that otherwise predisposes the C57BL/6 strain to age-related high frequency hearing loss has been corrected (Mianné et al., 2016).

A cohort consisting of 5 GTMK mice (3 females, aged 10–12 weeks) was used to pilot the task design. The main experiments involved 7 mice (GTMK: 2 females, 2 males; GTLA: 2 females, 1 male) that were trained with pure tones and 10 mice (GTMK ; GTLA) that were trained with noise bursts (Table 2). The mice were between 9 and 14 weeks old at the beginning of the study.

Name	Strain Description	Source (Identifier)
Ai95D	B6;129S-Gt(ROSA)26Sor$^{tm95.1(CAG-GCaMP6f)Hze}$/J GCaMP6f (T.-W. Chen et al., 2013) reporter line, Cre-dependent	The Jackson Laboratory (JAX:024105)
T29-1	B6.Cg-Tg(Camk2a-cre)T29-1Stl/J CAMK2A-Cre driver line	The Jackson Laboratory (JAX:005359)
B6ΔCdh23	C57BL/6NTac.Cdh23$^{753A>G}$ Corrected for age-related hearing loss (Mianné et al., 2016)	MRC Harwell
GTMK	Ai95D × T29-1	Internal colony
GTLA	Ai95D x T29-1 X B6 ΔCdh23	Internal colony

Table 2. Details the genetic strains of mice used in this study. The table includes double-transgenic GTMK mice, which express the GCaMP6f calcium indicator in CaMKIIα+ neurons, and triple-transgenic GTLA mice, which additionally include a genetic correction for age-related hearing loss associated with the C57BL/6 background. These strains were derived from internal breeding colonies using the parent strains Ai95D and T29-1, obtained from The Jackson Laboratory, and the B6ΔCdh23 strain from MRC Harwell. The table outlines the specific genetic modifications, source identifiers, and breeding strategies employed to generate the experimental mouse cohorts.

3.4.2 Localisation of auditory Cortical Fields through Wide-Field Imaging

A functional map of the auditory cortex was obtained for each mouse using wide-field calcium imaging (Figure 25) in a sound- and light-shielded chamber. Images were captured using a TL2x-SAP objective (Thorlabs) and a 340M-GE CCD camera (Thorlabs) controlled by Thorcam 3.3 software. Illumination was provided by a 470 nm LED (Thorlabs). The recording resolution was reduced to 128 × 96 pixels, and images were acquired at a rate of 10 Hz using a NI-DAQ I/O device (USB-6501, National Instruments).

Sinusoidally amplitude-modulated (SAM) tones with a 500 ms duration, ramped over 5ms and with 10 Hz modulation frequency were used for stimulation. These tones were produced at two sound pressure levels and four (two low- and two high-frequency sounds - typically 4 kHz / 5.7 kHz & 22.6 kHz / 32 kHz). These were generated by a custom LabView script, amplified (Avisoft Bioacustics) and delivered via a calibrated speaker (Avisoft Vifa). During imaging, mice were placed under the microscope objective within a tube, leaving only their heads exposed. SAM tones were presented sequentially, each separated by 3 seconds, and with three-second intervals between them and repeated at least ten times to ensure a sufficient signal-to-noise ratio.

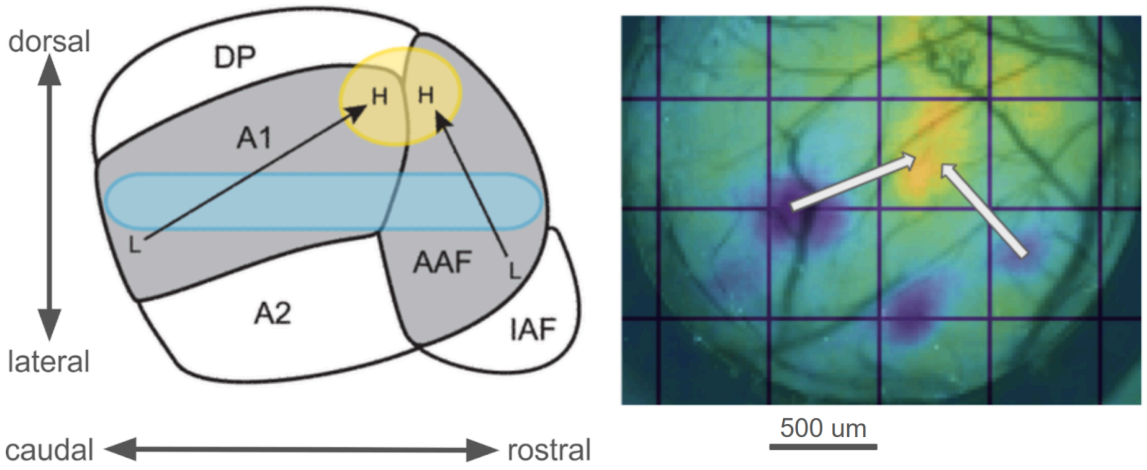

Figure 25. (Left): Schematic (Guo et al., 2012) illustrating the locations of auditory cortical fields (A1, AAF, A2, DP, and IAF) and the corresponding low (L) to high (H) frequency gradients. (Right): Widefield calcium imaging responses to high (yellow) and low-frequency (blue) SAM tones recorded in one example mouse. Arrows indicate low-to-high frequency gradients of A1 and AAF, respectively.

The collected data were analysed online so that additional data could be recorded immediately after acquisition, allowing additional repeats if needed. To remove slow drifts, a baseline F0 was computed for each pixel by low-pass filtering the fluorescence time series (using a LOWESS filter over 10 seconds). The signal was then normalised as $\Delta F/F = 100\times(F-F0)/F0$, and each frame was spatially smoothed with a small 2D filter (3×3 pixels). For quantifying tone-evoked responses, baseline mean and standard deviation were calculated over a 600 ms pre-stimulus window. The response in subsequent 200 ms bins was then expressed as a Z-score relative to this pre-stimulus baseline, thus providing a quick and reliable measure of tonotopic activation. This fast analysis pipeline enabled functional mapping of the cortical frequency gradient within minutes. Typically, 15 repeats per stimulus were sufficient. Slow drifts in the data were removed, and fluorescence signals were normalised. Responses were analysed by

computing the baseline and standard deviations prior to stimulus onset. Post-session, the mean response was calculated in time bins and Z-scored relative to the pre-stimulus baseline.

3.4.3 Two-Photon Calcium Imaging

Two-photon calcium imaging was employed to record the activity of groups of individual neurons in the auditory cortex. Images were acquired using a commercial two-photon laser-scanning microscope (B-Scope; Thorlabs, NJ, USA) equipped with a 16×/0.80W water immersion objective (CFI75 LWD 16X W, Nikon). The objective was immersed in diluted ultrasound gel (Anagel; Ana Wiz, Surrey, UK) for all recordings.

Excitation was provided by a pulse-modulated laser (Mai-Tai eHP, Spectra-Physics) at a wavelength of 930 nm, with a 70 fs pulse width and an 80 MHz repetition rate. The laser beam was deflected onto the brain via a galvo- (Y-axis) resonant (X-axis, bidirectional) scanner, facilitating image acquisition at approximately 30 frames per second over a 512 × 512 pixel field of view. Emitted photons from the fluorescent neurons were filtered through a 525 nm bandpass filter (50 nm window) and detected using GaAsP photomultiplier tubes (Hamamatsu).

Image acquisition was managed using the ScanImage toolbox for MATLAB [42, version 4]. To ensure synchronisation between imaging frames and behavioural data, a digital I/O device (USB-6501, National Instruments) was employed, as imaging and behavioural data were recorded on separate computers.

3.4.4 Session Validation and Series Generation

This experiment produced over 12 terabytes of raw calcium imaging data. To manage and analyse the generated data efficiently, a semi-automated pipeline was initially developed by Oliver Bredemeyer for the first experimental cohort and later revised by the author for subsequent experiments using MATLAB (R2020b, later R2022b; MathWorks, MA, USA). The pipeline was designed to streamline data validation, aggregation of sessions into series, and subsequent data processing.

Session and Series Definitions: A "session" refers to a single imaging block. Multiple sessions performed consecutively on the same day with consistent imaging parameters (e.g., unchanged microscope positioning and laser power) were combined into a "series". This aggregation of sessions into a series enhances statistical power and improves the accuracy of data analysis by maintaining consistency in imaging conditions.

3.4.4.1 Session Validation

Initially, the imaging files underwent an automated validation process to ensure structural integrity and detect potential corruption. Semantic validity was assessed by computing image statistics to flag any files with unusual characteristics for further review. Detailed reports were generated for flagged files, which were then manually inspected. Any recordings deemed unsatisfactory upon inspection were excluded from the dataset. An automated cross-validation was performed against the experimental records to identify inconsistencies, such as duplicate sessions, missing entries, subject mismatches, or timestamp discrepancies. Where discrepancies could be resolved, manual corrections were made; otherwise, the affected data were excluded from further analysis.

3.4.4.2 Series Generation

Sessions collected from the same mouse on a single day were combined into a series, provided they maintained consistent imaging parameters. This combination was contingent on meeting specific criteria: identical subject identity, the same session date, and recording coordinates within a tolerance of 10 μm in the xy-plane and 3 μm along the z-axis. Additionally, the excitation beam power and other imaging conditions had to remain unchanged to ensure data consistency. Following validation and series generation, the data were processed in batches using the suite2p software package (Pachitariu et al., 2017), incorporating rigid and non-rigid drift correction, automatic cell detection, and fluorescence signal correction to remove neuropil contamination (see below).

3.4.5 Raw Two-Photon Imaging Processing Pipeline

The raw imaging data (".tif" format) were processed using the suite2p toolbox (Pachitariu et al., 2017), which performed drift correction, region of interest (ROI) detection, and signal refinement to isolate true neuronal activity. Processing was executed in parallel across multiple computers, with verification procedures established by O.B.

Initial drift correction employed both rigid and non-rigid image registration to mitigate motion artifacts and ensure spatial alignment over time. Subsequently, ROIs were detected based on pixel intensity correlation maps. Although suite2p's clustering methods can identify various structures (e.g., cells, axons), only somata of layer II/III cortical neurons were extracted. To exclude irrelevant ROIs, a custom classifier was trained using manually annotated datasets, consisting of 2,922 ROIs from three sessions per mouse and validated with 1,210 ROIs. Performance was assessed against the default suite2p

classifier using ROC curves, with statistical significance tested via DeLong's test ($p < 0.05$). The superior classifier, chosen based on ROC-AUC, was used to filter ROIs by maximizing the Youden index (Youden, 1950), ensuring high positive predictive value.

Following ROI detection, neuropil correction was applied to minimise background fluorescence contamination. A scaled version of the neuropil signal (factor 0.7), calculated as a ring around each cell, was subtracted from the raw fluorescence to isolate neuronal activity. Lastly, $\Delta F/F$ was calculated for each neuron to quantify changes in fluorescence, normalising signals across cells and experimental conditions to facilitate comparison within and across sessions.

3.4.6 Stimuli used for 2-photon imaging

3.4.6.1 Auditory Stimuli

Auditory stimuli consisted of either 12 kHz pure tones (experiment 1) or broadband noise bursts (1–64 kHz) with 10 ms raised cosine onset and offset ramps. Both were presented for 50 ms at 60 dB SPL. Sounds were generated at at a 192 kHz sampling rate using the Psychophysics Toolbox extension in MATLAB (release R2015b; MathWorks, MA, USA) and a 16-bit sound card (Startech). The output was amplified (Portable Ultrasonic Power Amplifier; Avisoft Bioacoustics) and delivered via an ultrasonic speaker (Avisoft Vifa) positioned 23 cm in front of the mouse, slightly offset to the left. All sound stimuli were calibrated using a precision microphone (G.R.A.S. 40BE ¼") to ensure accurate sound pressure levels at the mouse's head during experiments.

3.4.6.2 Somatosensory (Whisker) Stimuli

Whisker deflections were delivered to the left whiskers using a custom-designed piezoelectric stimulator. For the pilot cohort, a stimulator based on Lohse et al. (2021) was used, consisting of a small brush with plastic hairs attached to a piezoelectric bimorph (PL122.10 PICMA Bender; Physik Instrumente). This device imparted a 1 mm vertical sinusoidal displacement at 20 Hz, corresponding to a peak whisker velocity of ~0.13 m/s. In its initial form, the stimulator delivered a single 50 ms cycle, but due to poor behavioral performance the stimulus was extended to three cycles (150 ms) at the same amplitude and frequency. Contact occurred as the brush pressed upward against the whisker pad, deflecting multiple whiskers in a smooth oscillatory manner. However, because the mice were often able to rapidly pull their whiskers free of the brush, the brush was replaced with a fine steel mesh grid to improve engagement. The mesh, mounted on the same piezoelectric bimorph, contained approximately square holes with a side length of ~0.7 mm in which individual whiskers became transiently trapped during deflection. This produced a more secure, though slightly more irregular, contact along the whisker shafts while still delivering the same displacement and velocity profile. The grid was positioned ~2 mm lateral to the left side of the face with whiskers fed through the openings. Waveforms were generated using the Psychophysics Toolbox in MATLAB and delivered to the piezoelectric bimorph via an E-650 piezo amplifier (Physik Instrumente). The specific stimulator type used depended on the experiment (details below).

3.4.6.3 Behavioural Monitoring and Control

Licking behaviour was monitored using a custom lickometer, comprising a conductive foil sheet and a metallic lick spout. This setup detected tongue contacts with high temporal precision (~1 ms), allowing for the precise correlation of licking behaviour with stimulus presentation. The timing of

stimulus delivery, behavioural monitoring, and reward dispensing was controlled using a MATLAB-based system with Psychophysics Toolbox extensions and interfaced with a data acquisition device (USB-6008; National Instruments, TX, USA).

3.4.7 Behavioural Training

3.4.7.1 Training Protocol

Mice were trained to associate either a somatosensory or an auditory stimulus with a water reward, aiming to induce a lick response. To ensure water acted as an effective motivator, the mice were placed on a fluid-restricted diet, reducing their body weight to 80–85% of baseline. Training was conducted under head fixation under the two-photon microscope setup, enabling simultaneous imaging of neuronal activity. In some mice visibility became too poor over time to continue imaging until the end of the experiment.

During each session, mice were presented with trials containing either a somatosensory stimulus (whisker-alone) or auditory stimulus (sound-alone) presented on its own or a multi-trial, in which both stimuli were at the same time. Additionally there were catch trials in which no stimulus was presented. Whisker-alone and sound-alone trials were presented with equal probability, multi-trials occurred with a variable probability depending on performance. Initially, multi-trial probability was 0 and increased up to 0.3 as detection performance improved. Catch trials occurred with a probability of 10%. Inter-stimulus intervals were drawn randomly from a shifted half-normal distribution with a standard deviation of 3 seconds and a minimum of 8 seconds to discourage anticipatory licking.

Licking the spout within the response window (this was typically set to 0-1s from stimulus onset but occasionally shortened to discourage excessive licking) of a sound-alone trial or whisker-alone trial triggered, depending on the task, the release of a drop of water. In contrast, licking during multi-trials was never rewarded to prevent forming an association between the off-target stimulus and the reward. To facilitate learning and maintain task engagement, in some sessions (mostly at the beginning of training or after a switch in reward contingency) rewards were delivered together with the target stimulus. These sessions were later excluded from the analysis. Training sessions were limited to a maximum of 2 hours or terminated sooner if the mouse stopped licking or exhibited signs of stress. Data collection including habituation took over 1300 hours.

3.4.7.2 Automated Record-Keeping and Performance Tracking

An automated system was developed by O.B. to track each behavioural training session, recording experimental parameters and behavioural data while assigning a unique serial number to each session. This system ensured precise identification of corresponding imaging and behavioural data, reducing human error and enabling the detection and resolution of discrepancies through timestamp and parameter comparisons. Consequently, none of the over 800 sessions conducted were rejected due to record-keeping ambiguity.

The automated system also allowed for performance-based progression by individually tracking each mouse's learning rate and advancing them through the experimental paradigms accordingly. An algorithm for progression was pre-defined and integrated into the system, which used the mouse's prior performance to recommend training parameters for the next day.

3.4.7.3 Real-Time Performance Monitoring

Performance was monitored in real time using the sensitivity index (d-prime), which was continuously calculated and updated during training sessions. This live feedback enabled the experimenter to detect and address sudden changes in task performance, such as reduced task engagement or technical issues. Additionally, the system provided on-demand graphical displays of performance metrics across the entire experiment, alongside relevant experimental parameters—information previously only available after completing the entire experiment.

IV Auditory Cortex Distinguishes Between Sound-Evoked and Spontaneous Movements

4.1 Introduction

Our perception of the world depends not only on how we process external stimuli but also on how we account for our own movements. Self-generated actions—such as locomotion in rodents—are known to modulate, and in some cases suppress, neural responses to external stimuli, as a sophisticated form of sensorimotor integration. In Chapters I and II, we examined the behavioural impact of a known crossmodal suppression circuit—one in which whisker stimulation suppresses sound-evoked cortical activity. Our findings underscored that this suppression is not merely a static gating mechanism but can be flexibly modulated by task requirements, suggesting that contextual factors (such as the relevance of tactile versus auditory cues) play a pivotal role in shaping how the circuit operates (Chapter III). However, these earlier chapters relied on externally applied (passive) whisker stimulation delivered concurrently with sounds. Under natural conditions, mice predominantly depend on active whisking—rather than purely passive deflections—to gather tactile information (Sofroniew & Svoboda, 2015). Consequently, the extent to which spontaneous whisker movements alter sound processing might differ substantially from the effects of passively imposed whisker stimulation.

Accumulating evidence suggests that whisking is more than merely tactile exploration; it also serves as a clear indicator of global brain state. Whisker movements frequently coincide with pupil dilation (McGinley et al., 2015), neuromodulatory activity in the locus coeruleus and basal forebrain, and cortical network "desynchronisation," reflecting shifts in arousal or attentional state (Collins et al., 2023; Eggermann et al., 2014; Musall et al., 2019; Reimer et al., 2016) see (McCormick et al., 2020) for a

review.). These whisking bouts often align with locomotion, broader internal state transitions and elevated cholinergic and noradrenergic signalling, which together suppress slow oscillations and enhance cortical responsiveness (Collins et al., 2023; Musall et al., 2019; Reimer et al., 2016). A recent large-scale imaging study has further shown that much of this cholinergic and noradrenergic axonal activity is driven by a unified global movement-related mechanism (Collins et al., 2023). Indeed, movements such as locomotion can substantially influence neuronal activity throughout sensory cortices, indicating that internal states and motor actions are modulators of sensory processing, often dominating neural responses more than external stimuli alone (Musall et al., 2019; Stringer et al., 2019). In the auditory system specifically, studies have demonstrated that locomotion robustly modulates subcortical activity (in the IC) (C. Chen & Song, 2019; Yang et al., 2020). Furthermore, whisker movements not only reflect shifts in arousal or exploratory behaviour but also exert significant influence over auditory processing. In addition to its role as an index of arousal, whisking, like other self-generated actions, can be accompanied by predictable reafferent signals (Eliades & Wang, 2003; Schneider et al., 2014; Singla et al., 2017). The central nervous system may use a corollary discharge to anticipate and modulate these signals, so as not to confound internally generated input with externally driven stimuli. In the auditory cortex, corollary discharge circuits help distinguish between self-made and external sounds, preventing self-generated noises from overwhelming relevant inputs. However, the influence of whisking on sound processing may be multifaceted: While local sensorimotor gating mechanisms could shape neural responses, the broader neuromodulatory changes tied to whisking (Collins et al., 2023; Flavell et al., 2022) can also alter excitability across wide regions of cortex.

Finally, auditory stimuli themselves can evoke motor responses including whisker movements. Loud or novel sounds, can trigger a startle or orienting response that includes whisking (Bimbard et al., 2023; Clayton et al., 2024; Schneider et al., 2018). In such cases, whisker movement may reflect both

corollary discharge (if the animal moves in response to its own intended action) and exafferent signals driven by the external cue. If auditory cortical neurons respond differently to whisker movements that are spontaneously initiated versus those triggered by a sound, this would support the context-dependent nature of sensorimotor integration (Collins et al., 2023; Sara & Bouret, 2012).

In this chapter, we explore how active whisker movements interact with concurrent auditory events in freely behaving but head-fixed mice. We simultaneously recorded whisker movements, pupil dilation, and locomotion, alongside two-photon calcium imaging of auditory cortical neurons Whereas previous work (Collins et al., 2023) demonstrated widespread coupling between movement and neuromodulatory activity, our study specifically examines neurons in the auditory cortex. By distinguishing spontaneous whisking from sound-evoked whisking, we reveal striking context dependence: similar whisker movements produce distinct neural responses depending on whether they are spontaneous or externally triggered. We address: (1) Do whisker movements consistently covary with other arousal-related behavioural signals like pupil dilation and locomotion or can they occur independently? (2) Can pure tones reliably elicit whisker movements, and do these differ from spontaneous movements in amplitude or timing? (3) Do auditory cortical neurons differentially encode spontaneous versus sound-triggered whisker movements, thus distinguishing self-generated exploratory actions from externally cued responses.

4.2 Results

In the preceding chapters, we investigated how externally applied (passive) whisker stimulation can modulate sound-evoked responses in the auditory pathway. However, mice ordinarily engage in active whisking to gather tactile stimuli, and self-generated whisker movements may interact differently with auditory processing than externally imposed whisker deflections. Consequently, we set out to examine how active whisking relates to other behavioural and physiological measures—such as pupil dilation and locomotion—and how these relationships might ultimately impact neuronal activity in the auditory cortex.

To enable precise measurements of whisking, we placed head-fixed mice on a freely rotating polystyrene wheel, allowing them to walk, run, or remain still with minimal constraint on body movements. At the same time, we recorded high-speed videos of one eye and the whiskers (Figure 26, A–C). By carrying out these experiments in a moderately lit rather than dark chamber, we avoided permanent pupil dilation and thus ceiling effects in the pupil diameter measurement.

We used DeepLabCut (Mathis et al., 2018) to track both the pupil rim and one to three individual whisker tips in each session. Specifically, the pupil diameter was calculated from four tracked points along its edge, and whisker movements were defined by changes in tip position from frame to frame (Figure 26, C). Averaging across all tracked whiskers yielded a single continuous measure of whisking activity. Throughout the recording sessions, we also monitored the rotational speed of the wheel to capture locomotion. In this way, we obtained a comprehensive view of how whisking, pupil dilation, and locomotor behaviour fluctuate over time—measurements that would allow us to dissect the interplay between whisker movements and brain state in subsequent analyses.

Having established methods for tracking whisker-tip positions to quantify whisking and simultaneously measuring other behavioural variables (locomotion and pupil diameter) we then went on to characterize how whisking co-varies with these variables.

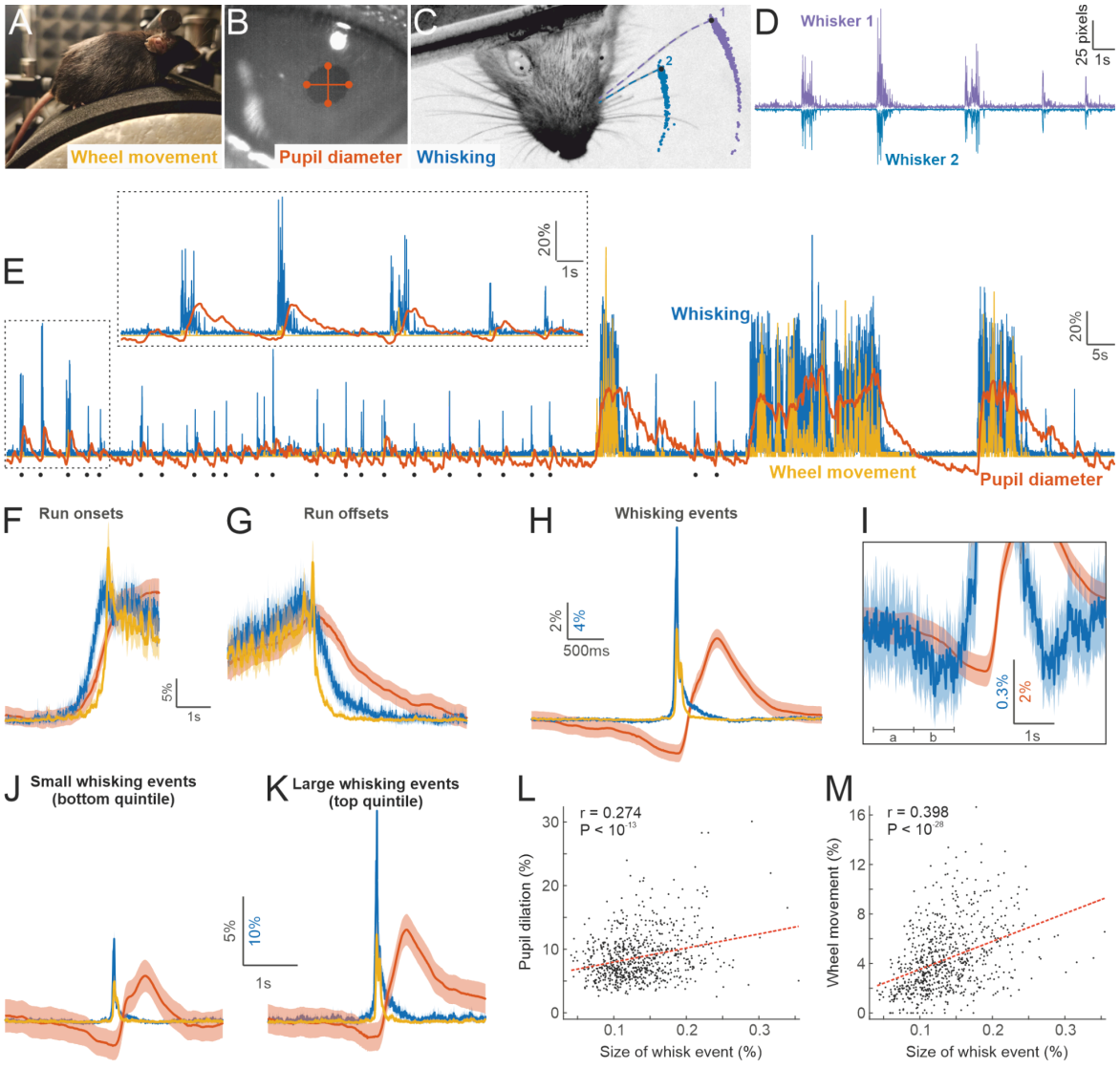

Figure 26. Relationship between whisking and behavioural state measures. (A) Head-fixed mice on a polystyrene wheel enabled measurement of locomotion and subtle body movements. **(B-C)** Pupil diameter and individual whisker movements were tracked via video recordings. In **(C)**, dots mark whisker

tip positions of two tracked whiskers during a 17-second example video. **(D)** Frame-by-frame displacement of these whisker tips from the same video segment. **(E)** Example traces of whisking, wheel movement, and pupil diameter measured across 200 seconds, normalised to session maxima. Insets highlight a zoomed segment corresponding to **(D)**. Black dots indicate whisking events outside of locomotion periods and more than 10 seconds after locomotion. **(F-G)** Mean traces for whisking, pupil diameter, and wheel movement aligned to run onsets (n = 111) and run offsets (n = 110). **(H)** Mean traces for the same variables during 743 whisking events outside locomotion, with **(I)** showing expanded axes of whisking and pupil responses. Preceding whisking events, a significant drop in median whisking activity is observed between consecutive 500 ms time segments 'a' and 'b' ($p < 10^{-12}$, Wilcoxon signed-rank test). **(J-K)** Whisking, pupil, and wheel traces for the smallest (bottom quintile, n = 148) and largest (top quintile, n = 148) whisking events. Shaded areas in **(F-K)** indicate 95% confidence intervals. **(L-M)** Correlations between whisking event size and pupil dilation ($r = 0.274$, $p < 10^{-10}$) or wheel movement ($r = 0.398$, $p < 10^{-38}$), respectively, show larger whisking events associate with increased pupil dilation and locomotion.

4.2.1.1 Whisking Tightly Correlates with Locomotion

Mice tended to whisk vigorously whenever they initiated locomotion. This relationship can be seen in the example recording of whisking, wheel movement, and pupil diameter traces (Figure 26, E) and was further quantified by aligning the other variable to the locomotion onsets and offsets (Figure 26, F&G). Interestingly, whisking usually preceded the onset of locomotion by a few hundred milliseconds (Figure 1, F) and persisted briefly—typically for a few hundred milliseconds— after locomotion had ceased (Figure 26, G).

4.2.1.2 Whisking Also Occurs Outside Locomotor Episodes

Mice did not only whisk when they ran; they also produced whisker movements during stationary periods. These "whisking events," indicated by dots in Figure 26 E, typically lasted a few hundred milliseconds and were not followed by sustained locomotion. Although these events were brief compared to whisking during running, they still showed robust pupil dilation (albeit smaller than during locomotion) and corresponded to a small, short-lived signal in the wheel's motion sensor. This slight wheel movement ("twitch") suggests that whisking bouts are often accompanied by a whole-body motor pattern, even though the mouse remains essentially in one place.

4.2.1.3 Coupled Changes in Pupil Diameter

Consistent with prior work (Reimer et al., 2014), the pupil dilated in tandem with whisking—and especially with locomotion (Figures 26, E–G). At run onset, the pupil reached its maximum diameter after whisking reached its maximum amplitude (Figure 1F). Once the mouse stopped, the pupil took longer to contract, with its final return to a smaller diameter trailing the offset of whisking (Figure 26, G). During the brief, spontaneous whisking events that occurred outside of locomotion, pupil diameter again rose, though less dramatically, reaching its peak ~0.5 s after the whisker and wheel motion (Figure 26, H).

Interestingly, a pre-twitch pupil constriction could be seen when plotting the pupil diameter aligned to the whisking events. On a more expanded scale (Figure 26, I) it became evident that the whisker movement trace similarly dipped very slightly prior to the whisking event, yielding a statistically significant but small decrease in whisking immediately preceding the whisking event. These observations hint at a short, transient shift in autonomic state just before the mouse twitches.

4.2.1.4 Scaling and Common Mechanisms

Comparisons of small vs. large whisking bouts revealed that the amplitude of whisker movements correlates positively with both wheel movement and pupil dilation (Figure 26, J–M). Stronger whisking events tended to co-occur with larger body twitches and more pronounced dilations of the pupil. This scaling effect supports the idea that all three signals—whisking, pupil diameter, and subtle whole-body movements—represent different readouts of a common underlying state change in the brain, likely linked to elevated arousal or exploratory drive.

4.2.2 Sound-Triggered Whisking

Next, we examined whether pure tones could elicit whisker movements in head-fixed mice, particularly under conditions where locomotion was absent. By presenting pure tones of varying frequencies and intensities, we found that sound-evoked whisking was not only reliable but also influenced by the tone's acoustic properties and stimulus history.

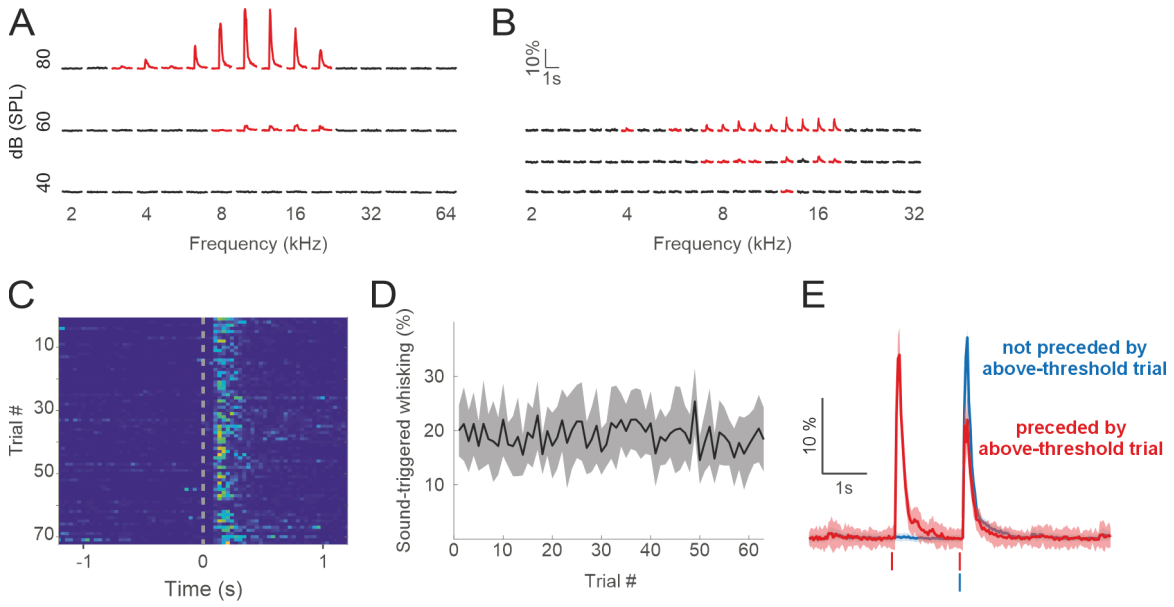

Figure 27. Pure tones trigger whisking. (A) Mean sound-triggered whisking as a function of pure tone frequency and intensity for three mice. Red traces indicate frequency-intensity combinations that elicited a statistically significant increase in whisking ($p < 0.05$, Wilcoxon signed-rank test; n = 234–254 sound presentations per frequency-intensity combination across all mice and sessions). Sound presentations occurring during locomotion periods were excluded. **(B)** Same as **(A)** for two different mice presented with lower-intensity pure tones. Sessions were conducted with mice restrained in an acrylic body tube (n = 120 sound presentations per frequency-intensity combination across all mice and sessions). **(C)** Trial-by-trial whisking aligned to sound onset (gray line) for a single session involving 80 dB SPL tones with particularly effective frequencies (8–20.15 kHz; $p < 0.001$). **(D)** Mean sound-triggered whisking across 14 recording sessions (three mice) as a function of trial number, showing stable whisking amplitudes for "above-threshold" tones. Shaded area indicates the 95% confidence interval. **(E)** Comparison of mean whisking amplitudes for above-threshold trials that were (red, n = 191) or were not (blue, n = 1,288) preceded by another above-threshold trial. Whisking responses were significantly smaller when preceded by an above-threshold trial ($p < 10^{-6}$, Wilcoxon rank

sum test). Vertical lines indicate sound onset. Shaded areas in **(D)** and **(E)** indicate 95% confidence intervals.

4.2.2.1 Frequency and Intensity Dependence of Sound-Triggered Whisking

Consistent with the mouse's auditory sensitivity, pure tones elicited significant whisking when their frequency ranged between approximately 3 kHz and 18 kHz—the region of highest hearing sensitivity. Whisking responses became more robust as sound intensity increased, with the largest whisking amplitudes occurring during presentations of medium-frequency tones (e.g., 8–16 kHz) at 80 dB SPL (Figure 27, A). At lower sound intensities (e.g., 50 dB SPL), significant whisking events were still observed but were less pronounced (Figure 27, B).

4.2.2.2 Persistence and Stability of Sound-Triggered Whisking Across Trials

To assess whether repeated sound presentations diminished the whisking response due to habituation, we analysed whisking amplitudes over successive trials. For frequencies and intensities that reliably triggered whisking (termed "above-threshold trials"), we observed consistent whisking responses throughout the recording session (Figure 27, C) occurring with a latency of approximately 100 ms (3 video frames, Figure 27, C&D). Mean sound-triggered whisking amplitude remained stable across trials (Figure 27, D), indicating that mice did not habituate to the tones within a session and that pure tones retained their efficacy as a stimulus for eliciting whisking. However, their amplitude was influenced by the immediate trial history. Whisking amplitudes were significantly reduced when a loud tone trial was immediately preceded by another loud tone trial ($p < 10^{-6}$, Wilcoxon rank sum test; Figure 2E). This reduction in response suggests that sound-triggered whisking is subject to short-term adaptation or suppression, with the immediate stimulus history modulating the amplitude of the behavioural response.

While recent studies have emphasized the capacity of broadband sounds to elicit body and facial movements (Bimbard et al., 2023; Clayton et al., 2024), our findings extend this understanding by showing that pure tones, particularly those within the mouse's most sensitive hearing range, are also capable of reliably triggering whisking.

4.2.3 Neural Activity Correlates with Spontaneous and Sound-Triggered Whisking

Building on the observation that pure tones can reliably trigger whisking (Figure 27), we classified whisking events into two categories: spontaneous whisking and sound-triggered whisking. A whisking event was considered sound-triggered if it occurred within 1 s of an 80 dB SPL tone of a frequency that reliably triggered whisking ($p < 0.001$, 'above-threshold trials'). A whisking event was considered spontaneous if it did not occur close to a tone of a frequency-level combination that caused a significant increase in whisking ($p < 0.05$).

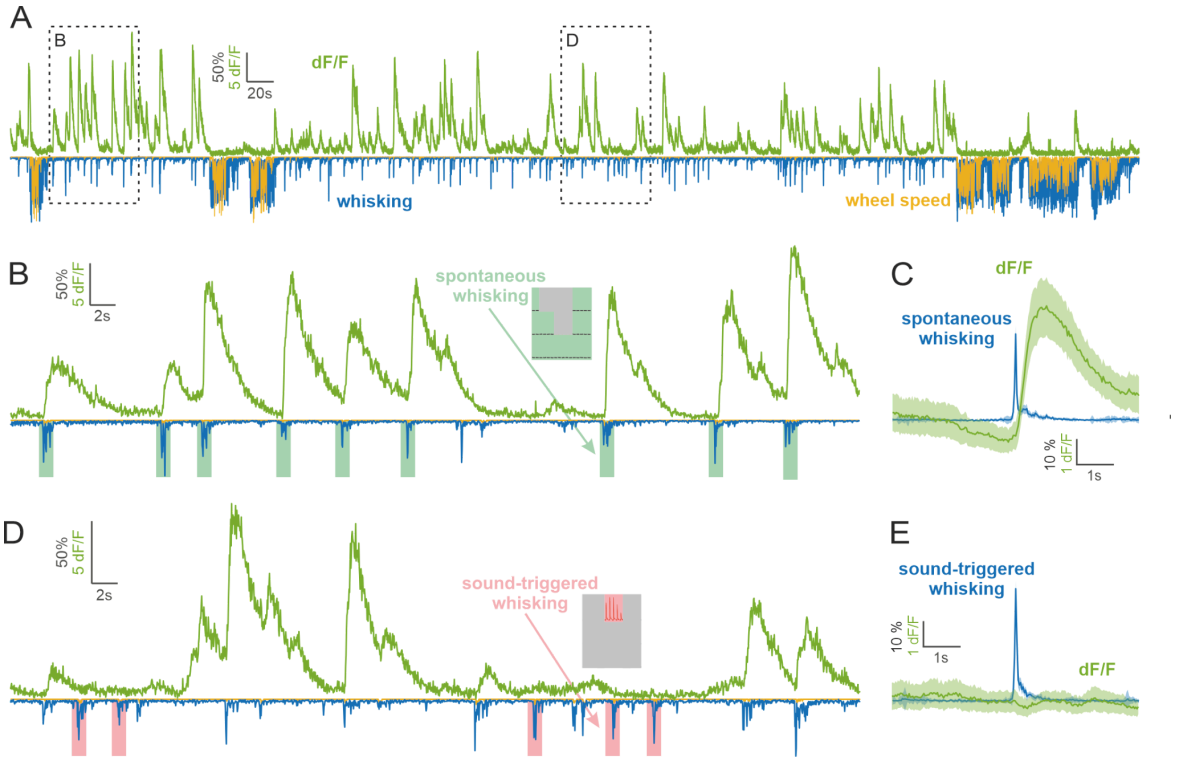

Figure 28. Relationship between whisking and neural activity for one example neuron in the auditory cortex. (A) Example trace showing whisking (blue), wheel movement (yellow), and fluorescence responses (green) for one neuron recorded in the auditory cortex. **(B)** Expanded segment of the trace from **(A)** marked by the small 'B.' Light green rectangles denote spontaneous whisking events. **(C)** Mean fluorescence response (green) for the same example neurons aligned to the spontaneous whisking events (blue) (n = 78 events). **(D)** Expanded segment of the trace from **(A)** marked by the small 'D.' Light red rectangles denote sound-triggered whisking events. **(E)** Mean fluorescence response (green) aligned to the sound-triggered whisking events (blue) for the same example neuron (n = 53 events). Shaded areas in **(C)** and **(E)** indicate 95% confidence intervals. The comparison between **(C)** and **(E)** illustrates distinct neuronal response profiles to spontaneous and sound-triggered whisking.

To isolate neuronal responses that are specifically coupled to whisking, we analysed activity exclusively around spontaneous whisking events. Figure 28 A-C illustrates the relationship between whisking, locomotion, and the activity of one example neuron recorded in the auditory cortex. Spontaneous whisking events (indicated by light green rectangles in Figure 38 B) elicited robust fluorescence responses in this neuron, as shown by the average trace aligned to spontaneous whisking events (Figure 3C).

Across all recordings, we identified 6,359 neurons from 14 sessions in three mice. Of these, 399 neurons (6.3%) were significantly modulated by spontaneous whisking ($p < 0.001$, Kruskal-Wallis test). The proportion of whisking-modulated neurons was comparable between the primary auditory cortex (A1, 6%, range: 0.5–13%) and the anterior auditory field (AAF, 9%, range: 1–15.5%), suggesting a consistent representation of whisking across auditory cortical areas.

4.2.4 Neural Activity During Sound-Triggered Whisking vs Spontaneous Whisking

We next asked whether these same spontaneous whisking-modulated neurons responded similarly when whisking was triggered by an external tone. Figures 28 B and 28 C illustrate an example neuron strongly excited by spontaneous whisking yet showing no significant modulation during tone-evoked whisking (Figure 28, D&E). Across the full population of spontaneous whisking–modulated neurons, 40% did not exhibit significant activity changes during sound-triggered whisking. Notably, this was not attributable to weaker whisker movements: on average, sound-triggered whisking bouts were slightly larger in amplitude than spontaneous ones ($p < 10^{-24}$, Wilcoxon rank sum test; Figure 29 C). Instead, the difference appears tied to the context of movement (i.e., spontaneous vs. externally induced).

A similar context dependence was evident in neurons suppressed by spontaneous whisking: among those cells, 32% remained suppressed, 50% lost their modulation, and 18% switched to excitation during sound-triggered whisking (Figure 29, B). Taken together, these results underscore that whisking in the auditory cortex is encoded differently depending on whether it arises spontaneously or is elicited by a sound—a pattern indicative of more complex, context-dependent sensorimotor integration than can be explained by whisking amplitude alone. The negative correlation between neural activity related to spontaneous whisking and neural activity related to sound-triggered whisking (Figure 29, D) indicates that neurons that are strongly activated by spontaneous whisking tend to not be strongly activated by sound-triggered whisking and vice versa.

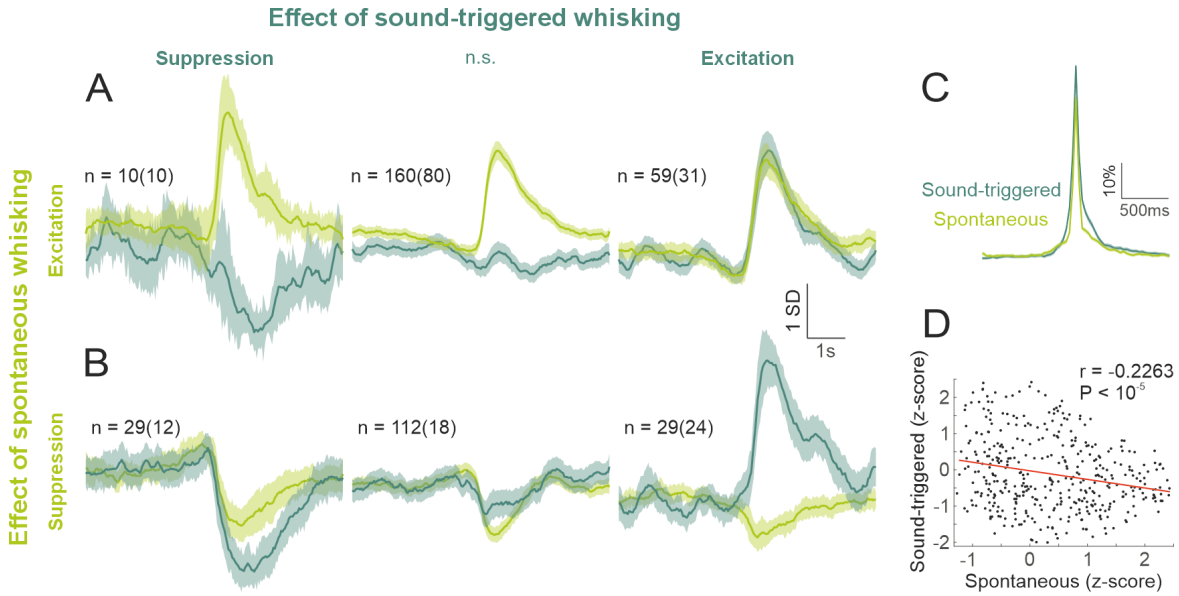

Figure 29. Whisking-sensitive neurons in the auditory cortex encode spontaneous whisking differently from sound-triggered whisking. (A) Mean fluorescence responses of neurons excited by spontaneous whisking (light green) across three groups based on their responses to sound-triggered whisking (dark green): suppression (left, n = 10), no significant modulation (middle, n = 160), and excitation (right, n = 59). (B) Mean fluorescence responses of neurons suppressed by spontaneous

whisking (light green) across three groups based on their responses to sound-triggered whisking (dark green): suppression (left, n = 29), no significant modulation (middle, n = 112), and excitation (right, n = 29). (C) Mean whisking amplitude for sound-triggered (dark green) and spontaneous (light green) whisking events, demonstrating a statistically significant but small increase in amplitude for sound-triggered whisking ($p < 10^{-24}$, Wilcoxon rank sum test). Shaded areas indicate 95% confidence intervals. (D) Neural response (mean of the z-scored fluorescence 0 to 1.5 s after whisking event) related to spontaneous whisking vs neural response to sound-triggered whisking. 'r' indicates the Pearson correlation coefficient.

These results reveal that whisking-sensitive neurons in the auditory cortex respond differently between spontaneous and sound-triggered whisking, suggesting that the neural circuits involved in processing whisker movements are dynamically influenced by the behavioural and sensory context in which those movements occur. To evaluate whether individual neurons distinguish between spontaneous and sound-triggered whisking, we performed a receiver operating characteristic (ROC) analysis on the distributions of neural activity measured around each whisking type. At every imaging frame, we generated an ROC curve comparing spontaneous and sound-triggered event-related activity, and we then computed the area under the curve (AUC; Figure 30, A&B). We defined an AUC as significant if it exceeded the mean shuffle-based AUC by at least three standard deviations (Gilad & Helmchen, 2020). By this criterion, 175 out of 399 neurons showed reliably different responses to spontaneous versus sound-triggered whisking (Figure 30 C).

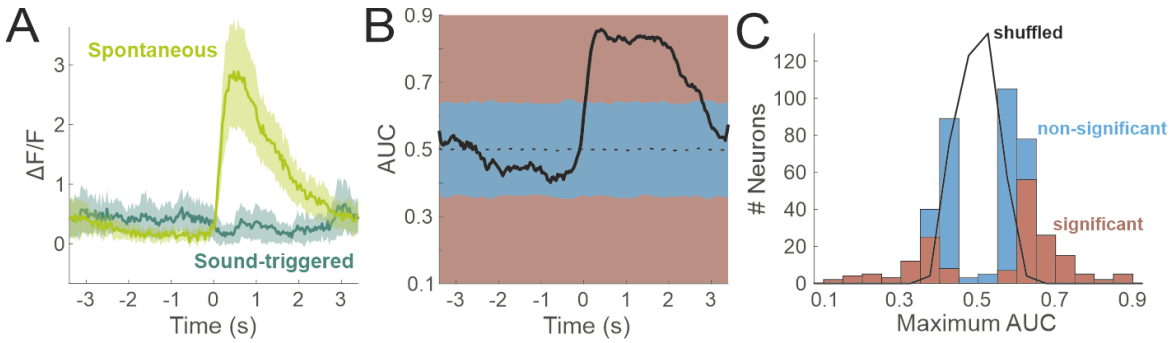

Figure 30. Receiver operating characteristic (ROC) analysis identifies neurons that discriminate between spontaneous and sound-triggered whisking. (A) Mean fluorescence response from a single example neuron, averaged separately for spontaneous (light green) and sound-triggered (dark green) whisking events. The shaded areas denote 95% confidence intervals, and time 0 marks the whisking event. **(B)** The area under the curve (AUC) is plotted for each time point from the ROC analysis applied to the responses in **(A)**. The dotted line represents the mean AUC obtained from shuffled data, and the blue shaded region indicates ±3 standard deviations around this mean. **(C)** Distribution of each neuron's maximum AUC (defined as its largest deviation from 0.5). Neurons whose maximum AUC exceeded the shuffled mean by 3 standard deviations are shown in brown (significant), while the remaining neurons are shown in blue (non-significant). The black line represents the distribution of maximum AUCs derived from the shuffle.

Finally, because locomotion is known to modulate sound-evoked activity in the auditory cortex, we asked whether whisker movements might similarly alter these responses. To address this, we examined tone-evoked activity during whisking—excluding trials where the tone itself triggered whisking—and compared it to the same stimuli when the mouse was not whisking. In contrast to locomotion, which significantly suppressed sound-evoked neural responses (Figure 31, A), whisking alone produced no measurable effect (Figure 31, B).

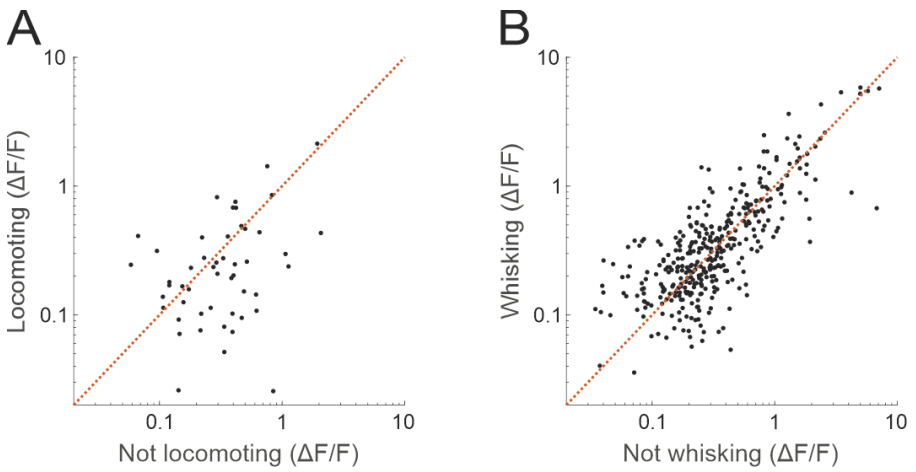

Figure 31. Locomotion suppresses sound-evoked activity in the auditory cortex, but whisking alone does not. Scatter plots comparing neuronal responses (ΔF/F) to sound stimuli during different behavioural states. **(A)** Responses during locomotion are plotted against responses to identical sounds during stationary periods, demonstrating significant suppression of sound-evoked neural activity during locomotion (Median ± IQR: locomoting, 0.17 ± 0.23 ΔF/F; not locomoting, 0.34 ± 0.34 ΔF/F; p = 0.002, n = 59 neurons). **(B)** By contrast, sound-evoked responses during whisking periods show no significant difference compared to non-whisking periods (whisking, 0.31 ± 0.37 ΔF/F; not whisking, 0.28 ± 0.37 ΔF/F; p = 0.14, n = 367 neurons). The red dotted line represents unity (equal response magnitudes).

4.3 Discussion

4.3.1 Whisking in Varying Behavioural Contexts

Our findings reveal several distinct relationships between whisker movements, behavioural states, and auditory cortical activity. Initially, we confirmed that whisking strongly correlates with locomotion, typically preceding locomotion onset and persisting briefly after movement cessation, highlighting whisking as an exploratory behaviour. Mice also exhibited spontaneous whisker movements during stationary periods, accompanied by transient pupil dilation and subtle body movements. These spontaneous whisking events may indicate smaller-scale shifts in brain state, potentially changes in arousal and are, as described previously (Collins et al., 2023; Eggermann et al., 2014), associated with neuromodulatory events, particularly increases in basal forebrain cholinergic activity and locus coeruleus noradrenergic activity.

Further, we showed that pure tones, particularly at higher intensities and within the mouse's most sensitive hearing range, reliably triggered whisker movements in line with recent findings demonstrating that salient or broadband auditory stimuli can elicit measurable facial or body movements (Clayton et al., 2024). The ability of loud pure tones to evoke whisking did not diminish over the course of a recording session. However, we observed modest reductions in whisking amplitude when loud tones occurred in rapid succession, akin to known short-term adaptation mechanisms in startle responses (Zheng & Schmid, 2023).

4.3.2 Brain State and Corollary Discharge in the Auditory Cortex

Importantly, we found that auditory cortical neurons exhibited distinct modulation depending on whisking context: 40% of neurons modulated by spontaneous whisking were not significantly responsive during sound-triggered whisking, even though sound-triggered whisking bouts were, on average, significantly larger in amplitude than spontaneous ones. ROC analysis further confirmed that nearly half of the whisking-modulated neurons reliably discriminated between these contexts. A parsimonious alternative is that a conventional sound-evoked response is simply added on top of the same whisking-locked drive present during spontaneous movements. If so, the shape of the response during sound-triggered whisking should closely match the neuron's sound-only response, just shifted (or stretched) by the whisking latency. In other words, the early, sharply time-locked component, typically peaking within ~50–100 ms of tone onset, would be identical in both conditions, while any additional activity aligned to the peak of whisker motion would be shared with spontaneous bouts.

If the temporal profiles do match, it becomes difficult to avoid the conclusion that direct auditory responses, rather than true context-dependent gating of the whisking signal, explain the apparent differences between spontaneous and sound-triggered conditions. Conversely, systematic changes in latency, rise-time, or response width would argue for genuine modulation of the whisking-related component and future analysis should therefore compare response kinetics across these conditions to disentangle additive from interactive effects.

Neurons selectively modulated by spontaneous whisking—but unaffected by sound-triggered whisking—likely reflect sensitivity to these neuromodulatory state changes. Since sound-triggered whisking is not rooted in a shift in brain state, these neurons show minimal or no modulation in this condition.

Neurons that are equally modulated by both spontaneous and sound-triggered whisking could reflect motor-related corollary discharge signals. Given that both forms of whisking involve active movement, these neurons may receive common movement-related predictive signals, independent of brain state changes (Nelson & Mooney, 2016; Schneider et al., 2014).

The activity of neurons exhibiting opposite modulation depending on whisking context (excited by spontaneous whisking and suppressed by sound-triggered whisking, or vice versa) could reflect integration of both brain state and corollary discharge signals, with one of the signals having an excitatory and the other having an inhibitory effect. Alternatively, direct neural responses to the sounds associated with sound-triggered whisking could account for differences in modulation between sound triggered and spontaneous whisking in these cells.

Viewed collectively, these differences align with predictive coding frameworks, suggesting that the auditory cortex selectively attenuates sensory inputs resulting from self-generated actions and amplifies unexpected external stimuli (Audette et al., 2022; Schneider et al., 2014). However, the observed heterogeneity across auditory cortical neurons indicates the coexistence of multiple parallel gating processes, perhaps enabling the brain to simultaneously predict self-generated sensory consequences while remaining sensitive to novel or unexpected external events but further research is needed to fully understand the mechanisms underlying and functional consequences of the different response profiles.

In addition to cortical circuitry, subcortical structures likely contribute significantly to this sensorimotor integration. For example, the superior colliculus integrates motor signals with sensory inputs, coordinating orienting behaviours (Martín-Cortecero et al., 2023). Thalamic nuclei such as the medial geniculate body and posterior medial nucleus may gate or amplify reafferent signals depending on whisker movement context. Similarly, basal ganglia regions, including the dorsolateral striatum, are

known to modulate sensory processing based on motor context, attenuating responses to self-generated stimuli (de la Torre-Martinez et al., 2023).

4.3.3 Future Directions and Methodological Extension

These hypothesized explanations suggest several promising next steps for understanding how whisker movements influence auditory cortical processing. Firstly, it will be important to directly measure the cholinergic and noradrenergic inputs to the whisking-sensitive neurons in the auditory cortex, following the approach recently described by Collins et al. (2023). Two-photon calcium imaging could be combined with the expression of GCaMP6s in cholinergic axons from the basal forebrain (BF-ACh) and noradrenergic axons from the locus coeruleus (LC-NA). This would enable the investigation of whether neuromodulatory signaling differences between spontaneous and sound-triggered whisking can explain the neuronal response patterns observed. It could be insightful to investigate whether local cortical cholinergic interneurons (VIP-ChAT interneurons, or VCINs), in addition to the long-range basal forebrain projections (BF-ACh), modulate auditory cortical responses differently. Collins et al. (2023) demonstrated that VCINs exhibit substantial heterogeneity, with only subsets reliably tracking behavioural state changes. This raises the possibility that local cholinergic signals may provide more spatially and functionally targeted modulation compared to the global arousal signals delivered by long-range cholinergic inputs.

Secondly, further experiments tracing connections at the circuit level could help clarify how motor signals reach the auditory cortex. Using anterograde and retrograde viral tracing from motor cortical areas responsible for whisker movement (vibrissal motor cortex) and from subcortical motor-related structures (e.g., superior colliculus, dorsolateral striatum) could enable identification of the precise pathways carrying potential corollary discharge signals. In addition, applying optogenetic

manipulations to these identified pathways could help establish causal evidence for their involvement in the observed auditory cortical responses, distinguishing whether particular circuits primarily convey predictive motor signals or are influenced more by neuromodulatory changes.

4.4 Methods

4.4.1 Animals

All procedures were approved by the Committee on Animal Care and Ethical Review at the University of Oxford and were licensed by the UK Home Office under the Animal Scientific Procedures Act (1986, amended 2012). Mice were maintained on a 12-hour light/dark cycle, housed at temperatures between 20–24°C, and kept at 45–65% relative humidity. All animals were approximately three months old at the time of data collection.

A total of eight mice were used for this study, representing three distinct genotypes: Figure 1 cohort: Three male double-transgenic mice generated by crossing the Ai95D line (B6;129S-Gt(ROSA)26Sor^tm95.1(CAG-GCaMP6f)Hze/J; JAX 024105) with Slc32a1^tm2(cre)Lowl/J (JAX 016962). Figure 2B cohort: Two female double-transgenic mice generated by crossing Ai95D (JAX 024105) with B6.Cg-Tg(Camk2a-cre)T29-1Stl/J (JAX 005359). Remaining figures (2A,C–E; 3–4): Three male mice generated by crossing B6;DBA-Tg(tetO-GCaMP6s)2Niell/J (JAX 024742) X B6.Cg-Igs7$^{tm94.1(tetO\text{-}GCaMP6s)Hze}$Tg(Camk2a-tTA)1Mmay/J (JAX 024115).

4.4.2 Surgery

Mice were premedicated prior to surgery with intraperitoneal injections of dexamethasone (Dexadreson, 4 mg), atropine (Atrocare, 1 mg), and carprofen (Rimadyl, 0.15 mg). Anesthesia was induced and maintained with 1.5–2% isoflurane in oxygen. Postoperative analgesia (buprenorphine, Vetergesic, 1 ml/kg) was administered following surgery.

For all surgical procedures, the animal's head was secured in a stereotaxic frame (Model 900LS, David Kopf Instruments). Body temperature was maintained at 37°C using a heating mat (FHC) controlled by a DC temperature controller and monitored via a rectal temperature probe. The scalp area over the right auditory cortex was shaved and disinfected (Chlorhexidine), and a small incision was made to expose the skull.

Head bar implantation (behaviour-only mice): A ~1 cm diameter flap of skin was removed over the skull's midline. The skull surface was cleared and dried, and a custom steel post was affixed to the skull using dental acrylic (Super-Bond C&B, Sun Medical).

Cranial window implantation (imaging mice): A slightly larger scalp area was removed on the right side of the skull, and part of the right temporal muscle was carefully removed to expose the auditory cortex. A 4 mm craniotomy was made over the auditory cortex. A glass coverslip (4 mm diameter) fused to a ~1 mm tall steel cylinder (0.5 mm wall thickness) was placed into the craniotomy such that the glass gently pushed down on the brain surface to minimise motion. The assembly was then secured using gel adhesive (Pattex Ultra Gel, Henkel) around the edges of the skull. Mice recovered for at least 7 days prior to recording.

4.4.3 Whisker, Pupil, and Locomotion Recording

Whisker and pupil movements were recorded using high-speed cameras (DMK 37BUX287 and DMK 23UV024, respectively; The Imaging Source). Frame rates were 200 Hz for the dataset in Figure 1 and 30 Hz for Figures 2–4. Two infrared illuminators (LIU780A, Thorlabs) were positioned in front of the mouse to enhance whisker visibility. One to three whisker tips were tracked, and the pupil was monitored by imaging the eye from a lateral viewpoint. The chamber was moderately lit in sessions that did not involve two-photon imaging to avoid persistent pupil dilation.

Locomotion: Mice were placed either in a transparent acrylic tube (Figures 2B) or on a freely rotating polystyrene wheel (coated with isolation foam to reduce noise) for the other experiments. When using the wheel, locomotion speed was measured by monitoring wheel rotation with an optical mouse (Logitech G500) connected to custom LabVIEW scripts (LabVIEW 2017, National Instruments). Locomotion was defined as wheel speed exceeding 3 cm/s.

Video synchronisation: Video frame timestamps were synchronised to the neural or behavioural data streams (e.g., sound presentations) via a NI-DAQ I/O device (USB-6501, National Instruments).

4.4.4 Sound Presentation

Pure tone stimuli of 200 ms duration (with 10 ms raised cosine onset and offset ramps) were generated at a 192 kHz sampling rate (custom LabView programs, National Instruments) through a 16-bit sound card (Xonar X7, Asus) and an ultrasonic amplifier (Portable Ultrasonic Power Amplified; Avisoft Bioacoustics). Tones were broadcast via an ultrasonic speaker (Avisoft Vifa) positioned ~10 cm

from the mouse's head. The speaker output was calibrated to maintain a flat (±2 dB) response over the frequencies used.

Stimulus sets: For Figures 2A, 2C–E, 3, and 4: Tones ranged from 2–64 kHz in ⅓-octave steps at intensities of 40, 60, and 80 dB SPL, with an interstimulus interval of 1.5 s. For Figure 2B: Tones ranged from 2–32 kHz in ⅙-octave steps at intensities of 40, 50, and 60 dB SPL, with an interstimulus interval of 1 s.

Widefield imaging stimuli: Sinusoidally amplitude-modulated (SAM) tones of 500 ms duration and 10 Hz modulation frequency (100% modulation depth) were presented at 55 or 65 dB SPL. Carrier frequencies tested were ~4.0 kHz, 5.04 kHz, 25.4 kHz, and 32.0 kHz. A 3 s interval separated successive stimuli, each repeated 10–15 times.

4.4.5 Calcium Imaging

To identify auditory cortical subfields before two-photon imaging, we performed widefield calcium imaging using a 340M-GE CCD camera (Thorlabs) fitted with a TL2x-SAP objective (Thorlabs) and a 470 nm LED (M470L4, Thorlabs). The system was attached to a rotating Bergamo II microscope body (Thorlabs). Data were acquired at 10–15 frames per second, and the acquired images were processed with custom MATLAB scripts to compute ΔF/F and generate tonotopic maps of the auditory cortex.

For higher-resolution measurements, two-photon calcium imaging was carried out in the primary auditory cortex (A1) and the anterior auditory field (AAF) at depths of 150–250 μm (layer II/III). We used a rotating multiphoton laser-scanning microscope (Bergamo II, Thorlabs) equipped with a

20×/1.00 NA water-immersion objective (Olympus). Excitation light at 940 nm was produced by a femtosecond laser (Chameleon Discovery, Coherent) at ~1300 mW, and directed onto the brain by an 8 kHz resonant scanner (for the fast axis) and a galvanometric scanner (for the slow axis). Typical laser power under the objective ranged from 45–70 mW. Images were acquired in bidirectional mode at ~30 Hz, with a resolution of 512 × 512 pixels over a 600 × 600 μm field of view. Scanner noise near the mouse's head was <40 dB SPL.

Pre-processing of two-photon data included rigid and non-rigid motion correction, segmentation of regions of interest (ROIs), and neuropil correction using the Python version of suite2p (Pachitariu et al., 2017). Neuropil contamination was subtracted from raw fluorescence signals using a coefficient of 0.7 (Kerlin et al., 2010; Pachitariu et al., 2017). The baseline fluorescence (F_0) for $\Delta F/F$ calculation was taken as the median of the fluorescence trace, excluding values below the 10th percentile and above the 70th percentile, following previously published approaches.

4.4.6 Data Analysis

4.4.6.1 Whisker and Pupil Tracking

We used DeepLabCut (Mathis et al., 2018) with a ResNet50 architecture to track (i) four points on the pupil's rim and (ii) the tips of 1–3 whiskers. Around 120 manually annotated video frames per animal were used for training, and additional frames were added if needed to improve labeling accuracy. Each mouse had a unique network to accommodate individual whisker identities. Post-training validation confirmed an average Euclidean tracking error below 7 pixels. Pupil diameter was estimated by averaging the distances between opposing pupil-rim markers, while whisker movements were quantified by frame-to-frame displacement of tracked whisker tips.

4.4.6.2 Whisking, Pupil, and Locomotion Quantification

For each session, whisker movements (displacements) were smoothed (100 ms Gaussian window) and averaged across all tracked whiskers to yield a single continuous whisker trace. Locomotion was defined by wheel speeds exceeding 3 cm/s. Pupil diameter, whisker movement, and wheel speed were normalised as percentages of their respective session-wide maxima (except when explicitly stated otherwise).

Identification of Whisking Events: To focus on whisking events outside locomotion, we excluded data recorded during locomotion and up to 10 s after the mouse stopped running. Using the MATLAB findpeaks function on the smoothed whisker trace, local maxima were identified as whisking events. Events occurring within 2 s of one another were excluded to isolate distinct bouts.

4.4.6.3 Sound-Triggered Whisking

A whisking event was considered spontaneous if it did not occur within 1 s after the onset of a tone with a frequency-level combination that causes a significant ($p < 0.05$, Wilcoxon signed rank test) increase in mean whisker movement from the 500 ms before sound onset to the 300 ms after sound onset. This approach excluded, as desired, even frequency-level combinations that cause very small increases in whisking. To find whisking events that are sound-triggered we applied very stringent criteria ($p < 0.001$, only 80 dB SPL sounds) in order to identify frequency-level combinations that reliably triggered whisking. Whisking events that occurred within 1 s of the onset of a tone with such a frequency-level combination were considered sound-triggered. Neurons were considered to be modulated by spontaneous (sound-triggered) whisking if they exhibited a statistically significant difference ($p < 0.001$, Kruskal-Wallis test) in fluorescence among the 70 frames (from 20 frames before

the whisking event peak to 50 frames after) surrounding spontaneous (sound-triggered) whisking events. Confidence intervals were typically computed at the 95% level, and significance thresholds are reported in the figure legends. All data analysis and figure generation were performed using custom routines in MATLAB (R2022b, MathWorks) and Python 3.8.

4.4.6.4 Assessing Behavioural State-Dependent Sound Responses

To determine whether individual neurons distinguish between spontaneous and sound-triggered whisking, we computed Receiver Operating Characteristic (ROC) curves for the neural activity distributions surrounding each whisking type. For each imaging frame from 100 frames before a whisking event to 100 frames after, we generated an ROC curve contrasting spontaneous-event activity with sound-triggered-event activity. We then considered the area under the curve (AUC) significant if it exceeded the mean shuffled (100 iterations) AUC by at least three standard deviations (Gilad & Helmchen, 2020).

To compare neural responses to tones delivered during locomotion versus stationary periods—and likewise during whisking versus non-whisking periods—we considered only frequency-level combinations that did not elicit significant whisking. We then selected neurons that showed a significant difference (one-way ANOVA, $p < 0.001$) in their responses across thesefrequency-level combinations. From each neuron, we identified the specific frequency-level pairs with mean responses at least 1.5 standard deviations above that neuron's overall mean. Next, for locomotion analyses, we only included neurons with ≥5 stimulus presentations during locomotion and ≥5 matching presentations (i.e., same frequency-level combination) during stationary periods. Analogously, to evaluate whisking effects, we included neurons with ≥5 presentations during whisking and ≥5 during non-whisking. Here, whisking was defined as any 1 s Gaussian-smoothed whisker trace

exceeding 5% of the session's maximum whisker movement; lower values were classified as 'not whisking.' Because of these multiple selection criteria, the number of neurons ultimately analysed—particularly in the locomotion condition—was substantially reduced.

V General Discussion

We opened this thesis with the question of what the behavioural relevance of whisker-driven suppression of sound responses in mouse auditory cortex is and the suspicion that it is not necessarily a hard-wired connection but a flexible computation that can be tuned by behavioural demands, stimulus history and motor context. Over three experimental chapters I followed that idea from complementary angles—perceptual implications, stimulus relevance, and sensorimotor-state coupling.

5.1 Summary of Findings

5.1.1 Whisker-driven suppression is graded by behavioural relevance

Throughout Chapters II-III, I showed that whisker stimulation produced the canonical divisive scaling of tone responses in A1/AAF. While we did not attempt to replicate any subcortical or S1 recordings, this observation is in accordance with the action of the S1 → shell-IC → MGBv gate first described by Lohse and colleagues. Crucially, the magnitude of that scaling correlated with the behavioural value as measured by animal reward collection performance: it was largest when tones were rewarded and whiskers irrelevant, weakened when whiskers carried the reward, and became minimal after prolonged exposure to unrewarded whisker trials. These observations suggest that the subcortical loop functions not as an immutable "multisensory veto" but as a filter that can be reweighted by learning or motivation. The fact that the somatosensory cortex - well known for experience-dependent plasticity - sits at the top of the loop makes such flexibility plausible. Indeed, the very involvement of somatosensory cortex—a structure capable of experience-dependent and cross-modal plasticity—implies that the gate is built for adjustment rather than rigidity similar to what is known in other sensory domains (Ewall et al.,

2021; Huang et al., 2024). Hence, one interpretation is that the gate is re-weighted by increased or decreased activity in S1 which tightens or loosens the IC gating, with a less likely alternative being reweighting in the S1 → IC shell → MGB link itself, with corticofugal feedback or neuromodulators adjusting the inhibitory drive that reaches MGBv.

5.1.2 Behavioural impact of whisker input depends on task history and demands

In our experiments, the behavioural impact of whisker stimulation varied with training history and task demands. In our near-threshold pure-tone detection task, where whisker stimulation was present from the outset but never rewarded, mice likely eventually learned to ignore it. Under these familiar conditions, concurrent whisker deflections slightly shortened reaction times without affecting sensitivity. In contrast, when the same whisker stimulus was introduced only at test—during a broadband-noise detection task in which premature licks were penalised—it lengthened response times, again without altering detection performance. This bidirectional latency effect implies that behavioural history determines whether the same somatosensory stimulus acts as a distractor, a cue, or is treated as noise that can be ignored. Notably, because detection sensitivity was unchanged in either case, it seems that whisker stimulation neither facilitated nor impaired tone detectability per se; rather, its influence lies in modulating the speed of decisions under conditions of uncertainty.

A different picture emerged in the frequency discrimination task requiring mice to distinguish between two tones spaced closely in frequency: Here, the addition of whisker stimulation significantly improved discrimination performance, without altering reaction times. Our interpretation is that suppression reduced the overlap between frequency-tuned channels, thereby enhancing the contrast between neural responses to nearby tones. In other words, divisive scaling may decorrelate overlapping

representations, improving discriminability without necessarily impairing detectability, aligning with theoretical predictions and prior work showing that reduced response magnitude can, under the right conditions, lead to sharper tuning and improved perceptual resolution (Christensen et al., 2019) .

Converging human, ferret and rodent work shows that pairing sound with touch can improve perception well beyond either modality alone. Near-threshold vibrotactile pulses become easier to detect and elicit faster responses when accompanied by a brief click, just as tones enhance tactile detectability in the reverse configuration (Gillmeister & Eimer, 2007; Schürmann et al., 2004; Sperdin et al., 2009). The facilitation is greatest when the temporal or spectral content of the two stimuli matches: performance gains of up to two-fold and ~20 ms reaction-time reductions are observed when auditory and vibrotactile carriers share the same flutter or vibration frequency (Ro et al., 2009; Wilson et al., 2010). Crossmodal coupling also inflates perceived magnitude—tones make taps feel "louder" and touch increases judged sound intensity (Gillmeister & Eimer, 2007). Such benefits generalise to ecological settings: trumpet players display selectively amplified cortical responses and quicker cue extraction for lip vibrations paired with the instrument's pitch (Schulz et al., 2003), and mice lick sooner and more accurately when a fore-paw tap is accompanied by broadband noise in a tactile detection task (Godenzini et al., 2021).

In ferrets, adding a spatially aligned light to an auditory location cue barely improved the speed of reflexive head-orienting, with benefits fully explained by probability summation; yet the same audiovisual pairing markedly sharpened accuracy and reduced latency when the animals had to approach the correct loudspeaker for reward—performance exceeded race-model bounds and therefore signified true multisensory integration (Hammond-Kenny et al., 2017). In humans, a visual flicker that is temporally coherent with one of two competing sound streams—while carrying no explicit information about target identity—selectively enhances timbre discrimination, likewise surpassing statistical facilitation and demonstrating that temporal concordance alone can bias auditory scene analysis (Atilgan et al., 2018;

Atilgan & Bizley, 2021). By contrast, the whisker stimulus in our frequency-discrimination task was spatially non-specific, identical on every trial, and behaviourally meaningless; thus neither spatial cue binding nor probability summation can account for the observed improvement. Instead, the data align with divisive suppression sharpening auditory frequency channels, reducing overlap and enhancing discrimination through circuit-level re-weighting rather than through explicit cue integration.

At the cortical level, there is now ample evidence that a reduction in mean firing can increase the sensory information that a population carries, in line with our own observations. In ferret auditory cortex, spatially congruent light flashes depress overall spike rates, but mutual-information analyses show that the same neurons convey more precise azimuthal information under audiovisual stimulation, an effect that is strongest in non-primary fields (Bizley & King, 2008, 2009). A follow-up study demonstrated a comparable "informational boost" for pitch and timbre: concurrent visual input suppressed peak firing yet decorrelated responses across neurons, thereby sharpening neural resolution along non-spatial dimensions (Atilgan et al., 2018). Similar suppression-with-gain motifs have been reported for naturalistic audiovisual movies in macaque A1, where divisive down-scaling is accompanied by a ~30 % rise in information about the soundtrack (Kayser et al., 2010), and for cross-modally evoked inhibition in mouse V1, which narrows orientation tuning and improves single-trial decoding (Ibrahim et al., 2016; Kayser et al., 2010), as mentioned in the discussion of Chapter II.

Taken together, these studies suggest that divisive suppression can re-allocate dynamic range, quenching noise correlations, reducing channel overlap and thus enhancing discriminability, without harming simple detectability. Our observation that whisker-driven suppression in core fields (A1/AAF) improves fine frequency judgements therefore fits a broader multisensory pattern: what appears as "response dampening" at the level of spike counts can translate into a net informational gain at the population level. Critically, the ferret data also warn that the magnitude of this gain depends on cortical

field: increases in spatial information were modest in A1 but pronounced in belt/anterodorsal regions where visual inputs are densest. If an analogous architecture exists for somatosensory inputs in mice, the most substantial sharpening may occur downstream of A1/AAF, a compelling reason to extend future recordings to higher-order auditory areas.

5.1.3 Brain-state dependent whisker signals in auditory cortex

Roughly 6 % of neurons in auditory cortex were modulated by spontaneous whisking and many of them exhibited different modulation or no modulation at all when the whisking was triggered by a sound presentation. One candidate for driving whisker-related neural activity in the auditory cortex are corollary discharge signals from motor areas. However, given that whisker movements that happen spontaneously and whisker movements that are triggered by sounds are near-identical and therefore likely to be accompanied by very similar corollary discharge signals, the notion of corollary discharge as the (sole) driver of whisker-related modulation is difficult to reconcile with the fact that a large proportion of neurons are modulated by spontaneous whisking and not (or differently) modulated by sound-triggered whisking. Given the fact that we found spontaneous whisking to be tightly coupled to pupil dilation and that others have demonstrated that spontaneous whisking is accompanied by a rise in cortical cholinergic and noradrenergic activity (Collins et al., 2023), the most parsimonious explanation is therefore that spontaneous whisking is accompanied (and perhaps even triggered) by a change in brain state (arousal) whereas sound-triggered whisking is not, resulting in different neural signatures in many neurons for spontaneous and sound-triggered whisking. Furthermore, some neurons that are modulated by spontaneous whisking may be excited or suppressed directly by the sounds that cause sound-triggered

whisking and therefore exhibit different activity in relation to sound-triggered whisking as opposed to spontaneous whisking.

5.2 Limitations

Despite the breadth of behavioural approaches used in this thesis, several constraints remain. The most salient shortcomings, how they affect interpretation of the data, and how they could matter for the next generation of experiments are:

5.2.1 Lack of causal perturbation

The first—and arguably most important—limitation is the absence of causal perturbations. Every dataset is observational: two-photon calcium imaging was combined with behavioural read-outs, but no projection-specific opto/chemogenetic silencing or activation was applied. Consequently, the S1 → lateral-colliculus-of-the-inferior-colliculus (LCIC) → MGB gate that we invoke to explain differences in behaviour, while well-supported by passive stimulation studies (Lohse et al., 2021) remains a likely hypothesis in our behavioural experiments rather than a demonstrated necessity. Likewise, we never transiently silenced or activated auditory-cortical circuits. In passively listening mice, Lohse et al showed that whisker-driven suppression does not involve GABAergic interneurons in A1, indicating that cortical inhibition is not sufficient, on its own, to generate the effect. What remains unknown is whether cortical interneurons can modulate the strength of the subcortical signal when the animal is alert, learning or moving - conditions under which neuromodulatory gain-control motifs are commonly engaged (Ferguson & Cardin, 2020). Because the magnitude of suppression tracked reward contingency in our

data, we cannot yet rule out whether A1 echoes a “decision” taken subcortically or whether it is an adjustable stage in the gain control hierarchy.

5.2.2 Restricted Systems Coverage

Recordings were confined to layers 2/3 of A1 and AAF. No data were collected in other cortical layers, the cochlear nucleus, IC, MGB, thalamic reticular nucleus or higher-order belt/temporal association fields where multisensory convergence is often strongest. The thesis infers the locus of suppression from prior literature rather than demonstrating it in the context of our experiments directly. This omission leaves open the possibility that the first tactile-auditory interaction happens earlier in the somatosensory or auditory pathways than assumed, and that higher-order fields exhibit very different balances of suppression and facilitation.

5.2.3 Limited Sensory and Behavioural Space

Related is the limited sensory space we sampled. External somatosensory drive was delivered with a brush or mesh piezo that simultaneously deflected large swathes of the vibrissa array at a fixed 20 Hz, 1 mm amplitude. These global, rhythmic, vertical deflections do not reproduce the sparse, texture-specific and phase-locked contacts that occur when a freely-moving mouse palpates an object and it appears natural that the greatest sensitivity of whiskers should be in the direction of active whisking, where most resistance to objects would occur - along the sagittal axis. Active whisking was monitored but never manipulated; no single-whisker flicks, no phase-locked perturbations, and no causal disruption of motor timing were attempted. The behavioural repertoire studied here was equally circumscribed. All tasks

were single-cue, Go/No-Go paradigms under head fixation. Although they revealed nuanced interactions between reward contingencies and sensory gain, they do not mimic real-world situations in which whisker and auditory cues evolve continuously and must be integrated on the fly—as during obstacle avoidance, social approach, or prey capture. Nor was the mouse ever allowed to choose between auditory and tactile information sources; thus, circuit mechanisms that adjudicate cue weighting in free choice remain unexplored.

5.2.4 Biological Breadth

The cohort was restricted to young adult mice. No juveniles, aged animals or females were examined, so potential developmental, ageing-related or sex-specific differences in multisensory gating are unknown. Likewise, all conclusions are mouse-specific; whether rats, ferrets or primates exploit an analogous LCIC gate, or rely more heavily on cortical integration, is an open question.

5.2.5 Statistical Structure and recordings

From a statistical perspective, datasets were pooled across sessions and animals without hierarchical modelling of inter-mouse variance. Because high-yield animals contributed disproportionately to the pooled estimates, true population effect sizes may be smaller (or occasionally larger) than reported. We tried to make this transparent through showing data from individual animals and sessions in addition to the combined data throughout.

These limitations do not, however, undercut the main empirical findings: 1) whisker stimulation leaves sound detection intact, 2) whisker stimulation improves sound discrimination, 3) for our specific stimuli, mice take far longer to learn a whisker stimulation detection task than a sound detection task (at least when using whole whisker pad stimulations), 4) the amplitude of suppression, as well as whisker- and sound-alone responses, scales with the reward contingency, 5) pure tones can effectively and reliably trigger whisking, 6) Spontaneous whisking bouts—accompanied by pupil dilation—activate a subset of auditory-cortical neurons that often show little or even opposite modulation when the very same whisking kinematics are triggered by a sound.

5.3 Future Directions

While numerous, these limitations have the benefit of presenting many paths for a future wave of experiments.

5.3.1 Neuromodulatory context: distinguishing arousal from corollary discharge

Our data show that many A1/AAF neurons respond to spontaneous whisking but not to sound-triggered whisking of similar amplitude, implying a hidden state variable—most likely cholinergic or noradrenergic tone. Future work should pair GCaMP imaging of auditory cortex neurons with axonal GCaMP recordings from basal-forebrain ACh and locus-coeruleus NA fibres (Collins et al., 2023). Closed-loop optogenetic stimulation of BF-ACh or LC-NA during sound-triggered whisking can then reveal whether boosting arousal reinstates the “spontaneous” response profile. Conversely, suppressing these modulators should test whether whisking modulation vanishes entirely. By isolating arousal

contributions, such experiments will clarify when cortical neurons treat whisking as a motor prediction versus a state change.

5.3.2 Cross-Species Test of Somatosensory-Auditory Gating

It is entirely plausible that the powerful, whisker-specific down-scaling of auditory responses we describe reflects a strategy adapted to the ethological niche of small, burrowing rodents rather than a universal brain design. In mice, vibrissae are the dominant near-field sensor, so a dedicated S1 → IC-shell → MGBv gate that silences sound-evoked cortical activity whenever the whiskers are touched could simply be the whisker analogue of saccadic suppression in foveate species. The first step toward testing that idea—before leaping to broader conclusions—should stick with the mouse as the animal model but swap the tactile channel: does stroking a glabrous digit, the pinna, or a tail hair produce the same divisive suppression of lemniscal sound responses? If the effect disappears (or reverses) with non-whisker somatosensation, we are dealing with a vibrissa-specialised circuit.

If, however, any salient tactile cue that coincides with a sound engages the same S1 → midbrain → thalamus pathway, the mechanism is likely more general (Lesicko et al., 2016)—and worth hunting for in species whose principal tactile organ is not a whisker. Several lines of anatomical evidence hint that such cross-modal corticofugal gates are at least possible:

In ferrets, retrograde tracers injected in the IC fill large layer-5 pyramidal neurons not only across auditory cortex (A1) but also in the neighbouring somatosensory sulci; the labelled axons arborise densely in the IC's lateral nucleus and dorsal cortex, the subdivisions homologous to the mouse "shell" (Bajo et al., 2006). Classic cat work complements this: Cooper & Young (1976) demonstrated that

injections in the same IC shell zones retrogradely label pyramidal cells in primary somatosensory cortex (S1) as well as A1, confirming a direct S1 → IC pathway in a non-whisker carnivore (Cooper & Young, 1976). In Marsupials, Robards (1979) used tiny HRP deposits confined to the IC external (lateral) nucleus and showed ipsilateral labelling of layer-5 neurons along the caudal parietotemporal strip that includes S1, proving that a tactile corticocollicular channel exists outside the eutherian lineage (Robards, 1979).

5.4 Conclusion: Transthalamic Circuits: Dynamic Bridges for Cross-Cortical Communication

The S1 → IC → MGB → A1 loop that motivates much of this thesis exemplifies a broader, transthalamic strategy for inter-areal dialogue: instead of relying solely on long corticocortical axons, functionally distinct cortical fields can "talk" to one another through cascades that detour via midbrain and higher-order thalamus, thereby inserting an extra layer of filtering, gating and amplification (Sherman and Usrey 2024). Anatomical work has mapped these circuits in exquisite detail, yet we still know remarkably little about what they do when an animal is awake, learning or moving. By pairing high-precision psychophysics with simultaneous recordings and perturbations across the loop, the present studies show how behavioural context shapes both the direction and the sign of cross-modal influences, and they reveal concrete, testable mechanisms—divisive suppression, state-dependent gain and interneuron-mediated disinhibition—through which subcortical hubs can retune cortical representations on the fly. In doing so, they move the field from descriptive anatomy toward causal physiology, and set the stage for future work that will probe, in real time, how transthalamic pathways integrate multisensory evidence and relay task-relevant predictions between distributed cortical modules.

VI References

Ansorge, J., Wu, C., Shore, S. E., & Krieger, P. (2021). Audiotactile interactions in the mouse cochlear nucleus. *Scientific Reports*, *11*(1), 6887.

Atiani, S., Elhilali, M., David, S. V., Fritz, J. B., & Shamma, S. A. (2009). Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron*, *61*(3), 467–480.

Atilgan, H., & Bizley, J. K. (2021). Training enhances the ability of listeners to exploit visual information for auditory scene analysis. *Cognition*, *208*, 104529.

Atilgan, H., Town, S. M., Wood, K. C., Jones, G. P., Maddox, R. K., Lee, A. K. C., & Bizley, J. K. (2018). Integration of Visual Information in Auditory Cortex Promotes Auditory Scene Analysis through Multisensory Binding. *Neuron*, *97*(3). https://doi.org/10.1016/j.neuron.2017.12.034

Audette, N. J., Zhou, W., La Chioma, A., & Schneider, D. M. (2022). Precise movement-based predictions in the mouse auditory cortex. *Current Biology : CB*, *32*(22), 4925–4940.e6.

Bajo, V. M., & King, A. J. (2012). Cortical modulation of auditory processing in the midbrain. *Frontiers in Neural Circuits*, *6*, 114.

Bajo, V. M., Nodal, F. R., Bizley, J. K., Moore, D. R., & King, A. J. (2006). The Ferret Auditory Cortex: Descending Projections to the Inferior Colliculus. *Cerebral Cortex (New York, N.Y. : 1991)*, *17*(2), 475.

Bartlett, E. L. (2013). The organization and physiology of the auditory thalamus and its role in processing acoustic features important for speech perception. *Brain and Language*, *126*(1). https://doi.org/10.1016/j.bandl.2013.03.003

Bigelow, J., Morrill, R. J., Olsen, T., & Hasenstaub, A. R. (2022). Visual modulation of firing and spectrotemporal receptive fields in mouse auditory cortex. *Current Research in Neurobiology*, *3*,

100040.

Bimbard, C., Sit, T. P. H., Lebedeva, A., Reddy, C. B., Harris, K. D., & Carandini, M. (2023). Behavioral origin of sound-evoked activity in mouse visual cortex. *Nature Neuroscience*, *26*(2), 251–258.

Bizley, J. K., Jones, G. P., & Town, S. M. (2016). Where are multisensory signals combined for perceptual decision-making? *Current Opinion in Neurobiology*, *40*, 31–37.

Bizley, J. K., & King, A. J. (2008). Visual-auditory spatial processing in auditory cortical neurons. *Brain Research*, *1242*, 24–36.

Bizley, J. K., & King, A. J. (2009). Visual influences on ferret auditory cortex. *Hearing Research*, *258*(1-2), 55–63.

Bizley, J. K., Maddox, R. K., & Lee, A. K. C. (2016). Defining auditory-visual objects: Behavioral tests and physiological mechanisms. *Trends in Neurosciences*, *39*(2), 74–85.

Bordi, F., & LeDoux, J. E. (1994). Response properties of single units in areas of rat auditory thalamus that project to the amygdala. II. Cells receiving convergent auditory and somatosensory inputs and cells antidromically activated by amygdala stimulation. *Experimental Brain Research*, *98*(2). https://doi.org/10.1007/BF00228415

Brungart, D. S., & Rabinowitz, W. M. (1999). Auditory localization of nearby sources. Head-related transfer functions. *The Journal of the Acoustical Society of America*, *106*(3 Pt 1), 1465–1479.

Bruns, P., Spence, C., & Röder, B. (2011). Tactile recalibration of auditory spatial representations. *Experimental Brain Research*, *209*(3), 333–344.

Budinger, E., Heil, P., Hess, A., & Scheich, H. (2006). Multisensory processing via early cortical stages: Connections of the primary auditory cortical field with other sensory systems. *Neuroscience*, *143*(4), 1065–1083.

Carvell, G. E., & Simons, D. J. (1990). Biometric analyses of vibrissal tactile discrimination in the rat. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *10*(8), 2638–2648.

Ceballo, S., Piwkowska, Z., Bourg, J., Daret, A., & Bathellier, B. (2019). Targeted Cortical Manipulation of Auditory Perception. *Neuron*, *104*(6), 1168–1179.e5.

Chang, S., Xu, J., Zheng, M., Keniston, L., Zhou, X., Zhang, J., & Yu, L. (2022). Integrating Visual Information into the Auditory Cortex Promotes Sound Discrimination through Choice-Related Multisensory Integration. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *42*(45), 8556–8568.

Chen, C., & Song, S. (2019). Differential cell-type dependent brain state modulations of sensory representations in the non-lemniscal mouse inferior colliculus. *Communications Biology*, *2*, 356.

Chen, T.-W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., Schreiter, E. R., Kerr, R. A., Orger, M. B., Jayaraman, V., Looger, L. L., Svoboda, K., & Kim, D. S. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature*, *499*(7458), 295–300.

Christensen, R. K., Lindén, H., Nakamura, M., & Barkat, T. R. (2019). White Noise Background Improves Tone Discrimination by Suppressing Cortical Tuning Curves. *Cell Reports*, *29*(7), 2041–2053.e4.

Clayton, K. K., Stecyk, K. S., Guo, A. A., Chambers, A. R., Chen, K., Hancock, K. E., & Polley, D. B. (2024). Sound elicits stereotyped facial movements that provide a sensitive index of hearing abilities in mice. *Current Biology : CB*, *34*(8). https://doi.org/10.1016/j.cub.2024.02.057

Collins, L., Francis, J., Emanuel, B., & McCormick, D. A. (2023). *Cholinergic and noradrenergic axonal activity contains a behavioral-state signal that is coordinated across the dorsal cortex*. https://doi.org/10.7554/eLife.81826

Cooper, M. H., & Young, P. A. (1976). Cortical projections to the inferior colliculus of the cat. *Experimental Neurology*, *51*(2), 488–502.

Crapse, T. B., & Sommer, M. A. (2008). Corollary discharge across the animal kingdom. *Nature Reviews. Neuroscience*, *9*(8), 587–600.

Crochet, S., & Petersen, C. C. H. (2006). Correlating whisker behavior with membrane potential in barrel

cortex of awake mice. *Nature Neuroscience*, *9*(5), 608–610.

de la Torre-Martinez, R., Ketzef, M., & Silberberg, G. (2023). Ongoing movement controls sensory integration in the dorsolateral striatum. *Nature Communications*, *14*(1), 1004.

Diederich, A., & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics*, *66*(8), 1388–1404.

Dragoi, V., Sharma, J., Miller, E. K., & Sur, M. (2002). Dynamics of neuronal sensitivity in visual cortex and local feature discrimination. *Nature Neuroscience*, *5*(9), 883–891.

Dragoi, V., Sharma, J., & Sur, M. (2000). Adaptation-induced plasticity of orientation tuning in adult visual cortex. *Neuron*, *28*(1), 287–298.

Efron, B., Ntelezos, A., Katz, Y., & Lampl, I. (2025). Detection and neural encoding of whisker-generated sounds in mice. *Current Biology : CB*, *35*(6), 1211–1226.e8.

Eggermann, E., Kremer, Y., Crochet, S., & Petersen, C. C. H. (2014). Cholinergic signals in mouse barrel cortex during active whisker sensing. *Cell Reports*, *9*(5), 1654–1660.

Eliades, S. J., & Wang, X. (2003). Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *Journal of Neurophysiology*, *89*(4). https://doi.org/10.1152/jn.00627.2002

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*(4). https://doi.org/10.1016/j.tics.2004.02.002

Ewall, G., Parkins, S., Lin, A., Jaoui, Y., & Lee, H.-K. (2021). Cortical and Subcortical Circuits for Cross-Modal Plasticity Induced by Loss of Vision. *Frontiers in Neural Circuits*, *15*, 665009.

Ferguson, K. A., & Cardin, J. A. (2020). Mechanisms underlying gain modulation in the cortex. *Nature Reviews. Neuroscience*, *21*(2), 80.

Finkel, E. A., Chang, Y.-T., Dasgupta, R., Lubin, E. E., Xu, D., Minamisawa, G., Chang, A. J., Cohen, J. Y., & O'Connor, D. H. (2024). Tactile processing in mouse cortex depends on action context. *Cell Reports*, *43*(4), 113991.

Flavell, S. W., Gogolla, N., Lovett-Barron, M., & Zelikowsky, M. (2022). The emergence and influence of internal states. *Neuron*, *110*(16). https://doi.org/10.1016/j.neuron.2022.04.030

Fu, K.-M. G., Johnston, T. A., Shah, A. S., Arnold, L., Smiley, J., Hackett, T. A., Garraghty, P. E., & Schroeder, C. E. (2003). Auditory cortical neurons respond to somatosensory stimulation. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *23*(20), 7510–7515.

Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, *10*(6), 278–285.

Gilad, A., & Helmchen, F. (2020). Spatiotemporal refinement of signal flow through association cortex during learning. *Nature Communications*, *11*(1), 1–14.

Gillmeister, H., & Eimer, M. (2007). Tactile enhancement of auditory detection and perceived loudness. *Brain Research*, *1160*, 58–68.

Godenzini, L., Alwis, D., Guzulaitis, R., Honnuraiah, S., Stuart, G. J., & Palmer, L. M. (2021). Auditory input enhances somatosensory encoding and tactile goal-directed behavior. *Nature Communications*, *12*(1), 4509.

Graziano, M. S. A., & Cooke, D. F. (2006). Parieto-frontal interactions, personal space, and defensive behavior. *Neuropsychologia*, *44*(6), 845–859.

Graziano, M. S., Reiss, L. A., & Gross, C. G. (1999). A neuronal representation of the location of nearby sounds. *Nature*, *397*(6718), 428–430.

Gruters, K. G., & Groh, J. M. (2012). Sounds and beyond: multisensory and other non-auditory signals in the inferior colliculus. *Frontiers in Neural Circuits*, *6*, 96.

Guillery, R. W., Feig, S. L., & Van Lieshout, D. P. (2001). Connections of higher order visual relays in the thalamus: a study of corticothalamic pathways in cats. *The Journal of Comparative Neurology*, *438*(1), 66–85.

Hammond-Kenny, A., Bajo, V. M., King, A. J., & Nodal, F. R. (2017). Behavioural benefits of multisensory

processing in ferrets. *The European Journal of Neuroscience*, *45*(2). https://doi.org/10.1111/ejn.13440

Hasegawa, M., Huang, Z., Paricio-Montesinos, R., & Gründemann, J. (2024). Network state changes in sensory thalamus represent learned outcomes. *Nature Communications*, *15*, 7830.

Huang, L., Hardyman, F., Edwards, M., & Galliano, E. (2024). Deprivation-Induced Plasticity in the Early Central Circuits of the Rodent Visual, Auditory, and Olfactory Systems. *eNeuro*, *11*(2), ENEURO.0435–23.2023.

Ibrahim, L. A., Mesik, L., Ji, X.-Y., Fang, Q., Li, H.-F., Li, Y.-T., Zingg, B., Zhang, L. I., & Tao, H. W. (2016). Cross-Modality Sharpening of Visual Cortical Processing through Layer-1-Mediated Inhibition and Disinhibition. *Neuron*, *89*(5), 1031–1045.

Iurilli, G., Ghezzi, D., Olcese, U., Lassi, G., Nazzaro, C., Tonini, R., Tucci, V., Benfenati, F., & Medini, P. (2012). Sound-driven synaptic inhibition in primary visual cortex. *Neuron*, *73*(4), 814–828.

Kato, D. D., & Bruno, R. M. (2025). Stability of cross-sensory input to primary somatosensory cortex across experience. *Neuron*, *113*(2), 291–306.e7.

Kato, H. K., Gillet, S. N., & Isaacson, J. S. (2015). Flexible sensory representations in auditory cortex driven by behavioral relevance. *Neuron*, *88*(5), 1027.

Kayser, C., Logothetis, N. K., & Panzeri, S. (2010). Visual enhancement of the information representation in auditory cortex. *Current Biology: CB*, *20*(1), 19–24.

Kayser, C., Petkov, C. I., Augath, M., & Logothetis, N. K. (2005). Integration of touch and sound in auditory cortex. *Neuron*, *48*(2), 373–384.

Kerlin, A. M., Andermann, M. L., Berezovskii, V. K., & Reid, R. C. (2010). Broadly tuned response properties of diverse inhibitory neuron subtypes in mouse visual cortex. *Neuron*, *67*(5). https://doi.org/10.1016/j.neuron.2010.08.002

Khorevin, V. I. (1980). Effect of electrodermal stimulation on single unit responses to acoustic stimulation in the parvocellular part of the medial geniculate body. *Neurophysiology*, *12*(2), 129–134.

Khorevin, V. I., & Shelest, I. I. (1998). Interaction of influences from the auditory and somatosensory afferent systems on neurons of the primary auditory cortex. *Neurophysiology*, *30*(6), 383–385.

Kimura, A., & Imbe, H. (2018). Robust Subthreshold Cross-modal Modulation of Auditory Response by Cutaneous Electrical Stimulation in First- and Higher-order Auditory Thalamic Nuclei. *Neuroscience*, *372*, 161–180.

Kleinfeld, D., & Deschênes, M. (2011). Neuronal basis for object location in the vibrissa scanning sensorimotor system. *Neuron*, *72*(3), 455–468.

Kohn, A. (2007). Visual adaptation: physiology, mechanisms, and functional benefits. *Journal of Neurophysiology*, *97*(5), 3155–3164.

Kolarik, A. J., Moore, B. C. J., Zahorik, P., Cirstea, S., & Pardhan, S. (2016). Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss. *Attention, Perception & Psychophysics*, *78*(2), 373–395.

Kuchibhotla, K. V., Gill, J. V., Lindsay, G. W., Papadoyannis, E. S., Field, R. E., Sten, T. A., Miller, K. D., & Froemke, R. C. (2017). Parallel processing by cortical inhibition enables context-dependent behavior. *Nature Neuroscience*, *20*(1). https://doi.org/10.1038/nn.4436

Lakatos, P., Chen, C.-M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron*, *53*(2), 279–292.

Lee, S.-H., Kwan, A. C., Zhang, S., Phoumthipphavong, V., Flannery, J. G., Masmanidis, S. C., Taniguchi, H., Huang, Z. J., Zhang, F., Boyden, E. S., Deisseroth, K., & Dan, Y. (2012). Activation of specific interneurons improves V1 feature selectivity and visual perception. *Nature*, *488*(7411), 379–383.

Lee, T. Y., Weissenberger, Y., King, A. J., & Dahmen, J. C. (2024). Midbrain encodes sound detection behavior without auditory cortex. *eLife*, *12*. https://doi.org/10.7554/eLife.89950

Lesicko, A. M. H., Hristova, T. S., Maigler, K. C., & Llano, D. A. (2016). Connectional modularity of top-down and bottom-up multimodal inputs to the lateral cortex of the mouse inferior colliculus. *The*

Journal of Neuroscience: The Official Journal of the Society for Neuroscience, *36*(43), 11037–11050.

Lohse, M., Dahmen, J. C., Bajo, V. M., & King, A. J. (2021). Subcortical circuits mediate communication between primary sensory cortical areas in mice. *Nature Communications*, *12*(1), 3916.

Lohse, M., King, A. J., & Willmore, B. D. B. (2024). Subcortical origin of nonlinear sound encoding in auditory cortex. *Current Biology : CB*, *34*(15). https://doi.org/10.1016/j.cub.2024.06.057

Lohse, M., Zimmer-Harwood, P., Dahmen, J. C., & King, A. J. (2022). Integration of somatosensory and motor-related information in the auditory system. *Frontiers in Neuroscience*, *16*, 1010211.

Martín-Cortecero, J., Isaías-Camacho, E. U., Boztepe, B., Ziegler, K., Mease, R. A., & Groh, A. (2023). Monosynaptic trans-collicular pathways link mouse whisker circuits to integrate somatosensory and motor cortical signals. *PLoS Biology*, *21*(5). https://doi.org/10.1371/journal.pbio.3002126

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, *21*(9), 1281–1289.

McCormick, D. A., Nestvogel, D. B., & He, B. J. (2020). Neuromodulation of Brain State and Behavior. *Annual Review of Neuroscience*, *43*. https://doi.org/10.1146/annurev-neuro-100219-105424

McGinley, M. J., David, S. V., & McCormick, D. A. (2015). Cortical membrane potential signature of optimal states for sensory signal detection. *Neuron*, *87*(1), 179–192.

Meredith, M. A., & Allman, B. L. (2015). Single-unit analysis of somatosensory processing in the core auditory cortex of hearing ferrets. *The European Journal of Neuroscience*, *41*(5), 686–698.

Mianné, J., Chessum, L., Kumar, S., Aguilar, C., Codner, G., Hutchison, M., Parker, A., Mallon, A.-M., Wells, S., Simon, M. M., Teboul, L., Brown, S. D. M., & Bowl, M. R. (2016). Correction of the auditory phenotype in C57BL/6N mice via CRISPR/Cas9-mediated homology directed repair. *Genome Medicine*, *8*(1), 16.

Miyashita, T., & Feldman, D. E. (2013). Behavioral detection of passive whisker stimuli requires

somatosensory cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, *23*(7), 1655–1662.

Mo, C., & Murray Sherman, S. (2019). A Sensorimotor Pathway via Higher-Order Thalamus. *Journal of Neuroscience*, *39*(4), 692–704.

Moore, D. R., & King, A. J. (1999). Auditory perception: The near and far of sound localization. *Current Biology: CB*, *9*(10), R361–R363.

Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S., & Churchland, A. K. (2019). Single-trial neural dynamics are dominated by richly varied movements. *Nature Neuroscience*, *22*(10), 1677.

Nelson, A., & Mooney, R. (2016). The Basal Forebrain and Motor Cortex Provide Convergent yet Distinct Movement-Related Inputs to the Auditory Cortex. *Neuron*, *90*(3). https://doi.org/10.1016/j.neuron.2016.03.031

Occelli, V., Spence, C., & Zampini, M. (2011). Audiotactile interactions in front and rear space. *Neuroscience and Biobehavioral Reviews*, *35*(3). https://doi.org/10.1016/j.neubiorev.2010.07.004

Ohshiro, T., Angelaki, D. E., & DeAngelis, G. C. (2011). A normalization model of multisensory integration. *Nature Neuroscience*, *14*(6), 775–782.

Ohshiro, T., Angelaki, D. E., & DeAngelis, G. C. (2017). A Neural Signature of Divisive Normalization at the Level of Multisensory Integration in Primate Cortex. *Neuron*, *95*(2), 399–411.e8.

Olthof, B. M. J., Rees, A., & Gartside, S. E. (2019). Multiple Nonauditory Cortical Regions Innervate the Auditory Midbrain. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *39*(45), 8916–8928.

Opoku-Baah, C., Schoenhaut, A. M., Vassall, S. G., Tovar, D. A., Ramachandran, R., & Wallace, M. T. (2021). Visual influences on auditory behavioral, neural, and perceptual processes: A review. *Journal of the Association for Research in Otolaryngology: JARO*, *22*(4), 365–386.

O'Sullivan, C., Weible, A. P., & Wehr, M. (2019). Auditory Cortex Contributes to Discrimination of Pure Tones. *eNeuro*, *6*(5), ENEURO.0340–19.2019.

Otazu, G. H., Tai, L. H., Yang, Y., & Zador, A. M. (2009). Engaging in an auditory task suppresses responses in auditory cortex. *Nature Neuroscience*, *12*(5). https://doi.org/10.1038/nn.2306

Pachitariu, M., Stringer, C., Dipoppa, M., Schröder, S., Federico Rossi, L., Dalgleish, H., Carandini, M., & Harris, K. D. (2017). Suite2p: beyond 10,000 neurons with standard two-photon microscopy. In *bioRxiv* (p. 061507). https://doi.org/10.1101/061507

Petty, G. H., & Bruno, R. M. (2024). Attentional modulation of secondary somatosensory and visual thalamus of mice. *eLife*, *13*. https://doi.org/10.7554/eLife.97188.2

Podachin, V. P., Naumova, T. S., Babichenko, I. I., & Kartseva, O. M. (1979). [Interaction between auditory and trigeminal afferent volleys in the dorsal cochlear nucleus of rats]. *Neirofiziologiia = Neurophysiology*, *11*(2), 179–182.

Rabinowitz, N. C., Willmore, B. D. B., King, A. J., & Schnupp, J. W. H. (2013). Constructing noise-invariant representations of sound in the auditory pathway. *PLoS Biology*, *11*(11), e1001710.

Rao, R. P., Mielke, F., Bobrov, E., & Brecht, M. (2014). Vocalization-whisking coordination and multisensory integration of social signals in rat auditory cortex. *eLife*, *3*, e03185.

Reimer, J., Froudarakis, E., Cadwell, C. R., Yatsenko, D., Denfield, G. H., & Tolias, A. S. (2014). Pupil fluctuations track fast switching of cortical states during quiet wakefulness. *Neuron*, *84*(2), 355–362.

Reimer, J., McGinley, M. J., Liu, Y., Rodenkirch, C., Wang, Q., McCormick, D. A., & Tolias, A. S. (2016). Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nature Communications*, *7*, 13289.

Robards, M. J. (1979). Somatic neurons in the brainstem and neocortex projecting to the external nucleus of the inferior colliculus: an anatomical study in the opossum. *The Journal of Comparative Neurology*, *184*(3), 547–565.

Ro, T., Hsu, J., Yasar, N. E., Elmore, L. C., & Beauchamp, M. S. (2009). Sound enhances touch perception. *Experimental Brain Research*, *195*(1), 135–143.

Rowland, B. A., Quessy, S., Stanford, T. R., & Stein, B. E. (2007). Multisensory integration shortens physiological response latencies. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *27*(22), 5879–5884.

Sara, S. J., & Bouret, S. (2012). Orienting and reorienting: the locus coeruleus mediates cognition through arousal. *Neuron*, *76*(1). https://doi.org/10.1016/j.neuron.2012.09.011

Schneider, D. M., Nelson, A., & Mooney, R. (2014). A synaptic and circuit basis for corollary discharge in the auditory cortex. *Nature*, *513*(7517), 189–194.

Schneider, D. M., Sundararajan, J., & Mooney, R. (2018). A cortical filter that learns to suppress the acoustic consequences of movement. *Nature*, *561*(7723), 391–395.

Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S., & Puce, A. (2008). Neuronal oscillations and visual amplification of speech. *Trends in Cognitive Sciences*, *12*(3). https://doi.org/10.1016/j.tics.2008.01.002

Schulz, M., Ross, B., & Pantev, C. (2003). Evidence for training-induced crossmodal reorganization of cortical functions in trumpet players. *Neuroreport*, *14*(1), 157–161.

Schürmann, M., Caetano, G., Jousmäki, V., & Hari, R. (2004). Hands help hearing: facilitatory audiotactile interaction at low sound-intensity levels. *The Journal of the Acoustical Society of America*, *115*(2). https://doi.org/10.1121/1.1639909

Seifritz, E., Neuhoff, J. G., Bilecen, D., Scheffler, K., Mustovic, H., Schächinger, H., Elefante, R., & Di Salle, F. (2002). Neural processing of auditory looming in the human brain. *Current Biology: CB*, *12*(24), 2147–2151.

Serino, A. (2019). Peripersonal space (PPS) as a multisensory interface between the individual and the environment, defining the space of the self. *Neuroscience and Biobehavioral Reviews*, *99*, 138–159.

Sherman, S. M., & Usrey, W. M. (2024). Transthalamic Pathways for Cortical Function. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *44*(35). https://doi.org/10.1523/JNEUROSCI.0909-24.2024

Shore, S. E., & Zhou, J. (2006). Somatosensory influence on the cochlear nucleus and beyond. *Hearing Research*, *216-217*, 90–99.

Singla, S., Dempsey, C., Warren, R., Enikolopov, A. G., & Sawtell, N. B. (2017). A cerebellum-like circuit in the auditory system cancels responses to self-generated sounds. *Nature Neuroscience*, *20*(7), 943–950.

Sofroniew, N. J., & Svoboda, K. (2015). Whisking. *Current Biology : CB*, *25*(4). https://doi.org/10.1016/j.cub.2015.01.008

Soto-Faraco, S., & Deco, G. (2009). Multisensory contributions to the perception of vibrotactile events. *Behavioural Brain Research*, *196*(2). https://doi.org/10.1016/j.bbr.2008.09.018

Soto-Faraco, S., Spence, C., & Kingstone, A. (2004). Congruency effects between auditory and tactile motion: extending the phenomenon of cross-modal dynamic capture. *Cognitive, Affective & Behavioral Neuroscience*, *4*(2), 208–217.

Sperdin, H. F., Cappe, C., Foxe, J. J., & Murray, M. M. (2009). Early, low-level auditory-somatosensory multisensory interactions impact reaction time speed. *Frontiers in Integrative Neuroscience*, *3*, 2.

Sreenivasan, V., Esmaeili, V., Kiritani, T., Galan, K., Crochet, S., & Petersen, C. C. H. (2016). Movement Initiation Signals in Mouse Whisker Motor Cortex. *Neuron*, *92*(6), 1368.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society, Inc*, *31*(1), 137–149.

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews. Neuroscience*, *9*(4), 255–266.

Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C. B., Carandini, M., & Harris, K. D. (2019). Spontaneous behaviors drive multidimensional, brainwide activity. *Science (New York, N.Y.)*, *364*(6437), 255.

Suga, N., Gao, E., Zhang, Y., Ma, X., & Olsen, J. F. (2000). The corticofugal system for hearing: recent progress. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(22),

11807–11814.

Suga, N., Xiao, Z., Ma, X., & Ji, W. (2002). Plasticity and corticofugal modulation for hearing in adult animals. *Neuron*, *36*(1), 9–18.

Tang, J., Yang, W., & Suga, N. (2012). Modulation of thalamic auditory neurons by the primary auditory cortex. *Journal of Neurophysiology*, *108*(3), 935–942.

Teneggi, C., Canzoneri, E., di Pellegrino, G., & Serino, A. (2013). Social modulation of peripersonal space boundaries. *Current Biology: CB*, *23*(5), 406–411.

Town, S. M., Wood, K. C., & Bizley, J. K. (2018). Sound identity is represented robustly in auditory cortex during perceptual constancy. *Nature Communications*, *9*(1), 1–15.

Von Bekesy, G. (1959). Similarities between hearing and skin sensations. *Psychological Review*, *66*(1), 1–22.

Watson, C., Paxinos, G., & Puelles, L. (2011). *The Mouse Nervous System*. Academic Press.

Williams, A. M., Angeloni, C. F., & Geffen, M. N. (2023). Sound Improves Neuronal Encoding of Visual Stimuli in Mouse Primary Visual Cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *43*(16), 2885–2906.

Williamson, R. S., Hancock, K. E., Shinn-Cunningham, B. G., & Polley, D. B. (2015). Locomotion and Task Demands Differentially Modulate Thalamic Audiovisual Processing during Active Search. *Current Biology : CB*, *25*(14), 1885–1891.

Wilson, E. C., Reed, C. M., & Braida, L. D. (2010). Integration of auditory and vibrotactile stimuli: effects of frequency. *The Journal of the Acoustical Society of America*, *127*(5), 3044–3059.

Winer, J. A., & Schreiner, C. E. (2005). *The Inferior Colliculus*. Springer Science & Business Media.

Wu, C., Stefanescu, R. A., Martel, D. T., & Shore, S. E. (2015). Listening to another sense: somatosensory integration in the auditory system. *Cell and Tissue Research*, *361*(1), 233–250.

Yang, Y., Lee, J., & Kim, G. (2020). Integration of locomotion and auditory signals in the mouse inferior colliculus. *eLife*, *9*. https://doi.org/10.7554/eLife.52228

Youden, W. Y. (1950). Index for rating diagnostic tests. *Cancer*, *3*(1). https://doi.org/10.1002/1097-0142(1950)3:1<32::aid-cncr2820030106>3.0.co;2-3

Yu, L., Cuppini, C., Xu, J., Rowland, B. A., & Stein, B. E. (2019). Cross-Modal Competition: The Default Computation for Multisensory Processing. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *39*(8), 1374–1385.

Zagha, E., Erlich, J. C., Lee, S., Lur, G., O'Connor, D. H., Steinmetz, N. A., Stringer, C., & Yang, H. (2022). The Importance of Accounting for Movement When Relating Neuronal Activity to Sensory and Cognitive Processes. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *42*(8). https://doi.org/10.1523/JNEUROSCI.1919-21.2021

Zhang, M., Kwon, S. E., Ben-Johny, M., O'Connor, D. H., & Issa, J. B. (2020). Spectral hallmark of auditory-tactile interactions in the mouse somatosensory cortex. *Communications Biology*, *3*(1), 64.

Zheng, A., & Schmid, S. (2023). A review of the neural basis underlying the acoustic startle response with a focus on recent developments in mammals. *Neuroscience and Biobehavioral Reviews*, *148*, 105129.

VII Appendix

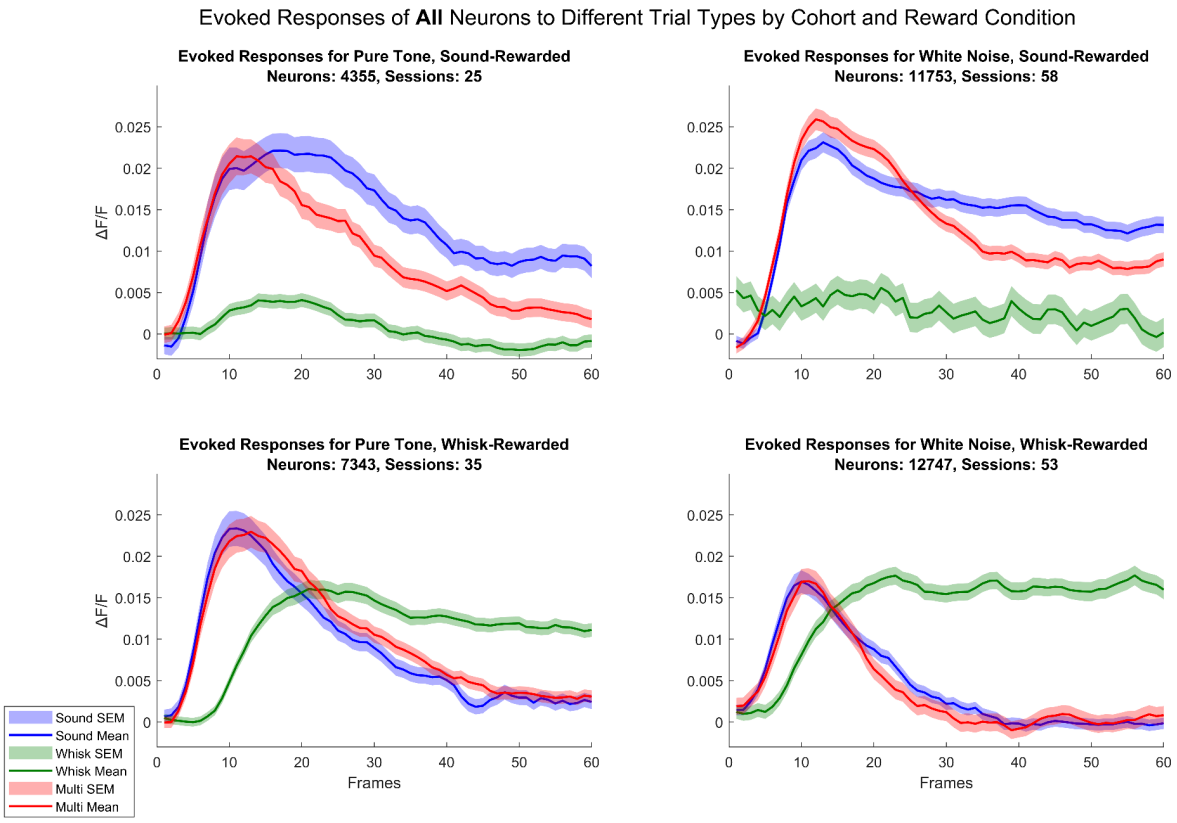

Figure S1. Evoked Responses for All Neurons are Dominated by Reward Related Activity. Evoked responses of all neurons to different trial types by cohort and reward condition. This figure shows the average evoked ΔF/F responses of all detected neurons (classified by the suite2p custom classifier) across both rewarded and missed trials in sessions with either sound-rewarded or whisker-rewarded conditions. Each plot represents a specific combination of sound stimulus type (pure tone or broadband noise) and reward condition. Responses to sound-only, whisker-only, and multi-sensory (sound + whisker) trials are shown in blue, green, and red lines, respectively, with shaded areas indicating the standard error of the mean (SEM). In sound-rewarded sessions (top row), sustained neural responses are observed for sound trials (blue line) after stimulus onset, suggesting persistent activation linked to the rewarded sound stimulus. In whisker-rewarded sessions (bottom row), sustained

responses are observed for whisker trials (green line), albeit with a slight delay. This sustained activity likely reflects neural processes associated with reward receipt, licking-related motor activity (Musall et al., 2019), and self-generated sounds.

7.1 Reaction Latency Differences Across Task Conditions

In chapter 1, we found that sound detection accuracy was unchanged during multi-trials (sound plus whisker stimulation) but that reaction times were significantly faster. To explore whether this effect depends on task context, we next compared detection latencies in tone- versus multi-trials under both tone-rewarded and whisker-rewarded conditions.

In noise-rewarded sessions, the trend is reversed. Mice exhibited slightly faster reaction times in whisker-only trials compared to multi-trials (Figure S21, bottom left panel). The combined analysis (bottom right panel) shows a significant difference ($p = 0.042$, Mann-Whitney U test), with an average difference of 2.88 ms favouring whisker-only trials.

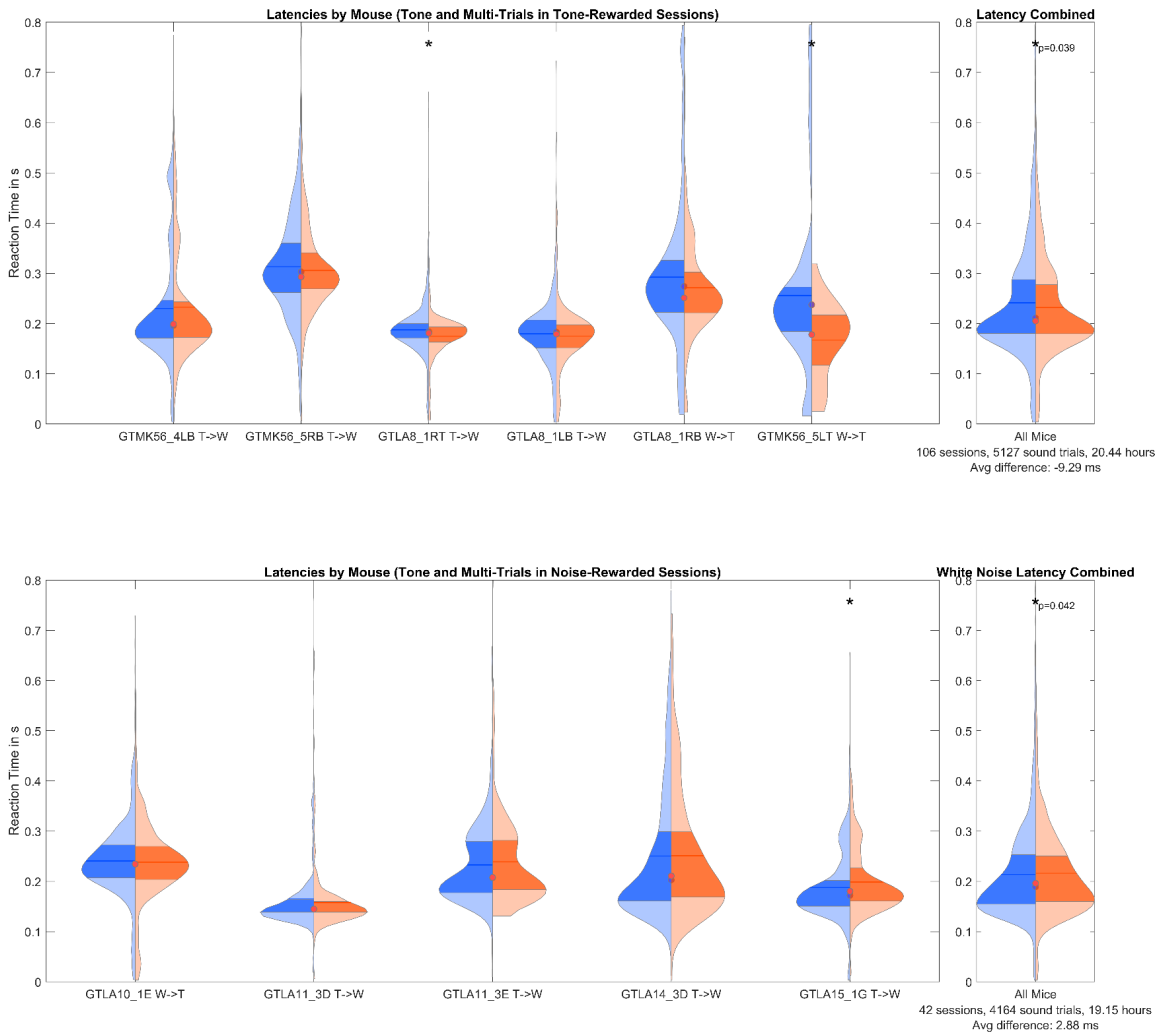

Figure S2: Detection Latencies for Tone and Multi-Trials Across Tone-Rewarded and Noise-Rewarded Sessions. Violin plots show reaction times for individual mice during tone (blue) or multi-trials (red). In tone-rewarded sessions (top panel), mice responded significantly faster in multi-trials compared to tone trials (combined analysis: p = 0.039, average difference of -9.29 ms). In noise-rewarded sessions (bottom panel), the trend reversed, with faster whisker detection in whisker-only trials compared to multi-trials (combined analysis: p = 0.042, average difference of 2.88 ms). Statistical significance was assessed using the Mann-Whitney U test, with significant differences marked by an asterisk (*). The combined panels on the far right summarize data across all sessions.

When we grouped mice by their initial reward condition (sound-rewarded vs. whisker-rewarded; Figure S3), those starting on sound-rewarded sessions showed no significant latency difference between sound and multi-trials (0.49 ms difference, $p > 0.05$). In contrast, the smaller group that began with whisker-rewarded sessions ($n = 3$) displayed overall slower latencies, but a significant −15.96 ms difference favouring multi-trials over sound-only trials ($p < 0.05$). Thus, while starting on sound-rewarded tasks may confer a long-term advantage in sound detection, initial reward condition alone does not fully account for the observed multi-trial latency differences.

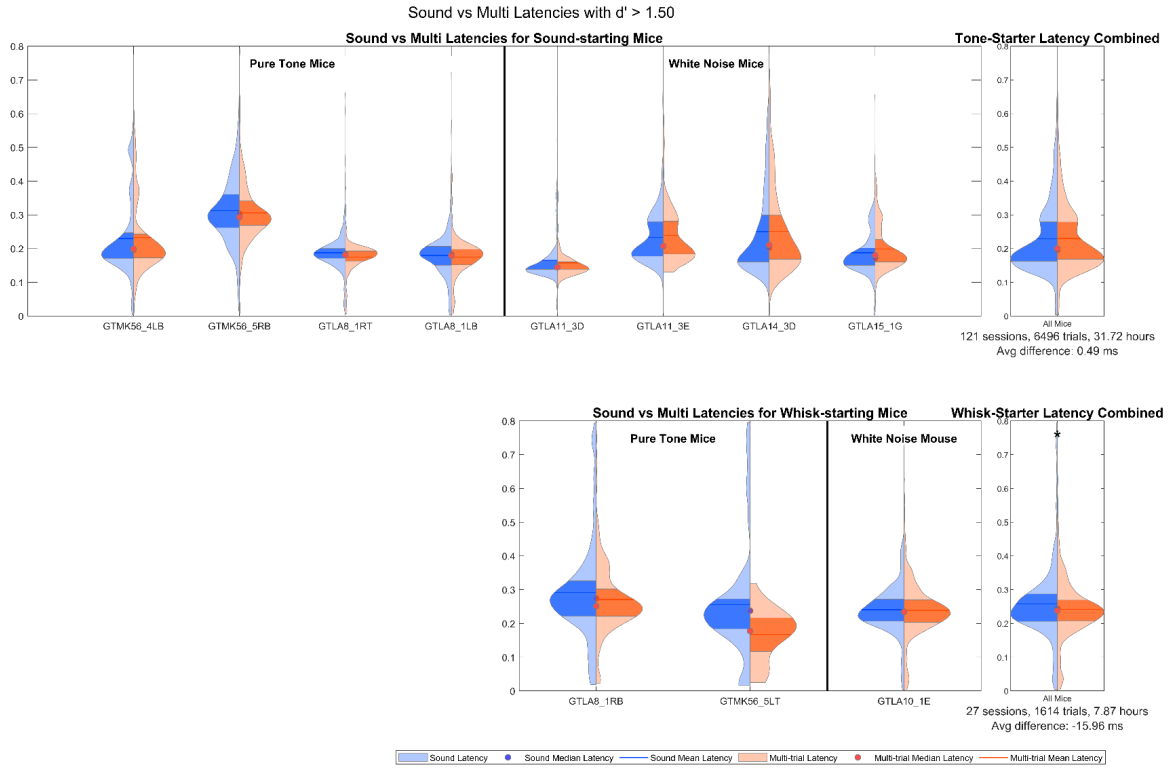

Figure S3: Reaction Latencies for Sound-only and Multi-Trials Based on Starting Reward Condition. Violin plots compare reaction times for sound trials (blue) and multi-trials (red) across mice grouped by their initial training condition. Top Row: Mice that started on sound-rewarded trials show no differences between sound and multi-trial latencies. The combined analysis (top right panel) indicates no significant difference ($p > 0.05$). Bottom Row: Whisker-starting mice exhibit slower latencies overall but show a significant latency difference in favour of multi-trials ($p < 0.05$). The combined analysis (far-right panel) reveals an average difference of -15.96 ms.

Finally, we examined whisker-rewarded sessions to see if whisker-only and multi-trials differed in detection speed (Figure S4). Despite the limited number of mice per condition, both sound-starting and whisker-starting mice were faster on multi-trials than whisker-only trials ($p < 0.05$ in both groups).

Overall, sound-starting mice tended to have shorter sound-detection latencies, whereas whisker-starting mice were generally faster at detecting whisker stimuli—indicating that early training may bias how each animal prioritizes subsequent sensory inputs.

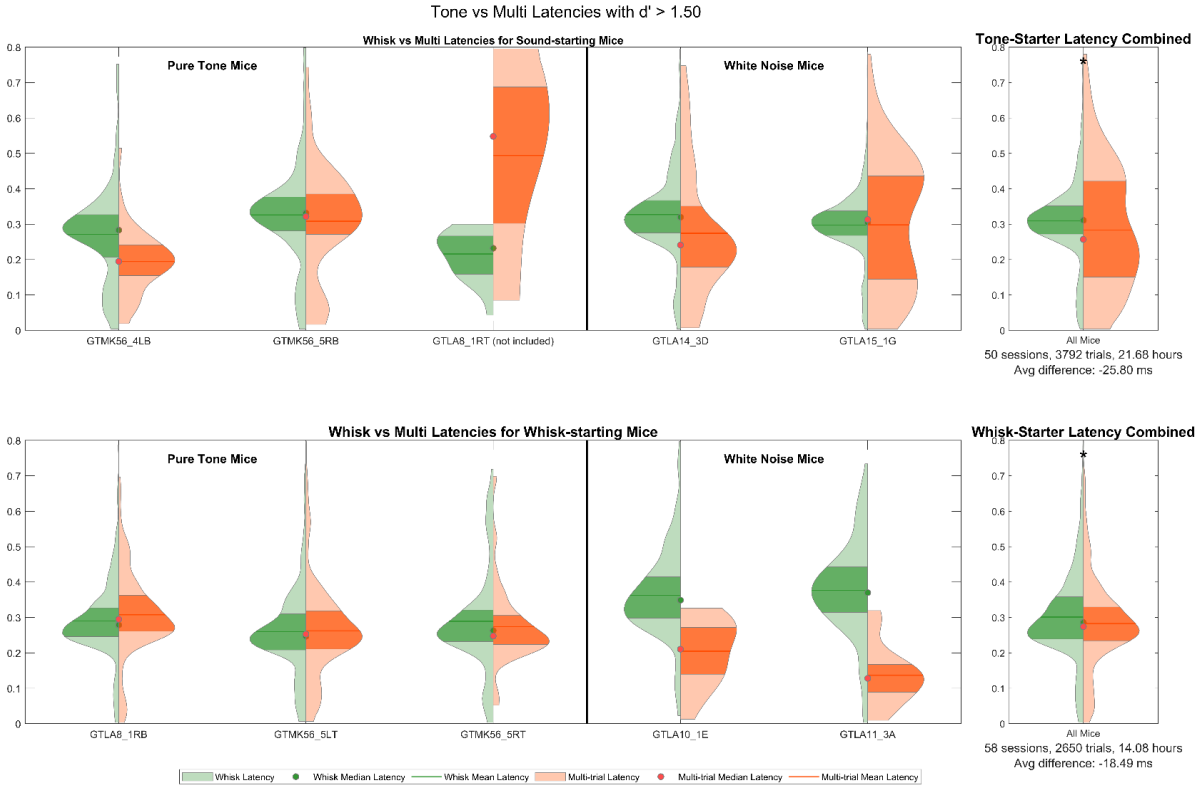

Figure S4: Reaction Latencies for Whisker-only and Multi-Trials in Whisker-Rewarded Sessions. Violin plots compare reaction times for whisker trials (green) and multi-trials (orange) across mice grouped by their initial training condition. Top Row: Mice that began their training on sound-rewarded trials show faster reaction times for multi-trials compared to whisker-only trials, which combines to a significant latency difference ($p < 0.05$, difference = 25.80 ms) (top right panel). Bottom Row: Whisker-starting mice also exhibit faster reaction times for multi-trials compared to whisker-only trials, with a combined latency difference of -18.5ms ($p < 0.05$).