# Visual, auditory, and motor signals in the mouse superior colliculus

**Flóra Takács**

A thesis submitted for the degree of

Doctor of Philosophy

University College London

I, Flóra Takács, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

*Last night you told me tomorrow,*
*we shall have to think up signs,*
*sketch a landscape, fabricate a plan*
*on the double page of day and paper.*
*Tomorrow, we shall have to invent,*
*once more,*
*the reality of this world.*
*/Octavio Paz, January first, translated by Elizabeth Bishop/*

# Abstract

The superior colliculus (SC) contains visual, auditory, and motor maps that are thought to support sensorimotor transformations. Do mouse SC neurons contain all component signals necessary for an audiovisual localisation task? Do auditory, visual, and movement signals mix at the level of individual neurons? Is the mouse SC necessary for movements instructed by both visual and auditory stimuli?

I recorded the responses of ~2,500 SC neurons in awake mice to checkerboard images and pink noise sound bursts presented at varying azimuths, alone or in combination. Of these neurons, ~1,700 were recorded in untrained mice (N=10), and the rest in mice trained in an audiovisual localisation task using the same stimuli in the task and in a passive context (N=11). In a third cohort of mice, I used optogenetics to unilaterally (N=9) and bilaterally (N=3) inactivate the SC while mice performed the task.

As expected, I observed neurons responding to visual stimuli in superficial SC and to auditory stimuli in deeper layers, both in untrained and trained mice. Neurons responding to both modalities were rare (~5 %). SC neurons could predict upcoming choices in the task. Auditory stimuli evoked uninstructed movements, and in some sessions, uninstructed movements were also predictive of upcoming choice. Using linear regression, I found that visual, auditory, movement, and choice signals are in separate neurons.

Unilateral inactivation of the mouse SC promoted ipsiversive movements, both during auditory and visual choices. Bilateral inactivation restored behaviour. Effects of inactivation could be explained by injecting a constant bias by inactivation to the behaviour.

These results indicate that the mouse SC guides both auditory and visual sensorimotor transformations, but auditory, visual, and motor signals are in separate pools of neurons in the SC, and the SC is unlikely to resolve the audiovisual localisation task locally.

## Impact statement

The impact of the work that has gone into this PhD is threefold:

First, it contributed to understanding audiovisual integration and decision-making in the mouse superior colliculus. For example, we used to think that tasks that require mice to turn the wheel based on visual and auditory stimulus location could be solved by the superior colliculus alone. I showed that visual, auditory, and motor maps rarely interact, and these maps may be recruited independently during task execution instead. To my knowledge, this is also the first study, where unilateral and bilateral superior colliculus inactivation was performed in the same mice while they were performing a decision-making task. These experiments recapitulated the Sprague effect and confirmed signals for choice alternatives compete in the superior colliculus.

Second, this work produced tools, including recoverable Neuropixels probes, and parallelized acquisition systems, that enabled us to collect data of ~1 million audiovisual decisions of mice, providing data for further studies beyond the scope of this thesis.

Third, together with the audiovisual working group, we made a non-quantifiable contribution to our local community, and we now understand much more about the demands and benefits of highly collaborative project structures than before.

The work presented in this thesis was presented at numerous local and international conferences (SfN, FENS), and has contributed to two publications (Van Beest et al. 2024; Bimbard et al. 2024).

# Acknowledgements

I learned the above quote from a book[1] that my friend, Irene, gave me after my excruciating description of all the maps of the superior colliculus. I think it sums up my PhD well. The truth is, I often thought there was nothing new to understand about the superior colliculus during this PhD — it just makes us turn toward the light. But then I was forced to think of new perspectives because I was surrounded by inspiring people, and we interpreted and reinterpreted many things we read and experiments we performed. Now, these people each mean so much to me, and hopefully, this acknowledgement can express that.

Pip. I think the first time Pip and I connected was when we installed SpikeGLX, and I kept reading the manual immersing myself in *SpikeGLX latest*, while he kept going around taking photos of other rigs' SpikeGLX settings, and this striking personality difference made us laugh. Indeed, Pip made me laugh a lot throughout the years. Equally importantly, he got me out of many difficult situations and was always ready to give me his time and take my "cold" and "combative" opinions with much more calmness than anyone else. Numerous times it is only thanks to his argumentative assertiveness that an action plan replaced my fury and hatred of the world.

Célian. One cannot ask for a more knowledgeable collaborator than Célian, and he is always fast-paced, whether it is about fixing my bugs or responding to my ideas or worries. I especially treasure that he has a book, paper, story, or life experience to share for every situation, so I always leave an interaction with Célian having to think and read more. Additionally, practically, this PhD thesis would not exist if Célian wouldn't have listened to my mid-writing crises and did not repeatedly tell me that I need to get it done. Also, I tend to give up early if things don't work, and Célian found a way to teach me to expand my perseverance thresholds, or at least I hope he did.

Charu. I spent endless hours in Charu's office on mini-therapy sessions, and Charu listened. I would say more, but I think this sentence encompasses best her generosity and her unique ability to make a change and build bridges. And Charu always kept me on my toes so that I aspired to be a better mentor, scientist, and human and encouraged me to stop being a timid & shy girl and stand on my own feet.

Kenneth. I did this PhD at large because of the clarity of Kenneth's lectures during my undergraduate degree that made me want to understand the world as clearly as him, and so I figured I should also do this systems neuroscience thing (I don't think he knows, but so Kenneth, open my thesis). And Kenneth kept making these heroic entries into my academic life: always unexpected, and always with a big impact, ready to provide long-sought clarity, a new idea, or an annoying but plausible alternative hypothesis.

And finally, dear Matteo. When I became Matteo's student, I thought I understood how he explains stats and brain-related concepts, and indeed he was always way ahead of me in understanding what things *meant*. But he had such a different style and worldview from my then-self that I recall telling my sister that either he will change me somehow, or it will be a complete disaster. I am pleased to report

[1] Peter Turchi: Maps of Imagination: The writer as a cartographer. Trinity University Press, San Antonio, Texas, 2004

that we managed to ride it out. Matteo made me more confident and independent, and I came to appreciate that he tells me to project my voice because he cannot hear me. And now, I am bloody loud.

I realize that at some point I need to stop making paragraphs for everybody, though there have been many others in the lab who provided ideas, listened, encouraged, supported, shared experiences, helped, and resolved issues, and I am so grateful for their collegiality and friendship. I am especially indebted to the collaborators I worked with: George Booth collected much of the optogenetic data and proofread this thesis, and Magdalena Robacha assisted in collecting the chronic electrophysiology data. I am also grateful for Tim Sit whose code I relied on for kernel fitting and for the extraction of motion energy, for Max Shinn for helping me implement the pyDDM, and for Karolina Socha for piloting which optogenetic inhibitor to use in the SC. Also, thanks to Michael Krumin who has been my reliable last resort in any troubleshooting problem. I think it is worth mentioning that my most essential collaborators were ultimately my mice. I also cannot be thankful enough that large language models hit the mainstream before the writing of this thesis.

I have also been blessed with supportive mentors as I went through this PhD, and it is hard to overstate their meta-role in the actual completion of this project. First, thanks to Marius and Carsen, without whom I would have never believed that I could even do anything remotely computational. I have also been lucky to have done my undergrad at UCL and many friends and professors have supported me here well beyond their job descriptions. Amongst them, I especially want to thank Martina Wicklein, who grabbed me in Lidl at the lowest point of my PhD and met me for countless coffees, until I figured out how to take ownership of my academic progress. I am also grateful to my mentor, Dani Léderer, and all my colleagues and students at Milestone, who kept giving me perspective when I was fed up with the selfishness of academia.

My gratitude also goes to my friends who for some reason unbeknownst to me decided that they would not let my books and poetry protect me, and have provided me with fun, comfort, support, and new ideas; I decided to not list them because the nature of this text would create an artificial hierarchy that would be all too reductive of their unique contributions to my life. Aside from such great people, the rest of my mental stability was a result of fencing and more recently rowing during this PhD, and thanks to all the coaches and friends who built great communities both in London Fencing Club, Central London Fencing Club, and Putney Town Rowing Club.

No PhD acknowledgement would be complete without thanking my family. Anya, Apa, köszönöm, hogy mindig támogattatok, es hittetek bennem. Meg hogy mindig meghallgattatok, bármi is volt éppen az eszemben. Szerintem ha nem olyan családban növök fel, ahol a gondolataimat a szüleim egyenértekunek veszik a sajátjukkal, soha nem lennék kutató. Szóval köszönöm, hogy mindig részt vehettem a vitákban es mindig megosztottátok azt, ami irant épp érdeklődtetek. Mindig nagyon büszke voltam rátok, hogy mindketten olyan sokat elértetek, es köszönöm, hogy nekem is ilyen sok lehetőséget teremtettetek. És Luca. Olyan fantasztikus, olyan más mint én, olyan inspiraló és érdekes vagy, es nélkuled mindiannyian az elefántcsonttornyunkban ülnénk a sok okoskodással. Köszönöm, hogy *bárhol*, es bármikor meg tudom veled beszélni a világ folyását.

## Ethical Approval

Procedures on mice were performed in accordance with the UK Animals (Scientific Procedures) Act 1986. Experiments were performed at University College London in accordance with its animal care committee's regulation under personal and project licenses issued by the UK Home Office.

## Table of Contents

# List of figures

# Chapter 1 — General introduction

## Why study visual, auditory, and motor signals together?

The main objective of sensory neuroscience is to understand how we process and interpret the external environment to guide our actions. In this, traditionally, the field has pursued a feedforward approach (Buzsáki, 2019): sensory stimuli cause responses in neurons, which in turn are connected to other neurons that integrate more and more information; ultimately neurons at the top of the hierarchy are thought to create abstract representations of the world. This hierarchical feedforward model has been particularly successful in describing early sensory processing of images, sounds, and other senses (Hubel and Wiesel 1962; Carr and Konishi 1990). Even higher-order sensory areas are often described as feedforward integrators (Siegle et al. 2021); for example, higher visual areas are thought to integrate inputs from primary sensory areas to represent visual concepts, such as faces or spatial locations. Furthermore, hierarchical feedforward models inspired machine learning methods that can solve complex input-output mapping problems, such as image classification (Krizhevsky et al. 2017).

The main objective of motor neuroscience, on the other hand, is to understand how the brain generates, selects, and integrates movements into a coherent action plan. Motor actions are also thought to be hierarchically organised: 1) In the spinal cord, single motor neurons can activate a set of muscles together, forming motor units. 2) Also in the spinal cord, motor neurons, and interneurons coordinate rhythmic movements, such as locomotion, breathing, or swallowing, through circuits often referred to as central pattern generators. 3) Brainstem structures initiate various motor programs by innervating neurons in the spinal cord, while higher-order motor areas, such as the cerebellum, motor cortex, and the basal ganglia coordinate the finer aspects of a motor program and can initiate entire sequences of movement, often in a goal-directed manner.

But how are sensory and motor systems interconnected? In the simplest cases, a hierarchical circuit model is sufficient to describe a sensorimotor transformation: close to the periphery, sensory neurons are connected to motor neurons, forming reflexive loops that can explain how we withdraw our limbs in response to a noxious stimulus (Sherrington 1910). However, these sensorimotor transformations are often more complex. First, the sensory environment might be more nuanced, with relevant stimuli embedded in a complex sensory scene. Second, executing a motor program might involve a complex sequence of movements rather than the simple motion of a single body part. Third, the nervous systems can incorporate prior expectations, or access information about the external world from memory before committing to an action.

A crucial step to understanding how sensations translate into voluntary actions is to summarise the process with simple equations. In the perceptual decision-making field, several theoretical frameworks have been developed to describe these transformations. Psychophysics (Fechner 1870), which was later formalised by signal detection theory (SDT), was one of the earliest frameworks applied to perceptual decisions (Green and Swets 1966). SDT is based on the idea that perceptual decisions are influenced by detection accuracy in a noisy sensory environment, suggesting that limitations in performance are due to difficulties in perceiving the stimulus. Another prominent modelling tradition

is the sequential sampling framework, which proposes that agents accumulate noisy evidence over time, until reaching a decision threshold (Ratcliff and Mckoon 2008). While this model accounts for sensory uncertainty, it also allows for decisions to be influenced by non-sensory factors — such as differences in the ease of reaching one decision threshold over another. More recent quantitative approaches include Bayesian methods (Ernst and Banks 2002; Knill and Pouget 2004) and the application of logistic classification to perceptual choices (Busse et al. 2011). Many of these approaches are mathematically similar under certain conditions (Carandini 2024) and can be combined with other areas of decision-making research such as economic decision-making models (Summerfield & Tsetsos 2012; Lak et al. 2020). Overall, these quantitative frameworks have been successful in explaining behavioural choices across various perceptual decision-making paradigms. However, one of the major challenges in neuroscience remains understanding how any of these models might map onto mechanisms in the brain (Ma et al. 2006; Gold and Shadlen 2007; Hanks and Summerfield 2017).

At an implementational level, the DDM has perhaps been the most promising in explaining perceptual decisions. In a random dot motion task, where monkeys are asked to saccade to two alternative targets based on whether the dots on the screen move predominantly towards the left or the right, several brain regions, including the lateral intraparietal area (LIP) (Roitman and Shadlen 2002), were found to exhibit signals that ramp up with increasing motion coherence, suggesting that these areas encode perceptual evidence in the task.

However, the degree to which the DDM may map onto actual brain mechanisms has been extensively questioned. For example, recordings from neuronal populations revealed that LIP neurons exhibit a diverse range of dynamics and not all appear to act as evidence accumulators (Park et al. 2014). Furthermore, inactivating LIP did not seem to alter monkey choices (Katz et al. 2016).

Similarly, in other paradigms or modelling frameworks, behaviourally plausible models often fail to explain behaviours at an implementational level. For example, multi-area recordings have shown that numerous variables, such as choice, value, or bias which were initially thought to be encoded in specific brain regions, are, in fact, encoded across multiple areas (Steinmetz et al. 2019; Ottenheimer et al. 2023; Findling et al. 2023; Chen et al. 2024). Indeed, there is a lack of theoretical frameworks, that can effectively link our understanding of behaviour through psychology with measurements from the brain. One potential explanation for this gap is that variables used to explain behaviour — such as prior, evidence, value, bias or choice — may be less meaningful from a neural perspective or might be embodied differently depending on the physical nature of the decision. Therefore, a key goal in the neuroscience of perceptual decision-making is to identify quantitative models that can be implemented by neural circuits in the brain.

In this thesis, I focus on the joint study of auditory and visual decisions using a well-quantifiable audiovisual decision-making task. My goal is to map quantitative models of this task to plausible neural mechanisms, thereby understanding how the brain integrates various senses into a unified framework to guide coherent actions. Studying multisensory decisions is useful for understanding the building blocks of sensory decision-making: each sense relies on a different peripheral processing system. For example, the visual system represents the location of the stimulus immediately in the retina, while the

auditory system builds the representation of space using interaural and potentially spectral cues (Carlini et al. 2024). Moreover, the timing of the responses is also often different between senses: the auditory system has a higher temporal resolution than the visual system(Yang and Zador 2012), and in the somatosensory system, peripheral nerves conduct impulses at speeds that can vary by two orders of magnitude (Handler and Ginty 2021). Thus, integrating multiple senses requires the synthesis of diverse encoding schemes into a common mechanism.

On one extreme, it is possible that multiple senses only integrate at the level of action generation: for example, auditory and visual neurons may connect to distinct pools of motor neurons, and the only integratory step occurs at the level of muscles, thus, each sensorimotor loop can influence the same sets of muscles, giving rise to a coherent action. In the multisensory literature, this is often referred to as the race model. Such a race-model-based encoding scheme may be advantageous when actions need to be adjusted rapidly: for example, during the generation of defence reflexes, when multiple receptors from multiple body parts need to rapidly generate a coherent action.

On the other extreme, it is also possible that unisensory signals from each sense are integrated into a multisensory representation at an early sensory processing stage, such that for example, visual responses in the retina might theoretically be different in a silent vs. loud environment. This scenario is often referred to as the integratory mode, which might be useful for providing flexibility and adaptability by using the same set of pathways to generate multiple types of action sequences in a diverse set of environments (Figure 1).

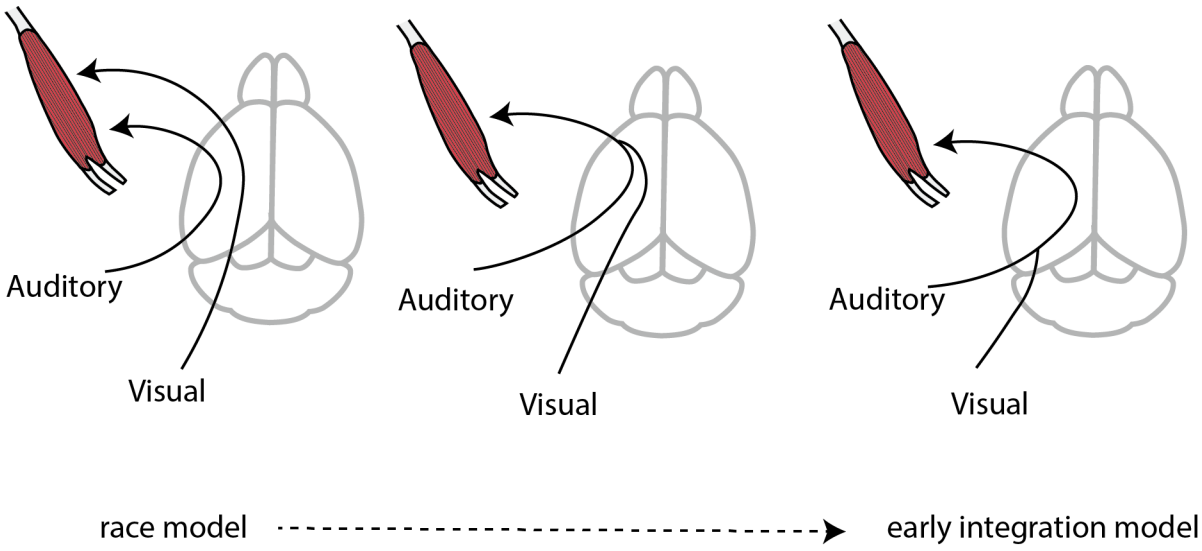



***FIGURE 1. THE RACE VERSUS THE INTEGRATORY MODEL OF MULTISENSORY PROCESSING.***

*Cartoons illustrate the differences between the race model and the integration model for audiovisual integration. In the most extreme version of the race model parallel auditory and visual pathways may influence the same muscles, whereas in the most extreme version of the integration model, audition and vision are already integrated in primary sensory areas. Muscle and brain images were downloaded from scidraw.io.*

It is unknown how any of these models are implemented in the brain. However, both neural and psychophysical evidence suggest the existence of an integratory coding scheme. Perceptual illusions

in humans, such as the McGurk effect (McGurk and MacDonald 1976) suggest that some percepts are generated by considering auditory and visual signals together. In mice, decisions are made faster when auditory and visual evidence supports the same decision, compared to when each modality supports competing actions, suggesting that integration is resolved before actions are generated (Coen et al. 2023).

Neurophysiological measurements typically consider the following evidence as support for the integratory model: 1) the existence of neurons that respond to stimuli from various modalities or where multisensory presentation of a stimulus modulates the unisensory response (Wu et al. 2015) 2) multisensory neurons that represent something meaningful, such as, for example, a spatial map that is invariant to stimulus modality; and 3) multisensory responses that cannot be predicted from the additive sum of unisensory responses, indicating a qualitative change in the multisensory processing mode. Such non-linear mixing of signals at a neural level is thought to provide easy readouts for motor systems at a diverse set of conditions, and thus non-linear integration of multiple senses may support flexible behaviour (Fusi et al. 2016).

Historically, neurophysiological measurements appeared to support an early integration model: neurons responding to multiple senses have been described in primary sensory areas (Ghazanfar and Schroeder 2006) as well as in an early visual structure, the superior colliculus. In the superior colliculus, furthermore, audiovisual neurons appeared to integrate unisensory signals non-linearly: they may exhibit non-linear properties such as multisensory enhancement that has been thought to be the mechanistic implementation for improving perceptual sensitivity by using multiple senses.

However, when neural approaches have been directly combined with psychophysical measures, evidence for the early integration scheme has been questioned. For example, Gu, Angelaki and DeAngelis (2008) found that the combining visual and vestibular cues is additive both behaviourally, and at the neural level in a higher-order visual association area, MSTd (medial superior temporal area) in primates. In rodents, (Raposo et al. 2014) found that non-linearity increased in higher-order association areas and was negligible in primary areas. Furthermore, Coen et al. (2023) found that primary cortical areas play a unisensory role in multisensory decisions, and that frontal cortex neurons combine audiovisual signals additively, suggesting that the audiovisual integratory step may arise later in the sensorimotor hierarchy than previously thought.

Thus, in this thesis, I aim to record and inactivate a well-known multisensory integratory structure, the superior colliculus. This brain structure has historically been considered both an early sensory region, but equally a highly integratory brain area that participates in both cue combination and decision-making. And yet, this region has hitherto been studied separately as a possible site of audiovisual integration and of perceptual decision-making. Thus, I will try to combine these two approaches and answer a single question:

How does the superior colliculus contribute to solving an audiovisual localisation task?

## Why the superior colliculus?

### Anatomy

The superior colliculus is a mesencephalic structure found in all vertebrates. Its primary function is to generate orienting actions towards salient stimuli. As vertebrates evolved and telencephalic structures developed a greater size and computational capabilities (Figure 2), the relative size of the superior colliculus decreased. Colliculus function is likely to have diverged, and in mammals, the superior colliculus is thought to also contribute to more cognitive computations, such as target selection and decision-making.

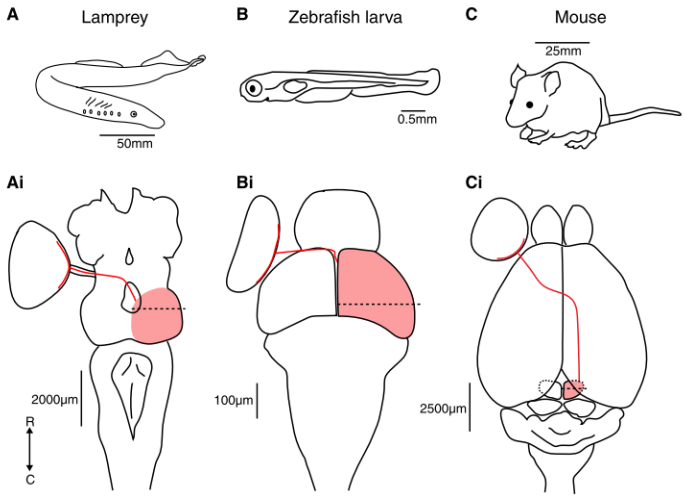



***FIGURE 2. LOCATION AND RELATIVE SIZE OF THE SUPERIOR COLLICULUS ACROSS SPECIES.***

*Reproduced from Isa et al. (2021) .*

#### *Mesoscale connectivity*

In mammals, the superior colliculus has extensive bottom-up, as well as top-down connections. Bottom-up inputs reach the SC from virtually all peripheral sensory systems. Specifically, the SC receives its main visual inputs from the retina and its main auditory inputs from the inferior colliculus, and it is also directly innervated by projection neurons from the spinal cord. The relative importance of these inputs may vary across species; for example, in nocturnal predators such as the barn owl, the auditory inputs may be more precise compared to prey animals such as mice, reflecting different ecological demands (Ito and Feldheim 2018; Isa et al. 2021).

The key top-down inputs of the SC include inputs from the endpoint of the basal ganglia bottleneck, i.e. the substantia nigra (Lee and Sabatini 2021). Furthermore, the mouse SC is highly connected with both the primary and the association cortices, including the primary visual and auditory cortices (Wang and Burkhalter 2013; Zhao et al. 2014; Zingg et al. 2017; Benavidez et al. 2021).

The main outputs of the SC are downstream: it projects to the spinal cord via the crossed and uncrossed pathways, which are thought to mediate approach and avoidance-like orienting behaviour (Isa et al. 2020). Additionally, the SC outputs upstream to higher-order thalamic nuclei, especially the pulvinar, and is interconnected with prethalamic nuclei such as the zona incerta (Benavidez et al. 2021) (Figure 3).

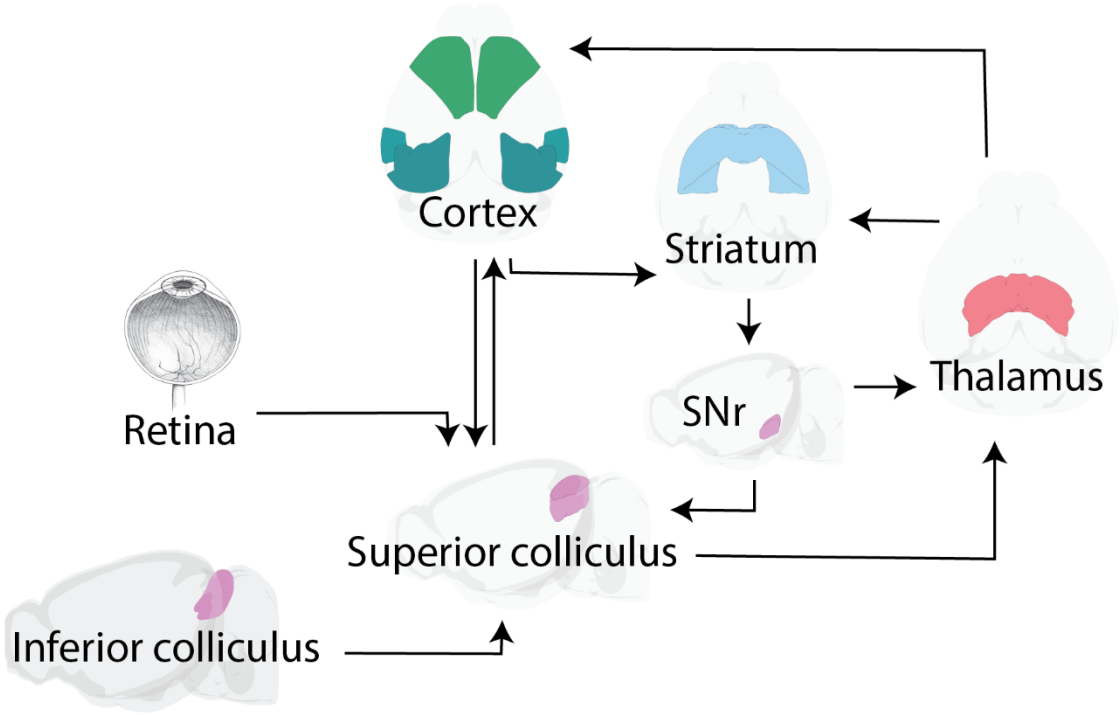



***FIGURE 3. MACROSCOPIC CONNECTIONS OF THE SUPERIOR COLLICULUS.***

*Alongside the extensive bottom-up input the SC receives (e.g. from the retina and the inferior colliculus), it also receives top-down inputs from sensory, motor, and association cortices as well as the basal ganglia. In turn, the SC projects to the thalamus, but also to cortical areas. Retina image from the database of National Institute of Diabetes and Digestive and Kidney Diseases, NIH. Other images were created with the brainrender-GUI (Claudi et al. 2021).*

### *Local functional connectivity*

The SC is organised into functionally specialised layers, which have historically been characterised via anatomical staining that assesses the degree of 'granulation' (Figure 4A). The superficial layers (SCs) specialise in processing visual information, while intermediate and deeper layers (collectively denoted SCm, which stands for motor-related SC) are more integrative. They receive inputs from multiple senses and contain motor neurons that drive orienting movements (Ito and Feldheim 2018). Indeed, the SCm was the first brain structure where audiovisual neurons — single neurons that respond to multiple modalities — were recorded (Meredith and Stein 1983). The SCm is also the primary recipient of top-down connections.

Anatomical and in vitro characterizations suggest that the superficial and the deep layers are locally connected. However, the function of these connections remains unclear (Isa and Hall 2009). Moreover, the SCm is thought to connect the two hemispheres via commissural fibres (Edwards 1977), and it is the primary structure that sends motor outputs relating to orienting movements. It is unclear whether all these functional properties of the SCm are in separate neurons, or whether the SC neurons perform highly integratory computations that bind these functions together.

In addition to these layer-based subdivisions, another subdivision, defined along the mediolateral (ML) axis has also been characterised (Figure 4B). This SC parcellation was inspired by the type of actions the SC can control; the SC can mediate both escape- and approach-like orienting behaviours in rodents and such circuits might be topographically wired, such that the upper visual field (medial SC) is linked to avoidance-like actions, especially in terrestrial animals who need to escape from aerial predators (Wheatcroft et al. 2022). This suggests that the topographical map of the visual space in the SC has an ethological meaning along the ML axis, such that the more medial parts monitor the environment for objects to avoid while the more lateral parts monitor objects to approach.

Further support for such functional specialization along the ML axis came from the thorough analysis of SC inputs and outputs, which suggested that the same parts of the SC, basal ganglia, thalamic nuclei, and cortex are connected, forming loops that might allow a set of continuous transformations of the same information. Each of these brain regions might specialise for distinct functions, illustrating the homotopical organization of information across brain regions in topological terms (Benavidez et al. 2021).

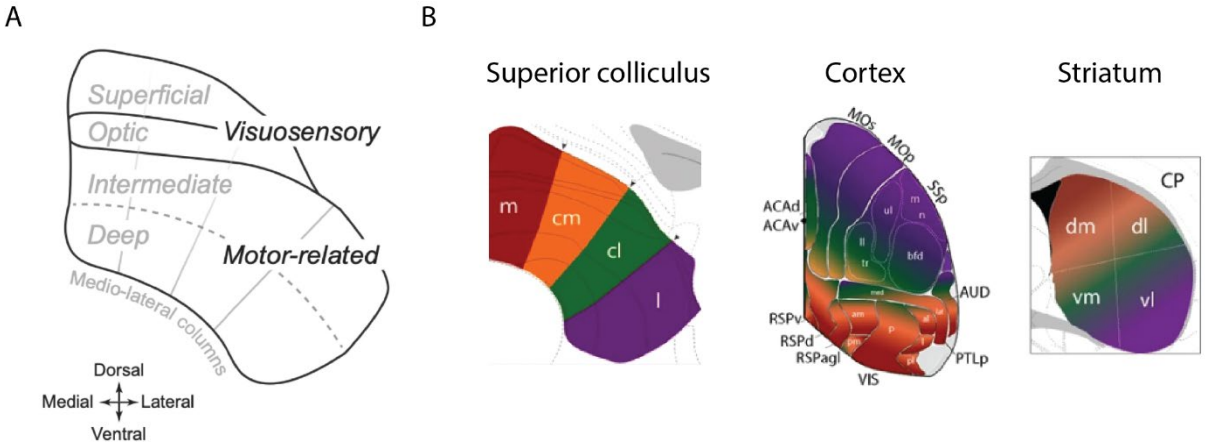



***FIGURE 4. SUBDIVISIONS OF THE SUPERIOR COLLICULUS.***

*(A) SC subdivisions along the dorsoventral (DV) axis, i.e. the superior colliculus layers. Adapted from* Wheatcroft et al. (2022) *(B) SC subdivisions along the mediolateral (ML) axis. The colours indicate which part of each of these structures are the most densely connected, suggesting that all structures follow a homotopical organization. Reproduced from* Benavidez et al. (2021).

## Function

Both the sensory and motor aspects of the SC have been extensively studied. The SCs is a predominantly visual structure, performing a variety of computations on retinal inputs (see for example Li and Meister 2023). The SCm, in turn, is thought to be a sensory integrative and motor structure; on one hand, it receives input from numerous sensory systems beyond vision, such as audition and somatosensation; on the other hand, it also has a well-established role in saccade generation and orienting in various species. In addition, numerous movements appear to have topographic maps in the SCm, like the visual map in SCs. This observation inspired numerous models that suggest local connections between the superficial and the deep layers might be able to resolve certain sensorimotor transformations. Furthermore, the SCm has been indicated to be involved in movements beyond saccades and orienting, such as forelimb movements and the generation of

complex motor programs for approach and avoidance, suggesting that the SCm might have a more general function in implementing perceptual decisions.

This thesis focuses on the role of SC in integrative mechanisms during multisensory decisions. That is, how the sensory and motor aspects of the SC might work *together* while animals make perceptual decisions. Thus here, I will focus on reviewing the SC from such an integratory perspective. More specifically, I will discuss three topics further: (1) the alignment of different sensory and motor maps of space, (2) whether the sensory maps functionally integrate at the level of single neurons in the SC forming a modality-invariant sensory map, and (3) our current understanding of SC's role in the generation of movements during perceptual decisions.

### *Alignment of sensory and motor maps*

An important finding that elevated the SC to key prominence as a substrate for sensorimotor transformations is the alignment of sensory and gaze shift maps. This idea originated from Julia Apter who measured extracellular potential in cat superficial layer SC and found that increasing the excitability of patches of the more ventral part of the SC results in spontaneous gaze shifts to the location where the more dorsal SC appears to respond most to visual stimulation (Apter 1946). These measurements inspired Pitts and McCullogh (1947) to suggest a model of local sensorimotor transformation from the sensory discharge to action (Figure 5). Goldberg and Wurtz (1972) measured neural activity in both the superficial and the deep layers while monkeys made saccadic eye movements to visual targets, described the receptive fields of SCs neurons, and found motor neurons that discharged in close temporal relationship with saccadic movements (Wurtz and Goldberg 1972; Goldberg and Wurtz 1972). The careful measurements of neural activity inspired several hypotheses about what this layer might compute, including a saccade vector that ought to describe either the direction of movement (foveation hypothesis, Schiller and Stryker 1972) or the end-goal target of the movement.

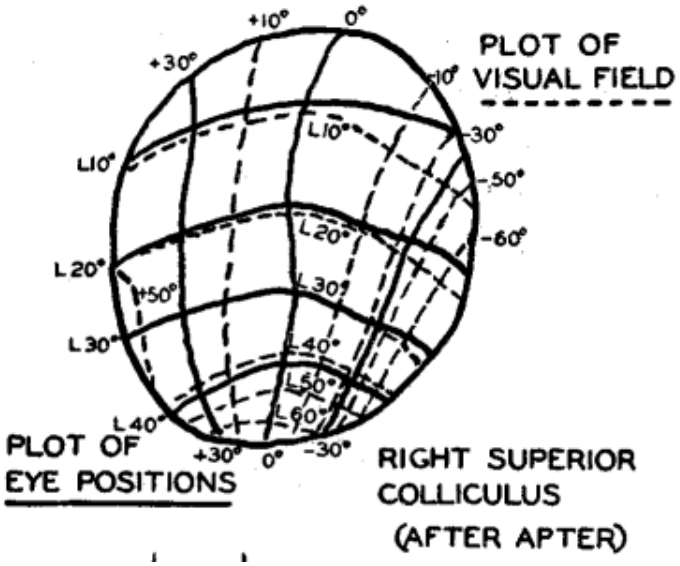



***FIGURE 5. FUNCTIONAL ORGANISATION OF THE SUPERIOR COLLICULUS RELIES ON TOPOGRAPHICAL MAPS.***

*The overlapping visual and movement fields of the superior colliculus, adapted from Pitts and McCullogh (1947).*

More questions arose as more maps were discovered: we now know that the SC contains numerous sensory maps beyond representing visual space; SC neurons also contain signals to represent maps of audition (Palmer and King 1982; Middlebrooks and Knudsen 1984; King and Hutchings 1987;), and somatosensory maps, including, for example, touches of the whiskers in rodents (Drager and Hubel 1975). The motor maps are also not confined to gaze shifts, and SC motor maps also include the maps of the head, tongue, and the entire body. These maps all seem to each be topographical, but do these maps interact? More specifically, are they all brought together in a common reference frame, i.e. a modality-invariant representation of egocentric-space that guides actions?

### *Role in cue integration*

The SCm has been identified as a key area where several sensory modalities converge on single neurons, including vision and infrared sensation, vision and audition, and vision and the whisker somatosensory system (Newman and Hartline 1981; Meredith and Stein 1983; Gharaei et al. 2018). The interaction between the auditory and the visual maps has perhaps been the most extensively studied. The first single neuron that responded to both auditory and visual stimuli was observed in the SC of the cat (Stein and Meredith 1983). This line of literature has put substantial emphasis on the possible emergent integratory computations (cf. Figure 1A) in the superior colliculus by describing audiovisual enhancement, i.e. supralinear responses during multisensory stimulus presentation, that are bigger than the additive sum neural signal in response to unisensory stimuli. Indeed, audiovisual non-linearities at the neural level have been guiding principles in the field, with numerous studies reporting supra- and subadditive neurons across senses and species, including most recently the mouse (Ito et al. 2021).

However, as the number of recorded neurons within each study has grown, there has been increasing doubt about the prevalence and indeed the conceptual importance of non-linearity. First, (Stanford et al. 2005) noted, that in large populations, most audiovisual neurons appear near-additive even under anaesthesia. Others have also reported mostly additive neurons in awake monkeys during auditory and visual saccade generation in the awake primates (Populin and Yin 2002). Furthermore, it has been noted, that behaviourally, visual and somatosensory cue combinations appear to fit additive models in humans(Ernst and Banks 2002). Angelaki, De Angelis, and colleagues started studying cue combinations to bridge the gap between psychophysical and neural approaches and found that visual-vestibular cue combination is additive both at a behavioural and a neural level in the cortex (Gu et al. 2008). Later, they extended their model to fit audiovisual interactions recorded in the superior colliculus, noting that the small fraction of non-linearity can be potentially explained by a divisive normalization (Ohshiro et al. 2011; Fetsch et al. 2013). This form of non-linearity is considered a ‘mild’ version of non-linearity, where non-linearity arises from the non-linear nature of action potential generation (i.e. sub-threshold activity and saturation), as opposed to dendritic spike generation or other types of non-linearities that could emerge from a single neuron implementing some form of audiovisual binding (Figure 6). Indeed, if this is the case, it is also possible that numerous studies that have historically made observations about non-linearity could do so because the measurements were done under anesthesia which might have decreased the overall excitability of SC cells, pushing evoked activity closer to the subthreshold range (Populin and Yin 2002).

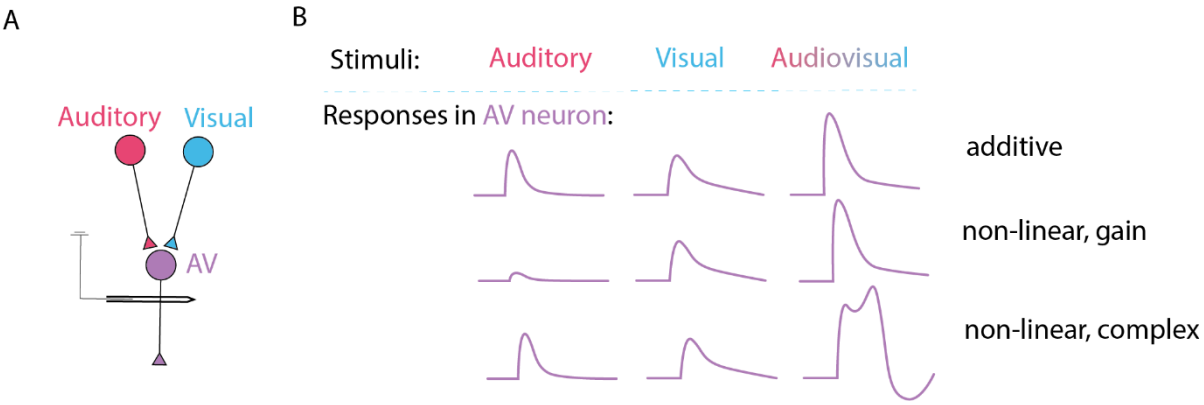



***FIGURE 6. THE ADDITIVE VERSUS NON-LINEAR COMPUTATIONS FOR AUDIOVISUAL INTEGRATION.***

*(A) Cartoon example indicating that electrophysiological measurements of audiovisual (AV) neurons can test the computational rules underlying the integration of inputs within the AV neuron. (B) Comparing responses to unisensory auditory, visual, and multisensory audiovisual stimuli can allow us to infer the computational rules underlying integration within the AV neuron: additivity, gain non-linearity, or complex non-linearity.*

In summary, the extent of audiovisual cue combination in the superior colliculus has been debated. It is unclear whether SC audiovisual neurons have emergent computational properties, such as audiovisual enhancement, and how the SC audiovisual signals contribute to audiovisual decisions.

### *Role in movement generation and decision-making*

The superior colliculus has long been recognised as to contribute to movement generation. Indeed, the SC contains a topographic movement map, and neurons in the intermediate and deep layers increase their firing rate even before monkey's saccade to a specific target location (Wurtz and Goldberg 1972). In line with this, initial studies suggested that the overall role of SC is to adjust eye positions towards visual targets rapidly (Apter 1946; Schiller and Stkyker 1972). These first studies in the SC mostly focused on eye movements in the monkey. However, over the years, more and more studies suggested that the SC may have a more generic role in movement control, including deliberate movements that may be classified as decisions.

First, in numerous species, the SC does not only control eye movements but also movements of other body parts, such as the tongue, the whiskers, the pinnae, and the neck muscles controlling the movement of the head (Gandhi and Katnani 2011; Masullo et al. 2019; Cooper and Mcpeek 2021). This has led to the idea that the superior colliculus might control orienting movements in general. Other studies went beyond this, suggesting that the SC might also control non-orienting movements of the forelimb and that it might be responsible for the execution of entire motor programs such as approach during prey capture or escape avoidance (Dean et al. 1988; Hoy et al. 2019; Shang et al. 2019; Evans et al. 2018; Campagner et al. 2023).

Second, it has been long appreciated that the SC does not simply issue a premotor command (Basso and May 2017). For example, when stimulating the SCm at a particular focal point, the monkey always appears to saccade toward a specific goal location, independent of the starting position of the eye, instead of just simply making a saccade in a particular movement direction (Robinson 1972).

Furthermore, others have also found that inactivating the SC does not just affect movement direction, but monkeys choose alternative targets, instead of those in the deactivated movement area (McPeek and Keller 2004). Furthermore, inactivating the SC only causes deficits in the affected hemifield if there are distractors in the opposing hemifield (Lovejoy and Krauzlis 2010; Krauzlis et al. 2013). In addition, some SC neurons correlate with choice directions in a rule-switching task, i.e. when animals need to turn either towards or away from the stimulus depending on an initial cue (Duan, Pagan, et al. 2021), suggesting that SC has a more cognitive role in rodents.

Finally, the SC was also found to participate in non-visual tasks: SC neurons correlate with choice directions an olfactory 2-AFC task too (Felsen and Mainen 2008; Felsen and Mainen 2012), and inactivating the SC in this task appears to alter choices (Essig et al. 2021). Yet, olfactory stimuli are not processed by the SC directly, suggesting that the SC might not only make movements to visual targets, but instead, it might have a more generic role in spatial decisions

Thus, the SC probably plays a crucial role in making voluntary decisions. However, there are still numerous open questions regarding the precise role of SC in generating movements: What do SC neurons that correlate with movement encode? Which body parts does the SC control? What decisions can the SC influence?

### *How can the SC help to understand audiovisual decisions?*

I have introduced substantial evidence that the SC plays a role in auditory- and visual-spatial processing, audiovisual integration, and decision-making. Indeed, the SC contains signals that can, in principle, solve spatial audiovisual decision-making problems that rely on the spatial localisation of stimuli. But how do SC neurons actually do it? Do the auditory and visual maps interact to represent egocentric space additively or non-linearly? Do these interactions change when mice learn spatial audiovisual decision-making tasks? Is the SC necessary in an audiovisual localisation task?

In this thesis, I try to bridge the fields of audiovisual integration and decision-making in the superior colliculus. I began studying this topic using electrophysiology, and in the next chapter, I describe the general methods I used.

To begin, I made measurements of superior colliculus activity in response to visual and auditory stimuli in naïve mice. Then, I trained a separate cohort of mice to perform an audiovisual spatial decision-making task using the same stimuli, and I measured the colliculus activity in these mice both during task execution and during passive stimulus presentation after task execution.

I broke down the results from these measurements into four chapters. In Chapter 3, I discuss the innate representation of space by auditory and visual signals in the SC in awake naïve mice to understand how these sensory signals can be potentially used to solve the spatial audiovisual task. However, the existence of neurons capable of solving spatial localisation tasks does not guarantee their active participation. Hence in Chapter 4, I explore the activity of superior colliculus neurons while mice perform the audiovisual task and assess whether neurons can predict the choice made during visual, auditory, and audiovisual decisions.

Chapters 3 and 4 present results regarding sensory responses and choice measures. However, these findings could be misinterpreted if the observed neural correlates of vision, audition, and choice are consequences of body movements rather than the direct processing of sensory signals. Chapter 5 addresses this by exploring how visual and auditory stimuli, along with animal choices, relate to body movements, and how these signals are spread across the SC.

Armed with an understanding of vision, audition, choice, and uninstructed body movements, Chapter 6 will explore these signals simultaneously. Key questions will be addressed regarding the interaction of these signals in single neurons and across the population: Are these signals in separate neurons, or are they mixed within neurons? If they are mixed, what computations determine this mixing?

The conclusions I make from my electrophysiological measurements are complemented with both unilateral and bilateral optogenetic inactivation in the SC while mice make audiovisual decisions. All results from SC inactivation are summarised in Chapter 7.

Finally, in Chapter 8, I discuss how the results of both the electrophysiological measurements and the inactivation studies can be reconciled with each other as well as the current literature and discuss possible mechanistic models of audiovisual decision making given the role of the superior colliculus.

# Chapter 2 — Methods for studying audiovisual localisation in the SC

## Sensory neuroscience and psychophysics

In this thesis, I focused on understanding spatial audiovisual localisation. Visual and auditory stimuli were delivered at different locations, both alone, and in combination. Cues were presented both in passive conditions and in an audiovisual localisation task, in which mice were trained to respond to the same auditory and visual cues to obtain reward. Below, I will describe the methods of delivering each of these stimuli, the considerations while delivering them together, and the audiovisual localisation task.

### Animals

Decision-making at a circuit level is often studied in monkeys or more recently rodents. Here, I chose mice as an experimental strain, because an audiovisual localisation task has been readily developed for mice at the start of my thesis (Coen et al. 2023), and further studies indicated that the SC might play a critical role in similar decision-making tasks (Steinmetz et al. 2019). Although at the start of my thesis, it was not clear whether audiovisual integration occurs in mice, audiovisual integration appeared to be widespread across species, for example, cats (Meredith and Stein 1983), barn owls (Zahar et al. 2009), and guinea pigs (King and Palmer 1985). Therefore, I reasoned that SC might contain audiovisual neurons in mice too.

The most used mouse strain is the C576B/L. However, this strain suffers from age-related high-frequency hearing loss after the age of ~12 weeks, meaning that in addition to the non-frequency dependent age-related hearing loss, C57 mice especially lose their audition in the >12kHz range (Parham and Willott 1988; Yin Zheng et al. 1999). During the original design of the audiovisual localisation task, this was accounted for by playing sounds in a low-frequency range (8-16kHz), in which age-related hearing loss is much less profound. Initially, in my PhD therefore, I used the C57BL strain. In ~2020, my lab managed to obtain a hearing-loss corrected strain, in which the point mutation that causes causing the hearing loss is corrected (Mianné et al. 2016). I used this strain afterwards. Although we know that mice can localise sounds after 12 weeks if the auditory stimulus in the intermediate frequency range even without hearing-loss correction, the auditory maps may be altered in mice that suffer from hearing loss, but much less so in the hearing-loss corrected strains (Si et al. 2022).

Altogether, I used 30 mice for this thesis, 22 of which were hearing loss corrected. The age of mice ranged from 10 to 45 weeks during data acquisition; notably, mice were often recorded for a long time using chronic Neuropixels recordings or inactivated daily for an extensive period. Altogether I used 22 Male and 8 female mice; this sex imbalance was also due to chronic recordings, as the early chronic implants used during my PhD were heavier. However, I participated in the development of lightweight chronic implants for Neuropixels probes, which allows the implantation of lighter-weight (~20g) mice (Bimbard et al. 2024). Further details about each mouse can be found in the Appendix A.

### Visual stimuli

Visual stimuli were presented on three iPad screens (Adafruit, LP097QX1) positioned ~11 cm from the mouse (Figure 7A). The screens covered a range of ±135° azimuth and ±37.5° elevation. To equalise

luminance across viewing angles, I covered the screens with Fresnel lenses (Wuxi Bohai Optics, BHPA220-2-5). All screens were further coated with a scattering film to reduce reflections. The screens refreshed at 60Hz, and luminance (39±11.6 cd/$m^2$) was equalised across the three screens (it might have differed across screens). Screens were gamma-calibrated to equal contrast values across all screens.

Throughout this study, I delivered two types of visual stimuli. Firstly, sparse noise stimuli, to map the receptive fields of visual neurons. Secondly, a checkerboard pattern was presented during passive viewing of the audiovisual stimuli and during the audiovisual task (Figure 7B).

For the sparse noise stimuli, white squares of 7.5°-size were displayed on a black background. The location of each white square was selected randomly from a spatially uniform distribution updating the selection at 20Hz. Each white square appeared for 1/6 s.

The checkerboard pattern consisted of four squares along the azimuthal axis and ten squares along the elevation axis, each square was 7.5° wide. The squares had alternating greyscale values varying between white and black. I used this checkerboard pattern, as opposed to Gabor patches that are more commonly used in similar tasks, to increase the chances of being able to record from neurons that respond to visual stimuli by covering a large field of view.

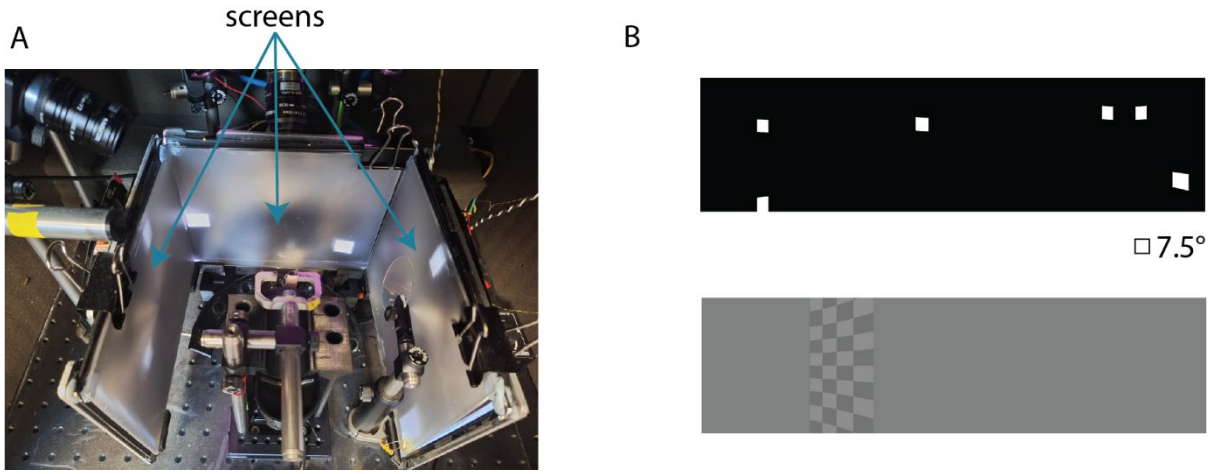



***FIGURE 7. THE VISUAL STIMULI.***

*(A) The screen setup, playing the sparse noise protocol.*

*(B) The two types of stimuli used: sparse noise (top) and checkerboard images (bottom).*

## Auditory stimuli

Seven loudspeakers (102-1299-ND, Digikey) were evenly spaced between ±90° azimuth, positioned just below the screen (Figure 8A). These speakers were driven with an internal sound card (STRIX SOAR, ASUS) and a custom amplifier as described before (Coen et al. 2023). To calibrate the speakers, I recorded 2-20 kHz white noise bursts placed at the mouse's approximate location (with GRAS 40BF 1/4" Ext. Polarised Free-field Microphone) and constructed a compensatory filter that ensured equal loudness across the frequency spectrum.

I used 8 – 16 kHz pink noise bursts as auditory stimuli during the task and the passive recordings. These noise bursts were filtered using ten pre-selected random filters that amplified or suppressed different frequency components. This adjustment aimed to prevent the identification of individual speakers based on residual differences in their frequency content, ensuring that the speakers were only identifiable by their location (Figure 8B-C).

Controlling background sounds and echoes in the experimental chamber is not trivial for audiovisual experiments. Echoes can be created by the screens used to create visual stimuli. Other necessary equipment (cameras, microscopes, etc.) can also distort sounds and create echoes. To minimise background sounds, I performed acute electrophysiological recordings in a sound isolation chamber. However, there were numerous computers located inside the chamber. Subsequently, we developed chronic recording stations that did not contain computers inside the chamber. Additionally, I also played a constant masking sound to provide a static ambient auditory background (Figure 8C).

To further confirm that sounds displayed as expected, I recorded sounds during experiments using an ultrasonic microphone (Ultramic, UMK200K, Dodotronic). The microphone was sampled at 96kHz. The signal was transformed to the frequency domain using the MATLAB spectrogram function with Kaiser window for visualisation (Figure 8C).

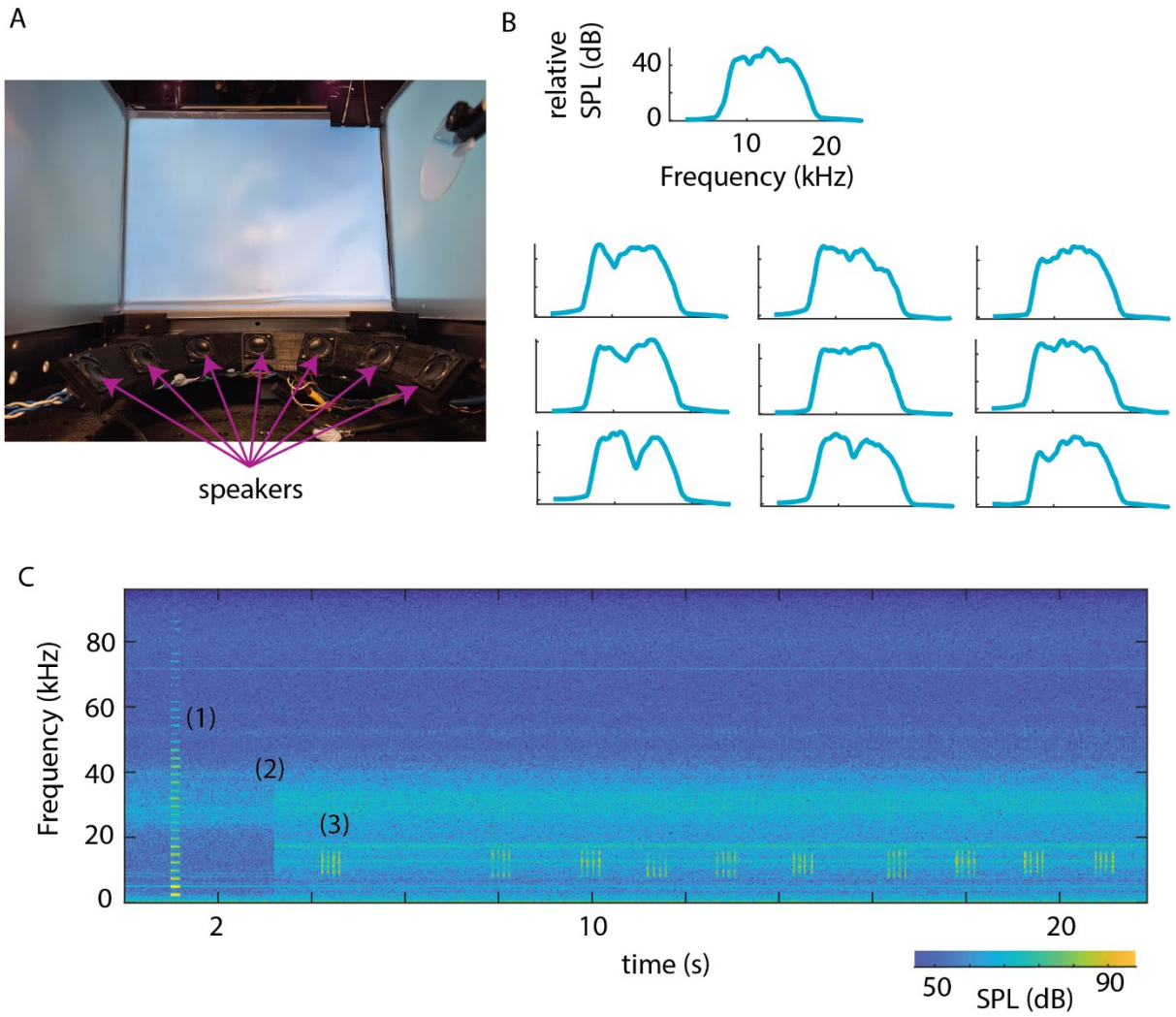



***FIGURE 8. AUDITORY STIMULI.***

*(A) Speaker location in the measurement setup. (B) The average frequency content of the generated sound clicks, with the 10 random filters used in the task and passive replay. (C) Example microphone recording of the sound bursts presented during passive stimulus replay, where three components are visible: (1) the buzzer sound (pure tones at 2.5 kHz intervals) used to align the microphone, (2) the initiation of the constant background noise played during the experiments, and (3) the 8– 16 kHz pink noise bursts.*

## Delivering auditory and visual stimuli together

Ensuring the synchronised delivery of auditory and visual stimuli presents a significant challenge. Over time, various methods have been explored to achieve spatial congruence. In the earliest studies that combined audiovisual stimulus delivery and electrophysiology, LEDs were used as visual stimuli that could vary in luminance. As LEDs are compact, speakers could be placed adjacent to LEDs (Stanford et al. 2005). This setup allows limited control over visual stimuli compared to controlling visual stimuli on LCD screens. On the other hand, LCD screens create uncontrolled echoes for the auditory stimuli, so LEDs are ideal in minimising echoes.

Another solution has been to mimic spatial auditory stimuli through head-related transfer functions measured in the ear of average animal heads *post-mortem* and deliver sounds directly into each ear with the frequency and timing content that would arrive at that ear (e.g. Ito et al., 2020, 2021). This method notably allows for bigger screens; however, it can be prone to systematic errors and does not account for some aspects of natural hearing (cf. pinnae movements). In this thesis, I used screens, and speakers located just below the screens.

Similarly, temporal congruence is hard to achieve in hardware because it requires the temporal synchrony of two different display streams. In this study, the trigger of the auditory stimulus was followed by the trigger of the visual display change with a delay of 13±7 ms. It is important to note that some studies have used an electrophysiology-guided approach when studying audiovisual integration to achieve peak temporal coherence in spike timing (i.e. the auditory stimuli were delivered later because the early auditory system processes information faster than the early visual system, see for example Rowland et al. 2007). However, this approach can only test the physiological capabilities of neurons, rather than audiovisual integration processing in a world where the temporal properties of visual and auditory information are inherently different both before they reach the body (cf. the speed of light versus the speed of sound), and after (cf. the conversion of photon input versus sound wave to action potentials).

### *Stimulus combinations*

Auditory and visual stimuli could be presented at seven different azimuthal positions (−90°, −60°, −30°, 0°, 30°, 60°, 90°). Both types of stimuli could also be varied in intensity (sound pressure level (SPL): 50 dB, 65 dB, 75 dB, and visual contrast: 5 %, 10 %, 16 %, 25 %, 40 %, 80 %, 100 %). Due to time limitations and the need to obtain enough repeats for statistical power, it was not possible to vary all these parameters in a single experiment (which would be 7*7*3*7+7*3+7*7 = ~1100 stimulus combinations), and only a subset of all possible combinations was used in both passive and experiments and during decision-making. Appendix A details the types of experiments and stimulus combination each mouse was presented. Most typically, in passive experiments, mice were presented with stimuli at all spatial locations in unisensory and congruent conditions, and with a subset of incongruent stimuli (Figure 9A). In the task, mice were typically presented with stimuli at three positions at three contrasts, and the SPL was not varied (Figure 9B).

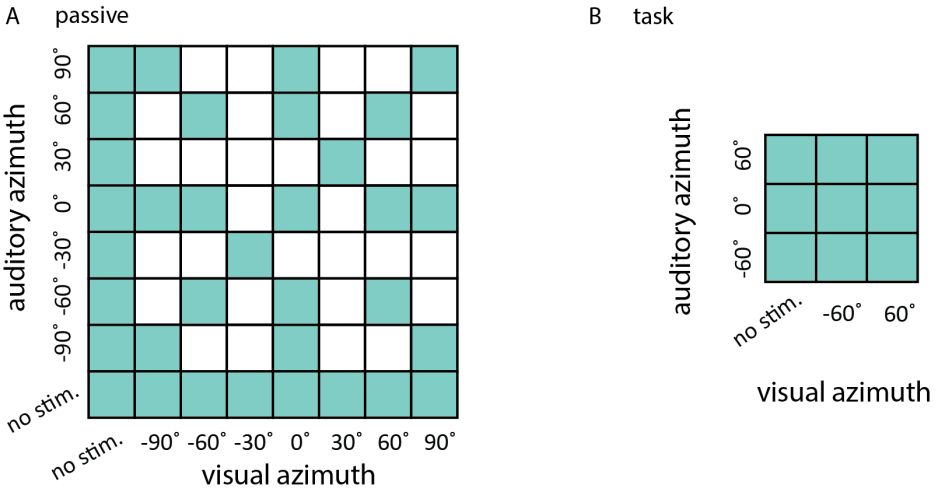



***FIGURE 9. STIMULUS LOCATION COMBINATIONS.***

*On the x-axis, I show the possible visual stimuli options: no stimulus, and stimuli at each azimuth. The y-axis represents the same for auditory stimuli. Green grids are typically the location of stimuli that I selected to play. (A) Stimulus locations in passive experiments. The bottom row represents the visual unisensory stimuli, the leftmost column represents the auditory unisensory stimuli, and the rest of the matrix represents the audiovisual stimuli. (B) Stimulus locations during the audiovisual localisation task.*

#### *Duration of stimuli*

In acute naïve experiments, stimuli were presented for 200−300 ms (fixed for each session) and both auditory and visual stimuli were flashed only once. In the audiovisual-decision-making task, sounds and images were flashed at 8Hz for 1.5 seconds in the final version of the task unless the mouse made a choice earlier. During passive experiments in trained mice, and in naïve mice recorded chronically (i.e. at a later stage of my PhD compared to acute recordings), stimuli were presented for four bursts (e.g. Figure 8C).

## Audiovisual decision-making task

### Task structure and training

The audiovisual two−alternative forced choice (2−AFC) task design was adapted from Coen et al. (2023). Mice were trained to indicate the auditory and visual stimulus location by turning a Lego wheel (Figure 10A). Turning of the wheel was measured with a rotary encoder (Kubler 05.2400.1122.0360) attached to the wheel. Mice could initiate the trial by holding the wheel for a random quiescent period (100−250ms, exponential distribution), after which a stimulus came on until mice made a choice by turning the wheel, such that the stimulus is brought in the middle of the screen (Figure 10B). Upon a correct choice, mice were rewarded with sucrose water (Figure 10C).

I made a few changes compared to the task originally described: (1) I presented checkerboard stimuli that spanned the entire screen in elevation to maximise the visual stimulus responses in visual areas such as the colliculus. As this altered the visibility of the stimulus, I lowered the visual contrast. (2) omitted the open loop period and the go cue (3) after the mouse learnt the full task, fixed the stimulus in place, thus the stimulus was not yoked to wheel movement in the recording phase of the task. Mice were shaped to perform this behaviour using a standardised training protocol (similar to the International Brain Laboratory et al. 2021). For this, task parameters were adaptively changed based

on various performance criteria being met, including total number of trials performed, left−right bias, and proportion of correct responses on various trial types. Some parameters were updated continuously (reward size, response window, and wheel gain determining the amount of wheel turn necessary to bring the stimulus to the centre), and some were updated in stages. Once mice reached the final stage of training, the probability of the wheel being yoked to the stimulus was continuously decreased, reaching zero before electrophysiological recording or inactivation (Figure 10D-G).

In the final stage, stimuli were randomly selected, such that auditory, visual, coherent, and conflict trials were presented in a 3:3:2:1 ratio. Mice were rewarded upon successful wheel turns 100 % of the time during unisensory and coherent trials. Conflict trials were rewarded randomly (50 %). In the final training stage, the stimulus was not yoked to the wheel. Fully trained mice had 1.5 s to respond. Minimum inter−trial interval (ITI) was 2 s, during which wheel holding could not initiate a trial yet. If mice responded incorrectly, they were subjected to a longer timeout (up to 3s), and incorrect responses during the easiest trials (highest visual contrast) were repeated, but not counted towards performance metrics.

### Analysis of wheel movements during the audiovisual task

I analysed wheel movements, as described previously (Steinmetz et al. 2019). Briefly, the momentary wheel velocity was calculated by integrating the change in wheel position over a 50 ms window and I detected times when the velocity trace crossed a threshold to obtain possible transition times from non-movement to movement. Of these times, those crossing times were selected that occurred before the completion of the choice but after the stimulus. Then, to identify the time that choice movement was initiated, I looked sequentially backward from the recorded reward time to find the first time when the wheel position deviated more than 0.2 mm from 0 within a given trial.

While this wheel movement analysis is standard in obtaining a time when wheel movements were initiated, notably it suffers from bias in ballistics: movements initiated with high velocity might be detected earlier than slower movements where the difference in velocity change is smaller in each time bin, making the velocity threshold detection more precise. For neural data analysis, it is crucial to estimate the initiation time of the choice, so I used the initiation of the wheel movement as the time of choice throughout the neural data analyses. However, for behaviour-only models, I derived reaction times from the times when the decision threshold was reached, i.e. the time when the mouse brought the stimulus to the middle of the screen to account for variations in reaction times as the choice is being made (that manifest in how fast the wheel is being turned).

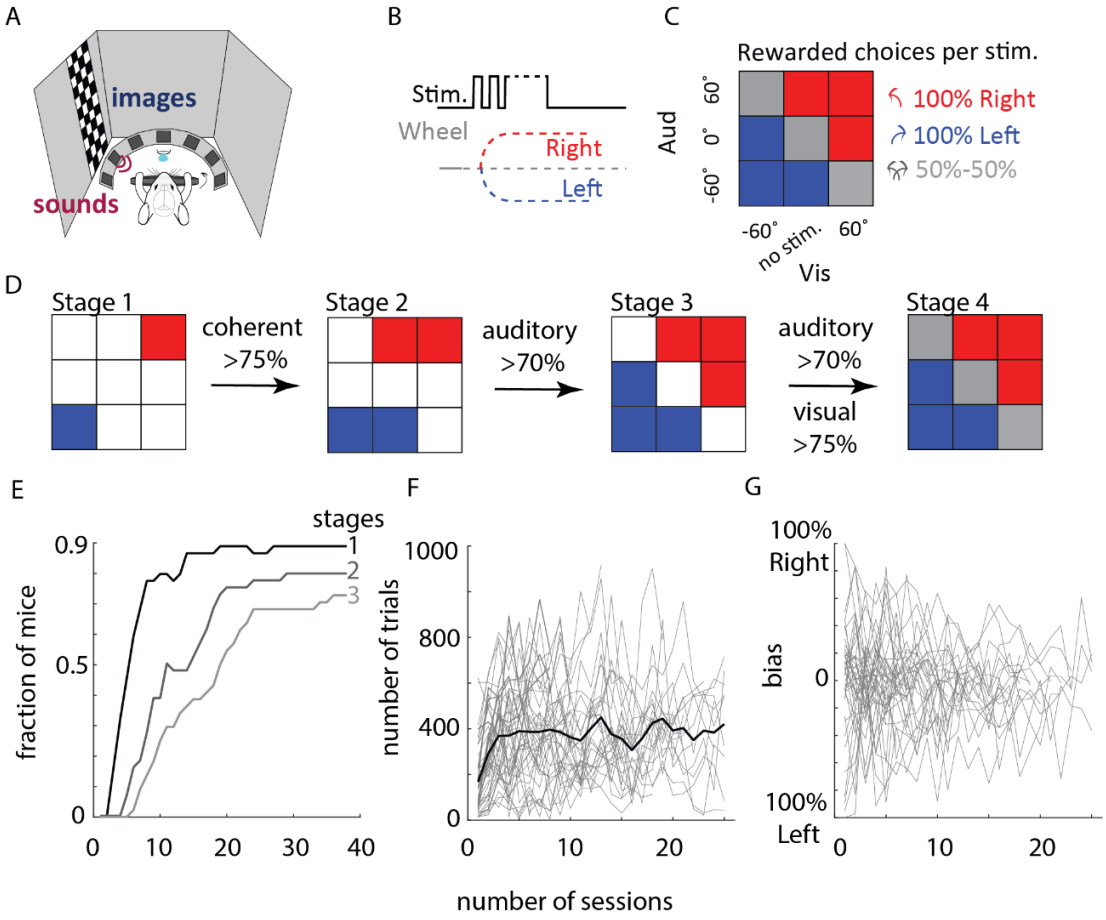



***FIGURE 10. AN AUDIOVISUAL SPATIAL LOCALISATION TASK.***

*(A) Illustration of the task setup. (B) Temporal trial structure. (C) Rewarded wheel turn directions for each stimulus condition in the final version of the task. (D) Automated training strategy. Initially, only coherent trials were introduced (stage 1), followed by auditory (stage 2), visual (stage 3), and then conflict trials (stage 4). Mice proceeded to the next stage if they passed threshold performance criteria and performed >200 trials. (E) Fraction of mice that learnt stages 1-3 across training sessions (n=44 mice). Occasionally, the experimenter downgraded the training stage, which typically occurred after holidays when mice were not trained for a prolonged period. (F) Number of trials performed across training sessions. For representative purposes, only the first 25 sessions are shown, which is a typical number of sessions to reach a fully trained stage. Light grey: various individual mice, black: average. (G) Left-right bias across training.*

## Models of choice in audiovisual task

In 2−AFC tasks, behaviour is often measured by the ratio of choices made for each option, which is the fraction of rightward turns made by the mouse in each stimulus condition for our audiovisual task. This relationship between performance and stimuli is often referred to as the psychometric curve, which empirically is most often a sigmoid (Figure 11A). The choices during the audiovisual task appear to also follow this trend and can be thus modelled using a logistic classification model (Raposo et al. 2014; Coen et al. 2023; Carandini, 2024). In this framework, the choices of the mouse are predicted by the logistic sigmoid function, such that the probability of rightward choices, $\mathrm{p(R)}$, is dependent on stimulus input $S$ weighted by $\alpha$ and a bias $B$:

$$\mathrm{p(R)} = \frac{1}{1 + \ exp^{-(\alpha S + B)}}$$

*Equation 1*

Instead of fitting $\mathrm{p(R)}$, it is often useful to fit the odds as it lends itself to linear regression methods, as:

$$\log\left(\frac{\mathrm{p(R)}}{1 - p(R)}\right) = \alpha S + \mathrm{B}$$

*Equation 2*

In this framework, the stimulus related term $\alpha S$ can be broken down to auditory input, visual input or even a possible audiovisual interaction term. However, empirically, it has been found that the interaction term is not needed to describe mouse behaviour in this audiovisual task, and so the choices of the mice were described by an additive combination of auditory and visual input (Figure 11B):

$$\log\left(\frac{\mathrm{p(R)}}{p(L)}\right) = v_R V_R^{\gamma} - v_L V_L^{\gamma} + a_R A_R - a_L A_L + \mathrm{B}$$

*Equation 3*

Where:

- $p(L)$, i.e. the fraction of the leftward choices, which equals $1 - p(R)$,
- $A_R, A_L, V_R$ and $V_L$ are task variables indicating auditory and visual stimulus strength on each side,
- $\gamma$ is a contrast scaling parameter,
- $a, a_L, v_R$ and $v_L$ indicate how much the mouse weights each sensory stimulus while making the choice.

To perform the fitting, I minimised the negative $\mathrm{Log_2}$likelihood using the *fmincon* package in MATLAB. I also implemented the same fitting procedure in Python, where I again minimise the negative $\mathrm{Log_2}$likelihood using the *scipy.optimize* function. To evaluate the goodness of fit, I assess the total negative $\mathrm{Log_2}$likelihood on a held-out dataset. To represent all trials as equally as possible in both the training and the held-out set, the splitting procedure followed the logic of the *StratifiedShuffleSplit* function in *scikit-learn*; that is, to classify each trial into a trial type category based on visual stimulus azimuth, auditory stimulus azimuth and the choice direction made, and then split each trial type separately to the training and the held-out set.

Notably, logistic classification is not the only way to model choices. In multisensory decision-making tasks, signal detection theory (SDT) models have historically been more popular. In the SDT framework, the sigmoid curve is modelled using cumulative Gaussians. Indeed, often there is no difference in fits when using these two methods. However, the logistic classification method allows for fitting more data points and is also more expandable; it can be used to predict from more than two predictors and can also predict an arbitrary number of choices (for a more comprehensive review, see Carandini, 2024). Indeed, for example, this model allows to easily include additional neural predictors (see Chapter 4)

and also allowed me to test the effects of inhibition by adding additional parameters that might be introduced by external perturbations (see Chapter 7).

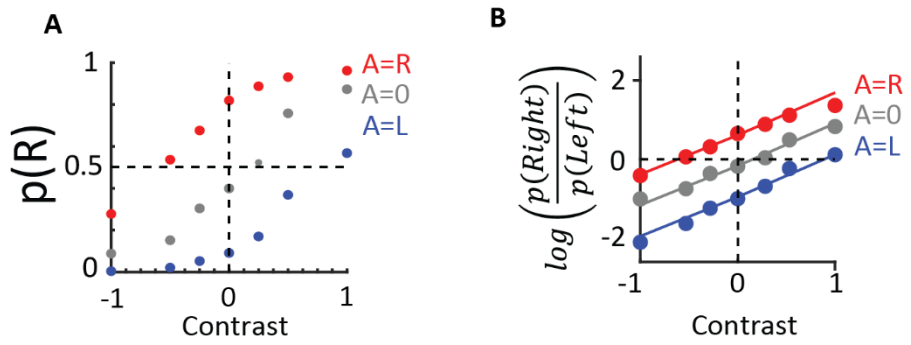



***FIGURE 11 LOG-ODDS MODEL DESCRIBES THE RELATIONSHIP BETWEEN STIMULUS AND CHOICE IN THE TASK.***

*(A) Example performance as measured by mouse choices in the audiovisual task, plotted by the fraction of rightward choices as a function of normalised visual contrast on the left (negative) vs the right side for each auditory stimulus location. (B) The same performance measures as in (A) but plotted in log-scale (i.e. by plotting the ratio of the fraction of left vs right choices instead of the fraction of right choices alone), fitted by the logistic classification model (Equation 3). If the choices are all in the same direction (e.g. 100 % rightward), the data point in the average log-Odds space would not be plottable, in these cases I add one trial type for each trial condition for visualisation purposes.*

## Electrophysiology

### Surgeries

On all mice, two surgeries were performed under isoflurane anaesthesia in total. First, a steel or titanium head plate was implanted on the mouse as described previously (for example, Bimbard et al., 2024). Typically, mice recovered from the first surgery were habituated to the rigs or trained in the task before the second surgery. During the second surgery, the mouse was either prepared for acute recording or implanted chronically as described before. For acute recordings, two 1.5−3 mm craniotomies were made bilaterally using a biopsy punch or a drill. The craniotomies were covered by duragel. For implanting 1.0 Neuropixels probes chronically, I made a single small (~.2 mm) craniotomy. To implant 2.0 probes, I made one or two 1.5 mm craniotomies bilaterally using a biopsy punch.

### Acute recordings with Neuropixels 2.0 probes

For acute Neuropixels recordings four-shank 2.0 Neuropixels (Steinmetz et al. 2021) were fixed on a custom rod and inserted using a manipulator (Sensapex). The probes were grounded by placing an Ag/AgCl wire into the water bath that covered the craniotomies and referenced externally through the same wire. Though the Neuropixels 2.0 probes contain four shanks, each with 5,120 recording sites, only 384 recording sites can be selected at a time for a recording and only in limited types of configurations (Figure 12).

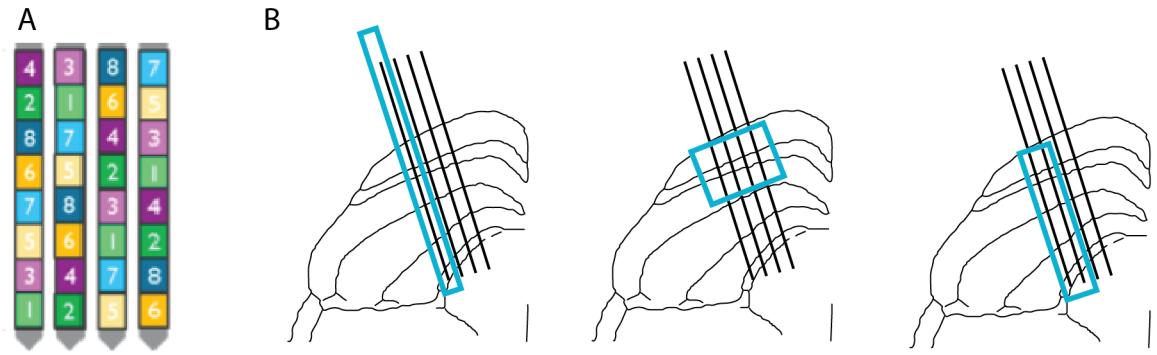



***FIGURE 12. ELECTRODE SELECTION IN NEUROPIXELS 2.0 PROBES.***

*(A) Illustration of the possible electrode selection principle of the Neuropixels 2.0 probe, adapted from the 2.0 user manual. There are 1280 passive electrodes on each probe, but only 384 acquisition channels. Each electrode is connected to an acquisition channel, but the electrodes illustrated with the same colour and number are connected to the same set of acquisition channels, and thus cannot be simultaneously selected for recording. (B) Illustrating the three main ways that I have used throughout this thesis to select electrodes to cover the SC. Left, a single shank configuration, which I typically used to locate the surface of the superior colliculus and thus guide the anatomical localisation of the SC. Middle, a four-shank configuration that I typically used during audiovisual experiments, since this yields the best coverage of SC. Right, a two-shank configuration that I occasionally used, but it is not possible to configure the probe like this at all depths, but only starting from the tip or 720 µm above the tip and so on, so this configuration adjusting the depth of the probe to the target region.*

## Chronic implants

### *Neuropixels 1.0 probes*

I initially implanted Neuropixels 1.0 probes (Jun et al. 2017) chronically, using a non-recoverable method (Figure 13A), as described previously. Briefly, for this, the Neuropixels probe is inserted and cemented onto the skull during the second surgery through a small craniotomy (Okun et al. 2016). The probe was grounded to the headplate externally or to the tip of the probe internally depending on which configuration showed less electronic noise.

### *Neuropixels 2.0 probes*

I have also participated in developing a recoverable chronic implant for Neuropixels 2.0 probes(Steinmetz et al. 2021), which, thanks to the configuration options of the Neuropixels 2.0 probes allows better coverage of the SC compared to 1.0 probes (Figure 13B-D). The detailed surgical strategy for this implantation is described in (Bimbard et al. 2024) . Briefly, the 2.0 probe is prepared in a two-compartment module that consists of a payload module, holding the Neuropixels probes, and is screwed together with a docking module that will be cemented to the skull after insertion. Upon recovery, the screws connecting the docking and the payload modules are removed, and the payload module together with the probes is recovered from the brain, while the docking module remains fixed to the skull (Figure 13B).

### *Advantages and disadvantages of the chronic technique in SC*

I started applying chronic implantations instead of acute recordings because chronic recordings do not require much preparation before recording (i.e. the process of probe insertion). This results in several advantages. First, with chronic recordings, it is easy to record several mice at the same time. Second, the mice do not need to undergo a lengthy procedure before executing the audiovisual task and thus are typically able to achieve higher trial numbers compared to acute recordings.

Additionally, in chronic recordings, the same neurons can be tracked across days, which might reveal common or changing patterns by various behavioural states, or physiological changes over time (Van Beest et al. 2024). Notably, initially, in my PhD, I tried to exploit chronic recordings to study changes during learning too. However, I found that the number of recorded units logarithmically decreases by the time more are fully trained in the task (Figure 13C). Furthermore, anecdotally, implants in more posterior positions, such as the SC may vary more in yield than more frontal implants (Figure 13D). To this end, I have not been able to analyse the results of tracking neurons across days in the context of this thesis. Indeed, in this the thesis, when a chronic recording was included, I ensured that a given part of the probe was only included once in the analysis.

The main disadvantage of the chronic technique is that it does not allow reinsertion and thus sampling of a different population within a given mouse. Therefore, the total yield of chronic recordings is lower from the same mouse compared to acute recordings.

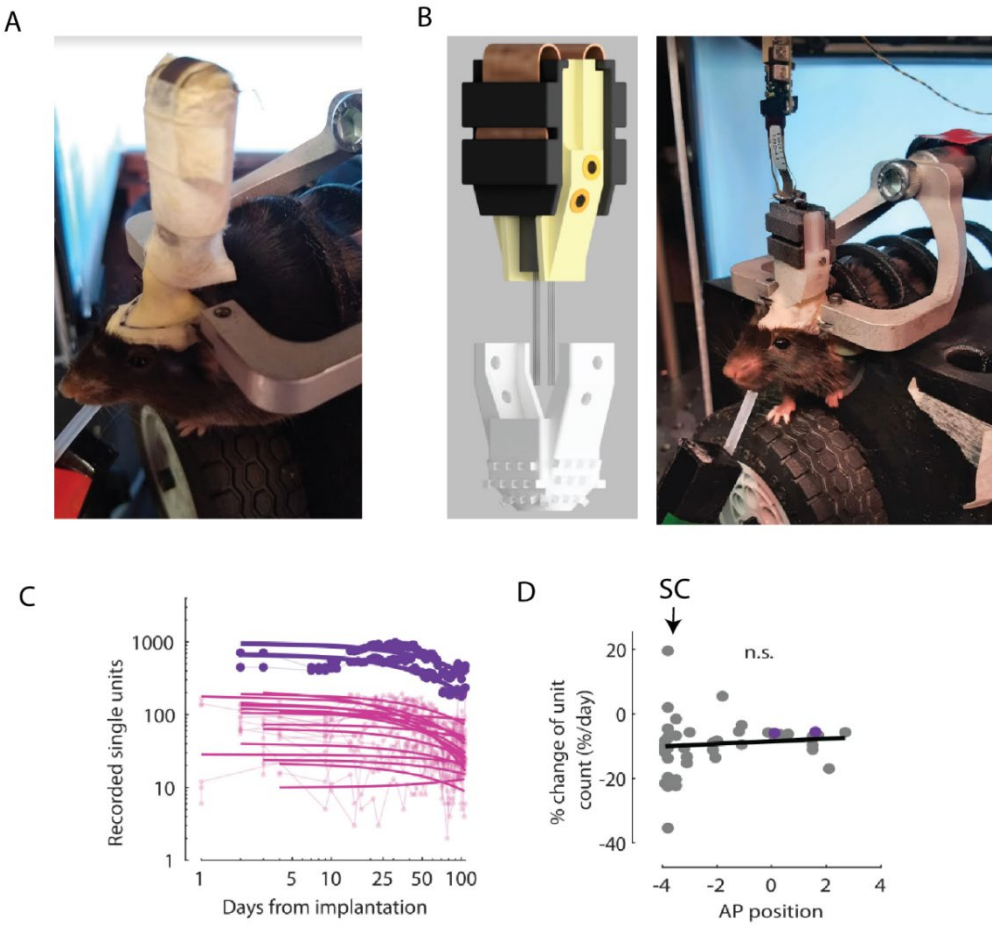



***FIGURE 13. CHRONIC NEUROPIXELS RECORDINGS.***

*(A) Example implant prep of a non-recoverable 1.0 probe implant. (B) The Apollo implant for recoverable 2.0 probe implants. Left: illustration (by Pip Coen) of the implantation strategy: the probes are attached to the payload module held by the white docking module cemented to the skull. The payload module can host up to two 2.0 probes. Right: example implant prep.(C) Number of recorded units per recording bank (pink) or across the entire probe (purple) for a given implantation using the Apollo implant. Lines are logarithmic fits. (D) Change in single unit count as a function of AP position relative to bregma, where the SC is ~3.8mm posterior to bregma. Both (C) and (D) are adapted from (Bimbard et al. 2024).*

### *Mice with chronic recordings*

Altogether I recorded 14 male and 5 female mice chronically. This sex imbalance was because the weight of the implant could not exceed 12 % of the body weight at the time of implantation. Furthermore, when implanting mice trained in the task, I performed the implantation on water-restricted mice. Male mice were more likely to pass the criterion set by the weight limit.

## Locating the SC surface

### *During probe insertion*

After each insertion, approximate SC depth was determined using sparse noise stimuli. For this, the raw data were high-pass filtered with a seventh-degree Butterworth filter at 600Hz and averaged every

10 ms, and across eight neighbouring channels. Subsequently, the stimulus-response average was computed for every square location 50−150ms post-stimulus onset (Figure 14).

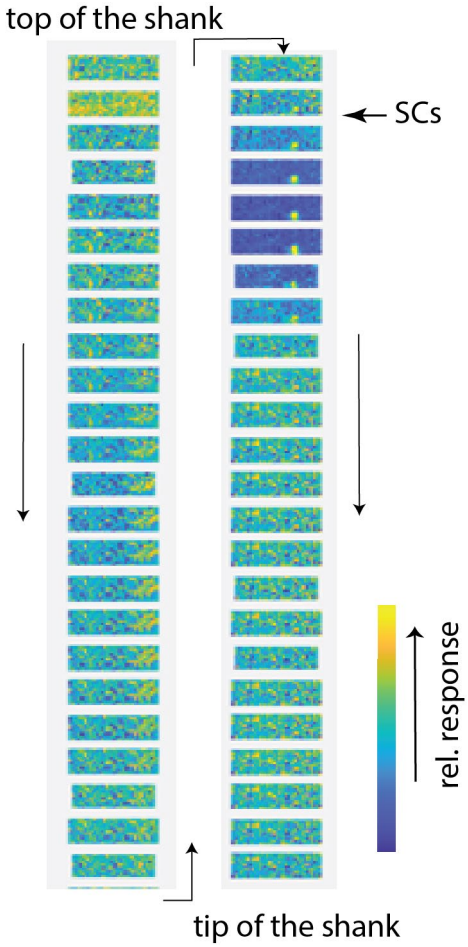



***FIGURE 14. LOCATING THE SUPERIOR COLLICULUS SURFACE AS ACUTE RECORDING IS PERFORMED.***

*Example filtered responses to 5 minutes of receptive field mapping in single shank configuration. Each square is the normalised average response to screen, averaged across eight channels (60 µm). Upon successfully locating the SC based on the receptive fields, I typically selected the two- or four-shank configuration at the appropriate position to ensure optimal SC coverage.*

### *Post-hoc*

I characterised the visual receptive field of both single units and multiunit activity at 60 µm bins (Figure 15A). For this, I used a two-dimensional Gaussian tuning curve (Figure 15B) to predict the average responses after each white square onset (in a 60 ms window, 10 ms after stimulus onset, baseline subtracted) during the sparse noise stimulus display, such that:

$$r_{\gamma,\varphi} = a * e^{-\left(\frac{(\varphi-\varphi_{pref})^2}{2\sigma_{\varphi}^2} + \frac{(\gamma-\gamma_{pref})^2}{2\sigma_{\gamma}^2}\right)}$$

*Equation 4*

Where:

- $\varphi$ denotes azimuthal position of the stimulus square,

- $\gamma$ denotes the elevation of the stimulus square,
- $r_{\gamma,\varphi}$ denotes the neural response measured when a white square turned on at a given elevation ($\gamma$) and azimuth ($\varphi$).
- $\varphi_{pref}$ is the fitted parameter describing the preferred azimuth of a given neuron,
- $\gamma_{pref}$ is a fitted parameter describing the preferred elevation,
- $\sigma_{\varphi}$ and $\sigma_{\gamma}$ are fitted parameters describing the width of the tuning curves in azimuth and elevation,
- $a$ is a constant scaling factor that accounts for the magnitude of the response at the preferred azimuth.

Fitting was initialised, using data-driven parameters, and the fitting was performed using non-linear least squares via the *scipy.optimize.curve_fit* function (Virtanen et al. 2020). Fits were evaluated using the $R^2$ score during cross-validation. Significance was determined using the permutation test. For this, both the azimuthal ($\varphi$) and elevation ($\gamma$) labels were shuffled 20 times, and the tuning curves were refitted each time. A neuron was considered spatially tuned if its $R^2$ score was higher than any of the $R^2$ scores obtained after shuffling. Fitting was performed as described in the one-dimensional case, except both the azimuthal ($\varphi$) and elevation ($\gamma$) labels were shuffled 20 times during the permutation test (i.e. $p<0.05$).

Visual receptive fields are densest in the SCs. Therefore, the location of most well-fitted units can indicate probe location along the dorsoventral axis. Conversely, the preferred azimuth and elevation are defined by the anatomical location of the unit along the mediolateral to the anterior-posterior plane. Therefore, the visual receptive fields were used to guide the anatomical registration of the probe insertions (Figure 15C).

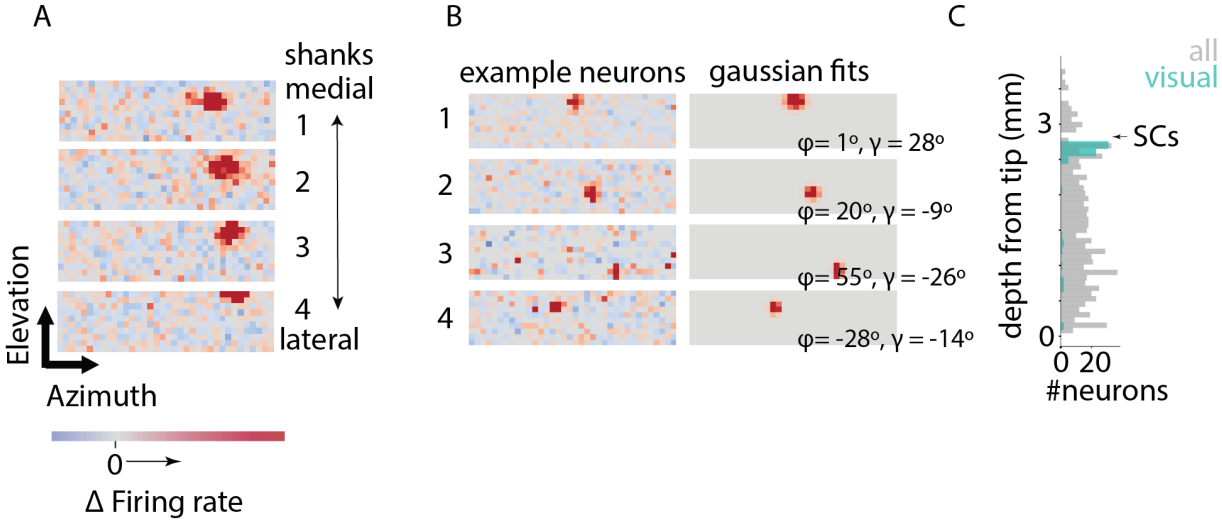



***FIGURE 15. DETERMINING THE VISUAL RECEPTIVE FIELD.***

*Multiunit responses show a receptive field from each shank from a 2.0 probe inserted into the SC, indicating that the probe orientation in the SC can be determined using the relative positions of the receptive fields. (B) Left: responses of isolated neurons. Right: prediction of the 2D Gaussian with its preferred azimuth and elevation. (C) Distribution of units per probe depth. Grey: all units. Green: units with a significant visual receptive field, as determined by the 2D Gaussian fitting.*

## Preprocessing electrophysiological data

### *Spike sorting*

All recordings were sorted with pyKilosort.
(https://github.com/MouseLand/Kilosort/releases/tag/v2.0) (Pachitariu et al. 2024).

### *Quality control*

Then I applied quantitative quality metrics to identify well-isolated units using Bombcell (Fabre et al. 2023). Criteria for unit selection were essentially the default criteria, which aim to discard noise and multiunit activity based on a set of thresholds applied on waveform properties and the temporal distribution of spikes of a particular cluster during the recording.

### *Registering recording site position to the Allen Common Coordinate Framework*

To reconstruct probe tracks, probes were coated with CM- DiI (ThermoFisher, C7000) or Qtracker 625 (ThermoFisher, A10198) dyes before insertion to track the probe location. All mice were perfused transcardially, and the brain was removed and post-fixed in 4 % formalin-PBS for 24 hrs, and subsequently stored in PBS. Brains were then mounted into agarose, imaged using two-photon serial tomography (Ragan et al. 2012), on a custom-built microscope, and pre-processed with custom software (https://bakingtray.mouse.vision/). For imaging probe tracks, the excitation wavelength was 780 nm or 920nm, and brains were imaged in two or three channels (Red: ET570lp, Green: ET525/50m, Blue: FF01-450/70) typically imaged at 10x10x25 µm resolution. All images were downsampled to 25*25*25 µm voxel size, aligned to the Allen Mouse Brain Atlas (Q. Wang et al. 2020) using Brainreg and probe tracks were manually traced in atlas space using Brainreg-Segment (Niedworok et al. 2016) (Claudi et al. 2020) (Tyson et al. 2022) (Figure 16A-B). To correctly localise the location/depth of each shank, Neuropixels channels were aligned to the atlas using the IBL atlaselectrophysiology-GUI (https://github.com/int-brain-lab/iblapps/tree/master/atlaselectrophysiology), typically using the ventricle between the SC and the cortex as a landmark, as well as the typical strong spontaneous correlation across channels in the superficial SC.

As typically the four-shank configuration of the 2.0 probes does not allow capturing electrophysiological landmarks, an entire bank was recorded separately of all shanks, each was aligned to the Allen Common Coordinate Framework (CCF) separately, and the channel location during the four-shank configuration was subsequently inferred from its location during this recording. In chronic recordings, I aligned each probe to the Allen Atlas only once, typically at the earliest available full shank recordings, and I used the channel location on these days as approximate channel locations across all days, assuming no drift. Correct reconstruction was validated by checking the location of neurons with reproducible visual receptive fields (Figure 16C).

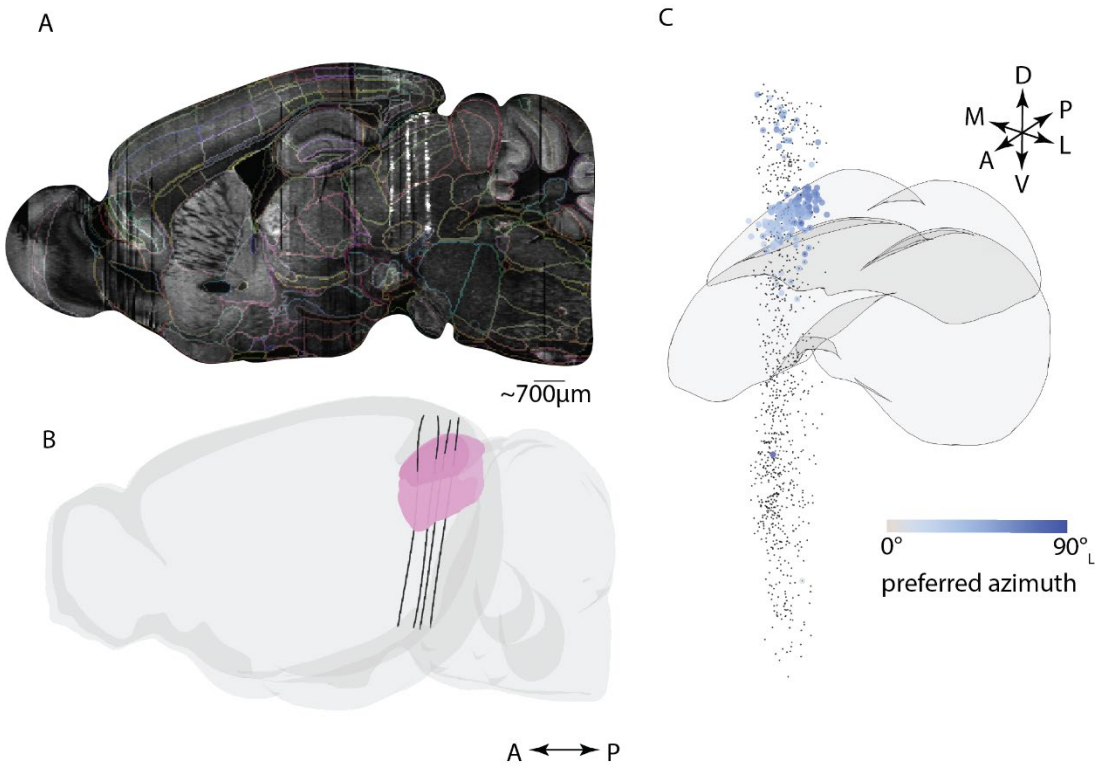



***FIGURE 16. REGISTERING THE POSITION OF NEURONS TO THE ALLEN COMMON COORDINATE FRAMEWORK.***
*(A) Example image showing a chronic four-shank probe passing through the SC, imaged with serial tomography. Contours denote area borders defined by the 25µm Allen Atlas. (B) Reconstructed probe tracks in Allen CCF space. (C) Three-dimensional location of neurons in an example insertion. Grey: all neurons. Coloured: neurons with significant receptive fields on a single shank recording. The hue of the colour denotes the approximated preferred azimuthal position on the left side of the screen (i.e. contralateral to the recorded site).*

### *Calculating firing rate*

Unless otherwise specified, the firing rate was estimated by binning the spike times into 10 ms bins and smoothing the spike counts using a causal half-Gaussian filter with a standard deviation of 25 ms. To estimate baseline subtracted firing rates, the baseline was calculated as the average firing rate in a 600ms window before the stimulus.

# Chapter 3 — Auditory and visual signals in the SC

We know that mice can spatially localise the checkerboard images and pink noise burst sounds used in the 2-AFC task (Coen et al. 2023). The previously described auditory (Palmer and King 1982; King and Hutchings 1987; Ito et al. 2020; Middlebrooks and Knudsen 1984) and visual (Apter 1946) spatial maps in the SC lend themselves to represent these task stimuli even in naïve mice and may contribute to the learning and/or the execution of the spatial audiovisual localisation task. This chapter aims to provide a coarse description of how task stimuli are represented in the SC in naïve, awake mice. Thus, I aim to answer: 1) are the checkerboard images represented topographically, forming a visual map? 2) Are the pink noise bursts represented topographically, forming an auditory map? 3) Are these maps aligned anatomically? 4) Are there neurons that respond to both sounds and images and do the audiovisual neurons also form a topographic map?

## Methods

### Experiments

In this section, I pooled data across 10 acutely and chronically recorded mice (6 male and 4 female) that were not yet trained in the task, of which 8 were from the C57BL/6J strain, while 2 were high-frequency hearing loss corrected ($Cdh23^{+}$). The mice were presented with stimuli with a variety of contrasts and SPLs. I recorded 1,747 well-isolated neurons in 14 experimental sessions across the 10 mice in total. In this chapter, I focused on analysing the unisensory responses for audition and vision in single neurons.

### Testing the responsivity of each neuron for auditory and visual stimuli

To detect auditory and visual stimulus-evoked responses presented at various azimuths, I performed a permutation test. For this, I took the average baseline subtracted firing rate at each in a 200 ms response window and took the maximum of the absolute value (to account for inhibition) of the response ($R_{max}$). Subsequently, the test statistic was obtained by subtracting $R_{max}$ on blank trials from $R_{max}$ on stimulus trials. The null distribution of test statistic was obtained by shuffling blank and stimulus trials 2000 times, and p-values were Bonferroni corrected for the 14 azimuths presented.

### Calculating preferred azimuth

Spatial tuning of neurons with significant auditory or visual responsivity was assessed (Figure 17). For this, I obtained one number ($r_\varphi$) to summarise the response at each of the 7 azimuths. Most frequently $r_\varphi$ is calculated by taking the baseline-subtracted the average response in each time window for a particular azimuth. However, I observed that especially for auditory responses, this time window is not simple to determine, as auditory responses might often have complex temporal profiles (Figure 17B). Therefore, I reduced the temporal dimension using singular value decomposition or a trial-by-time matrix, where time was instead by taking a 200 ms response window following stimulus onset. Responses were baseline subtracted, and the weight of the first singular value along the trial axes was averaged for each of the 7 azimuths to obtain $r_\varphi$.

This method is very similar to taking just the average response at each azimuth in a response window; however, this method ensures the temporal components of the trials are optimally summed up by one

number and in a sense is more optimal than taking a temporal average (cf. Eckart-Young theorem, suggesting that the optimal rank-r approximation to a matrix is given by the rank-r singular value decomposition, Brunton and Kutz 2019).

To obtain the tuning curves, the responses at each azimuth ($r_\varphi$) were 100-fold upsampled, and fitted using Gaussian function, such that:

$$r_\varphi = r_{min} + (r_{max} - r_{min}) * e^{-\frac{(\varphi - \varphi_{pref})^2}{2\sigma^2}}$$

*Equation 5*

Where:

- $\varphi$ denotes azimuthal position of the stimulus,
- $r_\varphi$ denotes the neural response at a given stimulus,
- $\varphi_{pref}$ is the fitted parameter describing the preferred azimuth of a given neuron,
- $\sigma_\varphi$ is a fitted parameter describing the width of the tuning curve,
- $r_{min}$ and $r_{max}$ are the measured minimum and maximum of the neural response across all azimuths, and they account for the baseline firing rate and the response magnitude in the equation.

Fitting was performed similarly to the two-dimensional Gaussian curve fitting (Equation 4). For this, fitting was initialised, using data-driven parameters, such that $y_{min} = \min(r_\varphi)$, $y_{max} = \max(r_\varphi)$ and $\sigma = \frac{10 * (y_{max} - y_{min})}{(y_{max} + y_{min})}$. Then, fitting was performed using non-linear least squares via the *scipy.optimize.curve_fit* function (Virtanen et al. 2020), and evaluated using the $R^2$ score. Significance was determined using the permutation test, for which azimuthal labels were shuffled 20 times, and the tuning curves were refitted each time. A neuron was considered spatially tuned if its $R^2$ score was higher than any of the $R^2$ scores obtained after shuffling.

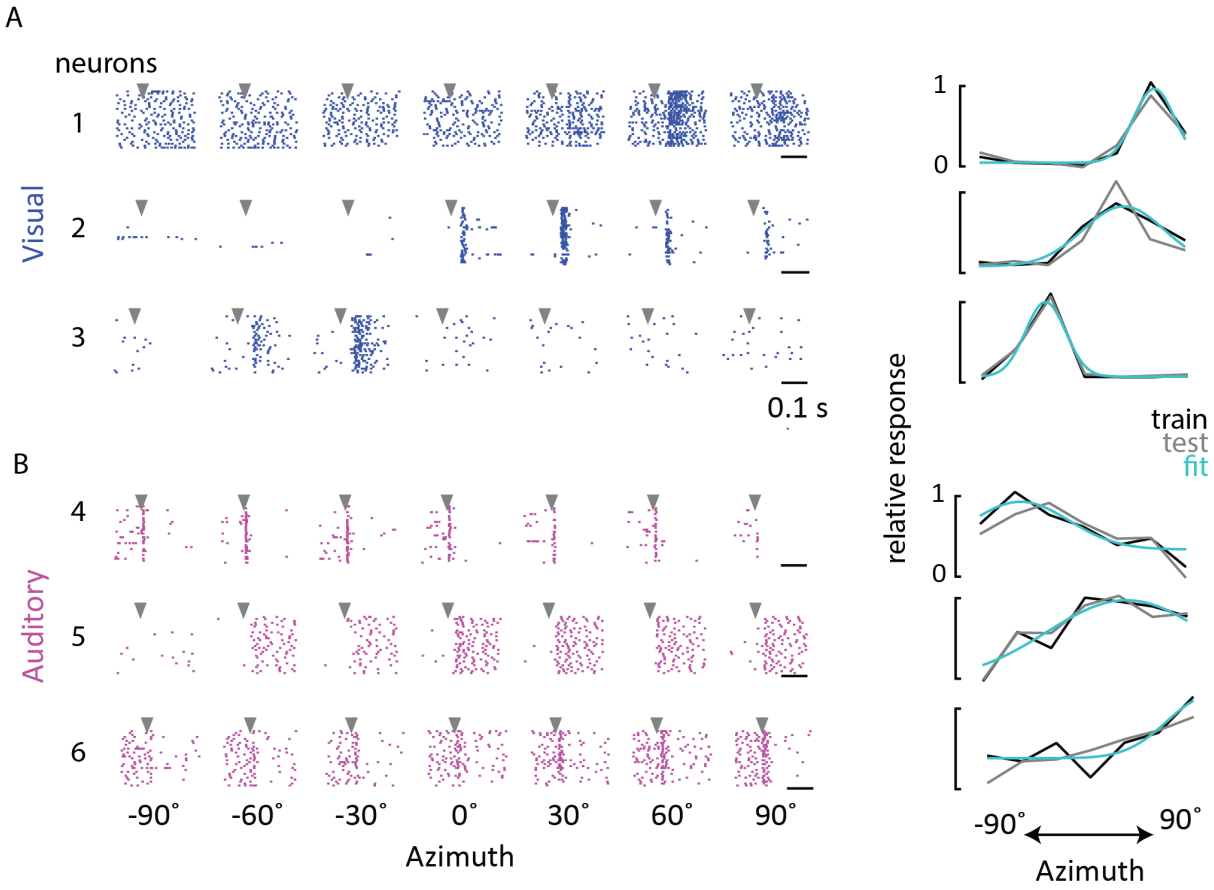



***FIGURE 17. VISUAL AND AUDITORY RESPONSES IN THE SUPERIOR COLLICULUS OF NAÏVE MICE.***

*Raster plots of example visual (A) and auditory (B) neurons responses across 7 different azimuths and their Gaussian tuning fits, shown together with the cross-validated tuning curves (left).*

## Results

### The visual spatial tuning

As expected, I found neurons responding to the checkerboard images across the SC (Figure 17A). 31 % of the recorded neurons showed some response to visual stimuli at any location, and 20 % of all recorded neurons showed reliable spatial tuning for vision in the azimuthal space that I explored (Figure 18A). Visual neurons were distributed across layers, but more neurons responded to visual stimuli in SCs compared to SCm, as expected (Figure 18B). Neurons spanned the entire azimuthal space (although fewer neurons were sampled at 90° azimuth) and were typically narrowly tuned with a median sigma of 16.1˚(Figure 18C), which is around the limit of what we could resolve with the checkerboard stimulus size used in this paradigm (30° wide). The actual receptive field sized of visual neurons, especially in the SCs tend to be 4°-8° which may vary by cell type (Hoy et al. 2019; Li and Meister 2023).

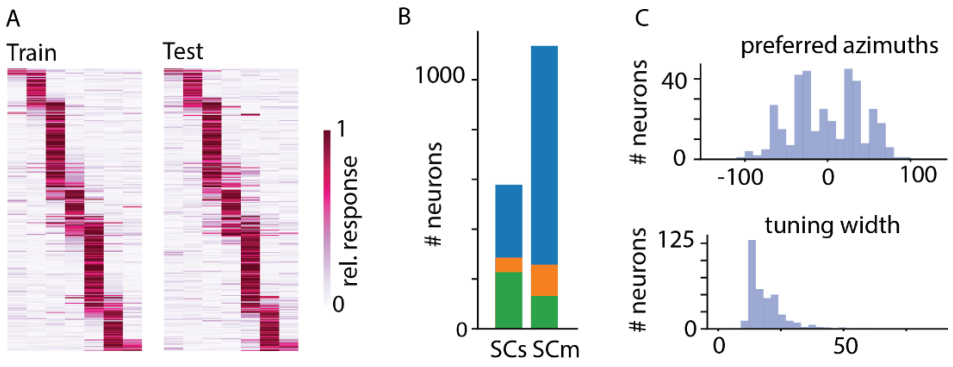



***FIGURE 18. THE VISUAL SPATIAL MAP.***

*(A) Relative response across azimuths for each visually tuned neuron in the training and test set. (B) Comparison of the total number of neurons recorded in each layer of the SC (blue) with the number of responsive (orange) and spatially tuned neurons (green). (C) distribution of the preferred azimuths and the width of tuning curves of each spatially tuned visual neuron.*

### The auditory spatial tuning

26 % of recorded neurons responded to auditory stimuli, but only 8 % showed reliable spatial auditory tuning (Figure 19A). Most auditory neurons were in the SCm, as expected (Figure 19B). Neurons spanned the entire azimuthal space explored but were more enriched towards the lateral edges of the azimuthal space, contrary to visual neurons (Figure 19C). Auditory spatially tuned neurons were more broadly tuned than visual neurons, with a median sigma of 31.2°. Notably, many neurons showed a monotonic tuning curve, suggesting that their preferred azimuth is bigger than the measured azimuthal range (±90°), and the preferred azimuth and tuning width of the auditory neurons is, therefore, less certain compared to neurons that peaked in the explored range.

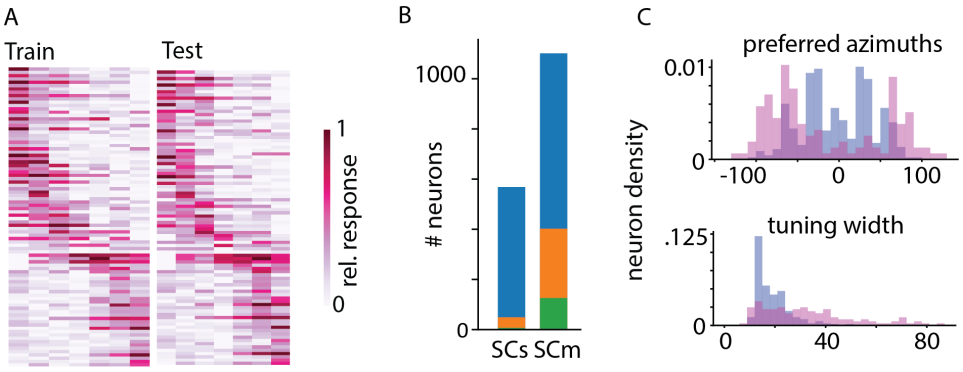



***FIGURE 19. THE AUDITORY SPATIAL MAP.***

*Both (A), (B), and (C) are the same as for visual neurons. In (C) the auditory neurons are plotted in purple, and the visual neurons are plotted in blue for comparison.*

## The alignment of the visual and the auditory map

To determine whether the either the visual or the auditory neurons form a spatial map, I registered all recorded neuron's location to the Allen CCF (Figure 20A). Then, I compared the average preferred visual vs auditory tuning of neurons in 150 µm bins along the anteroposterior axis (Figure 20B). I found that the visual map was more refined than the auditory map. Indeed, although the maps were in rough alignment (Figure 20C), however, the auditory map appears coarser.

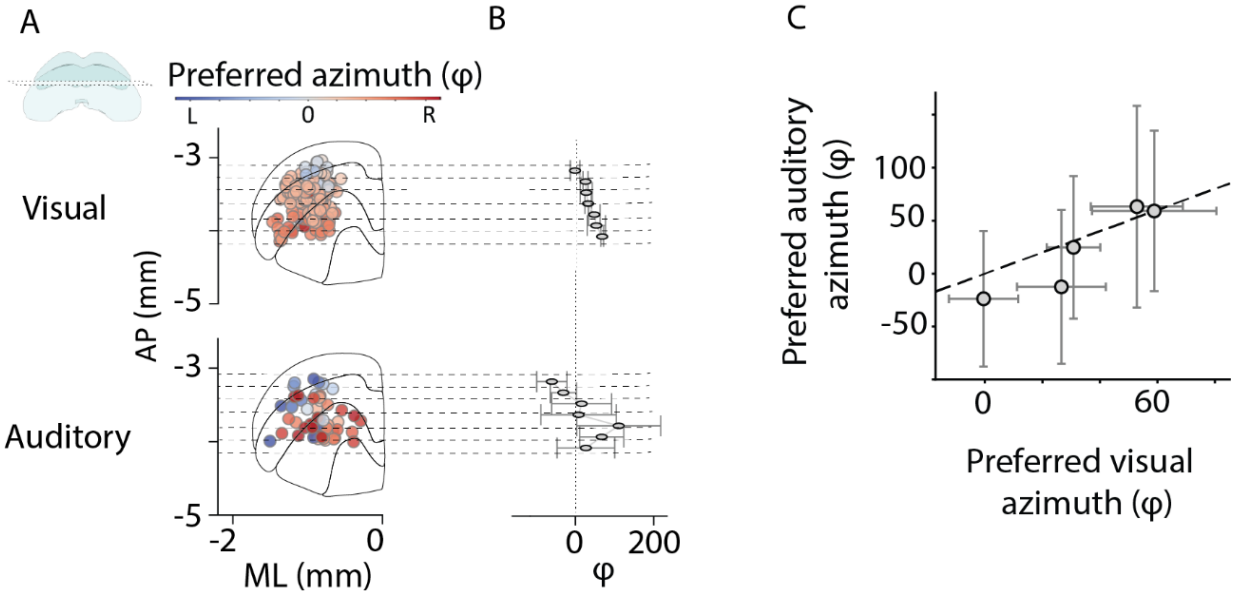



***FIGURE 20. THE ALIGNMENT OF VISUAL AND AUDITORY MAPS IN THE SUPERIOR COLLICULUS.***

*The location of the visual (top) and the auditory (bottom) spatially tuned neurons in the medio-lateral-antero-posterior plane, coloured by the preferred azimuth of each neuron. Neurons across all depths of the SC were pooled together and projected onto this plane for which the contours are at -2mm from bregma along the dorsoventral axis. For representational purposes only, I jittered the location of neurons with a random number from a Gaussian distribution with 40µm in both the mediolateral and anteroposterior axes. (B) the average preferred tuning of visual (top) and auditory (bottom) spatial neurons in 150 µm bins along the anteroposterior axis. (C) preferred azimuth of visual vs auditory neurons in the same topographical location (r=0.73).*

## Audiovisual neurons

To determine whether there is a distinct functional set of neurons – the integratory audiovisual neurons -- in the SC that could, in principle solve the spatial audiovisual task, I examined neurons that were spatially tuned for both audition and vision (Figure 21). I found that this was a mere 28 neurons, i.e. 1.6 % of the entire recorded population. Some of these neurons had clearly aligned receptive fields, but in other cases the receptive fields appeared misaligned.

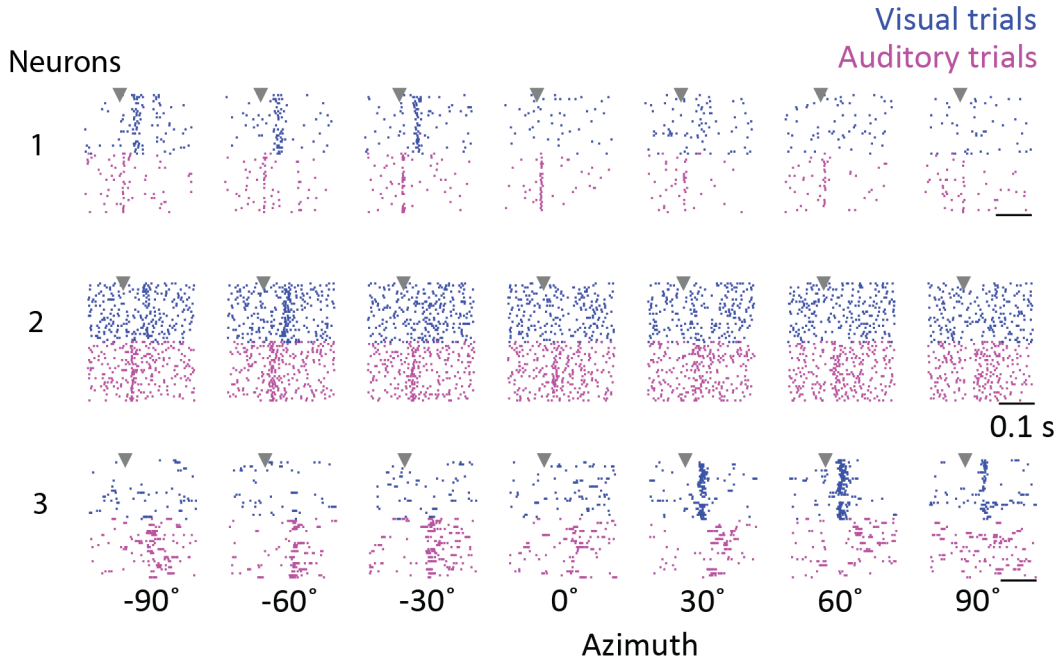



***FIGURE 21. EXAMPLE AUDIOVISUAL NEURONS.***

*Raster plots of the responses of example audiovisual neurons. Magenta denotes auditory trials while blue denotes visual trials.*

## Discussion

My results agree with many others who studied the superior colliculus in various species. There exists a visual layer of the SC in the superficial layers, and auditory spatial information is topographically represented in the deep layers. Notably, this is the first time that the auditory map of SC has been shown in mice using free-field stimuli to my knowledge. It is possible that the stimuli I chose don't drive SC neurons the most ideally; however, previous studies also did not use complex stimuli, but only LEDs with different luminance and sound bursts.

My results contradict the previous literature in two aspects: first, I found the auditory map is less refined than the visual map. Indeed, there is not a clear 1:1 mapping between the auditory and visual maps that would allow co-registration. Second, I did not find many audiovisual neurons. Here, I will attempt to critically assess why this might be the case, and then I will reflect upon the initial question: given the representation of auditory and visual-spatial signals used in the audiovisual task, can the SC possibly solve the audiovisual task?

### The auditory spatial map and co-registration

The idea that the visual map interacts with the auditory map in a refined manner comes from the barn owl literature, where primarily the auditory representation of space was studied. Perhaps the most influential experiment coming from this line of work was that of the prism experiment, i.e. when juvenile barn owls were raised with a prism in front of their eyes, such that their field of vision was systematically distorted (Knudsen and Knudsen 1989; Knudsen and Brainard 1991; King 2002). In these owls, the collicular auditory map was equally distorted in adulthood, giving rise to the idea that potentially the visual and the auditory maps functionally interact and that they are co-registered. However, crucially, these experiments were performed in barn owls – a species in which eye movements are restricted (~2°, Harmening and Wagner 2011) – and thus in the owl, both visual and auditory maps are in a head-centred coordinate system. Indeed, in mammals, that can move their eyes, visual information is retina-centred, and therefore it requires a step of transformation to co-register the auditory and visual maps into a common reference frame, a phenomenon referred to as Pöppel's paradox (Harris et al. 1980). Indeed, it is possible that this co-registration is not a necessary computation for the colliculus function, but rather the visual input serves as a scaffold for the development of deeper layers. Indeed, collicular neurons are multisensory during development, but largely unisensory in adulthood in the mouse (Guillamón-Vivancos et al. 2022).

It is possible that I was unable to describe refined co-registration due to the technical limitations I described in Chapter 2 regarding auditory stimulus presentation which might have prevented me from describing a more refined auditory map. Therefore, it's possible that I measured only a noisy representation of auditory space in the SC, as the chamber wasn't anechoic due to the presence of screens, cameras, and other equipment in the rig. There is also potential noise I have not accounted for that is the result of performing measurements in awake mice: pinnae or eye movements can introduce untracked variability in the spatial tuning data.

It is also possible that the stimuli selected did not allow me to describing cleaner maps. First, I presented auditory stimuli only in the -90° to 90° azimuthal range. In this range, numerous auditory neurons appeared monotonically tuned, in which case their preferred tuning is harder to identify. Indeed, numerous studies that investigated the auditory spatial tuning of SC neurons often used the entire larger azimuthal range (360°) and moved a speaker on a motorised hoop for auditory stimulus presentation (for example King and Palmer 1983); or presented a larger dynamic range using HRTFs (Ito et al. 2020; Si et al. 2022). Second, for example, the receptive fields of auditory neurons might be shaped by their spectral preferences, especially when the sound source is located at the front of the animal (Ito et al. 2020). Therefore, my ability to find neurons in the frontal auditory fields might depend on the spectrum of stimuli I use, and it is possible that spatial auditory tuning is more apparent at a high-frequency range. Finally, loudness and the distance of the stimuli might also influence the auditory responses, and therefore my ability to infer spatial tuning of auditory neurons (Palmer and King 1982; King and Palmer 1983; n.b. Middlebrooks and Knudsen 1984).

Another potential explanation for why we saw a less refined auditory map is related to the animal strain. Indeed, most of these measurements were made in the C57 mouse strain which is known to suffer age-related hearing loss at the high-frequency range. Even though in this thesis, I used sound stimuli that are low frequency (8-16 kHz), hearing loss could have prevented the formation or maintenance of auditory spatial maps; indeed (Si et al. 2022) reported that frontal auditory receptive fields are missing in C57BL/6 mice; instead frontal receptive fields are intact C57 mice that are corrected for hearing loss (i.e. in the $Cdh23^{+}$ mouse line I used in later experiments).

All in all, this chapter does not aim to question previous reports of a possibly more refined map of auditory space in other species or even in mice. However, it does aim to suggest that in conditions in which we know mice can behaviourally discriminate the sound locations (Coen et al. 2023), the auditory map is not refined, and the visual and auditory maps do not appear to be recruited in a manner where co-registration is detectable.

### Lack of audiovisual neurons

In this chapter, I defined audiovisual neurons that have spatial tuning to both audition and vision when being presented with unimodal stimuli. These measures are hard to compare to early studies in the SC used microelectrodes until the experimenter found an audiovisual neuron, therefore it is hard to estimate the fraction of audiovisual neurons in these recordings. However, a recent audiovisual study in mice also observed only 5 % of recorded neurons to be bimodal by the same definition (Ito et al. 2021). If my measurements of the fraction of auditory and visual neurons are correct (which is hard to know because spike sorting is biased towards finding neurons that spike more), we would expect to find 0.2 *0.08 = 1.6 % of the neurons to be audiovisual by chance, exactly as I measured. This would mean that the auditory and visual neural signals are mixing approximately at chance level, suggesting that there is no explicit evolutionary pressure that binds auditory and visual maps together into an audiovisual representation.

It is also possible to define audiovisual neurons by other metrics, that may perhaps be less restrictive; indeed, a neuron may appear unimodal when presented by auditory or visual stimuli individually, but when these two stimuli are presented together, the neural responses appear modulated compared to the unimodal case. By this definition, for example, Ito et al. (2021) found a more compelling 32% SCm neurons to be audiovisual. In this chapter, I have not analysed responses for multisensory stimuli (to which I will turn to in Chapter 6). However, anecdotally, I found that multisensory modulations are often not reproducible in cross-validation. Nevertheless, it is possible that a looser definition of audiovisual neurons, or a more careful account of other modulation sources such as whisker or pinnae movements could provide further clarity on the extent of interactions of auditory and visual signals in single SC neurons.

Furthermore, it is also possible that just like with the measurement of the auditory maps, the measurement of a richer audiovisual interaction would have been possible using different stimuli. For example, the speakers were located just below the screens in my setup. Although in principle, since auditory receptive fields are bigger than visual receptive fields, this should not prevent interaction; however, in practice, audiovisual integration might be more prevalent if stimuli are perfectly colocalised. It is also possible that we would have observed a different extent of joint auditory and visual unimodal responses in single SC neurons if I had presented different visual and auditory stimuli (for example looming visual stimuli or higher-frequency content auditory stimuli, that are thought to drive SC neurons more extensively).

Nevertheless, in this Chapter, I show that SC neurons are spatially tuned for auditory and visual stimuli used in an audiovisual localisation task that mice can solve even in naïve mice. I also show that task stimuli are not frequently integrated in audiovisual neurons. Thus, SC auditory and visual neurons may contribute to the task, but it is unclear how; maybe audiovisual neurons are more prevalent in mice that learn the task, or sensory signals from SC sensory neurons are integrated downstream. Alternatively, it is also possible that the SC does not participate in the audiovisual localisation task, despite its rich representation of both auditory and visual space.

# Chapter 4 — Signals in the SC during an audiovisual task

In Chapter 3, I demonstrated that even in naïve mice, neurons in the superior colliculus (SC) represent visual and auditory stimulus location. Thus, in principle, even a small subset of SC neurons could solve the spatial audiovisual localisation task. However, that does not mean that these collicular neurons are actively recruited to solve the task. In this chapter, I begin to investigate the activity of SC neurons while the mouse performs the audiovisual localisation task. Specifically, I describe the response profiles of single neurons during task performance and determine whether each neuron is selective for visual or auditory stimulus position, or choice direction. Additionally, I will also assess the predictive power of SC neurons in decoding choices during task performance by testing whether SC neurons better predict decisions than the log-odds model described in Chapter 2.

## Methods

### Experiments

I recorded neural activity using chronically implanted Neuropixels probes in 11 Cdh23$^{+}$ mice while they performed the audiovisual task. 9 of these mice were male, and 2 were female. The first 8 mice were shown visual stimuli with 5, 10, 25 % contrast and auditory stimuli and 75 dB SPL. For the remaining 3 mice, I adjusted the stimuli to 10, 20, and 40 % contrast and reduced the auditory SPL to 65 dB. All probes targeted the SC, but I also recorded nearby brain regions on the same probe trajectory. These additional regions were used for comparative statistics alongside the SC data. In total, I recorded the 11 mice in 135 experimental sessions, only including sessions where the mice completed at least good trials 150 trials (Appendix A, Table 3 for details). I considered all trials good, except 1) trials that were repeat trials after an incorrect choice, 2) trials where a choice was initiated less than 30ms after stimulus onset, and 3) Trials where a choice was not initiated within the response window.

First, I analysed how selective each neuron is to different task conditions, such as the upcoming choice. This analysis is computed separately for each neuron, and thus it requires me to ensure that I don't double-count the neurons in the summary statistics. For this, I selected one session for each 720µm section in the 4-shank configuration (Figure 12B) to select neurons from non-overlapping areas, though I did not account for potential electrode drift. This process yielded 3,762 neurons across 45 sessions (Appendix A, Table 3, behaviour + unique ephys positions) with 1,127 located in the SC (Figure 22). Brain regions with fewer than 30 neurons were excluded from the summary plots.

The second analysis focused on whether SC neural activity could predict the animal's choices, computing summary statistics by session rather than by neuron, but requiring sufficient neurons from the same region for a fair selection pool. Therefore, for this, I computed summary statistics for each selected region of interest (ROI) – the SCs, SCm, MRN, and RSPd – by selecting sessions that contained a minimum of 20 neurons from a given ROI, and then I used a linear mixed effects model to obtain summary statistics. In total, I selected 20 sessions for the SCs from 6 mice, 55 sessions for SCm from 9 mice, 13 sessions from the MRN from 6 mice, 16 sessions from the RSPd from 4 mice.

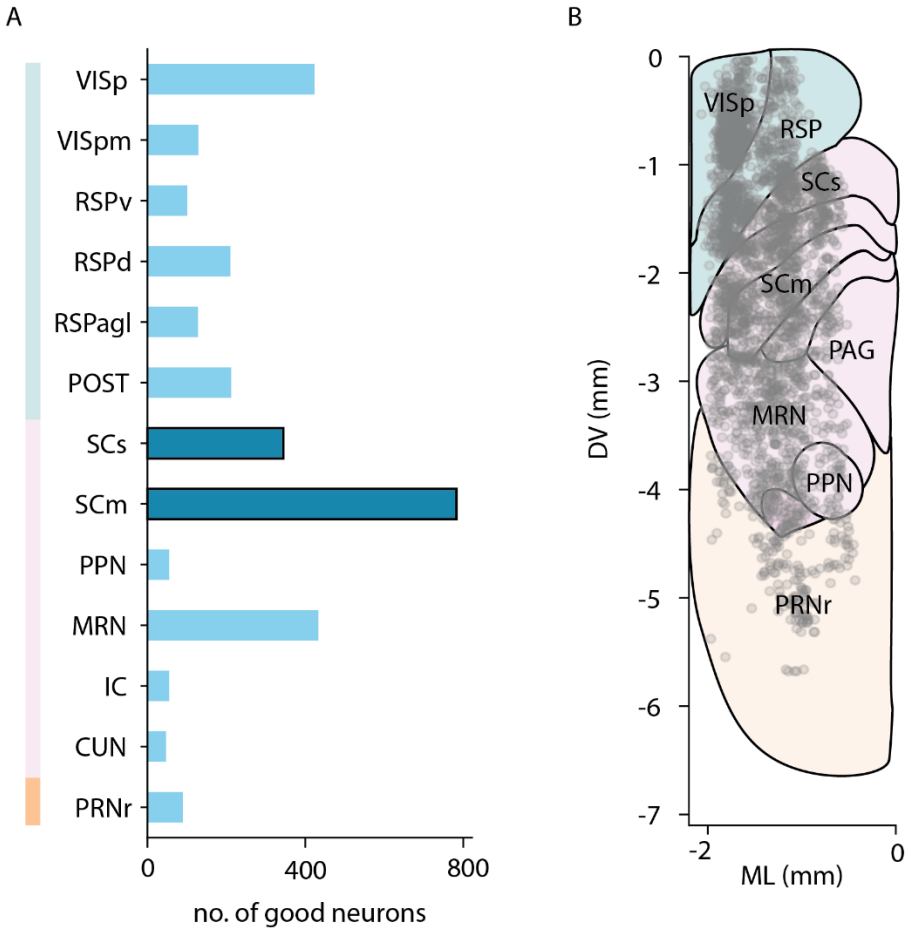



***FIGURE 22. NEURONS RECORDED IN TRAINED MICE.***

*Acronyms indicated are according to the Beryl annotation of the Allen Atlas Framework, in order of depth. Green indicates cortical areas; pink indicates midbrain areas and orange indicates hindbrain areas. VISp: primary visual cortex. VISpm: posteromedial visual areas. RSPv: ventral retrosplenial cortex. RSPd: dorsal retrosplenial cortex. RSPagl: anterior retrosplenial cortex, POST: postsubiculum, SCs: sensory layer of the superior colliculus (superficial), SCm: motor-related superior colliculus (intermediate & deep), PPN: Penduncolopontine nucleus, MRN: midbrain reticular nucleus, IC: Inferior colliculus, CUN: Cuneate nucleus, PRNr: Pontine reticular nucleus. (A) Number of good neurons that passed quality metric scores across regions. (B) Distribution of good neurons projected onto the coronal plane represented at -3.6mm in the anteroposterior position from Bregma. X-axis indicates mediolateral (ML) axis, while y-axis indicates the dorsoventral (DV) axis. Grey dots indicate the neurons. For representational purposes, I added a Gaussian noise (σ=80 μm) to the neurons' location along the ML axis.*

## Data analysis

I investigated whether relevant auditory or visual signals can be decoded from neural activity using the combined condition visual stimulus localisation (ccVP), auditory stimulus localisation (ccAP). Additionally, I investigated whether SC neural activity could predict choices choice using two complementary methods:

(1) **The combined choice probability (ccCP):** This method assesses how well stimulus and choice conditions can be decoded from SC neural activity, similar to ccVP and ccAP.

(2) **The neurometric Log-Odds model:** This method tests whether SC neural activity predicts choice on each session.

These methods have different assumptions: the Log-Odds model assumes a logistic relationship between neural activity and choice, requires fewer parameters and more data, and it does not need both leftward and rightward choices for all trial conditions. The ccCP method, on the other hand, is nonparametric.

### *The combined condition localisation and choice probabilities for each neuron*

To calculate a neuron's selectivity for choice, I examined how its firing rate differed between leftward and rightward choices under the same stimulus conditions, for example, on blank trials. This can be measured using the area under the receiver operating characteristic score (auROC score, Fawcett, 2006). The auROC score is closely related to the Mann-Whitney U-statistic (also called rank-sum, Mason and Graham, 2002), which is a non-parametric measure of differences between two groups, such as firing rates for leftward vs. rightward choices.

Since there often few blank trials, calculating the auROC score for these trials alone would lack statistical power. To address this, I also calculated auROC scores for other trial types and combined them. This combined-conditions testing method (Steinmetz et al. 2019), aggregates Mann-Whitney U-statistics across conditions, such that, at each stimulus condition, we calculate the U-statistic at each condition $i$, such that

$$U_i = R - \frac{n(n+1)}{2}$$

Equation *6*

Where $R$ denotes the sum of the ranks of the neural activity in one group, e.g. rightward choices, and $n$ the total number of trials within that group. Then, the U-statistics across conditions are combined into a single metric, the combined U, by

$$combined\ U\ = \frac{\sum U_i}{N}$$

*Equation 7*

Where N denotes the total number of trials. To estimate statistical significance of the ccCP measure, I used the shuffle test in which the trial labels (e.g. left versus right choice) were permuted within each condition 2000 times, and the combined condition U-statistic was recomputed for each shuffle.

In the context of the task, we can compute several combined conditions scores for each neuron: we can test not only the neuron's selectivity for choice (ccCP), but also their selectivity for stimuli, i.e. whether we can discriminate the auditory (ccAP) or visual position (ccVP) or the from the neural activity, by computing the auROC scores on left vs right auditory only trials when the animal made the same choices.

To obtain a number for the neural activity at each trial, I summed the total number of spikes in a 150 ms window after stimulus for ccAP and ccVP. I calculated ccCP in two 150 ms time windows: before the making of the choice and before stimulus onset.

### *The neurometric Log-Odds model*

I wanted to assess whether the SC neural activity contributes to mouse decisions in the audiovisual task *beyond* simply encoding the visual or auditory stimulus location. This would reveal correlational evidence that the SC is recruited to solve the task. Such decoding of choice can stem from various sources, as shown by the example neurons: it is possible that sensory signals are different depending on the choices the animal is going to make, or that the SC already carries a bias signal even before stimulus presentation, representing a form of habitual, or slow bias. The SC may also contain choice-related signals after stimulus onset but before the decision, indicating that either the SC partakes in more abstract decisions or motor execution. To analyse these 3 broad possibilities, I used an extension of the log-Odds model to assess whether neural activity can improve prediction on behaviour. For this, I essentially add a neural term to Equation 3:

$$\log\left(\frac{\mathrm{p(R)}}{p(L)}\right) = v_R V_R^{\gamma} - v_L V_L^{\gamma} + a_R A_R - a_L A_L + \mathrm{B} + \sum_i w_i r_i$$

*Equation 8*

Where $r_i$ is one number representing the activity of neuron $i$ on each trial and $w_i$ is the learned weight for each neuron. To get this number, I z-scored the firing rate of a neuron across trials in a 150 ms time bin before stimulus onset (pre-stimulus) or before choice onset. I also wanted to assess whether sensory stimulus signals are potentially already altered by the upcoming choice. For this, I took the average neural activity 150 ms after stimulus onset on trials where choice was not yet initiated in this 150 ms period (i.e. trials where the mouse was slower to respond). I will refer to this period as the post-stimulus time window.

I fitted the model using *scikit-learn* (Pedregosa et al. 2011). For this, I built a pipeline that divided predictors to neural and stimulus predictors. First, I fitted a model using only the stimulus predictors. Then I fixed the gamma parameter and fitted the neural model. For this, I first used a feature selector on the neural predictors. This consisted of a variance threshold (set to 0.001), and a sequential forward selector, which iteratively finds the best neural predictors until the cross-validated score did not improve any further. For the cross-validated score, I used the negative $\mathrm{Log_2}$Likelihood, and I set a minimum improvement threshold, normalised across trials. I determined this threshold using grid search, which was typically 0.015-0.035 bits averaged across trials. To assess whether the neurometric model improves the overall fit, and does not overfit, I refitted the model with all the predictor neurons identified and compared the cross-validated -$\mathrm{Log_2}$Likelihood score to the behaviour-only model.

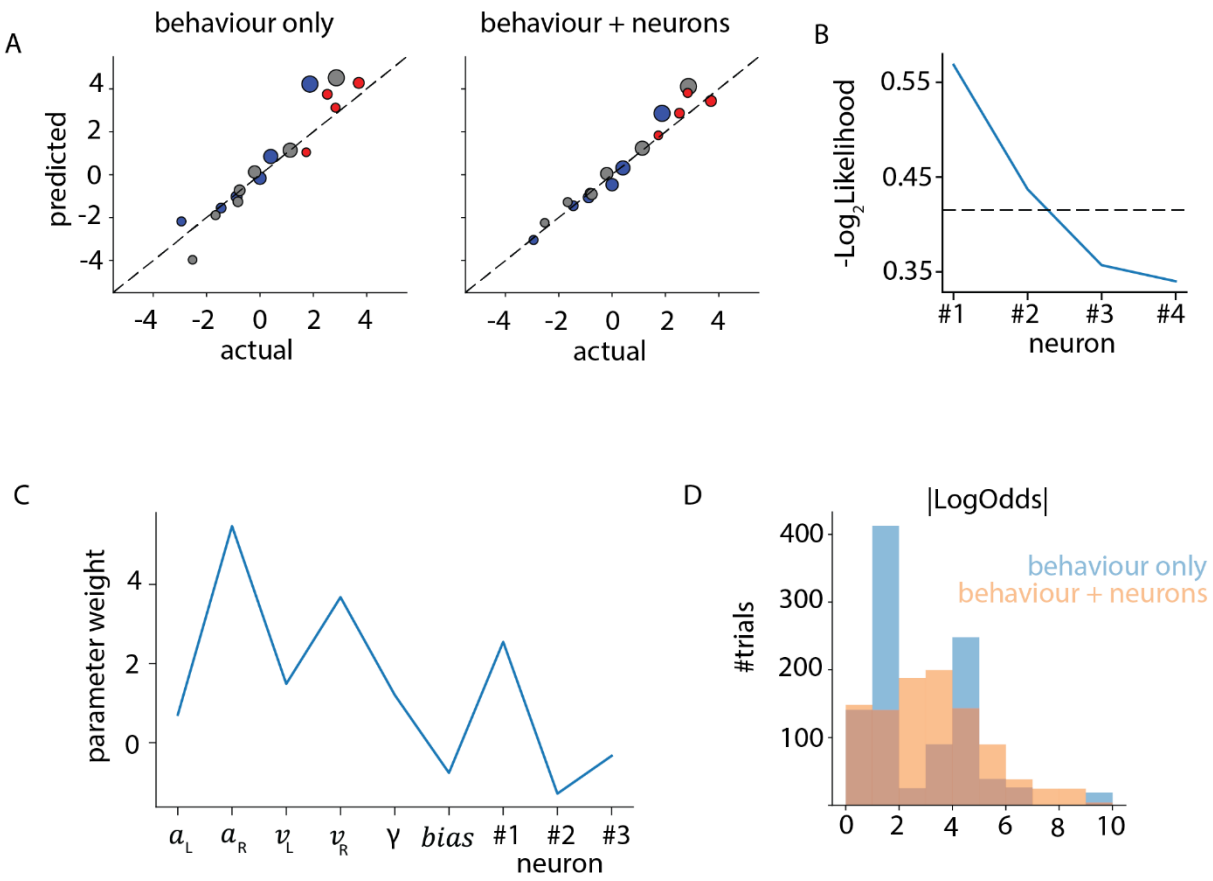



***FIGURE 23. FITTING A NEUROMETRIC LOG- ODDS MODEL.***

*(A) Actual vs predicted average log-odds of an example session without (left) and with the neural predictors (right). Colours denote different auditory azimuthal positions (blue: left, grey: middle, red: right). These are the same points as the data points on the log-odds plots. The size of the dot is scaled by sided contrast, such that the highest contrast stimulus on the left side is the smallest dot. (B) Decrease in -Log₂Likelihood as neurons are added to the model in the sequential forward selection procedure. (C) Learned weights for both the behavioural and the neural predictors. (D) Distribution of log-Odds across trials for the final neurometric model (yellow) vs the behavioural only model (grey). Log-Odds values may be interpreted as a measure of how confident the model is to assign a left vs right decision to each trial (Roy et al. 2018).*

## Results

### Example SC neurons responsive in the task

Among the units recorded during the task, I found SC neurons with varied response profiles. Some neurons appeared visual (Figure 24A); others auditory (Figure 24B). Some neurons responded to all stimuli equally (Figure 24C). Other neurons increased their firing rate as the choice movements were initiated (Figure 24D) or differentiated between leftward and rightward movements (Figure 24E-F), or appeared to contain a mix of auditory and choice-related signals (Figure 24G).

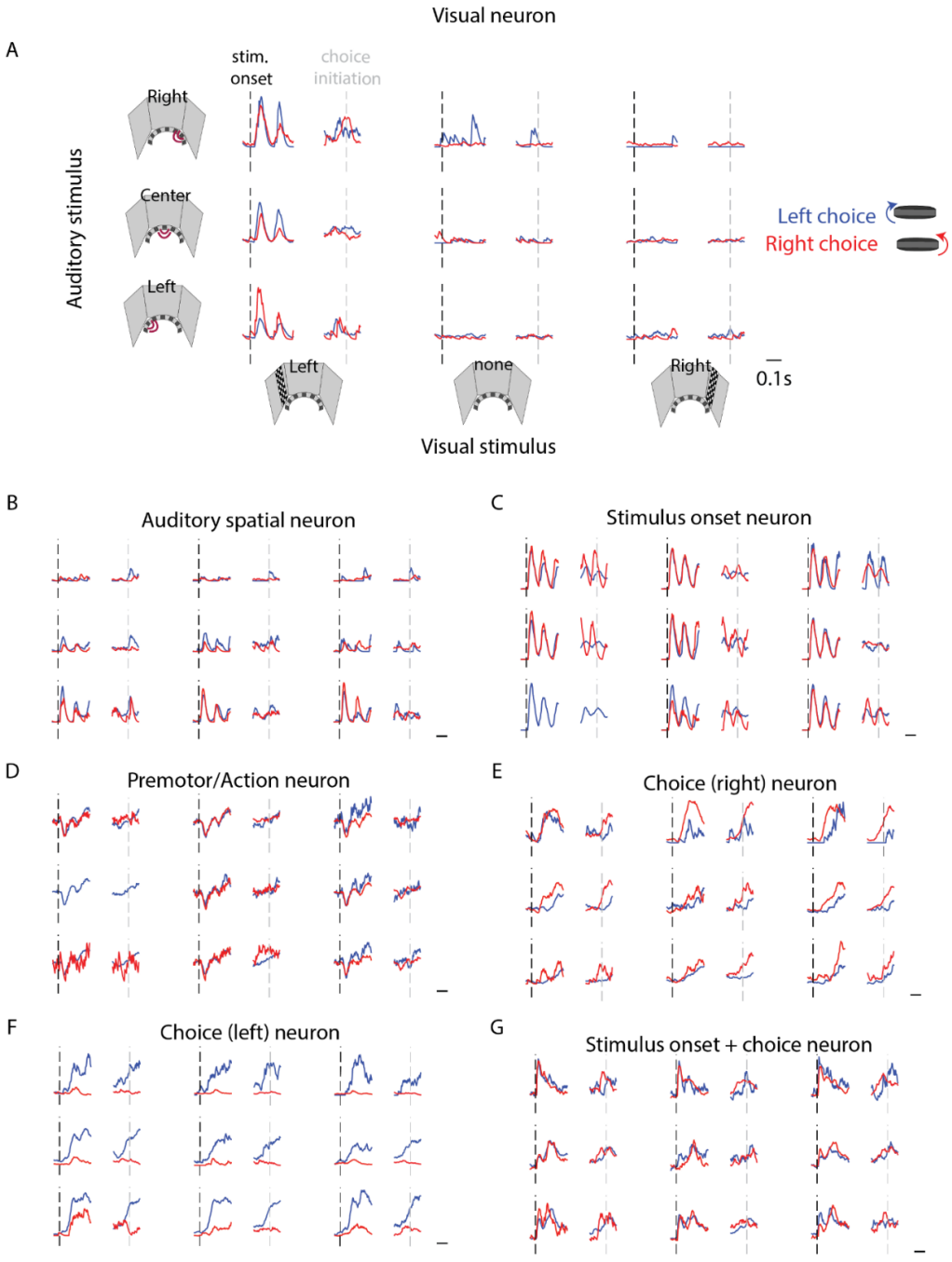



***FIGURE 24. EXAMPLE NEURONS FROM THE SC DURING AUDIOVISUAL TASK PERFORMANCE.***

*Each plot is arranged into 9 subplots, that denote each azimuthal condition, i.e. the 3 visual stimulus locations plotted against the 3 auditory azimuthal locations. Each subplot consists of two plots: the peri-stimulus time histogram (PSTH) aligned to auditory stimulus onset (left), and one aligned to the choice initiation preceding the decision (right). Neural activity on trials when mice make rightward choices are coloured red while leftward choices are coloured blue. At times, some trial types were missing because the mouse did not perform those trials (typically error trials). In these plots, I do not distinguish different conditions of visual contrasts. (A) Example visual neuron (B) Example auditory neuron (C) Example neuron that responds to stimulus onsets at all positions equally. (D) Example movement neuron (E-F) Example rightward and leftward choice neurons (H) Example auditory neuron that also appears modulated by choice.*

## Decoding stimulus location and choice

To assess whether SC neurons are selective to auditory or visual stimulus location, or choice, I applied the combined condition probability (ccP) statistic on the neural activity. As expected, neurons selective for visual stimulus location were most enriched in the primary visual cortex, VISp, followed by the superficial (SCs) and motor layers of superior colliculus (SCm). Auditory stimulus location selective neurons were most prevalent in motor-related midbrain areas (SCm, as well as the cuneiform nucleus (CUN), the midbrain reticular nucleus (MRN)) and in the inferior colliculus (IC). This confirms that visual and auditory localisation signal is prevalent amongst the recorded SC neurons, as expected. The same brain regions were also enriched in choice-selective neurons, especially prior to moving to indicate the choice (Figure 25).

Consistent with results from naïve mice, audiovisual neurons were rare. Only 6 neurons (<1 %) were selective for both auditory and visual stimulus location. Similar separation of the choice-related neural population could be observed: only 2 neurons were selective for auditory location and choice before movement, while 1 neuron was selective to visual stimulus position and choice.

To further test the prevalence of choice-related signals in SC, I applied the neurometric log-odds model to sessions with enough neurons from a single region. The neurometric model improved on predictions of choice when neurons were sampled from the SCm in the pre-choice period, suggesting that the SC participates in making the choice. The predictive power from SCs was not significant. MRN neural activity could predict mouse choice in this period, suggesting that strong choice signals do not singularly arise in the SCm. In the pre-stimulus or post-stimulus widow, neural activity was not significantly predictive of choice, suggesting that the strongest signals that are predictive of choice arise after the initial auditory or visual response (Figure 26). Importantly neurons can only improve predictions compared to stimulus predictors if the animals make mistakes and on blank trials. Indeed, improvements by the neurometric model were the largest on blank and conflict trials. Furthermore, the possible the neurometric model could provide larger improvements in sessions where the animals made fewer correct choices on average. In addition, the improvements also depended on the overall bias – indeed, if the mice were highly biased towards one direction, the LogOdds model effectively gets very little data to train on; as a consequence possible improvements are marginal on highly biased sessions.

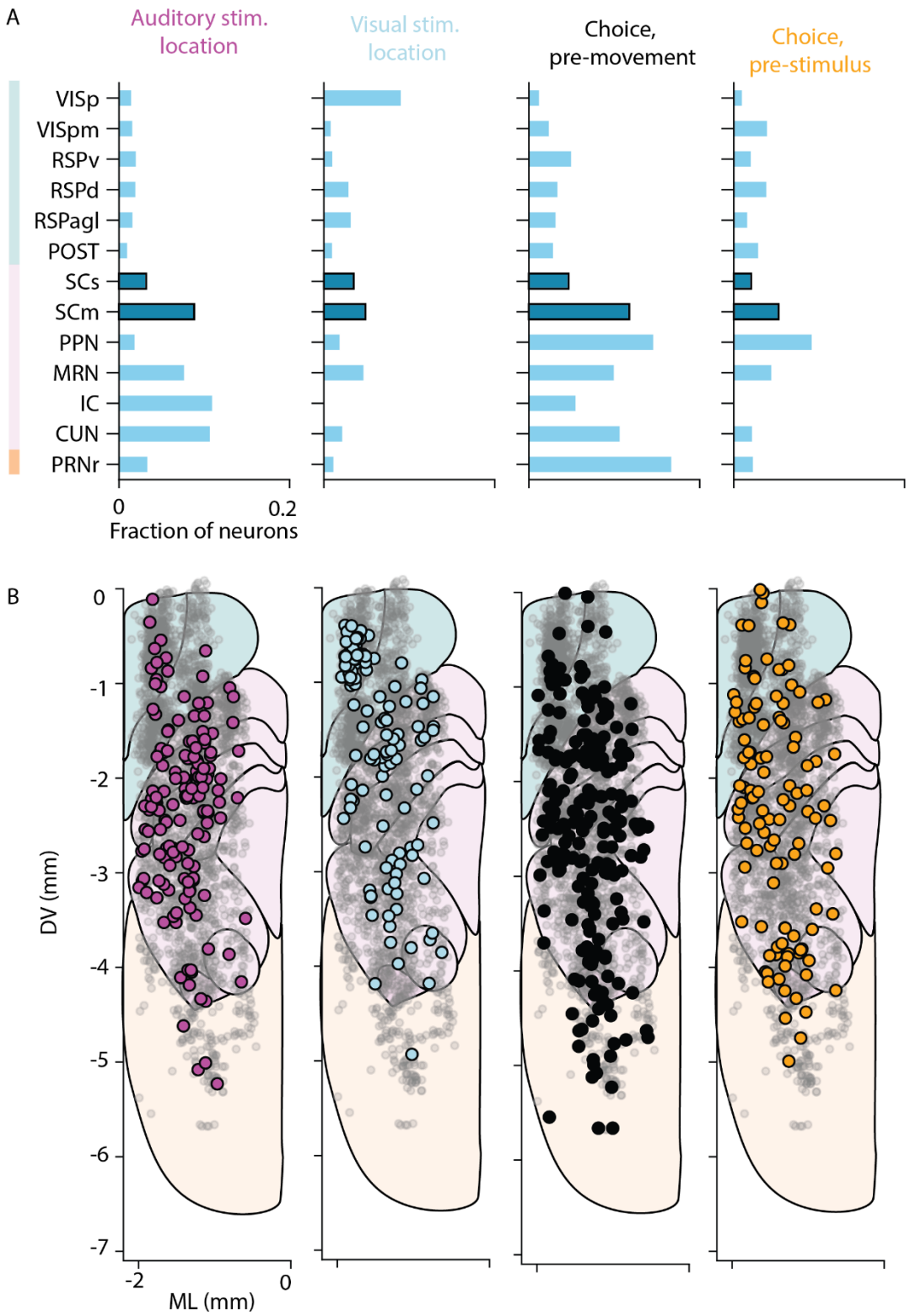



***FIGURE 25. NEURONS SELECTIVE FOR TASK VARIABLES.***

*(A) Left to right: Neurons with reliable encoding of auditory stimulus location, visual stimulus location, and the upcoming choice while mice perform an audiovisual task. Fraction of neurons with significant preference for the auditory stimulus location across brain regions as defined by the combined condition probability ($p<0.01$ for significance after 2000 shuffles). (B) Same as (A) but plotted in the Allen CCF projected to the coronal view, as in Figure 22. Grey dots: all good neurons. Coloured dots: neurons that were significantly selective for auditory stimulus location, visual stimulus location or choice in the pre-movement and pre-stimulus period.*

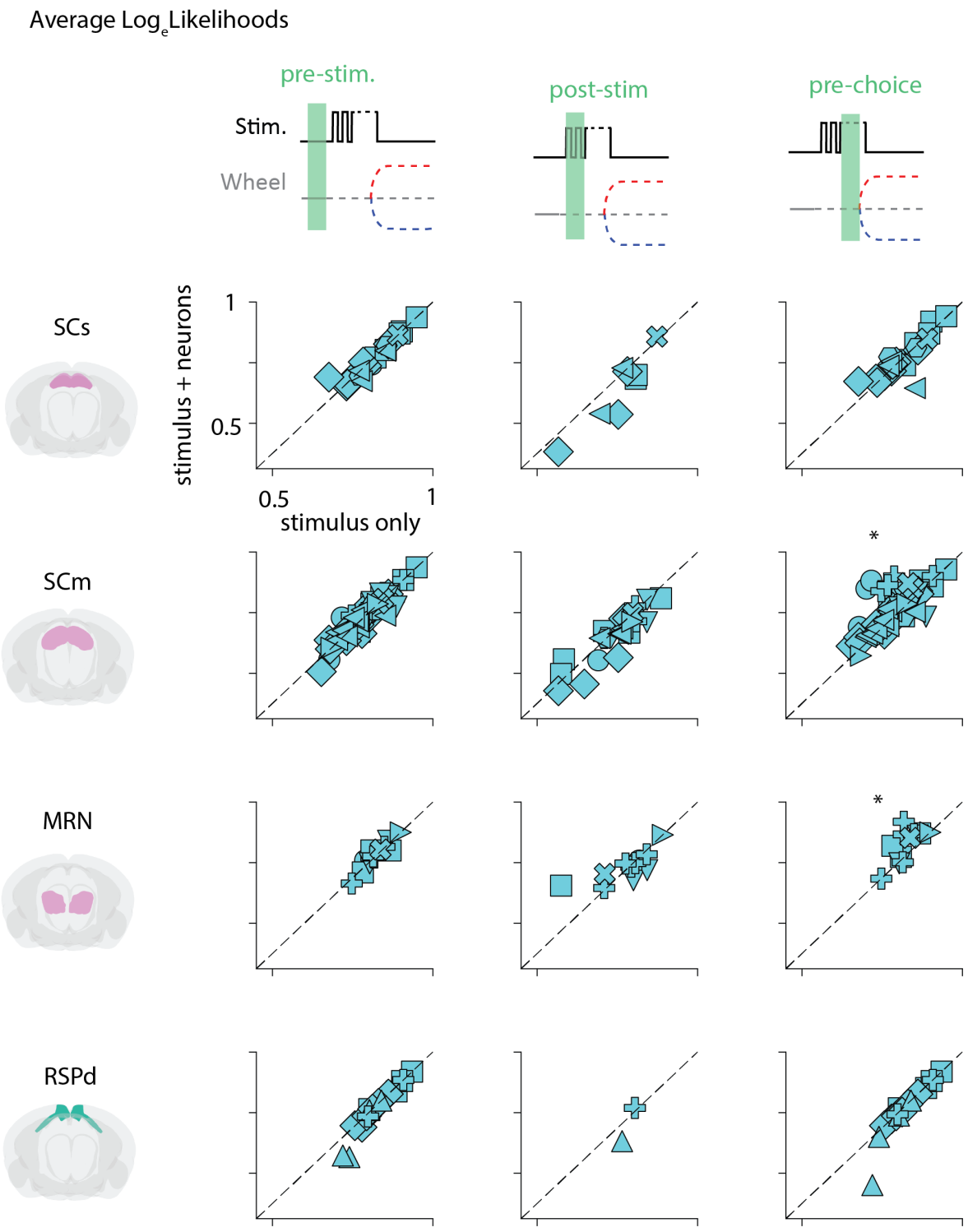



***FIGURE 26. SC NEURONS IMPROVE THE PREDICTIONS OF THE LOG-ODDS MODEL.***

*Average Log$_e$likelihood across trials for each session. Each plot shows the Log-Likelihood of the stimulus-only model, plotted against the log-Likelihood of the model that included neural predictors, for each session. Data is compared based on which brain region and what timeframe neural predictors are extracted taken from. Neural firing was extracted in three different timeframes represented by columns – prior to stimulus onset, after stimulus onset but before choice initiation, and prior to choice. Each row indicates a brain region: SCs, SCm, MRN, and RSP. Each shape indicates a different mouse. Asterisk (*) denote significant difference, using a linear mixed effects model (LogLikelihood ~ model type + (1|mouse)).*

## Discussion

Consistent with results in the naïve mice, during task performance, I recorded neurons that represent visual and auditory azimuthal location in the task. Audiovisual neurons remained sparse, suggesting that learning an audiovisual localisation task did not lead to the binding of auditory and visual localisation signals at a sensory level.

Both the combined condition choice probability (ccCP) and the neurometric log-Odds model suggested that the SCm is enriched in signals that can predict upcoming choice. This is consistent with recordings in similar 2-AFC tasks, where SC neural activity was shown to predict upcoming choices before decision onset in a variety of task settings, such as visual saccades, directional licking, wheel turning, or reorienting the body (Duan, Pagan, et al. 2021; Stine et al. 2023; Steinmetz et al. 2019; Felsen and Mainen 2008). Notably, several studies have observed choice signal enrichment in the lateral part of the SC (Thomas et al. 2023; Chen et al. 2024); here I found that choice neurons are distributed in the deep layers of the SC. In the above tasks, however, choices were executed by movements of the tongue, while in this task, choices are executed by movements of the forelimb. Choice signals may also be organised into motor maps depending on the body part choices with which choices are executed.

Analyses from both ccCP and the log-Odds model suggested that representations of sensory stimulus and upcoming choice do not appear to mix. Very few neurons were selective to both choice direction and stimulus location, and choices could not be decoded from SC neural activity in the stimulus sampling period. This suggests that the spatial sensory localisation signals are not encoding sensory evidence for localisation that eventually guides choices, but rather provide a reliable, but behavior-invariant representation of sensory space. In turn, the choice network also appears to operate independently to guide movements. This contrasts with the choice neurons in the SC observed in primates, which were modulated both by choice and sensory evidence, particularly early in the trial (Stine et al. 2023) . Further analysis could focus on choice-related neurons and examine their modulation by sensory variables such as stimulus location and visual stimulus contrast, to confirm the lack of modulation.

The ccCP also suggested that some neurons encode choice in the pre-stimulus period, which would indicate that the SC holds prior information that may influence choices beyond sensory representation. Such a prior may arise from assigning higher value to certain decisions, for example by preferring rightward over leftward choices. However, the Log-Odds model did not support SC signals containing such a prior. In similar tasks, where priors have been experimentally controlled by manipulating the reward probability, such that higher value is associated with certain choices, the SCm has been found to contain signals before stimulus onset related to prior information that can differentiate how much value is associated with rightward vs leftward choices (Findling et al. 2023; Lak et al. 2020). In our task, we did not experimentally manipulate the prior information, so it is possible that I did not see significant improvements because most choices were simply driven by stimuli rather than internal prior beliefs. Alternatively, the neurometric Log-Odds model is currently interpreted with stringent statistics, as we calculate improvement of the model on all trials; however, some trials (blank, conflict) are more likely to be driven by internal states than stimuli, and interpretation of the neurometric Log-Odds model may be more nuanced if I looked at specific trial types.

# Chapter 5 — Uninstructed movement signals across the SC

## Introduction

In recent years, it has been an increasingly appreciated that correlates of the uninstructed movements of the body, as well as various interoceptive signals (e.g. pupil size) influence the firing of neurons across the brain (Stringer et al., 2019). Uninstructed movements are defined as movements that have not been instructed to be performed by the animal due to task contingency. Such movements may be correlated with external events, both in the task but also during passive recordings. For example, sensory stimuli may cause animals to move even during passive viewing (Socha et al. 2018; Bimbard et al. 2023). Similarly, during task execution, choices might be preceded by body movements or small postural adjustments. For example, in tasks when decisions are indicated by licking a waterspout, small adjustments of the jaw precede movements of the tongue, and these are predictive of the upcoming choice (Mayrhofer et al. 2019). Body movements that are predictive of choice need not even be associated with the motor execution of the choice. Movement patterns on a slower timescale, for example, pupil size, might be predictive of general brain states, and thus, in turn, could be predictive of choices in decision-making tasks (Steinmetz et al. 2019).

These widespread correlates of movement signals cause a conceptual problem: neural responses that appear to be caused by a measured process (such as onset of sensory stimuli, choice initiation, reward consumption), might be in fact correlates of the body movements that were caused by these processes (Figure 27). This has historically resulted in misinterpretations of the nature of sensory and cognitive signals across the brain (Zagha et al. 2022). Thus, it is key to distinguish correlates of body movements from other processes, including auditory and visual responses, as well as choice-related neural activity.

.

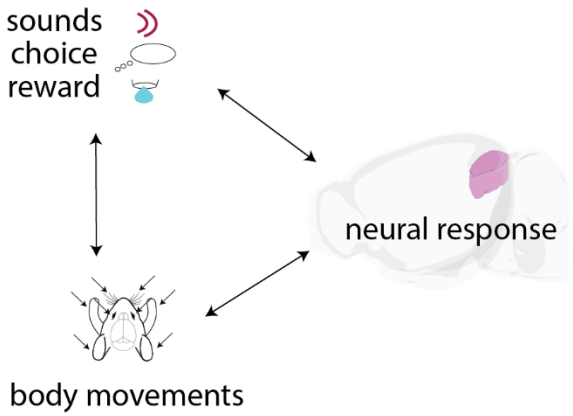



***FIGURE 27. NEURAL RESPONSES RELATED TO STIMULI AND TASK EVENTS MAY REPRESENT BODY MOVEMENTS.***
*A diagram that illustrates how body movement signals can be conflated with responses to external events, such as sensory stimuli, if not measured. Adapted from (the preprint version of Bimbard et al., 2023)*

Indeed, neural activity in the superior colliculus is expected to be modulated by uninstructed movements. Correlates of locomotion have been found in the medial layers of SC (Savier et al. 2019;

Chinta and Pluta 2024) but even the SCs's activity is thought to be modulated by arousal and eye movements (De Franceschi and Solomon 2018; Schröder et al. 2020).

In this chapter, I characterise the uninstructed body movements during passive viewing both in naïve and trained mice as well as during task performance and identify neural signals correlating with such uninstructed movements. Finally, I discuss whether correlates of body movement and other sensory processes can be disambiguated, given sufficient variability.

## Methods

### Recording movements

Mice were illuminated by infrared light and recorded at 30-60fps using high-speed cameras (30 fps in acute (DMK 23U618, The Imaging source), 60 fps in chronic recordings (CM3-U3-13Y3M-CS or BFS-U3-13Y3M-C, FLIR)). During naïve recordings, the body was typically recorded with a single camera, while in chronic recordings the body was monitored with three cameras viewing the mouse from different angles (Figure 28A). Due to technical difficulties, a large fraction of this camera data was not saved properly.

### Data analysis

#### *Quantifying uninstructed movements*

To describe the amount of movement that I observe, I used a previously well-described metric, motion energy (Figure 28B), which is defined as the sum pixel intensity change from frame to frame in a defined ROI. I focused this ROI on the whiskers and the face using Facemap (https://github.com/MouseLand/facemap (Stringer et al., 2019)). In a subset of videos, this ROI was manually determined, but in most videos, the eye position was automatically detected using DeepLabcut (Mathis et al. 2018), and the position of the whiskers and the face were inferred relative to the eye (Figure 28A). Videos were aligned to electrophysiological data using light flashes simultaneously acquired on the video and a DAQ. In recordings with several camera views, I picked only the side camera view for further analysis, unless it was missing due to the data-saving issue. If the side view was not available, I used the eye camera view, and if neither of these were available, I resorted to the front camera. Altogether I included 15 sessions in naïve recordings, 126 sessions in passive recordings from trained animals, and 55 sessions while mice performed the audiovisual task.

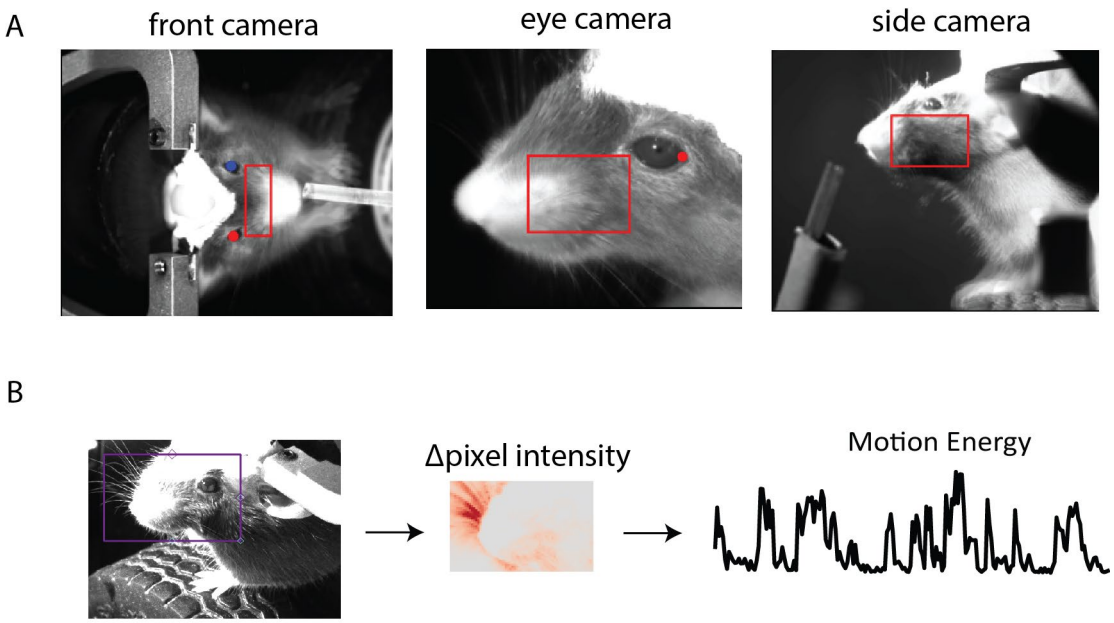



***FIGURE 28. MEASURING UNINSTRUCTED FACIAL MOVEMENTS.***

*(A) camera views with typical ROIs determined by DeepLabCut based on detecting key landmarks, such as they eye. (B) Process of calculating motion energy. To quantify the amount of movement, we sum the pixel intensity changes across frames in the video.*

### *Statistical test to relate uninstructed facial movement amplitude to stimulus and choice.*

To test whether movement amplitude is dependent on the location of the stimuli, I calculated the difference between the average movement amplitude on the left versus right using a permutation test in recordings with passive viewings.

In recordings during the performance of the task I applied combined conditions testing, as described in Chapter 4 (Steinmetz et al. 2019). For this, I calculated the ccAP, ccVP in a 150 ms window post auditory and visual stimulus onset, and the ccCP score before a 150 ms before choice movement onset for the uninstructed movement amplitude.

### *Relating neural activity to uninstructed movements*

I recorded videographic data both alongside the naïve recordings described in Chapter 3, as well as during the audiovisual task (Chapter 4). In this chapter, I related the neural activity in those recordings to uninstructed facial movements. In this chapter, I analysed 1811 neurons in naïve recordings out of which 1471 neurons were in the SC. In trained mice during passive viewing, I analysed 3561 good neurons from which 1280 were in the SC. In recordings during behaviour, I analysed 2077 neurons, including 710 neurons in the SC (this number is lower than the number of neurons recorded in trained mice during passive viewing because the inclusion criterion for active recordings also depended on task performance, as described in the Chapter 4).

To relate neural firing rate to uninstructed facial movements, I correlated motion energy with neural activity in the inter-trial intervals in passive recordings and during task performance. For this, I calculated firing rates for each neuron across the entire session in 100 ms bins, and interpolated the

motion energy values at each of the corresponding time points using the *scipy.interpolate.interp1d* function (Virtanen et al. 2020). Then, I removed time points near task events (100 ms before stimulus onset to 700 ms after stimulus offset or reward delivery) and calculated the linear correlation coefficient. To determine which neurons significantly correlated with movement, I used a session permutation method (Harris 2020). For this, I determined the correlation with motion energy acquired on other recording sessions and considered a neuron to be significantly correlating with movement if its correlation exceeded the correlation values obtained during session shuffles ($p<0.001$).

## Results

### Uninstructed movements during audiovisual recordings

Auditory stimuli during passive recordings sessions reliably evoked facial movements, but visual stimuli did not result in similar movements (Figure 29A). The auditory stimulus-evoked responses were reliable in both naïve mice and in mice that were trained in the audiovisual task (Figure 29A-B), suggesting that habitual exposure to auditory stimuli does not result in the extinction of uninstructed movement responses. Movement responses followed auditory stimulus onset within 10 ms (Figure 29C). Furthermore, in 33 % of naïve recordings and 38 % of recordings from trained mice, there was a significant difference between the magnitude of uninstructed movements depending on auditory stimulus location – the exact auditory stimulus location that caused the most movements varied across mice, suggesting that this effect is not simply due to a poorly calibrated speaker in my setup.

Uninstructed facial movements also accompanied the execution of the audiovisual task (Figure 30). In 24 % of the recorded sessions, the upcoming choice could be predicted from facial movements before the initiation choice as measured by the wheel movement. In 12 % of recordings, the auditory stimulus location was also predictable from facial movements, while the visual stimulus location appeared to be predictable only in 3 % of the sessions.

### Movement correlates across the SC

Uninstructed facial movements were correlated with neural activity (Figure 31A). Neurons with significant modulation by uninstructed movements were widespread across regions, including in the superior colliculus, both in naïve and in trained mice (Figure 31B-C). Trained mice had a higher number of neurons correlating with movement compared to naïve, probably because trained mice are typically more awake during recordings even in passive viewing conditions. Neurons in the SCm were more correlated with movement than neurons in the SCs on average in both naïve and trained mice (Figure 31D) (Kolgomorov-Smirnov (KS) test, $p<0.001$).

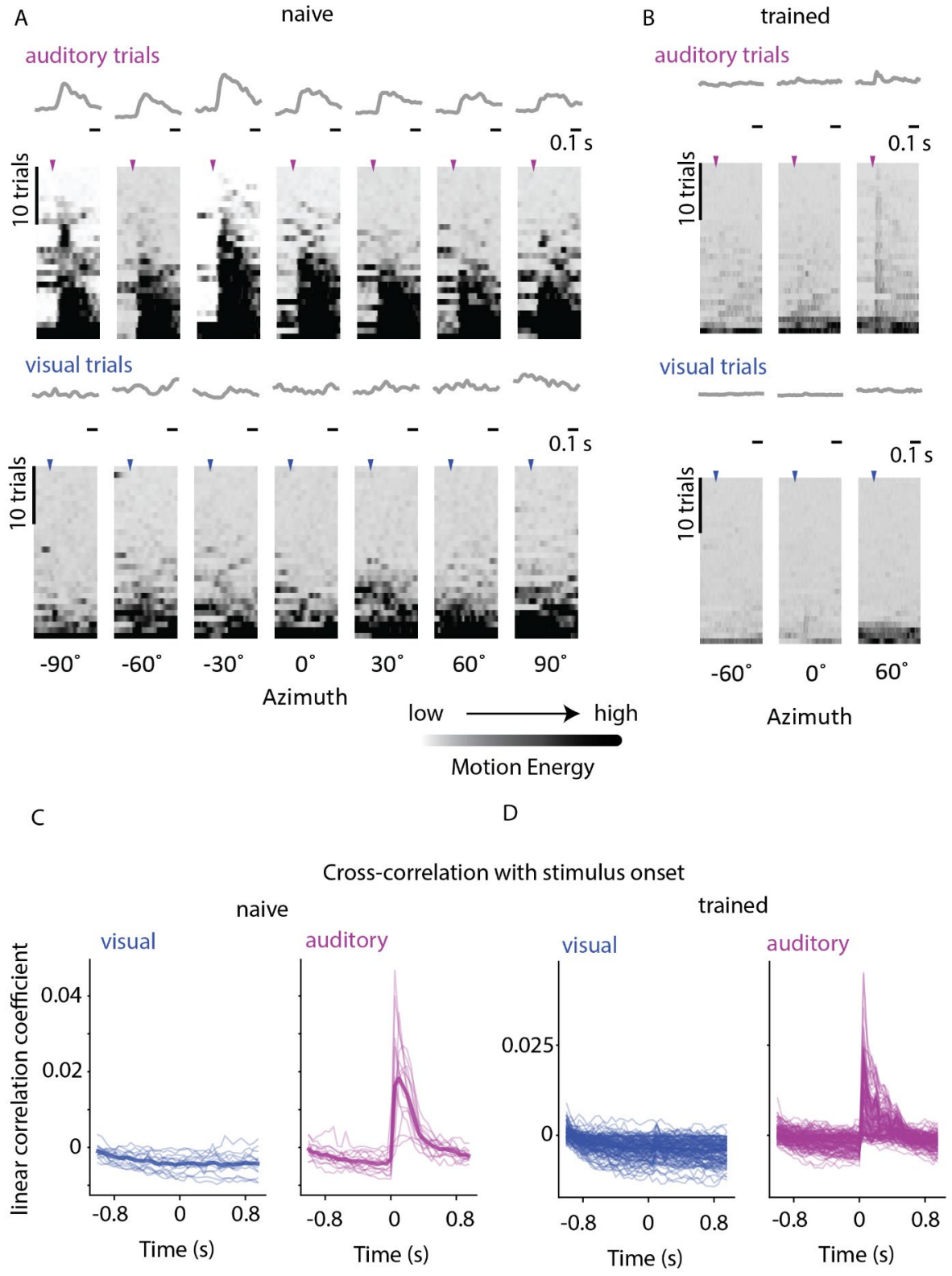



***FIGURE 29. AUDITORY STIMULI EVOKE UNINSTRUCTED FACIAL MOVEMENTS.***

*(A) Example triggered movement in a naïve mouse. Bottom: raster of movements sorted by the total amplitude of the movement on each trial. Top: average across trials. (B) Example triggered movements during passive recordings in a mouse trained on the audiovisual task. (C) Cross-correlation of motion energy and visual (left) and auditory stimulus onset (right) in naïve mice (n= 15 sessions). (D) Same as C for trained mice (n=126 sessions).*

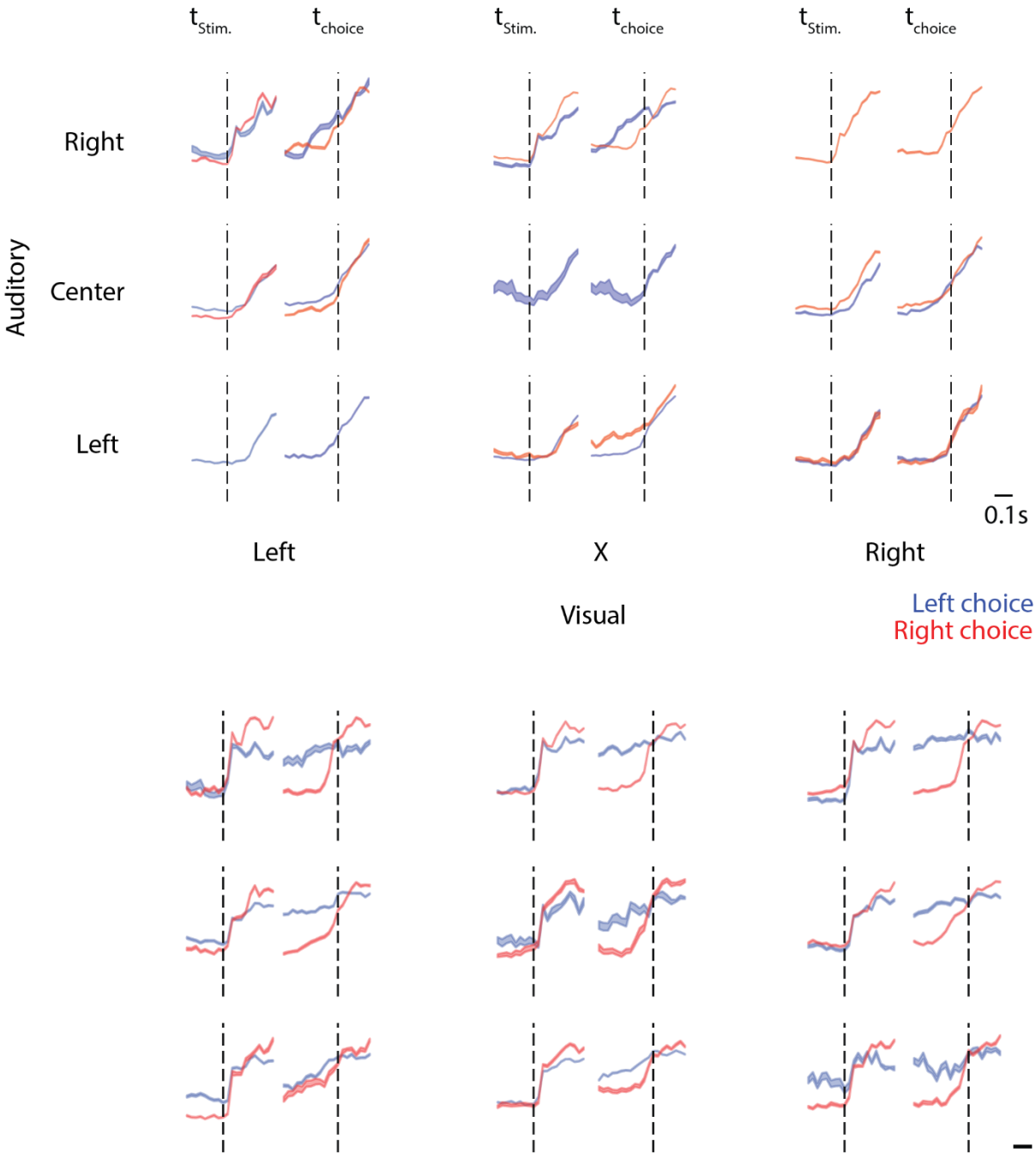



***FIGURE 30. UNINSTRUCTED MOVEMENTS MAY PREDICT STIMULUS LOCATION AND CHOICE.***

*Average facial movement amplitude at each stimulus & choice condition during the execution of the audiovisual task. Plots are arranged as in Figure 24, but instead of neural activity, motion energy is plotted here for the 9 combinations of auditory and visual stimulus positions. At each condition, motion energy traces are aligned to stimulus onset ($t_{stim.}$), and to the choice movement initiation of the wheel preceding the decision ($t_{choice}$). Motion energy on trials when mice make rightward choices are coloured red while leftward choices are coloured blue. Top: Example session where motion energy is selective of auditory stimulus location. Bottom: example session with motion energy selective of the upcoming choice.*

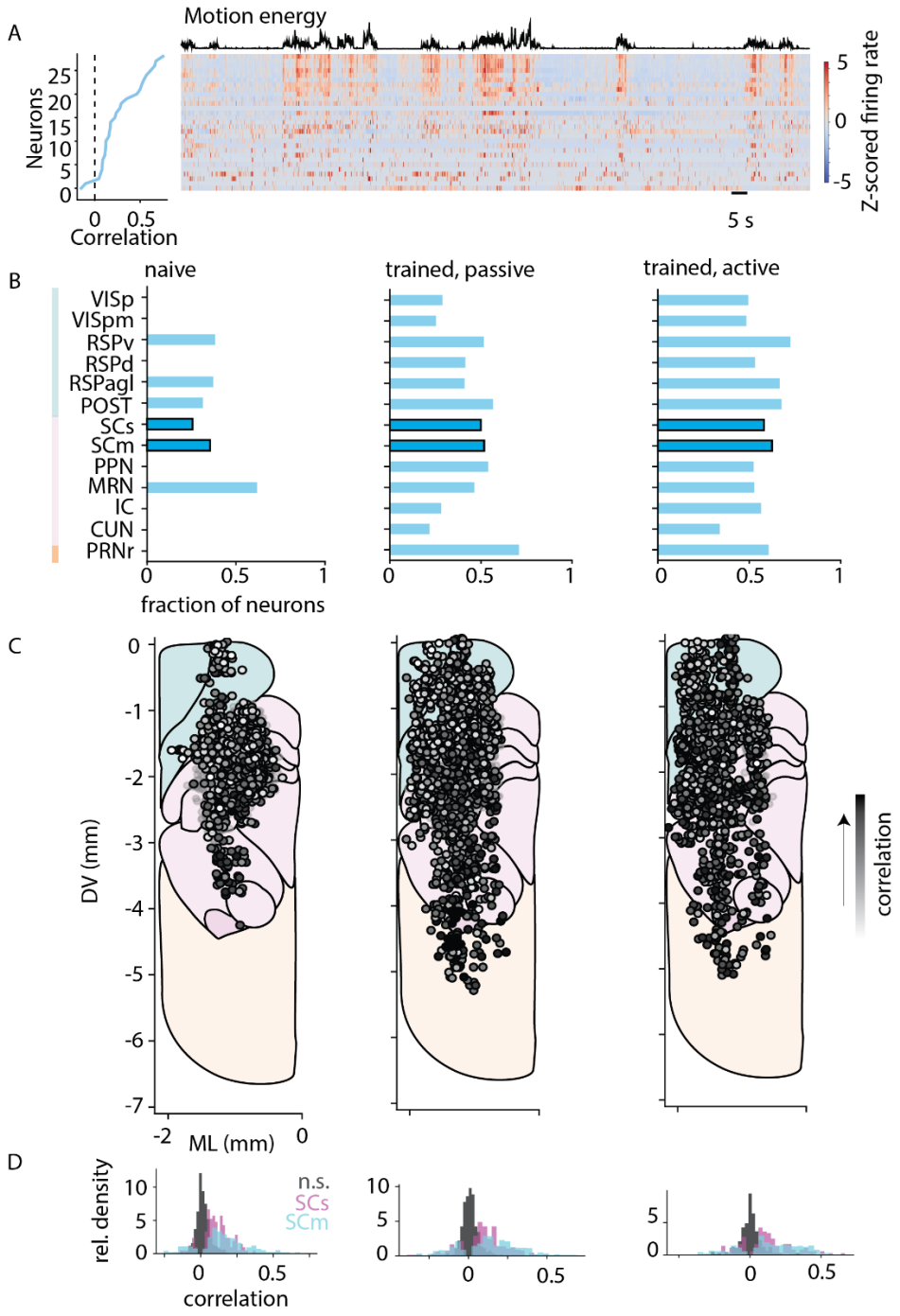



***FIGURE 31. CORRELATES OF FACIAL MOVEMENTS ARE WIDESPREAD.***

*(A) Example motion energy during a passive recording (top) and neurons that significantly correlate with motion energy during the recording (bottom) with their corresponding correlation values (left). Significance was calculated using a permutation test ($p<0.0001$). (B) Fraction of neurons that significantly correlate with facial motion energy across regions in naïve mice and trained mice during passive viewing (middle) and task execution (right). Missing bars indicate that the region was not recorded in this dataset. (C) Relative density of correlation values of non-significant neurons and significant neurons that were recorded in the SCs and the SCm across the different training conditions, as organised in (B). SCm and SCs distributions are significantly different in all cases (Kolmogorov-Smirnov (KS) test, $p<0.001$) (D) Anatomical location of correlated neurons. Neurons' location was projected to a coronal plane. Darker neurons indicate higher correlation values with facial motion energy.*

## Discussion

Here, I found that stimuli and auditory stimuli in particular evoked uninstructed movements of the face. In some experiments, stimulus-evoked movement responses are slower than stimulus-evoked neural responses (Williams et al. 2023; Oude Lohuis et al. 2024), making it possible to separate auditory stimulus-evoked neural responses from body movement-related neural activity. However, in my recordings, I found that body movements may occur within 10 ms after stimulus onset, which is not uncommon in other auditory protocols (Clayton et al. 2024); thus, in these recordings, it is difficult to separate movement-related neural activity from stimulus-evoked activity on a temporal basis.

An alternative to determine whether neural response is stimulus-evoked or related to body movements is testing whether neural activity represents purely stimulus onsets, or may represent further stimulus features, such as spatial tuning (Coen et al. 2023). However, I found that in many sessions, auditory stimulus location could also be predicted by uninstructed movements, suggesting that spatial tuning of neural responses does not eliminate the possibility that neural activity is related to body movements.

As with other decision-making paradigms (Gallivan et al. 2018; Steinmetz et al. 2019; Wang et al. 2023), body movements often predicted upcoming choices in our audiovisual localisation task. This predictive power from videos during the task might mean several cognitive or reflexive processes: it might simply signify that decisions are made earlier than the wheel is turned, and mice make body adjustments before they initiate a wheel turn. Alternatively, the mice are biased in making certain movements over others e.g. because they use a single paw on the wheel with which a leftward turn might be easier to execute that a rightward turn, requiring differential motion patterns when making the decision.

Movement correlates are widespread across all regions, as expected (Stringer et al. 2019). However, movement signals were not equally strong across areas, and movement-related signals were more enriched in the dorsal areas compared to the ventral areas while not following the borders defined by the anatomical landmarks. Perhaps neurons are more correlated with movements downstream in neural assemblies dedicated for motor execution (Chen et al. 2024).

Notably, here I used total facial motion energy as a measure of uninstructed movement. This measure is rather crude, and the interpretation of movement correlates could be more nuanced by taking a higher dimensional measure of body movement. On one hand, body movement may correlate signal a general state of arousal, which has a low dimensional representation across many brain regions, including the superior colliculus (Schröder et al. 2020). On the other hand, SC movement neurons may contain high-dimensional signals about the details of body movements.

Indeed, analysing body movements in more detail would potentially reveal a more nuanced representation both within the SC and across regions (Wang et al. 2023). For instance, the movement of different body parts may be differentially represented across brain regions; for example, the lateral SC is thought to be particularly enriched in signals related to the movements of the tongue (Thomas et al. 2023; Ito et al. 2024). Therefore, breaking down the one-dimensional motion energy signal to e.g. the movement of separate body parts, might show a less widespread representation of each body part across the SC. It might also be advantageous to consider a higher dimensional measure of body

movements by finding structure in the sequences of movements the animals perform; indeed, a set of movement patterns that are not directly related to specific events can reveal behavioural patterns that are predictive of task performance on a slower timescale (Yin et al. 2023). Finally, I also restricted my measures to movements of the face; although facial movements are thought to explain a particularly high fraction of variance in spontaneous activity, other body parts such as the pinnae and the eyes might contribute substantially to movement-related neural signals spread across the SC and may also influence responses in an audiovisual setting.

It is also possible that facial movements that I analysed here, are better interpreted a sensory rather than motor related processes, i.e. that the movements of the whiskers are associated with active somatosensory sampling of the whiskers rather than its motor components. Indeed, the superior colliculus can also represent somatosensory signals, that may integrate with other sensory or motor components (Chinta and Pluta 2024; Gharaei et al. 2018).

It remains possible that neurons that were historically measured to be audiovisual might be visual neurons modulated by auditory stimulus-evoked movements or changes in brain states. Indeed, in naïve mice, most auditory neurons I characterised had no spatial tuning but responded universally to all stimuli; these neurons activity may be related to body movements that each auditory stimulus may cause. Although some of the audiovisual neurons have been characterised under anaesthesia, the use of anaesthesia does not rule out that these stimuli trigger changes in brain states (Kramis et al. 1975; Klausberger et al. 2003).

# Chapter 6 — Combining vision, audition, movement and choice in SC neurons

## Introduction

In previous chapters, I described four sources of signals that can explain neural activity in the superior colliculus: visual stimuli, auditory stimuli, choices during the audiovisual task, and uninstructed facial movements. In Chapter 5, I highlighted the problem of partial correlations, i.e. that some of that auditory and visual stimuli are partially correlated with uninstructed movements and choices.

In this chapter, I attempt to understand how all these signals shape neural activity together using a linear regression method to predict neural activity from behavioural variables on each trial, i.e. an encoding model. The advantage of this method is that it accounts for variability across trials: mice will make uninstructed movements on some but not all trials (Figure 32A), and thus activity related to movements can be subtracted, i.e. regressed out (Musall et al. 2019). This way, the stimulus, body movement, and choice-related activity can be isolated. Here, I analyse both the passive data in naïve and trained mice and the active data during task execution and examine how these four signals: audition, vision, uninstructed movement, and choice to each other in single neurons across the SC. Notably, I analyse neural activity on unisensory and multisensory trials together, to investigate modulation by multisensory binding (cf. Figure 6).

## Methods

### Experiments

This chapter contains the same data as Chapters 3-5, except this data required the simultaneous quantification of body movements, while mice were recorded passively or while performing the audiovisual localisation task. Altogether, I selected 1673 neurons from 9 mice from naïve recordings, 3991 neurons from 10 mice trained in the task but recorded in passive conditions, and 2800 neurons from 10 mice recorded during the task.

### Kernel regression

During passive viewing, I used linear regression to predict neural activity from visual and auditory stimulus properties and uninstructed movements of the face. For this, I predicted the firing rate for a given neuron $r$ at each time point $t$ over the entire recording session, using a set of visual predictors, and set of auditory predictors, and motion energy as a single predictor.

Both the visual and auditory predictors were constructed as temporal predictors around the stimulus onset. I created unity matrices around the time bin for the stimulus onset to form the predictor matrix, and these temporal predictors are often referred to as design matrices. The weights learn on these predictors are often referred to as kernels. In mathematical terms,

$$r(t) = \sum_{v} V_{\varphi(v),p(v)}(t - t_v) + \sum_{a} A_{\varphi(a),p(a)}(t - t_a) + \alpha M(t) + R$$

*Equation 9*

Where

- $v$ denotes the visual variables, including a temporal kernel supported around visual stimulus onset $t_v$, visual stimulus location $\varphi(v)$ and contrast $p(v)$,
- $a$ denotes auditory variables, including a temporal kernel supported around auditory stimulus onset $t_a$, location $\varphi(a)$ and SPL $p(a)$.
- $M(t)$ represents motion energy at time t,
- $V$ denotes fitted visual kernels,
- $A$ denotes fitted auditory kernels,
- $\alpha$ denotes fitted weight for motion energy,
- $R$ is a fitted weight multiplied by 1 to learn the baseline firing rate (Figure 32B).

In naïve recordings, I presented visual and auditory stimuli at 7 azimuths, and at one or two contrast and SPL levels. In these recordings, each stimulus lasted 300 ms, therefore the stimulus kernels are also supported from 50 ms before to 300 ms after stimulus onset.

In passive recordings in trained mice, I typically presented the stimuli at 3 azimuths and varied the contrast such that the same contrasts were played in passive as in trained mice. Furthermore, the stimuli lasted longer. Therefore, stimulus kernels were supported from 50ms before to 600ms after stimulus onset in these sessions.

During the execution of the audiovisual task, there are more trial types to fit (as each stimulus condition can be associated with a different choice), and the trial numbers are smaller compared to passive recordings. Therefore, minimizing the number of parameters to avoid overfitting is necessary. Therefore, to fit the linear encoder to the active data, I had to introduce three reparameterizations.

First, I added a kernel related to choices, that contained a choice movement onset kernel, and a directional kernel that adds negative values to leftward choices and positive values to rightward choices. To construct the design matrix for the choice direction kernel, instead of a unity matrix around movement onset, I created a Toeplitz matrix consisting of -1 or 1 depending on the direction of the choice.

Second, I modified the visual kernel, such that it was only predictive of contralateral visual stimuli (as opposed to all azimuths). The visual kernel was also scaled by contrast using a fixed power of 0.5, as opposed to having a separate kernel for each visual contrast.

Third, I also reparametrized the auditory kernel. During the audiovisual task, the SPL of the auditory stimulus is not varied, and there is an auditory stimulus onset in every trial (as auditory stimuli are presented in the centre during visually informative trials (Figure 9B). Since most auditory neurons in the SC are expected to respond to auditory stimuli independent of stimulus location, a substantial component of neural activity will be related to auditory onsets that are not useful for solving the task, since they cannot identify auditory stimulus location. Therefore, I reparametrized the auditory kernels with an onset and a directional kernel, like the choice kernel. The onset kernel represented a global auditory stimulus onset kernel activated at each trial. The auditory direction kernel was weighted negatively on left auditory trials and positively on right auditory trials, and thus it could be used to

determine whether a neuron is selective for auditory stimulus location. Altogether, the kernel model during the audiovisual task was parameterized as follows:

$$r(t) = \sum_{v} V(t - t_v) + \sum_{a} A(t - t_a) + \sum_{a} A_{dir}(t - t_a) +$$
$$+ \sum_{c} C(t - t_c) + \sum_{c} C_{dir}(t - t_c) + \alpha M(t) + R$$

*Equation 10*

Where,

- $v$ denotes the visual variable that are the time points creating the temporal kernel around visual stimulus onset $t_v$
- $a$ denotes auditory variable that are the time points creating the temporal kernel around auditory stimulus onset $t_a$
- $c$ denotes the choice variable that are the time points creating the temporal kernel around choice stimulus onset $t_c$
- $M(t)$ represents the momentary motion energy,
- $V$ denotes visual weights,
- $A$ denotes auditory stimulus onset weights,
- $A_{dir}$ denotes the directional kernel weights distinguishing left versus right auditory stimulus location.
- $C$ denotes the weights related to initiating a choice-related movement,
- $C_{dir}$ is the direction kernel weights distinguishing left versus right choices,
- $\alpha$ is the weighing of the motion energy on neural activity.
- $R$ represents a constant baseline weight (Figure 32C).

In this model stimulus kernels were supported for maximally 0-500 ms around stimulus onset, but if the choice was made earlier stimulus kernels cease to be supported, as the stimulus presentation ends once the choice is made.

All models were fitted using Ridge regression with the regularization strength of α=1 using *scikit-learn*. To assess individual neuron selectivity, I evaluated the contribution of visual, auditory, choice, and movement kernels, by calculating the explained variance score on the residuals using cross-validation. For this, I calculated the prediction of all kernels except the kernel in question (termed leave-one-out kernel, for example, all auditory predictors), and I subtracted this prediction from the actual firing rates $r(t)$ to obtain the residual, which represents all aspects of the neural activity that are not explained by other predictors. Then, I retrained the model with the leave-one-out kernel and calculated the explained variance score (*scikit-learn*) predicted by this model on the test set. Neurons with >2 % explained variance score were deemed selective for the kernel in question, which is an arbitrary threshold often used in similar analysis (Steinmetz et al. 2019). Using different thresholds did not seem to change the conclusions. Trials were split to train and test set using the *StratifiedShuffleSplit* method as described previously, to ensure balanced representation of trial types across folds.

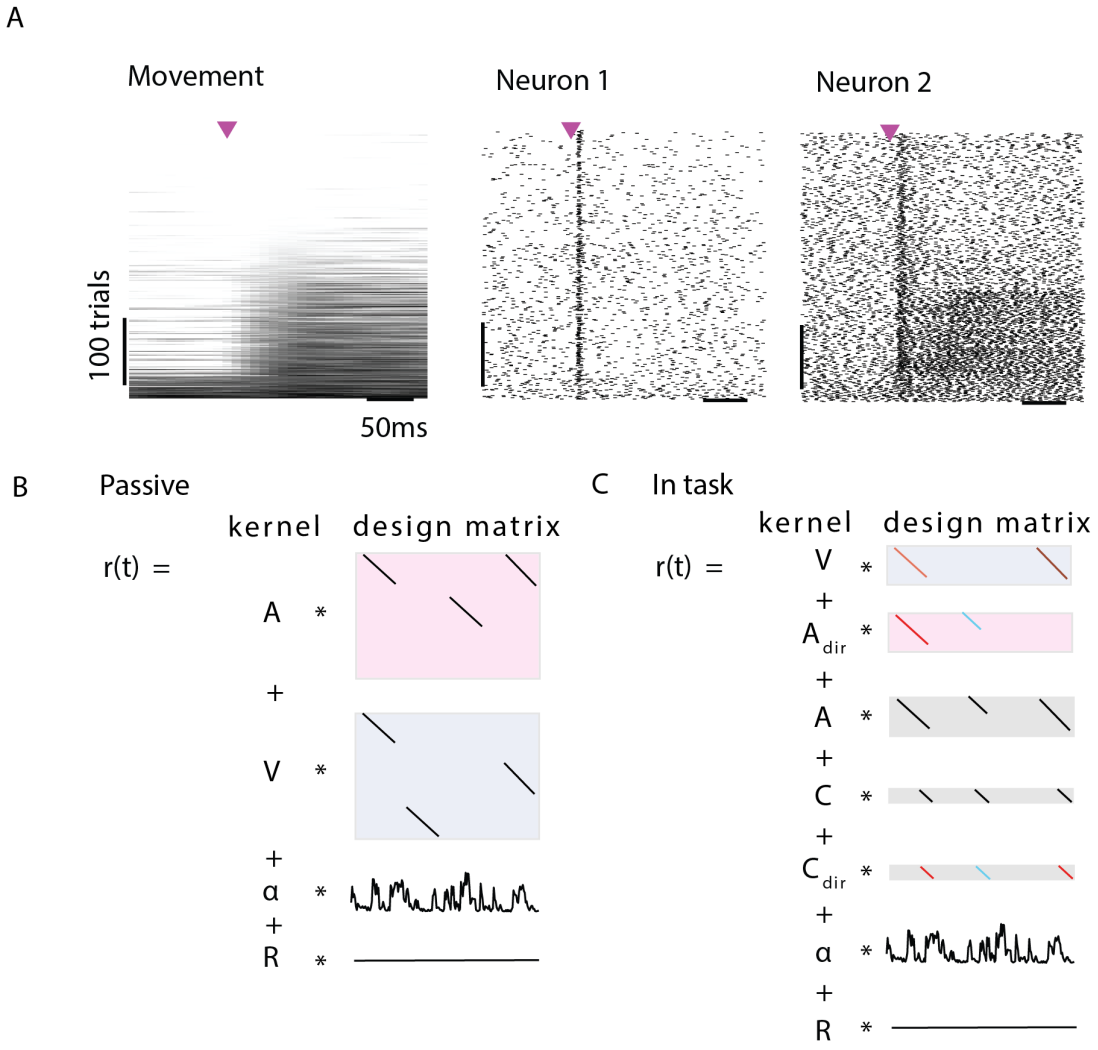



***FIGURE 32. KERNEL REGRESSION METHOD.***

*(A) Illustration of how the trial-by-trial models, such as kernel regression can pull apart auditory stimulus-related responses from responses related to auditory stimulus-evoked body movements. Left: motion energy on each auditory trial in an example passive recording, aligned to auditory stimulus onset and sorted by the amplitude of the total motion energy in the 200 ms window after stimulus onset. Right: two example neurons from the same recording, also aligned to auditory stimulus onset and sorted by the same order as the motion energy on the right. Neuron 1 thus responds to auditory stimuli independent of movements, while neuron 2 also responds to auditory stimuli, but is also modulated by movements of the body. (B) Visual representation of the linear encoder model described to fit passive responses. Boxes denote the design matrices, which are multiplied by a set of weights, referred to as kernels, as described in Equation 9 to predict the neural response vector r(t). In the design matrix time is on the x axis and the lines in the boxes indicate times around which the temporal kernels were supported, i.e. the design matrix contained a 1. Other elements of the design matrix are 0. (C) Kernel model to fit the neural responses during the audiovisual task (Equation 10). The symbols for the design matrix are the same as in (B). Coloured lines in the design matrix indicate directional kernels that allowed for the same weight to be subtracted from neural activity on leftward stimuli/choices and added to neural activity on rightward stimuli/choices, essentially allowing to fit a differential neural response depending on stimulus or choice location. For this, the design matrix was filled with -1 on left trials and +1 on right trials. Red shades for V indicate the contrast scaling parameters with which the supported timepoints on the visual kernel were scaled according to the visual stimulus contrast on a given trial.*

## Results

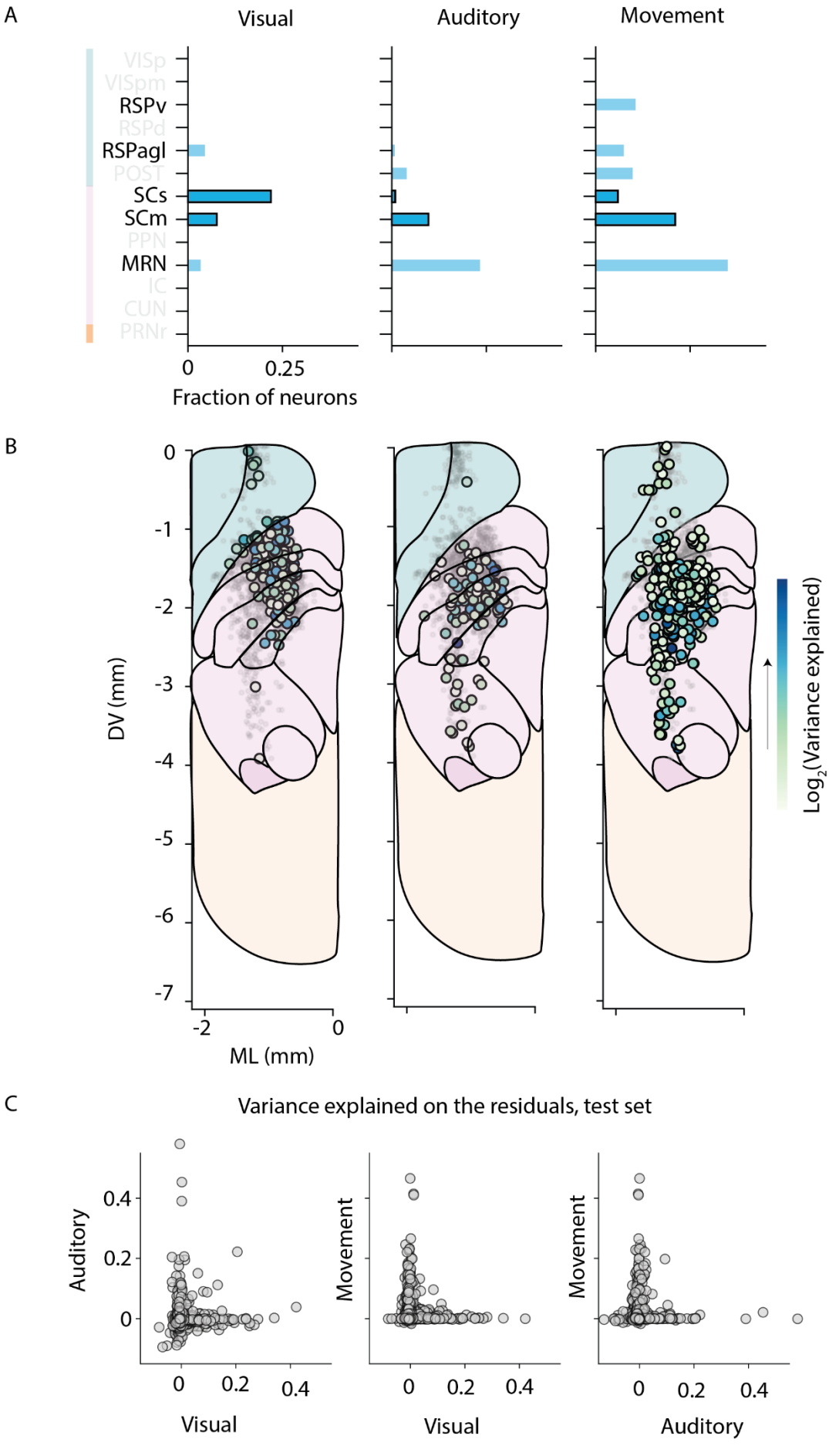
As expected, visual neurons were the most prevalent in the SCs in naïve mice, but they were also present in the SCm. Auditory neurons were restricted to the SCm. Neurons modulated by movement were spread across the entire SC, although more prevalent in the SCm (Figure 33A-B).

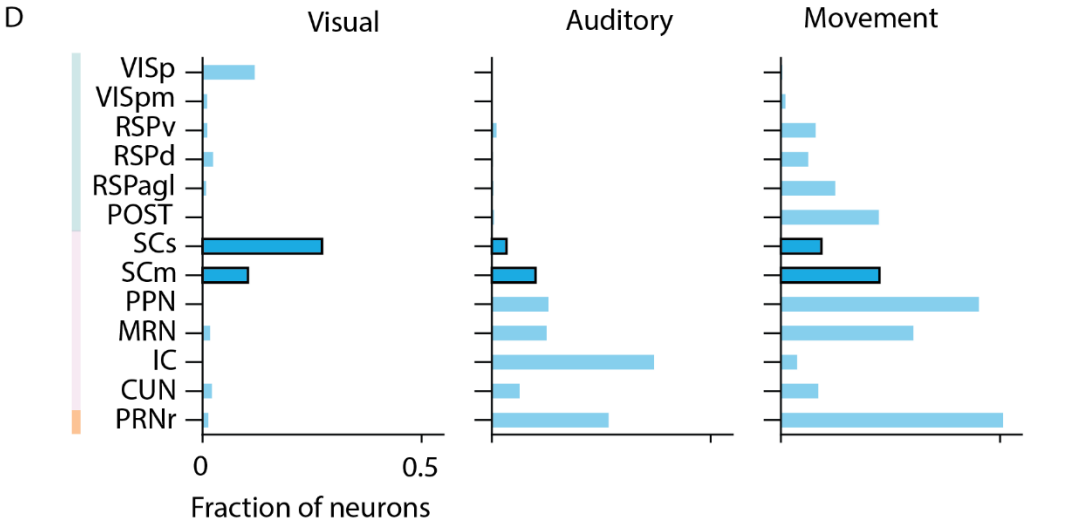
Consistent with results when analysing unisensory auditory and visual responses in naïve mice (Chapter 3), very few neurons had both auditory and visual kernel contributions even when including the audiovisual trials, suggesting that audiovisual neurons are a rarity in the superior colliculus. Interestingly, movement-related activity was also in an independent set of neurons at large, suggesting that SC neurons segregate into functional cell types representing visual, auditory, and movement signals, as opposed to mixing in individual neurons (Figure 33C).

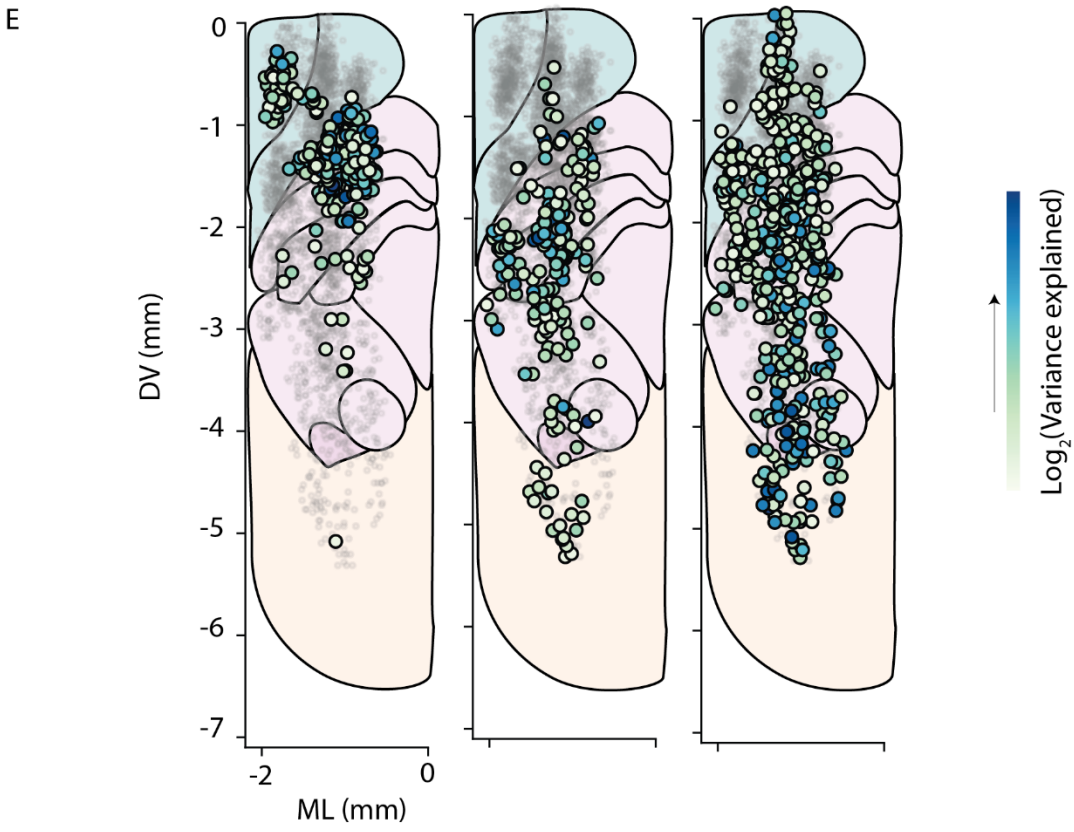
Results were qualitatively similar in trained mice during passive viewing, suggesting that training in an audiovisual task does not result in increased interactions of auditory, visual, and uninstructed movements, implying that learning this audiovisual localisation task does not require plasticity in the SC (Figure 33D-E).

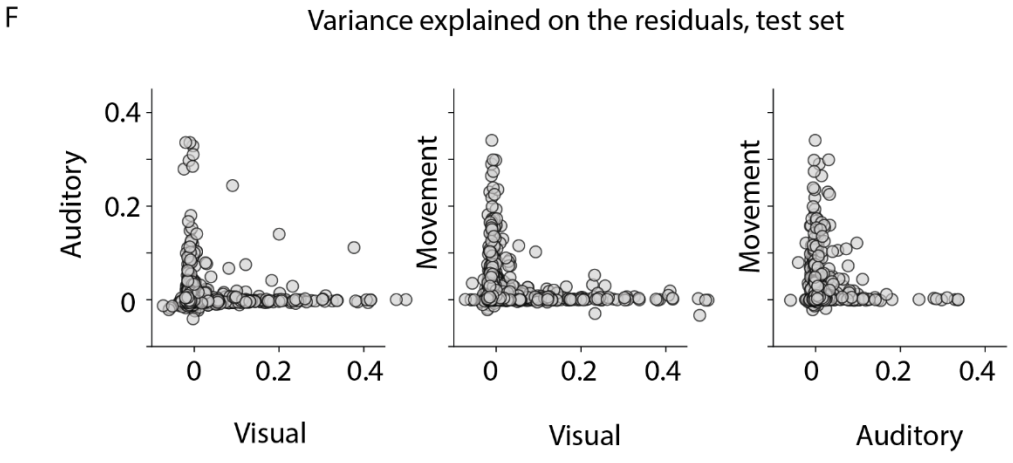
During task execution, I observed altogether fewer neurons passing the significance threshold, which is likely a result of reduced trial numbers and an increased number of fitted parameters. Most neurons appeared to respond to auditory stimulus onsets and uninstructed movements and were not tuned to auditory stimulus location or choice. Still, the signals representing auditory stimulus location as prevalent in the SCm as in most other midbrain regions. Although a few neurons appeared to have choice selectivity, these numbers were much lower than with the ccCP analysis which did not account for body movements, suggesting that body movements may explain a large fraction of choice-related neural activity (Figure 33G-H).

Even during task performance, I observed auditory and visual signals in separate sets of neurons. Auditory stimulus onset and location signals appeared to be in the same neurons, confirming that in principle, the kernel regression method can find neurons that have strong predictors in more than one kernel. Notably, visual and auditory stimulus location signal were in neurons that did not contain movement or choice signals suggesting that functional segregation is maintained in the SC during sensorimotor transformations and that stimulus representations are themselves not modulated by what the upcoming choice of the animal is going to be(Figure 33I).


A
Visual
Auditory
Movement
VISp
VISpm
RSPv
RSPd
RSPagl
POST
SCs
SCm
PPN
MRN
IC
CUN
PRNr
0
0.25
Fraction of neurons
B
0
-1
-2
-3
-4
-5
-6
-7
DV (mm)
-2
0
ML (mm)
Log2(Variance explained)
C
Variance explained on the residuals, test set
Auditory
Movement
Movement
0.4
0.2
0
0
0.2
0.4
Visual
Visual
Auditory


D
Visual
Auditory
Movement
VISp
VISpm
RSPv
RSPd
RSPagl
POST
SCs
SCm
PPN
MRN
IC
CUN
PRNr
0
0.5
Fraction of neurons


E
0
-1
-2
-3
-4
-5
-6
-7
DV (mm)
-2
0
ML (mm)
Log2(Variance explained)


F
Variance explained on the residuals, test set
Auditory
Movement
Movement
0.4
0.2
0
0
0.2
0.4
Visual
Visual
Auditory

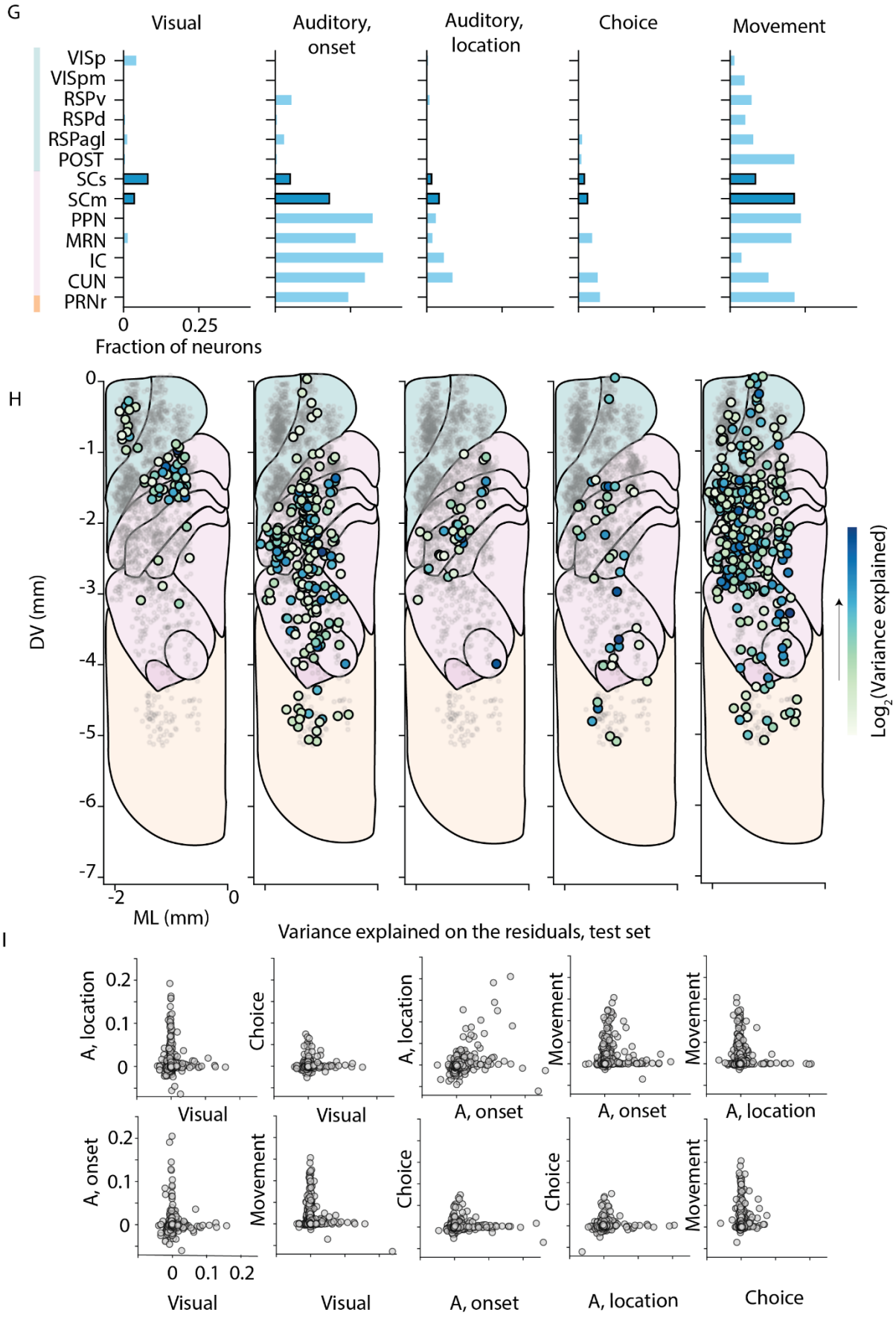

G
Visual
Auditory, onset
Auditory, location
Choice
Movement
VISp
VISpm
RSPv
RSPd
RSPagl
POST
SCs
SCm
PPN
MRN
IC
CUN
PRNr
0
0.25
Fraction of neurons
H
DV (mm)
ML (mm)
Log2(Variance explained)
I
Variance explained on the residuals, test set
A, location
A, onset
Choice
Movement
Visual

***FIGURE 33. THE MIXING OF AUDITORY, VISUAL, CHOICE, AND MOVEMENT SIGNALS IN SINGLE NEURONS.***

*(previous three pages)*

*(A) Fraction of neurons in which the visual (right), auditory (middle), and uninstructed movement (left) kernels explain >2 % residual variance on the test set across regions in naïve mice. Missing bars indicate that the region was not recorded in this dataset. (B) Anatomical location of recorded neurons. Neurons' location was projected to a coronal plane. Grey transparent dots indicate all the neurons. Coloured dots indicate neurons that passed the 2 % threshold in (A). Darker neurons indicate a higher % variance explained. (C) Pair plots of variance explained by each kernel for each recorded SC neuron (altogether 1020 neurons). Spearman correlation on each of these was <0.05). (D), (E) & (F) Same as (A)-(C) but during passive recordings of mice that were trained in the audiovisual task. (G), (H) & (I) Same as (A)-(C) but during task execution, where also the used kernel parameterisation was different compared to passive conditions. Therefore, here, the kernels included from left to right: visual kernel, auditory stimulus onset kernel (A, onset), auditory stimulus location kernel (A, location), choice direction kernel, and uninstructed movement kernel. A choice onset kernel was also fitted, but it did not have any explanatory power in neurons and therefore it is not represented on this plot.*

## Discussion

Superior colliculus neurons are sensitive to visual stimulus location, auditory stimulus location and uninstructed movement signal in separate neurons in both naïve mice and mice that have been trained in an audiovisual localisation task.

This analysis confirms results from Chapters 3 and 4: the binding of auditory and visual information in individual neurons is rare, even when analysing multisensory and unisensory trials together. Consequently, SC neurons do not contain a stimulus-modality-invariant representation of egocentric space. It is still possible that sensory neurons may undergo some weak gain modulation that is invariant of stimulus location. However, such modulation is unlikely to reflect audiovisual binding of spatial information, but rather a low-dimensional attentional modulation. It is also possible that such gain modulations are not even direct consequences of auditory inputs, but rather reflect the multiplicative gain modulations sensory neurons commonly undergo because of auditory stimulus-evoked movements and brain state changes, as has been observed in the cortex as well as the SCs (Niell and Stryker 2010; McGinley et al. 2015; Vinck et al. 2015; Schröder et al. 2020). Future work will aim to dissect possible gain state changes because of auditory stimuli or uninstructed body movements in sensory neurons across the SC.

In addition to auditory and visual signals, neurons that are additively modulated by uninstructed movements are also in separate neural populations in the SC. Thus, the SC does not exhibit mixed selectivity, which is thought to be prevalent in the posterior parietal cortex (PPC) (Raposo et al. 2014). This might be the cause because studies in PPC did not regress out uninstructed movements. Non-linear mixed selectivity is thought to implement flexible behaviours (Rigotti et al. 2013), and alternatively, it is also possible that cortical structures exclusively have a more integratory role, while SC sensory and motor neurons are essentially separate pools of neurons that can influence upstream regions for sensory modulation and downstream regions to implement orienting, however, the signals don't interact in the SC. Indeed, recent theories have questioned that mixed selectivity would be a dominant solution in discrimination tasks, and most brain regions may employ a feature-specific encoding scheme (Whittington et al. 2022).

Once I regressed out uninstructed movement-related activity, I found very few neurons that were selective of choice. This might be because the SC represents embodied choice-related activity (Selen et al. 2012), or that choice-related activity is embedded in the uninstructed movement-related activity in a more nuanced manner, and kernel regression is too stringent in regressing out such structures (Hasnain et al. 2023). The rare choice-related neurons appear to lie on the SCs-SCm border: it is indeed possible that exclusively this layer integrates ongoing motor information with sensory signals to make choices (Mohler and Wurtz 1976).

A key limitation of this analysis is that despite its power in pulling apart contributions to neural activity at a single neuron level, it does not analyse how neurons function together as a population. Future work should address the co-fluctuations across auditory, visual, choice, and movement neurons to understand their shared dynamics (e.g. Stringer et al., 2019).

# Chapter 7 — Effects of SC inactivation during an audiovisual task

## Introduction

In Chapters 3-6 I demonstrated that SC neurons contain visual, auditory, and choice signals while the animals are making decisions. In this chapter, I will investigate the impact of superior colliculus inactivation on behaviour during the audiovisual localisation task.

To my knowledge, such inactivation has not yet been carried out in an audiovisual task. Historically, (Burnett et al. 2004) unilaterally lesioned the SC and observed deficits in cats' ability to orient towards audiovisual targets, while orienting towards unisensory targets remained unaffected. However, interpreting these lesion studies is challenging due to the potential post-surgical plasticity during the recovery period. The emergence of optogenetics offers a promising avenue to address such caveats (Chen et al. 2022).

Indeed, the SC has been inactivated using optogenetics in various unisensory decision-making tasks both in rodents (Duan et al. 2015; Huda et al. 2020; L. Wang et al. 2020; Essig et al. 2021; Duan, Pagan, et al. 2021; Thomas et al. 2023) and in monkeys (Jun et al. 2021; Stine et al. 2023). However, the important stimulus feature that the animals must detect/discriminate in these tasks was not auditory or visual stimulus location, which is what the SC is thought to represent. Rather, the above tasks involved the detection of differences in auditory stimulus frequency, odour content, or pole position around the whiskers, which are thought to be encoded in the auditory, piriform, or somatosensory cortices.

Here, I inactivated the SC both unilaterally and bilaterally using optogenetics in our location-based task. This manipulation induced changes in both choice outcomes and reaction times. I assessed these changes using various computational models to understand the possible mechanisms that SC may partake in during audiovisual decision-making.

## Methods

### Hardware

For optogenetic inactivation, I used a light-activatable chloride pump halorhodopsin, eNpHR3.0 (Gradinaru et al. 2010) . Following the training in the audiovisual task, I performed unilateral or bilateral viral injection into the superior colliculus (SC) of mice.

Briefly, I made one or two 1mm craniotomies using a biopsy punch and injected the SC unilaterally or bilaterally with pAAV-hSyn-eNpHR3.0-eYFP (gift from Karl Deisseroth, Addgene plasmid #26972). In one mouse I injected AAV2-EF1a-DIO-eNpHR3.0-mCherry (a gift from the Sainsbury Wellcome Centre Viral Vector Core facility). I used a $2.5*10^{12}$ GC/ml titre throughout all experiments. In one hemisphere, I typically injected 110-140 nl virus diluted in ACSF at 6 positions (Figure 34A), to ensure widespread coverage across the SC.

Following the injections, a cannula (MFC_400/430_4mm_MF1.25_C60, Doric Lenses) was placed at the top of the injection spots and cemented to the skull; In most mice, the cannula was aimed at 3.8

mm AP, 1.1mm ML position with 0-degree tilt, but to inactivate bilaterally, I changed the tilt to 10 degrees (to ensure the patch cords fit comfortably next to each other).

The implanted cannulas (Figure 34B) were conical (C60), designed to ensure that most light remains concentrated around the target location. Cannulas were coupled to an LED (M565F3 Fibre Coupled LED, Thorlabs) using a patch-cord (FP400URT, NA=0.48, Thorlabs). We ensured that no light escapes into the visual field of the mouse.

Power was controlled by an analogue pulse (LEDD1B Driver, Thorlabs). To calibrate the outputted power, I used a cannula with a flat tip which had the same optical parameters as the implanted cannulas. Such flat-tipped cannulas project most of the light unidirectionally into the power sensor (S130C, Thorlabs), thereby making it easier to estimate the total outputted power.

I observed early in the experiments, that efficacy of inhibition varied across mice. Consequently, I used a combination of powers, to prevent saturation of the behavioural effects. Typically, I used the combination of 16,40,80,135 mW/mm$^2$ powers, measured at the tip. In some experiments, the lower powers were not used. Thus, for analyses that did not require a combination of powers, I selected the power that contained the highest number of inactivation trials.

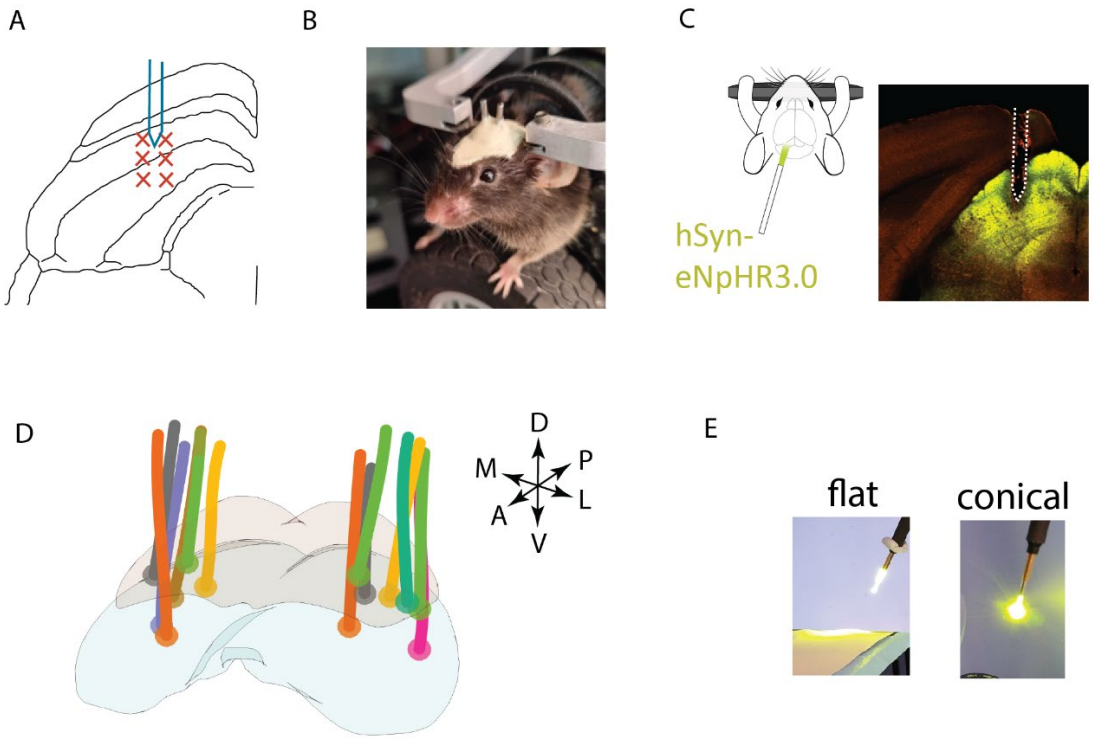



***FIGURE 34. METHODS FOR OPTOGENETIC INACTIVATION OF THE SC.***

*(A) Injection strategy. Crosses illustrate the positions of injection, while the blue object is the cannula, represented over a coronal section of the right SC. (B) Examples skull prep for bilateral implantation (i.e. portrait of AV041, courtesy of George M. Booth). (C) example image of the spread of the virus across SC after injection (marked by eYFP). (D) Illustration of all cannula tracks after they have been registered to the Allen CCF. Each colour represents a different mouse (n=9 mice in total). (E) Image demonstrating the differences in light spread when using a flat versus conical cannula tip. Flat-tipped cannulas project the light downwards, and thus most light coming out from the flat-tipped fibber, scatters on the paper placed below. Conical tips create a bolus of light around the tip.*

To verify that the virus injection was successful, I imaged the brains *postmortem* using 2P-tomography. This is essentially the same method I used for reconstructing the probe tacks described in Chapter 2. Imaging was performed optimally for eYFP imaging (Figure 34C). For this, I used a 920 nm excitation wavelength and imaged at two channels (red and green) at 10x10x25 µm resolution. For a subset of mice, I imaged the brains at higher, 1x1x3 µm, resolution with the hope of visualizing infected cell bodies versus vs axons. I later realised that eNpHR3.0 is targeted to the membrane as opposed to being distributed in the cytosol, rendering infected cell bodies invisible regardless of imaging resolution.

To verify the affected region, I reconstructed the cannula location after imaging (Figure 34D). Cannulas typically left a characteristic damage in the tissue, and so their position was identified based on this damage. The location of 3 cannula insertions location could not be identified. Anecdotally, I found that the brain tissue needs to be well-fixed before cannula removal for optimal reconstruction.

Subsequently, I attempted to understand how much tissue around the cannula might be affected. For this, I first simulated how much the light could spread around the cannula, which I estimated to be ~1mm. Then, I calculated the AAV injection success in this potentially affected area from the images around the cannula tip.

To understand the extent of the inactivation effect around the cannula, I estimated the spread of effective power. For this, I simulated the light decay using Monte Carlo simulations (Appendix A, Figure 45A) (Stujenske et al. 2015), and I calculated the distance at which the applied light reaches the effective power density ($EPD_{50}$) of eNpHR3.0, i.e. the power at which the inducible photocurrent is 50 % of a fully saturated state. For eNpHR3.0, $EPD_{50}$=5mW/mm$^2$ (Mattis et al. 2012) . Thus, at power levels of 16,40,80, 135mW/mm$^2$, one reaches eNpHR3.0 $EPD_{50}$ after a decrease to 30 %,12.5 %,6 %, 4 % from the tip, i.e. ~0.7,0.8,1,1.2 mm below the tip at most. However, the Monte Carlo simulations assume a flat cannula tip, while I used a conical tip, which restricts light the area of interest to a greater extent (Figure 34E).

To quantify the AAV infection success in the affected area, I quantified the fluorescence measured by eYFP in the green imaging channel. For this, I summed the pixel intensity values in a ~1mm$^2$ area around the tip. Subsequently, I normalised this value by the average autofluorescence (auto) calculated by summing the pixel intensities in same-sized squares 2,3, & 4mms below the SC.

### Software

Inactivation was only performed in mice that were fully trained in the task. I introduced inactivation in 33-37 %, of trials, the trials inhibited were randomly selected. Inactivation trials were never repeated, even after an incorrect choice. The LED was turned on after the end of quiescent period but before the stimulus. To prevent mice from learning the timing of the stimulus relative to the LED, the time gap between LED onset and stimulus onset was randomised. Notably, mice typically continued holding the wheel until stimulus onset even when the LED was turned on. The LED was turned on for a fixed period (1.5s) and gradually ramped down (in 400 ms) to prevent recurrent excitation, resulting in a total of 1.9s long inactivation period (Figure 35). This inactivation period was fixed such that the mouse could not learn to perform the task to terminate the inactivation.

In a subset of experiments, an additional trial type was introduced, where there was no sensory stimulus. In these trials, neither visual nor auditory stimuli were delivered, but the quiescent period was still enforced, and LED was turned on in 50 % of these trials. This trial type was introduced to evaluate the effects of LED activation on wheel turns following a period of forced quiescence, independent of stimulus presentation. This trial type was randomly interleaved during task performance, and it occurred in 4 % of the total trials.

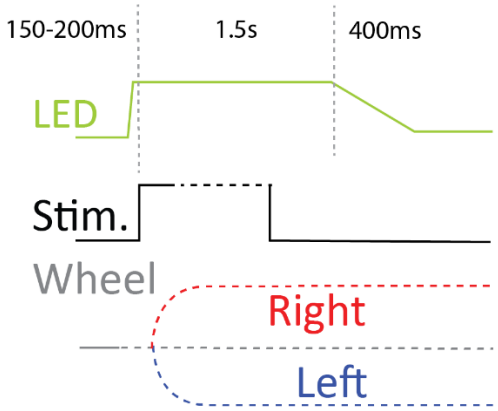



***FIGURE 35. INACTIVATION TIMINGS.***

*In optogenetic experiments, the inhibitory LED was turned on at randomised times following the quiescent period but before the onset of the stimulus. The LED was turned on for a fixed period of 1.9s in total.*

## Mice

I used 7 male and 2 female mice from the Cdh23$^{+}$ strain to perform these experiments. In the first 3 mice I tested, I initially tested only the highest power my LED system could output (135mW/mm$^2$) to assess any behavioural change could be induced. Subsequently, in 2 out of 3 mice (AV031 & AV033), I reduced the power where I observed strong effects. In the remaining mouse (AV029) the effect was weaker and thus was hard to see on single sessions.

In subsequent mice, I adapted a more sophisticated control system to enhance experimental flexibility. This system allowed for the randomisation of power selection and the interleaving of unilateral or bilateral inactivation while mice performed the task (https://github.com/cortex-lab/PinkRigs/blob/main/Zelda-Ephys/Opto/LaserController.m).

In the initial bilateral implantations (AV036, AV038), I did not manage to perform bilateral inactivation, due to the proximity of the cannulas, which prevented the simultaneous connection of the two patch cords. In subsequent bilaterally implanted mice (AV041, AV044, AV047), I successfully rectified this issue by introducing a ~10° tilt to the cannulas along the mediolateral axis during implantation.

## Results

### Basic qualitative overview

Inactivation of the superior colliculus (SC) altered mouse choices and reaction times. Mouse choices were altered differently on sessions where the left or the right SC or both colliculi were inactivated (Figure 36A), suggesting that SC inactivation can successfully cause transient changes in behaviour specifically during task performance.

Unilateral inactivation affected both auditory, visual, and audiovisual trials. Inactivation of the SC resulted in more frequent ipsilateral choices relative to control trials. Changes in choices were temporally limited to the inactivated the trials and did not affect the control trials following that followed inactivation trials (Figure 36B).

Unilateral SC inactivation sped up ipsilateral choices and slowed down contralateral choices and increased the number of timeouts. The timeouts remained unaltered on control trials that followed the inactivation trials, while the ipsilateral choices were slightly slower, and the contralateral choices were faster on the subsequent trials after inactivation (Figure 36C-D).

I implanted a subset of mice bilaterally and inactivated opposite hemispheres and then both hemispheres simultaneously on alternating days. Unilateral inactivation biased choices ipsilaterally as expected. Bilateral inactivation of SC in the same mice appeared to overall restore the behaviour (Figure 36A); the fraction of ipsilateral choices was lower on bilateral inactivation compared to unilateral inactivation in the same mice (Figure 36E). Similarly, reaction times appeared to be restored (Figure 36F). However, the fraction of timeouts still increased by unilateral inactivation (Figure 36G).

Overall, SC inactivation appears to affect choices on all trial types and changes in reaction times and timeouts depend on the direction of choice. However, from this analysis it is unclear whether inactivation differentially affects sensory sensitivity or affects all choices and how changes in timings affect the reaction time distributions. In the next sessions, I turn to modelling approaches; the Log-Odds model and a model that predicts both choices and reaction times: drift-diffusion.

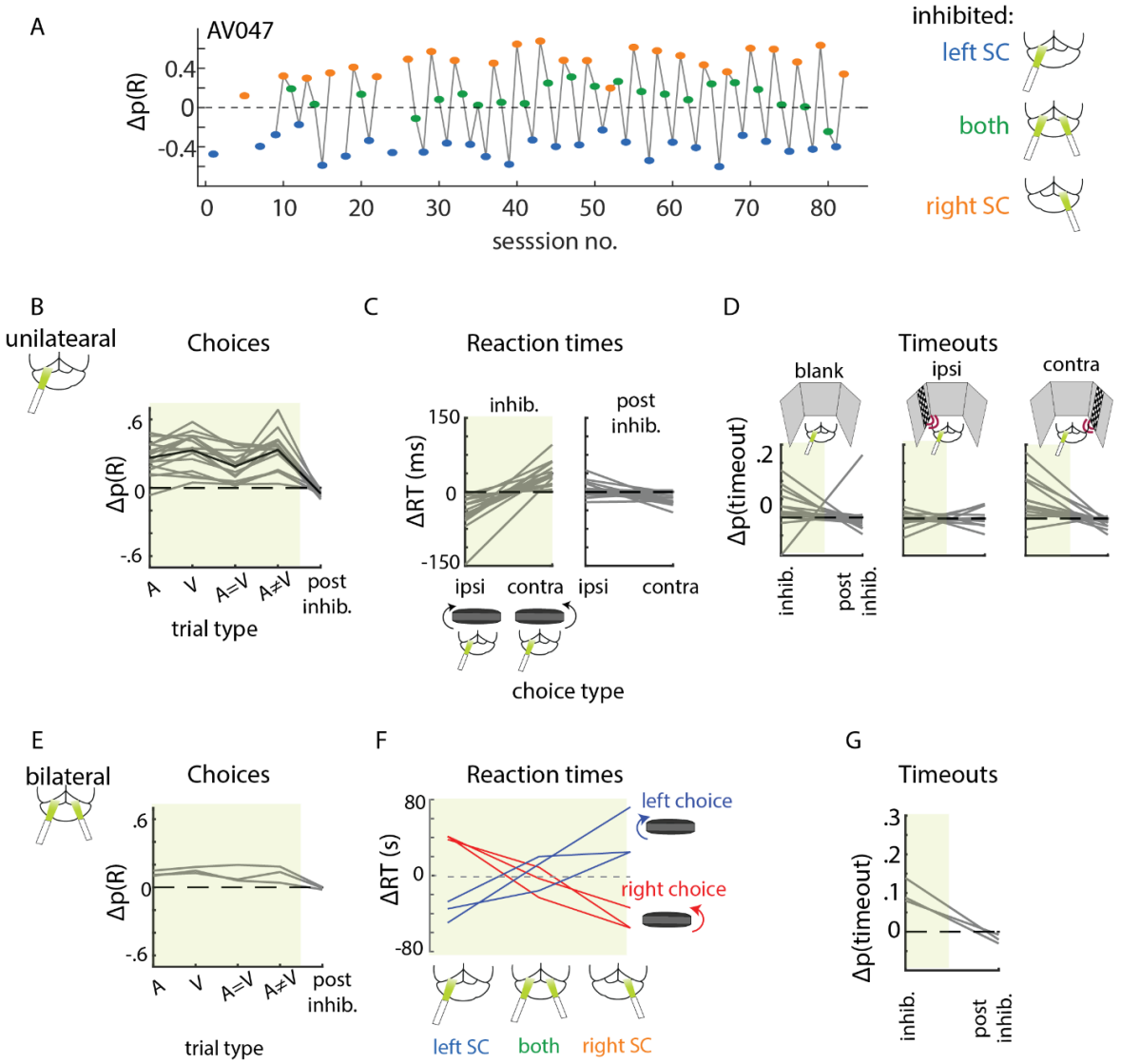



***FIGURE 36. SC INACTIVATION CHANGES CHOICES, REACTION TIMES, AND TIMEOUTS.***

*(A) Changes in the average fraction of rightward choices induced by SC inactivation in consecutive sessions in an example mouse. Missing data points indicate that the minimum number of 5 trials per sensory stimulus condition was not reached in any of the conditions on that session. (B) Effect of unilateral inactivation on choices. Changes in choices are quantified by changes of the average fraction of rightward choices per trial type on inactivated trials vs control trials (Δp(R)). Δp(R) shown on auditory, visual, coherent, and conflict trials and on control trials that occurred immediately after laser trials (post-inhibition trials). (C) Effect of unilateral inactivation on reaction times quantified as the average change of reaction time across all trial types (ΔRT). ΔRT is shown separately for ipsiversive and contraversive choices relative to the inactivated hemisphere on inactivated trials and on post-inhibition trials (paired t-test, * means p<0.01). (D) Effect of unilateral inactivation on timeouts quantified as the average change in the fraction of timeouts ((Δp(timeout)). Δp(timeout) is shown separately for blank trials and on trials when the stimulus was on the ipsilateral and the contralateral side relative to the inactivated SC hemisphere. (paired t-test, * means p<0.01).(E) Same as (B) but for bilateral inactivation (F) Effect of bilateral inactivation on reaction times, contrasted to changes in reaction times when the left or the right SC was inactivated only in the same mouse. (G) Effect of bilateral inactivation on timeouts. Changes in inactivation compared to control trials are calculated separately in each sensory stimulus condition and are subsequently averaged across all conditions.*

## Modelling effects of inactivation on choice using the Log-Odds model

To further characterise the effects of inactivation on choice, I used the Log-Odds model. First, I characterised how different parameters of the Log-Odds model – the visual, auditory, or bias – resulted in the most significant changes both during unilateral and bilateral inactivation. Then, I confirmed the results using a variation of the log-Odds model, i.e. a linear mixed effects model, which does not require the averaging over several sessions, and can thereby control for the behavioural changes mice naturally exhibit across sessions, revealing more subtle effects. Finally, I compared the parameter changes during SC inactivation with the parameter changes during the inactivation of other cortical brain regions to establish a comparative relationship between how SC may impact choices compared to other brain regions.

### *Log-Odds model to assess inactivation effects*

The log-Odds model was fitted on the choice direction on trials where animals responded within the 1.5s response window. For each session, trials were filtered to contain inactivation trials of a single power only, typically the 80 mW/mm$^2$, unless this power type contained <600 inactivation trials (2 mice). The effect of inactivation was characterised as described before (Coen et al. 2023). Briefly, mouse choices were fitted on control trials, using an additive model described in Equation 3:

$$\log\left(\frac{\mathrm{p}(R)}{p(L)}\right) = (v_L V_L^{\gamma} - v_R V_R^{\gamma} + a_L A_L - a_R A_R + \mathrm{B})$$

*Equation 311*

Where:

- $L$ represents the left side,
- $R$ represents the right side,
- $\log\left(\frac{\mathrm{p}(R)}{p(L)}\right)$ is the odds of choosing the right side compared to the left side,
- $V$ represents visual stimulus contrast,
- $\gamma$ represents a contrast scaling parameter,
- $A$ represents auditory stimulus azimuth,
- $v$ and $a$ denote how visual and auditory stimuli are weighted by the mouse,
- $B$ denotes how biased the mouse is toward the ipsi vs the contraversive choice.

Unilateral SC inactivation shifted all log-Odds curves towards more ipsilateral choices, i.e. inactivating the left SC resulted in more leftward choices, and inactivating the right SC resulted in more rightward choices. Bilateral inactivation appeared to restore mouse behaviour. In some mice, the change of the slope appeared more pronounced than in others (Figure 37, and Appendix B, Figure 46).

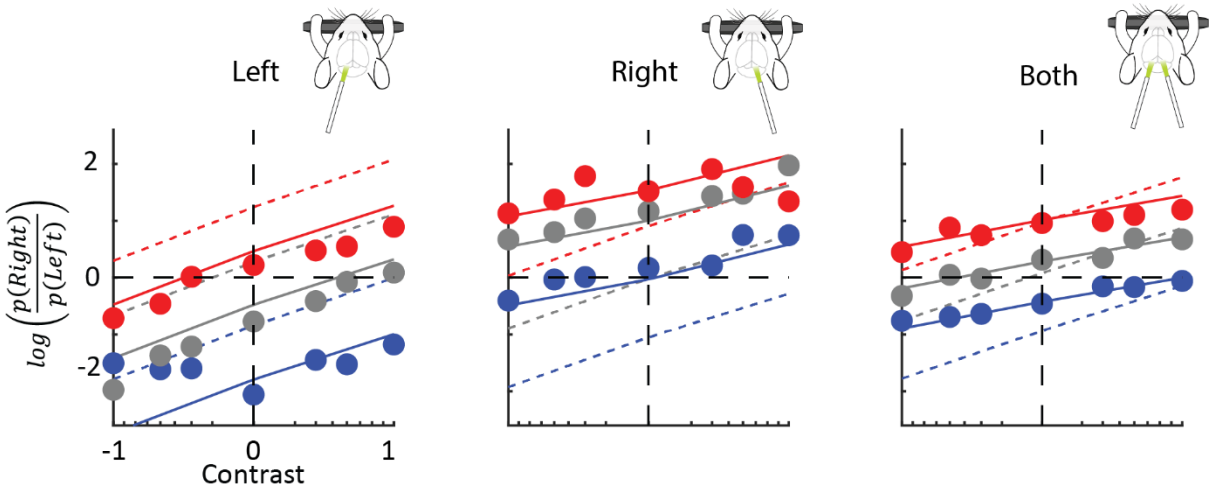



***FIGURE 37. LOG-ODDS MODEL FITS FOR INACTIVATION IN AN EXAMPLE MOUSE.***

*Log-Odds model fits on an example mouse (AV044) where SC was inactivated both in the left and the right hemisphere as well as bilaterally The lines are predictions of the log-odds model, as in Figure 11. Dashed lines: model fit on control trials. Solid lines: model fit on inactivation trials. Points correspond to the measured Log-odds on inactivation trials. The measured Log-odds on control trials are not plotted for visibility, but the prediction from the Log-odds model typically provides a good fit. As in Figure 11, an additional data point of each stimulus-choice combination was added to these plots such that the average log-odds are plottable.*

I first analysed the effects of unilateral SC inactivation by the Log-Odds model. To be able to interpret the unilateral inactivation results more easily, I renamed the model parameters such that they correspond to ipsilateral and contralateral effects relative to the site of inactivation, resulting in the following model formula:

$$\log\left(\frac{\mathrm{p}(i)}{p(c)}\right) = (v_i V_i^{\gamma} - v_c V_c^{\gamma} + a_i A_i - a_c A_c + \mathrm{B})$$

*Equation 12*

Where:

- $i$ represents the ipsilateral side relative to the site of inactivation,
- $c$ represents the contralateral side.

To describe the effects of unilateral inactivation, I compared the parameters of the Log-Odds model on control versus inactivation trials in each mouse. The sensory parameters $v_i$, $a_i$ and $a_c$ remained unaltered by inactivation. However, the parameter $v_c$ decreased, suggesting the SC inactivation reduced sensitivity to contralateral visual stimuli. The bias parameter $B$ was also significantly affected by unilateral SC inactivation resulting in an increased bias towards making ipsiversive choices (Figure 38A).

When fitting the Log-Odds model, multiple parameters can influence a single data point, making it possible for some parameters to change during inactivation trials because of overfitting as opposed to

a genuine effect. To address this concern, I assessed the contribution of each parameter, $v_i$, $v_c$, $a_i$, and $a_c$ and $B$ contributed to the average goodness of fit using cross-validation.

In high-dimensional models with a large number of parameters, this assessment is conducted by addressing two questions about each parameter (Musall et al. 2019):

1) how much does the goodness of fit increase by adding only the parameter in question compared to control trials?

2) How much does a goodness of fit decrease if I use all parameters, except the assessed one?

I refer to models addressing the first question as "gain" models, and those addressing the second question as "loss" models. To evaluate the goodness of fit of both gain and loss models I use the – $Log_2$Likelihood on the held-out data. For comparison I assess the relative – $Log_2$Likelihood of both the gain and the loss models min-max normalised by the – $Log_2$Likelihood obtained on the control model and on the full model, where all inactivation parameters are simultaneously fitted (Figure 38B).

The effects of unilateral SC inactivation seemed to be driven by changes in the bias parameter, B. Specifically, when analysing the gain models, I found that the bias parameter B alone can explain ~95 % improvement in negative $Log_2$Likelihood. Moreover, the analysis of the loss models indicated that >5 % losses in negative $Log_2$Likelihood only occurred when the bias was allowed to be altered by inactivation. This suggests that although there might be some subtle effects of inactivation on the sensory parameters, unilateral SC inactivation causes the most significant shifts in bias (Figure 38C).

It is plausible that the strength the inactivation effects depends on which part of the superior colliculus was inhibited. To investigate this, I assessed the distance between the tip of the cannula and the topographic position where the visual stimulus was represented in SCs along the anteroposterior axis. To estimate this topographic position, I used the results of receptive field mapping during electrophysiological recordings described in Chapter 2 and estimated where 45-75° azimuthal position would be represented relative to the cannulas (Figure 38D). I found that the strength of the change in the bias parameter $B$ depended on the anatomical position of the cannula relative such that the closed the cannula were to the area where the stimuli were represented, the more bias was induced by inactivation. I also observed a relationship between cannula position and $a_C$, such that contralateral auditory stimulus sensitivity was decreased to a greater extent when the cannulas were closer to the stimulus representation. Effects on $v_C$ visual sensitivity did not appear significant. The extent of the dependency between the extent of infection (measured by quantifying the relative eYFP intensity around the cannula, Figure 38F) and the changes in parameters appeared less pronounced compared to the location of the cannula (Figure 38G).

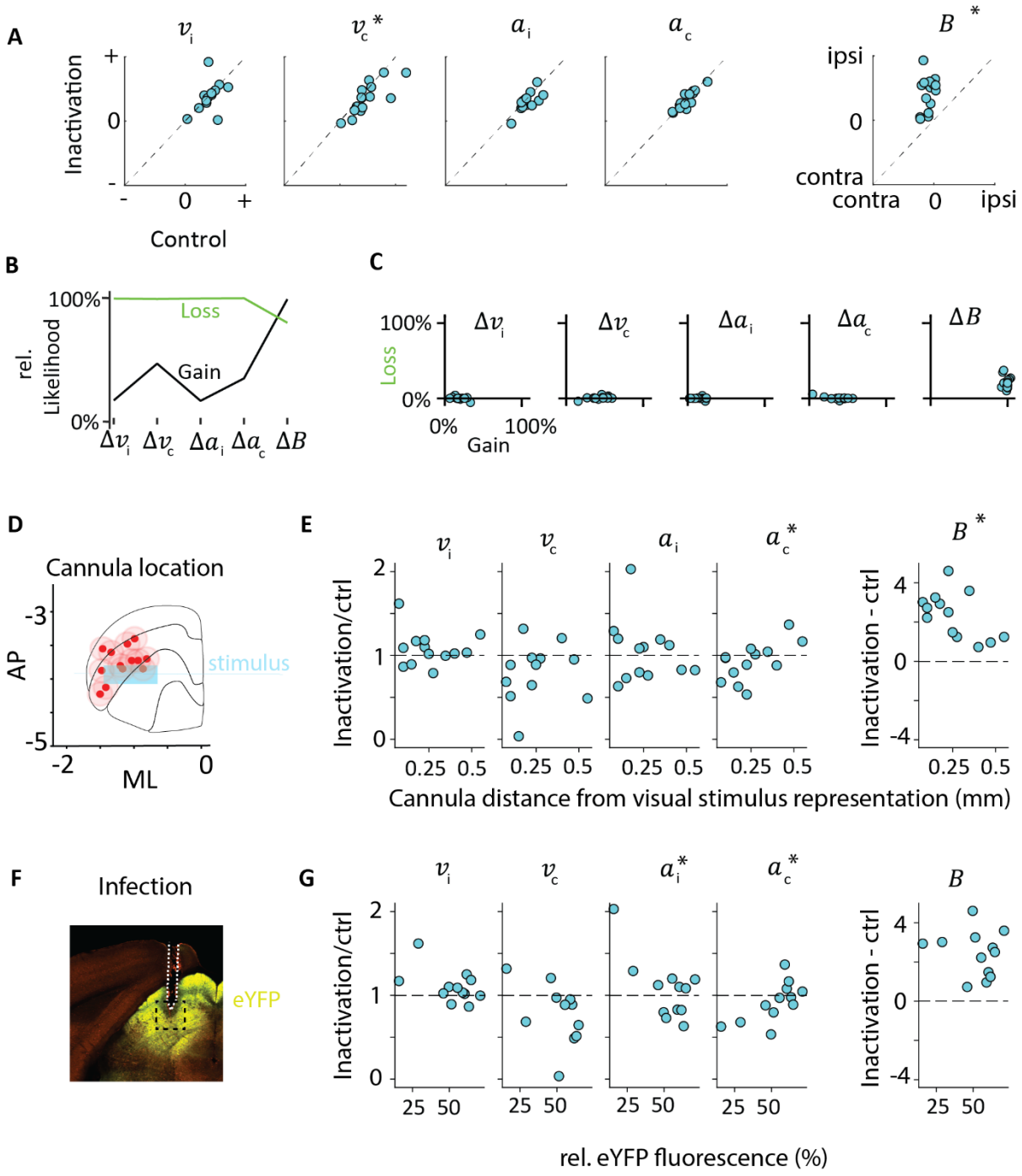



***FIGURE 38. UNILATERAL INACTIVATION OF THE SC CHANGES THE BIAS AND THE CONTRALATERAL VISUAL SLOPE.***

*(A) Values $v_i, v_c, a_i, a_c$ and $B$ from the models fitted on control vs inactivation trials for all mice. If a mouse was inactivated in both hemispheres unilaterally, it appears twice on this plot. (n=15 hemispheres form 9 mice, confirmatory analysis: paired t-test for each mouse, for two hemispheres the average was taken, * denotes p<0.05). (B) Assessment of the contribution of each parameter to fitting the changes caused by inactivation on an example mouse as measured by the relative – $Log_2$Likelihood compared to the control model and the full model when all parameters can fit the inactivation trials simultaneously. Black line: performance of gain models showing how much each parameter improves the relative – $Log_2$Likelihood on the held-out data (5-fold cross-validated). Green line: performance of loss models showing how much relative – $Log_2$Likelihood is lost if we fix a given parameter at the control value while fitting the inactivation trials. (C) Summary of the performance of gain and loss-models of each parameter Values $v_i, v_c, a_i, a_c$ and $B$ for all mice. (D) Location of the cannulas (red dots)*

*relative to the estimated area of the representation of visual stimulus registered to the Allen Atlas (blue rectangle). (E) Changes in parameter values as a function of distance from the approximated representation of the visual stimulus in SCs the anteroposterior axis. Since the sensory parameters can only take positive values, changes are represented as the ratio of parameter value on inactivation fits vs control fits. Meanwhile, bias can take both positive and negative values, and changes were expressed as the difference between inactivation and control fits (Confirmatory analysis: linear mixed effects model controlling for random effects of mouse and hemisphere identity, where model formula was parameter ~ distance + 1|mouseID + 1|hemisphere,* denotes p<.05). (F) Illustration of how I quantified the extent of infection measured by the relative intensity of eYFP in a ~1mm$^2$ area around the cannula, illustrated by the dashed square below the cannula. (G) Same as E but changes in parameter values are plotted as a function of relative eYFP fluorescence.*

Bilateral inactivation of the SC impacted both the sensory and the bias parameters, although the changes in the bias parameter $B$ were less pronounced compared to unilateral SC inactivation (Figure 39A-B). To investigate whether the overall bias change during bilateral inactivation stemmed from differences in inactivation efficacy during the unilateral inactivation, I examined the changes in $\mathrm{B}$ across various power combinations of inactivation.

I found that the greater the difference in the power applied on each hemisphere, the more biased behaviour was during bilateral inactivation (Figure 39C). This suggests that the SC may not have been fully inactivated at lower powers of inactivation, allowing further increased efficacy as the power difference increased. Moreover, the bias during bilateral inactivation could be predicted from the difference between the biases during the unilateral inactivation of each SC hemisphere, and this relationship was additive (Figure 39D). This additivity suggests that the difference in the activity of SC hemispheres is read out by a downstream integrator and that non-linear competition is unlikely to be implemented within the SC via commissural fibres.

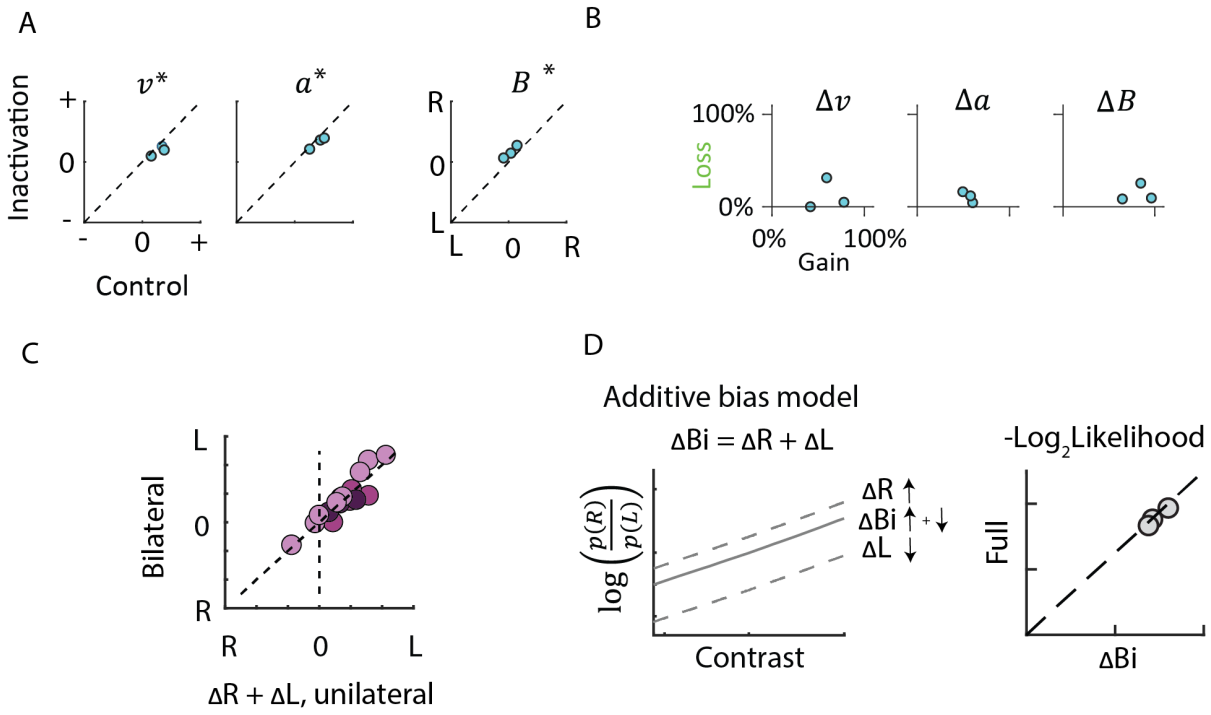



***FIGURE 39. ADDING THE BIAS CHANGES DURING UNILATERAL INACTIVATION PREDICTS BILATERAL EFFECTS.***

*(A) Visual, auditory, and bias parameters values when the Log-Odds model was applied to the control versus the inactivation trials. (paired t-test, * p<0.05).(B) Parameter contribution measured by gain and loss modes during bilateral inactivation to capture the changes by bilateral inactivation. (C) Relationship between the bias changes during unilateral vs bilateral inactivation at different power combinations. The x-axis represents the sum of the bias parameter changes induced by the right and left SC inactivation, while the y-axis represents the actual bias change during bilateral inactivation. Each dot corresponds to a different combination of power (e.g. 10 mW on the left & 5 mW on the right hemisphere). Each colour denotes a different mouse. (D) Left: Cartoon of the additive bias model of bilateral inactivation effects where adding the signed bias changes during the left and right hemisphere inactivation can predict the bias in the bilateral inactivation. Right: the – Log$_2$Likelihood of the additive bias model vs the full model of for the bilateral inactivation.*

### *Linear Mixed Effects Model*

While the Log-Odds model fitted on the pooled data reveals an effect on the $bias$ term by SC inactivation, the effects on the sensory terms $v$ and $a$ remained ambiguous. One known disadvantage of the pooling across all sessions is that it does not account for the variability induced by sessions, because it is fitted on the average choices across all sessions. However, there appears to be some variability in both the control behaviour and the inactivation effects across sessions, and thus it is possible that by chance, the sessions with more inactivation trials are more biased in general, which could bias the Log-Odds model. Therefore, I also applied a linear mixed effects model, that allows to test for effects of interaction between auditory, visual, and bias terms while controlling for the common effects of mouse and session.

More formally:

$$p(R) = logit(aA + vV + B + a_oOA + v_oOV + B_oO + random\ effects)$$

*Equation 13*

Where:

- $V$ is the visual contrast raised on a fixed power which is the average gamma power across mice fitted by the Log-Odds model.
- $A$ indicates auditory stimulus azimuth.
- O is a boolean indicating whether a given trial was inactivated or not
- $a, v, B$ are parameters defining the auditory weight, the visual weight, a behavioural bias
- $a_o, v_o, B_o$, and the optogenetics effects on $a, v, B$.
- $random\ effects$ indicate a combination of all the additional regressors that built from of mouse ID and session ID, and their interactions $V, A$ and $O$ to account for variability across sessions and mice.

To fit this model, I used the '**fitglme'** function in MATLAB with the following formula:

```
choice ~ opto+ aud + vis + aud*opto + vis*opto + …
        (1|mouse) + (1|mouseID:sessionID) + …
        (-1+vis|mouseID) + (-1+vis|mouseID:sessionID) + …
        (-1+aud|mouseID) + (-1+aud|mouseID:sessionID) + …
        (-1+opto|mouseID) + (-1+opto|mouseID:sessionID)
```

I found that the linear mixed effects model, showed a positive $B_o$ parameter confirming the significant shift towards more ipsilateral choices induced by bilateral SC inactivation. Additionally, $v_o$ but also to a lesser extent $a_o$ was negative showing a reduction in both auditory and visual sensory sensitivity in case of unilateral and bilateral inactivation. Effects of all these parameters $a_o, v_o, B_o$ were significant predictors in the model, suggesting that SC inactivation might minimally alter sensory sensitivity too, in addition to bias (Figure 40).

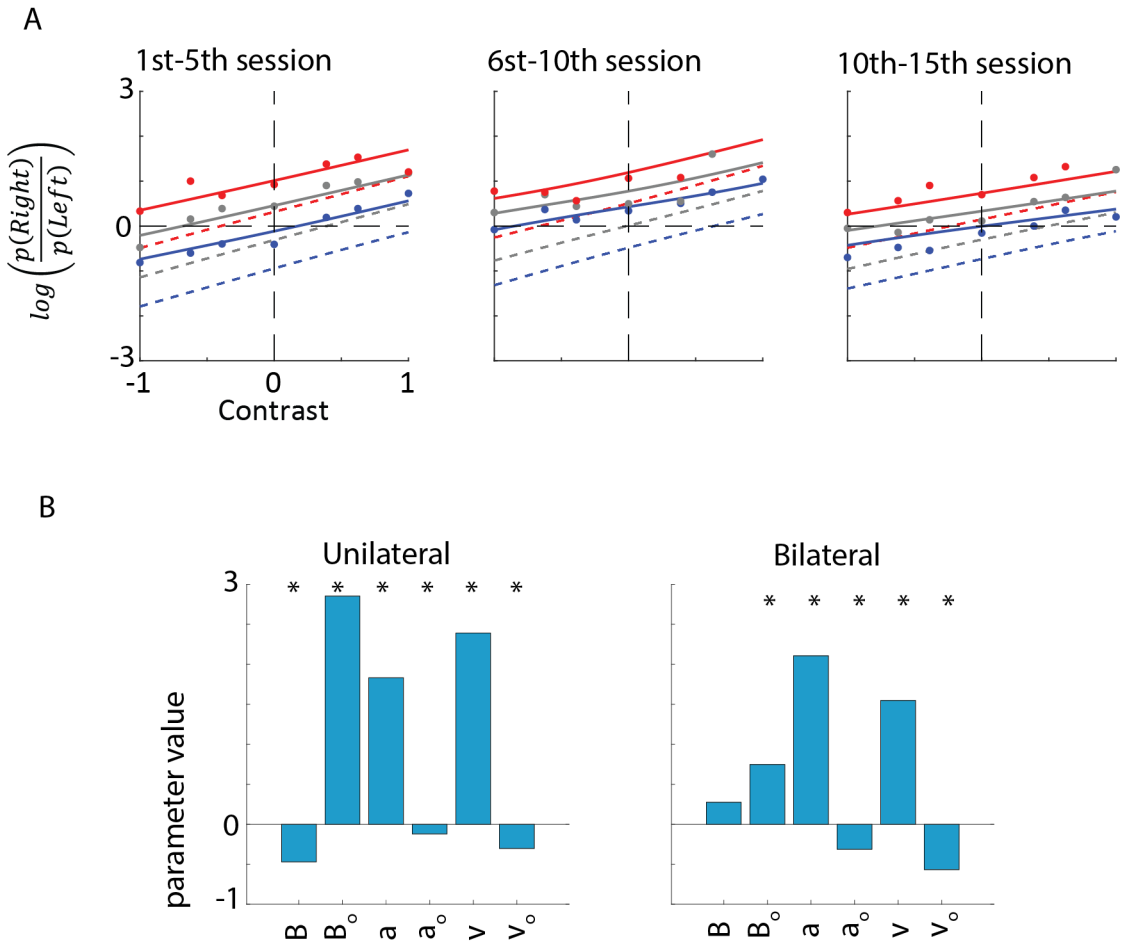



***FIGURE 40. EVALUATION OF OPTOGENETIC INACTIVATION WITH THE LINEAR MIXED-EFFECTS MODEL.***

*(A) Fit of the linear mixed effects model on an example mouse, averaged over sessions. Prediction across sessions was also averaged. Missing data points indicate that mice made rightward choices and when the Log-Odds could not be taken. (B) Fixed effect parameters of the linear mixed effects model for unilateral and bilateral inactivation. Parameters with significant predictive power ($p<0.05$) are marked with an asterisk (*). Left: unilateral inactivation. Right: bilateral inactivation.*

*Comparison to inactivating other brain regions*

To make quantitative comparisons of inactivating SC with inactivating other brain regions, I reanalyzed the publicly available data obtained by (Coen et al. 2023). In this dataset, the frontal, visual (VISp), and lateral (possibly auditory) cortex were unilaterally inactivated. In a subset of mice, the frontal, visual, and parietal cortices were also inactivated bilaterally, although only one visual contrast (5-6 %) was used in these experiments. Inactivation was performed by exciting cortical inhibitory (parvalbumin cells) using Channelrhodopsin (ChR2). I refitted the Log-Odds model on each mouse and assessed the magnitude of parameter change for each region and mouse.

Unilateral inactivation of each of these brain regions appeared qualitatively different: inactivating VISp only affected choices contralateral to the inactivated hemisphere, while inactivation of the frontal and lateral cortices as well as the SC appeared to affect all choices (Figure 41A). These differences were reflected in the parameter changes across mice. While the changes in $v_i$, $a_i$ and $a_c$ did not differ significantly across regions, the bias parameter changed the most substantially in the SC, while $v_c$ changed the most in VISp (Figure 41B-C).

Bilateral inactivation of both the frontal cortex and the SC appeared to restore behaviour, such that the changes in bias were not different anymore across regions. Meanwhile, bilateral inactivation of both visual regions, PPC and VISp, reduced the visual parameter $v$ such that the changes in visual parameters still differed significantly across regions (Figure 41D-E).

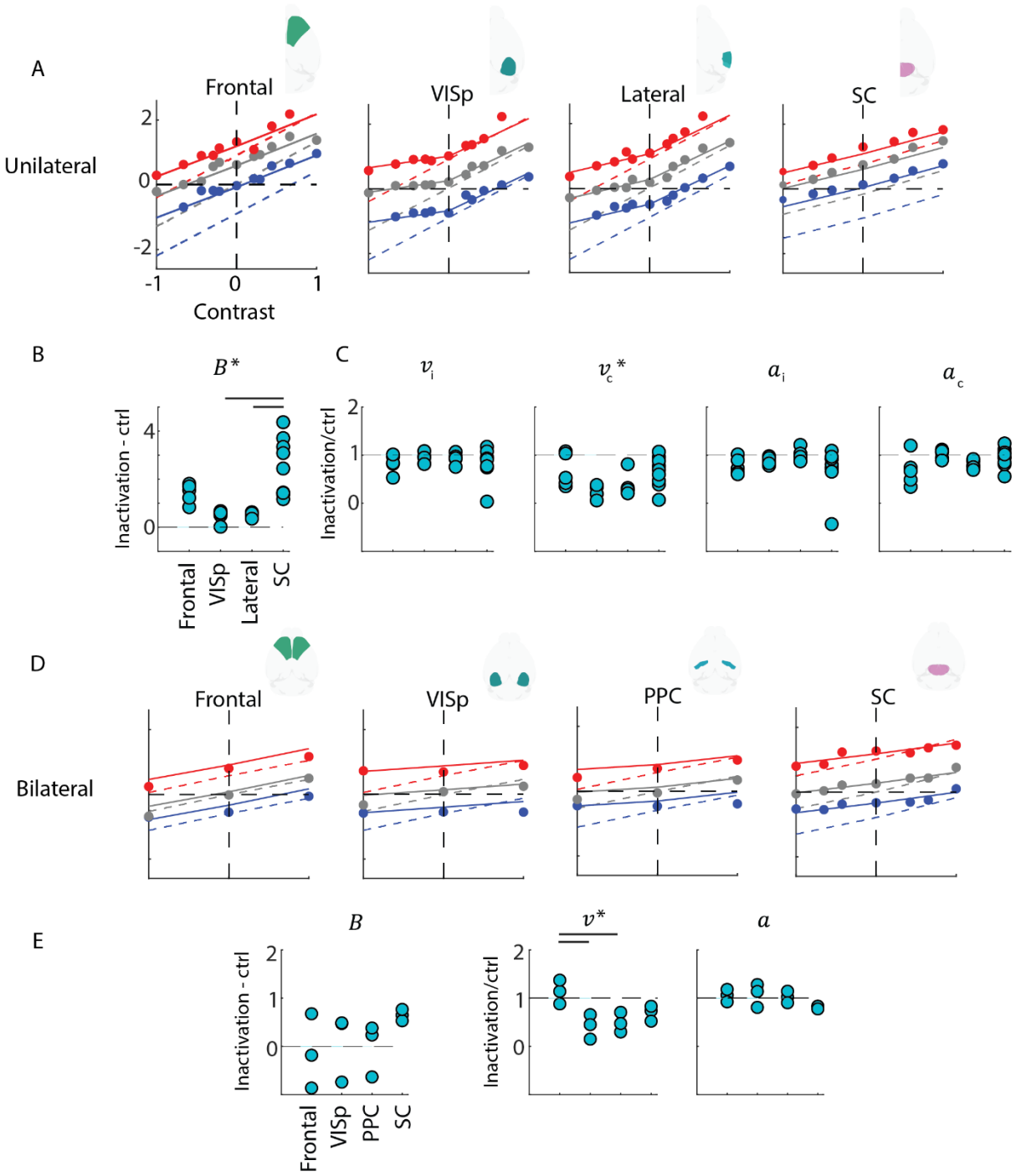



***FIGURE 41 SC INACTIVATION CAUSES STRONGER BIAS EFFECTS THAN THE INACTIVATION OF CORTICAL REGIONS.***

*(A) Effects on mouse choices because of unilateral inactivation of the frontal cortex (n=5 mice), visual cortex (n=5 mice), auditory cortex (n=5 mice), and superior colliculus (n=9 mice). Trials were pooled across all mice and all inactivated hemispheres. (B) Change in bias parameter (inactivated – control) across unilaterally inactivated regions. (C) Changes in sensory parameter values (inactivated parameter/control parameter) for each mouse when frontal, visual, or lateral cortex or the SC was inactivated. (D) Effects of bilateral inactivation of the frontal cortex (n=3 mice), visual cortex (n=3 mice), parietal cortex (n=3 mice), and the superior colliculus (n=3 mice) on mouse choices. Trials were pooled across all mice and all inactivated hemispheres. (E) Changes in sensory and bias parameter values during bilateral inactivation. (F) For confirmatory analysis, I used ANOVA with Tuckey's HSD test for posthoc comparisons throughout, * denotes p<0.05 for panels (B),(C), and (E).*

### The relationship between wheel movements and SC inactivation without task stimuli

One potential explanation for the changes observed in choices due to SC inactivation could be attributed to reflexive wheel movements rather than a direct alteration of the perceptual decision-making process Therefore, I explored whether SC inactivation onset triggers wheel movements in the absence of visual or auditory stimuli. To investigate this, I interleaved fake trials while the animal performed the task, where the quiescent period could be followed by the inactivation onset, but not by any auditory or visual stimuli. I called these “no-stimulus” trials (Figure 42A). On no-stimulus trials SC inactivation occurred with 50 % chance, such that effectively there were also “no-stimulus” control trials, where nothing happened after a period of enforced quiescence (Figure 42B).

In the absence of a sensory stimuli, both unilateral and bilateral inactivation resulted in earlier wheel movements than on control trials (Figure 42C-D). However, the wheel movement induced was not immediate, but even in the case of inactivation trials, it was >150ms after the inactivation onset suggesting that SC inactivation did not simply trigger a reflexive movement.

Moreover, during unilateral inactivation, the wheel turns were biased (Figure 42E). Specifically, when the SC was inactivated, wheel turns were preferably executed ipsilaterally towards the inactivated side, even in the absence of a stimulus.

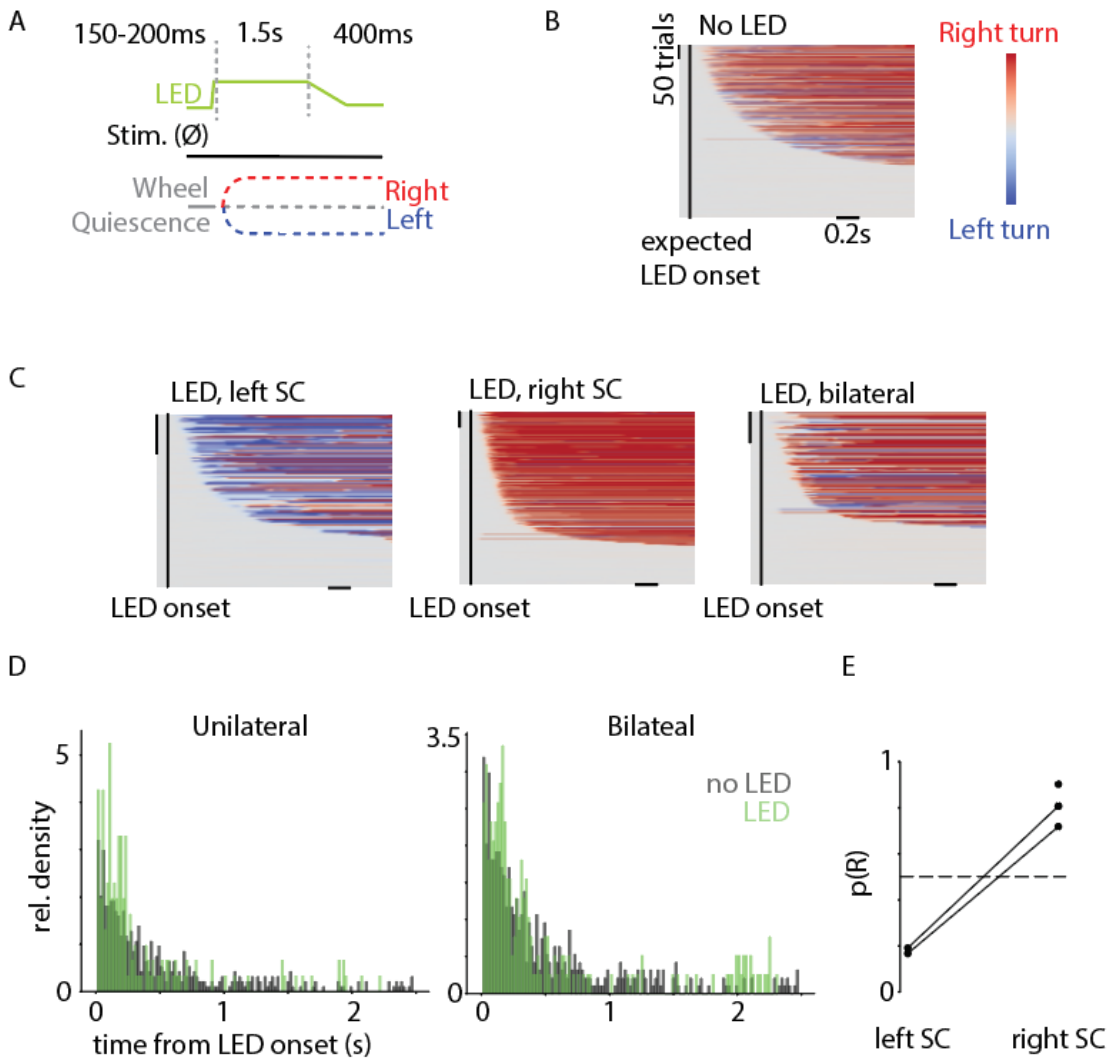



***FIGURE 42. SC INACTIVATION BIASES WHEEL TURNS IN THE ABSENCE OF A STIMULUS.***

*(A) Trial structure. On 4 % of the trials, I introduced the LED-only trial in 3 mice. On this trial type, I enforced a random quiescent period after which the LED was turned on in 50 % of trials but was not followed by an auditory or visual stimulus onset (no stimulus trials). (B) A raster plot of wheel turns on no stimulus trials when the LED was not turned on. Wheel turns are sorted by their onset time relative to the end of the enforced quiescent period. (C) The wheel turns no stimulus trials when SC was inactivated. Left: left SC inactivated, middle: right SC inactivated. Right: both superior colliculi inactivated. (D) Distribution reaction times, i.e. wheel turn onset times relative to the end of the quiescent period on no stimulus trials. Grey distribution is without inactivation (i.e. same data as in B, median & median absolute deviation: 286 ± 214 ms), and green distribution indicates inactivation trials. Left: unilateral inactivation (180 ± 114 ms). Right: bilateral inactivation (215 ± 139 ms). (E) Fraction of rightward turns when the left vs the right SC hemisphere was inactivated on no stimulus trials. Connected dots indicate that these measurements come from the same mice.*

## Modelling effects of inactivation on both choices and reaction times

While assessing models that can predict mouse choice, I showed that unilateral SC inactivation results in profound changes in bias. Moreover, I showed that these bias changes are not simply the result of reflexive wheel movements that the inactivation may induce. Here, I investigate what this bias change might mean in a cognitive decision-making framework: the drift-diffusion model (DDM). In the DDM framework, choices are made by accumulating evidence from a given starting point over time until the decision threshold is reached. Therefore, this model captures not only the direction of the choice but also predicts the time it takes to make the choice (Ratcliff 2006). Choice bias can stem from two sources: firstly, by altering the initial distance from the decision threshold either by changing the starting point or changing the decision threshold itself, and secondly, by introducing a constant drift while accumulating evidence. A shift in the starting point is often interpreted as response bias that results from preferring a certain response (e.g. a predilection towards making rightward movements) whereas adding a constant drift is often interpreted as preferentially accumulating evidence from certain stimulus sources (White and Poldrack 2014; Urai et al. 2019). Such drift-diffusion models have been applied to perceptual decision-making tasks while the SC was inactivated in monkeys. While (Jun et al. 2021) found that SC inactivation appears to be consistent with adding a constant drift, (Stine et al. 2023) suggested that SC inactivation increases decision bound.

Here, I first simulated how the reaction time distributions change when changing the starting point, the decision threshold (a.k.a. bound), or adding a constant drift. Then, I assessed whether changes in any of these parameters explain the inactivation data the best by fitting an additive drift-diffusion model to mouse choices and reaction times. Notably, such a drift-diffusion models have not been applied to rodent decision-making while inactivating the superior colliculus to my knowledge.

### Simulations

To understand the dynamics of the DDM, I simulated a simple version of the DDM using pyDDM (Shinn et al. 2020). This reduced model consists of a starting point $x_0$, and random drift is accumulated until either the lower or the upper boundary is reached, denoting rightward and leftward decisions. Keeping all parameters fixed except the starting point, the bound, or the constant drift, I simulated how altering these values would impact reaction time distributions.

Modifying the starting point or the bound influences the peak of the resulting distribution. Notably, the starting point and the bound are interdependent parameters, making it challenging to adjust them simultaneously without imposing constraints, as changes in one can be counterbalanced by adjustments in the other.

Modifying the constant drift, on the other hand, does not impact the peak, but only affects the tail of the reaction time distribution (Figure 43). Thus, the DDM offers a framework to discern whether bias changes stem from response bias, or evidence accumulation bias, by comparing how the reaction time distribution changes by unilateral SC inactivation. To explore this distinction, I compared model fits where the effects of inactivation were parameterised by changing the starting point or the constant drift and assessed how well these models fit empirical data.

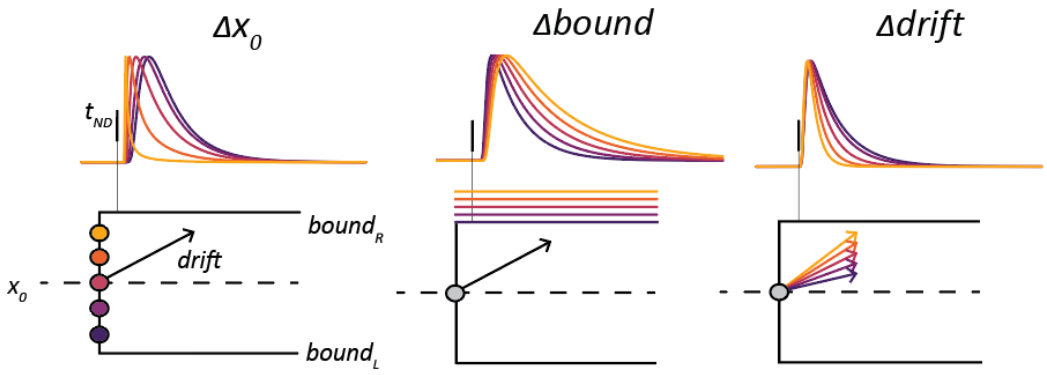



***FIGURE 43. SIMULATING THE EFFECTS OF STARTING POINT, BOUND, AND DRIFT CHANGE IN THE DDM.***

*The simulation model consisted of 4 parameters: the drift parameter denoting how much momentary value is added to the decision evidence at each timepoint during the accumulation process, a decision bound that denotes the upper and the lower threshold boundary, denoting commitment to rightward and leftward choices ($bound_R$ and $bound_L$) , a starting point ($x_0$) determining the ratio of the start position for evidence accumulation in between the bounds and a non-decision time ($t_{ND}$), that is a constant determining the minimum time to execute the decision.*

*Simulations with (A) changing starting point (B) increasing bound (C) increasing the constant drift. Each plot shows a cartoon of how the relative position of parameter values is simulated, and the top shows the resulting rightward decision reaction time distributions normalised by their peak value.*

### Fitting the additive DDM

I fitted the mouse choices and reaction times simultaneously with generalised DDM models. Reaction times were defined as the duration to reach decision threshold from inactivation (LED) onset. This approach enabled the fitting process to capture potential changes induced by the inactivation even before the stimulus onset, such as alterations in the non-decision time. Changes in non-decision time would suggest that SC inactivation is altering motor execution rather than perceptual decision-making processes triggered by the sensory stimulus.

Generalised DDMs were implemented with the pyDDM package (Shinn et al. 2020). To speed up the fitting procedure, I parallelised fitting on the UCL cluster, Myriad. The fullest model consisted of 19 parameters (Table 1.).

| **Task variables** | |
|---|---|
| $A_L, A_R$ | Auditory evidence strength on the left and the right ranging between 0 and 1 |
| $V_L, V_R$ | Visual evidence strength on the left and the right ranging between 0 and 1 |
| $O$ | 0 = control trials; 1=inactivation trials |
| **Parameters** | |
| $\sigma$ | Instantaneous noise during the evidence accumulation |
| $\alpha_L, \alpha_R, \nu_L, \nu_R$ | Auditory and visual drift magnitude |
| $\alpha_{L_O}, \alpha_{R_O}, \nu_{L_O}, \nu_{R_O}$ | Auditory and visual drift magnitude change caused by inactivation |
| γ | Visual contrast scaling parameter |
| $B, B_o$ | Drift bias magnitude |

| | |
|---|---|
| $x_0$ | Starting point |
| $p_{mixture}$ | Mixture coefficient |
| **Miscellaneous** | |
| x | Decision variable |
| W | Wiener process |
| $I_{O=1}$ | Function indicating the presence of inactivated trials |

***TABLE 1. SUMMARY OF NOTATION FOR DESCRIBING THE DRIFT DIFFUSION MODEL.***

Briefly, momentary evidence x was modelled using a drift term μ and a noise term σ, such that

$$dx(t) = \mu(A_L, A_R, V_L, V_R, O)dt + \sigma dW$$

*Equation 14*

Where, drift is essentially given at each timepoint t, from signed auditory and visual evidence and a constant drift term, such that:

$$\begin{aligned}\mu(A_I, A_C, V_I, V_C, O)\mathrm{dt} = \quad & (\alpha_L + I_{o=1}\alpha_{R_O})\mathrm{A}_R - (\alpha_\mathrm{L} + I_{o=1}\alpha_{L_O})\mathrm{A}_L + \\ & (\nu_\mathrm{R} + I_{o=1}\nu_{L_O})V_R^\gamma - (\nu_\mathrm{L} + I_{o=1}\nu_{L_O})V_L^\gamma + \\ & B + I_{O=1}B_{O.}\end{aligned}$$

*Equation 15*

Here, $dx$ is integrated from the starting point of $\chi_0$, and the decision is terminated when the evidence $|x(t)|$ crosses the bound.

The model's reaction time distribution is given as a mixture model between the first passage time distribution across this boundary and distribution of lapses sampled from an exponential distribution, where the ratio between the first passage time distribution coming from the drift process and the lapse distribution determined by the parameter, $p_{mixture}$. I also include a non-decision time $t_{ND}$, that may shift the total values of the entire reaction time distribution but will not change the shape of the distribution.

While model testing, I each of these parameters $\chi_0$, bound, $p_{mixture}$ and $t_{ND}$ were tested to account for changes by optogenetic inactivation in addition to a drift parameter by adding an inactivation parameter corresponding to each that could additively change the control values on optogenetic trials indicated by $I_{O=1}$.

Trials were split into training and test sets using the *StratifiedShuffleSplit* method in *scikit-learn* (Pedregosa et al. 2011) which splits trials into training and test set randomly while ensuring that each trial type is represented in all folds equally. Model performance was evaluated using the negative $\mathrm{Log_2}$Likelihood.

### Fit results

Fitting the DDM, I aimed to determine whether inactivation parameters significantly contributed to fitting the model to the data. This assessment mirrored the evaluation process of the Log-Odds model. Specifically, I calculated the relative negative $\mathrm{Log_2}$Likelihood on gain and loss models of each parameter, i.e. the change in the goodness of fit by allowing only the parameter in question to account

for all the changes by inactivation, and the goodness of fit when the parameter in question is set to 0 while all other inactivation parameters can improve the fit.

I found that the majority of inactivation parameters were unnecessary, and most changes could be accounted for by parameters that affected the overall bias, consistent with the results of the Log-Odds model. Particularly, introducing a constant drift by inactivation led to better fits compared to allowing changes in the starting point on average (Figure 44). In one instance, the starting point change yielded to a better fit than the change in the constant drift. In another instance, the non-decision time was the primary parameter affected; however, it remained unchanged in all other mice (Appendix B, Figure 47). This suggests that the inactivation of the SC primarily alters the perceptual decision-making process rather than inducing wheel movements.

Furthermore, I tested a model that explicitly models the role of SC as a decision threshold, as proposed by (Stine et al. 2023). In this model, I allowed both the bound and the starting point to vary, but the bound was constrained to only increase by the inactivation parameter. However, this model did not fit as well as the constant drift model (Figure 44C). Qualitatively, if the SC encoded the decision threshold itself, unilateral inactivation of the SC would be expected slow down contralateral choices while leaving ipsilateral choices unchanged, contrary to what I observed.

All in all, I found that injecting an ipsilateral constant drift by the inactivation of the SC explains reacting time changes the best.

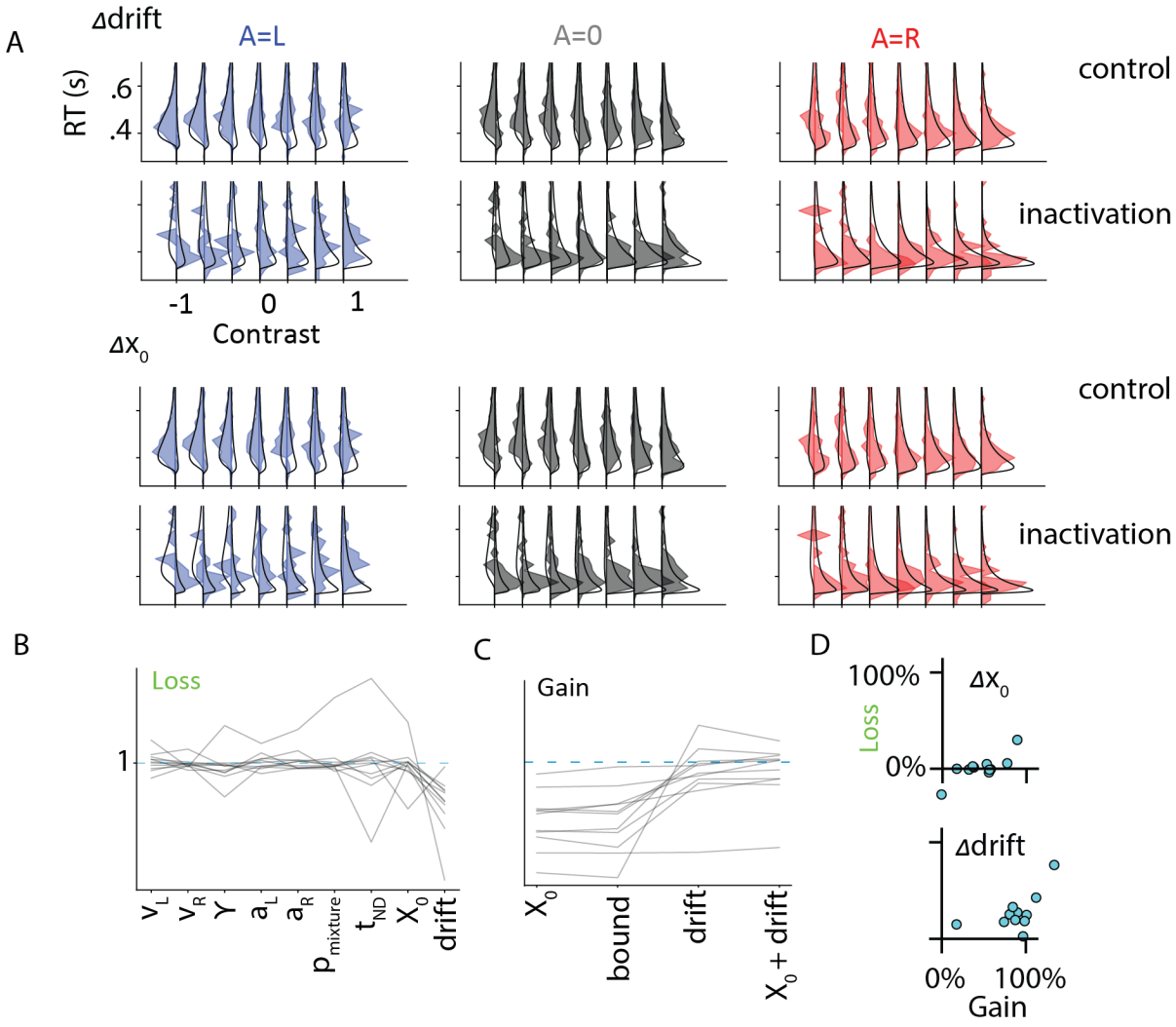



***FIGURE 44. CONSTANT DRIFT EXPLAINS BEST THE UNILATERAL INACTIVATION EFFECTS.***

*(A) Example fit by the constant drift and the starting point model of to the same mouse. (B) Normalised gain of -Log2Likelihood by each parameter if being altered unilaterally. (C) Assessment of gain and loss by the constant drift, the starting point, and the growing bound model.*

## Discussion

Unilateral optogenetic inactivation of the SC biases mice to making more ipsiversive choices. This bias effect is independent of the sensory stimulus, and wheel turns are biased even in the absence of sensory stimuli, suggesting that SC plays a role in action selection, independent of stimulus perception. Crucially, wheel turns do not appear on a reflexive timescale during SC inactivation, and SC inactivation does not reduce the non-decision time, or the peak of the reaction time distribution. Therefore, it is unlikely that SC inactivation is simply altering the motor execution of choices. Rather SC appears to play a role in perceptual decision making.

Surprisingly, the superior colliculus plays a minimal role in setting the sensory sensitivity parameters both for audition and vision. The effects on the bias parameter may be so strong that we simply cannot observe any further nuances; that is, we saturated the psychometric curves such that mice always

make ipsiversive choices. However, we can reject this hypothesis because we can further increase the bias by applying higher powers of inactivation.

It is also possible that I would observe stronger effects on visual sensitivity if I specifically targeted the superficial layers of the SC (SCs). Indeed, most cannula tips were in the SCm rather than the SCs. SCs neurons are thought to increase in responsivity (Steinmetz et al. 2019) during task execution and are modulated by reward during task performance (Baruchin et al. 2023). Furthermore, and various types of SCs neurons play a differential role in prey-capture (Hoy et al. 2019), suggesting that SCs might have a play in executing visual tasks.

However still, the SCm also contains visually responsive neurons and plays a role in stimulus detection both in natural (Evans et al. 2018) and learnt behaviours (L. Wang et al. 2020). Furthermore, the primary visual cortex, VISp, is thought to be able to modulate contralateral sensitivity for visual detection via ipsilateral corticotectal projections to the SCm (Ruediger and Scanziani 2020). Yet, I find that visual detection sensitivity is affected to a much lesser extent than the bias parameter when inactivating the SC unilaterally. Furthermore, contralateral detection sensitivity is similarly or more affected by inactivating VISp than by inactivating the SCm both during unilateral or bilateral inactivation. I speculate some interpretational discrepancies arise because the previous studies were conducted in Go/NoGo tasks, whereas my task has a 2-AFC design. Indeed, when looking at Go/NoGos, I also find that the fraction of ipsilateral NoGos appear unaffected, while the fraction of contralateral NoGos increases, similarly to previous studies. But due to the nature of the NoGo “choice”, further differences cannot be asserted in the Go/NoGo tasks, whereas I could observe that in fact, ipsilateral choices are sped up, suggesting that SC inactivation does not simply result in the inability to perceive contralateral stimuli. Nevertheless, contralateral visual sensitivity was consistently affected, and bilateral inactivation of the SC affected visual sensitivity more than bilateral inactivation of MOs. Indeed, I speculate that the VISp plays a primary role in setting perceptual sensitivity, while the SCm computes possibly “Go” cues for possible perceptual inputs to provide a landscape of possible actions to select from.

It has never been tested before whether the SCm can manipulate the sensitivity to detecting spatial location of the auditory stimulus to my knowledge. I found that the auditory parameters, particularly the contralateral appear mildly affected, suggesting that the SC can also receive auditory stimulus-related go cues. Importantly, however, performance was not disproportionately affected on coherent trials, consistent with my electrophysiological results that SCm neurons do not specifically bind auditory and visual stimulus encoding. Indeed, the previous lesion study might have found effects specifically on coherent trials (Burnett et al. 2004) because the overall perceptual sensitivity might have decreased due to lesions, masking potential further decreases in orienting accuracy.

Consistent with primary effects on biasing action selection, bilateral SC inactivation appears to restore behaviour. This is like the Sprague effect, which suggest that the effects of primary visual cortex lesion can be diminished by performing a further lesion of the contralateral superior colliculus, or by a collicular commissurotomy (Sprague 1966; Lomber and Payne 1996). Indeed, classical models of superior colliculus function suggest that the two hemispheres compete for potential actions, that

ultimately result in choice through a winner-take-it-all mechanism. In the monkey saccade literature, this has been extended to further sets of actions beyond the two hemisphere, such that even within a hemisphere several action options compete simultaneously until a winner action is selected (McPeek and Keller 2004; Kim and Basso 2010). Consistent with this idea, I also found that the extent of bias effect depends on the exact location of the cannula along the anteroposterior axis of the SC, suggesting that the different competing actions may be topographically represented in the SC.

Terminating the action selection competition may be implemented within the SC, or by an external brain area. Indeed, influential models of perceptual decision-making suggest that the winner-it-all mechanism is implemented via the commissural fibres of the superior colliculus (Machens et al. 2005; Lo and Wang 2006). Further evidence supporting the SC's role in decision termination comes from unilateral SC inactivation of monkeys during a visual decision-making task, which suggests that the SC may implement a fixed bound in the decision-making process (Stine et al. 2023). Contrary to these results, I find that a constant bias injection into the evidence better explains the effects of unilateral inactivation as opposed to a change in a fixed starting point or bound. These results are like other findings in the monkey perceptual decision-making literature (Jun et al. 2021). It is possible that in different task paradigms, the thresholding mechanisms are implemented differently. Indeed, the drift-diffusion model offers an overparameterised model of decision-making, where temporally non-linearly effects on choices may be implemented by changing a constant drift, but similar effects could, in theory, be mimicked by simultaneously changing the noise of the evidence accumulation and the decision bound. It is also possible that (Stine et al. 2023) simply had a smaller tail in the distribution of reaction times and thus they could not observe any further decrease in slow responses.

Bilateral SC inactivation also does not support the view that the SC directly implements the termination procedure. I find that the effects of bilateral inactivation appear additive. This suggests that one SC hemisphere does not boost the other one in a non-linear fashion during the decision-making process, but rather provides weights additively to a threshold mechanism downstream. Indeed, I speculate that the action options compete in the SC until an external area terminates the competition via the winner-take-it-all mechanism. In avians and reptiles, this region is suspected to be the *nucleus isthmi pars magnocellularis* (Mysore and Knudsen 2012; Fenk et al. 2023), the homolog of the parabigeminal nucleus in mice.

Effects on bias have been similarly observed in by the unilateral inactivation of the frontal cortex. Indeed, action selection signals appear in the frontal cortex and the superior colliculus in similar timescales (Steinmetz et al. 2019), and the inactivation of the superior colliculus appears to influence the activity of choice-related neurons in the frontal cortex by increasing activity to promote ipsiversive choices (Thomas et al. 2023). These observations lead to suggestions that the SC and the frontal cortex play a role in action selection in a cooperative fashion.

However, there are key differences between frontal cortex and SC unilateral inactivation. While inactivating both the frontal cortex or the SC unilaterally promotes ipsiversive choices, they have contrasting impact on reaction times and timeouts. Frontal cortex inactivation slows down both ipsiversive and contraversive choices while SC inactivation instead speeds up ipsiversive choices and

slows down contraversive choices. Similarly, timeouts remain unaffected when inactivating the SC ipsilateral to the sensory stimulus, while inactivating the frontal cortex in the same condition results in increased timeouts. These changes in timing are consistent with the spatial distribution of choice representation in these two areas, as in the choice representation in SC appears lateralised, while in frontal cortex the choice representation appears mixed within a single hemisphere (Steinmetz et al. 2019; Coen et al. 2023). It has been suggested that frontal cortex still promotes ipsiversive choices upon unilateral inactivation because pyramidal-tract (PT) cortical neurons preferentially have contraversive choice selectivity, contrary to intratelencephalic (IT) cortical neurons (Li et al. 2015).

While inactivating both the frontal cortex and the SC appears to affect the bias parameter, the effects in SC appear slightly more pronounced. It is possible that this might be simply a result of the difference in inactivation techniques. Indeed, parvalbumin cells are soma targeting inhibitory neurons that affect all responses in a divisive manner (Wilson et al. 2012), while halorhodopsin is thought to affect neurons' resting potentially equally in all compartments, that makes it more likely to have result in additive effects. Alternatively, the SC is downstream the frontal cortex in the decision-making process, and as it accumulates evidence, its effect on slope is smaller as opposed to upstream regions (Koulakov et al. 2005).

Overall, my SC inactivation experiments during the audiovisual task suggest that the primary role of the superior colliculus in a spatial audiovisual decision-making task is to drive action selection, probably in concert with PT neurons of the frontal cortex. It remains unclear whether the frontal cortex computes qualitatively different aspects of the choice that are then passed on to the downstream decision-maker regions, including the superior colliculus, the subcortical regions themselves compute aspects of the choice that can be influenced by frontal cortical input. It is also unclear whether the frontal cortex primarily cooperates with the SC via direct corticotectal projections (Huda et al. 2020; Duan, Pan, et al. 2021), or via the striatal loop (Lee and Sabatini 2021). Indeed, it would be interesting to understand the differential effects of the frontal cortex, the basal ganglia, the thalamic nuclei, and the superior colliculus by joint recordings and inactivation in these regions.

# Chapter 8 — General discussion

The main question this thesis aims to address is how the superior colliculus (SC) contributes to solving an audiovisual localization task. I found that the SC plays a role in audiovisual decision-making by both predicting behavioural choices from neural activity and through inactivation experiments during task performance. This finding aligns with previous research, which identified the SC as a key area for processing auditory and visual signals, integrating audiovisual information, and contributing to perceptual decision-making. In fact, the SC alone might be capable of solving simple localization tasks.

However, I found that in this task, the SC primarily plays a role in forming choices rather than being predominantly multisensory. Consistent with previous studies, I identified both visual and auditory maps in the SC, but these topographic maps didn't strongly interact, and sensory signals were separate from motor and choice signals within single neurons. This suggests that visual, auditory, choice and motor neurons form distinct populations in the SC.

When I inactivated the SC during the audiovisual task, I found that its strongest role was to bias choices independently of sensory stimuli. Inactivating the SC had only a minor effect on visual and auditory sensitivity, implying that the sensory population of the SC might not be actively involved in the task. Interestingly, bilateral inactivation of the SC seemed to counteract the behavioral deficits caused by unilateral inactivation, suggesting that the two SC hemispheres bias choices in a push-pull manner.

Using the Drift Diffusion Model (DDM), I identified that this general bias affects the evidence accumulation over time, rather than creating an initial bias toward a particular choice. This suggests that the SC's role may be to influence which part of space evidence is accumulated from during decision-making.

Why don't the sensory signals seem to participate in the task? The main evidence is two-fold: first, SC neurons can predict choices just before they are made, but not after the stimulus onset; second, inactivating the SC mainly adds a bias but doesn't clearly affect sensory processing. This suggests that sensory effects might be present but are minor. To detect more subtle effects, a larger trial numbers might be needed, such as tracking the same neurons across days as the animal performs the task across several days.

To further confirm the inactivation results, it could also be useful to specifically inactivate visual or auditory neurons. For example, the Ntsr1-Cre line could be used to target and inactivate a subset of visual neurons in the superficial layers (Hoy et al. 2019). However, for auditory neurons, it's less clear how to inactivate a functionally defined set during the task.

It's also possible that a different version of the audiovisual localization task could change the conclusions, suggesting that SC sensory maps are only recruited in certain tasks. For instance, some studies suggest that SC sensory neurons are more responsive to specific stimuli, like looming visual objects or high-frequency sounds, and the stimuli used in our task might be processed by other brain structures.

Another possibility is that the way the decision is executed could affect the way SC is involved in a task. For example, instead of turning a Lego wheel with the forelimbs, if the decision were made by licking left or right (Ito et al. 2024), or by orienting the head and body toward a target (e.g. Duan et al. 2015), SC neurons might be recruited differently during decision-making. In fact, Lee et al. (2022) found that even similar tasks can recruit different sets of neurons in the parietal cortex depending on whether choices are made by navigating in a T-maze or by turning a wheel with the forelimbs.

Alternatively, SC sensory neurons might only participate in non-linear computations, and since our task involved additive processing, audiovisual integratory structures in the SC might not have been engaged. For instance, in non-spatial detection tasks, audiovisual integration can be non-linear and dominated by auditory input (see e.g. Song et al. 2017), which might recruit the SC differently.

Is the SC an audiovisual integratory structure? While this thesis can't speak for all perceptual decisions, my findings suggest that in the specific audiovisual task tested, the SC does not seem predominantly audiovisual. This indicates that the SC may not be a structure capable of flexibly handling various decisions involving audiovisual localization.

In my experiments, visual and auditory stimuli did not appear to combine significantly in individual neurons, and I did not observe non-linear audiovisual integration. It’s possible that audiovisual integration in the SC only occurs under certain stimulus conditions, or perhaps it becomes most evident when stimuli are near the threshold for detection, making neural responses appear non-linear.

A crucial experiment to clarify the nature of audiovisual integration in the SC would involve recording the intracellular voltage of SC neurons while presenting a wide range of auditory and visual stimulus combinations, with precise control over colocalized stimuli. However, this type of experiment has not yet been conducted to my knowledge, likely because it's challenging to perfectly colocalize speakers without affecting visual stimuli and to patch individual neurons to measure voltage on a large scale. However, such experiment might be key in revealing important neural dynamics not visible in spiking data.

The multisensory integrative role of the SC may have been overemphasized in previous research. The extent of multisensory integration is still debated across various brain areas, including the PPC, frontal cortex, and primary sensory cortices. I believe this is due to two main reasons:

First, the definition of sensory integration is unclear. Some researchers view additive tasks, like the one in this thesis, as integrative, while others believe true integration requires the creation of a new percept by such that task can only be solved by the combination of visual and auditory stimuli (to my knowledge such task has not yet been created for rodents). Similarly, at the neural level, some define integration as gain modulation, while others expect neurons to perform complex non-linear operations.

Second, there’s a significant gap between decision-making and audiovisual research. Many signals identified as audiovisual haven’t been tested in a behavioural context, so it’s unclear how animals use these signals for behaviour. Therefore, future studies should investigate audiovisual integration

throughout the brain, including the SC, using tasks with clear quantitative measures of both behaviour and neural activity.

Finally, it's crucial to account for body movements when analyzing neural activity, as demonstrated in Chapter 5. Recent findings show that body movements can significantly impact brain activity, including in the superior colliculus. Some neural activity described as audiovisual might actually be influenced by interactions between body movements, brain state changes, and unisensory stimulus responses (Bimbard et al. 2023). By studying body movements more extensively, we might uncover a map of body parts in the superior colliculus that are differentially activated depending on how choices are made in decision-making tasks.

In this audiovisual localization task, the SC mainly seems to influence how evidence is accumulated. But how is this done? What other brain regions contribute to these biases? How do the dynamics between hemispheres or choice options work? And what is the neural code for choices on a trial-by-trial basis?

Future research could explore these questions in several ways. First, inactivating or recording from other brain regions like the frontal cortex, sensory cortices, SNr, and striatum together could help understand their shared dynamics and individual contributions to decision-making (Erlich et al. 2015; Hanks et al. 2015; Steinmetz et al. 2019; Zatka-Haas et al. 2021; Thomas et al. 2023).

Second, a multi-alternative task with more choice options or switchable stimulus-choice pairs, while recording bilaterally in the SC, could reveal whether SC choice preferences are due to hemispheric lateralization or a more complex role in target selection (McPeek and Keller 2004; Duan, Pagan, et al. 2021).

Finally, chronic recordings that track the same neurons across days could provide better insights into single-trial dynamics during decision-making. Overall, studying the SC in various tasks alongside other brain regions is essential to understand its role in audiovisual integration and decision-making.

# Appendix A — Further methods

| Mouse ID | Line | Sex | Age (weeks) | Experiment | Passive Azimuths | Contrasts (%) | SPLs (dB) | Targeted hemispheres |
|---|---|---|---|---|---|---|---|---|
| **FT008** | C54BL/6J | M | 10-11 | Acute – naïve | 7 | 100 | 50,65 | Both |
| **FT009** | C54BL/6J | M | 11 | Acute – naïve | 7 | 100 | 50,65 | Both |
| **FT010** | C54BL/6J | F | 16 | Acute – naïve | 7 | 20,100 | 65 | Both |
| **FT011** | C54BL/6J | F | 17 | Acute – naïve | 7 | 20, 100 | 65 | Both |
| **FT019** | C54BL/6J | M | 27 | Chronic – naïve | 7 | 25 | 75 | Right |
| **FT022** | C54BL/6J | M | 18 | Chronic – naïve | 7 | 25 | 75 | Right |
| **FT025** | C54BL/6J | M | 10 | Chronic – naïve | 7 | 25 | 75 | Right |
| **FT027** | C54BL/6J | F | 16 | Chronic – naïve | 7 | 25 | 75 | Right |
| **FT030** | Cdh23 | M | 17-19 | Chronic – trained | 3 | 5, 10, 25 | 75 | Right |
| **FT031** | Cdh23 | M | 17-19 | Chronic – trained | 3 | 5, 10, 25 | 75 | Right |
| **FT032** | Cdh23 | M | 22-23 | Chronic – trained | 3 | 5, 10, 25 | 75 | Left |
| **FT035** | Cdh23 | M | 23 | Chronic – trained | 3 | 5, 10, 25 | 75 | Right |
| **FT038** | Chd23 | F | 13-14 | Chronic – naïve | 7 | 40 | 75 | Right |
| **FT039** | Cdh23 | M | 12-14 | Chronic – naïve | 7 | 40 | 75 | Right |
| **AV005** | Cdh23 | M | 31-36 | Chronic – trained | 3 | 5, 10, 25 | 75 | Right |
| **AV008** | Cdh23 | M | 16-22 | Chronic – trained | 3 | 5, 10, 25 | 75 | Both |
| **AV014** | Cdh23 | M | 15-20 | Chronic – trained | 3 | 5, 10, 25 | 75 | Both |
| **AV020** | Cdh23 | M | 23-25 | Chronic – trained | 3 | 5, 10, 25 | 75 | Both |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **AV025** | Cdh23 | F | 23-26 | Chronic – trained | 7 | 10, 20, 40 | 65 | Both |
| **AV029** | Cdh23 | M | 20-22 | Optogenetics | - | 10, 20, 40 | 65 | Right |
| **AV030** | Cdh23 | M | 18-19 | Chronic – trained | 7 | 10, 20, 40 | 65 | Both |
| **AV031** | Cdh23 | M | 14-16 | Optogenetics | - | 10, 20, 40 | 65 | Left |
| **AV033** | Cdh23 | M | 18-19 | Optogenetics | - | 10, 20, 40 | 65 | Both - |
| **AV034** | Cdh23 | F | 18-19 | Chronic – trained | 7 | 10, 20, 40 | 65 | Left |
| **AV036** | Cdh23 | F | 30-34 | Optogenetics | - | 10, 20, 40 | 65 | Both *- |
| **AV038** | Cdh23 | F | 30-34 | Optogenetics | - | 10, 20, 40 | 65 | Both - |
| **AV041** | Cdh23 | M | 23-29 | Optogenetics | - | 10, 20, 40 | 65 | Both |
| **AV044** | Cdh23 | M | 23-38 | Optogenetics | - | 10, 20, 40 | 65 | Both |
| **AV046** | Cdh23 | M | 19-39 | Optogenetics | - | 10, 20, 40 | 65 | Right |
| **AV047** | Cdh23 | M | 28-45 | Optogenetics | - | 10, 20, 40 | 65 | Both |

**TABLE 2 DETAILS OF EXPERIMENTS PERFORMED ON MICE (N=30) THAT CONTRIBUTED TO THIS THESIS.**

*Annotation: M – male, F – female. Age indicates the age in weeks when the animal was recorded or inactivated. Azimuths indicate how many azimuths have been played in the passive experiments in the electrophysiological dataset. Targeted hemispheres indicate whether the electrophysiology recordings or optogenetic implants were in the left or right hemisphere, or both. * Indicates that the initial bilateral implant broke, and the dataset ended up unilateral. - marks indicates that the mouse was bilaterally implanted for optogenetics, but the cannulas were too close to each other, and the patch cords could not be inserted together for bilateral inactivation.*

| Subject | Behaviour + ephys | | Behaviour + unique ephys positions | | Behaviour + unique ephys + camera | | Passive + unique ephys + camera | |
|---|---|---|---|---|---|---|---|---|
| | *Sessions* | *Neurons* | *Sessions* | *Neurons* | *Sessions* | *Neurons* | *Sessions* | *Neurons* |
| **FT030** | 6 | 705 (604) | 1 | 221 (197) | 1 | 62 (60) | 1 | 110 (97) |
| **FT031** | 6 | 289 (76) | 1 | 76 (25) | 1 | 76 (25) | 1 | 76 (25) |
| **FT032** | 6 | 312 (171) | 4 | 201 (96) | 3 | 170 (96) | 4 | 201 (96) |
| **FT035** | 5 | 185 (69) | 4 | 162 (69) | - | - | - | - |
| **AV005** | 16 | 504 (373) | 2 | 40 (14) | 2 | 40 (14) | 5 | 282 (208) |
| **AV008** | 24 | 2850 (1078) | 7 | 650 (252) | 5 | 594 (211) | 8 | 846 (245) |
| **AV014** | 23 | 2593 (1010) | 6 | 237 (71) | 6 | 237 (71) | 8 | 915 (318) |
| **AV020** | 5 | 701 (27) | 1 | 148 (1) | 1 | 148 (1) | 3 | 200 (41) |
| **AV025** | 17 | 1019 (454) | 6 | 287 (119) | 6 | 287 (119) | 6 | 208 (74) |
| **AV030** | 13 | 2395 (427) | 8 | 1519 (157) | 5 | 978 (146) | 4 | 945 (146) |
| **AV034** | 10 | 354 (221) | 5 | 221 (132) | 5 | 208 (132) | 5 | 208 (132) |
| **Total** | **135** | **11907 (4510)** | **45** | **3762 (1127)** | **35** | **2800 (875)** | **45** | **3991 (1382)** |

***TABLE 3. THE NUMBER OF SESSIONS AND NEURONS ASSOCIATED WITH EACH METHOD OF DATA SELECTION IN TRAINED MICE RECORDED WITH CHRONIC NEUROPIXELS PROBES.***

*Data selection was performed according to the type of analysis I performed. The first number in each column indicates the total number, while the number in brackets indicates the numbers from SC. The neurometric logOdds model required good behavioural sessions with electrophysiological (ephys) recordings. The ccCP analysis was performed on each neuron and thus required unique neurons to be selected from chronic recordings. The kernel analysis in Chapter 6, in turn, also required camera data and was also performed on neurons that were recorded in the passive viewing condition after task performance.*

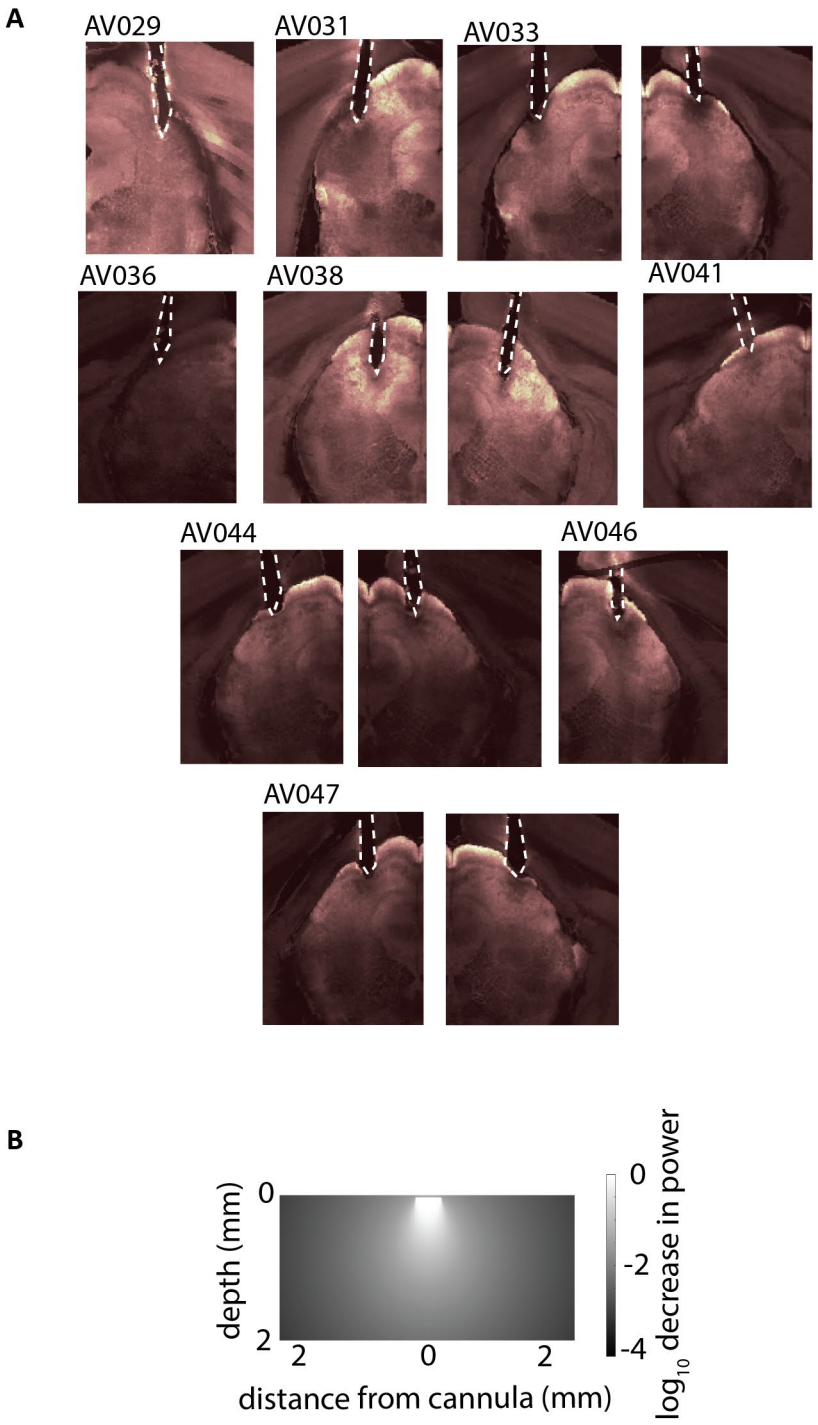
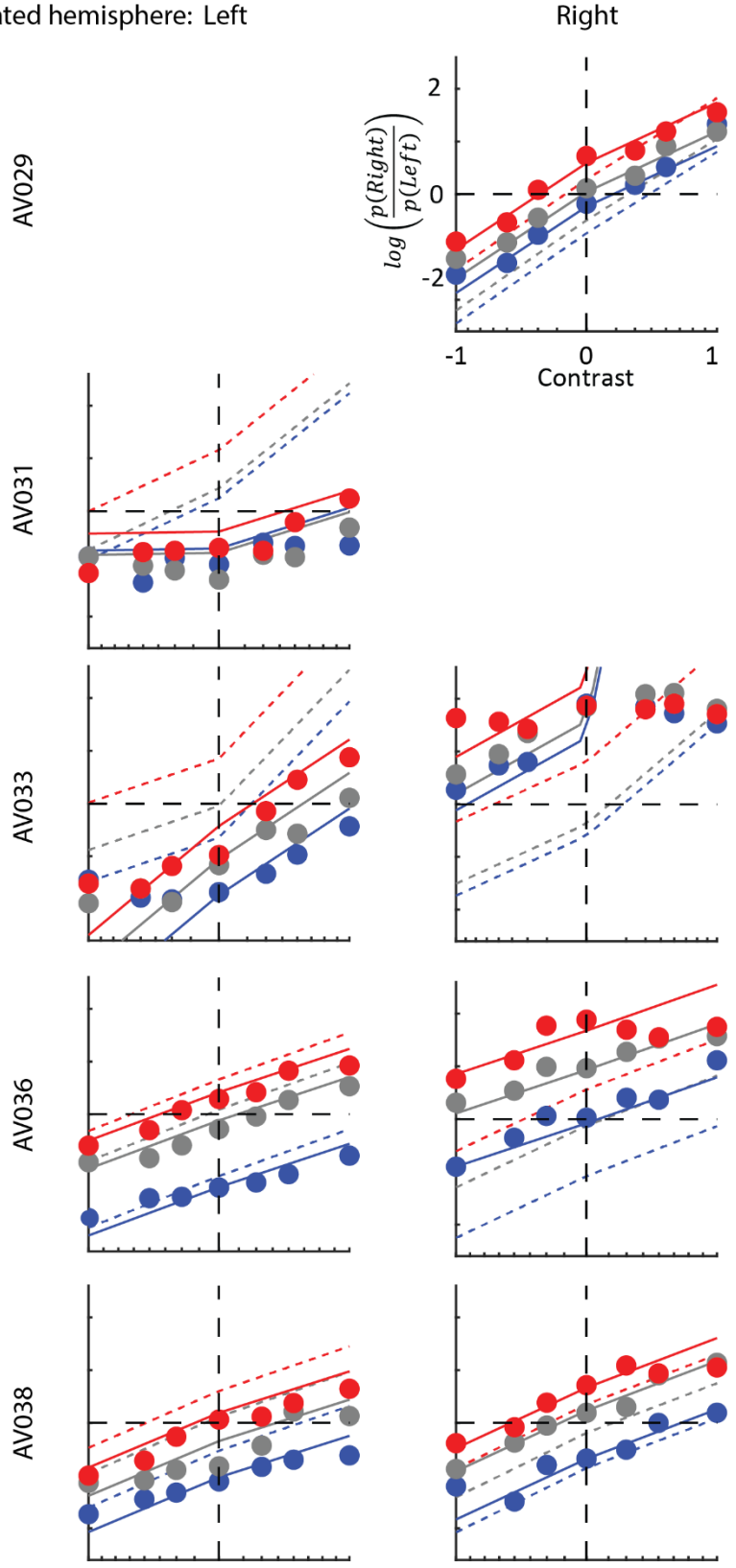



***FIGURE 45. FURTHER METHODS FOR OPTOGENETIC INACTIVATION.***

*(A) Cannula position in the SC in each mouse. (B) Monte Carlo simulation of the light spread from a cannula with a diameter of 0.4 mm using 565nm light.*

# Appendix B — Further results



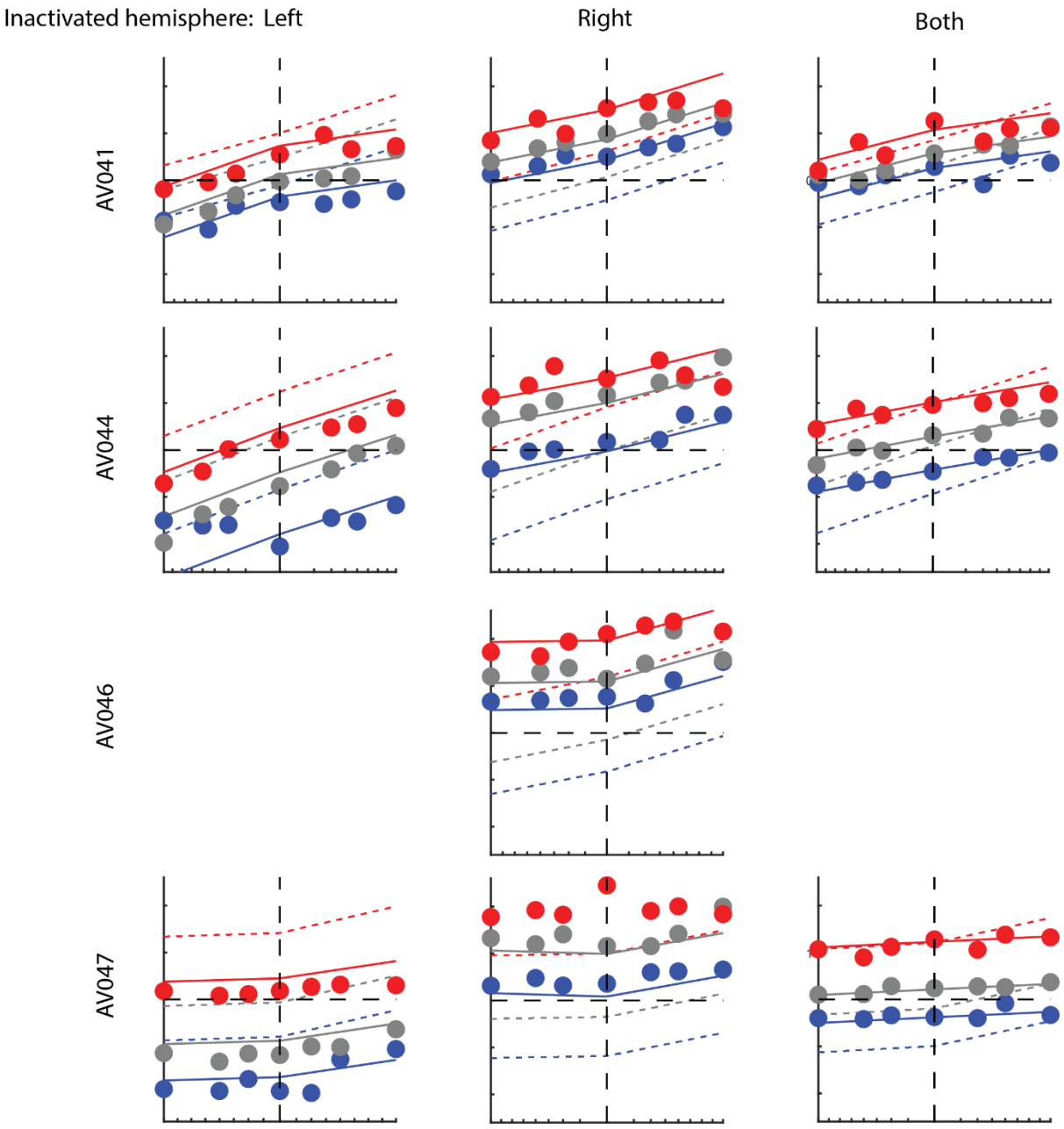
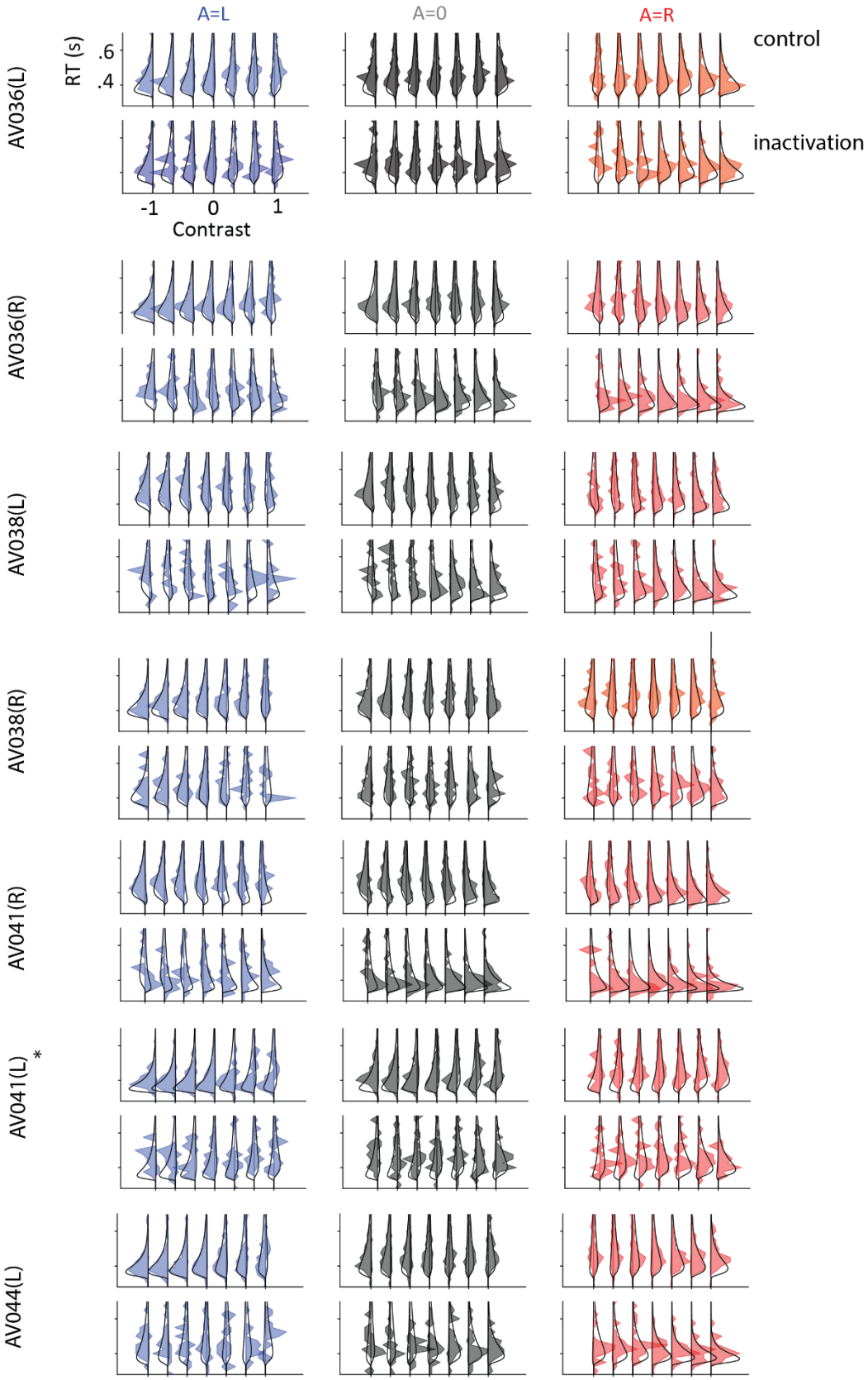



***FIGURE 46. LOG-ODDS MODEL OF INACTIVATION FOR ALL MICE.***

*Each column indicates a different condition for inactivation (left SC, right SC, and both SCs.). Each row indicates a different mouse. Dashed line is the prediction of the Log-Odds model fit to control trials while the solid line is the prediction of inactivation trials. Dots indicate the measured performance on inactivation trials (i.e. annotation is the same as in Figure 37.*


A=L
A=0
A=R
control
inactivation
RT (s)
.6
.4
-1
0
1
Contrast
AV036(L)
AV036(R)
AV038(L)
AV038(R)
AV041(R)
AV041(L) *
AV044(L)

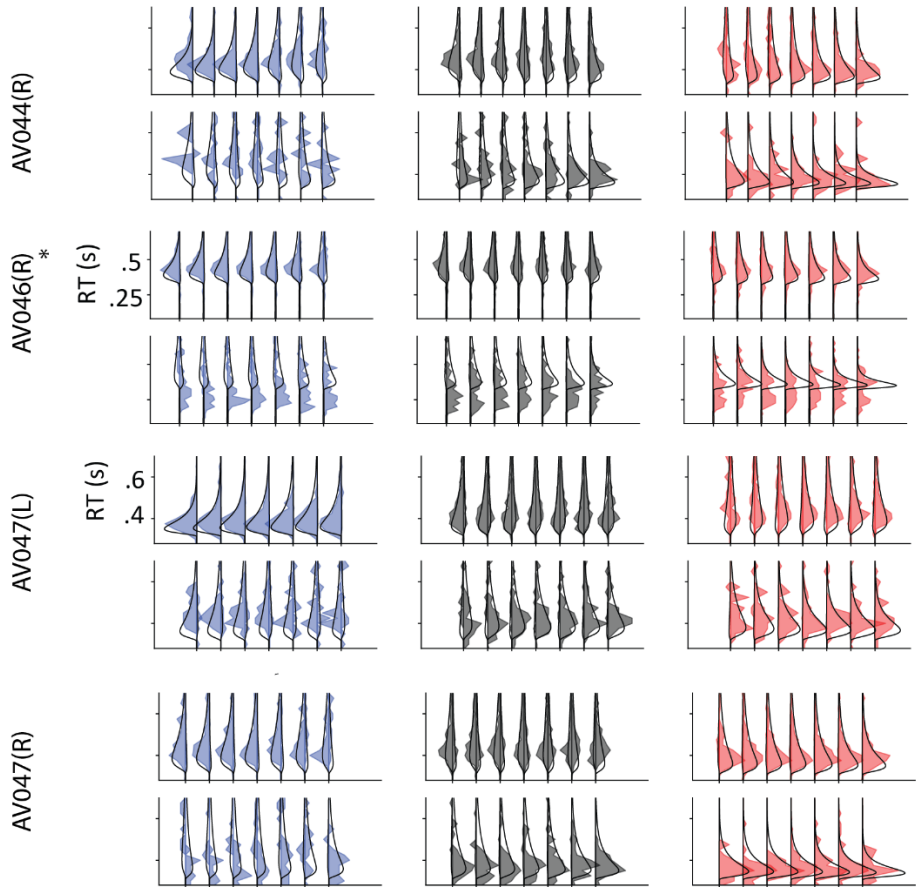



*FIGURE 47. FITS OF THE AUDIOVISUAL DDM WITH CONSTANT DRIFT ON EACH MOUSE.*

# References


Apter, J.T. (1946). Eye movements following the strychninization of the superior colliculus of cats. *Journal of Physiology*, pp.73–87.

Baruchin, L.J., Alleman, M. and Schröder, S. (2023). Reward modulates visual responses in the superficial superior colliculus of mice. *Journal of Neuroscience*, 43(50), pp.8663–8680.

Basso, M.A. and May, P.J. (2017). Circuits for Action and Cognition: A View from the Superior Colliculus. *Annual Review of Vision Science*, 3, pp.197–226.

Van Beest, E.H. et al. (2024). Tracking neurons across days with high-density probes. *bioRXiv*. [online]. Available from: https://doi.org/10.1101/2023.10.12.562040.

Benavidez, N.L. et al. (2021). Organization of the inputs and outputs of the mouse superior colliculus. *Nature Communications*, 12(4004), pp.1–20.

Bimbard, C. et al. (2024). An adaptable, reusable, and light implant for chronic Neuropixels probes. *bioRXiv*. [online]. Available from: https://doi.org/10.1101/2023.08.03.551752.

Bimbard, C. et al. (2023). Behavioral origin of sound-evoked activity in mouse visual cortex. *Nature Neuroscience*, 26(2), pp.251–258.

Brunton, S.L. and Kutz, J.N. (2019). Singular Value Decomposition. In *Data-Driven Science and Engineering, Machine Learning, Dynamical Systems and Control*. pp. 3–46.

Burnett, L.R. et al. (2004). Superior colliculus lesions preferentially disrupt multisensory orientation. *Neuroscience*, pp.535–547.

Busse, L. et al. (2011). The detection of visual contrast in the behaving mouse. *Journal of Neuroscience*, 31(31), pp.11351–11361.

Buzsáki, G. (2019). *The brain from inside out*. Oxford University Press.

Campagner, D. et al. (2023). A cortico-collicular circuit for orienting to shelter during escape. *Nature*, 613(7942), pp.111–119.

Carandini, M. (2024). Sensory choices as logistic classification. *Neuron*, 112, pp.1–15.

Carlini, A., Bordeau, C. and Ambard, M. (2024). Auditory localization: a comprehensive practical review. *Frontiers in Psychology*, 15.

Carr, C.E. and Konishi, M. (1990). A circuit for detection of interaural time differences in the brain stem of the barn owl. *The Journal of Neuroscience*, 70(10), pp.3227–3246.

Chen, S. et al. (2024). Brain-wide neural activity underlying memory-guided movement. *Cell*, 187(3), pp.676-691.e16.

Chen, W. et al. (2022). The roles of optogenetics and technology in neurobiology: a review. *Frontiers in Aging Neuroscience*, 14, p.867863.

Chinta, S. and Pluta, S.R. (2024). Internal monitoring of whisking and locomotion in the superior colliculus. *bioRXiv*. [online]. Available from: https://doi.org/10.1101/2024.03.16.585346.

Claudi, F. et al. (2020). BrainGlobe Atlas API: a common interface for neuroanatomical atlases. *Journal of Open Source Software*, 5(54), p.2668.

Claudi, F. et al. (2021). Visualizing anatomically registered data with brainrender. *eLife*, 10, p.e65751.

Clayton, K.K. et al. (2024). Sound elicits stereotyped facial movements that provide a sensitive index of hearing abilities in mice. *Current Biology*, 34, pp.1605–1620.

Coen, P. et al. (2023). Mouse frontal cortex mediates additive multisensory decisions. *Neuron*, 111(15), pp.2432-2447.e13.

Cooper, B. and Mcpeek, R.M. (2021). Role of the Superior Colliculus in Guiding Movements Not Made by the Eyes. *Annual Review of Vision Science*, 7, pp.279–300. [online]. Available from: https://doi.org/10.1146/annurev-vision-012521-.

Dean, P., Redgrave, P. and Westby, G.W.M. (1988). Event or emergency? Two response systems in the mammalian superior colliculus. *Trends in Neurosciences*, 85, pp.1340–1343.

Drager, U.C. and Hubel, D.H. (1975). Physiology of visual cells in mouse superior colliculus and correlation with somatosensory and auditory input. *Nature*, 253(2), pp.1233–1242.

Duan, C.A., Pan, Y., et al. (2021). A cortico-collicular pathway for motor planning in a memory-dependent perceptual decision task. *Nature Communications*, 12(2727), pp.1–16.

Duan, C.A., Pagan, M., et al. (2021). Collicular circuits for flexible sensorimotor routing. *Nature Neuroscience*, 24(8), pp.1110–1120.

Duan, C.A., Erlich, J.C. and Brody, C.D. (2015). Requirement of prefrontal and midbrain regions for rapid executive control of behavior in the rat. *Neuron*, 86(6), pp.1491–1503.

Edwards, S.B. (1977). The commissural projection of the superior colliculus in the cat. *Journal of Comparative Neurology*, 173(1), pp.23–40.

Erlich, J.C. et al. (2015). Distinct effects of prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat. *eLife*, 2015(4).

Ernst, M.O. and Banks, M.S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415, pp.429–433.

Essig, J., Hunt, J.B. and Felsen, G. (2021). Inhibitory neurons in the superior colliculus mediate selection of spatially-directed movements. *Communications Biology*, 4(719), pp.1–14.

Evans, D.A. et al. (2018). A synaptic threshold mechanism for computing escape decisions. *Nature*, 558(7711), pp.590–594.

Fabre, J.M.J.. et al. (2023). Bombcell: automated curation and cell classification of spike-sorted electrophysiology data. *Zenodo*. [online]. Available from: https://doi.org/10.5281/zenodo.8172821.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), pp.861–874.

Fechner, G. (1870). *Elements of Psychophysics*. Holt, Rinehart and Winston.

Felsen, G. and Mainen, Z.F. (2012). Midbrain contributions to sensorimotor decision making. *Journal of Nuerophysiology*, 108, pp.135–147.

Felsen, G. and Mainen, Z.F. (2008). Neural substrates of sensory-guided locomotor decisions in the rat superior colliculus. *Neuron*, 60(1), pp.137–148.

Fenk, L.A., Riquelme, J.L. and Laurent, G. (2023). Interhemispheric competition during sleep. *Nature*, 616(7956), pp.312–318.

Fetsch, C.R., Deangelis, G.C. and Angelaki, D.E. (2013). Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nature Reviews Neuroscience*, 14(6), pp.429–442.

Findling, C. et al. (2023). Brain-wide representations of prior information in mouse decision-making. *bioRXiv*. [online]. Available from: https://doi.org/10.1101/2023.07.04.547684.

De Franceschi, G. and Solomon, S.G. (2018). Visual response properties of neurons in the superficial layers of the superior colliculus of awake mouse. *Journal of Physiology*, 596(24), pp.6307–6332.

Fusi, S., Miller, E.K. and Rigotti, M. (2016). Why neurons mix: High dimensionality for higher cognition. *Current Opinion in Neurobiology*, 37, pp.66–74.

Gallivan, J.P. et al. (2018). Decision-making in sensorimotor control. *Nature Reviews Neuroscience*, 19(9), pp.519–534.

Gandhi, N.J. and Katnani, H.A. (2011). Motor functions of the superior colliculus. *Annual Review of Neuroscience*, 34, pp.205–231.

Gharaei, S., Arabzadeh, E. and Solomon, S.G. (2018). Integration of visual and whisker signals in rat superior colliculus. *Scientific Reports*, 8(1), p.16445.

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

Gold, J.I. and Shadlen, M.N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, 30, pp.535–574.

Goldberg, M.E. and Wurtz, R.H. (1972). Activity of superior colliculus in behaving monkey. I. Visual receptive fields of single neurons. *Journal of physiology*, pp.525–543.

Gradinaru, V. et al. (2010). Molecular and cellular approaches for diversifying and extending optogenetics. *Cell*, 141(1), pp.154–165.

Green, D.M. and Swets, J.A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley & Sons, Inc.

Gu, Y., Angelaki, D.E. and DeAngelis, G.C. (2008). Neural correlates of multisensory cue integration in macaque MSTd. *Nature Neuroscience*, 11(10), pp.1201–1210.

Guillamón-Vivancos, T. et al. (2022). Input-dependent segregation of visual and somatosensory circuits in the mouse superior colliculus. *Science*, 377(6608), pp.845–850.

Handler, A. and Ginty, D.D. (2021). The mechanosensory neurons of touch and their mechanisms of activation. *Nature Reviews Neuroscience*, 22(9), pp.521–537.

Hanks, T.D. et al. (2015). Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature*, 520(7546), pp.220–223.

Hanks, T.D. and Summerfield, C. (2017). Perceptual Decision Making in Rodents, Monkeys, and Humans. *Neuron*, 93(1), pp.15–31.

Harmening, W.M. and Wagner, H. (2011). From optics to attention: Visual perception in barn owls. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 197(11), pp.1031–1042.

Harris, K.D. (2020). Nonsense correlations in neuroscience. *BioRXiv*. [online]. Available from: https://doi.org/10.1101/2020.11.29.402719.

Harris, L.R., Blakemore, C. and Donaghy, M. (1980). Integration of visual and auditory space in the mammalian superior colliculus. *Nature*, 288(5786), pp.56–59.

Hasnain, M.A. et al. (2023). Separating cognitive and motor processes in the behaving mouse. *BioRXiv*. [online]. Available from: https://doi.org/10.1101/2023.08.23.554474.

Hoy, J.L., Bishop, H.I. and Niell, C.M. (2019). Defined cell types in superior colliculus make distinct contributions to prey capture behavior in the mouse. *Current Biology*, 29(23), pp.4130-4138.e5.

Hubel, D.H. and Wiesel, A.T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160, pp.106–154.

Huda, R. et al. (2020). Distinct prefrontal top-down circuits differentially modulate sensorimotor behavior. *Nature Communications*, 11(1), p.6007.

Isa, K. et al. (2020). Dissecting the tectal output channels for orienting and defense responses. *eNeuro*, 7(5), pp.1–18.

Isa, T. et al. (2021). The tectum/superior colliculus as the vertebrate solution for spatial sensory integration and action. *Current Biology*, 31(11), pp.R741–R762.

Isa, T. and Hall, W.C. (2009). Exploring the superior colliculus in vitro. *Journal of Neurophysiology*, 102(5), pp.2581–2593.

Ito, B.S. et al. (2024). A collicular map for touch-guided tongue control. In *bioRXiv*. [online]. Available from: https://doi.org/10.1101/2024.04.08.587629.

Ito, S. et al. (2021). Nonlinear visuoauditory integration in the mouse superior colliculus T. Yang, ed. *PLOS Computational Biology*, 17(11), p.e1009181.

Ito, S. et al. (2020). Spectral cues are necessary to encode azimuthal auditory space in the mouse superior colliculus. *Nature Communications*, 11, p.1087.

Ito, S. and Feldheim, D.A. (2018). The mouse superior colliculus: An emerging model for studying circuit formation and function. *Frontiers in Neural Circuits*, 12.

Jun, E.J. et al. (2021). Causal role for the primate superior colliculus in the computation of evidence for perceptual decisions. *Nature Neuroscience*, 24(8), pp.1121–1131.

Jun, J.J. et al. (2017). Fully integrated silicon probes for high-density recording of neural activity. *Nature*, 551(7679), pp.232–236.

Katz, L.N. et al. (2016). Dissociated functional significance of decision-related activity in the primate dorsal stream. *Nature*, 535(7611), pp.285–288.

Kim, B. and Basso, M.A. (2010). A probabilistic strategy for understanding action selection. *Journal of Neuroscience*, 30(6), pp.2340–2355.

King, A.J. (2002). Neural plasticity: how the eye tells the brain about sound location dispatch. *Current Biology*, 12, pp.393–395.

King, A.J. and Hutchings, M.E. (1987). Spatial response properties of acoustically responsive neurons in the superior colliculus of the ferret: a map of auditory space. *Journal of Neurophysiology*, 57(2), pp.596–624.

King, A.J. and Palmer, A.R. (1983). Cells responsive to free-field auditory stimuli in guinea-pig superior colliculus: distribution and response properties. *J. Phygiol*, 342, pp.361–381.

King, A.J. and Palmer, A.R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Exp Brain Res*, 60, pp.492–500.

Klausberger, T. et al. (2003). Brain-state-and cell-type-specific firing of hippocampal interneurons in vivo. *Nature*, 421, pp.844–848.

Knill, D.C. and Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12), pp.712–719.

Knudsen, E.I. and Brainard, M.S. (1991). Visual instruction of the neural map of auditory space in the developing optic tectum. *Science*, 253, pp.85–87.

Knudsen, E.I. and Knudsen, P.F. (1989). Vision calibrates sound localization in developing barn owls. *The Journal of Neuroscience*, 9(9), pp.3306–3313.

Koulakov, A.A., Rinberg, D.A. and Tsigankov, D.N. (2005). How to find decision makers in neural networks. *Biological Cybernetics*, 93(6), pp.447–462.

Kramis, R., Vanderwolf, C.H. and Bland, B.H. (1975). Two types of hippocampal rhythmical slow activity in both the rabbit and the rat: relations to behavior and effects of atropine, diethyl ether, urethane, and pentobarbital. *Experimental Neurology*, 49, pp.58–85.

Krauzlis, R.J., Lovejoy, L.P. and Zénon, A. (2013). Superior colliculus and visual spatial attention. *Annual Review of Neuroscience*, 36, pp.165–182.

Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), pp.84–90.

Lak, A. et al. (2020). Dopaminergic and prefrontal basis of learning from sensory confidence and reward value. *Neuron*, 105(4), pp.700-711.e6.

Lee, J. and Sabatini, B.L. (2021). Striatal indirect pathway mediates exploration via collicular competition. *Nature*, 599(7886), pp.645–649.

Li, N. et al. (2015). A motor cortex circuit for motor planning and movement. *Nature*, 519(7541), pp.51–56.

Li, Y.T. and Meister, M. (2023). Functional cell types in the mouse superior colliculus. *eLife*, 12, p.e82367.

Lo, C.-C. and Wang, X.-J. (2006). Cortico–basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nature Neuroscience*, 9(7), pp.956–963.

Lomber, S.G. and Payne, B.R. (1996). Removal of two halves restores the whole: Reversal of visual hemineglect during bilateral cortical or collicular inactivation in the cat. *Visual Neuroscience*, 13(6), pp.1143–1156.

Lovejoy, L.P. and Krauzlis, R.J. (2010). Inactivation of primate superior colliculus impairs covert selection of signals for perceptual judgments. *Nature Neuroscience*, 13(2), pp.261–266.

Ma, W.J. et al. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience*, 9(11), pp.1432–1438.

Machens, C.K., Romo, R. and Brody, C.D. (2005). Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science*, 307(5712), pp.1118–1121.

Mason, S.J. and Graham, N.E. (2002). Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. *Quarterly Journal of the Royal Meteorological Society*, 128(584 PART B), pp.2145–2166.

Masullo, L. et al. (2019). Genetically defined functional modules for spatial orienting in the mouse superior colliculus. *Current Biology*, 29(17), pp.2892-2904.e8.

Mathis, A. et al. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, 21(9), pp.1281–1289.

Mattis, J. et al. (2012). Principles for applying optogenetic tools derived from direct comparative analysis of microbial opsins. *Nature Methods*, 9(2), pp.159–172.

Mayrhofer, J.M. et al. (2019). Distinct Contributions of Whisker Sensory Cortex and Tongue-Jaw Motor Cortex in a Goal-Directed Sensorimotor Transformation. *Neuron*, 103(6), pp.1034-1043.e5.

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

McGurk, H. and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, pp.746–748.

McPeek, R.M. and Keller, E.L. (2004). Deficits in saccade target selection after inactivation of superior colliculus. *Nature Neuroscience*, 7(7), pp.757–763.

Meredith, M.A. and Stein, B.E. (1983). Interactions among converging sensory input in the superior colliculus. *Science*, 221, pp.389–391.

Mianné, J. et al. (2016). Correction of the auditory phenotype in C57BL/6N mice via CRISPR/Cas9-mediated homology directed repair. *Genome Medicine*, 15(1), pp.8–16.

Middlebrooks, J.C. and Knudsen, E.I. (1984). A neural code for auditory space in the cat's superior colliculus. *The Journal of Neuroscience*, 4(10), pp.2621–2634.

Mohler, C.W. and Wurtz, R.H. (1976). Organization of monkey superior colliculus: intermediate layer cells discharging before eye movements. *Journal of Neurophysiology*, 39(4), pp.722–744.

Musall, S. et al. (2019). Single-trial neural dynamics are dominated by richly varied movements. *Nature Neuroscience*, 22(10), pp.1677–1686.

Mysore, S.P. and Knudsen, E.I. (2012). Reciprocal Inhibition of Inhibition: A Circuit Motif for Flexible Categorization in Stimulus Selection. *Neuron*, 73(1), pp.193–205.

Newman, E.A. and Hartline, P.H. (1981). Integration of visual and infrared information in bimodal neurons of the rattlesnake optic tectum. *Science*, 213, pp.789–791.

Niedworok, C.J. et al. (2016). aMAP is a validated pipeline for registration and segmentation of high-resolution mouse brain data. *Nature Communications 2016 7:1*, 7(1), pp.1–9.

Niell, C.M. and Stryker, M.P. (2010). Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex. *Neuron*, 65(4), pp.472–479.

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

Okun, M. et al. (2016). Long term recordings with immobile silicon probes in the mouse cortex. *PLoS ONE*, 11(3), p.e0151180.

Ottenheimer, D.J. et al. (2023). A stable, distributed code for cue value in mouse cortex during reward learning. *eLife*, 12:RP84604, pp.1–30.

Oude Lohuis, M.N. et al. (2024). Triple dissociation of visual, auditory and motor processing in mouse primary visual cortex. *Nature Neuroscience*, 27, pp.758–771.

Pachitariu, M. et al. (2024). Spike sorting with Kilosort4. *Nature Methods*, 8, pp.1–8.

Palmer, A.R. and King, A.J. (1982). The representation of auditory space in the mammalian superior colliculus. *Nature*, 299, pp.248–249.

Parham, K. and Willott, J.F. (1988). Behavioral Neuroscience Acoustic Startle Response in Young and Aging C57BL/6J and CBA/J Mice. *Behavioral Neuroscience*, 102(6), pp.881–886.

Park, I.M. et al. (2014). Encoding and decoding in parietal cortex during sensorimotor decision-making. *Nature Neuroscience*, 17(10), pp.1395–1403.

Pedregosa, F. et al. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, pp.2825–2830. [online]. Available from: http://scikit-learn.sourceforge.net.

Pitts, W. and McCullogh, W.S. (1947). How we know universals: the perceptions of auditory and visual forms. *Bulletin of Mathematical Biophysics*, 9, pp.127–147.

Populin, L.C. and Yin, T.C.T. (2002). Bimodal interactions in the superior colliculus of the behaving cat. *Journal of Neuroscience*, 22(7), pp.2826–2834.

Ragan, T. et al. (2012). Serial two-photon tomography for automated ex vivo mouse brain imaging. *Nature Methods 2012 9:3*, 9(3), pp.255–258.

Raposo, D., Kaufman, M.T. and Churchland, A.K. (2014). A category-free neural population supports evolving demands during decision-making. *Nature Neuroscience*, 17(12), pp.1784–1792.

Ratcliff, R. (2006). Modeling response signal and response time data. *Cognitive Psychology*, 53(3), pp.195–237.

Ratcliff, R. and Mckoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, 20(4), pp.873–922.

Rigotti, M. et al. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature 2013 497:7451*, 497(7451), pp.585–590.

Robinson, D.A. (1972). Eye movements evoked by collicular stimulation in the alert monkey. *Vision Research*, 12, pp.1795–1808.

Roitman, J.D. and Shadlen, M.N. (2002). Response of Neurons in the Lateral Intraparietal Area during a Combined Visual Discrimination Reaction Time Task. *The Journal of Neuroscience*, 22(21), pp.9475–9489.

Rowland, B.A. et al. (2007). Multisensory integration shortens physiological response latencies. *Journal of Neuroscience*, 27(22), pp.5879–5884.

Roy, N.A. et al. (2018). Efficient inference for time-varying behavior during learning. *Advances in Neural Information Processing Systems*, 31, pp.5695–5705.

Ruediger, S. and Scanziani, M. (2020). Learning speed and detection sensitivity controlled by distinct cortico-fugal neurons in visual cortex. *eLife*, 9, p.e59247.

Savier, E.L., Chen, H. and Cang, J. (2019). Effects of locomotion on visual responses in the mouse superior colliculus. *Journal of Neuroscience*, 39(47), pp.9360–9368.

Schiller, P.H. and Stkyker, M. (1972). Single-Unit Recording and Stimulation in Superior Colliculus of the Alert Rhesus Monkey. *Journal of Physiology*, 35(6), pp.915–924.

Schiller, P.H. and Stryker, M. (1972). Single-unit recording and stimulation in the superior colliculus of the alert rhesus monkey. *Journal of neurophysiology*, 35(6), pp.915–925.

Schröder, S. et al. (2020). Arousal modulates retinal output. *Neuron*, 107(3), pp.487-495.e9.

Selen, L.P.J., Shadlen, M.N. and Wolpert, D.M. (2012). Deliberation in the motor system: Reflex gains track evolving evidence leading to a decision. *Journal of Neuroscience*, 32(7), pp.2276–2286.

Shang, C. et al. (2019). A subcortical excitatory circuit for sensory-triggered predatory hunting in mice. *Nature Neuroscience*, 22(6), pp.909–920.

Sherrington, C.S. (1910). Flexion-reflex of the limb, crossed extension-reflex, and reflex stepping and standing. *Journal of Physiology*, 40, pp.28–121.

Shinn, M., Lam, N.H. and Murray, J.D. (2020). A flexible framework for simulating and fitting generalized drift-diffusion models. *eLife*, 9, pp.1–27.

Si, Y. et al. (2022). High-frequency hearing is required to compute a topographic map of auditory space in the mouse superior colliculus. *eNeuro*, 9(3), pp.0513–0521.

Siegle, J.H. et al. (2021). Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature*, 592(7852), pp.86–92.

Socha, K. et al. (2018). Behavioral response to visual motion impacts population coding in the mouse visual thalamus. *bioRXiv*. [online]. Available from: https://doi.org/10.1101/382671.

Song, Y.H. et al. (2017). A Neural Circuit for Auditory Dominance over Visual Perception. *Neuron*, 93(4), pp.940-954.e6.

Sprague, J. (1966). Interaction of cortex and superior colliculus in mediation of visually guided behavior in the cat. *Science*, 153(3743), pp.1544–1547.

Stanford, T.R., Quessy, S. and Stein, B.E. (2005). Evaluating the operations underlying multisensory integration in the cat superior colliculus. *Journal of Neuroscience*, 25(28), pp.6499–6508.

Stein, B.E. and Meredith, M.A. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science*, 369(July), pp.1981–1983.

Steinmetz, N.A. et al. (2019). Distributed coding of choice, action and engagement across the mouse brain. *Nature*, 576(7786), pp.266–273.

Steinmetz, N.A. et al. (2021). Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. *Science*, 372(6539).

Stine, G.M. et al. (2023). A neural mechanism for terminating decisions. *Neuron*, 111(16), pp.2601-2613.e5.

Stringer, C. et al. (2019). Spontaneous behaviors drive multidimensional, brainwide activity. *Science*, 364(6437).

Stujenske, J.M., Spellman, T. and Gordon, J.A. (2015). Modeling the Spatiotemporal Dynamics of Light and Heat Propagation for InVivo Optogenetics. *Cell Reports*, 12(3), pp.525–534.

Summerfield, C. and Tsetsos, K. (2012). Building bridges between perceptual and economic decision-making: Neural and computational mechanisms. *Frontiers in Neuroscience*, 6, pp.1–20.

The International Brain Laboratory et al. (2021). Standardized and reproducible measurement of decision-making in mice. *eLife*, 10, p.e63711.

Thomas, A. et al. (2023). Superior colliculus bidirectionally modulates choice activity in frontal cortex. *Nature Communications*, 14(1), p.7358.

Tyson, A.L. et al. (2022). Accurate determination of marker location within whole-brain microscopy images. *Scientific Reports 2022 12:1*, 12(1), pp.1–8.

Urai, A.E. et al. (2019). Choice history biases subsequent evidence accumulation. *eLife*, 2(8), p.e46331.

Vinck, M. et al. (2015). Arousal and Locomotion Make Distinct Contributions to Cortical Activity Patterns and Visual Encoding. *Neuron*, 86(3), pp.740–754.

Virtanen, P. et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3), pp.261–272.

Wang, L. et al. (2020). A causal role for mouse superior colliculus in visual perceptual decision-making. *Journal of Neuroscience*, 40(19), pp.3768–3782.

Wang, Q. et al. (2020). The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. *Cell*, 181(4), pp.936-953.e20.

Wang, Q. and Burkhalter, A. (2013). Stream-related preferences of inputs to the superior colliculus from areas of dorsal and ventral streams of mouse visual cortex. *Journal of Neuroscience*, 33(4), pp.1696–1705.

Wang, Z.A. et al. (2023). Not everything, not everywhere, not all at once: a study of brain-wide encoding of movement. *bioRXiv*. [online]. Available from: https://doi.org/10.1101/2023.06.08.544257.

Wheatcroft, T., Saleem, A.B. and Solomon, S.G. (2022). Functional Organisation of the Mouse Superior Colliculus. *Frontiers in Neural Circuits*, 16, p.792959.

White, C.N. and Poldrack, R.A. (2014). Decomposing bias in different types of simple decisions. *Journal of Experimental Psychology: Learning Memory and Cognition*, 40(2), pp.385–398.

Whittington, J.C.R. et al. (2022). Disentanglement with biological constraints: a theory of functional cell types. *International Conference on Learning Representations*. [online]. Available from: https://arxiv.org/abs/2210.01768v2.

Williams, A.M., Angeloni, C.F. and Geffen, M.N. (2023). Sound improves neuronal encoding of visual stimuli in mouse primary visual cortex. *Journal of Neuroscience*, 43(16), pp.2885–2906.

Wilson, N.R. et al. (2012). Division and subtraction by distinct cortical inhibitory networks in vivo. *Nature*, 488(7411), pp.343–348.

Wu, C. et al. (2015). Listening to another sense: somatosensory integration in the auditory system. *Cell and Tissue Research*, 361(1), pp.233–250.

Wurtz, R.H. and Goldberg, M.E. (1972). Activity of superior colliculus in behaving monkey. III. Cells discharging before eye movements. *Journal of physiology*, pp.575–585.

Yang, Y. and Zador, A.M. (2012). Differences in sensitivity to neural timing among cortical areas. *Journal of Neuroscience*, 32(43), pp.15142–15147.

Yin, C. et al. (2023). Engaged decision-makers align spontaneous movements to stereotyped task demands. *BioRxiv*. [online]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/37425720.

Yin Zheng, Q., Johnson, K.R. and Erway, L.C. (1999). Assessment of hearing in 80 inbred strains of mice by ABR threshold analyses. *Hearing Research*, 120, pp.94–107.

Zagha, E. et al. (2022). The importance of accounting for movement when relating neuronal activity to sensory and cognitive processes. *Journal of Neuroscience*, 42(8), pp.1375–1382.

Zahar, Y., Reches, A. and Gutfreund, Y. (2009). Multisensory enhancement in the optic tectum of the barn owl: Spike count and spike timing. *Journal of Neurophysiology*, 101(5), pp.2380–2394.

Zatka-Haas, P. et al. (2021). Sensory coding and the causal impact of mouse cortex in a visual decision. *eLife*, 10.

Zhao, X., Liu, M. and Cang, J. (2014). Visual cortex modulates the magnitude but not the selectivity of looming-evoked responses in the superior colliculus of awake mice. *Neuron*, 84(1), pp.202–213.

Zingg, B. et al. (2017). AAV-Mediated Anterograde Transsynaptic Tagging: Mapping Corticocollicular Input-Defined Neural Pathways for Defense Behaviors. *Neuron*, 93(1), pp.33–47.