



**HAL**  
open science

# Attentional modulation of the auditory cortex activity in a sound discrimination task in mice

Clara Bardin

► **To cite this version:**

Clara Bardin. Attentional modulation of the auditory cortex activity in a sound discrimination task in mice. Neuroscience. Université Paris sciences et lettres, 2021. English. NNT : 2021UPSLE057 . tel-04008836

**HAL Id: tel-04008836**

**<https://theses.hal.science/tel-04008836>**

Submitted on 28 Feb 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**THÈSE DE DOCTORAT**

**DE L'UNIVERSITÉ PSL**

Préparée à l'École Normale Supérieure

**Modulation attentionnelle de l'activité du cortex auditif dans  
une tâche de discrimination sonore chez la souris**

**Attentional modulation of the auditory cortex activity in a sound  
discrimination task in mice**

Soutenue par

**Clara BARDIN**

Le 24 septembre 2021

École doctorale n° 158

**Cerveau, cognition,  
comportement (ED3C)**

Spécialités

**Les neurosciences et les  
sciences cognitives**



**Composition du jury :**

Valérie EGO-STENGEL Professeur, <i>Institut des Neurosciences de Paris-Saclay</i>	<i>Président du jury</i>
Kerry WALKER Associate Professor, <i>University of Oxford</i>	<i>Rapporteur</i>
Timo VAN KERKOERLE Directeur de Recherche, <i>Institut des Neurosciences de Paris-Saclay</i>	<i>Rapporteur</i>
Laurent BOURDIEU Directeur de Recherche, <i>École Normale Supérieure</i>	<i>Examineur</i>
Pierre BOURDILLON Praticien Hospitalier, <i>Hôpital Fondation Adolphe de Rothschild</i>	<i>Examineur</i>
Claire SERGENT Professeur, <i>Université de Paris</i>	<i>Examineur</i>
Jean-François LÉGER Directeur de Recherche, <i>École Normale Supérieure</i>	<i>Directeur de thèse</i>
Yves BOUBENEC Associate Professor, <i>École Normale Supérieure</i>	<i>Directeur de thèse</i>

## RÉSUMÉ

---

L'attention sélective auditive est une capacité cognitive permettant d'adapter le traitement des stimuli en fonction de leur pertinence dans un contexte donné. La compréhension de la dynamique neuronale sous-jacente à l'attention sélective auditive reste incomplète, diverse et même controversée dans la littérature. Cette étude vise à contribuer à la connaissance en comparant l'activité auditive corticale évoquée de sons similaires dans le cortex auditif primaire de souris (A1) entre un contexte d'écoute active et un contexte éveillé mais passif.

Pour l'étude, une tâche d'attention sélective Go/no-Go a été développée et adaptée aux souris, en alternant 4 blocs de discrimination sonore active avec 4 blocs d'exposition sonore passive. Grâce à une quantification exhaustive des performances dans les deux contextes, nous avons pu mettre en évidence une bonne compréhension du sens des stimulations sonores par les souris selon le contexte. Dans le contexte actif, les souris répondent presque exclusivement à la tonalité cible [avec un taux de réussite de 66 % et un temps de réaction moyen de 500 ms]. Dans le contexte passif, ils restent neutres aux sons présentés [dans 92% des essais passifs].

Une quantification millimétrique des mouvements des parties du corps de l'animal a permis de déterminer que l'état comportemental global de l'animal est modulé en fonction du contexte et du moment de l'essai. Ces mouvements (frétillements observables à l'œil) sont corrélés à la charge attentionnelle requise et varient dans cet ordre : les blocs passifs, l'instant précédant l'essai dans les blocs actifs et l'écoute attentive des sons de référence pendant le contexte actif.

Parmi les différences neuronales entre les contextes actif et passif, une suppression de l'activité de A1 est observée pendant le contexte actif pour l'activité spontanée et pour le traitement des sons de référence et cibles. Cependant, en plus de cette suppression, l'étude observe une concentration de l'activité dans un sous-groupe de neurones dont l'activité neuronale excède largement l'activité moyenne de la population. Ce mécanisme de concentration de l'activité est plus important dans le contexte actif comparé au contexte passif.

Les mécanismes de suppression et d'activation précise neuronale sont tous deux plus importants dans le contexte comportemental que dans le contexte passif. Ensemble, ces mécanismes conduisent à concentrer l'activité résiduelle dans un petit sous-groupe de neurones pendant que les souris sont engagées. Ces mécanismes coexistants pourraient prendre part dans plusieurs aspects de l'attention auditive pour des actions comportementales adaptées.

## MOTS CLÉS

---

Attention sélective, Comportement, Cortex Auditif, Microscopie à deux-photons



## ABSTRACT

---

Auditory selective attention is a cognitive ability enabling to adapt the processing of stimuli according to their relevance in a given context. The understanding of the neural dynamics underlying auditory selective attention remains incomplete, diverse and even controversial in the literature. This study aims to contribute to knowledge by comparing the evoked-cortical auditory activity of similar sounds in the primary auditory cortex of mice (A1) between an active listening context and an awake but passive context.

For the study, a Go / no-Go selective attention task was developed and adapted to mice, alternating 4 blocks of active sound discrimination with 4 blocks of passive sound exposure. Through an exhaustive quantification of the performances in both contexts, we were able to evidence a good understanding of the meaning of the sound stimulations by the mice according to the context. In the active context, they respond almost exclusively to the target tone [with a success rate of 66% and an average reaction time of 500 ms]. In the passive context, they remain neutral to the presented sounds [in 92% of passive trials].

A millimetric quantification of the movements of the animal's body parts enabled the study to determine that the overall behavioral state of the animal is modulated according to the context and the moment in the trial. These movements (wiggles observable by eye) are correlated with the required attentional load and vary in this order: the passive blocks, the moment preceding the trial in the active blocks and the attentive listening of the reference sounds during the active context.

While performing the attentional sound discrimination task, the study recorded the activity of the layer II-III of A1 with a 2-photon microscope. The comparison of the neuronal activity of A1 between the active blocks - where the mice actively recognize the target tone, and the passive blocks - where they remain passively exposed to the same sounds, made it possible to highlight the difference in the processing of sound information as early as from this first step of cortical treatment. These differences disappear when the mouse disengages in the task.

Among the neuronal differences between active and passive contexts, a suppression of A1's activity is observed during the active context for spontaneous activity and for the processing of the reference and target sounds. However, in addition to this suppression, the study observed a concentration of activity in a subset of neurons whose neuronal activity greatly exceeds the average activity of the population. This mechanism of concentration of activity is more important in the active context than in the passive context.

The suppression and sparse neuronal boosting mechanisms are both greater in the behaving context than in the passive one. Together, these mechanisms result in concentrating the residual activity in a small subgroup of neurons while mice are engaged. These coexisting mechanisms could underly several aspects of auditory attention for adapted behavioral actions.

## KEYWORDS

---

Selective attention, Behavior, Auditory Cortex, Two-photon microscopy



## Table of content

### List of Abbreviations

### List of Figures

### List of Annexes

## 1. Introduction

### 1.1. The concept of attention

#### 1.1.1. What is “attention”

#### 1.1.2. The importance of selective attention in education and attention deficit disorders

#### 1.1.3. Underlying mechanisms of active behavioural states

### 1.2. From a sound to its contextual meaning

#### 1.2.1. Overview of the auditory pathway and cortical processing

#### 1.2.2. Selective attention studies in the auditory cortex

#### 1.2.3. Different experimental paradigms lead to different conclusions

## 2. Experimental procedures

## 3. Attentional Go / no-Go auditory discrimination task

### 3.1. Shaping mouse behavior for an attentional Go / no-Go auditory discrimination task

### 3.2. Behavioral modulation according to the context

#### 3.2.1. Accurate performance according to the context

#### 3.2.2. Fast adaptation to context switching

#### 3.2.3. High engagement in the active context

#### 3.2.4. Behavioral state modulated by the context

### 3.3. Performance drop when mice are not motivated

### 3.4. Attentional Go / no-Go auditory discrimination task altering randomly two target frequencies

#### 3.4.1. Fast learning of a second target frequency

#### 3.4.2. Equivalent target discrimination

## 4. Imaging of the sensory representation of A1 during a context-changing behavioral task in mice

### 4.1. Passive pure tone sensory representations in A1 awake mice

### 4.2. Task engagement modulates sound representation in A1 between blocks of active listening and passive exposure

#### 4.2.1. Variations of evoked-cellular activity in A1 depend on task-engagement

#### 4.2.2. Active listening suppresses cellular activities in active trials compared to passive exposure

#### 4.2.3. A greater number of cells contrasts with the population in the active context

## 5. Discussion

## 6. Bibliography

## List of Abbreviations

<b>A</b>	active
<b>AC</b>	auditory cortex
<b>ACh</b>	acetylcholine
<b>ADHD</b>	attention deficit hyper activity disorders
<b>AFF</b>	anterior auditory field
<b>AUC</b>	area under the curve
<b>AVV</b>	Adeno-associated virus
<b>A1</b>	primary auditory cortex
<b>A2</b>	secondary auditory field and
<b>EEG</b>	electroencephalogram
<b>ERP</b>	event related potential
<b>FEF</b>	frontal eye field
<b>FOV</b>	field of view
<b>FR</b>	firing rate
<b>fMRI</b>	functional magnetic resonance imaging scanner
<b>LFP</b>	local-field potential
<b>mPFC</b>	medial prefrontal cortex
<b>mAChRs</b>	muscarinic acetylcholine receptor
<b>ms</b>	millisecond
<b>MGN</b>	medial geniculate nucleus
<b>NAc</b>	nucleus accubens
<b>NMDA</b>	N-methyl-D-aspartate receptor
<b>NREM</b>	non-rapid eye movement sleep
<b>P</b>	Passive
<b>PV cell</b>	parvalbumine inhibitory neuron
<b>REM</b>	rapid eye movement
<b>RF</b>	receptive field
<b>ROC curve</b>	receiver operating characteristic curve

<b>SOM cell</b>	somatostatin-expressing inhibitory neuron
<b>TORC</b>	temporally orthogonal ripple combination
<b>VIP cell</b>	vasoactive intestinal peptide inhibitory neuron
<b>5-CSRT</b>	5-choice serial reaction-time test
<b>5-HT</b>	serotonin

## List of Figures

**Figure 1.** Snapshot of a mouse performing the Go / no-Go task while A1 is being recorded

**Figure 2.** Signal extraction from a cell

**Figure 3.** Quantification of the responses for each stage of the learning protocol

**Figure 4.** Active and passive performances during the attentional auditory Go / No-Go task

**Figure 5.** Context change performances indicate fast adaptation to context switching

**Figure 6.** High target reactivity in the active context

**Figure 7.** Separable pressing paw movements between contexts indicate behavioral differences

**Figure 8.** Global body movements between contexts indicate different behavioral state

**Figure 9.** Active and passive performances when mice are not water restricted

**Figure 10.** Active performances for a second learning with a new target frequency at 4kHz.

**Figure 11.** Results for the task altering the target frequency randomly between 4 and 10 kHz.

**Figure 12.** Passive tone-evoked neuronal responses in A1

**Figure 13.** The sensory activity in A1 is modified according to the behavioral response between active and passive contexts.

## **List of Annexes**

**Annex 1.** Population activity patterns vary with cortical state

**Annex 2.** Schematic Diagram Showing the Key Circuits Involved in Regulating Brain States

**Annex 3.** Localization of A1 and during surgery and on post-mortem histology

## **1. Introduction**

### **1.1. The concept of attention**

#### **1.1.1. What is “attention”?**

The study of attention began in the psychology field in 1890 with the famous statement “Everyone knows what attention is”. To clarify this thought, William James the father of experimental psychology defined attention as: “It is the taking possession by the mind, in clear, and vivid form, of one out of what seems several simultaneously possible objects or trains of thought”. Since then, attention has become a subject of interest in many fields leading to a considerable amount of research. Many attempts have been made to further define, refine, quantify and explain this fundamental mental process. Taken together these studies led to a recent article, 130 years later, to respond to William James right from its titling, “No one knows what attention is” (Hommel et al., 2020). Indeed, the first notable problem is the variety of meanings and underlying concepts the single term “attention” can refer to. At first, we propose to review the main terms used in this research field and define the framework in which they will be referred to throughout this work.

To begin with, wakefulness, active behavioral state, arousal, vigilance, alertness, and sustained attention seem to be the first terms necessary to define and interweave. The first major behavioral states quantitatively separable are sleep and wakefulness. Indeed, each state has specific cortical and motor activity patterns (Harris, 2005). Sleep taken in itself, can be decomposed in several well understood and characterized stages (Watanabe., 2004). On the contrary, wakefulness is the starting point of multiple different forms of cortical and behavioral states (Okun et al., 2010). They are more difficult to categorize as no perfect physiological marker can differentiate them.

On the behavioral aspect, subjects who have sleep disorders or are taking sedative medication present lower levels of interaction with the environment. This degree of responsiveness to surroundings can be seen as the level of arousal (Lecea et al., 2013), enhanced by behavioral demands like motor and emotional activation or cognitive processing. The term of vigilance is used when

subjects are capable of maintaining behavioral activation for a required task over a prolonged period of time (Mackworth et al., 1964). Although this term has been employed differently by group of scientists, in this discussion as in a number of cognitive neuroscientific research (Oken et al., 2010; Ballard 2008; Sturm et al., 2001), vigilance will be used as previously defined and referred to as sustained attention. Importantly, arousal and sustained attention do not overlap as the Yerkes-Dodson curve (Yerkes and Dodson, 1908) and studies with high doses of psychostimulants testify (Wood et al., 2014). At high levels of arousal or stress situations, the ability to sustain attention and maintain performance decreases. Moreover, drowsy subjects can obtain high sustained-attention and performance when motivation and behavioral gain are increased (Oken et al., 2010). Alertness is another term widely used in this field. It can be divided into two separate processes: tonic alertness (or intrinsic alertness) and phasic alertness (Posner et al., 1990; Nebes et al., 1993). Tonic alertness is considered as the cognitive control of sustained attention and phasic alertness as the cognitive process increasing response readiness to specific stimuli.

Now that we have simply defined these terms, it is possible to consider their intersections and differences. Unfortunately, many studies label these distinct processes as “attention” (Lindsay et al., 2020). The overarching concept of attention is also used for many additional terms: focused attention, selective attention, involuntary attention, divided attention, alternating attention, feature integration ...etc. Even a broad definition cannot cover these different processes. For this reason, it is necessary to define and study them independently.

This particular study focuses on selective attention. Its most common scientific usage defines it as the cognitive process of focusing and enhancing relevant signals while suppressing irrelevant information (Lane et al., 1982; Sperling et al., 1978; Blaser et al., 1999). In several papers, this is described as a “gain control” mechanism (Francis et al., 2018). Turning off irrelevant information to reach concentration status includes both external and internal distractors such as thoughts and

emotions. This task can be very difficult especially for young subjects as selective attention develops with age (Enns et al., 1987).

### **1.1.2. The importance of selective attention in education and attention deficit disorders**

Human selective attention has been reported to be necessary for every day's life and development of academic skills such as language, literacy and mathematics (Stevens et al., 2012). Knowledge on sustained and selective attention has shown the importance for children to develop these skills for their adulthood. Training is possible to improve these skills for the better (O'Brian et al., 2013; Calmels et al., 2004). Application of these researches has been developed in special education-focused attention training programs for children with language-impairments (Stevens et al., 2008).

However, attention deficit hyper activity disorders (ADHD) is today the most frequent psychiatric disorder of childhood (Smalley, 1997) with a prevalence ranging from 3 to 7% of the population (Barkley., 1997). Often, these troubles persist into adulthood causing both disruption in professional and personal life along with other psychiatric disorders (Harpin et al., 2005). Consistently, subjects are easily distracted and unable to stay still. Standard treatments involve medication, education, skills training and psychological counseling (mayoclinic, 2021). The types of most commonly used medication are psychostimulants such as methylphenidate or amphetamine. Among other effects, psychostimulants appear to boost levels of neurotransmitters and have been reported to reduce impulsive responses to non-relevant stimuli (Bedard et al., 2004; Brodeur et al., 2000). Behavioral improvements obtained with psychostimulants are not well explained and are probably due to different causes depending on subject types (Satterfield et al., 1994; van der Meere et al., 1999). Possible explanations include cognitive and motivational boost, improved general perception or better motor control. Indeed, ADHD is a heterogeneous disorder resulting from a complex combination of genetic and environment factors (Grizenko et al., 2012).

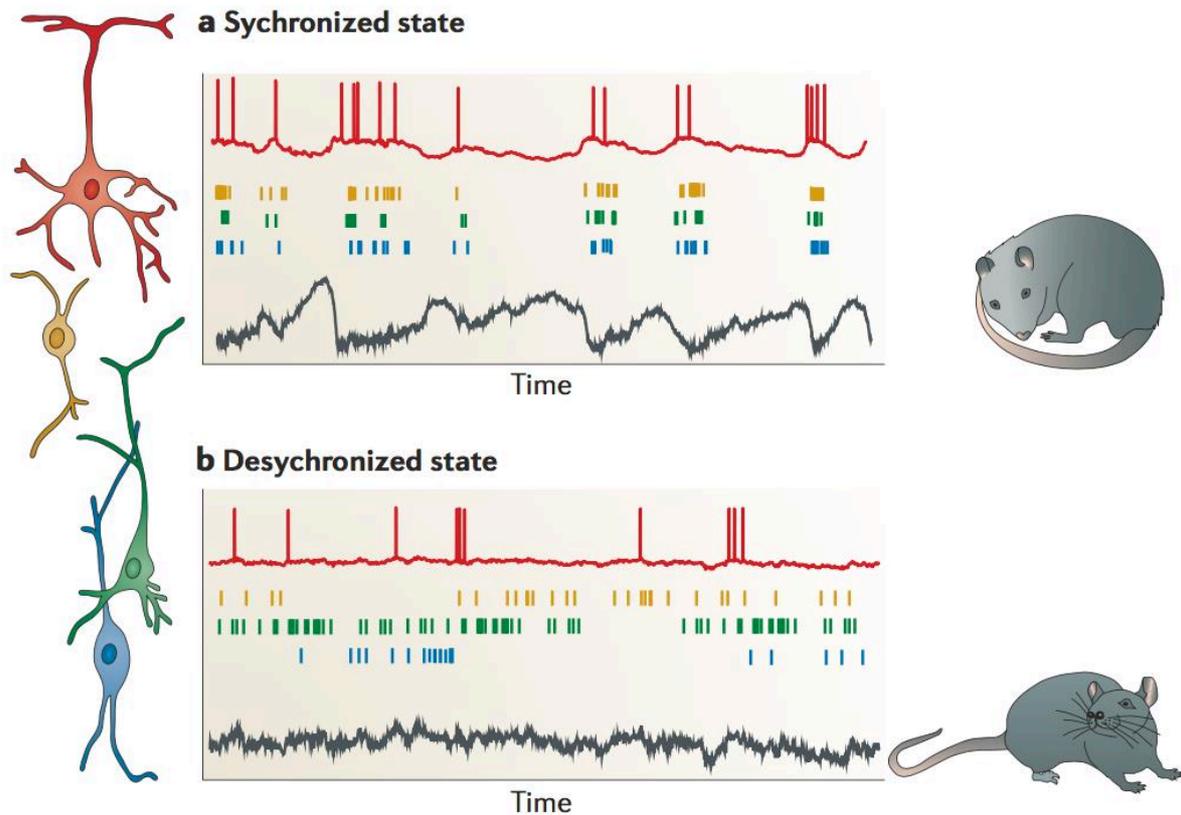
As for many other neurobiological disorders (depression, schizophrenia or bipolar disorders), only certain types of patients respond to treatment. Further research is necessary to understand better the different aspects and mechanisms under attentional processes for a personalized medicine to develop in this field (Arns et al., 2014). Personalized medicine will give rise to finer clustering of patients and more effective treatments targeting specifically the substract of the disorder.

Animals study are increasingly used in this field with the development of neuromanipulation techniques. The following discussion will mostly focus on selective attention and its neurological substract in a primary sensory cortical area to characterize its functioning during a selective attentional process. To study selective attention during a behavioral task, prerequisite include wakefulness and sustained attention.

### **1.1.3. Underlying mechanisms of active behavioral states**

#### **Active behavioral states**

Simultaneous behavioral and neuronal explorations have highlighted links between neuronal activity and behavioral states. Firstly, between sleep and wakefulness, obvious changes in the global pattern of neuronal activity can be observed widely across brain regions. These electrical differences were first described by the German psychiatrist Hans Berger (Berger, 1929). During non-rapid eye movement sleep (NREM), also known as quiescent sleep, the electroencephalogram (EEG) reports high-voltage, regular and slow patterns. With awake subjects, this activity changes to low-voltage fast electrical patterns. As the EEG signals are known to reflect underlying large population neuronal activity (Steriade et al., 1993a), quiescent sleep is referred to as “synchronized states” given that spontaneous low-frequency fluctuations activate together large populations of neurons. To the opposite, awake states are referred to as “desynchronized states”, as cortical signals have an irregular structure and length as illustrated in Annex 1 (Harris et al., 2011). More local brain activity measured with local-field-potential (LFP), confirms these findings and firing profile (Soltani et al., 2019). However, animals do not perform simple switches from awake to sleep - desynchronized to synchronized states. A continuum of states exists and can be characterized by the importance of variations between these two extreme states (Okun., 2010). For example, wakeful but drowsy animals can present slow-high voltage waves in their LFP signals similarly to sleeping animals (Poulet et al., 2008). The level of spontaneous desynchronization of cortical activity has also been reported to correlate with a high level of arousal and behavioral activity such as exploration and locomotion in rodents (Niell et al., 2010). The slow oscillations in sleeping states are translated at cellular level as several units firing simultaneously with low firing rates (Li et al., 2009). This synchrony between neurons is reflected by higher pairwise correlation values (Harris et al., 2011). On the opposite, a cortical desynchrony lowers pairwise correlation values hence modulates the neuronal population’s activity for behavioral actions.



### Annex 1. Population activity patterns vary with cortical state.

- a. In synchronized states, cortical populations show spontaneous common fluctuations in firing rate. During the up phase, all neuronal classes show a propensity to fire (shown by the coloured raster plots), whereas during the down phase spiking is reduced or absent. These phases are accompanied by corresponding depolarization and hyperpolarization in intracellular potentials (shown by the red trace). The deep-layer cortical local field potential (LFP) (shown by the black trace) shows slow negative waves accompanied by high-frequency activity in the up phase and smooth dome-shaped positive waves in the down phase. This type of activity is seen in drowsy or quiescent animals.
- b. In the desynchronized state, coordinated slow fluctuations in population activity are not seen, and low-frequency fluctuations in the LFP and membrane potentials are suppressed. This type of activity is seen in actively behaving, alert animals. Note that this figure does not show actual recordings from the neurons whose morphology is illustrated to the left, but is a drawing integrating the results of multiple studies.

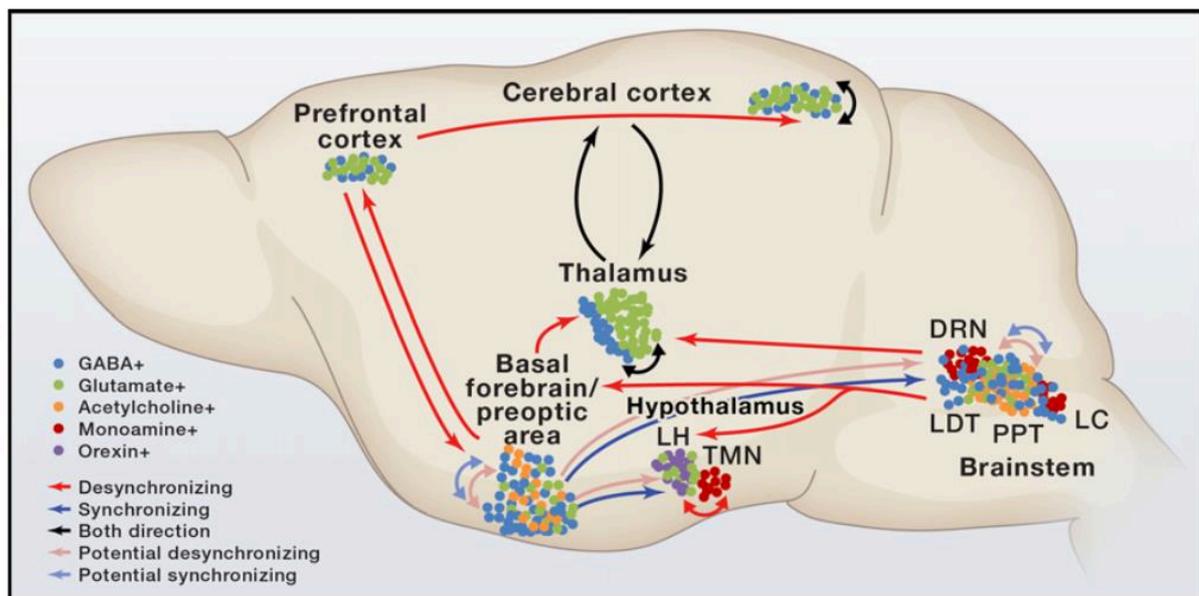
Source of illustration and legend: Harris, Kenneth D., and Alexander Thiele. "Cortical state and attention." *Nature reviews neuroscience* 12.9 (2011): 509-523.

## **Broad neuromodulatory mechanisms for active states**

The sleep-wake state depends on multiple brain-stem-thalamo-cortical-pathways involved in neuromodulatory mechanisms. The ascending reticular arousal system is known to promote cortical activation. This ascending activating system consists of several nuclei in the upper brain stem reticular formation and posterior hypothalamus, projecting to the thalamus, basal forebrain, and neocortex. They dispatch different neuromodulators such as: acetylcholine (ACh), monoamines (like noradrenaline), serotonin (5-HT) and histamine. These different mechanisms are highly intertwined and contribute to cortical desynchronization in direct or indirect ways (Kovalzon, 2016). In particular, projections from the parabrachial nucleus and precoeruleus region, relayed by the basal forebrain to the cerebral cortex, have been identified to play a critical role in desynchronized activated brain states (Fuller et al., 2012). The artificial activation of the basal forebrain also reduces low-frequency LFP power and pairwise correlations (Goard et al., 2009). Inversely, induced lesions affecting this main input of ACh to the cortex, result in an increase of cortical synchronization and low-frequency LFP power. Atropine, the antagonist of muscarinic ACh receptors (mAChRs), is sufficient to suppress cortical desynchrony in quiescent animals. However, behaviorally engaged motivated rats show atropine resistant desynchrony which is only suppressed when serotonergic blockers are combined with atropine (Vanderwolf et al., 1980). The innervation of the forebrain by the largest serotonergic nucleus (dorsal Raphe) is also part of the ascending reticular arousal system, and its stimulation increases desynchronization and levels of arousal (Vanderwolf et al., 1986). Interestingly, basal forebrain activation, ACh release and desynchronization have also been reported in rapid eye movement (REM) sleep, indicating that ACh is important for cortical activation but not exclusively correlated to active behavior (Lee et al., 2005a). Other neuromodulatory systems such as orexin neurons, noradrenergic inputs, histamine, GABAergic and glutamatergic neurons are also involved and work together to coordinate cognitive consciousness, cortical activation, and muscle availability for action, as illustrated in Annex 2 (Lee et al., 2012). Although the imbrication of these systems has been

studied extensively, further research remains necessary to continue exploring these complex activity patterns.

On top of arousal, selective attention requires orientation to a specific clue of interest. As for active behavior at a large cortical scale, selective visual attention in primates is also associated to a local cortical desynchronization. Maximum desynchronization and low-frequency LFP power are reported at a local level in the visual cortex near the neurons with receptive field (RF) responding to the targeted stimuli (Khayat et al., 2010; Cohen et al., 2009; Chalk et al., 2010). The variability of response and noise correlation are reduced in individual visual area V4 neurons processing the meaningful stimuli.



## Annex 2. Schematic Diagram Showing the Key Circuits Involved in Regulating Brain States.

Sagittal view of mouse brain. Arrows indicate major pathways connecting the brain areas. Red arrows, pathway inducing cortical desynchronization; blue arrows, pathway inducing cortical synchronization; black arrows, pathway that mediates both synchronization and desynchronization; light red arrows, possible pathway for desynchronization; light blue arrow, possible pathway for synchronization. Each cell type was schematically illustrated by colored dots in each brain area.

Source of illustration and legend: Lee, Seung-Hee, and Yang Dan. "Neuromodulation of brain states." *Neuron* 76.1 (2012): 209-222.

## **Parallel control networks**

ACh cortical innervation is reported to be required for selective attention, but is unlikely to be sufficiently targeted to modulate desynchrony in specific small local patches of cortex during selective attention phases. An additional mechanism is likely to compound with ACh to generate precise local desynchrony. With primates, feedback projections coming from the frontal eye field (FEF) are reported to spike during selective attention phases and orientated V4 responses (Moore et al., 2003). This region receives feedforward outputs from visual areas and subsequently controls saccades and eye movements (Vernet et al., 2014). In a visual detection task, enhanced functional coupling is reported between the corresponding areas of the FEF and V4 (Gregoriou et al., 2009). This coupling is time-shifted by about 8-13 ms and initiated by the FEF suggesting a top-down control mechanism. In addition, brief microstimulations of retinotopically FEF sites enhance RF responses in the corresponding V4 area, in a similar manner as during selective attention (Amstrong et al., 2003). These responsive neurons have also been reported to increase their activity during the preparatory times preceding target stimuli (Zhou et al., 2011). Preparation of response in the FEF is associated with larger and faster evoked responses to target stimuli within the RF of neurons (Noudoost et al., 2021). This mechanism can be seen as holding information for behavioral demands hence assimilated to working memory. Working memory reflects sustained attention, with both being closely related concepts impacted in the same way (Silver et al., 2005; Unsworth et al., 2020). These preparatory states for goal-directed actions facilitate behavior and lower reaction time (see part III). Several studies consider the FEF part of the frontoparietal network (Hart et al., 2020; Alahyane et al., 2014), mobilizing different structures in sustained attention (see onwards).

## **Mesolimbic reinforcement**

Drivers of active behavioral states are motivational-incentivized and require mesolimbic-processing. In particular, dopaminergic circuitry in the midbrain and nucleus accumbens (NAc) are

involved in memory formation (Adcock et al., 2006) and behavior selectivity (Knutson et al., 2001). The mesolimbic system is also involved in promoting ACh delivery through mesolimbic-cholinergic circuitry. The medial prefrontal cortex (mPFC) sends glutamatergic projections to the nucleus accumbens (NAc) and integrates top-down signals for goal-directed behavior (Gruber et al., 2009). In turn the NAc activates the basal forebrain that sends cholinergic projections to the cortex for active behavior. Activation with N-methyl-D-aspartate receptor (NMDA) of the NAc-cholinergic pathway increases visual discrimination performances especially in presence of distractors (Peters et al., 2011). In Peters' study, the removal of cholinergic projections to the mPFC reduced these performances. Unsurprisingly the removal of equivalent projections in the posterior parietal cortex also triggered similar impairments, as this region is considered part of the frontoparietal network.

### **The frontoparietal network**

The posterior parietal cortex and dorsolateral prefrontal cortex are activated simultaneously during sustained attention in humans (Corbetta and Shulman, 2002; Pardo et al., 1991). As generally low arousal levels induce poor behavioral performances (Ciria et al., 2021), arousal control is needed for sustained attention to maximize performance. In particular, a correlation has been described between pre-stimulus arousal, reaction-time to target and performance. The frontoparietal network seems to be involved in the top-down control of attentional demands, as it is only activated during detection tasks. High arousal levels reported during resting phases were no longer correlated with the activity of the frontoparietal network (Foucher et al., 2004). This network was particularly active during pre-stimulus preparatory times, leading to sustained attention during the tasks. This active phase preceding the stimuli helps the preparation of the response and improves performance and reaction time (Makeig and Jung, 1996; Makeig et al., 2000). The activation of this network is also increased under challenging conditions showing its importance in top-down control when subject to high attentional loads (Berry et al., 2017).

## **Selective attention**

For decades, selective attention has been defined as the mechanism to selectively enhance representations of target stimuli (Luck et al., 1997; Sperling et al., 1978; Blaser et al., 1999). These studies were essentially conducted in the visual cortex. More recent additional studies make it unclear whether this attentional selective enhancement mechanism applies to all cortices. During the engagement phase of learning a task, suppression of whisker-evoked responses was reported in the barrel cortex (Castro-Alamancos, 2004a). In another study on the auditory cortex of rats, decreased firing rates for behaviorally engaged conditions were observed at the presentation of the target stimuli compared to passive listening (Otazu et al., 2009).

This study focuses on auditory selective attention in the primary auditory cortex (A1) of mice and explores the extent to which the attentional meaning of a target sound stimulus affects its processing at this first cortical step.

## **1.2. From a sound to its contextual meaning**

### **1.2.1. Overview of the auditory pathway and cortical processing**

Auditory information originates from sounds in the environment. Sound are mechanical waves of molecular compressions and dilatations. In the air, these vibrations travel at  $340\text{m}\cdot\text{s}^{-1}$  and are characterized by their frequencies and amplitudes. Natural sounds are often highly complex and made up of a variety of frequencies, intensities and locations. Any complex sound can be decomposed into a sum of pure tone frequency signals of different intensity. Throughout mammalian development, the auditory pathway has developed to process this complex acoustic environment and send information to the brain.

#### **Auditory pathway**

The organ of hearing, the ear is the first to play a role in translating sound information from the environment to the brain. The visible part, the outer ear, is composed of the pinna and the ear canal. It receives the pressure waves and guides the vibrations to the eardrum (or tympanic membrane). The eardrum is mechanically linked to the middle ear composed of the three smallest bones in the body, the ossicular chain. In turn, the motions of the ossicular chain are transmitted to the fluids of the inner ear through the oval window membrane. The inner ear contains the cochlea responsible for transducing mechanical to electrical information. Fluidic movements in the cochlear canals cause the basilar membrane and the organ of Corti to vibrate. Stereocilia deflection depolarizes the hair cells at the top of the organ of Corti, triggering them to release neurotransmitters and excite auditory nerve fibers (White et al., 2020). According to the frequency of the mechanical sound wave, this process will be localized such that high frequencies cause activation in the base of the cochlea and low ones are processed at the apex. This organization is known as the tonotopy (Galambos et al., 1943; Robles et al., 2001). Through this functional mechanism, the cochlea is capable of decomposing sounds with precision. For humans, the perception abilities range from 20 Hz to 20 000 Hz. The encoding of

acoustic pressure is also very precise as it can make sense of sounds between 0 dB SPL to 120 dB SPL. The sound wave intensity information is captured through the firing rate of the fiber. Therefore, the firing rate of each fiber dedicated to a frequency band is dependent of the intensities of the frequencies in a sound (Johnson, 1980). These fibers constitute the cochleo-vestibular nerve (eighth cranial nerve) and project to the cochlear nuclei.

Once the information has reached the cochlear nuclei in the brain stem, auditory messages are conveyed to the brain through two pathways: the primary auditory pathway, specific to auditory information; and the reticular sensory pathway carrying all types of sensory processing.

The primary auditory pathway is a short and fast pathway (3 to 4 relays), composed of large myelinated fibers conveying information to A1. From the cochlear nuclei, fibers cross the midline and synapse in the second major relay in the brain stem: the superior olivary complex composed of 5 nuclei. A third ascending fiber then carries the message to the inferior colliculus in the midbrain. A last and fourth relay occurs in the medial geniculated body (MGB), also called thalamus, before reaching A1. This describes the lemniscal pathway, the core branch of the primary auditory pathway (Malmierca et al., 2012). A secondary branch, the non-lemniscal pathway, projects to caudal regions of MGB mainly sending information outside of A1 (cochlea.eu, 2020).

The reticular sensory pathway also integrates auditory information. Small fibers leave from the cochlear nuclei and join the ascending reticular formation. From the brainstem to the midbrain, several synapses occur as sensory information are integrated together and triaged. After the ascending reticular formation, fibers reach the non-specific thalamus and project to the polysensory cortex.

Both pathways possess multiple feedback connections influencing processing (Schofield et al., 2010). The rest of this discussion will be centered on describing A1 in the lemniscal primary auditory pathway.

## **Cortical A1 processing**

Information processed by a sensory system is reflected by sensory representations at each step of the pathway. The lemniscal primary auditory pathway has been described as reducing stimuli information redundancy between the midbrain and cortex (Chechik et al., 2006). It encodes for higher order dimensionality information from the cochlea to the auditory cortex (Atencio et al., 2012). Although the tonotopy organization is present from the cochlear nuclei to A1, frequency selectivity is less organized in the A1 than in subcortical areas (Kanold et al., 2014). This tonotopy organization starts to decrease within A1, between thalamo-recipient layer IV and the layer II-III (Montes-Lourido et al., 2021), and continues to decrease along the cortical hierarchy as complex nonlinear sound features are increasingly encoded (Issa et al., 2014; Sadagopan et al., 2009). Small differences to acoustic vocalization are also more pronounced in non-primary areas such as the suprarhinal auditory field compared as in A1 (Carruthers, 2015). The mechanisms by which the auditory cortex builds these complex representations are only beginning to be understood. A shared hypothesis suggests that representation is created through nonlinear computations between “building-blocks” encoding gradually complex sound features (Liu et al., 2019). This organization was artificially recreated through simulations using a multilayer deep learning network, which showed a boost of decoding performances in artificial tasks such as speech recognition (Kuchibholta et al., 2018; Akbari et al., 2019).

## **Cortical processing affected by contextual and behavioral meaning**

From the first stage of cortical processing, this encoding complexity is added to the contextual and behavioral meaning of sounds. For example, during rodent locomotion, the cortical processing in response to passive sounds is reduced (Zhou et al., 2014; Williamson et al., 2015). This decrease results from the recruitment of multiple circuits such as cholinergic fibers coming from the basal forebrain (Nelson et al., 2016) and top-down projections from the motor cortex (Schneider et al., 2014).

Cognitively demanding tasks also appear to modulate responses within the auditory cortex subject to the behavioral conditions. As described next, studies on selective attention with different species have reported modulated responses in the auditory cortex.

### **1.2.2. Selective attention studies in the auditory cortex**

#### **Auditory behavioral tasks with humans**

Selective auditory attention has been studied for over 50 years on humans. When subjects are attending to clicks, event related potential (ERP) activity has been reported as enhanced over the auditory cortex region compared to unattended situations (Picton et al., 1971). These findings were confirmed by a dichotic listening task whereas subjects responded to tone pips in an ear and ignored the same stimuli in the other. ERP peak activity was considerably larger for the hemisphere related to the attentive ear (Hillyard et al., 1973).

The dependency to sound meaning was assessed during intermodal selective attention studies, where subjects shifted attention between auditory and visual cues. The activity of the auditory cortex only increased during auditory cue processing (Shomstein et al., 2004). Increased cerebral blood flow was consistently observed during a dichotic task, with positron emission tomography in the hemisphere actively processing a sound (Alho et al., 1999). Neuromagnetic field recording also reported amplified signals for the auditory cortex attending stimuli. This amplification was detected as early as 20 ms after stimuli onset presentations (Woldorff et al., 1993; Fujiwara et al., 1998). This short latency suggests modulation may have already begun in sub-cortical processing stages. Although studies in the MGN show modulations during anticipation of auditory signals (Hoehnerman et al., 1990), this does not exclude the possibility of additional processing in the cortex. In addition to early-stage modulations, studies with magnetoencephalography have shown reorganized sound encoding. During a pitch discrimination task, when subjects were asked to differentiate high and low frequency, the distance of encoding between these signals was enhanced (Ozaki et al., 2004). These results suggest

selective attention can reshape the auditory cortex mapping response through short-term plasticity mechanisms. During an intermodal task, a precise localization of activity with functional magnetic resonance imaging scanner (fMRI), suggested that the labile attentional activation originates in the lateral auditory regions responsible for higher-order encoding of sound features (Petkov et al., 2004).

Taken together, these results suggest increased sound processing within the auditory cortex during attentional demands. Cited studies also found increased activity in frontal regions (Alho et al., 1999) where executive control mechanisms of attention take place (see part 1.1.3.).

### **Auditory behavioral tasks with primates**

Selective auditory attention has also been explored at the cellular level with invasive techniques on animals. Rhesus monkeys were trained to detect targets in an intermodal selective task, with visual and auditory cues coming from similar or different sides. Stimuli changed behavioral relevance depending on the cues block. Signals coming from auditory cues blocks have exhibited differences in response strength for 2/3 of the units recorded in the auditory cortex. Also, equal proportions of cells were enhanced or suppressed (Hoehnerman et al., 1976). Another intermodal selective attention study reported shortened latency of response and increased firing rate when attending to auditory stimuli (Miller et al., 1980).

Additional behavioral studies have also been carried out on rodents exploring selective attention mechanisms. The following part will concentrate on such studies on A1 to understand whether the very first step of auditory cortical processing is affected by attentional demands.

### **1.2.3. Different experimental paradigms lead to different conclusions**

Selective attention in rodents has been studied mainly during active listening and passive exposure to the same sound. Kato's study on mice started by habituating animals to pure tones. Single-cell activity was recorded during passive exposure with 2-photon imaging (Kato et al., 2015). Within 5 days, the excitatory activity of A1 cells was suppressed while inhibition was enhanced. The animals then learned to associate a given behavior to a target sound during a Go / no-Go task. Cellular signals in active blocks where animals had to discriminate target sounds for reward were compared to signals in passive blocks with no behavioral relevance. In behaving blocks, 51.6% of excited cells showed stronger activity compared to circumstances of passive exposure. The activity of somatostatin-expressing inhibitory neurons (SOM cells) was reduced in active conditions compared to learnt-passive and habituated-passive conditions. This decrease in activity observed in SOM cells resulted in less inhibition and therefore an increased activity at the cellular network level. Similar results were also obtained in Francis' study (Francis et al., 2018) where, with a similar behavioral paradigm, massive "gain control" to target tone was reported during active condition as 87% of cells were enhanced compared to passive condition.

However, another similar study comparing active and passive state in a Go / no-Go task showed contradictory results (Kuchibholta et al., 2017). Broad-based suppression of neural responses was observed during the active context. Network modeling to match responses in the active context to all recorded cell types (pyramidal neurons, SOM cells, parvalbumine - PV cells, vasoactive intestinal peptide -VIP cells) showed that this suppression is possible through the parallel recruitment of the inhibitory neurons (SOM cells, PV cells, VIP cells). An additional study on freely-moving rats with a Go / No-Go task electrically recorded A1 single and multi-unit activity and reported suppression during attended stimuli (Otazu et al., 2009). Moreover, during an intermodal auditory-olfactory task where animals performed interleaved blocks of auditory and olfactory cue discrimination, no differences were observed between the active auditory listening and the passive auditory exposure combined with

active olfactory detection (Otazu et al., 2009). This suggests that engagement in any discrimination task, either auditory or olfactory, impacts A1 activity independently of the sound behavioral meaning.

These opposing results could come from a “sensory centered approach”, where these various studies do not sufficiently characterize and take into account behavioral states and attentional demands. A quantification of states seems essential to study information processing, as arousal and behavioral states influence as much neuronal responses as sensory inputs (see part 1.1.3). Differences in attentional loads could also be a possible explanation. The A1 has been reported to be necessary for some tasks (Cooke et al., 2018) but unnecessary for others (Ceballo et al., 2019). In addition, lick responses and reward encoding (Francis et al., 2018) could induce changes not due to the processing of the sound.

To address these multiple challenges, we have developed and implemented a Go / no-Go auditory head-fixed task with several specificities. To begin with, the reference sounds used were a set of 30 broad-band noise-like stimuli (Temporal Orthogonal Ripple Combination, TORCs) including the target sound of 10 kHz (Klein et al., 2000; Klein et al., 2006). These complex sounds created ambiguous distractors to the target stimuli. In addition, the reference sounds had varying time lengths and were relatively long compared to previous studies (Kuchibholta et al., 2017). Through this design, selective attention is needed for the discrimination of the target stimuli as well as sustained attention within active blocks as they involve long distractors and block durations. In practice, animals were asked to press on a lever when they detected the target stimuli. This allowed to avoid motion artefacts due to the movements from licking observed in other experiments (Chen et al., 2013). Behavioral states between active and passive contexts were tracked by measuring the movements of the animal’s two paws and mouth. The animals were also trained as long as necessary so passive blocks led to passive behaviors.

In vivo 2 photon imaging allowed to record A1 activity. Analysis of active and passive blocks could be interpreted using a robust attentional task and behavioral state quantification.

## **2. Experimental Procedures**

All experiments were realized on CBA-Jackson male mice and carried out in accordance with the French and European legislations relative to the protection of animals used for experimental and other scientific purposes.

### **Experimental animal**

CBA Jackson mice (CBA/J – 000656) were used as experimental animals. CBA/J is an inbred mouse line created in 1948 and used in many research fields. It is particularly known for its blindness and long-lasting hearing abilities (jax, 2018). Indeed, the CBA/J strain does not carry any of the recessive genes responsible for the age-related hearing loss in BALB/cByJ and C57BL/6J. CBA/J are also described as a calm mice strain, with low aggression reported, low spontaneous locomotion activity, minimal evasive behavior and struggling during holding (conductscience, 2018).

Hence, the CBA/J mice strain was selected for our head-fixed auditory attentional task on the ground of its auditory abilities and the ease of manipulating the mice for several hours a day.

### **Behavioral training set up**

The behavioral set-up was designed and put together in-house, adapting and assembling manufactured and custom-made components.

Animal training is performed in a closed cubic 1 m<sup>3</sup> box, acoustically insulated with foam. The head-fixed support is composed of an aluminum base (15 cm x 10 cm) covered with a non-slip lining. This support is mounted on a rail, moving back and forth depending on the needs of each animal. The animals were kept head-fixed using a titanium head-post sealed on their skull (Ymetry, Bagnolet, France). This head-post is attachable to a dovetail mechanism placed at a height of 10 cm from the aluminum base. An ultra-sensitive lever (product number ENV-310) and lick-port mechanism were

sourced from Medical Associates. The lick-port was mechanically adapted to descend in less than a second in front of the animal's mouth. The water delivery is controlled by a miniature solenoid precision valve (LHD Series, Lee valve). For each correct response, a drop of 3-5  $\mu\text{L}$  is delivered. Three manual linear translation stages allow this block to move in the x,y and z axes with a millimetric precision.

The sound was delivered using a free-field electrostatic speaker (Fostex), positioned at a distance of 20 cm on the right of the animal's ear and on the contralateral side of the cranial window. The speaker was calibrated to cover a range from 2 to 70 kHz and to deliver a flat response ( $\pm 1$  dB). Infrared lights and camera (ace 2 Basic, Basler) were used to visualize and record the behavior of the animal during the sessions. The acquisition frequency was 40 Hz, triggered by the two-photon recording. A custom LabVIEW (Laboratory Virtual Engineering WorkBench) program was developed to pilot a National Instrument Card (PCIe6363 card) in real-time, which in turn sent 5 mV outputs to control the motors, valve and microphone. The program also recorded the lever pressing responses and associated reaction times through analog inputs.

### **Auditory stimulation**

Audio waveforms were digitally generated using Matlab (The MathWorks Inc., MA) at a sample rate of 80 kHz. They were converted into analog voltage through a PCIe6363 card (National Instrument). The target tone was a pure tone and the reference sounds noise-like stimuli (TORCs). We used 26 TORCs resulting from the addition of 6 moving ripples: broadband sounds covering 5 octaves (from 2 kHz to 32 kHz) with sinusoidal spectral envelopes. The ripple velocities (cycles/s) spanned from 1 to 8 Hz by step of 1 Hz drifting up or down the logarithmic frequency axis. As the temporal phase of each moving ripple was randomly chosen, each stimulus was temporally orthogonal to the other (Klein et al., 2000; Klein et al., 2006). During the task, 13 pairs of TORCs were used (each pair was composed of a TORC and its opposite for the direction of motion).

Sound pressure level was on average 70 dB. No equalization was applied for the changes in acoustic calibration. As the set-up had been acoustically insulated, the noise due to the laser and recording devices was limited. During the task, each reference sound was repeated 8 times - 4 times for each of the active and passive contexts, with a mean of 36 s per sound made of stimuli of 3 s (25 %), 4 s (50 %) and 7 s (25%, catch trials). Except for catch trials, the target sound followed the reference sound.

### **Animal preparation**

Surgery was performed on 11 to 13 weeks aged CBA/J mice. Buprenorphine (0.1 mg/kg, I.P) was injected to each animal before the procedure. Each animal was anesthetized with Isoflurane (3-4 % induction, 1-1,5 % maintenance) under a mix of oxygen (20 %) and nitrogen protoxyde (80 %). For intrinsic imaging sessions, isoflurane anesthesia was reduced to 0,5-1 % under oxygen (100 %). During surgery, paw withdrawal reflexes were suppressed and a heating blanket (Harvard Apparatus) maintained the rectally measured body temperature at 37°C. The head of the mouse was maintained by a custom-made fixation system, holding the nose and the incisors. The skin and temporal muscles overlying the AC were removed.

A1 was located with intrinsic imaging by visualizing the functional activity resulting from pure tones of 4 kHz and 10 kHz. After treatment of the bone with an acid and alcoholic cleaning procedure (Optibond FL, Kerr), a custom-made titanium implant (Ymetry, Bagnolet, France) was sealed on the skull of the mice with dental cement (Tetric Evoflow, Givadent). A small craniotomy (~3 mm diameter) was performed with a milling machine to expose the A1 on the left hemisphere. The dura mater was left intact and continuously hydrated with saline water. A solution containing an adeno-associated virus (AAV1-syn.jGCaMP7F.WPRE, Addgene) was injected in the lower frequency band of A1 to mediate the expression of the calcium probe protein GCaMP7f. A transparent coverslip was used to

seal the craniotomy and replace the skull. At the end of the surgery, mice were injected with Buprenorphine (0.05 mg/kg, I.P) and 1 mL of saline solution. Animals were left several days to recover.

### **Behavioral training procedures**

Animals were first accustomed to the experiment by being handled twice a day during 4 days for 5-10 minutes. Animals were then water-restricted in line with approved protocols by the French and European legislations. As no water was given and typically, animals dropped to 80-85% of their body weight within 48 hours. In parallel, mice were familiarized to the behavioral apparatus and head-fixed condition during 2 days twice a day for 10 minutes.

Animals were then trained to an attentional Go / no-Go auditory task following several steps. The task consisted in discriminating broad-band noise-like stimuli (TORCs) as reference signals from the pure tone as target signal. In each trial, a reference sound is followed by a target tone (except for the catch trials). The target frequency was fixed at 10 kHz. Responses to the reference signals were counted as wrong answers and were not reinforced. Correct responses to the target tone were rewarded by a ~3-5  $\mu$ L drop of water. Typically, animals underwent 3 training sessions a day of 100 trials each. At the end of each training day, animals received up to 1.2 mL of water to maintain body weight at around 80%.

To become able to perform the complete behavioral task, mice followed a preliminary training sequence of three steps. The first step consisted in learning to press on the lever. Animals were given 10 s of target tone listening to press on the lever. The animal was taken to the second phase of training once it had reached a >80 % performance over two subsequent sessions. The second step then introduced a short reference signal of 0.8 s before a reduced 3 s target tone. Responses to the reference signal were counted as false alarms (FAs). Once the correct answer (hit) rate reached >60%, the animal moved to the third step. It consisted in varying the duration of the reference signals between 0 and 4 s. The reference sound was followed by a target tone of 3 s. Once in five trials, a 7 s

catch trial of reference sound was introduced to prevent mice from learning to respond systematically at the end of the trial. This step was considered learned when the  $d'$  was above 1 (see 'Data Analysis' paragraph below).

After these preliminary training phases, the final task interleaved active and passive blocks of 25 to 29 trials each. During active blocks, the lever was accessible and hit trials were positively reinforced. During passive blocks, the lever was withdrawn and no reinforcement was applied. The animals were trained to perform the task in alternated blocks (A/P/A/P/P/A/P/A). At the end of their training, expert animals reached ~70% of hits during active blocks. Over passive blocks, animals learnt to disengage which meant they had understood the block structure of the task. This disengagement was initially evaluated visually by the experimenter. The passive exposure to sound was considered as learnt when the mouse remained motionless no later than after the 5<sup>th</sup> trial following a transition to a passive block.

The animal was considered 'expert' when it had reached these performance criteria, point at which two-photon calcium imaging recordings were performed. At the end of the recording sessions, water deprivation was interrupted and access to water provided for 1 hour, bringing back the mouse body weight at ~90%. Following this hydration, a behavioral session was recorded whereby it was observed that the hit rate dropped to ~10% during the active sessions, evidencing that animals were not motivated anymore.

For three mice, a second training cycle began the next day after no water was given overnight and the body weight had gone down to ~80%. After checking motivation through ~20-25 trials on the 10 kHz task, the target tone was changed to 4 kHz. In front of this change, hit rate decreased to ~2%. Hence and as mice did not associate target tones of different frequencies, a new training cycle with the same learning phases as described above was performed with 4 kHz as the target tone. The reference sounds and task structure were identical. After one week, the mice became 'expert' to the new target frequency as their performance increased to ~70%. We then run and recorded a task

randomly altering between the 4 kHz and 10 kHz target tones showing that the animals could discriminate both targets equally versus reference sounds.

### **Confirmation of imaging location in A1**

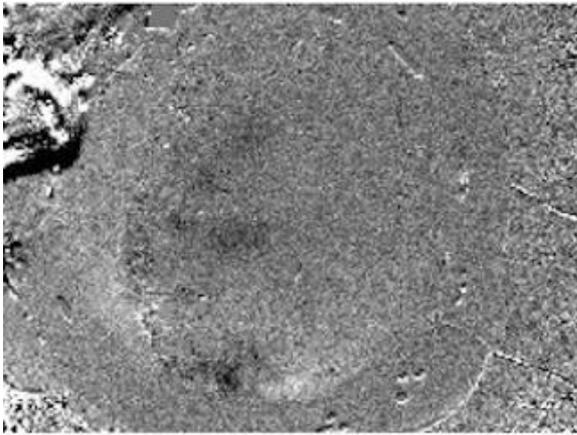
The location of the craniotomy was first determined spatially (0.5-1 mm anterior to Lambda; 6.5-7 mm lateral). A window of 5 mm diameter was created around the zone of interest and the skull was thinned for intrinsic imaging. 4 kHz and 10 kHz pure tones were played while mice were kept under 0.5-1 % isoflurane in oxygen (100 %) anesthesia. After tone onset, three dark spots appeared from rostral to caudal in AAF, A2 and A1, indicating an increase of the oxygen consumption. The craniotomy and virus injection were performed in the low frequency band region of A1 that responds to 4 kHz tone. After recovery from the surgery, another intrinsic imaging session was carried out through the cranial window to confirm the superposition of the signal with the fluorescence.

Using two-photon calcium imaging, we also confirmed the location of recordings in A1 based on onset response latencies and the response profile to pure tone. When recording at ~300  $\mu\text{m}$  depth, the fields of view (FOVs) showed latency of ~50 ms and a sustained activity in response to low frequency pure tone (2-8 kHz). The same analysis was performed at the end of the recordings to confirm no functional damage was done during the several weeks long experiments.

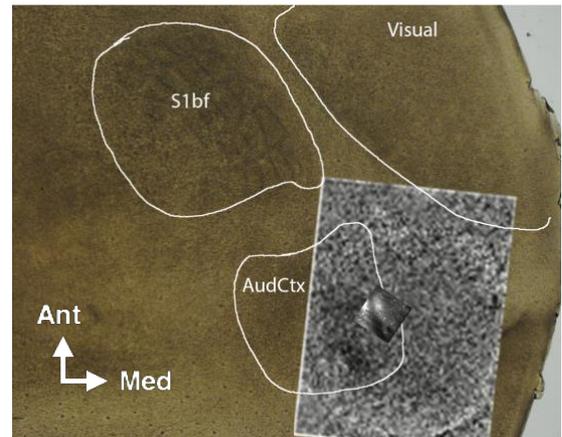
Finally, a post-mortem histology of the brain was conducted to locate anatomically AC, A1 and the recorded regions. While mice were deeply anesthetized with Ketamine (100 mg/kg, I.P) and Xylazine (10 mg/kg, I.P), a whole-body fixation was achieved by injecting fixatives (paraformaldehyde, PFA 4 %) in the left ventricle. After the fixatives had diffused into all tissues, the cortex of the animal was dissected. It was flattened and left overnight in PFA 4%. Sagittal slices of cortex were then cut from surface to depth by steps of 100  $\mu\text{m}$  using a vibrating microtome (Leica). These slices from surface to depth correspond to layers of the cortex. The boundaries of the AC were identified by the contours

of its granular layer IV made visible by using white light. Matching the blood vessels on the surface of the brain made it possible to overlay AC boundaries, intrinsic imaging signals and recording sites.

A.



B.



### Annex 3. Localization of A1 and during surgery and on post-mortem histology

- A. Imaging intrinsic signal for 4 kHz pure ton during surgery
- B. Localization of the AC with the contours of the granular layer IV. According to the blood vessels pattern at the surface of the cortex, the functional intrinsic imaging signal and the recording site were superimposed on the cortex slice.

## Two-photon calcium imaging

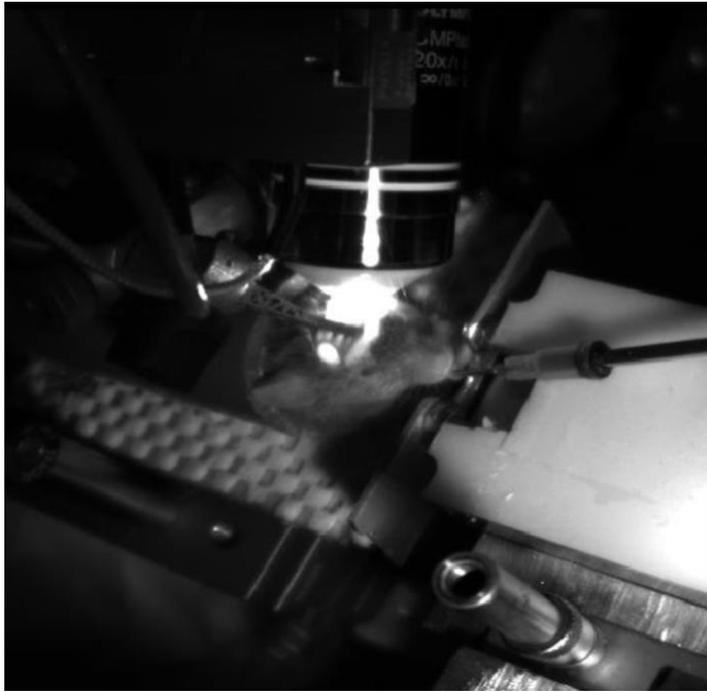
During the craniotomy, animals were injected with a solution containing an adeno-associated virus (pGP-AAV1-syn-jGCaMP7f-WPRE, Addgene) diluted in sterilized PSB. The initial virus concentration was  $1.8 \times 10^{13}$  GC/ml, which was diluted 20 times to reach a concentration of  $9 \times 10^{11}$  GC/ml. The injection was performed under continuous pressure with a picospritzer and glass capillaries (tip opening  $\sim 3 \mu\text{m}$ ). A red fluorescent dye (Alexa Fluor 594) allowed to monitor the injection so that it stays at a slow and constant pace. In average, the fluorescence was visible  $\sim 21$  days after injection.

The calcium fluorescence was recorded with a two-photon microscope. The custom-made two-photon microscope resulted from the coupling of an in vivo microscope (equipped with a water immersion objective Olympus x20 NA 0.95, XLUMPLFLN20XW) and an infrared pulsed laser (Ti:sapphire laser, Mai-Tai HP, Spectra Physics). The path was shielded from external light contamination.

At 250 and 400  $\mu\text{m}$  depth, the FOVs were recorded at different sizes: 150  $\mu\text{m}^2$ , 300  $\mu\text{m}^2$  and 500  $\mu\text{m}^2$ . The combination of a resonant and a galvanometric scanner allowed to record frames at an acquisition frequency of 40 Hz, representing a maximal temporal resolution of 25 ms. Very short pulses ( $10^{-13}$  s) of 920 nm were sent to the tissue at the rate of 80 MHz. The resulting energy applied to the focal plane was  $\sim 10$ -50 mW preventing the tissue from being damaged. The emitted photons (500-550 nm) were collected by a fiber bundle and counted by four GaAs photomultiplier (Hamamatsu H10770PA-40 + C9744).

Before each recording, the cover glass was cleaned with alcohol and optical gel was placed between the window and the objective to maximize the photon collection.

The figure 1 is a snapshot of a mouse performing the Go / no-Go auditory task while A1 is optically recorded. This snapshot illustrates the assembly and use of the techniques described above.



**Figure 1. Snapshot of a mouse performing the Go / no-Go task while A1 is being recorded**

The mouse has performed a hit and it is licking the lick-tube that had just arrived in front of it. The mouse has 4 s before the lick-port is retracted. Five seconds of inter-trial is counted before the next trial begins.

## Data analysis

### Behavioral data

All plots presenting shaded error bars are mean  $\pm$  standard error of the mean (SEM).

For learning curves (figure 3.B, 3.D, 10.A), steps of 45 trials were interpolated from the individual mice curves of the mean performance values of each run (~100 trials). Mean values and SEM were obtained by averaging rates every 45 trials. For each phase, the values retained for mice having finished training before the maximum training duration were their last performance values. This is represented on individual curves with dotted lines (figure 10.A).

$d'$  (figures 3.C, 6.B, 9.C, 10.C) is a measurement derived from signal detection theory used for the analysis of decision making in the presence of uncertainty (Stanislaw et al., 1999).  $d'$  values are calculated as  $d' = z\text{-score}_{\text{hit}} - z\text{-score}_{\text{false alarm}}$ .

Figures 4.B and 11.A representing the mean hit rate in active blocks and the mean FA rate in passive blocks were obtained by averaging individual mean performances of mice (mean of 6 sessions per mouse). During the active phase, catch trials were considered as hits when they were preceded or followed by a trial during which mice participated.

For active block analysis (figures 6.A, 10.B, 11.B), the hit rate was calculated as the number of correct responses divided by the total number of trials. Omission rate is the ratio between missed target tones and the total number of targets presented, as FA and catch trials are excluded, the omission rate is not equal to '1 – hit rate'. The FA rate is the number of FA performed divided by the total number of trials. The lever press rate indicated the level of engagement in the task and is equal to '1 – omission rate'. Correct rejection rate shows the proportion of catch trial mice correctly rejected (else, counted as FA).

The lever pressing probability density (figures 6.C, 9.A, 10.D, 11.C) is calculated in time steps of one tenth of a second. This probability density (%/s) was obtained by dividing the probability to

press by the length of the time step. The area under the curve is equal to the sum of the hit and FA rates (~0.77).

Instantaneous rates (figures 6.D, 10.E, 11.C) represent the mean hit and FA rate from trial onset. For each time step of half a second, the instantaneous hit rate is calculated as the number of effective hits divided by the number of potential hits. The same calculation was applied for FA rates. As TORCs last for at least 3 s, no hit can occur during the first three seconds from trial onset. Between the 3<sup>rd</sup> and 4<sup>th</sup> second, the lever press can be a hit or FA subject to the trial been target sound or TORC. The spontaneous FA continues up to 7 s which is the full duration of the catch trials.

For each animal, mean active and passive FA proportions were calculated over ~6 sessions for the first five trials after context transition (figure 5). The FA proportion obtained from the 6<sup>th</sup> trial after transition to the end of the blocks was considered as the reference FA rate.

### Camera data

Camera data were analyzed using a popular open source pose estimation toolbox DeepLabCut (Mathis et al., 2018). For each mouse, a deep neuronal network was trained (~200 labeled frames) to recognize paws and mouth. Incorrect labelling corrections were performed by hand for optimizing the network. Once the network reached sufficient levels of likelihood for each body part and no errors were observed on assessment videos, all corresponding videos were processed.

To estimate movements at each time point, a mean value position was calculated for each body part in each trial. The distance to the mean  $R = \sqrt{(x - x_{mean})^2 + (y - y_{mean})^2}$  was extracted for each time point within a trial. The x and y coordinates correspond to the horizontal and vertical positions of the body part in the image. As the labeling precision for the algorithm training is a few pixels (approximately the middle of the labeled body part) while DeepLabCut returns positions to the hundredth of pixel, an estimate of the precision of the measure was required. To that end, the STD of the  $R_{still}$  from quiescent moments was calculated for each animal, which dependent to the body parts

and mice, ranged from 1 to 4.5 pixels. For each assessed period in the trial, a Z-score representing the amplitude of movements was attributed using the formula  $Z\text{-score}_{\text{Behavior}} = \frac{STD(R)}{STD(R_{\text{still}})}$ . This measure is referred to as the index of movement.

### **Behavioral state analysis**

To further characterize behavior in the active and passive state, four conditions were studied. For each animal, the comparisons between the distributions of  $Z\text{-score}_{\text{Behavior}} = \frac{STD(R)}{STD(R_{\text{still}})}$  were made between contexts. Two-time intervals were studied: “Before” - 3 s preceding trial onsets, and “After” - 3 s following TORC onsets. These comparisons between contexts indicate whether animals are behaving similarly before trial onset and during TORC listening. Attentive trials were considered as active hit trials and non-attentive trials as passive correct trials. Only attentive active trials and non-attentive passive trials were included in the comparisons between distributions. Receiver operating characteristic curves (ROC curves) were calculated for each pair in each of the four conditions. The area under the curve (AUC) value was calculated and used as the indicator of disparity between each pair of distributions. Within each context, the difference of distributions before and after the trial onset was also studied to characterize the reactions of the animal to the TORC.

The statistical difference between each distribution pair is validated using a boot-strap method. This technique consists in randomly shuffling the distributions 250 times and calculating an AUC value for each shuffle. If one of these AUC values (1/250) calculated from the random distributions is greater than the one on the original distributions, they are considered as not statistically different ( $p > 0.004$ ).

## **Motion correction**

The two-photon images were corrected using a custom-written software in Matlab (MathWorks) for pixels to remain stable from one frame to another. We developed and applied a correlation-based correction with possible translations in the X-Y planes. At first, a mean image of a few hundred frames was computed. The maximum pixel-by-pixel correlation with the mean image was computed for each frame and frames were repositioned accordingly. This correction was performed routinely. After image correction, a film was generated of 40-by-40 frames stacked together. Any residual movement or loss of FOV could be detected by a low peak of correlation with the mean image. The minimum peak was taken as a threshold and any trial containing a 40-by-40 stacked frame with a correlation below this minimum value was excluded from the analysis. For each recording, a manual confirmation was performed to confirm the stability of frames after correction and the removal of artefacts. No residual movement was found.

This motion correction technic is only valid for mice with little brain movements visible in active conditions. This can be largely explained by the surgical technic that applies pressure on the surface of the brain with a cover-glass. Experience were only continued once preliminary recordings showed little movements in the X-Y planes and no Z-shifts.

## **$\Delta F/F$ extraction**

After correction, ImageJ and MatLab (MathWorks) were used for signal extraction. Mean images of recordings and several films of stacked images were used to identify the region of interest ( $ROI_{cell}$ ) where the activity takes place. Cellular regions of interest were manually drawn on ImageJ and exported as binary masks to MatLab.

Neuropil signal correction was performed to avoid background noise contamination. For each  $ROI_{cell}$ , a second mask was calculated with an expansion algorithm. This consisted in drawing a ring

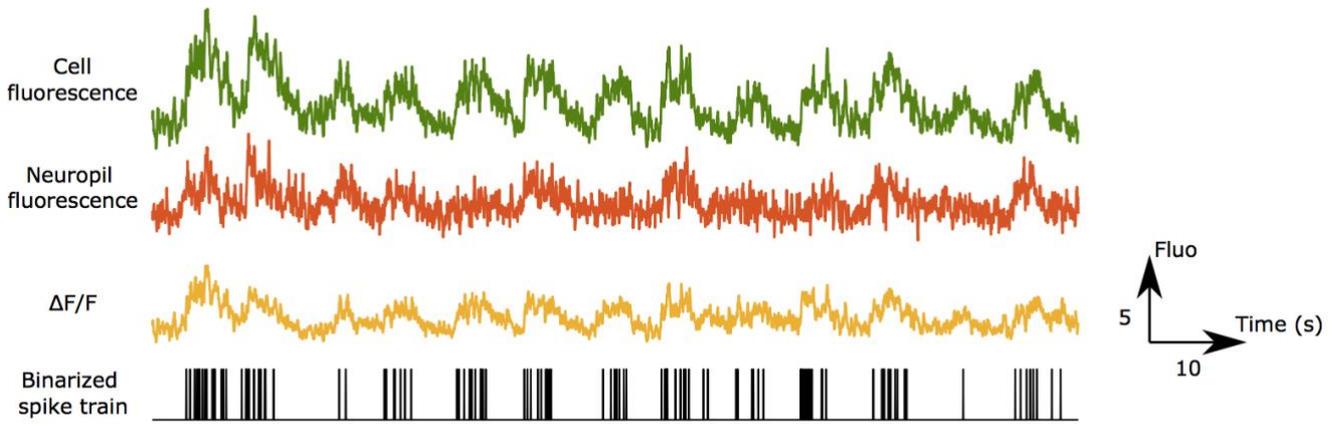
around the cell. This  $ROI_{Neuropil}$  had an equal number of pixels as its corresponding  $ROI_{cell}$ . Pixel contained in  $ROI_{cell}$  were ignored by the algorithm and could not be part of a  $ROI_{Neuropil}$ .

Fluorescence signals for the  $ROI_{cell}$  and  $ROI_{Neuropil}$  were extracted for each frame. For each time step, the fluorescence  $F$  was calculated by averaging the number of photons contained in a ROI. The fluorescence baseline signal  $B_0$  was estimated as the 8<sup>th</sup> percentile of the fluorescence intensity in a sliding window of 20 s centered on the frame (Panier et al., 2013). The residual fluorescence for  $ROI_{cell}$  and  $ROI_{Neuropil}$ , free from large baseline movements was calculated as  $[\Delta F/F] = [F - B_0 / B_0]$ .

Neuropil correction was obtained by subtracting  $[\Delta F/F]_{cell}$  to  $[\Delta F/F]_{Neuropil}$  weighed by a parameter  $\alpha$  as:  $[\Delta F/F] = [\Delta F/F]_{cell} - \alpha [\Delta F/F]_{Neuropil}$ . The  $\alpha$  parameter was cell dependent and inferred from a noise regression algorithm (Dipoppa et al., 2018).  $\alpha$  values were tested between 0 and 1 and chosen to minimize the noise contained in the distribution of each cell's signal. This was done by reducing the width of the left part  $[\Delta F/F]$  distributions, thus  $\alpha = \text{argmin}\{SD(\Delta F/F(t; \alpha) \leq 0)\}$ .

### **Spike extraction**

In order to extract the spike trains from  $\Delta F/F$  calcium signals, we used a Bayesian Deconvolution Algorithm: the Blind Sparse Deconvolution Algorithm (Tubiana et al., 2020). We used an adaptative version of the algorithm beginning with a template of the GCaMP7f response with  $\tau_{rise}$  (50 ms),  $\tau_{decay}$  (300 ms), and a  $\Delta F/F$  amplitude (0.1) for a single spike. A trace example is shown on figure 2.



**Figure 2. Signal extraction from a cell**

The raw fluorescence of a cell is shown (green) as well as its associated neuropil fluorescence (red).  $\Delta F/F$  was extracted from these signals by minimizing the signals noise (yellow).  $\Delta F/F$  were next deconvoluted to obtain the binarized spike train (black). These signals are plotted for one minute of recording.

### Pure tone analysis

To confirm the location of the recordings in A1 at the cellular level, pure tone recordings were performed for each mouse with two-photon microscopy. Three time intervals  $t$  were studied: 1 s from tone onset, total sound duration (3 s) to characterize sustained activity and 2 s after tone offset.

To determine the best frequency of each cell  $i$ , a Z-score representing the strength of response was calculated for each frequency  $f$  as:  $Z_{f,i,t} = \frac{[\langle \Delta F/F \rangle_i >_f - \langle \Delta F/F \rangle_i >_{SPONT}]}{std[\Delta F/F]_{i,SPONT}}$ . Spontaneous activity ( $\langle \Delta F/F \rangle_i >_{SPONT}$ ) corresponds to the 2 s recording preceding auditory stimuli onsets. The tone activity ( $\langle \Delta F/F \rangle_i >_f$ ) was calculated on the time interval  $t$ . For each  $t$ , the best frequency  $f_B$  was defined as the frequency  $f$  with a maximal  $Z_{f,i}$  and was noted as  $Z_{f_B,i}$ . Cells with a  $Z_{f_B,i} < 2$  were considered as not responding to sound. An example of a FOV recorded on a mouse is shown on figure 12.A. Responding cells are of the color that corresponds to their best frequency and grey cells are non-responsive cells. An example of a cell having a best frequency  $Z_{f_B,i} = 5.6$  kHz is shown on figure 12.B. The mean fluorescence traces computed on 10 repetitions for pure tones stimuli is given. An associated

heat-map of individual responses to each stimulus is presented on figure 12.C. For all mice ( $n=3$ ), best frequency of the responding cells is shown on figure 12.G.

Time onset responses are an important marker to confirm the recording of A1. For each cell  $i$ , for each frequency  $f$ , and for each repetition  $r$  of the stimulus, we extract  $T_{i,f,r}$ : the timing of the first spike following auditory stimuli onsets. A window of 0.5 s was added before the tone onset and a window of 1 s was added after the sound offset. A histogram binning  $T_{i,f,r}$  applied every 25 ms and is shown on figure 12.D. The mean firing rate (FR) for each frequency and for each cell was calculated during the total sound duration and is shown on figure 12.E.

As in previous studies, a mean reliability index was calculated from passive responses to pure tone stimuli (Bandyopadhyay et al., 2010; Francis et al., 2018; Kanold et al., 2014). For each frequency  $f$ , and for each cell  $i$ , the reliability index equates to the percentage of responding trials. Responsive repetition  $r$ , for each frequency  $f$ , and each cell  $i$  are identified as  $|Z_{r,f,i}| > 1$  with:  $Z_{r,f,i} = \frac{[\langle \Delta F/F \rangle_{i>r,f} - \langle \Delta F/F \rangle_{i>SPONT}]}{std[\Delta F/F]_{iSPONT}}$ . Figure 12.F. shows the cumulative density of the proportion of responsive cells in function of its responding trials %.

### Context modulation analysis

Among all cells (963 cells,  $n=4$  mice), we first defined responding cells to TORCs stimuli. The fluorescence variations for each cell  $i$ , between passive listening of TORCs and spontaneous activity, was calculated as  $D_i = \langle \Delta F/F \rangle_{i>TORC} - \langle \Delta F/F \rangle_{i>SPONT}$ .  $\langle \Delta F/F \rangle_{i>TORC}$  represents the mean over TORC length and  $\langle \Delta F/F \rangle_{i>SPONT}$  the spontaneous activity for the 3 s of recording before the auditory stimuli onsets. A bootstrap distribution of  $P(D_i)$ , randomizing the position of TORC onsets was built and  $\alpha$  was equal to the standard deviation of  $P(D_i)$ . Cells were considered as responding when they were facilitated  $D_i > \alpha$  or suppressed  $D_i < -\alpha$ . Only responding cells were included in the rest of the analysis.

Spike trains for all responding cells were extracted with the Blind Sparse Deconvolution Algorithm (Tubiana et al., 2020). Individual firing rates  $f_i$ , for the context  $C$  and for the sensory environment  $S$  were calculated and noted as  $f_{i,C,S}$ . The context  $C$  can be active  $f_{i,A,S}$  or passive  $f_{i,P,S}$ . The sensory environment was divided as following: spontaneous activity (3 s before stimuli onset), TORCs stimulation (from 3 to 7 s) and target tone onsets (0 to 0.2 s).

For both contexts, only correct behavioral responses trials were included in the analysis (per recording,  $90 \pm 7$  passive trials;  $69 \pm 5$  active trials). In order to avoid artefacts linked to the difference of trial number between contexts, we first calculated the mean response of individual TORCs in active and passive contexts., The mean  $f_{i,C,S}$  of TORCs was calculated subsequently. In the active condition, the mean behavioral response time to the target tone is 0.5 s. To avoid including any time step containing big paw movements related to lever pressing, the studied time window was from 0 to 0.2 s after target onset. For comparability purposes, the same target time window was considered for passive trials.

In the sequence of sensory appearance during trials (spontaneous, TORCs and target onset), the distribution of  $f_{i,C,S}$  for the active and passive contexts are shown on figure 13, panels A to C. The significance of the difference between distributions of responses was evaluated using the Student's  $t$ -test.

For each cell, the modulation index  $MI_{i,S} = \frac{[<f_{i,A,S} - f_{i,P,S}>]}{<f_{i,A,S} + f_{i,P,S}>}$  is calculated. Cells have a significantly different response between contexts when they are excluded from the confidence interval (CI) as shown on figure 13, panels D to F. The CI is the value of the standard deviation of the distribution of MI calculated on randomized active and passive assignation:  $CI_S = std (P (MI_{A_{ran,S}} - MI_{P_{ran,S}}))$ . Cells having MI close to zero are not represented.

Mean active  $f_{i,A,S}$  are plotted according to mean passive  $f_{i,P,S}$  as shown on figure 13, panels G, H and I. On these panels, each color represents the responding cells of a mouse. The CI was

calculated as previously, as the value of the standard deviation of the distribution of differences between randomized active and passive assignation:  $CI = \text{std} (P (f_{A\_ran,S} - f_{P\_ran,S}))$ .

Figure 13.J. shows an example of FOV calculated by stacking the entire recording. Arrows 1 and 2 point to two example cells. Figure 13.K shows their mean responses to TORCs onset in the active and passive contexts.

The control condition was performed on three mice who were not water deprived, hence not motivated during the behavioral task. The majority of active responses were misses (miss rate  $75.0 \pm 14.9\%$ ,  $n=3$  mice). As previously, the  $f_{i,c,S}$  were calculated. The significance of difference between sensory response distributions was evaluated using the Student's *t*-test.

### **3. Attentional Go / no-Go auditory discrimination task**

As mentioned in Part 1. Introduction, different forms of attention exist and are grounds for study. In this work, we are focusing on three attentional processes: 1) sustained attention over time; 2) selective attention through environment filtering; and 3) ability to adapt attention to the context. The 5-choice serial reaction-time test (5-CSRT) developed by Carli (Carli et al., 1983) is widely used to assess sustained and selective attention in rodents. In the 5-CSRT experimental task, rodents have to detect visual illuminations of ports and break the associated photo-beam to achieve reward. This freely moving task requires to maintain attention spatially to all ports present in the environment and respond correctly. Temporal attention is also needed as any premature response is considered as an error and not reinforced. To these ends, mice have to train and develop inhibitory control to reduce impulsivity and gain in attention (Robbins et al., 2002; Bari et al., 2008). Attentional load is increased by changing cues delay and adding distractors.

For 25 years, this task has been extensively used to study effects of drugs, lesions and attention deficit hyperactivity disorders (ADHD, Russell et al., 2015). Multiple variations of this task exist such as its adapted version for mice using auditory cues (Nilsson et al., 2016).

To determine how the behavioral meaning of sounds affects the cortical networks of sensory processing, this study has developed a specific attentional Go / no-Go auditory discrimination task. On the one hand, it mobilizes and tests the same attentional characteristics as the 5-CSRT in that it demands sustained attention and vigilance over time (~208 trials), selective attention through filtering of the target stimuli and inhibitory control through reduced motor impulsivity. On the other hand, a major difference with the 5-CSRT lies in the head-fixed condition. As much as head-fixed tasks are well-known to be well more complex and prolong learning (Juczewski et al., 2020), they are necessary and enable to record brain activity with two-photon calcium imaging microscopy. In addition, the ability to adapt attentional set was also added as mice were trained on a context-switching block structured task alternating active and passive blocks. The properties of the sounds presented in the active (A) and

passive (P) context were equivalent and played the role of distractors with high attentional load as they integrated the target frequency highly relevant to behavior into noise.

The analysis of behavioral results confirms correct learning and evidence of attentional processes through accurate responses according to context, low omission rate, low response latencies to target and low premature responses (Hoffman et al., 2015).

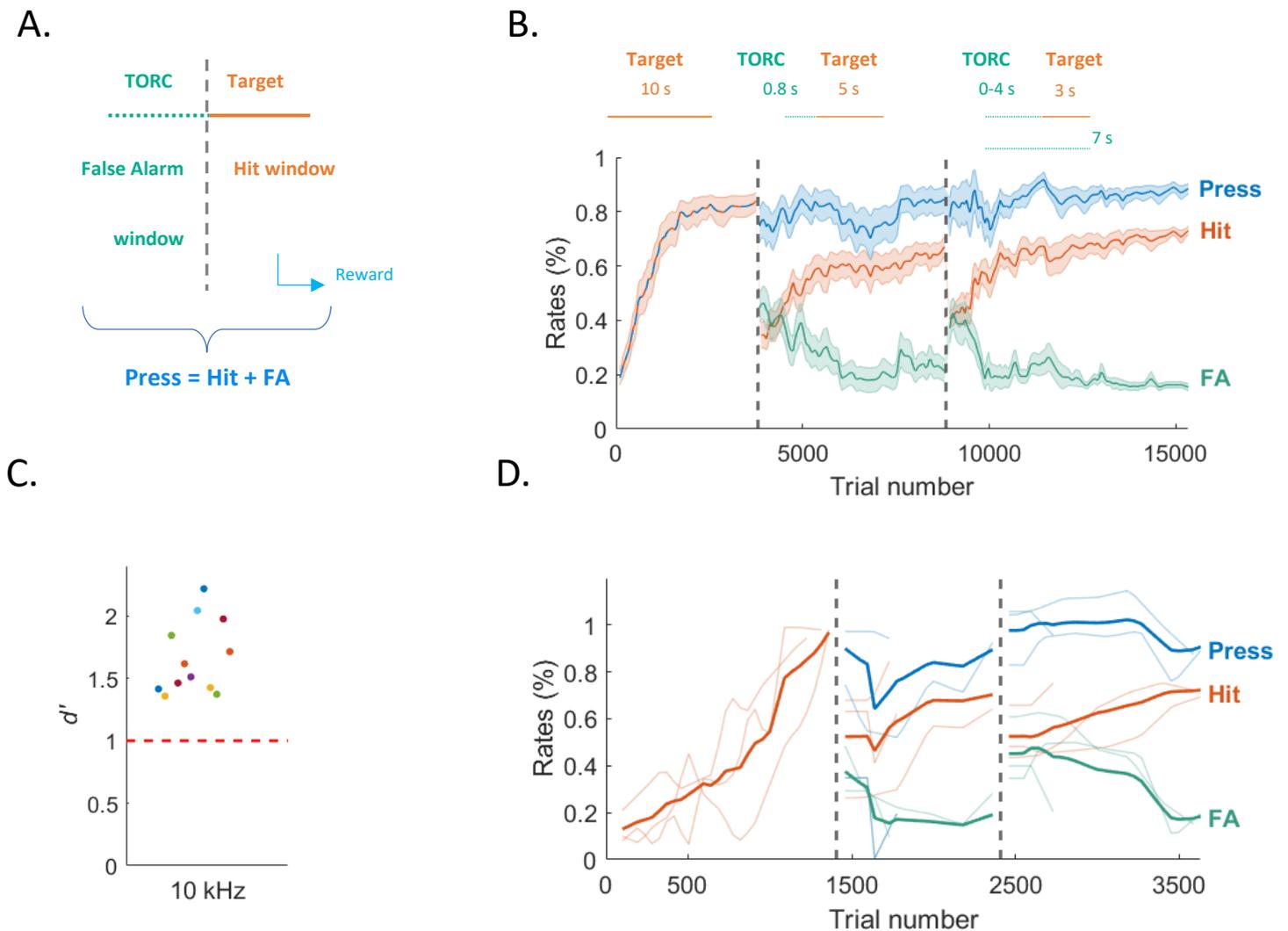
### **3.1. Shaping mouse behavior for an attentional Go / no-Go auditory task**

Twelve mice were trained to discriminate the target frequency (10 kHz for 3 s at 70 dB SPL) following the steps described in figure 3.B. The three first steps of behavioral shaping encompassed up to 15 500 trials per mouse, corresponding to approximately 51 full training days. The first step consisted in teaching mice to associate a lever press with a reward - with success requiring between 1 200 and 4 300 trials, 4 to 15 days subject to the animal.

The second step introduced a short reference sound preceding the target tone. At first, FA rate was superior to the hit rate as motor impulsivity led mice to press on the lever as soon as a tone was emitted. Response gradually became more specific to the target tone. At the end of this training step, the hit rate was >60 %.

The third training step alternated randomly 26 reference sounds lasting from 0 to 4 s. FA and hit rate started around 40 %. For each TORC duration, mice learnt to respond exclusively to the target frequency after around 10 000 trials over 10 days, subject to the animal. Among the fifteen mice enrolled in the training protocol, twelve correctly learned to respond to the target frequency following a reference sound lasting from 0 to 4 seconds (figure 3.C). This step was considered “learned” when the  $d'$  ( $1.6 \pm 0.1$ ) exceeded 1. Learning durations differed markedly from one mouse to another. The individual learning profiles of three fast learning “best mice” are shown in figure 3.D. Overall, the long

training duration compared to simple pure tone discrimination tasks, can be explained by the difficulty of the task required for attentional load (Rinberg et al., 2006).



**Figure 3. Quantification of the responses for each stage of the learning protocol**

- A. Diagram of response domains valid at each stage of learning. Correct responses (Hit) to the target are rewarded. Wrong answers (False Alarm) during the TORC result in the end of the trial.
- B. Description of the three stages of learning with their associated performances. During the first stage, hit and pressing rate are equivalent. The solid lines represent the mean rates ( $\pm$  SEM). Respectively for each learning stage:  $n=16$ ,  $n=15$ ,  $n=12$ .
- C. Discrimination index ( $d'$ ) between the target and reference sounds at the end of the learning. Each point represents an individual mouse,  $n=12$ .
- D. Individual learning curves for three “best mice” (~10 days of training). The bold line represents the mean rates,  $n=3$ .

Once this multi-step training protocol had been optimized and was routinely performed, the study took four mice to the final full context-switching block structured task for two-photon microscope recordings.

### **3.2. Behavioral modulation according to the context**

#### **3.2.1. Accurate performance according to the context**

Figure 4.A represents the active and passive contexts schematically. In the active context, the lever is accessible for the full duration of the trials. During the passive context, the lever is removed for the whole block. 104 trials were generated by randomly picking the duration of the TORC – among 3, 4 and 7 s in the required proportions (respectively 25 %, 50 %, 25 %), as well as the reference sound among the 26 TORCs. Each of the trials was conducted in the active and passive contexts. The target tone was presented after 3 and 4 s TORCs. Catch trials of 7 s TORCs with no target tone represented one out of four occurrences. In the active context, the mice performances assessed their ability to perform selective attention by responding to the target tone and proving inhibitory control by not responding to previous signals. During the passive context, mice had to remain quiescent. Passive FA was detected by paw movements mimicking hits in the active context. The absence of response in the passive context (figure 4.B) was completely voluntary on the part of the mouse as no reinforcement was associated with correct responses or FAs.

As the experiments were ran, mice correctly learnt to discriminate the contexts in ~1 week by refraining from reaching the lever in passive blocks. Performance in the active context remained stable. Once the animals reached correct performances in the passive blocks, camera behavioral and neural recordings were performed. These recordings involved four mice and last over a month each.

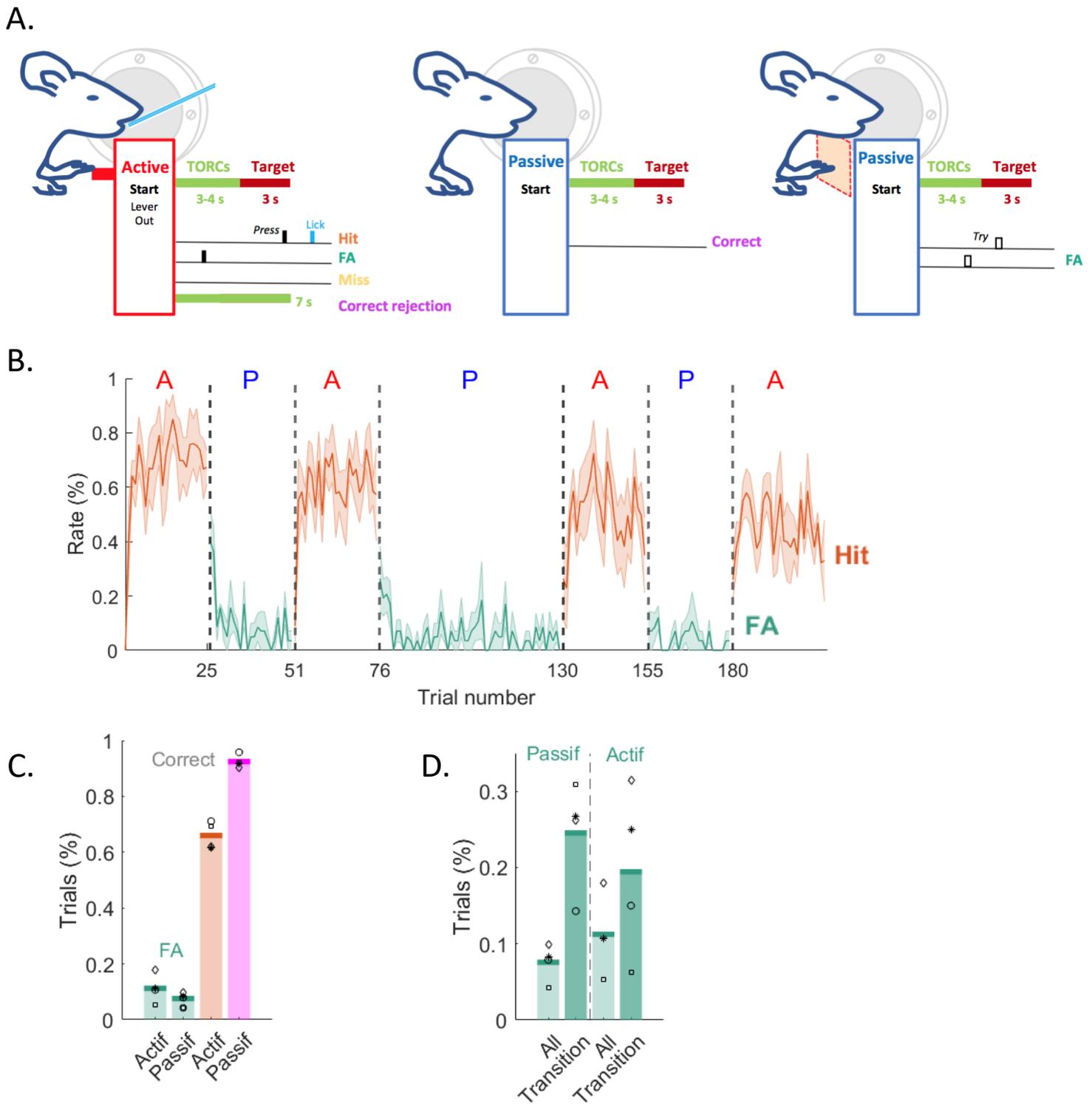
The complete task contained 8 blocks of 25-29 trials each (A/P/A/P/P/A/P/A). Figure 4.B shows the mean performances of the four recorded mice for each consecutive block (~6 sessions per mice,

n=4). The mean hit rate in the active blocks decreased slightly but steadily across blocks (minus  $7.4 \pm 0.04$  % at each block). Despite this slight motivation loss due to water intake, animals were able to maintain a sustained attention over time as FA rate remained stable.

Overall, as shown on figure 4.C the correct responses are respectively  $66.0 \pm 2.4\%$  (A) and  $92.4 \pm 1.2\%$  (P). The FAs are  $11.2 \pm 2.6\%$  (A) and  $7.6 \pm 1.2\%$  (P).

These results suggest that filtering and responding to the target tone is highly energy consuming for mice and the balance between their energy cost versus benefits lead the animals to change behavior between contexts. Consequently, subject to the context, the target tone has different behavioural meanings.

Whatever the context, FAs occurred more frequently after context change.



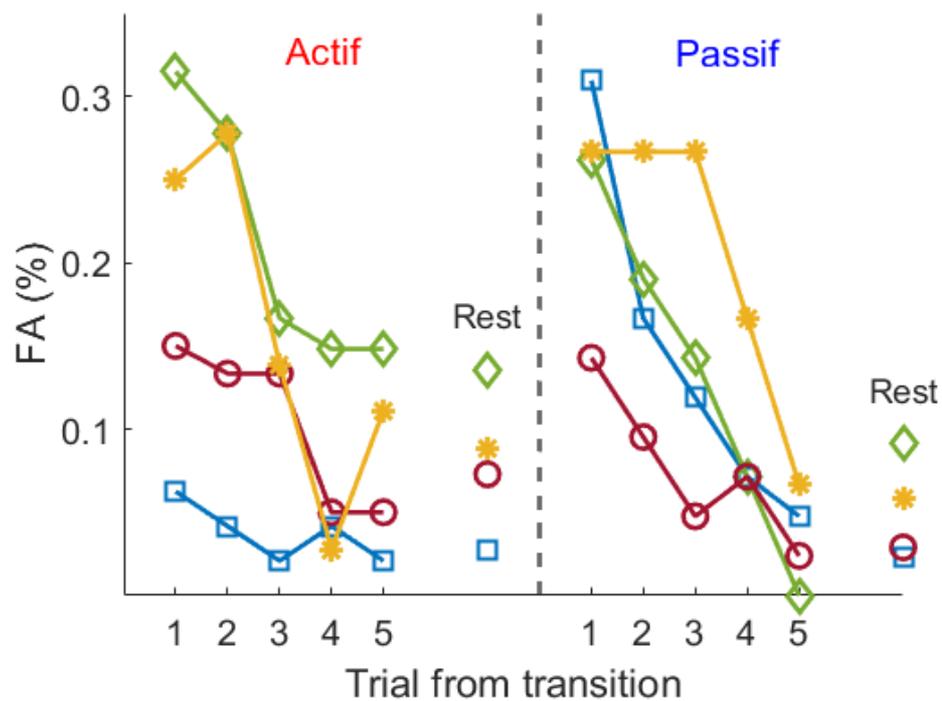
**Figure 4. Active and passive performances during the attentional auditory Go / No-Go task**

- Diagram of the active (left) and passive context (middle, right). In the active context, the lever is available. In the passive context, correct responses are counted when no paw coordinates are detected near the location of the lever (middle). Paw movements mimicking hits are identified as FA (right).
- Average performance for each trial during active and passive blocks. The hit rate (orange) is represented in the active context. The FA rate (green) is represented in the passive context. The solid lines represent the mean rates ( $\pm$  SEM),  $n=4$ .
- Total average of performances (FA and correct responses) for active and passive context. Individual averages are indicated by a marker,  $n=4$ .
- FA average for the first trial of each block compared to the global average of the block. Individual rates are indicated by a marker,  $n=4$ .

### 3.2.2. Fast adaptation to context switching

For the first trials right after context change, the proportion of FA is higher than the mean FA rate in that context (figure 5). This higher FA rate is quickly reduced and the error rate returns to normal after 2 to 3 trials depending on the animal.

This fast adaptation to context switching reflects an effective acquisition of the proper behaviors in each context and an efficient ability to adapt attention to the set.



**Figure 5. Context change performances indicate fast adaptation to context switching**

Average FA for each trial after context change. FA rates quickly drop to average values obtained from the sixth trial to the end. Solid lines represent individual averages,  $n=4$ .

### 3.2.3. High engagement in the active context

Among the four recorded mice, the hit rate represented  $66.0 \pm 2.4\%$  and the FA rate  $11.2 \pm 2.6\%$  (figure 6.A). Interestingly, the correct rejection rate (no response to the 7 s TORCs catch trial) of  $83.5 \pm 4.1\%$  rules out responses based on the timing from trial onset.

The omission rate of  $19.1 \pm 2.7\%$  evidenced that the task was well calibrated in terms of difficulty and length to mice behavioral demands and cost-benefit balance (Ortiz et al., 2020). As a methodological reminder, the omission rate represents the proportion of missed targets while catch trials are not included in the number of potential correct answers. The lever press rate is equal to '1 – omission'.

For each mouse, the psychometric measurement ( $d'$ ) is well-above 1 indicating the good discrimination of the target tone (figure 6.B).

During active trials, the mean reaction time is 500 ms showing that mice respond almost instantly after the target onset (figure 6.C). This reaction time is consistent with previous findings that reported reaction times with rats of 300-400 ms (Blockland et al., 1998). The instantaneous hit rate remains constant whether TORCs last 3 or 4 s (figure 6.D), showing no change of expectancy despite 4 s TORCs being more frequent (50%) than 3 s ones (25%). This rate decreases sharply after 4.75 s from trial onset showing that animals respond either very rapidly after the target tone onset or not at all.

Fast responses to targets have been reported to reflect internal motivation states and maximize gain rewards (Mosberger et al., 2016). Ripples variations for the reference sounds and their durations have been reported to drive the increase of the animals' attentional loads (Bushnell et al., 2009). The absence of variations to the low instantaneous FA rate (figure 6.D) around 3 and 4 s confirms no motor impulsivity or target anticipation behaviors from the animals.

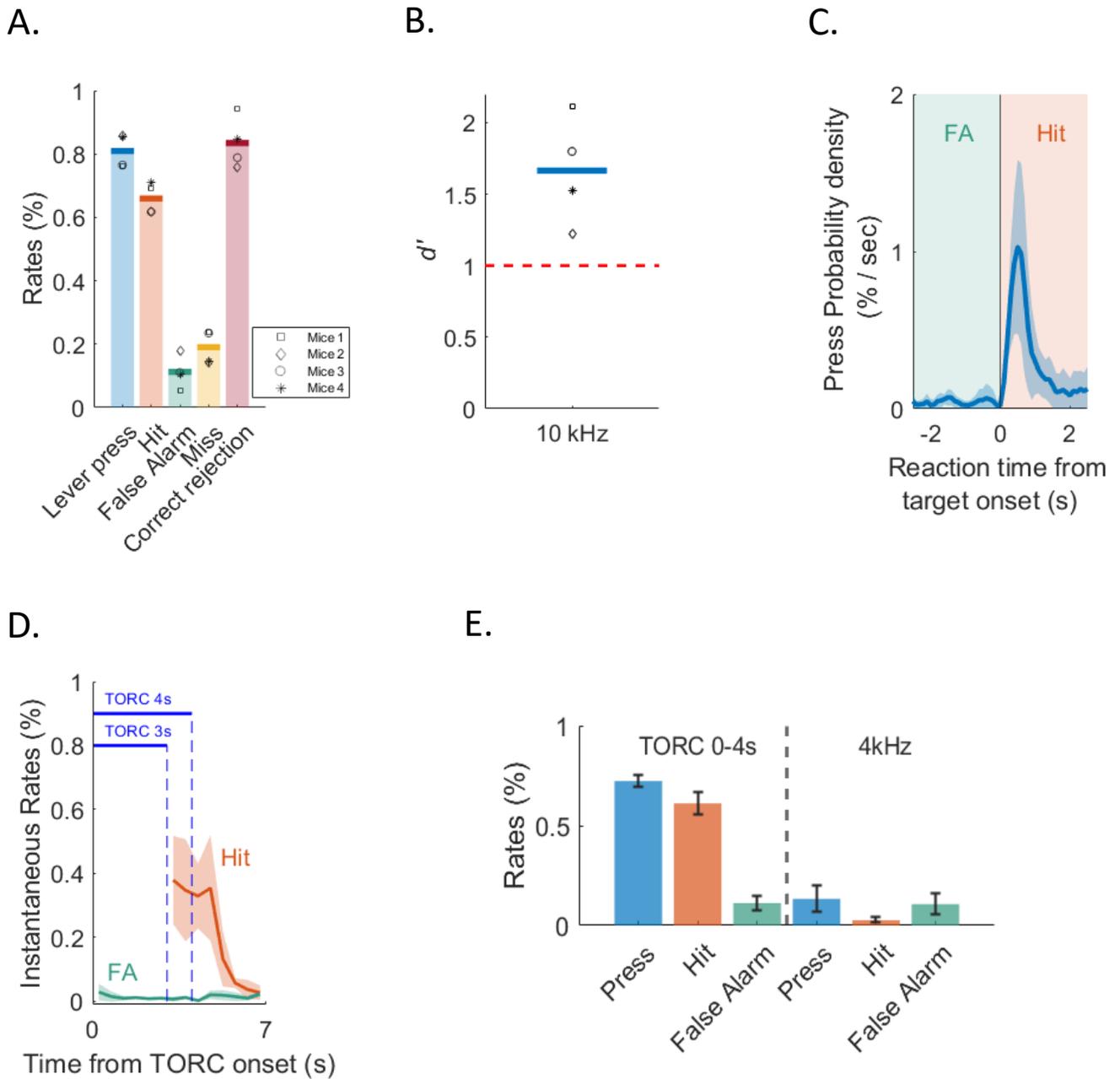
Together, these measures reflect a high-speed processing of information and target tone selectivity (Domenger et al., 2007). Such behavior necessarily requests an active filtering of sounds

from the trial onset evidencing selective attention. Consequently, mice selective attention is proven by active hits on a trial-to-trial basis.

To confirm attention selectivity to the target tone of 10 kHz, two controls were performed. A first control with TORCs ranging from 0 to 4 s confirmed that mice could also detect target tone in a different trial architecture with shorter reference sounds (similar hit rate,  $61.1 \pm 5.6\%$ ).

A second control kept a similar trial architecture but changed the target tone from 10 kHz to 4 kHz. After checking motivation with a few 10 kHz target tones, the target was changed to 4 kHz and led to task disengagement (fig 2.4.e). Performances dropped sharply ( $2.7 \pm 1.3\%$  of hit) and FA rate exceeded the hit rate.

Together these controls confirmed the major dependency of the response to the target frequency of 10 kHz. The responses were independent of the trial structure and the nature of the pure tone.



**Figure 6. High target reactivity in the active context**

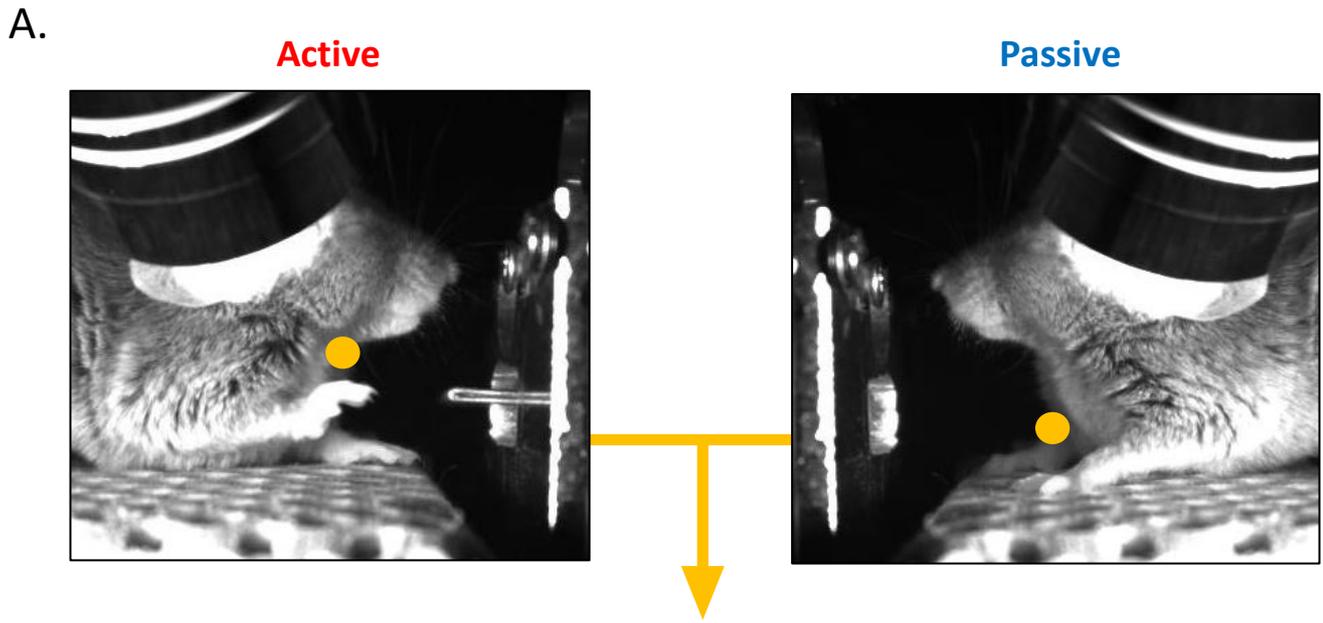
- A. Mean rates in the active context.
- B. Mean discrimination index ( $d'$ ). Individual rates are indicated by a marker,  $n=4$ .
- C. Mean press probability density around target onset (peak at 0.5 s from target onset).
- D. Mean instantaneous FA and hit rates from trial onset ( $\pm$  SEM). Hit rate starts with target presentation (3-4 seconds from trial onset).
- E. Mean performances ( $\pm$  SEM) for TORC ranging from 0 to 4 seconds. Mean performances for target frequency changed to 4 kHz.

### 3.2.4. Behavioral state modulated by the context

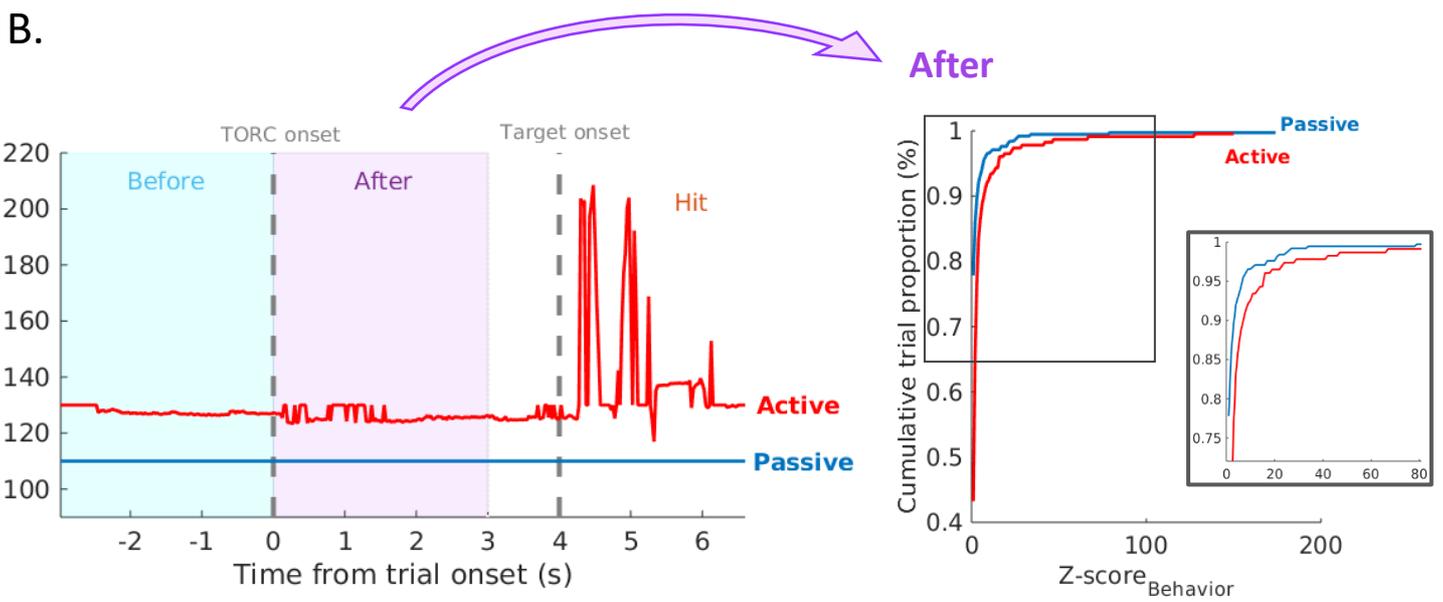
Sustained and selective attention requests the animal to adapt their brain state for behavioral arousal and demands (Harris et al., 2011). As described previously, the motor responses to targets in the active phases with fast reaction times are strong indicators of goal directed behaviors. Besides direct response analysis, variations of the pupil diameter are widely used to track the arousal level of mammals, including rodents and humans (Gee et al., 2020). This tracking was not available in our case as most of our CBA/j mice were blind at the age of recordings (Part 2. Experimental procedures) and closed their eyes when engaging in the task. As pupils and locomotion movements are reported to be correlated (McGinley et al., 2015), we investigated the existence of potential attentional behavioral markers through paws and mouth movements.

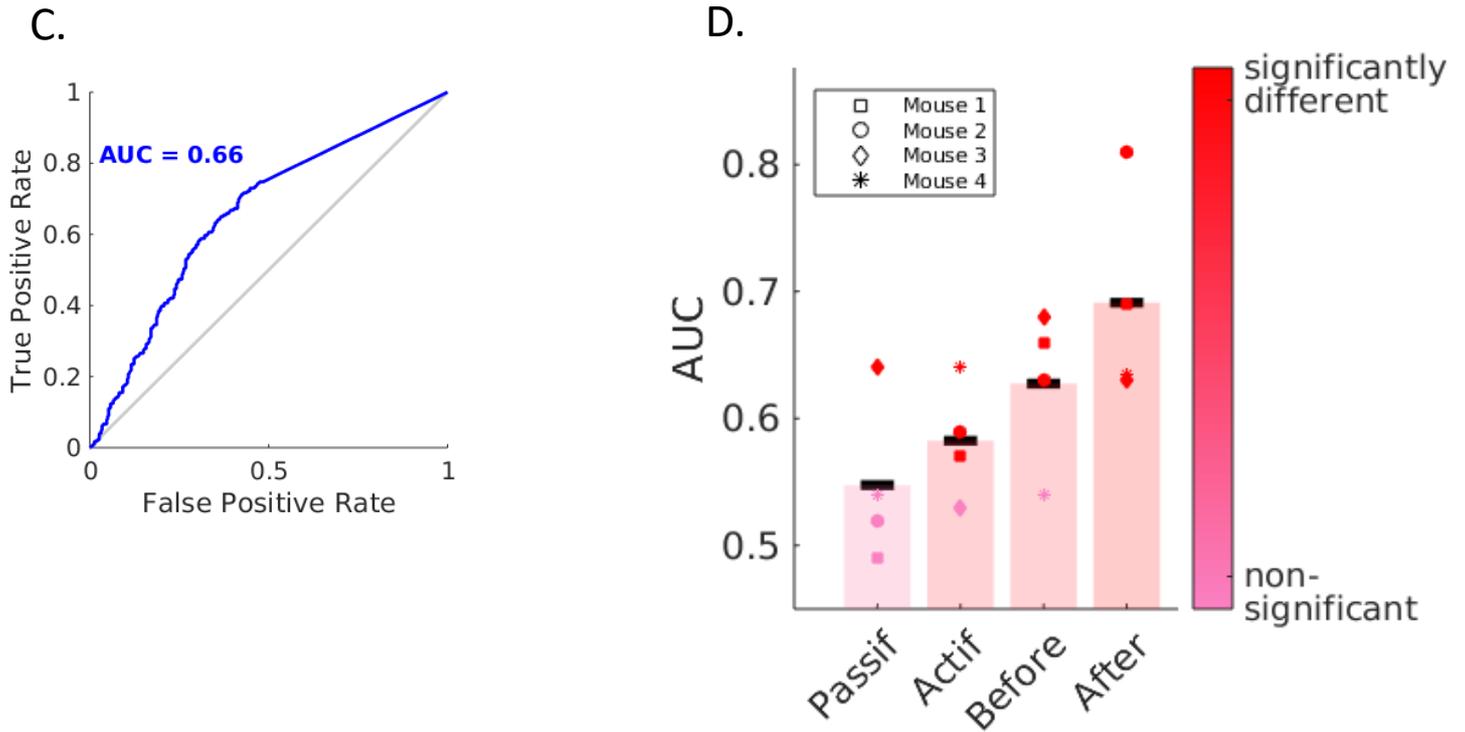
The tracking of specific body parts using the DeepLabCut deep neural network framework (Forsyth et al., 2020) captured the positions of both the paws and mouth during active and passive blocks (figures 7, 8). During hits, big movements of tens of pixels ( $42.0 \pm 8$  pixels on 2 s) were generated as the mouse was pressing the lever (fig 7.B.left) - compared to no movement observed in quiescent moments ( $1 \pm 0.5$  pixels).

As behavioral planning and sustained attention are required from the trial onset to get correct goal directed behaviors from the animal, the analysis of the behavior preceding response should already evidence attentional features. Attentive trials were considered as active hit trials and non-attentive trials as passive correct trials. Comparing the distribution of the indexes of movements before and after the sound onsets - within a given context and between different contexts – provides quantifications of the differences of behavior. Figure 7.B.right, presents a mouse's cumulative distribution of movements of the pressing paw in active and passive contexts during three seconds preceding trial onsets. The difference between distributions is evaluated by a ROC curve. The AUC associated value is 0.66 (bootstrap t-test,  $p < 0.004$ ) indicating a statistically significant difference between movement distributions in the active and passive contexts.



$$f(t) = (x_{\text{pixel}}, y_{\text{pixel}})_{\text{press paw}}$$





**Figure 7. Separable pressing paw movements between contexts indicate behavioral differences**

- Snapshot of a mouse performing the task under the microscope. The animal switches from active (left) to passive context (right). Coordinates from the pressing paw (yellow label) are extracted for each frame.
- Example of horizontal pixel coordinates of the pressing paw for an active and a passive trial for the same reference sound (left). The target onset is followed by a hit in the active context. Four conditions are studied: before (sky blue patch), after (mauve patch), active and passive (before/after). Example of cumulative proportions of the STD of movement for each correct trial during the TORC listening (right). Analysed trials provide from recordings of the same mouse (*active trials*:  $n=452$ , *passive trials*:  $n=760$ ).
- ROC curve indicating the performance of binary classifier to separate the two previous distributions. The area under the curve (AUC) is equal to 0.66 and significantly different from a random classifier (bootstrap t-test,  $p < 0.004$ ).
- Mean AUC values for each condition. Individual means are indicated by a red marker for significantly different distributions (bootstrap t-test,  $p < 0.004$ ), pink if non-significant.

Considering all tracked body parts and all comparisons of movement distributions, the lowest change appears between before and after the start of the trial in the passive context (mean AUC:  $56.1 \pm 2.5$ ). This result is consistent with the mice learnt disengagement in the passive context that leads them not to respond to sound.

The greater behavioral difference is visible during TORC listening between the active and passive contexts as the mean AUC classification value reaches  $0.67 \pm 0.05$ . This value reflects the difference of movement distribution of body parts, of which on average  $2.8 \pm 0.3$  show a significant contextual behavioral change (bootstrap *t*-test,  $p < 0.004$ , figure 8.B). These differences are visible through small body wiggles during TORC listening in active blocks. Importantly and to avoid any hit movement artefacts, the TORC listening analysis was restricted to the first three seconds from trial onset.

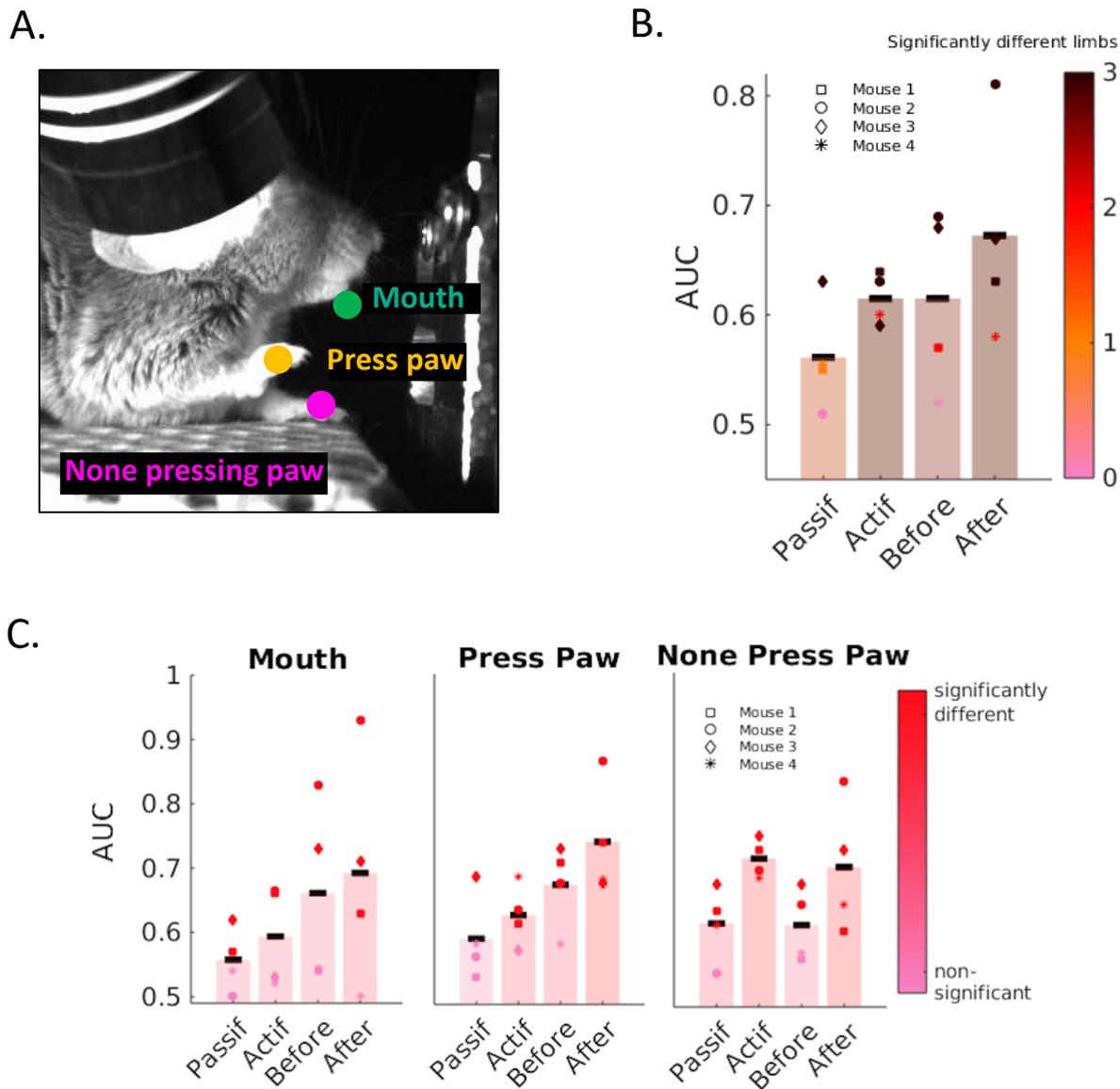
During the first three seconds preceding the start of the trial, body part movements were already largely differentiable between the two contexts, as on average  $2.0 \pm 0.7$  body parts per mouse present a significantly different distribution of indexes of movement (mean AUC:  $0.62 \pm 0.04$ ). These differences prior to trial onset confirm previous findings by which mice understood the block structure of the task (section 3.2.2), and adapted their behavioral state to respond to the active trials.

When comparing before and after the start of the trial in the active context, an average difference in the movements of  $2.5 \pm 0.5$  body parts was significant (mean AUC:  $0.62 \pm 0.01$ , bootstrap *t*-test,  $p < 0.004$ ). Therefore, the reaction of mice to the sound onset during active trials is detectable by their behavior.

In summary, within the passive context block, little variations of movements are visible as behavioral needs do not vary. In contrast, most body parts show a significant behavioral difference before the start of trials between the contexts, as a greater level of sustained attention is needed from the animal in anticipation of the trial onsets in the active context. Furthermore, within the active context, the change of behavior triggered by the TORC onset reflects a behavioral state suited to

attentive listening and a greater selective attentional load to detect the target tone. This is supported by the evidenced important differences reported between active and passive TORC listening.

Together these results indicate that levels of sustained and selective attention observed, in correct answers trials in passive and active blocks, are linked to and can be anticipated by the animal's body movements.



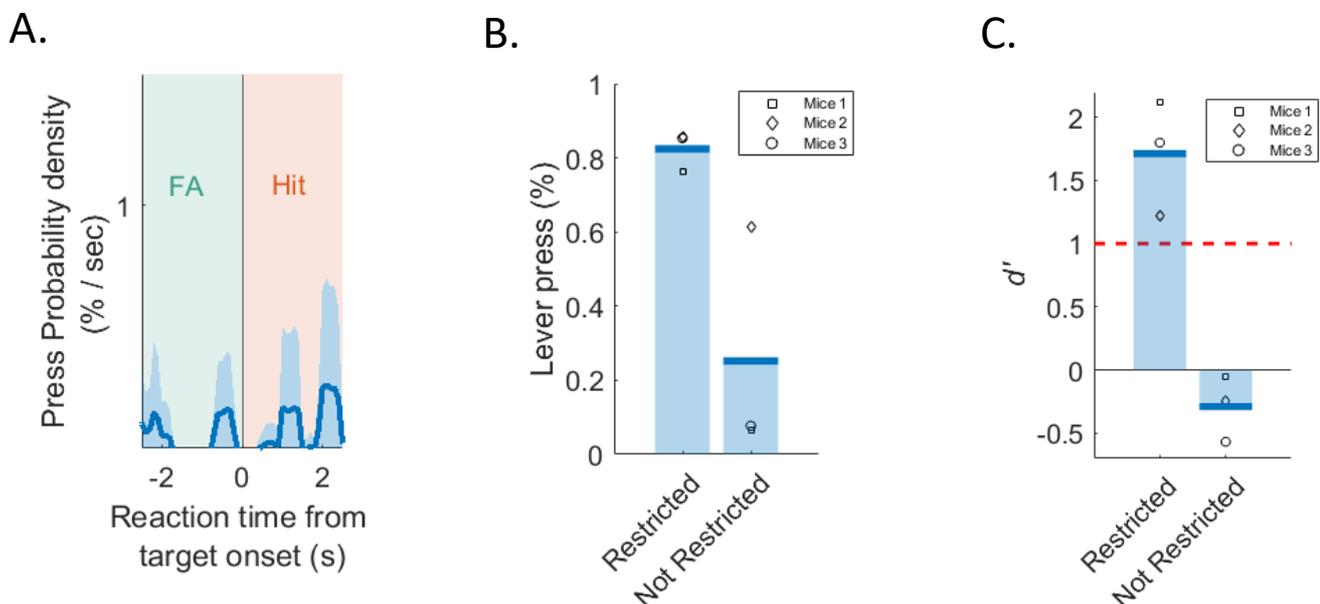
**Figure 8. Global body movements between contexts indicate different behavioral state**

- Snapshot of a mouse performing the task under the microscope. The pressing paw (yellow), the non-pressing paw (fuchsia) and mouth (green) are labelled for motion tracking.
- Mean AUC values for each condition. Individual means are indicated by a marker colour corresponding to how many body parts are significantly different in each pair of studied condition (bootstrap t-test,  $p < 0.004$ ).
- Mean AUC values for each condition and body part. Individual means are indicated by a red marker for significantly different distributions (bootstrap t-test,  $p < 0.004$ ), pink if non-significant.

### 3.3. Performance drop when mice are not motivated

Animal behavior are motivated by internal drives and thirst is reported to be one of the strongest (Allen., 2017). When reducing the level of thirst (water intake of 3-5 mL, with body weight increasing from 80 to 90% of the original weight), the engagement in the task dropped completely. As shown on figure 9.B, the lever pressing rate of  $82.4 \pm 3.1\%$  went down to  $25.2 \pm 18.2\%$  along with the  $d'$  moving from  $1.7 \pm 0.3$  to negative values (figure 9.C). Random and sparse reaction time for pressing the lever (figure 9.A) indicates the absence of goal directed behavior (Dickinson., 1994) and attentional selectivity.

When no internal drive pushes mice to differentiate context and sound, the recording of unmotivated disengaged animals offer a good control for fixed context representations. For example, context-dependent conditioned reflexes influencing cortical processing could also be evaluated (such as Palovian cues, Homayoun., 2009).



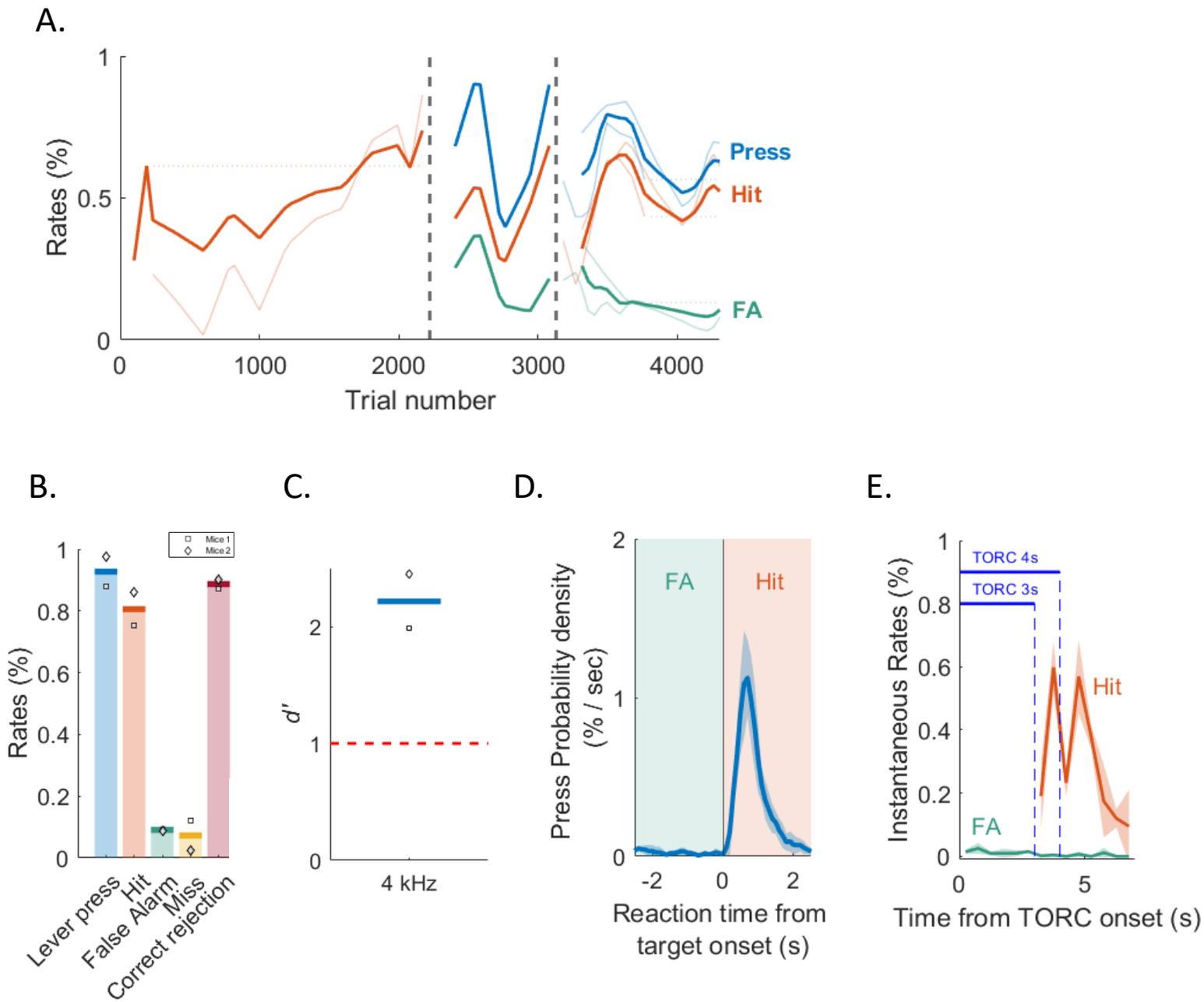
**Figure 9. Active and passive performances when mice are not water restricted**

- A. Mean press probability density around target onset,  $n=3$ .
- B. Mean lever press rate between the restricted ( $\sim 0.82$ ) and not restricted ( $\sim 0.25$ ) condition.
- C. Mean discrimination index ( $d'$ ) between the restricted and not restricted condition. Individual rates are indicated by a marker,  $n=3$ .

### **3.4. Attentional Go / noGo auditory discrimination task altering randomly two target frequencies**

#### **3.4.1. Fast learning of a second target frequency**

At the end of the recording sessions with the target frequency at 10 kHz, mice underwent a second training with a new 4 kHz target frequency. Data are presented for two mice in figure 10.A. Similar training steps were necessary as in the first training (see section 2.1) after it was observed mice did not respond instantaneously to the new target frequency (figure 6.E). The second training duration was reduced by a factor of 2.5 from ~51 to 20 days. Mice successfully learned to recognize the 4 kHz target frequency reaching  $d' > 2$  (figure 10.C). Similar performance levels and measures were reached as with the first training, which can be interpreted as in figure 4.



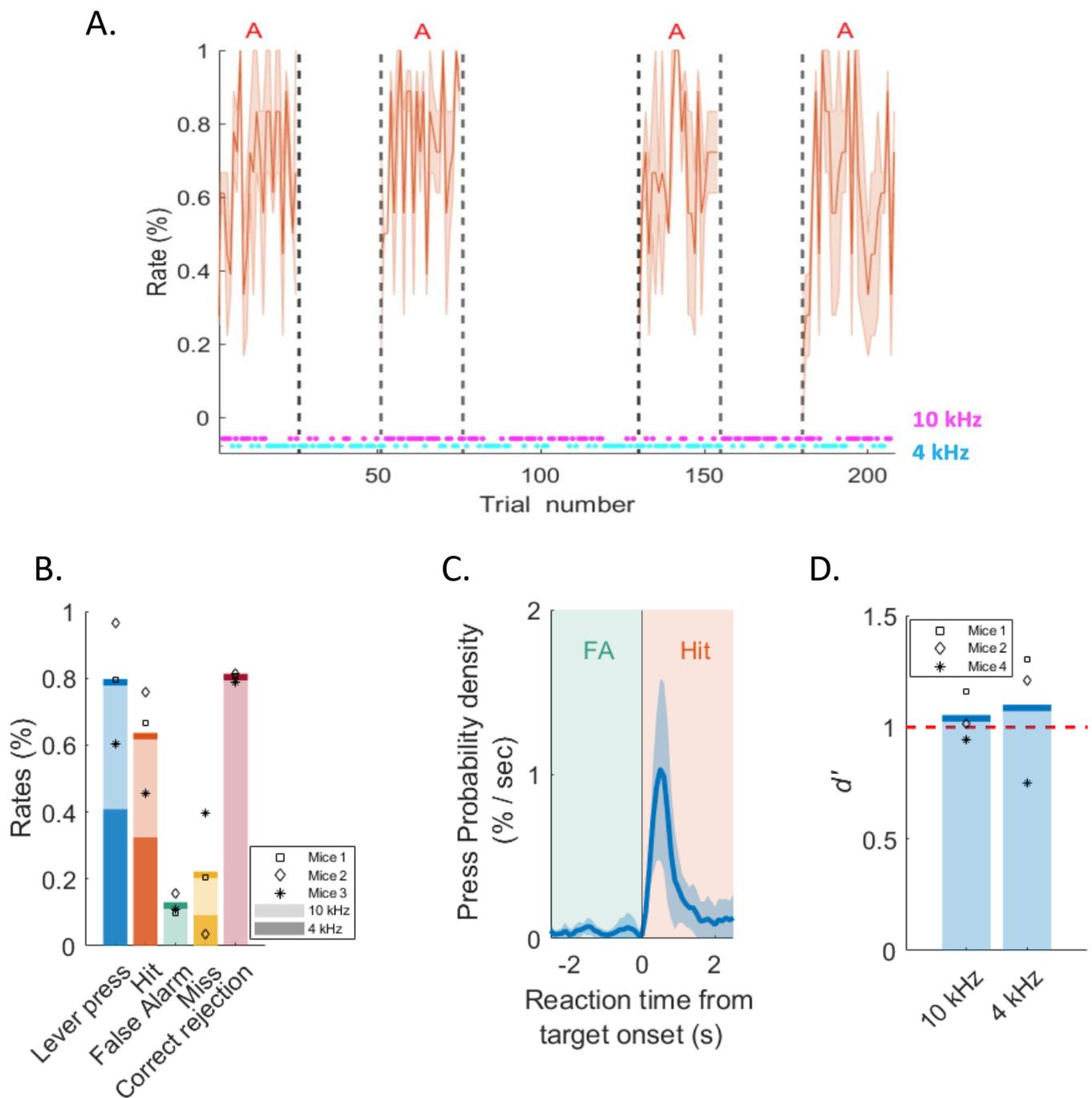
**Figure 10. Active performances for a second learning with a new target frequency at 4kHz.**

- A. Individual curves for the new target frequency at 4 kHz. Similar steps of learning as previously (figure 3),  $n=2$ .
- B. Mean rates in the active context,  $n=2$ .
- C. Mean discrimination index ( $d'$ ). Individual rates are indicated by a marker,  $n=2$ .
- D. Mean press probability density around target onset (peak at 0.7 s from target onset).
- E. Mean instantaneous FA and hit rates from trial onset ( $\pm$  SEM). Hit rate starts with target presentation (3-4 seconds from trial onset).

### **3.4.2. Equivalent target discrimination**

In order to assess the mice performance when exposed to the two target tones, they were subsequently tested with a task alternating randomly the target frequencies of 10 kHz and 4 kHz in both contexts. The results are presented for three mice in figure 11.A and show good performance with both target tones and no preference between them (figure 11.B-D).

This task is particularly interesting as alternating target tones has been reported to increase attentional load in ferrets (Fritz et al., 2003).



**Figure 11. Results for the task altering the target frequency randomly between 4 and 10 kHz.**

- A. Average hit rate (orange) for each trial during active blocks. The solid lines represent the mean rates ( $\pm$  SEM),  $n=3$ . The alternation of the target between 10 kHz and 4 kHz is represented at the bottom of the diagram for each trial.
- B. Mean rates in the active context for 10 kHz (light color bar) and 4 kHz (full color bar). Individual global rates are indicated by a marker,  $n=3$ .
- C. Mean press probability density around target onset (as previous, peak at 0.5 s from target onset).
- D. Mean discrimination index ( $d'$ ) for 10 kHz and 4 kHz. Individual rates are indicated by a marker,  $n=3$ .

#### **4. Imaging of the sensory representation of A1 during a context-changing behavioral task in mice**

##### **4.1. Passive pure tone sensory representations in A1 awake mice**

We combined two functional intrinsic imaging sessions (during the surgery and 3 weeks after, see Annex 3), with passive pure tone two-photon recordings and post-mortem A1 anatomical identification to confirm the localization of the recordings. Intrinsic imaging systematically highlights three auditory regions: A1, secondary auditory field (A2) and anterior auditory field (AFF). This robust auditory cortical organization is in line with previous topography studies in mice AC (Tsukano et al., 2016; Guo et al., 2012).

Adeno-associated virus (AAV) expressing the calcium probe GCaMP7f was injected in the caudal area of A1 being activated by low-frequency tones (Issa et al., 2014). To characterize the tuning properties of individual cells, pure tone recordings ranging from 2 kHz to 64 kHz were performed in awake passive mice with two-photon calcium imaging.

For the first second of pure tone, on average  $52 \pm 5\%$  cells revealed an increase of  $\Delta F/F$  activity (166,  $n=3$ ), consistent with previous findings (Bandyopadhyay et al., 2010). As the injection was performed in the lower frequency band of A1, the average best frequency was  $5.2 \text{ kHz} \pm 0.8 \text{ kHz}$ . Among responding cells, BFs ranging from 2 to 10 kHz represent 94 % of the cells. Only a few cells had BFs tuned to mid-frequency tones (from 10 kHz to 20 kHz: 4 cells) and high-frequency tones (from 20 kHz to 64 kHz: 6 cells).

This low-frequency tuning property shared by most of neighboring neurons in a small spatial area ( $400 \mu\text{m}$ ) is in line with the macroscopic intrinsic imaging results and A1 tonotopic organization already well described otherwise (Issa et al., 2014; Bowen 2020). In addition to their BF, cellular frequency response area can be narrow or span a large band of frequencies (Francis et al., 2018). An example cell responding to  $5.6 \pm 5.3 \text{ kHz}$  is shown on figure 12, panels B and C.

For the total duration of the sound (3 s), 58.9 % of cells had a sustained activity (189 cells,  $n=3$ ). The average sustained BF was  $6.8 \text{ kHz} \pm 0.9 \text{ kHz}$ . Among these sustained cells, BFs ranging from 2 to 10 kHz characterized 85.7% of the cells. Only 17 cells had BFs tuned to mid-frequency tones (from 10 kHz to 20 kHz) and high-frequency tones (from 20 kHz to 64 kHz, 10 cells). Unlike recordings in anesthetized animals, in awake conditions, a subset of A1 neurons sustain firing to their preferred stimuli (Wang et al., 2005; Kato et al., 2015).

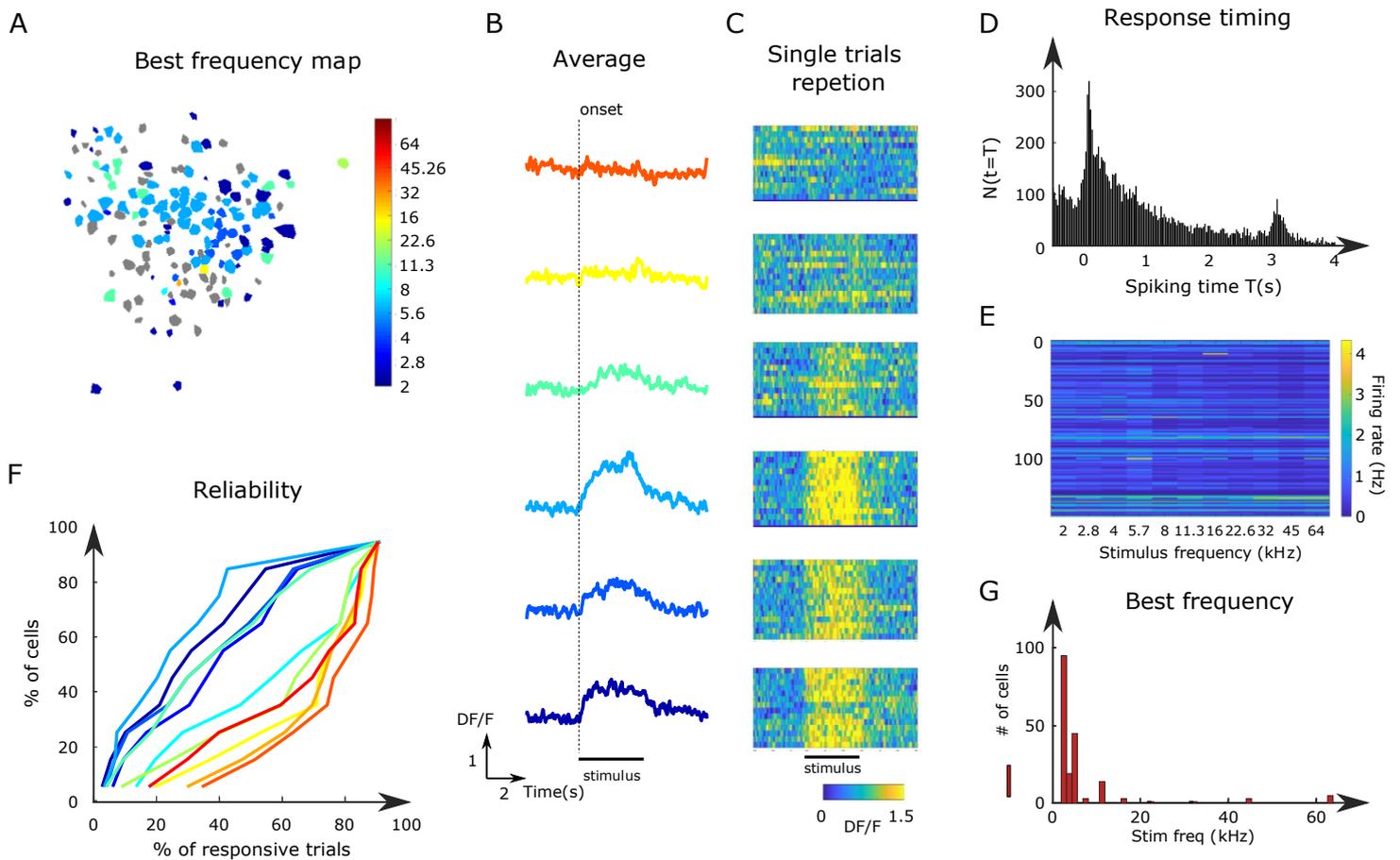
Spontaneous FR before stimulus onset was around 0.5 Hz. Spontaneous A1 activity has been previously reported to be around  $\sim 4 \text{ Hz}$  with electrical recordings (Zhou et al., 2014). This higher FR can be largely explained by the use of electrical recordings which tend to under sample cells with low FR (Hromádka et al., 2008). Offset responses were also visible in 40.1 % of cells and are part of A1's activity profiles (Bandyopadhyay et al., 2010). 13 % of offset responses were for high-frequency tones offsets (from 20 kHz to 64 kHz).

In awake mice, the latency of response peaked at 50 ms post stimuli onset as shown on figure 12.D. This quasi-instantaneous response is another functional indicator of A1 as other auditory regions show slower reaction times (Solyga et al., 2019; Bandyopadhyay et al., 2010).

The reliability of response was quantified by a reliability index reflecting the number of reliable trials for cells and frequencies. This reliability index to the set of stimuli varied from 0.23 to 0.76 depending on the frequency. The best reliability index was found for 5.657 kHz with a mean of 76% of responsive trials. The mean response reliability distribution was equal to 48% and is consistent with the literature (Bandyopadhyay et al., 2010; Francis et al., 2018).

Post-mortem histologic identification of A1 was done by locating the granular layer IV of the AC. The superimposition of the anatomic localization of A1, the functional activity maps of intrinsic imaging and the recording sites could be carried out by matching the blood vessels on the surface of the brain (see Annex 3).

Together these localization techniques and functional properties reliably confirmed that recordings were performed in A1.



**Figure 12. Passive tone-evoked neuronal responses in A1.**

- A. An example of a FOV recorded on a mouse (148 recorded cells). Cells are of the colour of their best frequencies shown on the colour bar. Grey cells are defined as non-responding (49 cells).
- B. An example of the average  $\Delta F/F$  of a cell for different frequencies indicated by the colour bar, BF = 5.6 kHz.
- C. Heat-map of single trial repetitions to frequencies presented in b. The signal strength for each time step is indicated on the colour bar.
- D. Histogram of the first spiking event for single cells and trial repetition (151 trials, 148 cells,  $n=3$ ).
- E. FR of responding cells shown for each frequency (148 cells,  $n=3$ ).
- F. Cumulative density function of responsive trials for each frequency for all cells.
- G. Histogram of best frequencies for all responding cells ( $321 \pm 20$  cells,  $n=3$ ).

## **4.2. Task engagement modulates sound representation in A1 between blocks of active listening and passive exposure**

As neural encodings depend on the behavioral response categories of hits, misses, errors (Gronskaya et al., 2019), only correct behavioral responses were included in the analysis. In the active context, this corresponds to hit trials where animals pressed on the lever in response to the target tone and to correct rejection trials in engaged sequences. Disengaged trials for the total duration of the sound constitute the passive context trials.

Spontaneous activity refers to the window of 3 s preceding trial onset. TORC activity includes the total duration of these reference sounds. In order to avoid hit movement artefacts, the target tone activity is calculated from 0 to 0.2 s after target onset.

Only responding cells to passive TORCs were included in the analysis - counted as 313 cells out of a total of 963 in 4 mice. Statistically significant differences between FR distributions were evaluated using the Student's *t*-test.

### **4.2.1. Variations of evoked-cellular activity in A1 depend on task-engagement**

When mice actively participated in the complete task as defined in part 3. "Attentional Go / no-Go auditory discrimination task", the FR distributions of responding cells were significantly different between contexts for the similar sensory environments of: spontaneous activity ( $p = 8.10^{-15}$ ), exposure to TORCs ( $p$ -value =  $1.10^{-11}$ ) and target tone sound ( $p$ -value =  $5.10^{-11}$ ).

A control condition was performed with un-motivated mice, not water deprived and ignoring the target tone during active blocks (miss rate  $75.0 \pm 14.9\%$ , 332 cells, 3 mice). No or very little statistically significant differences were found between the FR distributions of each context: spontaneous activity ( $p$ -value = 0.2), exposure to TORC ( $p$ -value = 0.7) and target tone sound ( $p$ -value = 0.01).

#### **4.2.2. Active listening suppresses cellular activities in active trials compared to passive exposure**

For each sensory environment, the mean cellular activity was reduced in the active blocks compared to the passive blocks. The average active and passive FR were calculated from the FR of the same population of cells.

#### **Spontaneous activity**

The spontaneous activity preceding the onset of trials where hit occurred was on average  $0.42 \text{ Hz} \pm 0.30 \text{ Hz}$  – in comparison to the passive trials average of  $0.54 \text{ Hz} \pm 0.30 \text{ Hz}$  (757 cells,  $n=4$ ). A statistically significant difference between FR distributions in active and passive contexts was reported ( $p$ -value =  $8.10^{-15}$ ). This difference to the mean FR could be explained as 80.78 % cells modulated their spontaneous activity depending on the context. In line with the higher average discharge rate in the passive context, 79.84 % of the context-modulated cells had higher FRs during passive trials. Only 20.16 % cells were more excited in the active context (figure 13.D).

The control condition did not show any significant difference between the FR distributions of cells ( $p$ -value = 0.2). During spontaneous activity, only 3.30 % of cells showed a significant MI between contexts.

#### **TORCs**

To begin with, within each context, the TORC response FR distributions were significantly different from previous spontaneous FR activity distributions (active  $p$ -value =  $1.10^{-24}$ ; passive  $p$ -value =  $2.10^{-30}$ ). This difference confirmed the modulation to sound of the recorded cells activities. Compared to spontaneous activity, the TORC FRs decreased for both contexts (active, reduction of  $0.31 \text{ Hz} \pm 0.02$ ; and passive, reduction of  $0.38 \text{ Hz} \pm 0.03$ ). As previously reported, white-noise like stimuli reduces the

activity of neurons in the AC (Christensen et al., 2019). A significant difference of FR distributions between contexts was shown ( $p$ -value =  $1.10^{-11}$ ). Cells responded more to TORC stimuli in the passive context, as the FR value was  $0.15 \text{ Hz} \pm 0.02 \text{ Hz}$  and on average  $0.04 \pm 0.01 \text{ Hz}$  higher than in the active context. A significant difference of activity was reported for 85.53 % of cells in response to TORCs between contexts. Among this proportion, 85.71 % of neurons were more active during the passive context. The other 14.29 % of cells enhanced their response to TORCs during the active context.

In the control condition, no significant difference was shown between the distributions of FR between contexts ( $p$ -value = 0.7). The number of cells modulating their responses between contexts dropped to 5 %.

### **Target tone**

The target tone FR was calculated from 0 to 0.2 s after target onset following the TORC broadband noise like stimuli. In both contexts, the FR distributions between the target and the TORC was significantly different indicating an effect of the target tone on cellular activity (active  $p$ -value =  $1.10^{-9}$ ; passive  $p$ -value =  $1.10^{-4}$ ). In the active context, the onset of the target tone caused a major suppression as the mean FR decreased to  $0.06 \text{ Hz} \pm 0.01 \text{ Hz}$  (reduction of 0.05 Hz compared to TORCs). This decrease was less visible in the passive context as the mean FR was  $0.15 \pm 0.02 \text{ Hz}$  during TORCs and  $0.12 \pm 0.02 \text{ Hz}$  during the target tone.

Between context, a significant difference in the distributions of FRs was statistically showed ( $p$ -value =  $5.10^{-11}$ ). 78.21 % of responding cells modulated their responses according to context of which 81.97 % were more active for target tones during the passive context.

Despite the massive FR suppression reported in the active context (mean FR rate:  $0.06 \text{ Hz} \pm 0.01 \text{ Hz}$ ), 18.03 % of context-modulated cells were more excited for the target tone during the active context. This proportion of context modulated cells preferentially responding to stimuli within active

trials is greater than for other auditory environments as 14.29 % of cells preferred TORCs in the active context.

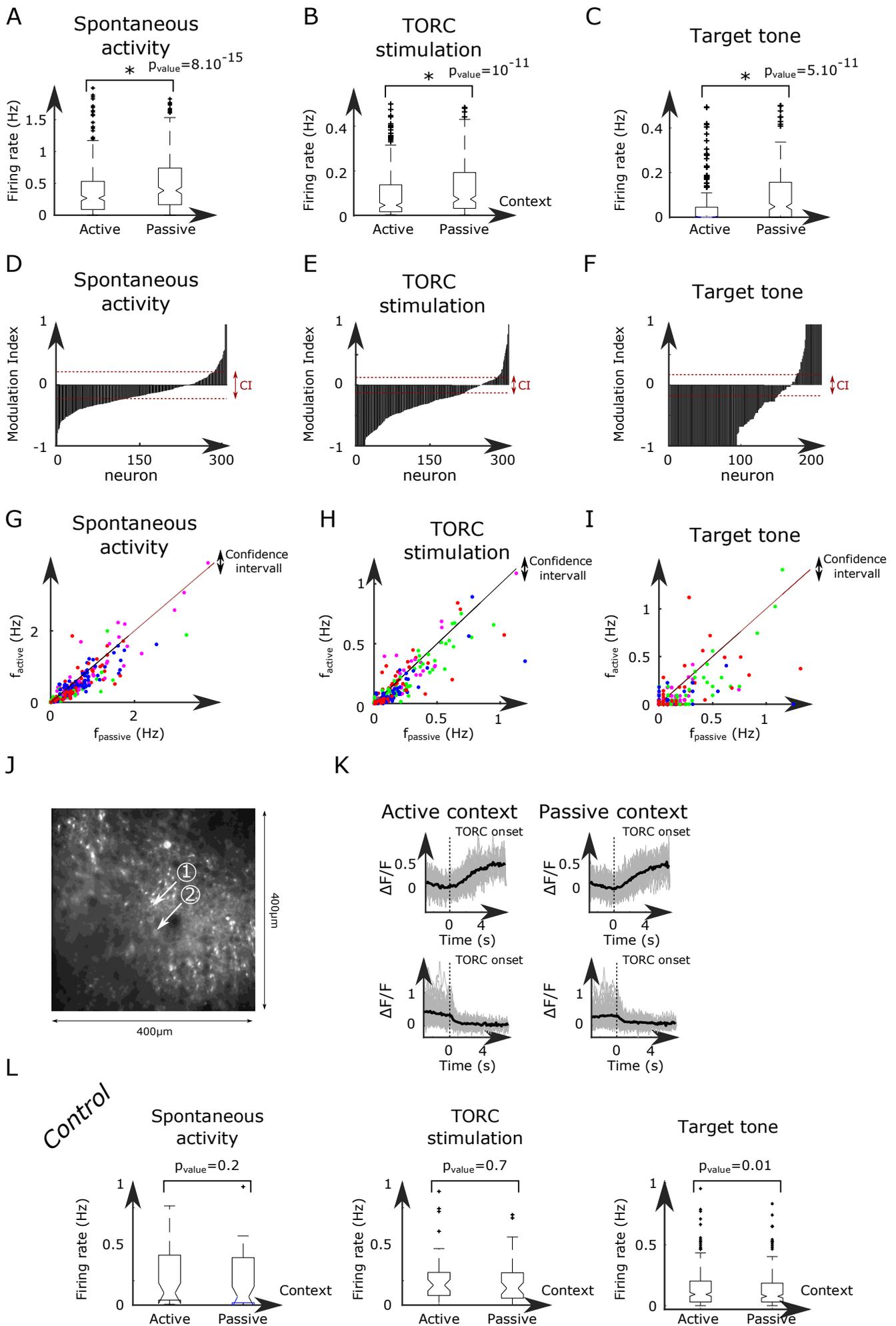
During the control condition, the proportion of cells modulated between contexts dropped to 5.00 % in response to target tone stimuli.

#### **4.2.3. A greater number of cells contrasts with the population in the active context**

Figure 13, from panels A to C, shows the outlier cells that deviate largely from the FR active and passive distributions.

During spontaneous activity, the number of outlier cells is relatively similar in the active and passive context, respectively 23 and 18 cells. The active TORC listening increases the number of outliers in the active context to 34 cells. The FR distribution of passive exposure to TORC shows 23 outliers. The onset of the pure tone has no effect on the number of outliers for the passive context (24 cells). In comparison, for pure tone detection in the active context, the number of outliers increases to 48 cells.

Together these results show a decreased average activity of A1 during the active context but with a larger sample of outlier cells. Similar small proportions of enhanced cells were also reported in other studies describing a suppression of activity in A1 during a selective attention discrimination (Kuchibholta et al., 2017; Otazu et al., 2009).



**Figure 13. The sensory activity in A1 is modified according to the behavioral response between active and passive contexts.**

A, B, C) Spontaneous, TORC and target tone FR distributions for both contexts (313 cells,  $n=4$ ). “\*” indicate  $p<0.001$ , Student  $t$ -test.

D, E, F) Spontaneous, TORC and target tone MI between the FR of contexts, (309 cells,  $n=4$ ). CI is the SD of the MI distribution calculated using randomized active/passive assignments.

G, H, I) Single-cell mean spontaneous, TORC and target tone active FR in function of the passive, (313 cells,  $n=4$ ). CI is the SD of the difference between  $FR_{Active} - FR_{Passive}$  with randomized active/passive assignments.

J) Example FOV (300  $\mu\text{m}$  x 300  $\mu\text{m}$ ) recorded during a behavioral session alternating active and passive contexts. Two cells are indicated by white arrows.

K) Mean signal examples of two previously pointed cells between the active and passive contexts. Gray region results from the superposition of individual single trials.

L) Spontaneous, TORC and target tone FR distributions for both contexts in the control condition where no behavioral difference was made between contexts (47 cells,  $n=3$ ). “\*” indicate  $p<0.001$ , Student  $t$ -test.

## 5. Discussion

### **Main contributions of the study**

When interpreting the neural dynamics underlining the attentional process during behavioral tasks, a major challenge is to determine the sensory features of interest to which the animal directs its attention to for action, as well as to characterize the animal's behavioral state during the different active and passive phases.

For our study, we developed a Go / no-Go auditory selective attention task alternating 4 blocks of active sound discrimination with 4 blocks of passive sound exposure. During active blocks, the animals based their detection of target stimuli on the frequency aspect of the sound whereby they remained neutral to the same sounds in the passive context. Once the meaning of both contexts had been well learned, a quantification of the movements of the animal's body parts enabled us to determine that its overall behavioral state is modulated according to the context.

The comparison of the neuronal activity of A1 between the active and passive blocks made it possible to highlight the difference in the processing of sound information as early as from this first step of cortical treatment. A suppression of A1's activity is observed during the active context when processing the reference and target sounds.

However, this task-engaged suppression mechanism comes with a subgroup of neurons whose activity is boosted during the active context. This boosted subset of neurons, whose activity is enhanced relative to the average activity of the population, is larger in the active context than in the passive context.

## **A complex attentional Go / no-Go auditory discrimination task to explore frequency detection in the active context**

### **Sound characteristics**

In order to interpret the neural encoding of sounds during auditory discrimination tasks, it is essential to determine the features of interest used by the animal to detect the target stimuli – to understand what the selective attention process is directed to. Indeed, auditory attentional tasks group the detection of different sound features such as: the onset or offset of a sound, its nature (pure tone or complex stimuli), its duration and its spatial location.

These differences in the importance attributed to sound characteristics may partly explain the apparent conflicting results between previous studies on selective auditory attention in rodents (see part I - Introduction). In Kato's study, the mouse has a 1 s answer period to respond to the target frequency offset, which varies from 5 to 9 seconds (Kato et al., 2015). In this task, animals performed correct responses when they recognized the offset of a sound, making it an offset-detection behavioral task.

Likewise, Francis' study rewards mice for detecting the pure tone from 0.5 to 2.5 s after its onset (Francis et al., 2018). The study also involves mice receiving a reward after 0.5 s from sound onset for correct responses in the active context – probably resulting in reward artefacts in the active context. As no reference sound has been introduced, different characteristics of the sound, such as its frequency or temporal composition, are probably not used to respond to target stimuli and identified as important by the animal.

In contrast, one of the behavioral tasks developed in Otazu's study asks the animal to detect a sustained broadband stimulus of 300 ms following a train of transient clicks (5 ms) of white noise bursts varying in number and repetition rate (Otazu et al., 2009). In this experiment, the detection of the target by the animal is based on the length of the sound.

In our study, we developed a learning based on the detection of the frequency component of the target tone. This was done through a long training protocol, consisting of three successive stages of behavioral shaping. The dependency of the response to the frequency property of the sound was tested using two controls: suppression or shortening of the reference sounds and shift of the target tone to 4 kHz (see part 3.2.3 - High engagement in the active context). These controls made it possible to validate our learning protocol that shapes the animal's behavior to detect frequency. For these reasons, our study is closer to the behavioral task set up in Kuchibholta's study, who also introduced a reference sound of one octave away from the target sound (Kuchibholta et al., 2017). During learning, the animal is only reinforced for the target tone which teaches it to adopt a single-tone strategy detection. This similarity of single-tone strategy detection with our study makes the neural results comparable. In addition, this task structure is the most suitable for selective attention studies which tries to understand the difference of responses to stimuli of interest compared to irrelevant stimuli.

### **Training and learning**

Moreover, the complexity of the task and the attentional load required for good performance is also a factor responsible for a modulation of cortical engagement and responses. Indeed, in previous head-fixed auditory tasks performed on mice, animal's expertise is reached after training over 4 to 8 days (Kato et al., 2015) or 5 to 15 days (Kuchibholta et al., 2017).

The task developed in our study requires a training of 18 to 51 days to obtain good performance in the active context. We would highlight two major differences and possible explanations of this longer learning period compared to the studies cited above. The first is the use of a lever to record responses, as the training starts with teaching the mouse the motor action allowing to press on the pedal to obtain a reward. This phase lasts about 25% of the overall learning time for the active phase, which indicates it is not the only reason responsible for a longer learning time.

Actually, it is the use of complex sounds modulated temporally and in their frequency composition (whose frequency range encompasses the 10 kHz target tone frequency) that creates ambiguous distractors to the target stimuli and complexifies the discrimination of the target tone. This complexity of discrimination has been studied by Ceballo, who showed that learning to discriminate two pure tones was on average 5 times faster than to achieve discrimination between a pure tone and a frequency-modulated sound (Ceballo et al., 2019).

In this study, once both tasks were correctly learnt, the silencing of the auditory cortex with optogenetic tools led to a decrease of performances to chance level (50%) for the discrimination between the pure tone and a frequency-modulated sound - while it was only responsible for a drop of 10% for the discrimination between two pure tones.

### **Other sound features**

In addition to the use of complex sounds, the reference sounds had varying time lengths and asked animals to produce inhibitory control for long time periods before responding to the target (from 3 to 7 s). These periods of reference sound listening are long compared to previous studies (1.25 s in Kuchiboltha et al., 2017).

## **Behavioral state change between the active and passive contexts**

### **Response in different contexts**

The video recording of the task enabled our study to evidence the good learning of the passive context. The absence of response to sounds in the passive context was completely voluntary on the part of the mouse who took approximately one week to remain quiescent during exposure to sound. Among the previous cited studies, only one mentions this point and provides some evidence of passive behavior in the passive context (Kuchiboltha et al., 2017) - which is an essential point when studying differences between active listening and passive exposure to sound.

Moreover, a millimetric quantification of the movements of the animal's body parts enabled us to determine that the overall behavioral state of the animal is modulated according to the context and the moment in the trial. These movements (wiggles observable by eye) are correlated with the required attentional load and vary in this order: the passive blocks, the activity preceding the 'trial' in the active blocks and the attentive listening of the reference sounds during the active context. This measure provides a second quantification evidencing that the task has been learned by the animal who modulates its behavior according to the active and passive contexts.

### **Selective and non-selective attention**

This leads to the following subsequent question: are the differences observed during engagement in a behavioral task due to differences in the arousal level or sustained attention between contexts? Or are these differences specific to selective auditory attention? To try to disentangle selective and non-selective attention effects, Otazu set up an intermodal auditory task in which he compared sound-evoked responses elicited in the auditory cortex when the animal was exposed to reference and target tones in three different contexts: passive, auditory-engaged and olfactory-engaged task (Otazu et al, 2009). This intermodal task allows to compare the passive encoding of sound

as the attention is not directed towards auditory targets, while the animal has a similar level of arousal and sustained attention as it responds to olfactory stimuli.

However, these intermodal tasks can introduce another bias: given that the animal's attention is directed towards olfactory stimuli, the integration of other sensory information can be filtered-out (Mozolic et al., 2008). To overcome these challenges, we determined to set up an auditory task alternating the target frequency according to the context. Indeed, during a first training the animal associates the target with 10 kHz and during a second one the animal associates the target with 4 kHz. At the end of the two trainings, the animal responds equally well to both frequencies (see part 3.4.2 - Equivalent target discrimination).

In order to progress further and complete the development of this task, it would be interesting to add a context learning whereby in one context, the animal learns to respond to 10 kHz and in a second setting, the animal learns to respond to 4 kHz. It would make it possible to study the selective auditory attention directed towards a given frequency while being decoupled from the differences in attention due to changes in task engagement or type of modality of interest.

### **Active listening suppresses responses in the primary auditory cortex compared to passive exposure to the same sounds**

In our Go / No-Go auditory task, a suppression of A1 activity is observed during the active context compared to the passive context - and this during the whole duration of the active trials: spontaneous activity, listening to the reference sounds and to the target tone. Engagement in sound discrimination reduced the number of spikes used for encoding stimuli in the auditory cortex for both target and non-target sounds. This suppression of evoked-responses was also found in Otazu's and Kuchiboltha's studies, for both the target and non-target tones.

By recording the auditory thalamus in two performing rats, Otazu also shows that the thalamic spontaneous activity is increased during the engaged condition compared to the passive

context while evoked thalamic responses to the target and reference sounds are equivalent between contexts (Otazu, 2009). These results suggest that a synaptic depression of the thalamocortical inputs could be partly responsible for the suppression of cortical evoked responses (Abbott et al., 1997). The increased spontaneous thalamic firing rate preceding sound onset leads thalamocortical synapses to be in a more “depleted” state, exhibiting decreased synaptic release probability. This would explain the relatively depressed evoked cortical responses in the active context. In the layer II-III of the auditory cortex of mice, thalamocortical inputs have been reported to directly innervate excitatory cells and PV neurons (Ji et al., 2015).

In Kuchiboltha’s study (Kuchiboltha et al., 2007), targeting the expression of the calcium probe to specific subsets of cells provided to identify sub-groups among the recorded cells (excitatory cells, SOM, VIP and PV interneurons). Excitatory and VIP neurons had a passive preferring whereas PV and SOM had an active preferring. During engaged blocks, optogenetic suppression of PV and SOM neurons increased excitatory outputs whereas the suppression of VIP interneurons led to decreased excitatory activity - as they are involved in disinhibition. The suppression of any of these types of interneurons during behavior led to impaired performances, indicating their important ‘regulatory’ role. In the auditory cortex, cholinergic agonist depolarizes mainly interneurons, suggesting that acetylcholine could be an important neuromodulator for interneurons context-dependent modulations.

Another set of experiments showed the importance of the activation of cholinergic projections coming from the nucleus basalis during active blocks. When applying atropine (muscarinic acetylcholine receptors), performances were impaired during behavior. Taken in their entirety, cholinergic projections drive engagement by targeting mainly interneurons - which results into an overall suppression of responses in the auditory cortex. These effects are contrary to those of cholinergic stimulation of the visual cortex, known to trigger an increase in cortical activity (Goard et al., 2009). A similar increase in cortical activity is also observed during engagement in visual attentive tasks (Blaser et al., 1999).

These top-down and bottom-up suppression mechanisms could explain the suppression in the active context described in our study and pave the way for future research using the tools and protocols developed to run the experiments.

The auditory suppression during active blocks compared to passive blocks, was also found in Otazu's intermodal task. Auditory evoked cortical suppression was similar during auditory-engaged and olfactory-engaged blocks.

This suppression mechanism could be the non-selective mechanism of attention where the activity of neurons irrelevant to the task is reduced. In addition, a second mechanism could coexist and boost the activity of neurons important in target representation.

### **Active listening boosts the activity of a subset of neurons compared to passive listening**

Our results show that this task-engaged suppression mechanism comes with a subgroup of neurons whose activity is boosted during the active context. This boosted subset of neurons, whose activity is enhanced relative to the average activity of the population, is larger in the active context than in the passive context - especially for sound-evoked cortical responses to target and TORC reference sounds. Similar small proportions of enhanced cells were also recorded in the previous studies reporting cortical suppression in A1 (Kuchibholta et al., 2017; Otazu et al., 2009).

Indeed, in Otazu's auditory discrimination task, a minority of single units showed an increase in their evoked activity (8/32 units). The sparseness of neural representation of sound was quantified in term of kurtosis of the firing rate distribution between contexts. This kurtosis measurement was of  $8.4 \pm 2.4$  in the engaged condition vs.  $4.1 \pm 0.7$  in the passive condition, indicating that the stimulus representation was sparser in the engaged condition. Unfortunately, the measurements of the kurtosis of the firing rate and the number of cells activated was not reported during the intermodal task where the animal is engaged in olfactory blocks and passively exposed to sound.

In Kuchibholta's study, some neurons also showed an increase in activity during active blocks compared to passive ones for both the target and non-target sounds (92/227 target-responsive cells and 64/210 foil-responsive cells). In proportion, the modulation to the target and non-target by behavioral contexts was similar and therefore not significant.

Together, our study and these previous results show a 'concentration' of stimuli-evoked activity in a smaller proportion of neurons during engaged blocks compared to the passive context. This boosted subset of neurons could reflect selective attentional effects sending precise appropriate information to a wide range of targets such as motor neurons to trigger appropriate actions in the active context.

### **Frequency tuning of neurons and rapid change according to the context**

Kuchibholta's study shows that the neurons whose best frequency was closer to the target or foil were more likely to be suppressed during active behavior. In contrast, neurons with best frequency tuning  $\geq 0.25$  octaves from the sounds were more likely to have enhanced responses during the active contexts (Kuchibholta et al., 2017). In Otazu's study, no correlation was found between the best frequency selectivity to the click rate and the neuronal modulation index between contexts (Otazu et al., 2009).

Moreover, in Fritz's study where ferrets had to discriminate pure tones compared to TORCs - in a similar way to our study -, the analysis of these complex sounds showed that the neurons responded preferentially to the target frequency nested in TORC during the active context compared to the passive context. These facilitative changes were found for most A1 neurons (>70%) suggesting a filtering mechanism that enhances responses to the target tones during behavior (Fritz et al., 2003). It should be noted that for several units, the enhanced representation of the target frequency and the maximum facilitative changes stood outside of their initial STRF. These results were obtained when comparing active blocks to previous passive exposures to sound. Among these cells, only 1/3 (13/39)

reverted their STRF back to their original shape in post-behavior passive blocks. In fact, most units showed “sensory memory” to the target tone after active behavior that could last for several hours.

As these referenced studies suggest, enhanced representation to the target stimuli would be seen in regions tangent to the best frequency tuning area of the target. In our study, we recorded A1’s low-frequency band area responding preferentially to 4-5 kHz with a target placed of 10 kHz. Moreover, our behavioral task started with an active block and alternated 4 active and 4 passive blocks that we respectively analyzed together, erasing behavioral changes that would not adapt quickly to context modulation.

## Further Research

This massive suppression accompanied by more strongly excited cells in the active contexts could be the signature of a selective attention mechanism, responding specifically to the frequency of the target tone. In order to verify this filtering of the pure tone during the active context, further analysis of TORCs should allow to evidence a change in the response of cellular activity at the target frequency depending on the context. Comparable studies in ferrets have shown this increased filtering of the target frequency during active context (Fritz et al., 2003). This analysis would allow to quantify the proportion of cells modulated according to the response to the target frequency in the different contexts. It would also be valuable to observe whether the cells involved in the frequency response of interest are the same more active ones during spontaneous activity - as well as during the onset of the target tone in the active context.

In addition, the study of the cellular network in different contexts is reported to be task-relevant (Mitchell et al., 2009). The response patterns between cells, the pairwise and noise correlation within the network would enable to understand its dynamics better. It would be interesting to learn more on the role played by the cells increasing their response in an active context within the selective attention network.

Furthermore, future experiments using the same selective attention task would make possible the quantification of the modulations of the different cortical layers, as well as their variations dependent on the frequency. This would lead to understand better the involvement of the different layers of A1 during selective attention, in particular characterizing the cortical processing between layer IV and the layers II / III (O'Connell et al., 2014) and the desynchronized activity generation in the deep layers V / VI (Sakata et al., 2009). These researches would further explain the mechanisms necessary for the selective attention process.

Further experiments using transgenic mouse lines would also allow a detailed unpicking of the cellular mechanisms involved in selective attention (pyramidal cells, PV, SOM, VIP) as well as their role in the different phases of behavior (anticipation, motor response, reward, sound discrimination).

In particular, replicating our experiments with transgenic mouse lines would allow to determine the excitatory or inhibitory nature of the “outlier” neurons described in our study. Our results suggest a correlation between a lower level of global activity in the population with a greater number and enhanced activity of outlier cells. This focused “concentration” of activity with a small fraction of cells mechanically increases their relative contribution to the sound processing in A1. This phenomenon is more pronounced in the active versus passive context and during the target versus the TORC presentations. Understanding the excitatory or inhibitory nature of these cells seems a key next step to further characterize the auditory selective process.

Moreover, additional cellular characterization could also lead to identify the cells in A1 that project to other brain regions for task-relevant executive functions. The greater number of outlier cells in the active context, especially during the target presentation, could trigger a filtering mechanism initiating the response actions.

Such experiments were obviously beyond the original scope of our study, but are now enabled through further uses of the equipment, tools and protocols we have developed throughout this research progression with these further steps in mind.

## 6. Bibliography

Adcock, R. Alison, et al. "Reward-motivated learning: mesolimbic activation precedes memory formation." *Neuron* 50.3 (2006): 507-517.

Akbari, Hassan, et al. "Towards reconstructing intelligible speech from the human auditory cortex." *Scientific reports* 9.1 (2019): 1-12.

Alahyane, Nadia, et al. "Developmental improvements in voluntary control of behavior: effect of preparation in the fronto-parietal network?." *NeuroImage* 98 (2014): 103-117.

Alho, Kimmo, et al. "Selective tuning of the left and right auditory cortices during spatially directed attention." *Cognitive Brain Research* 7.3 (1999): 335-341.

Allen, William E., et al. "Thirst-associated preoptic neurons encode an aversive motivational drive." *Science* 357.6356 (2017): 1149-1155.

Armstrong, Katherine M., Jamie K. Fitzgerald, and Tirin Moore. "Changes in visual receptive fields with microstimulation of frontal cortex." *Neuron* 50.5 (2006): 791-798.

Atencio, Craig A., Tatyana O. Sharpee, and Christoph E. Schreiner. "Receptive field dimensionality increases from the auditory midbrain to cortex." *Journal of neurophysiology* 107.10 (2012): 2594-2603.

Ballard, Joan C. "Computerized assessment of sustained attention: a review of factors affecting vigilance performance." *Journal of clinical and experimental neuropsychology* 18.6 (1996): 843-863.

Bandyopadhyay, Sharba, Shihab A. Shamma, and Patrick O. Kanold. "Dichotomy of functional organization in the mouse auditory cortex." *Nature neuroscience* 13.3 (2010): 361-368.

Bari, Andrea, Jeffrey W. Dalley, and Trevor W. Robbins. "The application of the 5-choice serial reaction time task for the assessment of visual attentional processes and impulse control in rats." *Nature protocols* 3.5 (2008): 759-767.

Barkley, Russell A. "Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD." *Psychological bulletin* 121.1 (1997): 65.

Bedard, Anne-Claude, et al. "Methylphenidate improves visual-spatial memory in children with attention-deficit/hyperactivity disorder." *Journal of the American Academy of Child & Adolescent Psychiatry* 43.3 (2004): 260-268.

Berry, Anne S., Martin Sarter, and Cindy Lustig. "Distinct frontoparietal networks underlying attentional effort and cognitive control." *Journal of Cognitive Neuroscience* 29.7 (2017): 1212-1225.

Blaser, Erik, George Sperling, and Zhong-Lin Lu. "Measuring the amplification of attention." *Proceedings of the National Academy of Sciences* 96.20 (1999): 11681-11686.

Blokland, Arjan. "Reaction time responding in rats." *Neuroscience & Biobehavioral Reviews* 22.6 (1998): 847-864.

Bowen, Zac, Daniel E. Winkowski, and Patrick O. Kanold. "Functional organization of mouse primary auditory cortex in adult C57BL/6 and F1 (CBAx57) mice." *Scientific reports* 10.1 (2020): 1-14.

Brodeur, Darlene A., and Miranda Pond. "The development of selective attention in children with attention deficit hyperactivity disorder." *Journal of Abnormal Child Psychology* 29.3 (2001): 229-239.

Bushnell, Philip J., and Barbara J. Strupp. "Assessing attention in rodents." (2009).

- Calmels, Claire, Christelle Berthoumieux, and Fabienne Fabienne d'Arripe-Longueville. "Effects of an imagery training program on selective attention of national softball players." *The Sport Psychologist* 18.3 (2004): 272-296.
- Castro-Alamancos, Manuel A. "Absence of rapid sensory adaptation in neocortex during information processing states." *Neuron* 41.3 (2004): 455-464.
- Carruthers, Isaac M., et al. "Emergence of invariant representation of vocalizations in the auditory cortex." *Journal of neurophysiology* 114.5 (2015): 2726-2740.
- Carli, M., et al. "Effects of lesions to ascending noradrenergic neurones on performance of a 5-choice serial reaction task in rats; implications for theories of dorsal noradrenergic bundle function based on selective attention and arousal." *Behavioural brain research* 9.3 (1983): 361-380.
- Ceballo, Sebastian, et al. "Cortical recruitment determines learning dynamics and strategy." *Nature communications* 10.1 (2019): 1-12.
- Chalk, Matthew, et al. "Attention reduces stimulus-driven gamma frequency oscillations and spike field coherence in V1." *Neuron* 66.1 (2010): 114-125.
- Chechik, Gal, et al. "Reduction of information redundancy in the ascending auditory pathway." *Neuron* 51.3 (2006): 359-368.
- Chen, Jerry L., et al. "Online correction of licking-induced brain motion during two-photon imaging with a tunable lens." *The Journal of physiology* 591.19 (2013): 4689-4698.
- Cheng-yu, T. Li, Mu-ming Poo, and Yang Dan. "Burst spiking of a single cortical neuron modifies global brain state." *Science* 324.5927 (2009): 643-646.
- Christensen, Rasmus Kordt, et al. "White noise background improves tone discrimination by suppressing cortical tuning curves." *Cell reports* 29.7 (2019): 2041-2053.
- Ciria, Alejandra, et al. "Predictive processing in cognitive robotics: a review." *Neural Computation* 33.5 (2021): 1402-1432.
- Cohen, Jeremiah Y., et al. "On the origin of event-related potentials indexing covert attentional selection during visual search." *Journal of Neurophysiology* 102.4 (2009): 2375-2386.
- Corbetta, Maurizio, and Gordon L. Shulman. "Control of goal-directed and stimulus-driven attention in the brain." *Nature reviews neuroscience* 3.3 (2002): 201-215.
- Cooke, James E., et al. "Contrast gain control in mouse auditory cortex." *Journal of neurophysiology* 120.4 (2018): 1872-1884.
- Dickinson, Anthony, and Bernard Balleine. "Motivational control of goal-directed action." *Animal Learning & Behavior* 22.1 (1994): 1-18.
- Domenger, Dorothee, and Rainer KW Schwarting. "Sequential behavior in the rat: role of skill and attention." *Experimental brain research* 182.2 (2007): 223-231.
- Dipoppa, Mario, et al. "Vision and locomotion shape the interactions between neuron types in mouse visual cortex." *Neuron* 98.3 (2018): 602-615.
- Enns, James T., and Sharon Cameron. "Selective attention in young children: The relations between visual search, filtering, and priming." *Journal of experimental child psychology* 44.1 (1987): 38-63.

Forys, Brandon J., et al. "Real-time selective markerless tracking of forepaws of head fixed mice using deep neural networks." *Eneuro* 7.3 (2020).

Foucher, J. R., H. Otzenberger, and D. Gounot. "Where arousal meets attention: a simultaneous fMRI and EEG recording study." *Neuroimage* 22.2 (2004): 688-697.

Francis, Nikolas A., et al. "Small networks encode decision-making in primary auditory cortex." *Neuron* 97.4 (2018): 885-897.

Fritz, Jonathan, et al. "Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex." *Nature neuroscience* 6.11 (2003): 1216-1223.

Fujiwara, Naohito, et al. "Role of the primary auditory cortex in auditory selective attention studied by whole-head neuromagnetometer." *Cognitive brain research* 7.2 (1998): 99-109.

Fuller, Patrick, et al. "Reassessment of the structural basis of the ascending arousal system." *Journal of Comparative Neurology* 519.5 (2011): 933-956.

Galambos, Robert, and Hallowell Davis. "The response of single auditory-nerve fibers to acoustic stimulation." *Journal of neurophysiology* 6.1 (1943): 39-57.

de Gee, Jan Willem, et al. "Pupil-linked phasic arousal predicts a reduction of choice bias across species and decision domains." *Elife* 9 (2020): e54014.

Goard, Michael, and Yang Dan. "Basal forebrain activation enhances cortical coding of natural scenes." *Nature neuroscience* 12.11 (2009): 1444-1449.

Gregoriou, Georgia G., et al. "High-frequency, long-range coupling between prefrontal and visual cortex during attention." *science* 324.5931 (2009): 1207-1210.

Grizenko, Natalie, et al. "Maternal stress during pregnancy, ADHD symptomatology in children and genotype: gene-environment interaction." *Journal of the Canadian Academy of Child and Adolescent Psychiatry* 21.1 (2012): 9.

Gronskaya, Elena, and Wolfger von der Behrens. "Evoked response strength in primary auditory cortex predicts performance in a spectro-spatial discrimination task in rats." *Journal of Neuroscience* 39.31 (2019): 6108-6121.

Gruber, Aaron J., Rifat J. Hussain, and Patricio O'Donnell. "The nucleus accumbens: a switchboard for goal-directed behaviors." *PLoS one* 4.4 (2009): e5062.

Guo, Wei, et al. "Robustness of cortical topography across fields, laminae, anesthetic states, and neurophysiological signal types." *Journal of Neuroscience* 32.27 (2012): 9159-9172.

Harpin, Valerie A. "The effect of ADHD on the life of an individual, their family, and community from preschool to adult life." *Archives of disease in childhood* 90.suppl 1 (2005): i2-i7.

Harris, Kenneth D., and Alexander Thiele. "Cortical state and attention." *Nature reviews neuroscience* 12.9 (2011): 509-523.

Hart, Eric, and Alexander C. Huk. "Recurrent circuit dynamics underlie persistent activity in the macaque frontoparietal network." *Elife* 9 (2020): e52460.

Hillyard, Steven A., et al. "Electrical signs of selective attention in the human brain." *Science* 182.4108 (1973): 177-180.

HOCHERMAN, SHRAGA, and RAZ YIRMIYA. "Neuronal activity in the medial geniculate nucleus and in the auditory cortex of the rhesus monkey reflects signal anticipation." *Brain* 113.6 (1990): 1707-1720.

Hocherman, S., et al. "Evoked unit activity in auditory cortex of monkeys performing a selective attention task." *Brain research* 117.1 (1976): 51-68.

Hoffman, Kurt Leroy. *Modeling neuropsychiatric disorders in laboratory animals*. Woodhead Publishing, 2015.

Homayoun, Houman, and Bitá Moghaddam. "Differential representation of Pavlovian–instrumental transfer by prefrontal cortex subregions and striatum." *European Journal of Neuroscience* 29.7 (2009): 1461-1476.

Hommel, Bernhard, et al. "No one knows what attention is." *Attention, Perception, & Psychophysics* 81.7 (2019): 2288-2303.

Hromádka, Tomáš, Michael R. DeWeese, and Anthony M. Zador. "Sparse representation of sounds in the unanesthetized auditory cortex." *PLoS biology* 6.1 (2008): e16.

Issa, John B., et al. "Multiscale optical Ca<sup>2+</sup> imaging of tonal organization in mouse auditory cortex." *Neuron* 83.4 (2014): 944-959.

Johnson, Don H. "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones." *The Journal of the Acoustical Society of America* 68.4 (1980): 1115-1122.

Juczewski, Konrad, et al. "Stress and behavioral correlates in the head-fixed method: stress measurements, habituation dynamics, locomotion, and motor-skill learning in mice." *Scientific reports* 10.1 (2020): 1-19.

Kanold, Patrick O., Israel Nelken, and Daniel B. Polley. "Local versus global scales of organization in auditory cortex." *Trends in neurosciences* 37.9 (2014): 502-510.

Kato, Hiroyuki K., Shea N. Gillet, and Jeffrey S. Isaacson. "Flexible sensory representations in auditory cortex driven by behavioral relevance." *Neuron* 88.5 (2015): 1027-1039.

Khayat, Paul S., Robert Niebergall, and Julio C. Martinez-Trujillo. "Frequency-dependent attentional modulation of local field potential signals in macaque area MT." *Journal of neuroscience* 30.20 (2010): 7037-7048.

Klein, David J., et al. "Robust spectrotemporal reverse correlation for the auditory system: optimizing stimulus design." *Journal of computational neuroscience* 9.1 (2000): 85-111.

Klein, David J., et al. "Stimulus-invariant processing and spectrotemporal reverse correlation in primary auditory cortex." *Journal of computational neuroscience* 20.2 (2006): 111-136.

Knutson, Brian, et al. "Anticipation of increasing monetary reward selectively recruits nucleus accumbens." *Journal of Neuroscience* 21.16 (2001): RC159-RC159.

Kovalzon, Vladimir M. "Ascending reticular activating system of the brain." *Translational Neuroscience and Clinics* 2.4 (2016): 275-285.

Kuchibhotla, Kishore, and Brice Bathellier. "Neural encoding of sensory and behavioral complexity in the auditory cortex." *Current opinion in neurobiology* 52 (2018): 65-71.

Kuchibhotla, Kishore V., et al. "Parallel processing by cortical inhibition enables context-dependent behavior." *Nature neuroscience* 20.1 (2017): 62-71.

Lane, David M., and Deborah A. Pearson. "The development of selective attention." *Merrill-Palmer Quarterly (1982-)* (1982): 317-337.

de Lecea, Luis, Matthew E. Carter, and Antoine Adamantidis. "Shining light on wakefulness and arousal." *Biological psychiatry* 71.12 (2012): 1046-1052.

Lee, Maan Gee, et al. "Cholinergic basal forebrain neurons burst with theta during waking and paradoxical sleep." *Journal of Neuroscience* 25.17 (2005): 4365-4369.

Lee, Seung-Hee, and Yang Dan. "Neuromodulation of brain states." *Neuron* 76.1 (2012): 209-222.

Lindsay, Grace W. "Attention in psychology, neuroscience, and machine learning." *Frontiers in computational neuroscience* 14 (2020): 29.

Liu, Shi Tong, et al. "Optimal features for auditory categorization." *Nature communications* 10.1 (2019): 1-14.

Luck, Steven J., et al. "Bridging the gap between monkey neurophysiology and human perception: An ambiguity resolution theory of visual selective attention." *Cognitive psychology* 33.1 (1997): 64-87.

Mackworth, Jane F. "Performance decrement in vigilance, threshold, and high-speed perceptual motor tasks." *Canadian Journal of Psychology/Revue canadienne de psychologie* 18.3 (1964): 209.

Makeig, Scott, and Tzyy-Ping Jung. "Tonic, phasic, and transient EEG correlates of auditory awareness in drowsiness." *cognitive brain research* 4.1 (1996): 15-25.

Makeig, Scott, Tzyy-Ping Jung, and Terrence J. Sejnowski. "Awareness during drowsiness: Dynamics and electrophysiological correlates." *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale* 54.4 (2000): 266.

Malmierca, Manuel S. "Auditory system." *The rat nervous system*. Academic Press, 2015. 865-946.

Mathis, Alexander, et al. "DeepLabCut: markerless pose estimation of user-defined body parts with deep learning." *Nature neuroscience* 21.9 (2018): 1281-1289.

McGinley, Matthew J., et al. "Waking state: rapid variations modulate neural and behavioral responses." *Neuron* 87.6 (2015): 1143-1161.

Miller, Josef M., et al. "Electrophysiologic studies of the auditory cortex in the awake monkey." *American Journal of Otolaryngology* 1.2 (1980): 119-130.

Mitchell, Jude F., Kristy A. Sundberg, and John H. Reynolds. "Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4." *Neuron* 63.6 (2009): 879-888.

Montes-Lourido, Pilar, et al. "Neuronal selectivity to complex vocalization features emerges in the superficial layers of primary auditory cortex." *PLoS biology* 19.6 (2021): e3001299.

Mosberger, Alice C., et al. "Motivational state, reward value, and Pavlovian cues differentially affect skilled forelimb grasping in rats." *Learning & Memory* 23.6 (2016): 289-302.

Moore, Tirin, and Katherine M. Armstrong. "Selective gating of visual signals by microstimulation of frontal cortex." *Nature* 421.6921 (2003): 370-373.

Nebes, Robert D., and Christopher B. Brady. "Phasic and tonic alertness in Alzheimer's disease." *Cortex* 29.1 (1993): 77-90.

Nelson, Anders, and Richard Mooney. "The basal forebrain and motor cortex provide convergent yet distinct movement-related inputs to the auditory cortex." *Neuron* 90.3 (2016): 635-648.

Niell, Cristopher M., and Michael P. Stryker. "Modulation of visual responses by behavioral state in mouse visual cortex." *Neuron* 65.4 (2010): 472-479.

Nilsson, Simon RO, et al. "A mouse model of the 15q13. 3 microdeletion syndrome shows prefrontal neurophysiological dysfunctions and attentional impairment." *Psychopharmacology* 233.11 (2016): 2151-2163.

Noudoost, Behrad, Kelsey Lynne Clark, and Tirin Moore. "Working memory gates visual input to primate prefrontal neurons." *Elife* 10 (2021): e64814.

O'Brien, Jennifer L., et al. "Cognitive training and selective attention in the aging brain: An electrophysiological study." *Clinical neurophysiology* 124.11 (2013): 2198-2208.

O'Connell, Monica Noelle, et al. "Layer specific sharpening of frequency tuning by selective attention in primary auditory cortex." *Journal of Neuroscience* 34.49 (2014): 16496-16508.

Okun, Michael, Amir Naim, and Ilan Lampl. "The subthreshold relation between cortical local field potential and neuronal firing unveiled by intracellular recordings in awake rats." *Journal of neuroscience* 30.12 (2010): 4440-4448.

Oken, Barry S., Martin C. Salinsky, and SM2865224 Elsas. "Vigilance, alertness, or sustained attention: physiological basis and measurement." *Clinical neurophysiology* 117.9 (2006): 1885-1901.

Okun, Michael, Amir Naim, and Ilan Lampl. "The subthreshold relation between cortical local field potential and neuronal firing unveiled by intracellular recordings in awake rats." *Journal of neuroscience* 30.12 (2010): 4440-4448.

Ortiz, Alexander V., David Aziz, and Shaul Hestrin. "Motivation and Engagement during Visually Guided Behavior." *Cell reports* 33.3 (2020): 108272.

Otazu, Gonzalo H., et al. "Engaging in an auditory task suppresses responses in auditory cortex." *Nature neuroscience* 12.5 (2009): 646-654.

Ozaki, Isamu, et al. "Rapid change of tonotopic maps in the human auditory cortex during pitch discrimination." *Clinical Neurophysiology* 115.7 (2004): 1592-1604.

Pardo, José V., Peter T. Fox, and Marcus E. Raichle. "Localization of a human system for sustained attention by positron emission tomography." *Nature* 349.6304 (1991): 61-64.

Peters, Megan St, et al. "Enhanced control of attention by stimulating mesolimbic–cortical cholinergic circuitry." *Journal of Neuroscience* 31.26 (2011): 9760-9771.

Petkov, Christopher I., et al. "Attentional modulation of human auditory cortex." *Nature neuroscience* 7.6 (2004): 658-663.

Picton, Terence W., et al. "Human auditory attention: a central or peripheral process?." *Science* 173.3994 (1971): 351-353.

Posner, Michael I., and Steven E. Petersen. "The attention system of the human brain." *Annual review of neuroscience* 13.1 (1990): 25-42.

Poulet, James FA, and Carl CH Petersen. "Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice." *Nature* 454.7206 (2008): 881-885.

Rinberg, Dmitry, Alexei Koulakov, and Alan Gelperin. "Speed-accuracy tradeoff in olfaction." *Neuron* 51.3 (2006): 351-358.

Robbins, T. "The 5-choice serial reaction time task: behavioural pharmacology and functional neurochemistry." *Psychopharmacology* 163.3 (2002): 362-380.

Robles, Luis, and Mario A. Ruggero. "Mechanics of the mammalian cochlea." *Physiological reviews* 81.3 (2001): 1305-1352.

Russell, Vivienne A., Terje Sagvolden, and Espen Borgå Johansen. "Animal models of attention-deficit hyperactivity disorder." *Behavioral and Brain functions* 1.1 (2005): 1-17.

Sadagopan, Srivatsun, and Xiaoqin Wang. "Nonlinear spectrotemporal interactions underlying selectivity for complex sounds in auditory cortex." *Journal of Neuroscience* 29.36 (2009): 11192-11202.

Sakata, Shuzo, and Kenneth D. Harris. "Laminar structure of spontaneous and sensory-evoked population activity in auditory cortex." *Neuron* 64.3 (2009): 404-418.

Satterfield, James H., Anne M. Schell, and Thomas Nicholas. "Preferential neural processing of attended stimuli in attention-deficit hyperactivity disorder and normal boys." *Psychophysiology* 31.1 (1994): 1-10.

Schneider, David M., Anders Nelson, and Richard Mooney. "A synaptic and circuit basis for corollary discharge in the auditory cortex." *Nature* 513.7517 (2014): 189-194.

Schofield, Brett R. "Structural organization of the descending auditory pathway." *The auditory brain* (2010): 43-64.

Shomstein, Sarah, and Steven Yantis. "Control of attention shifts between vision and audition in human cortex." *Journal of neuroscience* 24.47 (2004): 10702-10706.

Silver, Henry, and Pablo Feldman. "Evidence for sustained attention and working memory in schizophrenia sharing a common mechanism." *The Journal of neuropsychiatry and clinical neurosciences* 17.3 (2005): 391-398.

Smalley, S. L. (1997). Genetic influences in childhood-onset psychiatric disorders: autism and attention-deficit/hyperactivity disorder. *The American Journal of Human Genetics*, 60(6), 1276-1282.

Sołyga, Magdalena, and Tania Rinaldi Barkat. "Distinct processing of tone offset in two primary auditory cortices." *Scientific reports* 9.1 (2019): 1-12.

Soltani, Sara, et al. "Sleep-wake cycle in young and older mice." *Frontiers in systems neuroscience* 13 (2019): 51.

Sperling, George, and Melvin J. Melchner. "The attention operating characteristic: Examples from visual search." *Science* 202.4365 (1978): 315-318.

Stanislaw, Harold, and Natasha Todorov. "Calculation of signal detection theory measures." *Behavior research methods, instruments, & computers* 31.1 (1999): 137-149.

Stevens, Courtney, and Daphne Bavelier. "The role of selective attention on academic foundations: A cognitive neuroscience perspective." *Developmental cognitive neuroscience* 2 (2012): S30-S48.

Steriade, Mircea, David A. McCormick, and Terrence J. Sejnowski. "Thalamocortical oscillations in the sleeping and aroused brain." *Science* 262.5134 (1993): 679-685.

Stevens, Courtney, et al. "Neural mechanisms of selective auditory attention are enhanced by computerized training: Electrophysiological evidence from language-impaired and typically developing children." *Brain research* 1205 (2008): 55-69.

Sturm, Walter, and Klaus Willmes. "On the functional neuroanatomy of intrinsic and phasic alertness." *Neuroimage* 14.1 (2001): S76-S84.

Tubiana, Jérôme, et al. "Blind deconvolution for spike inference from fluorescence recordings." *Journal of neuroscience methods* 342 (2020): 108763.

Tudor, Mario et al. "Hans Berger (1873-1941)--povijest elektroencefalografije" [Hans Berger (1873-1941)--the history of electroencephalography]. *Acta medica Croatica : casopis Hrvatske akademije medicinskih znanosti* vol. 59,4 (2005): 307-13.

Tsukano, Hiroaki, et al. "Quantitative map of multiple auditory cortical regions with a stereotaxic fine-scale atlas of the mouse brain." *Scientific reports* 6.1 (2016): 1-12.

Unsworth, Nash, and Matthew K. Robison. "Working memory capacity and sustained attention: A cognitive-energetic perspective." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 46.1 (2020): 77.

Van Der Meere, Jaap, Boudewijn Gunning, and Nanke Stemerding. "The effect of methylphenidate and clonidine on response inhibition and state regulation in children with ADHD." *The Journal of Child Psychology and Psychiatry and Allied Disciplines* 40.2 (1999): 291-298.

Vanderwolf, C. H., and G. B. Baker. "Evidence that serotonin mediates non-cholinergic neocortical low voltage fast activity, non-cholinergic hippocampal rhythmical slow activity and contributes to intelligent behavior." *Brain research* 374.2 (1986): 342-356.

Vanderwolf, C. H., and Bruce A. Pappas. "Reserpine abolishes movement-correlated atropine-resistant neocortical low voltage fast activity." *Brain research* 202.1 (1980): 79-94.

Vernet, Marine, et al. "Frontal eye field, where art thou? Anatomy, function, and non-invasive manipulation of frontal regions involved in eye movements and associated cognitive operations." *Frontiers in integrative neuroscience* 8 (2014): 66.

Wang, Xiaoqin, et al. "Sustained firing in auditory cortex evoked by preferred stimuli." *Nature* 435.7040 (2005): 341-346.

Watanabe, Takashi, and Kajiro Watanabe. "Noncontact method for sleep stage estimation." *IEEE Transactions on biomedical engineering* 51.10 (2004): 1735-1748.

White, Hunter J., Muhammad Helwany, and Diana C. Peterson. "Anatomy, head and neck, ear organ of corti." *StatPearls [Internet]* (2020).

Williamson, Ross S., et al. "Locomotion and task demands differentially modulate thalamic audiovisual processing during active search." *Current Biology* 25.14 (2015): 1885-1891.

Woldorff, Marty G., et al. "Modulation of early sensory processing in human auditory cortex during auditory selective attention." *Proceedings of the National Academy of Sciences* 90.18 (1993): 8722-8726.

Wood, Suzanne, et al. "Psychostimulants and cognition: a continuum of behavioral and cognitive activation." *Pharmacological reviews* 66.1 (2014): 193-221.

Yerkes, Robert M., and John D. Dodson. "The relation of strength of stimulus to rapidity of habit-formation." *Punishment: Issues and experiments* (1908): 27-41.

Zhou, Huihui, and Robert Desimone. "Feature-based attention in the frontal eye field and area V4 during visual search." *Neuron* 70.6 (2011): 1205-1217.

Zhou, Mu, et al. "Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex." *Nature neuroscience* 17.6 (2014): 841-850.

<https://conductscience.com/maze/cba-j-mouse-strain/>

<http://www.cochlea.eu/cerveau-auditif>

<https://www.jax.org/strain/000656>

<https://www.mayoclinic.org/diseases-conditions/adhd/symptoms-causes/syc-20350889>