# The neural representation of multisensory percepts:

# Studying the integration of auditory and tactile stimuli in the somatosensory cortex




Séverin Limal

Brasenose College

University of Oxford

A thesis submitted for the degree of

*Doctor of Philosophy*

Trinity Term 2020

*Je dédie cette thèse à Papa Gilles,*

*et à Rose et Félicie.*

When I applied for a DPhil at the University of Oxford, I was warned that a good relationship with your supervisor and lab are sometimes even more important than the project itself. This does not cover the half of it. This Thesis would not be anywhere without the patience, support and ultimately friendship of not one, nor two but three fantastic mentors: Prof Kerry Walker, Dr Michael Kohl and Dr Mariangela Panniello.

Michael, thank you for the hours spent in the lab building, assembling, imaging, analysing and so many other -ings. Thank you for taking a chance on a fresh-faced student and encouraging me to challenge established methods and thinking outside the box. The trust, respect and friendship you give to the people you work with give me faith in the future of academic research.

Kerry, you made me want to return to Oxford in the first place and, while you could have done with 30 h days, you found the time to be present for my doubts, my questions, and my learning. You give another meaning to kindness and I look forward to still working together in the future.

A very special mention must be made for Mari. I first learned under your supervision when you were yourself a DPhil student, and you are the reason I wanted to stay in research. Your dedication, help and skills have taught me everything, and I will be forever grateful for our friendship, both through happiness and hardship.

I am grateful to my supervisor Prof Andrew King, for having enabled my research and supported my work through the years.

I am thankful for my college, Brasenose College for providing a tight-knit community and collegiality outside of the lab.

I feel amazingly lucky having made such amazing friends in the lab. Ana Bottura de Barros, you were here at the very beginning of the lab and it was a joy and heartwarming pleasure to spend these long days in the office with you. Liad Baruchin, thank you for your humour, your patience and support and thank you for being a great desk neighbour. Cristiana 'Kiki' Vagnoni, thank you for our coffee breaks, your laughter, your craziness too, sometimes! Yayoi Teramoto, your grounding reassurance and thoughtfulness carried me through hard days and doubts, thank you for being there. I would also like to thank past and present members of the Kohl and Walker labs: Suzanne, Inês, Sebastian, Jimmy, Tai, Filippo, Sara, Alex, Veronika, and so many others.

Thank you to my lifelong friends Alexandre and Jayrane for being such strong supports, lifelines and sources of joy throughout my whole studies. I am so lucky to be back in Paris with you.

I am grateful for the friends I made during my time at Oxford including Shyam, Harvey, Bradley, Sam, Aneka and Kristin but you know who you are! Thank you for making these years as rich and enjoyable as they were, I am lucky to count you among my friends. I am equally fortunate to have been supported by an amazing support network, both in France and the UK. Thank you to roof boiz for encouraging me to do this DPhil in the first place, Valentin, Myriam, Lucille, Clara and Nicola for the long-distance support and constant reminder of our friendship.

Thank you to my girlfriend and best friend, Rebecca. Thank you for enduring this DPhil alongside me, for getting up at ungodly hours because I had to go to the lab, for short-notice change of plans because of experiments gone wrong, for the relief and happiness in and out of Oxford. Thank you for always believing in me and encouraging me to pursue my ambitions.

Thank you to my family, for supporting me as aces (well, as jacteas), visiting me when you could, reminding me of the life outside of the bubble. Thank you for being you and having helped me grow and attempt to become half as good as you. To my parents, thank you for teaching me to let go, breathe, and accept that if I give my best, I will obtain my best. Thank you to all of Maman Ione's family, and to my grandparents Maïka, Dad and Papa Gilles - I was so lucky to spend my childhood with you around.

Thank you to Rose and Félicie for bringing such joy as you joined the ride along the way.

# Abstract

Multisensory integration is the constant process through which we combine information across different sensory modalities in order to better interpret our environment. In early models of multisensory integration, primary sensory areas were thought to almost exclusively treat information from their preferred sensory modality and contribute little to the integration process. However, a growing literature now describes cross-modal responses in primary sensory cortices and direct cortico-cortical connections between primary sensory cortices. Therefore, cortical multisensory integration may begin as early as the primary sensory regions.

I built on this literature by examining how training on a multisensory integration task may alter vibrissal primary somatosensory cortex (vS1) responses to vibrotactile whisker stimulation and amplitude modulated pure tones. Unlike most previous studies, ours required the animal to identify a unique combination of vibrotactile and auditory stimulation in order to perform the task correctly. I used chronic *in vivo* two-photon imaging of vS1 to measure neural responses to stimuli throughout training on the Go/No-Go task. Responses of the same neural populations were compared prior to, during, and following task acquisition.

Mice learned to correctly identify specific combinations of audio-tactile stimulation. I observed vS1 responses to vibrotactile stimuli and found a subset of neurons that were responsive to sounds. I also found neurons that responded only to a single multisensory combination. Training led to an increase in the proportion of responsive neurons during acquisition on all tasks and a decrease in the proportion of responsive neurons after training on the multisensory task.

Most tracked neurons became either responsive or unresponsive during or after training. In contrast, very few neurons maintained constant response properties. While it remains unclear whether or how training may remodel local circuits, my findings suggest that training and experience lead to the recruitment of different populations of neurons in vS1 in a modality-specific manner.

# Contents


**1 General Introduction** **1**

1.1 Behaving in a multisensory world . . . . . 1

1.1.1 Sensory dominance and multisensory illusions . . . . . 3

1.1.2 Processing tactile input from the whiskers . . . . . 6

1.1.3 Auditory perception in mice . . . . . 7

1.1.4 Multisensory perception in mice . . . . . 7

1.2 Measuring neural correlates of perception . . . . . 10

1.2.1 Neural substrates of tactile perception . . . . . 12

1.2.2 The barrel cortex is an inherently sensorimotor area . . . . . 17

1.2.3 Neural substrates of auditory perception . . . . . 19

1.3 Neural substrates of multisensory integration . . . . . 21

1.3.1 Subcortical responses to multisensory stimuli . . . . . 21

1.3.2 Cortical multisensory integration . . . . . 23

1.3.3 Multisensory integration in primary sensory cortices . . . . . 24

1.3.4 Circuits underlying multisensory integration . . . . . 27

1.3.5 Effects of training on neural responses . . . . . 30

1.4 General aims . . . . . 32

1.5 General hypotheses . . . . . 33

**2 Performance of mice on a novel multisensory learning task** **36**

2.1 Introduction . . . . . 36

2.1.1 Overview . . . . . 36

2.1.2 Objectives . . . 42
2.2 Methods . . . 42
2.2.1 Subjects . . . 42
2.2.2 Surgical procedures . . . 43
2.2.3 Behavioural habituation phase . . . 44
2.2.4 Reward-stimulus association phase . . . 45
2.2.5 Training phase . . . 45
2.2.6 Pole position stimuli . . . 47
2.2.7 Vibrotactile stimuli . . . 48
2.2.8 Auditory stimuli . . . 48
2.2.9 Multisensory discrimination stimuli . . . 49
2.2.10 Controlling for cross-modal interference . . . 51
2.2.11 Analysing behavioural data . . . 52
2.2.12 Two-photon imaging . . . 54
2.2.13 Analysing imaging data . . . 55
2.3 Results . . . 56
2.3.1 Pole position training . . . 56
2.3.2 Imaging with pole stimulus . . . 58
2.3.3 Limitations of the pole stimulus . . . 61
2.3.4 Vibrotactile training . . . 62
2.3.5 Unisensory auditory training . . . 66
2.3.6 Multisensory training . . . 67
2.3.7 Rate of acquisition on the auditory, vibrotactile, and multisensory tasks . . . 70
2.3.8 Lick patterns . . . 73
2.3.9 Within-session variations . . . 82
2.3.10 Modality preferences during multisensory training . . . 91
2.4 Discussion . . . 94
2.4.1 Overview . . . 94

2.4.2 On the importance of designing an appropriate stimulus . . . 95
2.4.3 Potential effects of hearing loss in mice . . . 97
2.4.4 On the importance of using an appropriate measure of behavioural choices . . . 97
2.4.5 Multisensory task difficulty: Cognitive or Perceptual? . . . 101
2.4.6 Conclusions . . . 104

**3 Effect of training on the neural representation of stimuli 105**
3.1 Introduction . . . 105
3.1.1 Tactile responses in vS1 . . . 106
3.1.2 Training effects in vS1 . . . 109
3.1.3 Multisensory encoding in primary sensory cortices . . . 110
3.1.4 Objectives . . . 112
3.2 Methods . . . 113
3.2.1 Subjects . . . 113
3.2.2 Surgical procedures . . . 114
3.2.3 Behavioural training . . . 114
3.2.4 Imaging . . . 115
3.2.5 Longitudinal tracking of neurons across imaging sessions . . . 116
3.2.6 Barrel cortex targeting with optical signal imaging . . . 119
3.2.7 Support vector machine classification . . . 121
3.3 Results . . . 122
3.3.1 Neural responses during passive stimulation in naïve mice . . . 122
3.3.2 Neural responses during passive stimulation in trained mice . 128
3.3.3 Changes to passive responses pre- and post-training . . . 131
3.3.4 Changes in neural selectivity for stimuli throughout multisensory training . . . 134
3.3.5 Response profile changes throughout training . . . 137
3.3.6 Stimulus decoding from vS1 responses . . . 148

3.3.7 Training alters the responses of most neurons to the trained stimuli . . . 159
3.3.8 Neurons in vS1 are modulated by licking . . . 163
3.3.9 Neurons in vS1 are modulated by locomotion. . . . 175
3.4 Discussion . . . 188
3.4.1 Summary of findings . . . 188
3.4.2 Characterising responsiveness to stimuli . . . 189
3.4.3 Effects of training on responses . . . 193
3.4.4 Sensorimotor integration . . . 198
3.4.5 Conclusion . . . 200

**4 Discussion** **202**
4.1 Summary . . . 202
4.2 The choice of a behavioural paradigm . . . 204
4.3 Defining a measure of responsiveness . . . 207
4.4 Limitations of two-photon imaging . . . 208
4.5 What networks support multisensory integration in vS1? . . . 209
4.6 Is the primary somatosensory cortex only sensory? . . . 211
4.7 General conclusion . . . 213

**Bibliography** **215**

# List of Figures


1.1 Necker Cube . . . . . 2
1.2 Whisker pad . . . . . 12
1.3 Barrels are mapped to the whisker pad . . . . . 14

2.1 Behavioural outline . . . . . 47
2.2 Multisensory congruency snapshots . . . . . 50
2.3 Frequency measurements . . . . . 52
2.4 Pole position performance . . . . . 57
2.5 Pole-position neural response . . . . . 60
2.6 Pole-position tuning . . . . . 61
2.7 Vibrotactile task performance . . . . . 64
2.8 Whisker-dependent task control . . . . . 65
2.9 Auditory task performance . . . . . 67
2.10 Multisensory task performance . . . . . 69
2.11 Rates of acquisition of unisensory and multisensory tasks . . . . . 73
2.12 Lick distributions . . . . . 76
2.13 Number of licks per trial . . . . . 79
2.14 Within-session variations, Hit v. FA . . . . . 84
2.15 Lick-inferred within-session variations . . . . . 90
2.16 Cumulative FA rates in trained and training mice . . . . . 92
2.17 Sensory bias during multisensory training . . . . . 93

3.1 Number of neurons . . . . . 118
3.2 Intrinsic optical signal imaging . . . . . 121

3.3 Sample traces . . . . . 124
3.4 Selectivity and tuning of neurons . . . . . 127
3.5 Selectivity and tuning of neurons . . . . . 131
3.6 Selectivity and tuning of neurons . . . . . 134
3.7 Multisensory sensitivity . . . . . 137
3.8 Response types . . . . . 139
3.9 Response profile to tactile stimuli . . . . . 141
3.10 Response profile to multisensory stimuli . . . . . 144
3.11 Response profile to multisensory stimuli (2) . . . . . 145
3.12 Tactile stimulus decoding . . . . . 151
3.13 Multisensory stimulus decoding in responsive and non-responsive neurons 154
3.14 Multisensory stimulus decoding in excited and suppressed neurons . . 155
3.15 Multisensory stimulus decoding in L2 and L3 neurons . . . . . 156
3.16 Tracked neuron responsiveness stability . . . . . 161
3.17 Licking and stimulus presence . . . . . 164
3.18 Responses of tactile sharp onset neurons to licks . . . . . 165
3.19 Responses of tactile amplitude-locked neurons to licks . . . . . 166
3.20 Responses of tactile delayed onset neurons to licks . . . . . 167
3.21 Responses of tactile suppressed neurons to licks . . . . . 168
3.22 Responses of non-responsive neurons to licks . . . . . 169
3.23 Responses of multisensory sharp onset neurons to licks . . . . . 170
3.24 Responses of multisensory amplitude-locked neurons to licks . . . . . 171
3.25 Responses of multisensory delayed onset neurons to licks . . . . . 172
3.26 Responses of multisensory suppressed neurons to licks . . . . . 173
3.27 Responses of non-responsive neurons to licks . . . . . 174
3.28 Running and stimulus presence . . . . . 176
3.29 Responses of tactile sharp onset neurons to locomotion . . . . . 178
3.30 Responses of tactile amplitude-locked neurons to locomotion . . . . . 179
3.31 Responses of tactile delayed onset neurons to locomotion . . . . . 180

3.32 Responses of tactile suppressed neurons to locomotion . . . . . 181
3.33 Responses of non-responsive neurons to locomotion . . . . . 182
3.34 Responses of multisensory sharp onset neurons to locomotion . . . . . 183
3.35 Responses of multisensory amplitude-locked neurons to locomotion . . 184
3.36 Responses of multisensory delayed onset neurons to locomotion . . . . 185
3.37 Responses of multisensory suppressed neurons to locomotion . . . . . 186
3.38 Responses of non-responsive neurons to locomotion . . . . . 187

# List of Tables


2.1 Multisensory trial types . . . . . 51
2.2 Mean number of licks per Go and No-Go trial . . . . . 76
2.3 Mean number of licks per Hit and FA trial . . . . . 80
2.4 Lick-rate inferred performance . . . . . 81
2.5 Within-session changes in performance . . . . . 82
2.6 Within-session changes in Hit and FA rates . . . . . 87
2.7 Lick-rate-inferred within-session changes in performance . . . . . 88

3.1 Number of neurons per day . . . . . 118
3.2 Passive stimulus responsiveness . . . . . 127
3.3 Passive stimulus responsiveness . . . . . 131
3.4 Proportion of unisensory- and multisensory-sensitive neurons across training . . . . . 136
3.5 Proportion of tactile responses . . . . . 142
3.6 Proportion of multisensory responses . . . . . 147
3.7 Unisensory vibrotactile classifier performance . . . . . 152
3.8 Multisensory classifier performance . . . . . 158
3.9 Stability of tracked neurons responses . . . . . 160
3.10 Stability of tracked neurons responses . . . . . 162

# List of abbreviations

**2-AFC** two alternative forced choice

**A1** primary auditory cortex

**ANOVA** analysis of variance

**BI** sensory bias index

**CR** Correct Rejection

**FA** false alarm

**IC** inferior colliculus

**IOSI** intrinsic optical signal imaging

**ITI** inter-trial interval

**M1** primary motor cortex

**POM** posterior medial nucleus of the thalamus

**ROI** region of interest

**S1** primary somatosensory cortex

**S2** secondary somatosensory cortex

**SAM** sinusoid amplitude modulation

**SC** superior colliculus

**SVM** support vector machine

**V1** primary visual cortex

**VPM** ventral posterior medial nucleus of the thalamus

**vS1** vibrissal primary somatosensory cortex

# Chapter 1

# General Introduction

> The senses perceive each other's special objects incidentally; [...] all form a unity.
>
> *Aristotle*
> *De Anima*

## 1.1 Behaving in a multisensory world

We live in a loud, complex and noisy environment. At every instant, we are bombarded with information collected through a wide array of modalities. We may not always be aware of separate senses, instead perceiving unified scenes. Strikingly, we sometimes are only aware of our reliance on multiple modalities when information from one is removed. For instance, it is typically harder to shave without a visual aid (mirror) or without auditory feedback (i.e. with ears plugged) [Foxe, 2009]. Senses are essential in our daily interactions with our own environment as well as in social situations. It is thanks to sight that we can find a friend in a loud party. Similarly, it is thanks to audition that we can talk with one another or hear a bus coming while crossing the road. The list of examples is endless, as most of our behaviours are sensory-guided and we rely on a constant flow of information collected through

our sensory organs to make sense of the environment in which we navigate and with which we interact.

Integrating information across multiple senses can enable an individual to perceive their surroundings more reliably and may help resolve sensory ambiguities. The Necker Cube is an example of such an ambiguity [Necker, 1832]. The representation of a three-dimensional cube in a two-dimensional drawing can lead to a bi-stable percept wherein observers report it being seen alternatively from above and below (Figure **1.1**).

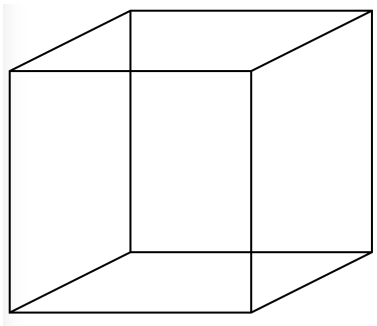


**Figure 1.1:** The Necker Cube: a two-dimensional drawing of a three-dimensional cube leading to a bi-stable percept of a cube seen from above and below.

Such an example is theoretical and, in a practical situation, haptic feedback (e.g. grabbing and handling a transparent cube) would help disambiguate the percept. It does however illustrate the fact that different modalities may carry different pieces of information which, combined, help us resolve our environments.

Throughout this Chapter, I will first describe behavioural and psychological observations on perception and multisensory integration. Subsequently, I will explain in more depth our current understanding of the neural substrates supporting such perception.

### 1.1.1 Sensory dominance and multisensory illusions

Senses are weighed depending on how reliable they are [Angelaki et al., 2009]. This is not only due to the intrinsically noisy nature of stimuli but also the fact that sensory organs are themselves limited and optimised for particular functions. For example, the human retina supports high visual acuity [Hirsch and Curcio, 1989] which in turn enables us to orient and direct ourselves in space as well as avoid obstacles. In comparison, the sense of touch requires physical contact. Needless to say, it is preferable to rely on vision than touch to find someone in a crowd. On the other hand, vision would prove quite useless in total darkness and we would in that case rely more heavily on audition and touch to find our bearings. This preferred reliance upon a modality over an other does not however appear to be a conscious percept, as illustrated in a study led in 1974 [Colavita, 1974]. Participants were asked to categorise whether they were presented with visual stimuli or auditory stimuli with similar equivalent intensities. In a subset of catch trials, the participants were presented with both auditory and visual stimuli. They consistently reported detecting the visual stimulus over the auditory stimulus, suggesting visual over auditory dominance in humans. Similarly, in a another task comparing visual and haptic performance [Rock and Victor, 1964], participants were asked to estimate the size of a block they either perceived by touching or by watching through a distorting lens. The findings of this study showed that participants relied more heavily on their visual perception, often without even being aware of a conflict between senses. A later study showed that increasing the noise level of visual stimuli, and correspondingly decreasing their reliability, boosted the relative haptic weight in the size estimate reported by participants [Ernst and Banks, 2002]. In short, humans rely the most on vision to resolve their environment but will rely on other modalities when vision is

not sufficient or informative enough.

This last study hinted toward the notion that information is integrated dynamically across senses, and one modality does not necessarily always take over another one. This is best observed through illusions based on incongruities between modalities. One such example is the famous McGurk effect [McGurk and Macdonald, 1976]. Briefly, presenting a viewer with a video of a person mouthing syllables but playing a recording of someone saying a different sound creates a confusing composite percept (i.e. perceiving the syllables 'da-da' when seeing 'ga-ga' and hearing 'ba-ba'), or new combination non-words (i.e. 'gabga'). This is an example in which one modality does not necessarily dominate the other, rather they are combined into a unified percept. This was also aptly summarised by Alais and Burr in a study of the processes underlying the ventriloquist effect [Alais and Burr, 2004]. A common example of the ventriloquist effect is the concept of visual capture of sound, in which we perceive speech as coming from an actor's lips on television, while the actual sound originates from speakers placed elsewhere. In this particular study, the experimenters varied the reliability of visual stimuli in an audio-visual localisation task and, similarly to Ernst and Banks in 2002, found that participants followed the location of auditory cues over visual ones when using the most blurry visual stimuli. In most conditions though, for instance using moderately blurry stimuli, participants integrated information from both visual and auditory cues and located stimuli more accurately than in visual-only or auditory-only trials. In a well characterised example of multisensory interaction, Jousmäki and Hari played back recordings of subjects rubbing their palms together and either accentuated or dampened high frequencies in the recording [Jousmäki and Hari, 1998]. Enhancing high-frequency feedback led to a perception of dry skin among the participants,

famously known as the 'parchment-skin' illusion.

While the main sense upon which humans rely is vision, this may not be the case across all animal species. Indeed, vision is most useful for us to navigate our surroundings. In contrast, nocturnal animals or species living underground may best perceive their environments through different modalities. One such species, the mouse (*Mus musculus*) has proven itself a useful animal model for behavioural training and concurrent imaging [Guo et al., 2014, Peron et al., 2015, Meijer et al., 2017]. This will be further discussed in the context of neural responses to stimuli and behaviour in Section **1.2**. The visual acuity of mice is substantially lower than that of human observers [Caves et al., 2018]. Mice rely heavily on audition and active touch to explore their environments. They emit complex ultrasonic vocalisations to communicate [Sales, 1972] and use their whiskers to navigate within their environments and explore objects [Prescott et al., 2011]. In a recent study on sensory dominance [Song et al., 2017], experimenters trained mice on a head-fixed Go/No-Go task. Some animals were trained to lick when presented with a sound (noise or pure tone) and withhold when presented with a flashing light while others were trained on the reverse version of the task. In some test trials, both sounds and lights were presented, therefore indicating contradictory signals to both lick and withhold. In both training cohorts, mice responded to these trials as if they were auditory (i.e. licked if auditory trials were rewarded or withheld if they were not).

To summarise, there are clear examples of cross-modal interactions between most senses and many examples of combination of information supporting a more accurate perception of multi-modal stimuli. More importantly, decades of experimenting have shown that modalities may interact and interfere with one another without an observer being even aware of it.

### 1.1.2 Processing tactile input from the whiskers

A large body of experimental evidence has shown that rodents can perform a vast array of sensory-guided behaviours by using their whiskers. For instance, rats have been found to reliably estimate the size of an aperture by using only their whiskers [Krupa et al., 2004]. Further, mice have been trained to discriminate between different azimuthal pole positions [O'Connor et al., 2010b, Peron et al., 2015], even with all but one whiskers trimmed [Guo et al., 2014]. Beyond spatial perception (i.e. size, position), mice have also been found to use their whiskers to discriminate between different textures of sandpaper [Chen et al., 2015, Wu et al., 2013]. A common marker of many of these tasks is the requirement for active whisker movement, or whisking, to support such behaviours. As noted as early as 1962 [Gibson, 1962], touching and being touched are not the same percepts. Indeed, in the aforementioned examples, it is not certain that rats or mice may be able to pinpoint the size of an aperture or the position of a pole without actively whisking against them. Further, it seems impossible for rodents to discriminate between textures of static sandpaper without actively brushing against it. Nonetheless, passive whisker deflections should not be discarded as uninformative stimuli. Indeed, a more recent study [Miyashita and Feldman, 2013a] showed rats' ability to learn a right-left whisker discrimination task (in which rats had to lick either a left or right drinking port depending on the lateral position of a piezoelectric actuator-driven whisker deflection) in a task in which tactile information was collected passively – that is, the whiskers were deflected by an exterior stimulus instead of actively brushed against a static object by the rats. Another set of studies showed that rats were able to discriminate between vibrotactile stimuli based on the product of their amplitude and frequency [Arabzadeh et al., 2003, Adibi et al., 2012]. In these studies,

the whiskers were either resting on a custom built ladder or actively positioned by the animals against a mesh. In both cases, the vibrotactile parameters were delivered through a piezoelectric driver and the tactile stimulus (i.e. whisker deflection) was not dependent on active whisking on the rats' part. Hence, whether they may need to actively whisk against objects or not, it is now clearly understood that rodents can learn to discriminate between different textures, vibration frequencies and amplitudes, or positions.

### 1.1.3 Auditory perception in mice

As previously mentioned, mice produce complex ultrasonic vocalisations which enable them to communicate with one another [Sales, 1972]. The mouse hearing range is far wider than that of human listeners, ranging from a conservatively low estimate of 1 kHz to upwards of 80 kHz [Heffner and Heffner, 2007]. They have been shown to discriminate between sounds with very close frequencies [De Hoz and Nelken, 2014] and even to maintain high levels of performance (over 70% correct discrimination) in a task reversal paradigm over tens of training sessions [Jaramillo et al., 2014]. Such observations have been also been made in pitch discrimination tasks [Tsukano et al., 2011]. In short, mice are well suited for auditory tasks and their ability to discriminate between several different sound features is today well understood.

### 1.1.4 Multisensory perception in mice

Both auditory- and tactile-guided behaviours of mice have been extensively used and characterised. More recently, studies have investigated the ability of mice to perform multisensory-guided tasks.

For instance, mice were found to have increased detection performance of audiovisual stimuli over unisensory auditory or visual stimuli [Meijer et al., 2018]. More specifically, the experimenters varied the amplitude of the stimuli (visual contrast and auditory loudness respectively) following adaptive staircase procedures. In short, they constantly updated the amplitude of the respective stimuli depending on an animal's performance in an ongoing session, by increasing the amplitude if the performance dropped and decreasing it if it increased. Eventually, the amplitude of both stimuli narrowed around the perceptual thresholds for each modality. Detection performance was low with low amplitude stimuli across unisensory and multisensory tasks and conversely consistently high across all trial types with full intensity stimuli. However, around the perceptual threshold for these stimuli, mice responded significantly more often to audiovisual stimuli than auditory or visual cues alone. Another study found similar results in a discrimination task [Siemann et al., 2015]. Briefly, mice were trained to selectively attend a port to obtain a reward based on a spatial light cue (a light in the port to attend) and an auditory cue (a pure tone or white noise indicated separate ports). They observed an increase in discrimination accuracy in multisensory trials compared to unisensory trials. Specifically, they presented trials of with varying stimulus lengths (ranging from 50 ms to 1 s) and observed the strongest multisensory gain (i.e. difference between multisensory performance and the best unisensory performance) at intermediate stimulus duration (100-300 ms). It could be argued that changing the length of stimuli is equivalent to changing their reliability, as the accuracy of mice increased overall with increased stimulus length.

These two studies illustrate very well the principle of inverse effectiveness, which states that 'the degree to which a multisensory response exceeds the response to

the most effective modality-specific stimulus component declines as the effectiveness of the modality-specific stimulus components increases' [Stein and Meredith, 1993]. To put it plainly, this principle underlines the fact that benefits from multisensory integration are most prevalent around the perceptual thresholds for individual modalities. Otherwise, information contained within a single modality may be sufficient to drive sensory-guided behaviour without the need to extract more information from other senses. These are typical examples in which adding a modality enhances performances, a term which will be referred to throughout this Thesis as 'multisensory enhancement'.

In such studies, it is not obvious whether animals rely on one modality over another. Indeed, an enhancement in performance in multisensory stimuli containing the same components around perceptual thresholds as unisensory trials does show an improvement, probably due to more information being available from different senses. It is not clear whether animals still rely preferentially on one modality and use another to achieve an incremental gain in performance or whether they rely equally on both modalities. However, multisensory integration can serve other purposes. For example, we regularly rely on cross-modal combinations of features, such as touch, vision and smell in order to pick fruits or vegetables. Fewer studies have examined how we learn to categorise objects in this manner, based on their combinations of features across multiple senses, a process known as 'multimodal combination identification'.

As noted in a recent review by Deroy and colleagues [Deroy et al., 2016], it is important to distinguish such cases in which 'unisensory consciousness [is] biased by the information provided by another modality, be it consciously perceived or not' from a different case in which multisensory objects are experienced as part of a multisensory scene. They referred to these examples to contrast cases of 'crossmodal bias of

unisensory awareness' with cases of 'multisensory scene perception'. I have already highlighted several studies in which multisensory stimuli help resolve ambiguities around perceptual thresholds – cases of crossmodal bias. On the other hand, we are all familiar with the fact that most of our percepts are unified – in fact you may right now not be separately aware of your position on a chair, the sound your computer makes and the taste and warmth of the cup of tea next to you separately. Instead, these percepts may be integrated into one unified multisensory scene. To my knowledge, this distinction between multisensory enhancement and multimodal combination identification has not been made in the study of multisensory integration and multisensory perception in mice.

As I will discuss in Section **2.4.5**, I did not aim to study the limits of perceptual discrimination and how multisensory stimuli may help enhance performance around these limits. Rather, I wanted to characterise the ability of mice to learn to categorise multisensory objects. This may not reflect the conscious experience of a multisensory percept but would, as suggested by Deroy and colleagues [Deroy et al., 2016], at the very least hint to the binding of two modalities into a unified experience, in which different modalities may remain unisensory but be co-attributed to a common experience. I subsequently aimed to assess how neural responses to unisensory parts of these objects may change with training and experience.

## 1.2 Measuring neural correlates of perception

The mouse is a useful and widely used animal model for measuring neural responses to sensory stimuli. In particular, two-photon calcium imaging

allows us to measure the responses of large numbers of individual neurons in the cortex simultaneously, bringing new insights into sensory processing that were not possible with electrophysiological techniques alone [Denk and Svoboda, 1997]. In the auditory cortex, this technique revealed the fine-scale heterogeneity of the tonotopic mapping of the primary auditory cortex (A1) not observed previously using electrophysiological recordings [Rothschild et al., 2010, Stiebler et al., 1997]. Several studies from the same research group characterised neural responses in the vibrissal primary somatosensory cortex (vS1) of mice to poles being placed at various positions relative to their whiskers [O'Connor et al., 2010c, Guo et al., 2014, Peron et al., 2015]. Thanks to the wide variety of available transgenic mouse lines, mice expressing calcium indicators under neuron-specific indicators can be imaged, allowing for a targeted identification of responses to stimuli based on cell types, for instance only imaging GABAergic interneurons [Prönneke et al., 2015]. Similarly, Cre-dependent targeting of auditory cortex neurons projecting to the primary visual cortex (V1) has allowed for the characterisation of projections between sensory cortices and cross-modal responses in V1 [Deneux et al., 2019]. The combination of behavioural training on a tactile discrimination task and imaging has enabled the tracking of responses of individual neurons in awake, behaving animals [Li et al., 2016, Peron et al., 2015, Ceballo et al., 2019]. The main point to note here is the suitability of the mouse as an animal model to investigate the neural representation of multisensory stimuli during training. The following sections will now describe in more depth our understanding of sensory pathways and the studies of neural responses to sensory stimuli in mice.

### 1.2.1 Neural substrates of tactile perception

Several textbooks and reviews have addressed and characterised the processing of tactile information from the periphery (the whiskers) to cortical processing [Petersen, 2007, Diamond and Arabzadeh, 2013, Helmchen et al., 2018]. While the scope of this Thesis will focus on the vibrissal primary somatosensory cortex (vS1), it is useful in this introduction to synthesise our current understanding of somatosensory processing before addressing recent findings in multisensory processing within the somatosensory pathway.

#### Peripheral and subcortical processing

The neural pathways for whisker information processing have been substantially studied and reviewed [Petersen, 2007]. Whiskers are deeply innervated structures [Ebara et al., 2002] arranged in a grid-like layout. Larger whiskers, known as macrovibrissae, are arranged in horizontal rows (A to E) and vertical columns (up to 12 follicles per row) [Brecht et al., 1997] (Figure **1.2**).

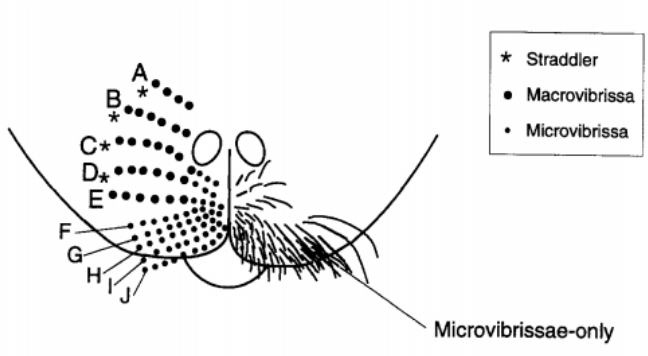



**Figure 1.2:** Schematic frontal view of a mouse whisker-pad. Macrovibrissae in rows A to E are shown as large dots, as well as small microvibrissae in rows A to J. Straddlers (asterisks) are caudal whiskers that are displaced relative to rows. Adapted from Brecht et al., 1997.

The way whiskers move during active whisking and how they brush against objects depending on the roughness of their surface has been studied and tracked substantially [Isett et al., 2018, Clack et al., 2012]. Whiskers glide and stop against irregular surfaces in a motion described as 'stick-slip' [Diamond et al., 2008, Chen et al., 2015], in which whiskers catch on surface irregularities ('stick') and, after a time during which the whiskers stay stuck and their curvature increases, they spring free ('slip'). These 'stick-slip' events have been found to support discrimination performance, as the performance of mice at texture discrimination decreases in trials without such events [Chen et al., 2015]. The exact nature and transduction mechanisms of receptors within the whisker follicles are unclear although it appears that trigeminal ganglion neurons respond to whisker motion, acceleration and even texture (i.e. 'stick-slip' events) [Arabzadeh et al., 2005, Arabzadeh et al., 2006]. These neurons possess receptor endings in peripheral branches within whisker follicles and project to the trigeminal nuclei of the brainstem. In the lemniscal pathway, trigeminal neurons project to the ventral posterior medial nucleus of the thalamus (VPM). Finally, thalamic neurons typically project to layer 4 (L4) of vS1 and, to a lesser extent, layer 6 (L6). In another pathway, the paralemniscal pathway, trigeminal neurons project to the posterior medial nucleus of the thalamus (POM) which itself predominantly projects to layers 1 and 5 of vS1 [Bureau et al., 2006]. At each stage of these pathways, neurons appear to be arranged in a somatotopic fashion, with each macrovibrissa corresponding to respectively a barrelette in the trigeminal nucleus [Veinante and Deschênes, 1999], barreloids in the VPM [Brecht and Sakmann, 2002] and finally barrels in vS1 L4 [Woolsey and Van der Loos, 1970] (See Figure **1.3**).

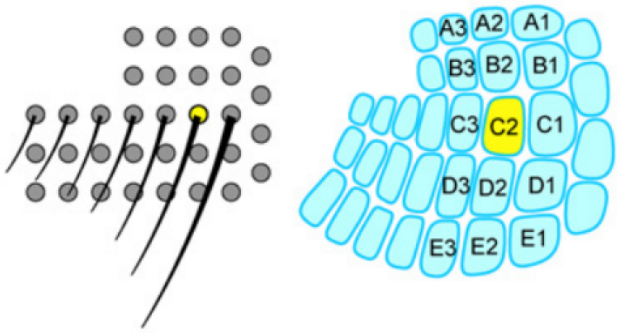



**Figure 1.3:** Schematic view of the organisation of vS1 in the mouse. Each vS1 barrel (right, blue and yellow) corresponds to a single macrovibrissa. The C2 barrel, shown in yellow on the right, correspond to the C2 whisker, shown on the left with the yellow base. Adapted from Petersen 2007

### Early cortical processing

The primary somatosensory cortex is, like the rest of the neocortex in mammals, arranged in six layers (L1 to L6). As mentioned above, the vibrissal area, vS1, is characterised by barrels, making the 'whisker area' of the somatosensory cortex readily identifiable [Woolsey and Van der Loos, 1970]. Inputs from the lemniscal and paralemniscal pathways terminate in L4 barrels and L1 and L5A respectively. Further mapping of inputs into rodent barrel cortex further highlighted a relative absence of direct thalamic inputs into L2/3 neurons. In fact, L3 receives lemniscal L4 inputs while L2 is more predominantly targeted by paralemniscal L5 neurons [Bureau et al., 2006].

The cortical neural correlates of tactile perception have been observed and studied for a long time, using a wide array of tactile stimuli to measure responses, such as positional pole stimuli [Peron et al., 2015, Guo et al., 2014], textured sandpaper [Chen et al., 2015, Isett et al., 2018] or piezoelectric drivers deflecting threaded

whiskers [Helmchen et al., 2018]. While trigeminal and thalamic neurons respond robustly and consistently to whisker deflections [Arabzadeh et al., 2005], we now understand that vS1 neurons are sparsely driven by tactile inputs. This could be explained by: 1) trial-to-trial variability, with different neural subsets responding to different trial presentations; 2) small but consistent proportions of responsive neurons; 3) increased proportion of responsive neurons with increased alertness or motivation; and 4) different fractions of highly-tuned neurons responding to specific stimuli only (for review, see [Barth and Poulet, 2012]).

The exact proportion of vS1 neurons responsive to touch has been a point of contention. Some findings suggest that a low proportion (20 to 50%) of neurons in vS1 typically respond to tactile stimuli [Peron et al., 2015]. In contrast, Chen and colleagues found up to 60% of active neurons to respond to tactile stimuli [Chen et al., 2015] while a more recent study identified only 16% of responsive neurons to passive whisker deflections [Zhang et al., 2020]. Another study found 64% of touch-responsive neurons in L2/3 and 57% in L5 of running mice [Ayaz et al., 2019]. Each of these studies investigated neural responses to different tactile stimuli (i.e. wall contact during locomotion or a spatial pole position) in experimental conditions ranging from anaesthetised imaging to engaged operant behaviour, which may help explain such discrepancies.

Considering the wide variety of tactile stimuli reported in the aforementioned study, what features may vS1 neurons actually respond to? Barrel cortex neurons have been found to respond predominantly to the sudden changes in speed and curvature associated to 'slips' [Jadhav and Feldman, 2010, Isett et al., 2018]. Further, some studies have focused on single-barrel responses to stimuli in animals in which all but one whisker had been removed, in effect leaving mice with one

functional whisker upon which to rely [Peron et al., 2015, Guo et al., 2014]. While it may have originally been thought that all neurons within a single barrel respond to deflection of its associated whisker, it appears in fact that, similarly to frequency preferences in A1, the broad somatotopic structure of vS1 still reveals more local, 'salt-and-pepper' distributions of whisker preferences [Clancy et al., 2015]. Clancy and colleagues used two-photon imaging to measure responses to whisker deflections and map the distribution of preferred whiskers (i.e. whiskers for which a deflection elicited the highest neural response for a given neuron) across vS1 of mice. They found that, while neurons located closer to the center of a barrel tended to respond preferentially to its associated whisker, neurons further toward the edge of barrels responded to deflections delivered to other, often neighbouring, whiskers.

The selectivity of these responses may further help inform on how vS1 neurons support tactile perception. Indeed, the fact that individual barrels are composed of spatially interspersed neurons with varied preferred whiskers [Clancy et al., 2015] suggests that the perception of some stimuli could be based on population responses (i.e. different neurons responding as whiskers swipe across a texture or an object sequentially). Further, Isett and colleagues reported a non-negligible proportion of neurons (18%) responding similarly preferentially to particular textures, corresponding to specific slip rates and amplitudes [Isett et al., 2018]. Hence, combined, these findings suggest that separate features of tactile perception, such as texture or location, are encoded on a population level with subpopulations of neurons exhibiting respective slip rates or whisker preferences. The inherent difference between active touch and passive touch as well as other behavioural factors such as locomotion may further modulate neural responses in vS1, which will be further discussed in Section **1.2.2**.

**Later cortical processing**

Higher cortical processing of tactile information lies somewhat beyond the scope of this Thesis. Briefly, S1 and S2 are densely and reciprocally connected [Aronoff et al., 2010]. While it has been a debated point [Menzel and Barth, 2005], S2 appears to exhibit some responses to multisensory stimuli, including stimuli from another modality altogether, such as auditory cues [Woolsey and Van der Loos, 1970, Zhang et al., 2020].

Another target of vS1 projections is the primary motor cortex (M1), which responds to somatosensory stimuli and reciprocates projections to vS1 [Aronoff et al., 2010, Petreanu et al., 2012]. In fact, M1 neurons appear to respond to sensory stimuli with a short delay (less than 10ms) compared to vS1 neurons [Ferezou et al., 2007]. The extent to which M1 and vS1 reciprocally connect is consistent with the sensorimotor nature of whisker-based tactile exploration in mice, based on active whisking and therefore a tight relationship between afferent sensory inputs and efferent motor outputs. The following Section will discuss in depth this feature of tactile perception.

### 1.2.2 The barrel cortex is an inherently sensorimotor area

It is now commonly accepted that vS1 plays a bigger role than simply responding to tactile stimuli. For example, targeted intracortical microstimulation in a vS1 barrel has been found to elicit retraction of its associated whisker [Matyas et al., 2010]. Beyond driving whisking, a study using a tactile virtual reality setup [Ayaz et al., 2019] found neurons whose responses were modulated by running and concurrent whisking, including some neurons modulated by whisking alone or

running alone. Whisking and locomotion are tightly related, to the extent that mice whisk whenever they run [Sofroniew et al., 2015], although the opposite (i.e. when mice do not run they do not whisk) is not true. Indeed, mice can actively whisk even when not running [Mayrhofer et al., 2013]. This is of particular relevance to the study of tactile information processing, as a previous study reported that observed neural responses to passive whisker deflections were stronger when animals were not whisking [Crochet and Petersen, 2006]. In another study, Petreanu and colleagues characterised M1 projections to vS1 [Petreanu et al., 2012], further highlighting the tight sensorimotor integration occurring as early as the primary sensory cortex in the tactile pathway. This further adds to previous findings of trans-thalamic sensorimotor pathways [Mo and Sherman, 2019] and emphasises on the intertwined efferent motor control and afferent sensory perception in the mouse somatosensory system.

These findings sum up the nature of active touch which differs from passive tactile stimulation. As mentioned previously, touching and being touched are indeed not the same percepts [Gibson, 1962], and these differences may be related to the variety of available tactile information collected through whiskers, including notably object location or texture perception. Many reciprocal connections have been found between the primary somatosensory cortex (S1) and the secondary somatosensory cortex (S2) as well as with the primary motor cortex (M1) [Aronoff et al., 2010]. Specifically, a study identifying vS1 neurons projecting to S2 and M1 found that more neurons projecting to M1 exhibited touch-related activity during object localisation whereas more neurons projecting to S2 were active during texture discrimination [Chen et al., 2013a]. Such specialisation may hint toward differential processing streams for tactile information on 'what' as opposed to 'where' [Diamond et al., 2008].

### 1.2.3 Neural substrates of auditory perception

#### Peripheral and subcortical processing

Auditory information is processed through a different pathway altogether. The exact structure of the auditory pathway will be briefly summarised here [Winer and Schreiner, 2005].

Sound waves are converted into neural signals in the cochlea by hair cells and carried to the cochlear nuclei. The hair cells are disposed along the basilar membrane, which runs through the cochlea. At its base, the membrane is thickest and becomes thinner as we reach the end of the cochlea. As such, the sound frequencies that generate basilar membrane oscillations and allow for subsequent sound transduction vary along the membrane. In short, just as tactile information is mapped on the whisker pad [Woolsey and Van der Loos, 1970] (somatotopic organisation), the cochlea provides a mapping of sound frequencies, or tonotopic arrangement.

Subcortical processing of auditory information is extensive and integrates information from both ears early on, a feature most necessary for sound localisation. The cochlear nuclei project excitatory outputs respectively to the ipsilateral superior olive complex and the medial nucleus of the trapezoid body, which itself projects inhibitory inputs to the contralateral olive complex. This complex has been shown to support sound localisation processes by integrating interaural level and time differences [Tollin, 2003, Tollin, 2005]. The inferior colliculus (IC), in the midbrain, is a site of extensive convergence of auditory signals [Adams, 1979] and the last node before the thalamus. The thalamic relay to the primary auditory cortex (A1) is the medial geniculate body of the thalamus. At each stage of auditory information processing, the involved structures are mapped to the cochlea and organised following

a gradient of best sound frequencies to which neurons respond. Of course, as mentioned above, there is more to sound information than simply its constituent frequencies such as sound location, a feature already processed subcortically.

### Cortical processing

Similarly to reports made by Clancy and colleagues in vS1 [Clancy et al., 2015], A1 is organised broadly following a tonotopic gradient of best frequencies but has a more local 'salt-and-pepper' layout of frequency tuning in neurons [Kanold et al., 2014, Rothschild et al., 2010]. Further studies have also examined the organisation of A1 neurons with regards to sound localisation cues, finding for example a clustering of neurons with similar preferences but no global scale organisation [Panniello et al., 2018]. A striking feature of many sounds, such as spoken syllables, is their amplitude modulation, which describes variations in a sound's amplitude envelope (i.e. the sound's frequency may remain but its amplitude varies). Amplitude modulation is understood to carry information useful for pitch perception and auditory scene analysis [Bregman and McAdams, 1994] and some auditory cortex neurons have been found to specifically respond to amplitude-modulated sounds [Liang et al., 2002]. Specifically, A1 neurons phase-lock to the amplitude modulation of sounds [Yin et al., 2011]. Hence, A1 neurons inform about a wide array of sound features through combinations of temporal and rate codes (i.e. phase-locking and timing of firing as responses to sound modulation as well as action potential firing rate as responses to, for instance, a preferred sound frequency). The importance of this feature in multisensory perception will be highlighted in Sections **1.3.3** and **2.3.3**.

Higher auditory fields, including the secondary auditory cortex, anterior

auditory fields and posterior auditory fields, do not only receive projections from A1 but as well from subcortical nuclei [Morel and Imig, 1987, Lee et al., 2008]. A functional deactivation study has nonetheless hinted toward specialised streams of auditory information processing, namely a frontal 'what' stream necessary for pattern discrimination and a dorsal 'where' stream necessary for sound localisation [Lomber and Malhotra, 2008].

Here, perhaps of most importance to the reader is the fact that most models of sensory processing highlight extensive subcortical networks and precisely identified primary sensory cortices, in which it was originally thought that information from only a single modality was processed. Before addressing such claims, I will now describe the extensive body of research that has identified multisensory integration within subcortical regions as well as in higher cortical areas.

## 1.3 Neural substrates of multisensory integration

Stein and Stanford defined multisensory integration as 'the neural processes that are involved in synthesising information from cross-modal stimuli' [Stein and Stanford, 2008]. Neural responses to multisensory stimuli have long been studied, with a particular distinction between subcortical processes and cortical integration.

### 1.3.1 Subcortical responses to multisensory stimuli

Research into the neural substrates of multisensory integration is not new. In fact, electrophysiological studies have shown since the 1970s that

multisensory integration occurs in subcortical areas [Drager and Hubel, 1975, Meredith and Stein, 1986, Meredith et al., 1987]. Most notably, these studies identified robust superadditive responses in superior colliculus (SC) neurons. Meredith and Stein identified neurons with an increase in the number of impulses during audiovisual stimulation of anaesthetised cats of up to 1200% compared to auditory or visual alone [Meredith and Stein, 1986]. As found in this study and later discussed in a review [Stein and Stanford, 2008], this enhancement of responses is most striking when cues are individually weak and evoke few responses – a neural illustration of the principle of inverse effectiveness mentioned in Section **1.1.4**. Similarly, another study in guinea pigs found superior colliculus (SC) neurons with enhanced audiovisual responses and other with depressed responses (over a 50% drop in spikes per trial) [King and Palmer, 1985]. Further, varying the relative timing of visual stimulus onsets to auditory stimuli modulated the degree of that enhancement, highlighting the importance of timing and synchronisation on the degree to which separate modalities are integrated. This will, as with amplitude modulation, be most relevant in Sections **1.3.3** and **2.3.3**. In the thalamus, a recent study in rats reported responses to cross-modal visuo-tactile stimuli in a predominantly tactile area (VPM) but not in the dorsal lateral geniculate nucleus, a region whose first order modality is understood to be vision [Bieler et al., 2018]. In comparison, an fMRI study in humans found that concurrent sound presentation enhanced activity in the lateral geniculate nucleus [Tyll et al., 2011]. It therefore appears that, although there are species- and modality-specific differences, subcortical structures support at least in part multisensory integration processes.

### 1.3.2 Cortical multisensory integration

Research into multisensory integration has also focused on cortical areas. Building on the aforementioned descriptions of unisensory processing pathways terminating in unisensory primary sensory cortices (i.e. sensory cortical parcellation), Wallace and colleagues repeatedly used electrophysiological recordings to assess responses of neurons to auditory, visual, somatosensory and multimodal stimuli across various sensory areas of the rat cortex [Wallace et al., 2004]. In areas commonly designated as unisensory (i.e. A1 or S1) they found a low proportion of multisensory neurons (i.e. exhibiting enhanced responses to multimodal stimuli). However, they identified a larger proportion of multisensory-tuned neurons in higher sensory areas, typically between different sensory cortices, which they subsequently defined as transitional multisensory areas. More recently, Olcese and colleagues focused on such an area between S1 and the primary visual cortex (V1) of the mouse [Olcese et al., 2013]. They found a high proportion of both unimodal and bimodal-tuned neurons in the region and more specifically found that the largest multisensory response enhancement was achieved when presented with non-preferred unisensory stimuli (i.e. stimulus parameters that did not elicit the strongest response), another example of the principle of inverse effectiveness.

I have so far presented a model for sensory information processing in which multisensory integration occurs at subcortical stages (mainly the SC and thalamus) and higher associative cortical areas. In such a model, primary sensory areas contain only neurons sensitive to a single modality (i.e. tactile-tuned neurons in vS1). However, further research has revealed that the picture may not be this clear-cut and multisensory integration does in fact occur as early as primary sensory cortices.

### 1.3.3 Multisensory integration in primary sensory cortices

A first hint to this observation was made by Bizley and colleagues in 2007 [Bizley et al., 2007]. They recorded from different fields in the auditory cortex of ferrets, including A1 and other fields, such as anterior dorsal or anterior ventral fields, that had been identified as higher sensory areas. They recorded spikes in neurons while animals were presented with auditory, visual and audiovisual stimuli. Depending on their responses (characterised by spike rates), neurons were characterised as either auditory, visual or bimodal. Bimodal neurons were those for which responses to bisensory stimuli were significantly different from either unisensory stimuli. Over 75% of the units from which they recorded in A1 were categorised as auditory while ~15% were either visual-responsive, responsive to sound and light or responsive to sounds but modulated by light. On the other hand, in the auditory ventral field, around 50% of units were light-responsive and ~25% were responsive to bimodal stimuli. While the main argument of this study resided in the report of multisensory responses in higher sensory fields, it did nonetheless highlight the fact that a small but significant proportion of neurons in A1 exhibited multisensory responses.

It is worth noting here that the aforementioned findings were collected in anaesthetised animals. In comparison, a more recent study investigated the nature of audiovisual modulation in V1 of awake mice [Meijer et al., 2017]. They used two-photon imaging to measure neural responses to auditory, visual and audiovisual stimuli. They found that just under 10% of the imaged neurons were responsive to tones, confirming Bizley and colleagues' previous findings of crossmodal responses in A1 were also verified in V1. They also compared responses of neurons to congruent

audiovisual stimuli (i.e. same temporal frequency of auditory and visual stimuli) and incongruent stimuli. They found that, while some did not change response in multimodal conditions compared to unimodal (roughly 25% of neurons tuned to visual stimuli), neurons responsive to visual stimuli exhibited greater enhancement when the audiovisual stimuli were congruent than incongruent. This observation is consistent with behavioural reports of enhanced performance in multisensory tasks using congruent stimuli and enhanced neural representation of congruent stimuli compared to incongruent ones [Kim et al., 2008, Atilgan et al., 2018], further discussed in Section **2.3.3**. Such a finding makes a lot of sense from a behavioural perspective. Indeed, self-motion may alter auditory and visual information, as we typically experience when peering out of a moving car or train. In such a situation, stimuli from different modalities would be coherently affected by motion and yet remain highly relevant for an animal to navigate its environment.

How can we explain the discrepancy in findings suggesting minimal multisensory integration in primary sensory cortices in some studies while others highlight complex, congruence-dependent integration in similar areas? Studies finding multisensory-responsive neurons have highlighted the importance of congruence of unisensory cues on their multisensory responses and, as early as 2007 [Lakatos et al., 2007], hinted at a role for neuronal oscillations in enabling or suppressing integration. Lakatos and colleagues used multielectrode recordings in awake macaque monkeys to probe auditory-somatosensory integration in A1. They presented brief auditory clicks and mild electrical stimulation at the wrists of the monkeys. Their recordings highlighted two main components of multisensory integration in A1. First, they found that the amplitude of the enhancement was greatest at lower auditory intensities, consistent with the numerous aforementioned

reports of the principle of inverse effectiveness. Second, the amplitude of enhancement varied depending on the stimulus onset asynchrony between auditory and somatosensory stimuli. Unsurprisingly, the largest enhancement occurred when the stimuli were simultaneously presented. However, they also noted asynchrony ranges corresponding to neural oscillations ($\gamma$-, theta- and delta-bands) at which multisensory enhancement proved significant. Hence, while it typically involves less than 25% of all imaged or recorded cells, multisensory integration does appear to occur as early as primary sensory cortices. More importantly, it is already clear that this integration is dependent on numerous factors, both internal (i.e. neural oscillations) and external (i.e. cue reliability, timing of respective onsets).

Most studies discussed so far have illustrated essential principles of multisensory integration in situations in which individual modalities may help inform on one another (for example ambiguous auditory and visual information during motion may be resolved through integration). As noted in Section **1.1.4**, these may therefore inform on neural principles underlining multisensory enhancement when one or both modalities are presented around an animal's perceptual limits.

### Audiotactile integration in the primary somatosensory cortex

As previously mentioned in Section **1.1.1**, audiotactile integration can lead to striking illusions, including that of drier skin [Jousmäki and Hari, 1998]. Several studies in human participants have investigated the nature of audio-tactile integration in somatosensory areas [Foxe et al., 2000, Pannunzi et al., 2011, Young et al., 2017]. As early as 2000, Foxe and colleagues reported auditory-somatosensory interactions in early sensory cortices [Foxe et al., 2000]. They recorded from scalp electrodes placed on participants presented with auditory-alone stimuli, median nerve stimulations

(somatosensory stimuli) or combined somatosensory-auditory stimuli. They found a significant voltage waveform difference between combined stimuli and the sum of unisensory stimuli alone in a scalp location close to the postcentral gyrus, corresponding to early somatosensory cortical areas. These findings constitute a first hint toward a role for S1 in audiotactile integration.

The main senses used by mice in their natural exploratory behaviours are touch and audition. It remained unclear how they are integrated in primary sensory cortices, such as vS1. However, a study recently investigated this integration, finding a modulation of responses to tactile stimuli when sounds were added [Zhang et al., 2020] but no auditory responses from S1 neurons. Similarly, Maruyama and Komai found a modulatory effect of sound presentations on S1 responses in rats [Maruyama and Komai, 2018]. Briefly, they performed whole-cell recordings of S1 neurons and found that sound presentations elicited subthreshold postsynaptic responses – that is, while sounds alone may not have been sufficient to elicit a response, they undeniably play a modulatory role on networks in S1.

Therefore, there is an extensive body of research investigating the correlates of multisensory integration in primary sensory cortices. While there have probably been fewer studies focusing on auditory-somatosensory interactions in S1, more recent evidence collected in mice suggests that it is indeed present.

### 1.3.4 Circuits underlying multisensory integration

In a set of recent studies, it has become apparent that a convergence of cortico-cortical inputs in primary sensory cortices may in part support multisensory processes [Teichert and Bolz, 2018]. First, anatomical tracing studies revealed the

existence of robust direct cortico-cortical connections between A1, V1 and S1 in monkeys [Cappe and Barone, 2005, Cappe et al., 2009, Meredith and Lomber, 2017], ferrets and rodents [Henschke et al., 2015, Teichert and Bolz, 2018, Iurilli et al., 2012]. In monkeys, some thalamic nuclei have been found to project to several primary sensory cortices [Cappe et al., 2009], prompting the suggestion that multisensory integration may be both a feedback and feed-forward process. As such, it would combine both projections from higher sensory areas and association cortices as well as subcortical inputs already containing multisensory information [Foxe and Schroeder, 2005].

Second, a recent study led by Ibrahim and colleagues [Ibrahim et al., 2016] revealed the laminar organisation of multisensory integration in V1 of mice. They recorded from V1 neurons while awake mice were presented with drifting sinusoidal gratings with some trials containing added white noise pulses. Neurons in L2/3 exhibited typical orientation preferences, characterised by increased responses around a preferred orientation. However, when sounds were added to stimuli, responses to preferred stimuli were enhanced while those to non-preferred orientations were suppressed. In comparison, L4 neurons did not show this tuning sharpening effect showing in fact identical response patterns to stimuli in both visual and audiovisual conditions. Finally, they imaged L1 neurons in the same conditions and observed that, while these exhibited similar tuning preferences as in other cortical layers, their responses were enhanced across all orientations with added sounds. In fact, sound alone was sufficient to excite L1 neurons. L4 neurons are the main target of thalamic input while L2/3 neurons typically receive cortico-cortical inputs. Based on previous work [Iurilli et al., 2012] and their own findings, Ibrahim and colleagues confirmed the presence of A1 inputs in L1 of V1. With this in mind, they optogenetically silenced

A1 neurons while recording from L2/3 V1 cells. They found that silencing inputs from A1 into L1 of V1 suppressed the sharpening of tuning observed when presenting audiovisual stimuli. The authors concluded by theorising a cortico-cortical circuit in which inhibitory L1 interneurons receive auditory inputs and inhibit L2/3 neurons more strongly when presented with a sound and a visual stimulus at their preferred orientation than when combined with a stimulus at a non-preferred orientation, subsequently sharpening the tuning of downstream L2/3 neurons.

It is worth noting that these direct cortico-cortical connections are not all symmetrical in rodents [Henschke et al., 2015]. For instance, primary cortico-cortical projections from V1 predominantly target S1 and only faintly A1, although it receives similarly dense projections from A1 and S1. In the case of S1, there are more neurons projecting to and from V1 than A1, and there are more projections from A1 to S1 than S1 to A1. Correspondingly, Iurilli and colleagues [Iurilli et al., 2012] found that different modalities did not lead to the same cross-modal response across primary sensory cortices. For example, they found that auditory stimuli hyperpolarise S1 neurons while visual stimuli elicited a depolarisation. In contrast, A1 neurons were not modulated by visual stimuli, but were depolarised by tactile stimuli. These findings, combined with aforementioned reports of cross-modal responses in tactile, but not visual thalamus [Bieler et al., 2018], suggest that the circuits underlying multisensory integration in primary sensory cortices are modality-specific. For each pair of modalities being integrated, the relative contribution of thalamo-cortical and cortico-cortical inputs may vary. These findings were reported in mice and, in agreement with the differences between [Bieler et al., 2018] and [Tyll et al., 2011], may reflect species-specific processes at play.

It is therefore clear that multisensory integration occurs within primary

sensory cortices in a species- and modality-specific manner. It is also understood that vS1 receives inputs from other primary sensory cortices. Based on studies in other areas, such as V1 [Ibrahim et al., 2016], recordings from cortico-cortical inputs have even been characterised, highlighting a role beyond simply receiving inputs that had already been integrated subcortically.

However, it remained unclear how training on a multisensory task and binding modalities in a behaviourally relevant fashion (i.e. stimuli that are meaningful to an animal) may alter the representation of multisensory stimuli at the primary sensory cortical level. In particular, it is not clear how training affects audio-tactile integration in vS1 of mice. While seminal findings may hint toward the nature of responses and cross-modal influences [Zhang et al., 2020, Maruyama and Komai, 2018, Iurilli et al., 2012], questions remain with regards to potential effects of training and the circuits underlying them.

### 1.3.5 Effects of training on neural responses

In recent years, the conjoined development of two-photon imaging and novel transgenic mouse lines has allowed for novel procedures to track the effects of training and practice on neural responses in behaving animals [Margolis et al., 2014, Peron et al., 2015, Minderer et al., 2012]. Most notably, using two-photon imaging allows for the imaging of hundreds, sometimes thousands, of neurons at the same time [Pachitariu et al., 2016]. Using it chronically in transgenic mouse lines allows for better tracking of individual cells over tens of days during training than electrophysiological methods [Margolis et al., 2014, Tolias et al., 2007]. Taken together, these two factors allow for improved tracking of up to hundreds of individual cells across days. Further, the imaged mouse lines may express genetically encoded

calcium indicators under identified promoters, allowing for the unbiased targeting of specific cell types [Chen et al., 2012, Chen et al., 2013b, Wachowiak et al., 2013].

Several studies have investigated how neural responses to stimuli changed with training [Peron et al., 2015, Rutkowski and Weinberger, 2005, Poort et al., 2015]. Poort and colleagues [Poort et al., 2015] used two-photon imaging to track V1 neurons while mice were training on a virtual-reality based visual discrimination task. In this study, they quantified each neuron's tuning selectivity as a measure of the differences between average responses to different grating angles. They found that more neurons were selective toward the end of training than the beginning. Moreover, they found that tuning selectivity became more stable (i.e. less likely to switch or be lost from one day to the next) as mice trained. In fact, they found a significant correlation between behavioural performance and population selectivity. Neural responses reflected more than sensory inputs though, as some responses were modulated by running while others reflected behavioural outputs rather than sensory inputs alone. For example, some cells responded more strongly to false alarms than correct rejections – two trials with identical stimuli but different outcomes. Such a finding is not isolated in fact and several studies have reported enhanced responses to conditioned stimuli in A1 [Rutkowski and Weinberger, 2005] or as an other example, changes to visual tuning in V1 monkeys [Schoups et al., 2001]. Similar findings were made in vS1 of mice, such as those reported by Peron and colleagues [Peron et al., 2015]. They studied neural responses to a pole's azimuthal position as mice trained on a spatial discrimination task. Specifically, they tracked responses from individual cells and found that, as training went along, a single cell's responses to stimuli predicted the next day of training's responses more accurately. However, in contrast with Poort and colleagues' findings, they reported high stimulus

discrimination ability based on neural data alone, even prior to training, suggesting that stimulus information is reliably encoded in vS1 regardless of behavioural performance. These findings were consistent with previous reports from the same research group [O'Connor et al., 2010c].

## 1.4 General aims

Several studies have combined sensory-guided training paradigms with chronic imaging in order to assess how experience and expertise may affect sensory responses at a primary sensory cortex level. However, to my knowledge, no study has yet investigated the nature of multisensory stimulus representation during training.

Specifically, as mentioned previously, there is some very early work assessing the nature of audio-tactile integration in vS1 of mice [Zhang et al., 2020]. Evidence for such integration is still scarce and has not yet assessed how it changes with training.

Furthermore, most research into multisensory integration has investigated the neural basis for multisensory enhancement. It remains unclear whether sensory binding, in which different unisensory components of a multisensory stimulus contribute to a broader multisensory percept [Deroy et al., 2016] occur at early cortical stages.

Therefore, this Thesis aimed to address three main aims:

**1**. I first set out to design a task in which readily discriminable auditory and tactile stimuli are paired to create a multisensory percept (Chapter **2**).

**2**. I then wanted to characterise the nature of multisensory integration in vS1 of mice using the stimuli designed in the aforementioned task (Chapter **3**).

**3**. I aimed to combine the two first objectives in order to assess how training on a multisensory task may alter the representation of multisensory stimuli in vS1 of mice (Chapter **3**).

## 1.5 General hypotheses

Based on the aforementioned literature on multisensory integration in general and audio-tactile integration in mice more specifically, I made several key hypotheses which my experiments would test.

First, I made predictions on the ability of mice to learn to discriminate between multisensory stimuli. Under appropriate control, I expected that mice could be trained to categorise audio-tactile combinations of stimuli [Meijer et al., 2018, Siemann et al., 2015]. I more specifically expected mice to take longer to learn the multisensory identification task than a unisensory version of the same task. [Juavinett et al., 2018]. Based on previous findings in mice, I also predicted that their discrimination performance would improve within individual sessions [Berditchevskaia et al., 2016].

Second, previous studies of audio-tactile integration in rodent vS1 [Maruyama and Komai, 2018, Zhang et al., 2020, Iurilli et al., 2012], found an inhibitory modulating effect [Maruyama and Komai, 2018] and hyperpolarisation of neurons following sound presentations [Iurilli et al., 2012]. I hypothesised that I would observe an inhibition of vS1 neurons when presented with auditory stimuli.

Based on the work by Bieler and colleagues [Bieler et al., 2018], I postulated that the thalamic inputs into vS1 would include multisensory information although this lies beyond the scope of the current Thesis. While I hypothesised that the main effect of auditory tones on vS1 responses would be inhibitory, I also predicted, given findings in other modalities [Ibrahim et al., 2016], that tones may sharpen a neuron's tuning. This would translate into a proportion of neurons responding specifically to precise audio-tactile combinations. If I did not find responses specific to a single pair, and instead responses to all stimuli with similar tactile identities,

Third, I predicted some effects of training on neural responses to stimuli. Specifically, Peron and colleagues found that when a mouse is trained on a tactile pole position discrimination task, a vS1 cell predicted its own future responses to task stimuli more accurately with advanced training [Peron et al., 2015]. However, they found that vS1 neuron responses were sufficient to discriminate between stimuli from the very beginning of training, suggesting that training did not enhance sensory representation as much as stabilise neural responses to stimuli. Therefore, I did not expect to find changes in the responsiveness or overall proportions of responsive neurons in vS1 with training. Instead, I predicted that a single neuron's response to stimuli would stabilise (i.e. remain responsive or unresponsive) over the course of training. In contrast, changes in responsiveness that could be observed include some neurons becoming responsive during training, others becoming unresponsive or developing long-lasting changes beyond the course of training, such as changes in response onset times. If such training effects are observed, they could reflect either changes in local circuitry or a change in the weights of inputs from thalamic nuclei or other cortical areas. Based on Peron and colleagues' aforementioned reports of high neurometric performance prior to training, I did not expect training to emphasise

stimulus response specificity, since stimuli could reliably be decoded from the onset. However, regarding multisensory stimuli specifically, I expected that training may reinforce cross-modal inputs and increase the proportion of neurons responsive to auditory and audio-tactile stimuli.

# Chapter 2

# Performance of mice on a novel multisensory learning task

## 2.1 Introduction

### 2.1.1 Overview

We live in a truly multisensory world, full of ambiguous and noisy inputs. At any instant, we combine information across senses. For example, the pitch and loudness of an engine sound combined with changes in the size and position of the car within our visual field help inform us of whether it is speeding up, and consequently guide our decision to cross the road. The integration of information across different modalities appears to be such a constant and automatic process that we usually do not notice it, even when information from our different senses 'clash'. A well studied example is the McGurk effect, described by Harry McGurk and John MacDonald [McGurk and Macdonald, 1976] (see Section **1.1.1**). In their experiments, subjects observed a speaker mouth a different syllable than the audio track of their voice, causing subjects to poetically 'hear lips and see voices'. Subjects perceived either words fusing both visual and auditory streams (i.e. perceiving the syllables 'da-da' when seeing 'ga-ga' and hearing 'ba-ba'), or new combination non-words (i.e. 'gabga').

Most studies of sensory processing in animals have presented stimuli that are near the threshold of perceptual detection, in order to measure just-noticeable differences or detection thresholds [De Hoz and Nelken, 2014, Meijer et al., 2018, Bizley et al., 2013]. Such stimuli include tones of varying amplitude or visual gratings at different contrast levels [Meijer et al., 2018]. Below that threshold, animals can no longer reliably detect stimuli at an above-chance level. In that regard, the few studies of multisensory integration that have been carried out in mice have typically focused on such thresholds and investigated whether cross-modal cues may help improve detection – a process called multisensory perceptual enhancement [Meijer et al., 2018]. However, despite this focus in the research to date, multisensory integration can serve other purposes. For example, we regularly rely on cross-modal combinations of features, for example touch, vision and smell in order to pick fruits or vegetables. Fewer studies have examined how we learn to categorise objects in this manner, based on their combinations of features across multiple senses. I will refer to this process as 'multimodal combination identification'. In this process, one modality does not necessarily improve the sensory acuity of another, but rather an object in our environment is identified reliably only when multiple sensory features are detected together.

Regarding audio-tactile integration more specifically, behavioural studies have already established strong interactions between auditory and tactile information in human subjects, characterised notably by the 'parchment-skin' illusion [Jousmäki and Hari, 1998]. In this study, experimenters played back recordings of subjects rubbing their palms together with accentuated or dampened high frequencies, which altered the perception of dryness reported by participants. A later study found similar effects in texture perception of abrasive surfaces [Guest et al., 2002].

However, while these examples highlight the interactions between auditory and tactile modalities and the resulting multimodal percepts they generate, there is to my knowledge no existing literature describing audio-tactile categorisation tasks that rely on multimodal combination identification, rather than multisensory perceptual enhancement.

The mouse (*Mus musculus*) is a popular animal model for imaging neural responses to sensory stimuli, in both auditory and tactile studies. Two-photon calcium imaging allows us to measure the responses of large numbers of individual neurons in the cortex simultaneously, bringing new insights into sensory processing that were not possible with microelectrode techniques alone. In the auditory domain, Rothschild and colleagues used two-photon imaging in anaesthetised mice in which they had injected the calcium indicator Fluo-4 in order to map frequency preferences of hundreds of A1 neurons [Rothschild et al., 2010]. This revealed the fine-scale heterogeneity in the tonotopic mapping of A1 not observed previously using electrophysiological recordings alone [Stiebler et al., 1997]. Similarly, Cre-dependent targeting of auditory cortex neurons projecting to V1 has allowed for the characterisation of projections between sensory cortices and cross-modal responses in V1 [Deneux et al., 2019]. Several studies originating from Karel Svoboda's group have imaged neural responses in somatosensory cortex to poles being placed at various azimuthal positions relative to mice whiskers [O'Connor et al., 2010c, Guo et al., 2014, Peron et al., 2015]. The combination of behavioural training on a tactile discrimination task and imaging has further enabled the tracking of responses of individual neurons in awake, behaving animals [Li et al., 2016, Peron et al., 2015, Ceballo et al., 2019]. Auditory stimuli can vary in their location, loudness, timbre and perceived pitch. Mice can learn to discriminate sounds based on these features,

for example between white noise and pure tones [Siemann et al., 2015], and different sound frequencies [De Hoz and Nelken, 2014]. Similarly, tactile stimuli have several features, including location, frequency and amplitude. Rodents can use their whiskers to localise objects [Guo et al., 2014, Peron et al., 2015], or to discriminate between textures [Arabzadeh et al., 2005, Chen et al., 2015]. Many studies have used Go/No-Go tactile discrimination tasks to investigate neural responses to either an object's position relative to a mouse's whiskers [O'Connor et al., 2010a, O'Connor et al., 2010b], or to different tactile textures [Chen et al., 2016]. Mice have also been shown to discriminate between vibrotactile stimuli presented with piezoelectric drivers on single threaded whiskers bilaterally [Mayrhofer et al., 2013]. Mice can also reliably combine a visual detection task with an auditory discrimination task [Wimmer et al., 2015] although, to my knowledge, no published study has trained mice on an audio-tactile discrimination task.

While a variety of tasks and stimuli have been used to examine somatosensory processing in mice [Helmchen et al., 2018, Mayrhofer et al., 2013, Sofroniew et al., 2015], one that has been adopted in several recent studies [Peron et al., 2015, Peng et al., 2015, Li et al., 2016] is a Go/No-Go task originally described by Guo and colleagues [Guo et al., 2014]. Their task required head-fixed mice to discriminate between poles presented at different azimuthal positions along their whisker pads. Mice were trained to lick for water rewards when the pole was presented at one of two possible positions. Licking on No-Go trials (false alarm (FA)) generated a 2-5s time-out. The head-fixed design is ideal for simultaneous two-photon calcium imaging, and mice readily learn this simple task [Guo et al., 2014]. I therefore began my behavioural experiments by implementing a similar task design, with the aims of developing it into a multisensory (audio-tactile) categorisation task (Section

**2.2.5**).

In previous discrimination tasks in mice, it has sometimes been unclear which modality the animal is relying on for its discrimination decisions. For example, one study investigated multisensory combination categorisation by training mice to respond at either a left or right lick port for a reward depending on a visual spatial cue (a light turning on above the corresponding port) and an auditory timbre cue (either a pure tone or white noise) [Siemann et al., 2015]. Combining auditory and visual stimuli enhanced discrimination performance compared to unimodal stimuli. However, the auditory and visual stimuli were always presented in the same pairs (i.e. the white noise was always presented with the visual cue over the right lick port), so these findings do not demonstrate conclusively that mice learned to categorise multisensory cues. Indeed, mice could learn to categorise visual cues and ignore the auditory stimuli altogether (or vice-versa). The enhanced performance in multisensory conditions compared to unisensory sessions could be explained by the addition of independent information conveyed by another sensory cue, overall increasing the total amount of available information, although not teaching mice that a particular combination would be rewarded. If the experimenters had added 'incongruent' trials, in which the auditory and visual cues were presented in 'incorrect' pairs, they could have tested if mice learned multisensory combinations.

This lack of stimulus controls across modalities is evident in a number of studies. More recently, a study investigating tactile responses in V1 of mice [Kandler et al., 2018] failed to replicate observations of neural responses to tactile stimuli after changing their recording set-up. Responses were originally collected by means of two-photon imaging in a visual and tactile virtual reality environment and replicated in total darkness, suggesting that the mice could in fact see the tactile

stimuli in the first instance. In another recent study investigating audio-tactile integration in the somatosensory cortex of mice [Zhang et al., 2020], experimenters used two-photon imaging to measure neural responses in mice passively presented with vibrotactile stimuli, sinusoid amplitude modulation (SAM) tones and combined audio-tactile stimuli. The authors recorded sounds from the piezoelectric driver they used to deflect whiskers at frequencies above 140 Hz. They found significant acoustical energy at resonant frequencies over 3kHz, which were within the mouse hearing range [Heffner and Heffner, 2007], so the neural responses to tactile stimuli may in fact have been to audible sounds. These examples highlight the need for a multisensory training protocol in which the tactile and auditory stimuli are clearly independent and free from cross-modal artifacts.

Here, I aimed to not study the limits of perceptual discrimination, but the ability of mice to learn to categorise a novel multisensory object. Therefore, I used stimulus values that are readily distinguishable in order to avoid making the task perceptually difficult for the mice. This corresponds to a large frequency difference between tones in the auditory domain, for example, or widely spaced azimuthal positions of a pole presented to the whiskers. This should assist learning, as mice have been found to train more slowly on discrimination tasks with similar stimuli than with more dissimilar pairs [Treviño et al., 2013]. By combining pairs of auditory and tactile stimuli as multisensory objects to be identified, I trained mice on a task that required cross-modal combination learning. The inherent difficulty in the task is the need to recognise the coincidence of specific stimulus values across different modalities. Each modality could therefore not be used in isolation to solve the task, unlike most previous studies of multisensory integration in animals [Siemann et al., 2015, Meijer et al., 2018, Angelaki et al., 2009].

### 2.1.2 Objectives

The work presented in this chapter aimed to:

1. Design and implement a simple behavioural task which could be used to train mice to categorise stimuli based on any stimulus feature, or combination of features.

2. Develop a tactile stimulus on the above task avoiding potential auditory or visual confounds.

3. Train mice to categorise stimuli on this task based on either an auditory feature, a tactile feature, or a combination of auditory and tactile features.

To my knowledge, mice have not been previously trained on a truly multisensory categorisation task. Given the plethora of examples of mice successfully training on unisensory and multisensory discrimination tasks, and the importance of this task for everyday survival, I predicted that these animals can learn to categorise audio-tactile pairs.

## 2.2 Methods

### 2.2.1 Subjects

All animal procedures were approved by the local ethical review committee and performed under license from the UK Home Office, in accordance with the Animals (Scientific Procedures) Act of 1986.

Nine male mice from a GCaMP6f reporter line were trained on the pole position task and a further five underwent two-photon imaging while presented

with a pole at different azimuthal pole positions. All mice used in this thesis had a C57BL/6J background. The GCaMP6f reporter line was generated by mating the floxed Ai95(RCL-GCaMP6f)-D line (Jackson Laboratories; stock number 024105) with the CaMKIIalpha-Cre T29-1 line (Jackson Laboratories; stock number: 005359). A further 22 mice (nine females) were trained on the vibrotactile ($N = 6$), auditory ($N = 7$) and audio-vibrotactile ($N = 11$) tasks. Two mice trained on the audio-vibrotactile task had been previously trained on either the auditory or vibrotactile task. These were obtained from GCaMP6s reporter line, generated by mating the floxed Ai94(TITL-GCaMP6s)-D line (Jackson Laboratories; stock number 024104) with the CamKII-tta (Jackson Laboratories; stock number 003010).

### 2.2.2 Surgical procedures

#### Headplates

Mice were implanted with custom-built titanium headplates for head-fixation. These flat implants allowed for head-fixation from both sides of the animals' heads for increased stability during imaging. Anaesthesia was induced using isoflurane (4% in 1.5 L/min $O_2$) and the animals were then placed into a stereotaxic frame (Kopf Instruments). Anaesthesia was maintained at 1-2% isoflurane in 0.5 L/min $O_2$. At the start of surgery, analgesia was induced by subcutaneous injections of meloxicam (2 mg/mL at 5 mg/kg) and buprenorphine (0.03 mg/mL at 0.1 mg/kg). Topical analgesia was induced under the scalp (marcaine 0.25% solution at 2 mg/kg). The temperature was maintained by means of a thermal blanket (FHC) and the mice's eyes were protected by a layer of eye ointment (Lacri-lube; Alergan, UK). A flap of skin was cut off and the exposed skull was cleaned (including removal of the periostium), scraped and dried. The headplates were then positioned and secured to the clean dry

skull by means of dental cement (Superbond C&B; Prestige Dental). After surgeries, mice were allowed at least seven days to recover prior to any further experimental manipulation (Figure **2.1 B**).

#### Chronic imaging windows

Five mice were imaged in passive conditions when presented with pole position stimuli (i.e. outside of any training protocol). Further, three mice trained on the tactile and audio-vibrotactile tasks as well as one mouse trained on the auditory task were concurrently imaged during training. During surgery, these mice also underwent craniotomies in order to allow for the implantation of chronic imaging windows.

Custom imaging windows were designed by stacking one 3 mm wide and one 4 mm wide round coverslips (total thickness 0.45 mm), kept together by means of an ultraviolet-cured optical glue (NOR-81; Norland, Edmund Optics). A 3 mm craniotomy was made over vS1 in the right hemisphere (position of centre relative to Bregma: 1.3 mm posterior and 3.2 mm lateral). An imaging window was placed, with the 3 mm wide stack fitting in the craniotomy and the 4 mm wide coverslip allowing it to be glued to the bone. The stacked coverslip replaced the bone and protected the brain from potential sources of infection or damage.

### 2.2.3 Behavioural habituation phase

Each mouse was first water-regulated and handled for three to five days in order to stabilise its weight around 85% of its initial weight and habituate it to handling. It was handled and given ~1-1.5 mL of water every day during this period. It was then habituated to the testing apparatus for the following three to five

days before training began. Each mouse was head-fixed on the wheel for gradually increasing amounts of time (ranging from a few seconds on the first day to up to five minutes on the last day of habituation) then given water after being taken off the wheel.

### 2.2.4 Reward-stimulus association phase

During the second stage of habituation, mice went through a reward-stimulus association phase. On each day, they were head-fixed on the wheel and a water delivering spout was placed within reach for them to lick. They could initiate trials by running for a second, although trials were also automatically generated every 20 s if they failed to do so. One of the stimuli, the Go stimulus, was automatically rewarded after every presentation (delivery of a 15-20 $\mu$L drop of water from a water spout positioned directly in front of the mouse's mouth). If the mouse licked on No-Go stimulus presentations (i.e. a false alarm (FA)), no reward or punishment were delivered.

### 2.2.5 Training phase

After three to five days of reward-stimulus pairing, mice started the training phase. A Hit was defined as licking within the 3 s response window following presentation of the pole in the Go position. A Correct Rejection (CR) was an omission of licking during the response window following presentation of the No-Go position. A Miss was a failure to lick within the response window on a Go trial, and a FA occurred when the mouse licked during the response window on a No-Go trial. Hits resulted in a ~15-20 $\mu$L water reward. Unless specified otherwise (see **2.3.1**), FAs

were followed by a 3 s time-out, during which no trial could be elicited. In order to avoid mice developing a strategy of continuous licking, the time-out restarted if they licked before the 3 s had gone. CRs and Misses did not lead to either a time-out or a reward delivery (Figure **2.1 C**).

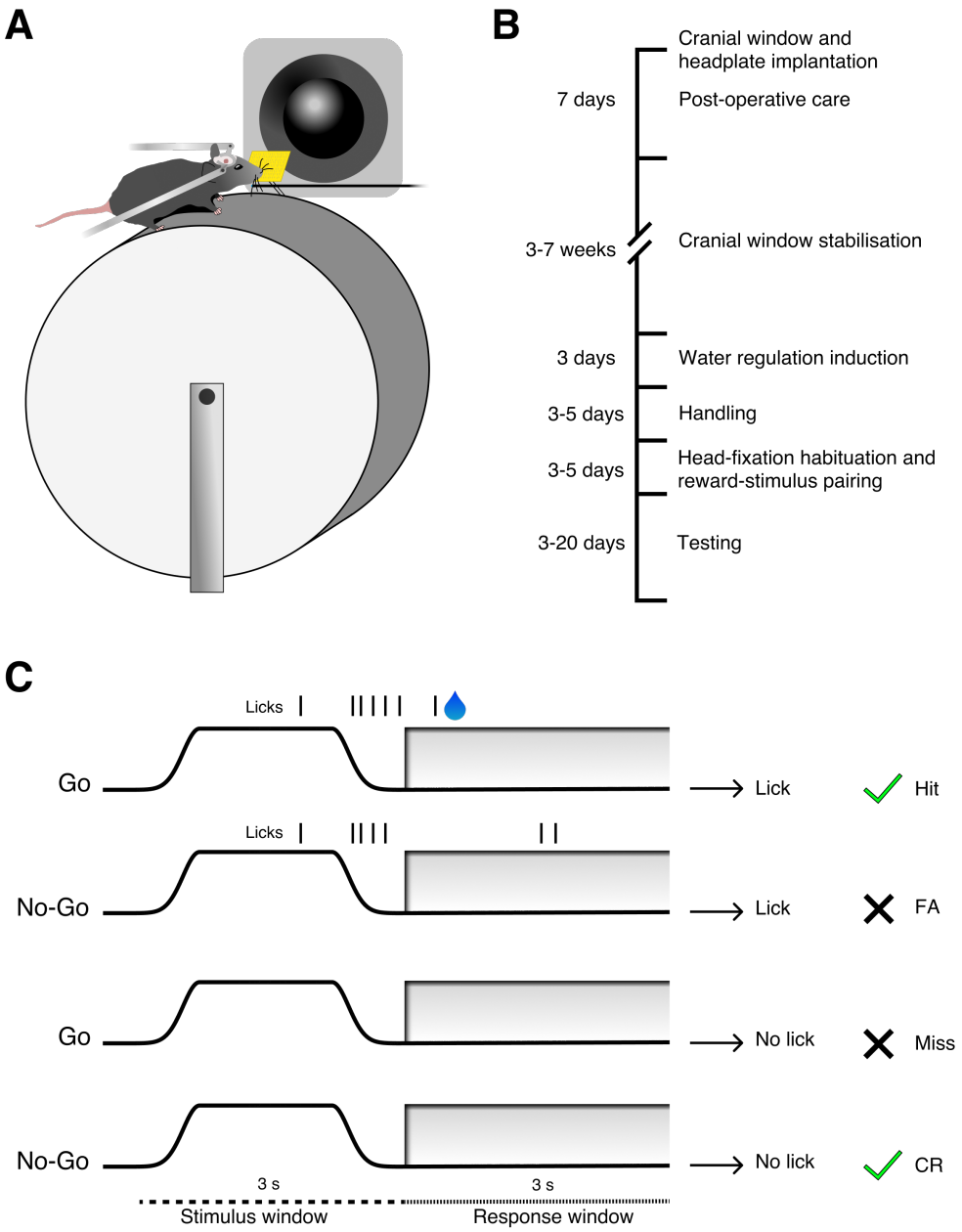


**Figure 2.1:** ***Figure on previous page.*** **A.** Behavioural training setup for auditory, vibrotactile and multisensory tasks. Mice were head-fixed into position on a running wheel. They were exposed to vibrotactile (Velcro pad drawn in yellow) and auditory (loudspeaker drawn in grey) stimuli coming from their left side. Mice could lick the spout located in front of them (shown in black) for water rewards delivered through the same spout. **B.** Summary of procedures from the cranial window implant until the end of training. **C.** Go/No-Go trial schematic. The time-course of Go and No-Go trials are shown. The thick black line represents the amplitude of the stimulus (0-250 $\mu$m for vibrotactile or 0-70 dB $SPL_C$ for auditory stimuli). Times of example licks are indicated as vertical lines above the stimulus trace. Licking during the response window (grey box) on a Go trial (a Hit) led to a reward delivery (indicated as a blue water drop). Failing to lick during this window on a Go trial was a Miss and resulted in no reward or time-out. On a No-Go trial, licking during the response window (a FA) led to a time-out, and withholding licking (a CR) led to no reward or punishment.

### 2.2.6 Pole position stimuli

In the pole position discrimination task, mice were trained to discriminate between two azimuthal positions of a pole presented unilaterally to their their left whiskers.

The two positions were anterior and posterior along the mouse's whiskers, separated by 7.2 mm. The Go position was either the anterior or posterior position for different mice but was always the same for each mouse. On each trial, the pole moved into position against the whiskers from 1.4 cm in front of the mouse's head using a rotating permanent magnet stepper motor (RS Pro). It followed the same path for all trials but stopped earlier when presented in the anterior position. The pole was kept in position for 1 s, and was then retracted away from the whiskers back to the front starting position.

### 2.2.7 Vibrotactile stimuli

In previous studies, it was shown that rats combine the frequency and amplitude of vibrotactile stimuli in a discrimination task [Arabzadeh et al., 2003, Adibi et al., 2012]. Most notably, these studies found that rats could not reliably discriminate between two different vibrotactile stimuli if the products of their amplitudes and frequencies were equal. In these tasks, the whiskers of rats were positioned either through a custom-built ladder or against a mesh. In the present study, I aimed to design a task that could be readily run simultaneously on different setups by a single experimenter. In order to achieve this, I did not want to thread each animal's whiskers daily into a ladder, so I instead designed a setup similar to the mesh stimuli against which the whiskers could rest and be passively vibrated.

I presented vibrotactile stimuli using a permanent magnetic shaker (LDS V201, Brüel&Kjær) fitted with a cap screw (SH6MS45; Thorlabs) and a Velcro pad (VWSE40724; Sellotape). The shaker was mounted on a damped post (DP14A/M, Thorlabs) to isolate the rest of the set-up from disruptive vibrations (Figure **2.1 A&C**). The stimuli consisted of 3s long vibrations at 93 or 129 Hz. These frequencies are well below the mouse hearing range [Heffner and Heffner, 2007]. The vibrotactile stimuli were 3 s in duration, including 1 s onset and offset ramps following a Gaussian function, with a peak amplitude of 250 $\mu$m . The ramps avoided auditory confounds at their onsets and offsets (Figure **2.1 C**).

### 2.2.8 Auditory stimuli

Auditory stimuli were presented through a pole-mounted speaker (Visaton FRS8 4Ω, Reichelt Electronics) positioned on the same pole as the aforementioned

shaker. Sounds consisted of either a 6.3 kHz pure tone carrier with 93 Hz (sinusoid amplitude modulation (SAM)), or a 12.5 kHz pure tone with 129 Hz SAM. The tones where 3 s in duration, with 1s onset and offset ramps, and a peak amplitude of 70 dB $SPL_C$ (Figure **2.1 A&C**). The tones were modulated by frequencies used in vibrotactile training conditions to increase the congruence between the auditory and vibrotactile components of multisensory stimulus pairs (See Section **2.2.9**).

### 2.2.9 Multisensory discrimination stimuli

The multisensory version of the task used the same vibrotactile (93 and 129 Hz) and auditory stimuli (6.3 and 12.5 kHz) as described above. For each stimulus presentation, one of the four possible pairs of audio-vibrotactile stimuli was presented. For each mouse, only one of the four presented stimuli was rewarded as the Go stimulus.

A study on human participants has shown that performance on a motion detection task was improved in temporally congruent audiovisual conditions compared to incongruent audiovisual conditions [Kim et al., 2008]. Further, a more recent study [Atilgan et al., 2018] has shown that neurons in the auditory cortex of ferrets respond preferentially to an auditory stream that is synchronised with a visual stream. Therefore, I encouraged multisensory fusion in my task by choosing Go stimuli with congruent amplitude modulation frequencies in the auditory and vibrotactile components. The Go stimulus was either the low-frequency congruent pair (6.3 kHz tone and 93 Hz vibrotactile stimulus) or the high-frequency pair (12.5 kHz tone and 129 Hz vibrotactile stimulus), counterbalanced across mice. Stimulus waveforms are shown in Figure **2.2**.

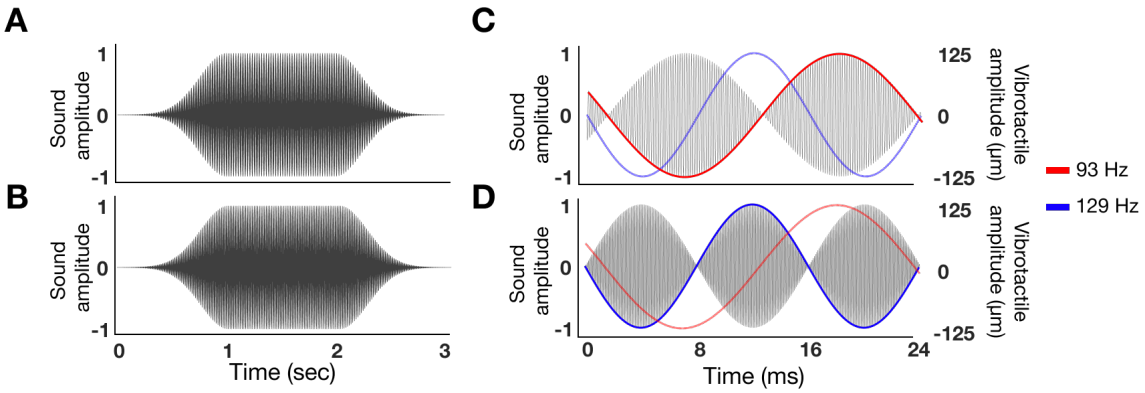



**Figure 2.2:** **A.** Waveform of the low frequency congruent auditory stimulus (6.3 kHz carrier, 93 Hz modulator). **B.** Waveform of the high frequency congruent auditory stimulus (12.5 kHz carrier, 129 Hz modulator). **C.** Snippet of the waveform shown in **A**, highlighting the cycles of amplitude modulation at 93 Hz, corresponding to the vibrotactile oscillations (bold red line) presented synchronously in congruent trials. In incongruent trials, the auditory stimulus was presented with the 129 Hz vibrotactile stimulus (blue line). **D.** Snippet of the waveform shown in **B**, highlighting the cycles of amplitude modulation at 129 Hz, corresponding to the vibrotactile oscillations (bold blue line) presented synchronously in congruent trials. In incongruent trials, the auditory stimulus was presented with the 93 Hz vibrotactile stimulus (red line).

The use of incongruent stimulus pairs as No-Go stimuli allowed me to measure potential sensory biases during training by tracking a mouse's probability of licking on a No-Go trial in which either the auditory or vibrotactile component was the same as the reinforced pair. A large bias towards one sensory modality may indicate that the mouse was disproportionate reliant on that modality. Table **2.1** illustrates the pairing of stimuli for mice trained with a low-frequency Go audio-vibrotactile stimulus.

| Auditory \ Tactile | Low | High |
|---|---|---|
| Low | Congruent Go ($A^+T^+$) | Ambiguous Auditory No-Go ($A^+T^-$) |
| High | Ambiguous Tactile No-Go ($A^-T^+$) | Congruent No-Go ($A^-T^-$) |

**Table 2.1:** Stimulus types for a mouse trained on the multisensory task with the low-frequency congruent pair set as the Go stimulus. $A^+$: Same auditory frequency as the Go pair. $A^-$ No-Go auditory frequency $T^+$: Go tactile frequency. $T^-$: No-Go tactile frequency. Only responses to the $A^+T^+$ stimulus were rewarded.

### 2.2.10 Controlling for cross-modal interference

In order to control for cross-modal confounds (e.g. mice hearing tactile stimuli) as well as confirm the precision of the frequencies delivered in both modalities, I recorded both vibrotactile and auditory stimuli using a handheld microphone (H5 Recorder, Zoom). I applied a fast Fourier transform to the recordings in order to reveal the frequencies that make up the delivered stimuli and potential resonant frequencies that may result from audible vibrotactile stimuli. There was no evidence of acoustical energy outside the intended frequencies presented (>30 dB $SPL_C$). The tactile stimuli did not generate audible resonance at frequencies within the mouse hearing range (conservatively as low as 1 kHz and as high as 85kHz, although the microphone used here could not sample above 40 kHz [Heffner and Heffner, 2007]) (Figure **2.3**).

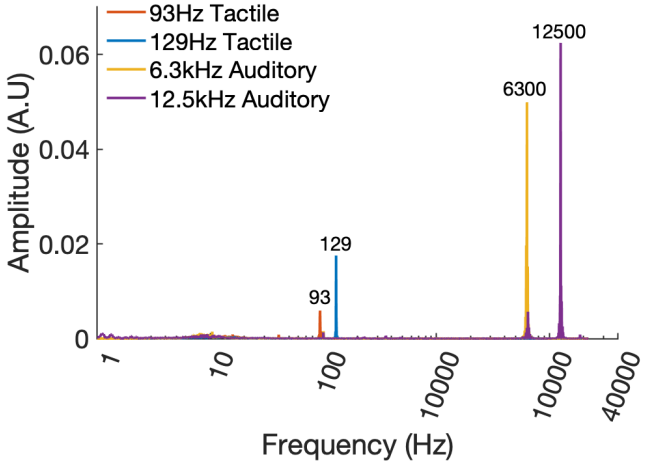



**Figure 2.3:** Frequency decomposition of both vibrotactile and auditory stimuli. Each auditory and vibrotactile stimulus was recorded in experimental conditions (i.e. the same conditions as mouse training) by means of a handheld microphone. Each coloured line depicts the frequencies that make up each stimulus, obtained through a fast Fourier transform. The weight of the frequencies making up each recording is expressed in normalised arbitrary units.

### 2.2.11 Analysing behavioural data

#### Performance

In the initial pole position discrimination task, the probability of presenting a Go trial was variable across training sessions, in an effort to motivate individual mice. For the sake of consistency and clarity, the proportion of Go trials in subsequent auditory, vibrotactile and audio-vibrotactile tasks was fixed at 50%. This did not appear to hinder the mice's ability to train on the task. Therefore, instead of expressing performance on a given session as the percentage of correct trials, I calculated a discrimination index. The discrimination index (d') takes into account both the Hit and FA rate in each session, and is less affected by the differences between Go and No-Go trials. The discrimination index (d') was defined as:

$$d' = Z(Hit) - Z(FA),$$

where Z(Hit) and Z(FA) are the inverse of the cumulative distribution function of the Gaussian distribution for the Hit and FA rates, respectively. However, in the subsequent tasks, both hits and CRs were considered correct trials. This meant that the percent of trials correctly executed and d' were equivalent measures of performance and I therefore used the more straightforward performance measure of total percent correct trials.

**Estimating sensory bias**

The multisensory task presented mice with three different No-Go trials, as described in Section **2.2.9**. This allowed me to compare their FA rates on trials with the same tactile component For each mouse, I calculated a daily sensory bias index (BI), defined as:

$$BI = \frac{\%FA_{\text{A}^+\text{T}^-} - \%FA_{\text{A}^-\text{T}^+}}{\%FA_{\text{A}^+\text{T}^-} + \%FA_{\text{A}^-\text{T}^+}},$$

where $\%FA_{\text{A}^+\text{T}^-}$ corresponds to the FA rate on all trials in which the auditory component of the stimulus is the same as that in the Go stimulus, and the tactile component is different from that in the Go stimulus. Conversely, $\%FA_{\text{A}^-\text{T}^+}$ is the FA rate on all trials in which the tactile component of the stimulus is the same as that in the reinforced pair, but the auditory component is different. Hence, a BI of 1 corresponds to a day in which all FAs on ambiguous stimuli took place on trials in which only the tactile component differed from the reinforced pair (i.e. an auditory ambiguous trial). This indicates that the animal is making response decisions based largely on the auditory component of the stimuli. Conversely, a BI of -1 corresponds to a mouse FAing on tactile ambiguous, but not auditory ambiguous, trials, indicating that decisions are largely based on tactile stimulation.

### 2.2.12 Two-photon imaging

While developing the behavioural task, it was essential to ensure that it could be run under a two-photon microscope. There is a plethora of literature illustrating that awake mice can be imaged in head-fixed conditions [Peron et al., 2015, Andermann et al., 2010, O'Connor et al., 2010a], even while they run on a stationary wheel [Poort et al., 2015, Ayaz et al., 2019]. To test the quality of imaging gathered on my set-up during whisker stimulation, I implanted cranial windows in five GCaMP6f-expressing mice, as described in Section **2.2.2**. I imaged calcium transients in vS1 neurons using a Sutter MOM two-photon microscope controlled by ScanImage v.5.3 software (Vidrio Technologies). A Mai-Tai laser (SpectraPhysics; 70 fs pulsewidth, 80 MHz repetition rate) emitted a light tuned to 900 nm to excite GCaMP during imaging. I used the resonant scanner in bidirectionnal mode at a resolution of 512 $\times$ 512 pixels. The multiplexed image acquisition was obtained by means of an acousto-optic deflector (DTSX-400-980, AA Optoelectronics), which allows switching between two axially separated fields of view at kHz rates, limited only by the frequency of the resonant scanner (8 kHz) and the image resolution (512 $\times$ 512) – allowing for quasi-simultaneous imaging of two fields of view. Images were acquired from two planes separated by 120 $\mu$m (at 200 and 320 $\mu$m respectively) quasi-simultaneously giving an acquisition rate per imaged plane of ~15Hz) [Chong et al., 2019]. I used a 16X/0.80W LWD i mmersion objective (Nikon) and GaAsP photomultipliers (Hamamatsu) to detect emitted photons. The field of view was 650 $\times$ 650 $\mu$m and was imaged up to 400 $\mu$m below the cortical surface.

I carried out two-photon calcium imaging of neural activity while the mice were head-restrained on a wheel, and exposed to a pole presented to the whiskers at

four different positions along the mouse's anterior-posterior axis. The most anterior and posterior positions correspond to the two positions used for training on the pole position discrimination task (Section **2.2.6**). This was a passive condition, so the mice were not performing a behavioural task. Stimuli were generated at the same frequency as during the stimulus-reward association phase of behavioural training (Section **2.2.4**).

### 2.2.13 Analysing imaging data

I used a pre-existing two-photon analysis package (Suite2P, [Pachitariu et al., 2016]) for image registration and ROI detection. The latter was done in a semi-automatic way, with manual verification after each ROI categorisation. A correction for neuropil signal contamination and the somatic signal was calculated as:

$$F_{\text{cell}}(t) = F(t) - F_{\text{neuropil}}(t) \times r,$$

where $F_{\text{neuropil}}(t)$ is the neuropil signal surrounding each cell. This was measured by averaging the signal of all pixels with a 20 $\mu$m region from cell centre (excluding other identified cells). The contamination ratio $r$ is a constant calculated as the ratio between the average fluorescence of blood vessels and that of neuropil, which here holds a value of 0.7. Changes in fluorescence were defined as:

$$\frac{\Delta f}{f}(t) = \frac{f(t) - f_0}{f_0},$$

where $f(t)$ represents the ROI's neuropil-corrected fluorescence at a time $t$ and $f_0$ is the baseline for that acquisition, defined as all fluorescence values that fall below the $70^{\text{th}}$ percentile of their distribution averaged across the session.

## 2.3 Results

### 2.3.1 Pole position training

I began by implementing a version of Guo and colleagues' Go/No-Go task [Guo et al., 2014]. This task is desirable for my research aims, as it is a simple task that could be straightforwardly modified into a multisensory categorisation task, and it has been used to study somatosensory cortex responses with simultaneous two-photon imaging [Peron et al., 2015]. In this task, a pole was positioned within the whiskers at either a frontal or more caudal position on each trial, and mice were trained to lick in response to only one of these pole positions for a water reward (Section **2.2.5**).

Previous studies have often used mild electrical shock as a positive punishment following FA trials in training mice [Miyashita and Feldman, 2013b]. However, somatosensory responses to the aversive punishment may confound our attempt to understand how neurons represent the test stimuli themselves. One common alternative is to instead present a time-out (a form of negative punishment) after FAs. The benefit of time-outs is that they do not involve the presentation of an additional sensory stimulus in the task, which may be incorporated into the animal's perceptual learning. However, the disadvantage is that the time-outs themselves slow down the rate of trials during a behavioural session, making training sessions longer. This is an important consideration for simultaneous two-photon calcium imaging, as neurons may photo-bleach after prolonged fluorescent excitation [Chen et al., 2013b]. Time-outs were used by Guo and colleagues [Guo et al., 2014], and I aimed to determine if they are necessary for mice to learn this task.

During the reward-stimulus pairing phase of habituation (Section **2.2.4**),

a water reward was automatically delivered following Go stimulus presentation, and there were no consequences of licking on No-Go trials. I assessed whether introducing a 3 s time-out after FA trials following the habituation period would result in improved discrimination in trained mice compared to mice trained without time-outs. I compared the changes in d' between the first and last days of training. Mice trained with a time-out following FAs achieved a significantly higher increase in d' (2.16 ± 0.79; mean ± std, $N = 4$) after 6 days than those trained without a time-out (-0.34 ± 1.92 std, $N = 5$; unpaired two-tailed $t$-test , t=-2.42, $p = 0.046$)(Figure **2.4**). In fact, there was no difference in the d' between days 1 and 6 for animals trained without time-outs, suggesting that they failed to learn the task during this period (paired $t$-test , t=0.40, $p = 0.71$). In contrast, the d' of mice trained with time-outs had significantly improved on day 6 (paired $t$-test , t=-5.48, $p = 0.012$). Based on these results and the frequent use of time-outs in previous studies [Guo et al., 2014, Busse et al., 2011, Li et al., 2016], I decided to include time-outs on FA trials in order to support learning on the task, especially as I expected the multisensory version of the task to be more challenging for animals to learn.

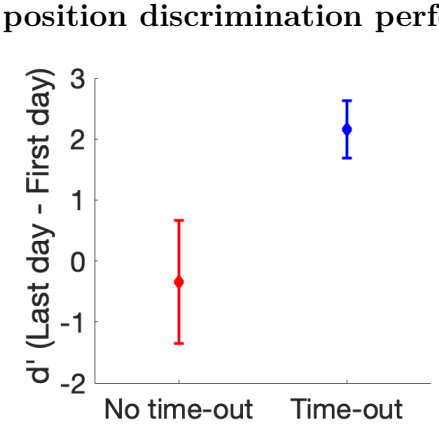



**Figure 2.4:** Change in performance across habituation and early training sessions. Average change in d' between the first day of training (day 1) and last day (day 6) (± std) of animals trained without (red; $N = 5$ mice) and with (blue; $N = 4$ mice) time-outs following FAs.

### 2.3.2 Imaging with pole stimulus

While I was developing the behavioural methods, I tested the quality of the cranial window implant and stability of the head fixation by performing two-photon calcium imaging of barrel cortex in five mice on the behavioural set-up used to train animals. I aimed here to simply confirm that I could image cortical neurons in awake, head-fixed mice while they were running on the wheel and stimulated on their whiskers. Fluorescence transients were measured while animals were presented with four possible pole positions: the caudal and frontal positions used in the discrimination task, and two intermediate positions. Responses from three example neurons imaged in three different mice are presented in Figure **2.5**.

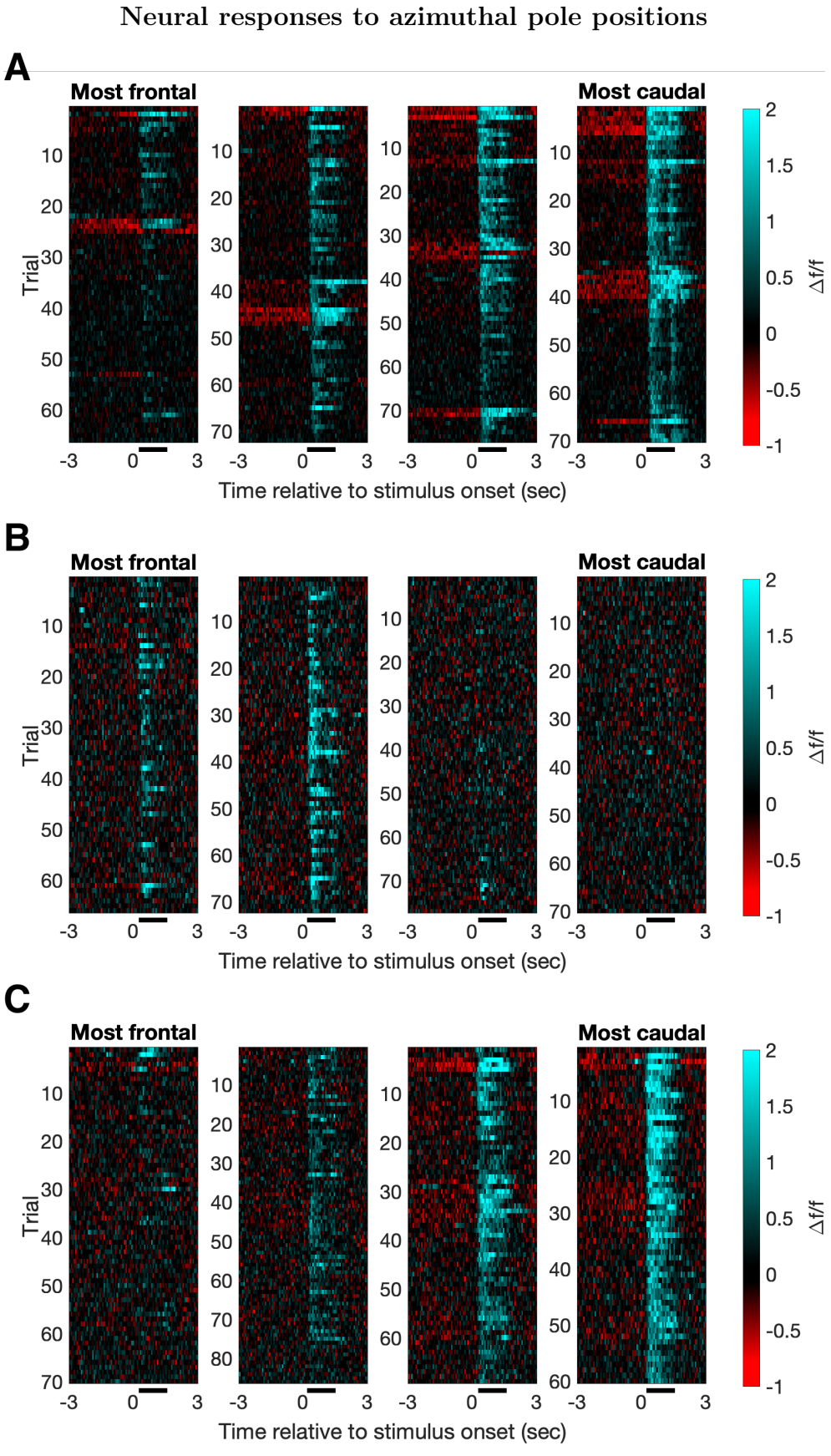

Neural responses to azimuthal pole positions
A
Most frontal
Most caudal
Trial
Time relative to stimulus onset (sec)
Δf/f
B
Most frontal
Most caudal
Trial
Time relative to stimulus onset (sec)
Δf/f
C
Most frontal
Most caudal
Trial
Time relative to stimulus onset (sec)
Δf/f

**Figure 2.5:** ***Figure on previous page.*** **A** Each plot shows the responses of an example neuron to a pole presented at one position, plotted throughout presentation time (x-axis) and across trials (y-axis). The thick black bar along the x-axis represents the time of pole presentation (1 s duration). Responses to poles presented at the four different positions are shown

in separate plots from left (frontal) to right (caudal). Responses are expressed as $\Delta f/f$ (colour bar, right). This neuron responds most strongly and reliably when the pole was presented at the most caudal position (far right plot). **B-C** Responses of two other example neurons imaged in other mice, plotted as in **A**. This neuron in **B** shows a preference for more frontal pole positions, while that in **C** responds in a sustained manner to caudal positions.

For each neuron, the response to a stimulus presentation was summarised by comparing its' maximum $\Delta f/f$ value during stimulus presentation to a pre-stimulus window (1 s immediate prior to stimulation). A neuron was considered responsive if its' trial-by-trial pre- and post-stimulus $\Delta f/f$ values differed for at least one of the four pole positions (paired $t$-test ). The results are shown in Figure **2.6**. Most imaged neurons were non-responsive ($t$-test ; $p$ >0.05) to all four pole positions (76.67% ± 10.09% std across five imaging fields). A total of 21.33% (± 8.94%) of all imaged neurons were responsive ($t$-test ; $p$ <0.05) to at least 1 pole position. More specifically, 4.46% of neurons (±3.67%) were tuned to at least one (or both) of the two frontal positions, whereas 5.86% of neurons (±4.47%) were tuned to one or both of the more caudal positions. Finally, 7.29% of neurons (±5.64%) responded non-selectively to all four pole positions. The proportion of responsive neurons was consistent with previous reports of responses to separate pole positions [Peron et al., 2015, O'Connor et al., 2010a].

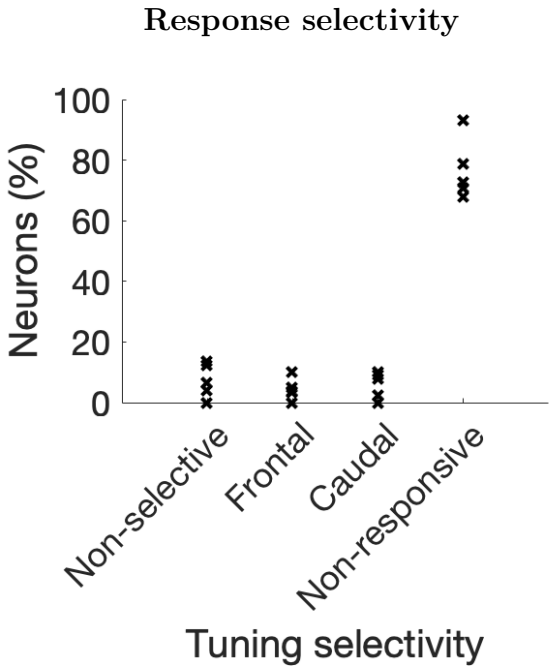



**Figure 2.6:** Proportion of non-selectively responsive, frontal-tuned responsive, caudal-tuned responsive and non-responsive neurons in response to passive presentation of a pole to the whiskers at different azimuthal positions relative to their whiskers ($N = 5$ mice).

To summarise, I confirmed the surgical implant and head fixation in awake animals in my behavioural apparatus were sufficient to measure the responses of somatosensory cortex neurons using two-photon calcium imaging. I found a small proportion of neurons responsive to tactile stimuli and was able to characterise the specificity of their tuning. Pilot imaging using the pole position stimuli therefore enabled me to validate the compatibility of my training set-up and imaging methods, so that I could adapting the stimuli for multisensory training and chronic imaging.

### 2.3.3 Limitations of the pole stimulus

Throughout the pilot experiments described above, it became clear that the pole position task introduced several confounds into the experiment, which would be especially problematic in a multisensory task.

First, the pole moved through the whiskers into its final position over a duration of about 180 ms. Therefore, the time of onset of the stimulus is difficult to pinpoint. Is it when the pole first hits a whisker, or when it is in its final position? Timing stimuli is essential for multisensory studies, as the onsets should be appropriately synchronised across modalities in order to form a unified percept of one perceptual object [Atilgan et al., 2018, Spence and Squire, 2003]. It was key that I achieved this synchrony, in order to present congruent stimuli, as discussed in Section **2.2.9**. Therefore, I required a tactile stimulus with precise onsets and offsets.

Second, it may be possible for mice to use non-tactile cues to perform the pole position discrimination task. In informal observations in our lab, we observed that trained mice would lick in response to the pole rotating position, even when the rotary motor was not in contact with the mouse's whiskers. This highlighted the fact that mice may see or hear the motor rotating the pole into position, making this intended 'unisensory' task a multisensory one. In fact, rotating the pole into position took approximately 90 ms longer for more caudal than rostral positions, which would provide an auditory duration cue if the pole motor made sounds audible to the mouse. This issue does not seem to have been considered in previous studies using pole position stimuli to study somatosensation [Guo et al., 2014, Peron et al., 2015, Li et al., 2016].

### 2.3.4 Vibrotactile training

I set out to design a novel tactile stimulus that mice could not hear nor see, with precise onset and offset times, and that would be stable throughout training. To this end, I used a Velcro pad attached to a vibrating shaker plate to deliver vibrotactile stimuli to the mouse's whiskers. I presented 3 s vibrotactile stimuli with an amplitude

of 250 $\mu$m at two different frequencies: 93 and 129 Hz. Vibrotactile stimuli had 1s onset and offset ramps. These stimuli did not generate sounds within the mouse hearing range (Section **2.2.10**), had no clear visual cues, and could be precisely temporally controlled. The frequency of the Go stimuli were counterbalanced across mice. The training protocol was the same as that of the pole-position task, with time-outs of 3 s after FA trials.

Prior to training mice, I checked for potential auditory artifacts by recording the acoustical output of the shaker while it produced the vibrotactile stimuli in the behavioural testing chamber. I found no evidence of auditory artifacts within the mouse hearing range, as described in Section **2.2.10**.

Six mice (three males) were trained on the vibrotactile discrimination task. Mice were first habituated for 4-6 days, as described for the pole position task, and were then trained using water rewards and time-outs (Section **2.2.5**). Training continued daily until mice reached the performance criterion of 70% correct for three consecutive days. The number of training sessions needed to reach 70% correct performance ranged from 1 to 15 sessions across mice (6.20 days $\pm$ 5.67, mean $\pm$ std). The performance of the six mice on the vibrotactile task is shown in Figure **2.7**.

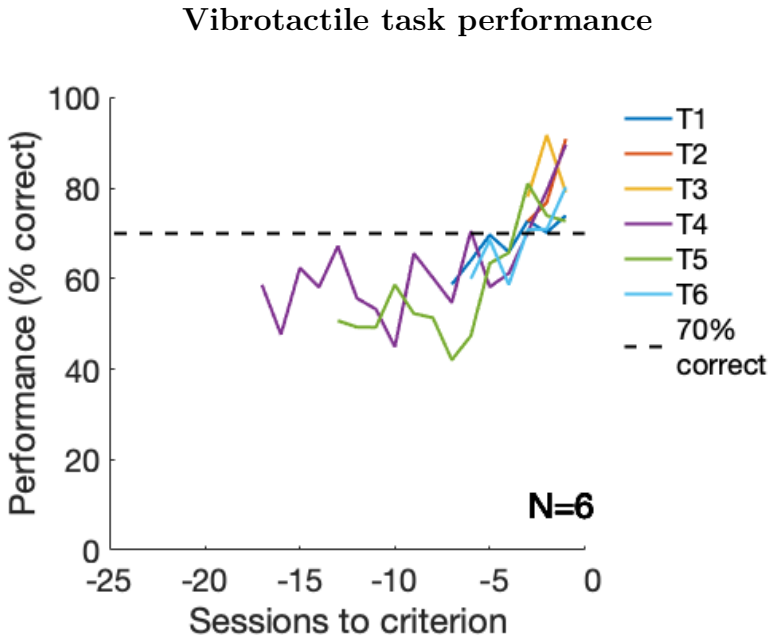



**Figure 2.7:** Percent correct performance on the vibrotactile task is plotted across daily sessions ($N = 6$ mice). The performance of each mouse is plotted as a separate coloured line. Session '0' is the session in which each animal reached the performance criterion (three consecutive days over 70%). The red dotted line shows the 70% threshold.

To investigate if mice could use visual or auditory cues to solve the vibrotactile task, I performed a control experiment in 2 mice on the day after the mice completed training. The mice were kept under water-regulation and the Velcro pad was positioned just out of reach of their whiskers in the testing box. The task otherwise remained the same and the shaker was still vibrating in the same position with respect to the animals' heads. During these sessions, the performance of both mice dropped back to chance (Figure **2.8**), suggesting that their performance did indeed depend on vibrotactile contact between their whiskers and the pad, and not any unintended visual, auditory or whole-body vibrations cues provided by the shaker.

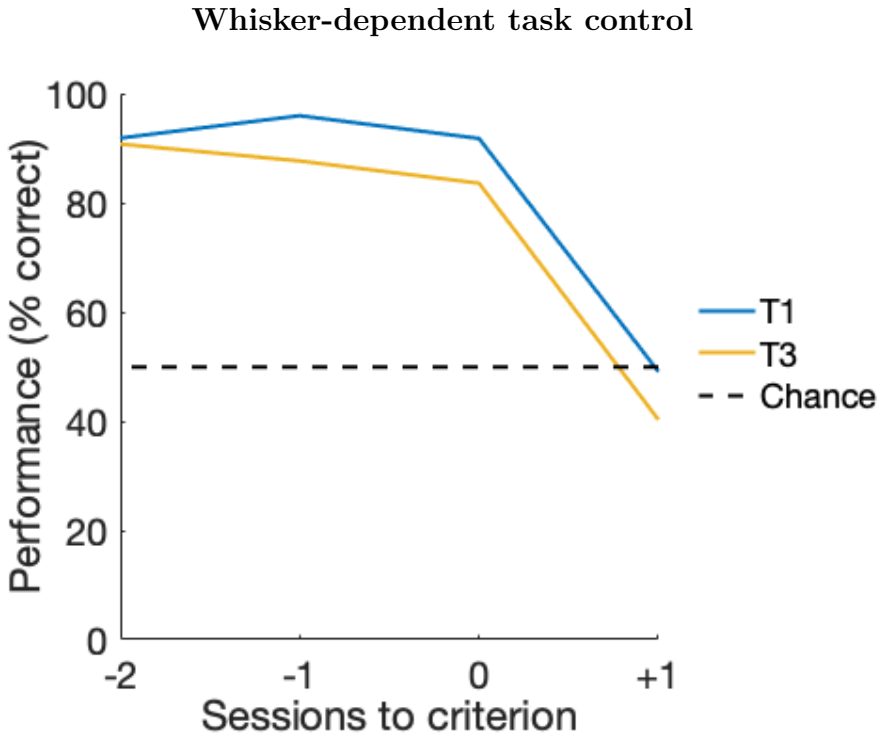



**Figure 2.8:** Percent correct performance of two mice (T1 & T3) on the vibrotactile task is plotted across daily sessions after crossing the performance threshold, including the post-training control session (1 session after the criterion was reached). They reached the performance criterion after three consecutive days above threshold. In the control session (+1 on the x-axis), the shaker was on, but the Velcro pad was out of reach of the whiskers. The horizontal dotted line indicates chance performance.

This vibrotactile stimulus was therefore well suited for incorporating into a multisensory Go/No-Go discrimination task. Mice could be trained to accurately (>70%) discriminate between vibrotactile frequencies across consecutive training days once they reached the performance threshold. The illustrated necessity for contact between whiskers and the vibrating pad to support performance demonstrates that mice depended on whisker inputs to perform this task, ruling out other sensory confounds.

### 2.3.5 Unisensory auditory training

I next aimed to demonstrate that mice could learn to discriminate between the auditory stimuli I planned to use on the multisensory task (described in Section **2.2.8**). I presented two auditory stimuli from a loudspeaker at 70 dB $SPL_C$: a 6.3 kHz tone with 93 Hz sinusoid amplitude modulation (SAM), and a 12.5 kHz tone with 129 Hz SAM. The sounds had the same overall duration (3 s) and Gaussian onset and offset ramps (1 s each) as the vibrotactile stimuli described above. Go stimuli were counterbalanced across mice, and the training protocol was identical to the vibrotactile discrimination task.

Seven mice (five males) were trained on the auditory discrimination task. After four to six days of habituation, mice were trained using water rewards and time-outs (Section **2.2.5**). Mice were trained daily until reaching the performance criterion of 70% correct for three consecutive days. Mice required between 1-14 days to reach the threshold performance of 70% (6.70 days ± 4.40, mean ± std). The performance of the seven mice on the auditory task is shown in Figure **2.9**.

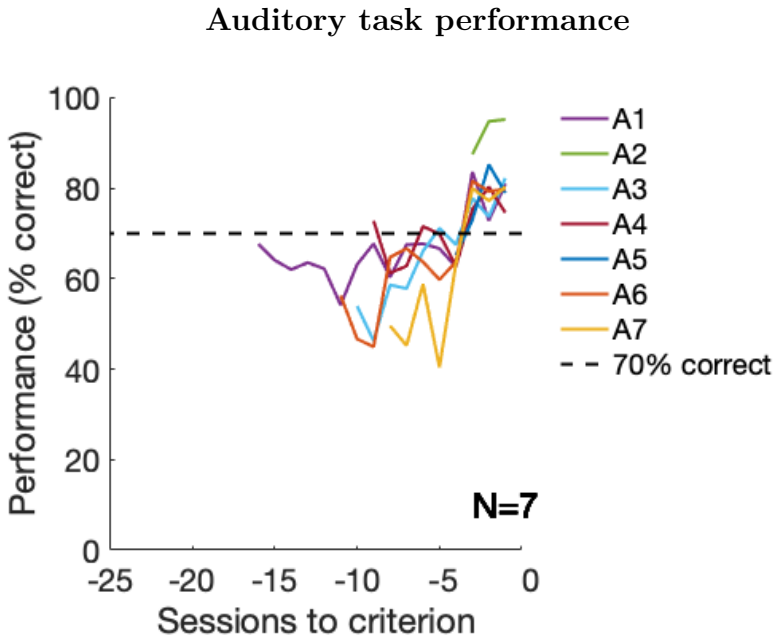



**Figure 2.9:** Percent correct performance on the auditory task is plotted across daily sessions ($N = 7$ mice). The performance of each mouse is plotted as a separate coloured line. Session '0' is the session in which each animal reached the performance criterion (3 consecutive days over 70%). The red dotted line shows the 70% threshold.

Therefore, mice were able to learn to discriminate between the two tones in the same apparatus and task design as used in the vibrotactile task, in a similar number of training sessions. The time required to acquire these two tasks will be quantitatively explored in Section **2.3.7**.

### 2.3.6 Multisensory training

Having confirmed that mice could learn to discriminate between the frequencies of both the vibrotactile and auditory stimuli in the context of unisensory tasks, I set out to train mice to categorise pairs of audio-vibrotactile stimuli. The multisensory stimuli were combined from the unisensory stimuli described above, and as described in Section **2.2.9**, giving four possible pairings of the two tones and two vibrotactile frequencies. Eleven mice (seven males) were trained on the multisensory

task. For each mouse, the Go audio-vibrotactile stimulus consisted of a tone with a modulation frequency equivalent to the vibrotactile frequency. For six mice, this comprised of the 6.3 kHz carrier tone (amplitude modulated at 93 Hz) presented synchronously with the 93 Hz vibrotactile stimulus. For the remaining five mice, the Go stimulus was the 12.5 kHz tone (129 Hz SAM) together with the 129 Hz vibrotactile stimulation. The apparatus, habituation and training protocols remained otherwise unchanged from the unisensory tasks described above.

Mice were first habituated for 4-6 days and were then trained, as described for the auditory and vibrotactile tasks (Section **2.2.5**), until reaching 70% correct for three consecutive days. It took between 4 to 23 days for nine mice to reach the 70% performance threshold (12.70 $\pm$ 5.66, mean $\pm$ std), while the two remaining mice failed to reach criterion after 30 days of training. The performance of the nine mice successfully trained on the multisensory task is shown in Figure **2.10**.

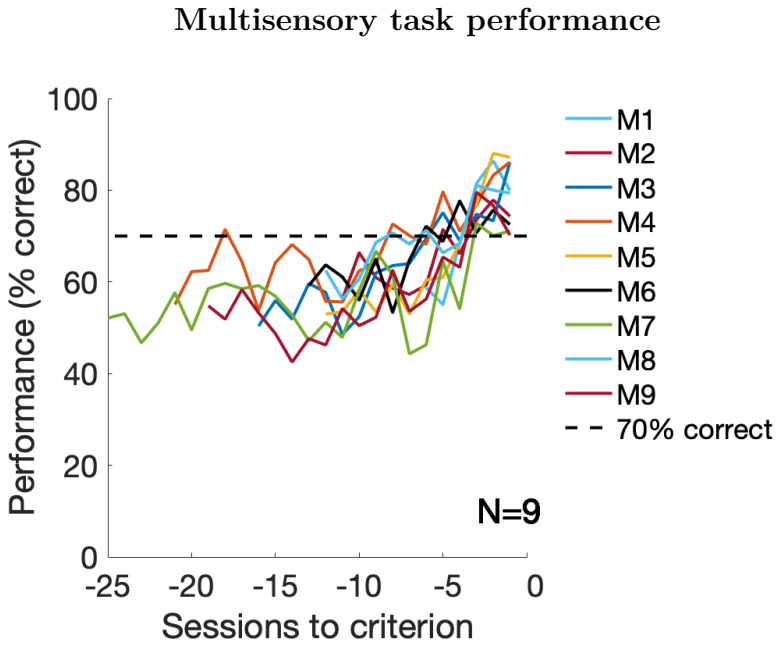



**Figure 2.10:** Percent correct performance on the audio-vibrotactile task is plotted across daily sessions ($N = 9$ mice). The performance of each mouse is plotted as a separate coloured line. Session '0' is the session in which each animal reached the performance criterion (three consecutive days over 70%). The red dotted line shows the 70% threshold.

Therefore, most mice were able to learn to discriminate between pairs of audio-vibrotactile stimuli. At the end of training, all nine mice were able to maintain their performance over threshold for at least three consecutive days. This demonstrates that mice can combine information from the auditory and tactile modalities to learn to classify multisensory stimuli.

The remaining results sections of this chapter will more fully describe and compare the behaviour of mice across the vibrotactile, auditory, and audio-vibrotactile versions of this task.

### 2.3.7 Rate of acquisition on the auditory, vibrotactile, and multisensory tasks

As an initial step in comparing the performance of mice trained on different versions of the task, I assessed whether there were differences in the time taken to reach criterion across the three tasks (auditory, vibrotactile, and audio-vibrotactile), and whether age or sex could explain these differences.

I used a one-way ANOVA to compare the time taken to reach the perfornance threshold for the first of three consecutive days across the three versions of the task. I found a significant effect of task type on the number of sessions required to reach threshold ($F_{(2,19)}$=3.67, $p$ = 0.0448; Figure **2.11 A**). Pairwise comparisons (Fisher's least significant difference) indicated that mice required more sessions to learn the multisensory task (12.7 $\pm$ 5.66 sessions) than either the tactile (6.2 $\pm$ 5.67 sessions; $p$ = 0.031) or auditory (6.7 $\pm$ 4.4 sessions; $p$ = 0.0379) versions of the task. There was no significant difference in the number of sessions required to reach threshold on the auditory and tactile unisensory tasks ($p$ = 0.8544).

Similarly, I found an effect of task type on the total number of trials performed before reaching threshold (one-way ANOVA, $F_{(2,19)}$=5.73, $p$ = 0.0113; Figure **2.11 B**). In particular, Fisher's least significant difference tests showed that mice trained on the multisensory task (1832.22 $\pm$ 912.68 trials) performed more trials before reaching threshold than mice trained in the auditory (840.29 $\pm$ 508.04 trials; $p$ = 0.0164) and tactile tasks (646.17 $\pm$ 566.17 trials; $p$ = 0.0068). There was no difference between the number of trials required to learn the tactile and auditory tasks ($p$ = 0.6306).

I carried out an unpaired $t$-test (two-tailed) to compare the learning rates between between males ($N$ = 15) and females ($N$ = 7). Here, performance was

combined across all 3 versions of the task. I found that males and females did not differ in the number of sessions required to reach threshold (males: 8.2 ±5.7 sessions, females: 10.7 ±6.4 sessions; t=-0.92, $p = 0.367$) (Figure **2.11 C**), nor did they differ in the number of trials to threshold (males: 1245 ±729 trials, females: 1142 ±937 trials; t=-0.25, $p = 0.802$) (Figure **2.11 D**).

As the ages of mice were not matched across tasks, age differences may explain the learning rate differences observed between tasks. I used a one-way ANOVA to compare the mean ages of mice at the start of training across the 3 versions of the task. I found significant differences in ages across the mice trained on the different tasks ($F_{(2,19)}$=10.72, $p = 0.0008$; Figure **2.11 E**). More specifically, post-hoc pairwise comparisons between the groups revealed that mice trained on the auditory task were younger than those trained on the tactile ($p = 0.0003$) and multisensory ($p = 0.0027$) tasks (auditory: 13.7±4.55 weeks, tactile: 24.5±4.82 weeks, multisensory: 21.4±3.96 weeks), but there was no significant difference between the ages of the mice trained on the tactile and multisensory tasks ($p = 0.1913$). Therefore, as the mice trained on the auditory task were younger than those trained on the multisensory task, it is possible that mice learned the multisensory task more slowly due to age-related effects on the auditory component of the task.

Finally, to further examine this possibility and also possible gender effects on task acquisition, I carried out a three-way ANOVA with the sessions to threshold as the dependent variable, the task type and mouse gender as categorical predictors, and mouse age at the start of training as a continuous predictor. In this instance, neither the task, gender, or the age at the start of training had a significant effect on the number of sessions to threshold (Figure **2.11 F**). Similarly, age, task and gender did not significantly predict the number of trials to threshold (Figure **2.11 G**).

**Rates of acquisition of the unisensory and multisensory tasks**

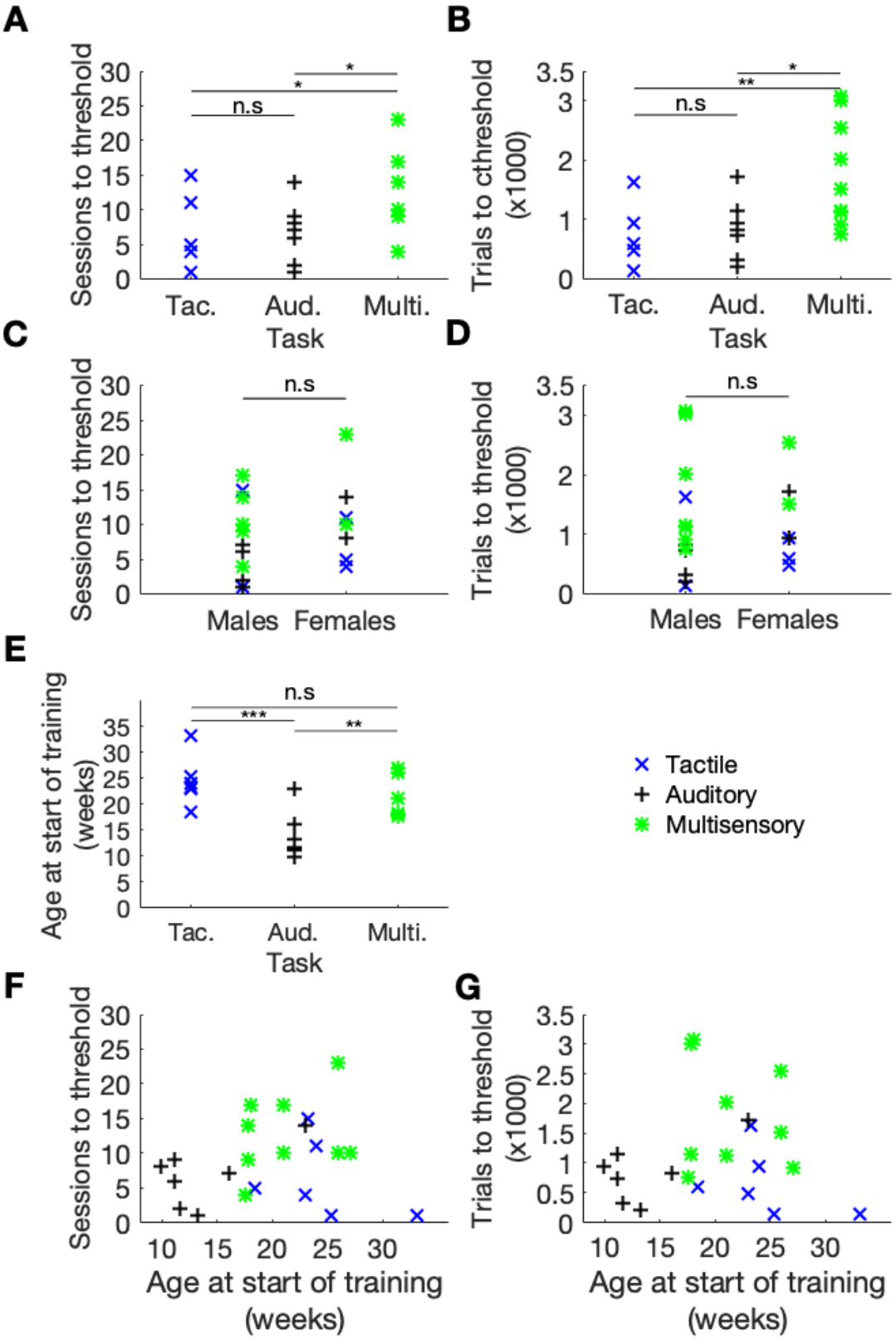


**Figure 2.11:** ***Figure on previous page.*** **A.** Number of sessions (days) to reach threshold for mice trained on the different versions of the task $** : p < 0.01$, $* : p < 0.05$, n.s: non-significant, post-hoc pairwise comparison (Fisher's least significant difference). **B.** Total number of trials performed before reaching performance threshold across the three tasks. $** : p < 0.01$, $* : p < 0.05$, n.s: non-significant, post-hoc pairwise comparison (Fisher's least significant difference). **C.** Number of sessions to threshold in males ($N = 15$) and females ($N = 7$). n.s: non-significant, independent two-tailed $t$-test . **D.** Number of trials to threshold in males ($N = 15$) and females ($N = 7$). n.s: non-significant, independent two-tailed $t$-test . **E.** Differences in age at the start of training between mice trained on the different versions of the task. Tactile (Tac.): blue crosses, Auditory (Aud.): black vertical crosses, Multisensory (Multi.): green stars. $*** : p < 0.001$, $** : p < 0.01$, n.s: non-significant, post-hoc pairwise comparison (Fisher's least significant difference). **F.** Number of sessions to threshold plotted against the age at the start of training for mice trained on each task. **G.** Number of trials to threshold plotted against the age at the start of training for mice trained on each task.

To summarise, mice that were trained on the multisensory task took longer to learn than those trained on the unisensory tasks. Mice trained on the auditory task were significantly younger than mice trained on the tactile and multisensory tasks, but there was no effect of age or gender on the rate of learning overall.

### 2.3.8 Lick patterns

Licking and lick patterns can be a good indication of an animal's behaviour. Beyond being used as a binary response measure (i.e. 'lick' or 'no-lick' as an operant behaviour measure), licking could also be used as a confidence indicator (a lick rate in Go or No-Go trials for example) or even as a marker for thirst levels (compulsive licks in more thirsty mice). In order to assess lick-related factors that may underlie changes in performance observed across mice, I compared licking patterns across the three tasks and at different stages of training. The stages included a 'training' stage, early in training, during which performance was low (<50% correct), and a 'trained'

stage, typically the last day of training and third day above performance criterion (>70% correct).

Two mice trained on the tactile task reached criterion on their first day of testing post-association, as well as a mouse trained on the auditory task. These were therefore not included in the 'training' cohort. Hence, four mice trained on the tactile task, six trained on the auditory task and all nine trained on the multisensory task were included in the training analysis.

In both training and trained sessions, I set out to compare whether patterns of licking changed between correct and incorrect licks (Hits and FAs respectively). As shown in Figure **2.12**, individual mice adopted different licking strategies at different stages of training. For instance, one mouse trained on the tactile version of the task (T2) licked for both Go and No-Go trials during training but already exhibited increased licking frequency in Go trials. However, once trained, it had learned to withhold licking altogether in No-Go trials and wait for the end of stimulus presentation before licking in Go trials. Other mice adopted different strategies. For example, another mouse trained on the auditory task (A7) never licked at the beginning of training, either in Go trials or No-Go trials, and licked during Go stimulus presentation when trained. Some mice never learned to completely withhold in No-Go trials, even when trained (for instance tactile-trained mouse T5, auditory-trained mouse A4 and multisensory-trained mouse M9). However, the rate of licking in No-Go trials was lower than in Go trials for these animals, and this differentiation of responses indicates that they can distinguish between Go and No-Go trials.

**Lick distribution**

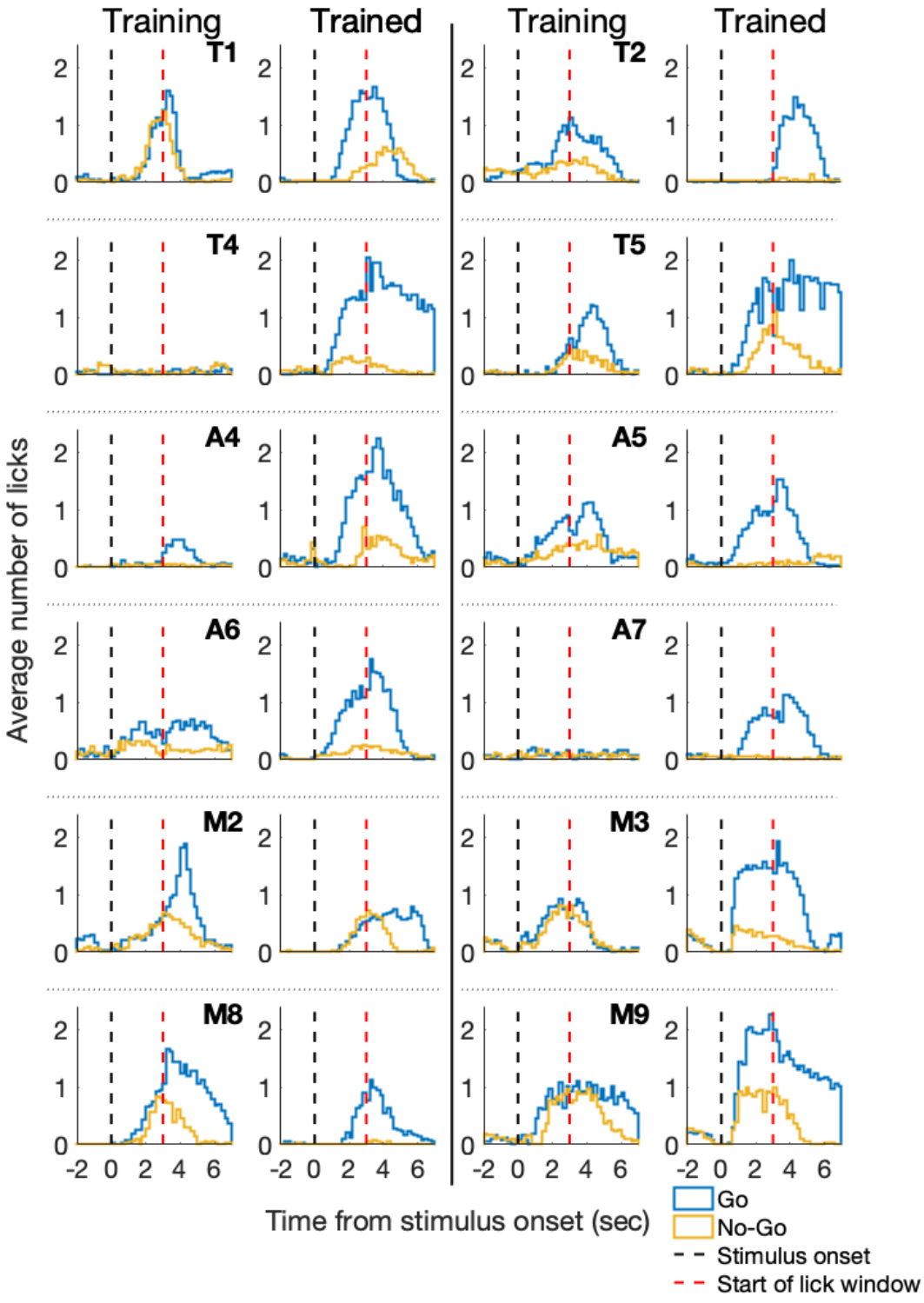


**Figure 2.12:** ***Figure on previous page.*** Distribution of licks on Go (blue) and No-Go (yellow) trials of 12 mice (**T** mice trained on the tactile task, **A** mice trained on the auditory task, **M** mice trained on the multisensory task) on a training day (<60% correct) and once trained (>70% correct). The average number of licks is plotted per 200 ms time bin across Go and No-Go trials.

Having observed that licking strategies may differ across animals, I subsequently aimed to define these differences and find consistent and reliable behavioural markers of learning across the training cohort. I defined the mean lick rate as the average number of licks per trial between one second preceding stimulus onset and the response window onset. Mean lick rates were calculated separately for individual animals and for Go and No-Go trials. The lick detection window in this analysis was set to run until the response window onset (corresponding to the end of the stimulus) as licks past that time in Go trials elicited a reward and subsequent drinking. This usually resulted in a transient increase in licking rate as the mouse drank the water, which does not necessarily reflect the animal's categorization of the stimulus. The resulting average lick rates are shown in Table **2.2**.

| Training | | | | | |
|---|---|---|---|---|---|
| Tactile | | Auditory | | Multi. | |
| Go | No-Go | Go | No-Go | Go | No-Go |
| 2.90±1.86 | 2.13±1.66 | 3.25±2.78 | 3.23±3.00 | 4.48±4.58 | 3.81±3.51 |
| Trained | | | | | |
| Tactile | | Auditory | | Multi. | |
| Go | No-Go | Go | No-Go | Go | No-Go |
| 6.15±4.74 | 2.33±2.91 | 9.90±4.68 | 2.31±1.26 | 12.76±6.31 | 5.46±3.37 |

**Table 2.2:** Mean number of licks (from stimulus onset to response window onset) per Go and No-Go trial during and after training (± std).

First, I used a three-way ANOVA with the average number of licks per trial as the predicted variable, and the task type (auditory, vibrotactile or multisensory), trial

type (Go or No-Go) and training stage (training or trained, as described above) as predictor variables. I found significant effects of trial type ($F_{(1,79)}$=13.26, $p = 0.0005$) and training stage ($F_{(1,79)}$=12.16, $p = 0.0008$) on the number of licks per trial, as well as a smaller but significant effect of task type ($F_{(2,79)}$=4.68, $p = 0.0124$).

A post-hoc pairwise comparison (Fisher's least significant difference) revealed no significant differences in the number of licks per Go and No-Go trial in training animals on either of the three tasks (Tactile: $p = 0.6201$; Auditory: $p = 0.6314$ and Multisensory: $p = 0.5406$). There were however significant differences in the trained mice (Tactile: $p = 0.0097$; Auditory: $p = 0.0002$ and Multisensory: $p = 3.20 \times 10^{-5}$)). Specifically, mice licked significantly more in Go trials than No-Go trials at the end of training, but not during training (Figure **2.13 A**).

**Number of licks per trial and lick-rate inferred performance**

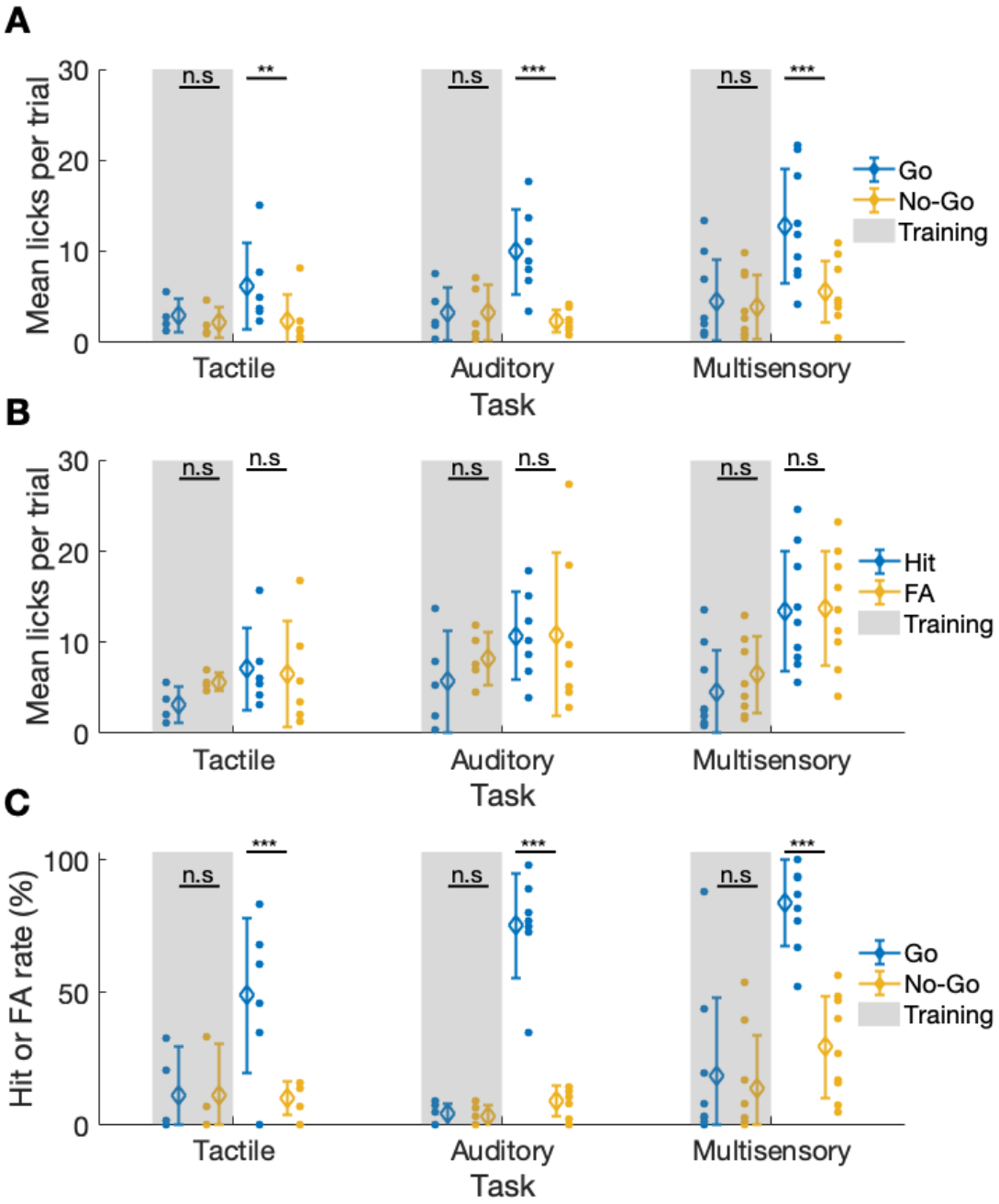


**Figure 2.13:** ***Figure on previous page.*** **A.** Mean number of licks (± std) on Go (blue) and No-Go (yellow) trials, during training (grey shadowing) and after training. Dots show data points for individual mice. The average number of licks was calculated across the duration of stimulus presentation and before any potential reward delivery. Tactile: training $N = 4$ mice, trained $N = 6$ mice; Auditory: training $N = 5$ mice, trained $N = 7$ mice; Multisensory: training $N = 7$ mice, trained $N = 7$ mice. $***$ : $p < 0.001$, $**$ : $p < 0.01$, n.s: non-significant, paired $t$-test . **B.** Mean number of licks (± std) per Hit and FA trials (blue and yellow bars respectively), during training (grey shadowing) and after training. Dots show data points for individual mice. **C.** Lick-rate-inferred Hit and FA rates (blue and yellow bars respectively; mean ± std), during training (grey shadowing) and after training. Dots show data points for individual mice. The 'Lick' threshold (i.e. the lick rate required for a trial to be categorised as 'Lick') was set at 2.5 Hz.

In trained mice, the FA rate was consistently lower than during training. Therefore, there was approximately the same number of Go and No-Go trials, but a substantially smaller number of No-Go trials in which the mice did lick (i.e. fewer FA trials than Hit trials). Figure **2.12** suggests that some mice had lower lick rates on false alarm trials than on hit trials. This may reflect a lower confidence in detecting the target stimulus on false alarm trials. To examine this observation statistically, I performed the same ANOVA analysis described above, but calculated the average number of licks on Hit and FA trials instead of Go and No-Go trials, and used this as the predicted variable. The lick rates on Hit and FA trials are summarised in Table **2.3**. In the 3-way ANOVA, trial outcome (Hit or FA) replaced trial type (Go and No-Go) as one of the predictors.

The ANOVA showed a significant effect of training stage ($F_{(1,79)}$=14.2, $p$ = 0.0003) and task type ($F_{(2,79)}$=3.5, $p$ = 0.0356) on the number of licks per trial. However, there was no effect of trial outcome on the lick rate ($F_{(1,79)}$=0.76, $p$ = 0.386). Post-hoc pairwise comparisons revealed that at the start of training, there was no differences between the number of licks in Hit and FA trials in mice trained

on all versions of the task (Tactile: $p = 0.487$, Auditory: $p = 0.363$, Multisensory: $p = 0.310$). Similarly, there were no lick rate differences in mice trained on the tactile ($p = 0.9267$), auditory ($p = 0.9388$) and multisensory ($p = 0.996$) versions of the task (Figure **2.13 B**). Hence, while mice quite logically licked less in No-Go trials than Go trials after training, they did not show differences in the number of licks on FA and Hit trials.

| Training | | | | | |
|---|---|---|---|---|---|
| Tactile | | Auditory | | Multi. | |
| Hit | FA | Hit | FA | Hit | FA |
| 3.15±1.92 | 5.66±0.96 | 5.81±5.44 | 8.23±2.86 | 4.53±4.63 | 6.49±4.20 |
| Trained | | | | | |
| Tactile | | Auditory | | Multi. | |
| Hit | FA | Hit | FA | Hit | FA |
| 7.11±4.53 | 6.50±5.86 | 10.67±4.83 | 10.84±8.96 | 13.46±6.61 | 13.73±6.25 |

**Table 2.3:** Mean number of licks (from stimulus onset to response window onset) per Hit and FA trial during and after training (± std).

Nonetheless, the differences in licking patterns between mice shown in Figure **2.12** suggest that there may be more reliable ways of measuring mouse decisions on a given trial than simply the absence or presence of licking, as it is difficult for some mice to inhibit licking entirely on No-Go trials. Therefore, I performed a post-hoc examination of an alternative behavioural performance measure. Instead of using a binary lick detection, in which a single lick during the response window is sufficient to qualify a trial as a Hit or FA), I re-calculated the animals' performance using a lick-rate threshold as a behavioural metric. For each trial, an average lick-rate above 2.5Hz during the last second of stimulus presentation was considered a lick response. This is a conservative estimate as mice lick at rates of up to 5 Hz when they are drinking [Murakami, 1977]. The resulting Hit and FA rates are summarised in Table

**2.4**.

In order to assess whether this lick-rate inferred performance may accurately reflect a mouse's performance on my task, I performed a three-way ANOVA with the task type, trial type (Go or No-Go), and training stage (training or trained) as predictor variables, and the Hit and FA rates (defined using the 2.5Hz lick rate rule) as the predicted variable. I found a significant effect of trial type on rate measurement ($F_{(1,79)}$=41.39, $p = 1.36 \times 10^{-8}$) as well as of training stage ($F_{(1,79)}$=53.38, $p = 3.45 \times 10^{-10}$) and task type ($F_{(2,79)}$=5.8, $p = 0.0047$). Pairwise post-hoc tests showed that early in training there was no significant differences between the Hit and FA rates on either the tactile ($p = 0.553$), auditory ($p = 0.315$) or multisensory ($p = 0.588$) task. However, as expected, Hit rates were higher than FA rates after training (tactile: $p = 4.19 \times 10^{-6}$; auditory: $p = 6.80 \times 10^{-10}$; multisensory: $p = 3.54 \times 10^{-10}$). Therefore, using lick-rates above 2.5Hz as a behavioural read-out of decisions on each trial can reliably inform us of an animal's performance on this task (Figure **2.13 C**), and this may be preferable to the less forgiving threshold of 1 lick per trial.

| | Training | | |
|---|---|---|---|
| | Tactile | Auditory | Multisensory |
| Hit rate | 11.28±18.37 | 3.99±3.75 | 18.44±29.72 |
| FA rate | 11.11±19.25 | 3.20±3.95 | 13.59±19.90 |
| | Trained | | |
| | Tactile | Auditory | Multisensory |
| Hit rate | 48.74±29.19 | 75.13±19.92 | 83.53±15.93 |
| FA rate | 9.89±6.38 | 8.74±5.77 | 29.30±19.21 |

**Table 2.4:** Inferred Hit and FA rates based on a 2.5Hz lick-rate threshold durng and after training (± std).

### 2.3.9 Within-session variations

Having examined the variety in licking performance across animals and training stages, I next investigated whether summarising an animal's performance over a whole session may mask variations in performance within a single training session. To study changes in performance within a session, I calculated the performance of each mouse in 3 session blocks: the early, middle and final third of trials within each session. I then averaged the performance in each session block across training (3 sessions in which performance was <70% correct) or trained sessions (3 final sessions with performance >70% correct). For each mouse, I therefore obtained a measure of average performance at the start, middle and end of sessions in both training and trained conditions (Table 2.5). I included four mice trained on the tactile task, six trained on the auditory task and all nine trained on the multisensory task in the 'training' cohort for analysis.

| Tactile | | | | | | |
|---|---|---|---|---|---|---|
| | Training | | | Trained | | |
| | Start | Middle | End | Start | Middle | End |
| % correct | 52.00 ±7.74 | 56.81 ±7.42 | 61.69 ±7.40 | 67.67 ±2.92 | 78.44 ±6.22 | 78.28 ±5.49 |
| **Auditory** | | | | | | |
| | Training | | | Trained | | |
| | Start | Middle | End | Start | Middle | End |
| % correct | 57.25 ±8.98 | 57.45 ±7.56 | 62.77 ±5.85 | 68.59 ±6.27 | 82.13 ±6.91 | 84.30 ±5.27 |
| **Multisensory** | | | | | | |
| | Training | | | Trained | | |
| | Start | Middle | End | Start | Middle | End |
| % correct | 55.00 ±5.12 | 60.24 ±4.82 | 62.42 ±5.66 | 72.85 ±12.24 | 76.48 ±5.97 | 79.11 ±4.57 |

**Table 2.5:** Performance at the start (beginning 1/3 of trials), middle (the second third of trials) and end (last third of the session) of training (performance <70% correct) and trained (>70%) sessions, for each task (% correct±std).

I used a one-way repeated-measures ANOVA to determine if performance changed across the early, middle and late blocks of a training sessions. I found that performance changed across the 3 session blocks ($F_{(2,32)}$=11.25, $p = 0.0002$)(Figure **2.14 A**). Similarly, in trained mice, performance also changed across these blocks of the sessions ($F_{(2,38)}$=29.86, $p = 1.6 \times 10^{-8}$)(Figure **2.14 B**).

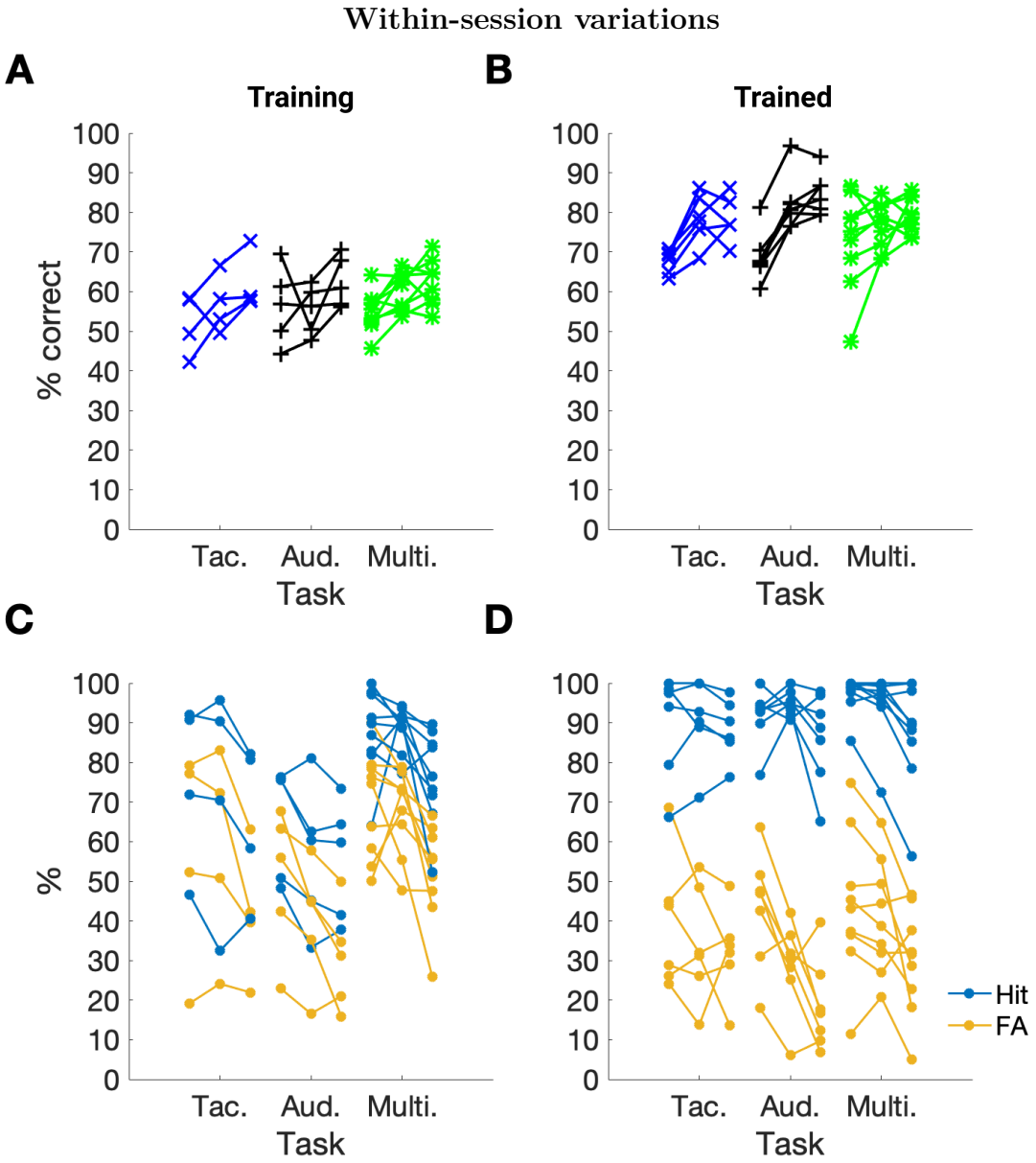



**Figure 2.14: A.** Within-session changes in performance (% correct) during training on each task (before crossing the 70% correct performance threshold). The leftmost points in each column correspond to the first third of trials across sessions. The middle points correspond to the second third, and the rightmost points correspond to the final third. Tactile task: $N = 4$; Auditory: $N = 5$; Multisensory: $N = 9$. **B.** Within-session changes in performance (% correct) after training on each task (performance over 70% correct). Tactile task: $N = 6$; Auditory: $N = 7$; Multisensory: $N = 9$. **C.** Within-session variations in Hit rates (blue, % of all Go trials) and FA rates (yellow, % of all No-Go trials) during training on all tasks. **D.** Within-session variations in Hit rates and FA rates in trained mice.

More specifically, during training there was no performance difference between the beginning and middle of sessions (post-hoc pairwise comparison, Fisher's least significant difference; $p = 0.086$). However, performance in the final third of the session was significantly higher than both the early and middle blocks ($p = 8.09 \times 10^{-5}$ and $p = 0.01$ respectively). In trained animals, performance at the beginning of sessions was significantly lower than in the middle ($p = 1.29 \times 10^{-6}$) and late periods of the sessions ($p = 1.08 \times 10^{-5}$), while it did not differ between the middle and end of sessions ($p = 0.26$). Therefore, during training, the performance of mice significantly improved at the end of each session, while expert mice improved faster, showing increased performance in the middle of sessions.

In order to better understand how the changes in rates of trial outcomes underlying these within-session performance improvements, I next examined changes in Hit rate (on Go trials) and FA rate (on No-Go trials) during training (Figure **2.14 C**) and in trained mice (Figure **2.14 D**). I used a one-way repeated-measures ANOVA with the session time blocks as the within-subject factor and task type as the between-subject factor. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated in this analysis ($\chi^2(2) = 8.85$, $p = 0.012$), and I therefore used a Greenhouse-Geisser correction. I found a significant effect of session block on FA rate during sessions ($F_{(1.38,22.14)}$=16.62, $p = 0.00016$). Post-hoc pairwise comparisons revealed that FA rates were lower at the end of a session compared to either the beginning or middle of sessions ($p = 0.0007$ and $p = 6.56 \times 10^{-5}$, respectively). There was no difference in FA rate between the early and middle blocks ($p = 0.248$).

While the FA rate consistently declined during training sessions, so did the Hit rate. I found a significant effect of session block time on Hit rates ($F_{(2,32)}$=6.89,

$p = 0.0032$). Post-hoc pairwise comparisons revealed that Hit rates were lower in the final third of sessions than in either the first or middle third of sessions ($p = 0.0008$ and $p = 0.0132$, respectively). Hit rates did not significantly differ between the beginning and middle of sessions ($p = 0.268$). Therefore, during training there was a decreased likelihood of mice licking in response to stimuli in the final third of a session, on both Go and No-Go trials.

Similarly, in trained mice, I assessed the effect of within-session timing on the FA rate. I used a one-way repeated-measures ANOVA as described above. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated ($\chi^2(2) = 9.66$, $p = 0.0079$) and I therefore used a Greenhouse-Geisser correction. I found that the FA rate changed significantly across session blocks ($F_{(1.41,26.85)}$=13.46, $p = 0.00034$). As described during training, the FA rate of trained animals consistently and significantly decreased throughout a given session (beginning v. end: $p = 0.00047$; beginning v. middle: $p = 0.0007$; middle v. end: $p = 0.02$).

Finally, I measured changes in Hit rates in trained mice throughout sessions. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated ($\chi^2(2) = 9.72$, $p = 0.0077$), so I used a Greenhouse-Geisser correction on a one-way repeated-measures ANOVA. I found a small but significant decrease in Hit rates within sessions ($F_{(1.41,26.81)}$=6.67, $p = 0.009$). Although minimal, the decrease in Hit rate was significant between the end of trained sessions and both the beginning and middle of sessions ($p = 0.03$ and $p = 0.002$, respectively), but not between the beginning and middle of sessions ($p = 0.629$).

| Tactile | | | | | | |
|---|---|---|---|---|---|---|
| | Training | | | Trained | | |
| | Start | Middle | End | Start | Middle | End |
| Hit rate | 75.33 | 72.25 | 65.42 | 89.28 | 90.50 | 88.37 |
| | ±21.23 | ±28.68 | ±19.83 | ±13.62 | ±10.62 | ±7.64 |
| FA rate | 56.96 | 57.54 | 41.74 | 39.41 | 34.16 | 32.17 |
| | ±28.02 | ±25.98 | ±16.87 | ±16.87 | ±14.65 | ±11.35 |
| Auditory | | | | | | |
| | Training | | | Trained | | |
| | Start | Middle | End | Start | Middle | End |
| Hit rate | 69.07 | 62.48 | 57.25 | 91.58 | 94.88 | 86.30 |
| | ±15.78 | ±21.95 | ±14.38 | ±7.20 | ±3.22 | ±11.68 |
| FA rate | 51.36 | 43.09 | 31.87 | 43.06 | 28.61 | 18.47 |
| | ±16.39 | ±15.78 | ±12.28 | ±14.76 | ±11.35 | ±11.33 |
| Multisensory | | | | | | |
| | Training | | | Trained | | |
| | Start | Middle | End | Start | Middle | End |
| Hit rate | 87.96 | 88.69 | 76.30 | 97.35 | 94.76 | 87.32 |
| | ±11.04 | ±5.64 | ±11.87 | ±4.75 | ±8.62 | ±13.63 |
| FA rate | 69.43 | 67.90 | 52.31 | 43.86 | 40.72 | 29.81 |
| | ±13.46 | ±10.41 | ±12.37 | ±18.40 | ±14.13 | ±13.23 |

**Table 2.6:** Hit and False Alarm rate at the start, middle and end of sessions during training then trained sessions across tasks (mean % FA & Hit ± std).

Hence, both during and after training, and throughout all versions of the task, I observed a drop in Hit and FA rates throughout each session. Conversely, this means that there was an increase in both Miss and Correct Rejection rates during sessions. Therefore, although the respective Hit and FA rates may have differed between training and expert sessions, within-session patterns characterised by a high lick-rate at the beginning of sessions (equivalent to high Hit and FA rates) decreasing toward the end (consistent with lowered Hit and FA rates). Nonetheless, a few notable differences arose between training and expert sessions. While Hit and FA rates only significantly dropped during the end of sessions in animals that had not yet learned the tasks, the FA rate of trained animals already significantly decreased between the

first and middle session blocks. This suggests that the FA rate dropped faster in trained animals than during training. This may occur because the trained animals are delivered water rewards on a higher proportion of early trials than untrained animals. Therefore, their thirst and motivation for licking may decrease more quickly throughout the session, especially during the initial trials.

If licking motivation is the primary factor driving the within-session changes observed, they may be more disruptive to performance when only a single lick on a No-Go trial is sufficient to determine a false alarm. To examine this further, I applied the lick-rate inferred performance measurement criterion described in Section **2.3.9** to assess whether it may help account for within-session variations in behaviour (inferred performance is shown in Table **2.7**).

| Tactile | | | | | | |
|---|---|---|---|---|---|---|
| | Training | | | Trained | | |
| | Start | Middle | End | Start | Middle | End |
| % correct | 50.49 ±28.36 | 52.23 ±34.45 | 39.58 ±23.23 | 49.17 ±29.35 | 56.87 ±19.59 | 51.42 ±13.64 |
| **Auditory** | | | | | | |
| | Training | | | Trained | | |
| | Start | Middle | End | Start | Middle | End |
| % correct | 20.40 ±13.65 | 20.59 ±15.25 | 16.04 ±15.32 | 40.98 ±10.10 | 46.86 ±7.16 | 40.60 ±12.54 |
| **Multisensory** | | | | | | |
| | Training | | | Trained | | |
| | Start | Middle | End | Start | Middle | End |
| % correct | 61.08 ±23.47 | 59.63 ±19.62 | 43.19 ±14.40 | 60.13 ±13.98 | 60.29 ±15.93 | 48.61 ±16.85 |

**Table 2.7:** Lick-rate-inferred performance at the start, middle and end of sessions during training (<70% correct) then trained (>70% correct) sessions across tasks (% correct ± std). A trial was categorised as a 'Lick' trial (i.e. 'Hit' or 'FA') if the lick-rate in the last second of stimulus presentation exceeded a 2.5 Hz threshold.

I used a one-way repeated-measures ANOVA to assess changes in performance

derived using the 2.5 Hz licking criteria across the three session blocks during training sessions (Figure **2.15A**). Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated ($\chi^2(2) = 9.25$, $p = 0.0098$) and I therefore applied a Greenhouse-Geisser correction. I found a significant effect of session block on within-session performance ($F_{(1.35,20.22)}$=5.87, $p = 0.017$)(Figure **2.15 A**). Similarly, in trained mice, I found that performance changed across blocks within a session ($F_{(2,38)}$=4.72, $p = 0.0148$)(Figure **2.15 B**).

Post-hoc pairwise comparisons indicated that during training, there was no performance difference between the beginning and middle block of sessions (Fisher's least significant difference; $p = 0.998$), nor between the first and last block within a session ($p = 0.09$). However, performance in the final third of sessions was significantly lower than the middle block ($p = 0.013$). Similar differences were observed in trained mice. I did not find a significant difference between the beginning and the end of trained sessions ($p = 0.313$). However, there was a modestly significant difference between the beginning and middle of trained sessions ($p = 0.049$) and a more robust difference between the middle and end blocks ($p = 0.0027$). I did not find any significant effect of task type on performance during training ($F_{(1.35,2.70)}$=0.80, $p = 0.495$) or in trained mice ($F_{(4,38)}$=1.50, $p = 0.221$).

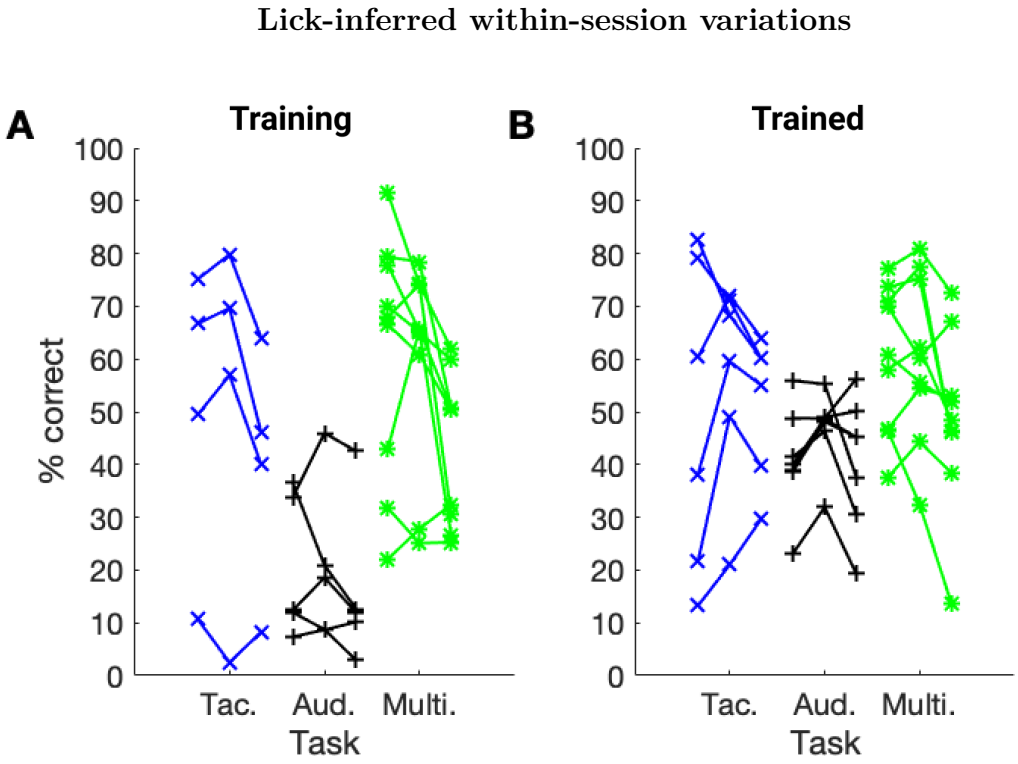



**Figure 2.15:** **A.** Within-session changes in performance (% correct) during training (performance <70% correct) inferred using a 2.5Hz lick rate threshold for a behavioural response. Performance is plotted separately for each of the 3 tasks. For each task, data are shown for the first (left dot), second (middle dot), and final (right dot) third of trials within a session. Tactile task: $N = 4$; Auditory: $N = 5$; Multisensory: $N = 9$. **B.** Within-session changes in lick-rate-inferred performance (% correct) after training on each task (performance >70% correct). Tactile task: $N = 6$; Auditory: $N = 7$; Multisensory: $N = 9$.

While using the adjusted lick-rate-inferred performance abolishes the initially observed early increase in performance during sessions, the adjusted metric also highlights a peak performance in the middle of both trained and training sessions. Further, using this criterion reveals the variability in licking patterns mentioned in Figure **2.12**, as the inferred performance of mice varied across individuals; even sometimes remaining below chance, which suggests higher lick-rates during No-Go trials than Go trials. Therefore, using an adjusted behavioural output metric appears to at least partly control for within-session improvements in performance, which are

likely linked to decreasing thirst during a session. However it also highlights the wide range of behavioural outputs exhibited by different individual animals.

### 2.3.10 Modality preferences during multisensory training

An underlying aim of the task design was to make a task that is not in itself perceptually difficult. Indeed, the individual sensory components of the audio-vibrotactile pairs are independently readily discriminable. A second critical aim was to ensure that mice learned to rely on both sensory modalities to learn to categorise multisensory stimuli. Thanks to the two incongruent pairs presented to mice among the No-Go trials, I was able to present mice with ambiguous stimuli in which one of the unisensory components was identical to that in the reinforced pair. In order to track whether mice trained on the multisensory task relied more on information acquired from one modality over another, I devised a sensory bias index (BI) (Section **2.2.11**) which allowed me to investigate whether mice performed more FAs on trials in which the auditory or vibrotactile component matched the Go stimulus. For example, a mouse relying exclusively on auditory information would perform FAs on trials with the same tone presented as the Go pair but not on those with the same vibrotactile frequency but a different tone. Such a mouse would have a BI of 1.

However, in trained mice, the FA rate was low, so the BI might not be a reliable measure for more subtle biases as there could be fewer than 10 FA trials. Hence, I defined a FA rate threshold for sessions to be included in the BI calculation. To choose an appropriate threshold, I examined the cumulative distributions of FA rates in the last 3 days of training, in which all mice had a performance over 70% correct, and compared this to the FA rates in the 7 days leading up to the mice

reaching 70% correct (Figure **2.16**). In trained mice, 50% of sessions had a total FA rate below 30%, while 20% of sessions in training mice were below that value. Therefore, on days in which the total FA rate was below 30% of all No-Go trials, I did not calculate a BI.

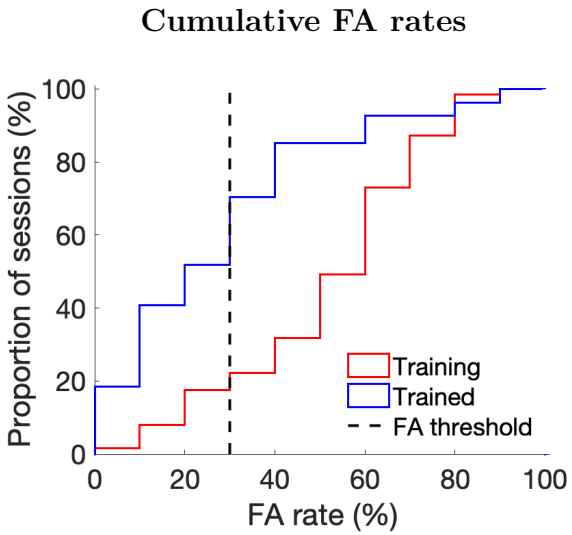



**Figure 2.16:** Cumulative FA rates during trained (performance >70& correct; blue) and training (performance <70% correct; red) sessions. The black line shows the 30% threshold, below which (left side) 50% of trained sessions and just under 20% of training sessions fell. $N = 9$ mice.

BI values fluctuated on a daily basis for every mouse. On the last day of training, three mice had no BI. The average BI was -0.126 $\pm$ 0.421 ($N = 6$ mice). I used a one-sample $t$-test to verify that the mean BI did not differ from 0 ($t$ =-0.73, $p$ = 0.496)(Figure **2.17**). Therefore, mice did not rely more heavily on auditory or tactile components of multisensory stimuli. However, this does not mean that individual mice did not develop a bias toward a frequency. Three out of six mice developed a bias toward the same modality for three days after reaching the performance threshold (one toward the auditory component). This was not consistently toward the same modality nor was it strong enough that the mice could not reach the performance threshold of 70% correct.

Hence, mice trained on the multisensory task did not develop a consistent bias toward one sensory modality in particular and relied equally on information from both the auditory and vibrotactile components of the stimuli. Two mice that had previously been trained on the tactile task (M1 and M4) followed a similar pattern of inconsistent biases fluctuating around 0, suggesting that previous training on a unisensory task did not make them bias towards that modality.

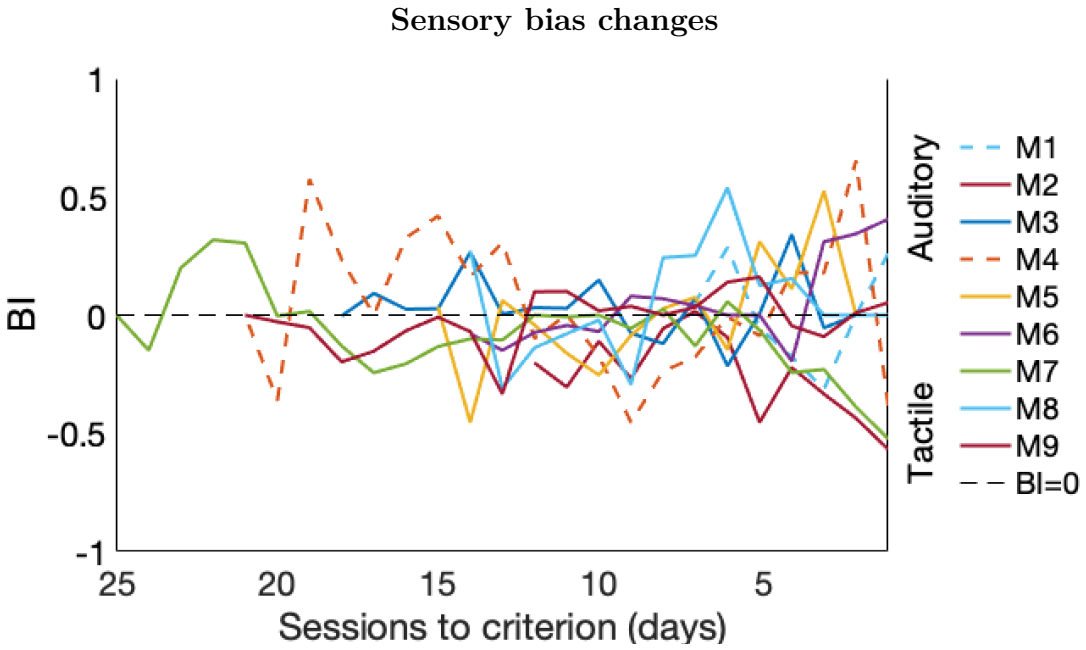



**Figure 2.17:** Sensory bias during multisensory training ($N = 9$ mice). Positive and negative bias values correspond to auditory and touch biases respectively. On days in which the FA rate was below 30% of all No-Go trials, no bias was calculated (plotted at BI=0 on the graph). Each coloured line corresponds to an individual mouse's bias over the course of training. Dashed lines represent mice trained previously on the unisensory vibrotactile version of the task ($N = 2$ mice).

Therefore, mice did not develop a strong overall bias toward one modality over an other and learned in fact to combine the information from separate modalities to identify pairs of stimuli. In other words, the fact that no unisensory stimulus component could reliably inform on the nature of audio-vibrotactile pairs was successful in training the mice to rely equally on separate modalities.

## 2.4 Discussion

### 2.4.1 Overview

In this Chapter, I developed a simple behavioural task on which mice can be quickly trained to categorise stimuli. First, I adapted a previously used Go/No-Go task structure (for instance, by assessing the use of time-outs after FA trials) for training on a multisensory task. I relied on this initial task structure to implement a novel vibrotactile stimulus and design a paradigm suited to compare unisensory and multisensory performance.

Second, I confirmed the ability of mice to discriminate between the auditory and vibrotactile frequencies used, as well as their ability to discriminate between pairs of these stimuli. Importantly, I confirmed the fact that the vibrotactile stimuli were indeed constrained within the somatosensory modality and that mice trained on the tactile version of the task relied on contact between the shaker and their whiskers. This was necessary to tightly control the separate unimodal constituents of the multisensory pairs.

Finally, I highlighted particular behavioural patterns during training, consistent across task types, showing both within- and across-sessions variations during training. I further noted the variety of ways in which attentional lapses and compulsive licking have been accounted for in previous studies. These have highlighted the inherent variability and difficulty associated to mouse training and I therefore suggested one possible way of reliably training water-regulated mice in a manner that would take into account their natural compulsion to lick.

### 2.4.2 On the importance of designing an appropriate stimulus

Initial training on a pole position discrimination task allowed me to adjust several parameters of the protocol, fine tune rules of the task (for instance adding a time-out after FA trials), and test the quality of cranial window implant and headplate fixation for imaging.

While the positional pole stimulus has been used on several occasions [Guo et al., 2014, Peron et al., 2015, O'Connor et al., 2010a], I illustrated several of its shortfalls, particularly for its use in research on multisensory integration. First, early observations in our lab revealed potential cross-modal perception of tactile stimuli in mice, relying either on audition or vision to respond to motor rotation even without contact against the whiskers. Second, as mentioned in Section **2.3.3**, the pole makes contact with the whiskers at different times for different locations. Some potential ways around these include using a separate motor to first move the pole into the right rostral-caudal position above the whiskers and then dropping it in place [Guo et al., 2014], thus avoiding brushing through whiskers. This would further help by rotating the motor to position the pole out of reach of mice hearing. However, the timing of contact with the pole can still vary by a few milliseconds depending on the mouse's whisking behaviour, or lack thereof (e.g. a whisking mouse may hit the pole with its whiskers with a small lag depending on whether it is protracting or retracting its whiskers). Presenting different surfaces for a texture discrimination task [Arabzadeh et al., 2005, Chen et al., 2015] also introduces a timing issue, as a non-whisking mouse would not perceive a surface's texture. In fact, as mentioned previously, stimulus synchrony is understood to support enhanced multisensory discrimination performance [Kim et al., 2008].

As illustrated in Figure **2.2**, the vibrotactile stimulus I designed avoids each of the above problems. It allows us to precisely synchronise the onset and offset of auditory and tactile stimuli. Recordings of the vibrotactile stimulus showed they they do not produce sound artifacts (Section **2.2.10**). Due to the technical range of our equipment, I was not able to record potential artifacts across the entire hearing range of the mice (>40kHz), but our control experiments in which we presented the vibrotactile stimulus outside the reach of the whiskers (Section **2.3.4**) further suggested that mice were not relying on any cross-modal artifacts. Hence, the use of a vibrotactile stimulus appears to be particularly well suited to audio-tactile tasks.

Another advantage of the vibrotactile stimulus is that since it was composed of a regular oscillation rate, I could use common modulation rate in my tones to increase congruence and synchrony across modalities. This aim was based on reports of increased discrimination performance and enhanced neural responses to congruent multisensory stimuli compared to incongruent pairs [Kim et al., 2008, Maddox et al., 2015, Atilgan et al., 2018]. This was useful in the present study as it was unclear if the multisensory categorisation task might prove too difficult for mice to learn. In previous examples of multisensory training in mice, such as [Siemann et al., 2015], the task design did not allow for this added congruence consideration. It would be interesting for future experiments to train mice to categorise audio-vibrotactile pairs that do not contain a common modulation frequency, as it remains unclear if this is a necessary component their performance on this task.

### 2.4.3 Potential effects of hearing loss in mice

Mice took longer to learn the multisensory task than the tactile or auditory tasks. This suggests that the multisensory task may either have been more difficult (either cognitively or perceptually) than its unisensory counterparts (see Section **2.4.5**), or one in which animals struggled to maintain attention. Given that mice trained on the auditory task were younger than trained on the vibrotactile and multisensory tasks, it is possible that the older animals struggled to learn the auditory component of the multisensory task. A perceptual difficulty could arise from early onset hearing loss, which the C57BL/6J strain of mice are known to experience from about 12 weeks onward, starting with high frequency sounds (above 30 kHz) [Ison et al., 2007]. However, given the relatively low sound frequencies I presented (6.3 and 12.5 kHz), I do not expect this to be an issue [Ison et al., 2007]. Furthermore, my sensory bias analysis did not indicate that mice were more likely to make auditory errors than tactile errors on the multisensory task. Still, the possible effects of age-related hearing loss could be tested in future experiments by training the same strain of mice on the multisensory task at an earlier age, or by training a strain of mice that are known to not experience early onset hearing loss, such as CBA mice [Kane et al., 2012].

### 2.4.4 On the importance of using an appropriate measure of behavioural choices

In this chapter, I assessed measured performance using a binary behavioural output on each trial: did the mouse lick – even once – or not in the 3 seconds following a stimulus presentation? This measure allowed me to track changes in performance

and characterise improvements in task acquisition across unisensory and multisensory tasks.

Based on a binary 'Lick/No-Lick' measure, I observed increases in the number of licks per Go trial compared to No-Go trials over the course of training as well as within-session improvements in performance (Sections **2.3.8** and **2.3.9**). I reported patterns of performance improvement within sessions, consistent with previously reported observations of 'overmotivation' resulting in high Hit and FA rates at the beginning of sessions and satiation resulting in higher Miss and CR rates toward the end [Berditchevskaia et al., 2016]. These changes could be attributed to increased thirst at the beginning of sessions, leading to a reduce restraint capacity and increased FA rate. As a mouse's behaviour evolves during a session, it would be sensible to design a behavioural output measure that varies to reflect this evolution.

Different strategies have been previously developed to account for changes in satiety and motivation during a behavioural session. For instance, Guo and colleagues [Guo et al., 2014] adjusted the position of lick spouts at the start of sessions in order to make it more difficult for mice to lick, thereby reducing the lick rate. They acknowledged, however, that this protocol could risk discouraging mice if the spout were positioned too far out of reach. Furthermore, this approach did not take into account within-session changes in satiety. Another strategy described in 2018 [Meijer et al., 2018] consists of interleaving task trials with blank trials, in which no stimulus is presented. Meijer and colleagues then calculated a sliding FA rate by calculating the number of blank trials on which a mouse licked. They excluded blocks of 100 trials with a FA rate over 50% from their data analysis. Finally, an earlier study introduced a no-lick period before stimulus generation in order to avoid initiating trials during bouts of compulsive licking [Mayrhofer et al., 2013].

The diversity in methods accounting for compulsive licking illustrates the universality of the issue in training mice (and other rodents) on a Go/No-Go task. Adjusting the spout position on a daily basis to manually account for impulsivity was at first a tempting approach. However, changing the relative position of the spout may in turn change the level of motivation required to respond to stimuli, introducing performance changes that are unrelated to stimulus categorisation learning. The alternative strategy of discounting blocks with high proportions of FA on blank trials can avoid issues with extended periods of compulsive licking. However, it would likely also discard trials in which an animal performed adequately, while including shorter bouts of compulsive licking that do not reach the FA exclusion threshold. Lastly, introducing a no-lick period before stimulus presentation is likely to prevent animals from initiating trials during bouts of spontaneous licking, but it faces another potential shortfall. It would not account for such compulsive licking to stimulus onset that might impair discrimination performance. Indeed, increased thirst may not simply correlate with spontaneous licking but also with impaired discrimination.

There was no significant difference between lick rates on Hits and FAs, even after the mice were trained (Section **2.3.8**). In other words, mice responded to No-Go stimuli in FA trials as if they were Go stimuli. One possible interpretation for this finding at the beginning of training, is that mice responded when they detected the onset of stimuli, as opposed to discriminating between different stimuli. This appears to be supported by the fact that licks did not occur before trial generation (Figure **2.12**). In this instance, the control used by Mayrhofer and colleagues' (described above; [Mayrhofer et al., 2013]) would not account for all compulsive FA trials. Similarly, Meijer and colleagues' [Meijer et al., 2018] strategy may account for some extended bouts of such trials although, as aforementioned, this would result

in a loss of data on trials in which animals engage in the task and include other shorter bouts of compulsive licking. The persistence of high-lick FA trials in trained mice could indicate that mice either failed to remember the categories, failed to inhibit their response due to high thirst motivation, or failed to discriminate between Go and No-Go stimuli. Given the observation that Hit and FA rates decreased throughout a given behavioural session in both training and trained mice (Figure **2.14**), and the fact that trained mice performed well overall, the second interpretation seems most likely.

In Sections **2.3.8** and **2.3.9**, I suggested using a 2.5 Hz lick rate threshold instead of a binary lick response (i.e. a threshold of 1 lick) as a measure of behavioural choice. Using the lick rate threshold of 2.5 Hz, we found many of the same improvements in performance (i.e. % correct) during training (Figure **2.13**). I demonstrated that the lick rate threshold compensated for excessive licking in early stages of the task (i.e. when mice could not refrain from licking) and avoided the observed within-session improvements in performance that may not be due to an improvement *per se*, but rather reduced lick compulsion. However, applying this threshold on data obtained in mice rewarded according to the single-lick criterion also highlighted that mice considered to be 'trained' did not necessarily perform better under this new criterion at the end of the training protocol.

Specifically, mice who learned for example to withhold licking until the end of a stimulus and lick only during the response window would be considered to perform poorly by using this 2.5Hz criterion. However, it is not ideal to analyse the animal's behaviour using a different choice metric than they were trained and rewarded with. To better examine the effects of licking metrics, further studies in the lab could apply a variable lick-rate threshold during training sessions. A higher threshold at the

beginning of sessions would account for high compulsive lick-rates and may better capture categorisation performance than a 1-lick criterion. This is because mice may occasionally fail to completely withhold licking on a No-go trial, but stop licking earlier than in a decisive lick response to a Go stimulus. Then, as a session progresses, the threshold could decrease in order to account for reduced thirst and compulsion. The lick threshold could be guided by FA rates, varying as in [Berditchevskaia et al., 2016]. The task would remain sensibly identical for mice throughout the session, but control for satiety-based motivation decreases. At all points in the task, the mouse would have to learn to deliberately lick to a Go stimulus and do their best to withhold from licking when presented with a No-Go stimulus. I would not expect it to be more difficult for mice to learn with this variable lick criterion, as Figure **2.12** highlights the fact that some even learned to withhold licking before the response window.

I have illustrated the ability of mice to learn both the unisensory and multisensory versions of the task, and it is key to note here that there were no apparent differences in within-session variations or licking patterns between mice trained on the different tasks. This suggests that these variables reflect general processes of motivation and behaviour in these animals, which will affect performance on any sensory discrimination task.

### 2.4.5 Multisensory task difficulty: Cognitive or Perceptual?

Overall, mice exhibited similar performance and licking behaviours on all three versions of the Go/No-go task. However, mice trained on the multisensory task took significantly longer to reach criterion than those trained on a unisensory task, both in terms of the total number of training sessions and trials. Why might this be?

A first instinctive answer is to argue for increased task difficulty. But there are a number of ways in which a sensory discrimination task may become more difficult, and I explore several of these different types of 'difficulty' here in turn.

Typically, previous studies investigating sensory discrimination have increased task difficulty by making the stimuli more perceptually similar to each other. This could be achieved by presenting more proximate pole positions to discriminate in somatosensory tasks [Guo et al., 2014, Peron et al., 2015], closer frequencies in auditory tone discrimination tasks [De Hoz and Nelken, 2014], or reducing the contrast of stimuli in a visual orientation discrimination task [Long et al., 2015]. In such tasks, animals can reliably learn a set of rules and perform at a high level with easily discriminated stimuli and their performance drops to chance when stimuli are made harder to distinguish – even though at no point does the task structure or rules change. These are examples of tasks being made perceptually difficult while remaining cognitively identical. In contrast, the stimuli used in my multisensory task were combinations of the same ones used in the unisensory tasks. They were readily perceptually discriminable when trained in a unisensory task, and should remain so in the multisensory version. Therefore, we cannot attribute the slower learning on the multisensory task to pure perceptual discrimination thresholds within a modality.

Nevertheless, the multisensory task did include stimuli that were perceptually ambiguous - namely, the No-go stimuli that resembled the Go stimulus in one modality ($A^+T^-$ and $A^-T^+$, see Table **2.1**). Therefore, while the perceptual 'distance' within a modality was equivalent in the multisensory and unisensory tasks, the multisensory stimuli were more similar in 'multisensory' space. One could control for this in future experiments by adding a second feature dimension to the unisensory tasks. For example, varying the tone across two frequencies and amplitudes, and only rewarding

one frequency and amplitude combination.

In the multisensory task, mice were required to combine information about stimuli across the auditory and tactile modalities. This could either involve perceptually 'fusing' them into a single multisensory 'object', or simply remembering the combination of both unisensory 'objects'. Either of these two processes would put an extra demand on the mice in order to perform the task correctly, and could account for the slower acquisition rates compared to the unisensory task. We are not able to determine if multisensory fusion took place with our given data. But we do observe that the mice successfully learned that neither modality alone was sufficient to predict the water reward, and that instead both sensory modalities were equally important to getting the reward. The sensory biases report in Section **2.3.10** highlight the fact that mice did not overwhelmingly or consistently rely on one sensory modality over an other. Instead, the BIs fluctuated around 0 during training, including in mice trained previously on a single modality. This lack of a consistent bias on FA trials suggests that mice do perceptually integrate information from the two modalities on our task, because if they treated the modalities as separate one would expect them to exhibit stronger or more consistent biases, especially when previously trained on a unisensory task with the same stimuli.

Finally, the multisensory task had a higher cognitive load than the unisensory tasks. Although in all three tasks mice had to learn that a single stimulus was reinforced, the mice were presented with four different stimulus pairs in the multisensory version, and only two in the unisensory tasks. Therefore, there were more stimulus categorisations to remember and more modalities to attend in the multisensory task, increasing the cognitive demand on the multisensory version. The suggestion above of adding a second feature dimension to the unisensory task (e.g.

stimulus amplitude) would also control for this increase in cognitive demand in the multisensory task.

While the multisensory task took longer for mice to learn, 9 of 11 mice successfully reached criterion in 13 days. Therefore, the aim of implementing a cognitively demanding multisensory task (i.e. the learning of a multimodal combination identification rule) was successfully achieved.

### 2.4.6 Conclusions

This novel multisensory task is grounded in previously established training methods and based on a newly established vibrotactile stimulus. It further builds upon them to enable a new approach to investigating neural responses to audio-vibrotactile stimuli while mice learn to categorise a novel multisensory object. In the next Chapter, I will therefore characterise neural responses to this novel stimulus and measure changes in responses to the different stimuli during training on both unisensory and multisensory versions of the task.

# Chapter 3

# Effects of training on the neural representations of unisensory and multisensory stimuli in somatosensory cortex

## 3.1 Introduction

In the previous Chapter, I have introduced a novel vibrotactile stimulus suited for the implementation of an audio-vibrotactile categorisation task in head-fixed mice. Mice were able to learn to categorise auditory, vibrotactile and audio-vibrotactile stimuli and several controls ensured that: 1) stimuli contained no cross-sensory artifacts; and 2) mice could learn to categorise multisensory stimuli without relying on one modality over another (see Section **2.3.10**).

Using this task, I now seek to investigate how these novel stimuli are represented in the vibrissal primary somatosensory cortex (vS1), and how those representations change with training. While our understanding of the neural responses to static tactile stimuli is extensive, we are only beginning to understand the neural responses to vibrotactile stimuli. Moreover, it remains entirely unclear how information from different sensory modalities is integrated in vS1 and how this

processing may change with learning.

### 3.1.1 Tactile responses in vS1

Several studies have characterised whisker motion during active whisking [Isett et al., 2018, Clack et al., 2012]. Whiskers glide and stop against irregular surfaces in a motion described as 'stick-slip' [Diamond et al., 2008, Chen et al., 2015], in which whiskers catch on surface irregularities ('stick') and spring free ('slip'). Our understanding of subcortical and cortical processing of tactile information is discussed in depth in Section **1.2.1**. The first region of cortical tactile processing, vS1, is laid out in a somatotopic fashion, characterised by barrel-like groups of neurons that respond to stimulation of individual positions on the whisker pads [Woolsey and Van der Loos, 1970].

The cortical correlates of tactile perception have been studied for a long time [Petersen, 2007] (see Section **1.2.1**), using a wide array of tactile stimuli to measure responses, such as positional pole stimuli [Peron et al., 2015, Guo et al., 2014], textured sandpaper [Chen et al., 2015], and piezoelectric drivers deflecting threaded whiskers [Helmchen et al., 2018]. While subcortical trigeminal and thalamic neurons respond robustly and consistently to whisker deflections [Arabzadeh et al., 2005], this does not appear to be the case of cortical neurons. It is now understood that vS1 neurons are sparsely driven by tactile inputs [Petersen, 2007, Jadhav and Feldman, 2010, Barth and Poulet, 2012]. This can be observed through the small proportion of responsive cells [Chen et al., 2015, Clancy et al., 2015] as well as the fact that responses to whisker deflections typically consist of a single action potential [De Kock et al., 2007] or the observation that single-cell stimulation in vS1 can bias an animal toward detecting a stimulus [Houweling and Brecht, 2008]. This

sparsity could be explained by: 1) Variability across trials (from cell-intrinsic or network properties) creates the appearance of sparseness that is diminished with longer time-intervals for analysis. 2) A small but constant subset of cells fires more reliably across many trials than the rest of the population. 3) The fraction of firing neurons can be changed by brain state, such as alertness or motivation [De Kock and Sakmann, 2009]. 4) Highly stimulus-specific neurons will create the appearance of sparseness; as stimulus space expands, more neurons will fire (for review, see [Barth and Poulet, 2012]).

The exact proportion of vS1 neurons responsive to whisker touch has been a point of contention. Some findings suggest that a low proportion (20 to 50%) of neurons in vS1 typically respond to tactile stimuli presented on whiskers [Peron et al., 2015]. In contrast, Chen and colleagues found up to 60% of active neurons to respond to tactile stimuli [Chen et al., 2015] while another study identified only 16% of responsive neurons to passive whisker deflections [Zhang et al., 2020]. Another study found 64% of touch-responsive neurons in L2/3 and 57% in L5 of running mice [Ayaz et al., 2019]. This discrepancy in findings across studies may partly result from the sensitivity of vS1 neurons to 'slips' [Jadhav and Feldman, 2010, Isett et al., 2018]. Chen and colleagues found more responsive neurons in a study in which awake mice learned to discriminate between textures, while Zhang and colleagues, who investigated responses to passive whisker deflections, report fewer responsive neurons [Chen et al., 2015, Zhang et al., 2020].

Another possible cause for these different results may be differences in the design of the aforementioned studies. For example, Peron and colleagues imaged mice with only one intact whisker [Peron et al., 2015], while other studies imaged mice with whole whisker pads intact [Zhang et al., 2020, Chen et al., 2015]. This is

of particular relevance here, since it was originally thought that all neurons within a single barrel respond to deflections of its associated whisker. However, the broad somatotopic structure of vS1 gives way to more local, 'salt-and-pepper' distributions of whisker preferences [Clancy et al., 2015, Kerr et al., 2007]. While neurons located closer to the center of a barrel may respond preferentially to its associated whisker, neurons further toward the edge of barrels could in fact respond more strongly to deflections delivered to other, often neighbouring, whiskers [Clancy et al., 2015].

In observing neural responses to tactile stimuli, one must bear in mind the fact that vS1 does not simply reflect tactile information alone, but rather a plethora of motor-related features [Crochet and Petersen, 2006, Ayaz et al., 2019, Saleem et al., 2013]. Whisker-guided tactile exploration is a sensorimotor process, during which mice actively whisk to identify and map their surroundings. The vibrissal area of M1 is directly connected to facial nucleus motor neurons [Grinevich et al., 2005] and direct stimulation of M1 is sufficient to elicit whisker movements [Brecht et al., 2003]. It is however worth noting that M1 is also extensively and reciprocally connected to vS1 [Aronoff et al., 2010, Petreanu et al., 2012]. As such, motor artifacts may affect vS1 responses to tactile stimuli. Using whole-cell patch-clamp recordings in vS1 of awake mice, Crochet and Petersen found that active whisking generated high frequency neural fluctuations that were correlated with whisker movements, even when these movements did not lead to any contact with an object (i.e. only motor action) [Crochet and Petersen, 2006]. In addition to whisking, running can also effect cortical responses to tactile stimuli. A recent study measuring neural responses in vS1 of mice using a tactile virtual reality setup found neurons whose responses were modulated by running and concurrent whisking [Ayaz et al., 2019], highlighting a stronger modulatory effect of whisking

with concurrent running. Locomotion has also been reported to modulate responses in the primary visual cortex (V1) [Saleem et al., 2013], and the primary auditory cortex (A1) [Schneider et al., 2014]. More specifically, the influence of locomotion on neural responses varies depending on the recorded cortical area. While Ayaz and colleagues found an increase in neural activity in S1 due to locomotion, it has in contrast been found to suppress responses in A1. Visual speed and locomotion were found to be integrated in V1. Therefore, I expected that vS1 responses to the novel vibrotactile stimulus would be modulated by behavioural features such as whisking or running.

### 3.1.2 Training effects in vS1

Despite the wide variety of neural recordings measuring responses to an array of tactile stimuli, it remains unclear how the vibrotactile information used by mice to train on the novel task described in Chapter **2** is processed in vS1. In rats, vS1 neurons have been found to respond to whisker deflections generated with custom-built ladders attached to a piezoelectric wafer at spike rates proportional to the product of the frequency and amplitude of deflection [Arabzadeh et al., 2003]. Rats have also been shown to discriminate between vibrotactile stimuli in a manner that is also dependent on the product of the amplitude and frequency of deflections [Adibi et al., 2012]. Mice have been successfully trained on vibrotactile discrimination tasks that use piezoelectric drivers to deflect threaded whiskers at different frequencies [Mayrhofer et al., 2013].

I aimed to measure how neural responses to auditory and vibrotactile stimuli changed with training by performing daily 2-photon calcium imaging on the same cell populations for several weeks. The ability to image the same cortical volume

on a chronic basis was first described in 2012 [Minderer et al., 2012], and has since been widely used [Margolis et al., 2014, Ceballo et al., 2019, Poort et al., 2015]. Several studies have investigated how vS1 responses to other types of tactile stimuli change with training [Peron et al., 2015, Mayrhofer et al., 2015, Chen et al., 2015, O'Connor et al., 2010b]. Chen and colleagues found that as mice were trained to discriminate between textures presented to the whiskers, the proportion of responsive neurons trended upwards over time with increased task performance [Chen et al., 2015]. The same year, Peron and colleagues studied the representation of a pole's azimuthal position within the whiskers as mice were trained on a spatial discrimination task [Peron et al., 2015]. They tracked responses from individual cells and found that throughout training, a cell's response to stimuli predicted its own responses on the next training day. The ability to observe population-wide changes and track individual vS1 cell responses to stimuli throughout tactile training is therefore well established.

### 3.1.3 Multisensory encoding in primary sensory cortices

Research into the neural correlates of multisensory integration has historically focused on subcortical processes – such as superadditive responses in superior colliculus (SC) neurons to audiovisual stimuli [Meredith and Stein, 1986] – or 'higher' associative cortical areas [Olcese et al., 2013]. However, over the past two decades multisensory responses have also been extensively described in primary sensory cortices. The majority of studies that describe multisensory or cross-sensory responses at a cellular level have examined audiovisual integration. For instance, as early as 2007, recordings in the A1 of ferrets found bimodal neurons, whose responses to spatial auditory cues were modulated by the position of a visual stimulus. Overall, just

under 15% of the recorded neurons exhibited either bimodal or crossmodal responses (that is responses to audiovisual or visual-only stimuli, respectively). Similarly, a recent study of V1 in awake mice found that a subset of neurons were more sharply tuned to the orientation of visual gratings when presented with temporally congruent auditory stimuli [Meijer et al., 2017]. These cross-modal interactions may be at least partly provided by direct cortico-cortical connections. There are direct anatomical inputs from A1 onto L1 interneurons in V1, which functionally sharpen the tuning curves of L2/3 V1 neurons [Ibrahim et al., 2016]. The effects of learning on multisensory responses in the primary sensory cortices are now beginning to be studied. A recent study tracked the responses of individual V1 neurons in mice and found that audio-visual pairing enhanced the responses of canonically visual neurons to auditory stimuli, a feature not observed when unpaired auditory stimuli were presented [Knöpfel et al., 2019]. They further found increased correlations between multimodal neurons after pairing stimuli, suggesting that changes occur within local V1 networks as a consequence of multisensory experience.

In comparison, the study of multisensory integration in vS1 remains preliminary. While a building body of research has described auditory-tactile integration in humans [Foxe et al., 2000, Pannunzi et al., 2011, Young et al., 2017] and higher animal models [Bizley et al., 2007, Lakatos et al., 2007] (see Section **1.3**), I will focus the scope of this Chapter on rodents, and particularly mice. As discussed in Chapter **2**, a recent study investigated the integration of auditory and tactile information in vS1 of mice using two-photon imaging [Zhang et al., 2020] and found a modulatory (mainly inhibitory) effect of sounds on somatosensory responses. However, a core limitation to that study is the fact that the piezoelectric driver used to deflect whiskers generated resonant sounds within the mouse hearing range.

Using similar methods, another study reported results from whole-cell recordings in rat and mouse S1 highlighting subthreshold postsynaptic responses to auditory stimuli, suggesting a modulatory effect of auditory stimuli on S1 neurons and, yet again, hinting to an absence of suprathreshold responses to auditory stimuli alone [Maruyama and Komai, 2018]. Audio-tactile integration in vS1 of mice is yet to be conclusively elucidated and, to my knowledge, no study has investigated how training on a multisensory task may alter neural responses in vS1 in mice.

Therefore, there is a wide body of research on tactile encoding, against which we can compare observations of neural responses to the vibrotactile stimulus presented in Chapter **2**. Furthermore, a growing body of evidence describes that multisensory integration occurs as early as primary sensory cortices, and that experience may shape these multisensory responses. Yet, to date, no study has tracked neural responses in vS1 to vibrotactile and auditory stimuli over time, nor described audio-vibrotactile interactions in this region. Finally, it is unknown how audio-vibrotactile responses may change with training on discrimination tasks.

### 3.1.4 Objectives

The work presented in this chapter aimed to:

**1**. Measure responses of vS1 neurons to a novel vibrotactile stimulus;

**2**. Measure responses of vS1 neurons to audio-vibrotactile stimuli; and

**3**. Track changes in vS1 responses on a population and individual neuron level over the course of multimodal combination identification training.

I used 2-photon calcium imaging in awake mice passively exposed to stimuli to measure the responses of neurons in this study. Given the results of the wide variety of studies characterising neural responses to somatosensory stimuli, I predicted that I would find sparse population responses to vibrotactile stimuli in vS1. The fraction of responsive neurons may vary from 20 to 50% of imaged cells. Based on previous findings in rats [Arabzadeh et al., 2003, Arabzadeh et al., 2005], I expected that some neurons would present frequency-tuning characteristics. Based on few previous studies available, I expected that the proportion of neurons responsive to auditory stimuli in the vS1 of mice would be very low and that sound presentation would instead modulate responses to vibrotactile stimuli. Finally, similar to Peron and colleagues' work [Peron et al., 2015], I hypothesised that training would have little effect on the overall population activity but stabilise the session-to-session variability of select individual neurons.

## 3.2 Methods

### 3.2.1 Subjects

Seven mice from the training cohort described in the previous chapter (Section **2.2.1**) were imaged (mice T4, T5 and T6 trained on the tactile task; mouse A1 trained on the auditory task; and mice M5, M8 and M9 trained on the multisensory task). All mice were imaged in one passive session before training. In these sessions, the mice were not required to perform the task and were not given a water reward or time-out upon presentation of any stimulus. The water delivery spout was kept out of reach in these sessions. Trained mice then underwent a second passive imaging session after training was complete. At this point, they were off water-regulation and

the spout was kept out of reach, so that again no behavioural task was performed. Two mice were imaged in the initial passive conditions but subsequently failed to reach criterion on the task. These were therefore included in data acquired in the pre-training passive conditions but are not part of the longitudinal analysis cohort (i.e. mice imaged across training until reaching performance criterion).

### 3.2.2 Surgical procedures

All surgical procedures were identical to those described in Section **2.2.2**. Briefly, mice were implanted with custom-made titanium headplates for head-fixation. Craniotomies were performed and chronic imaging windows were implanted under general anaesthesia. Customised imaging windows were designed by stacking one 3 mm wide and one 4 mm wide round coverslips (total thickness 0.45 mm), held together with an ultraviolet-cured optical glue (NOR-81; Norland, Edmund Optics). A 3 mm diameter craniotomy was made over vS1 in the right hemisphere (position of centre relative to Bregma: 1.3 mm posterior and 3.2 mm lateral). The imaging window was made by placing the coverslip into the craniotomy, with the 3 mm coverslip fitting into the craniotomy and the 4 mm coverslip glued around the lower edge to the surrounding bone. The stacked coverslip replaced the bone, stabilising the surface of the brain and protecting the brain from potential sources of infection or damage.

### 3.2.3 Behavioural training

The behavioural protocol was identical to that described in Section **2.2.5**. Briefly, each mouse was first water-regulated and handled for three to five days in order to stabilise its weight around 85% of its initial weight and habituate it to

handling. It was handled and given ~1-1.5 mL of water every day. It was then habituated to the testing apparatus for three to five days before training began. During the reward-stimulus pairing, either an auditory, vibrotactile or multisensory stimulus was presented to the mouse (vibrotactile stimuli consisted in vibrations against mouse whiskers). One of the stimuli, the Go stimulus, was automatically rewarded after every presentation (delivery of a 15-20 $\mu$L drop of water from a water spout positioned directly in front of the mouse's mouth). If the mouse licked on No-Go stimulus presentations (FA), no reward or punishment were delivered. After three to five days of stimulus-reward pairing, mice started the training phase. A Hit was defined as licking within the 3 s window following presentation of the Go stimulus. A Correct Rejection (CR) was an omission of licking during the response window following a No-Go stimulus. A Miss was a failure to lick within this window in response to the Go stimulus, and a FA occurred when the mouse licked during the response window on a No-Go trial. Hits resulted in a ~15-20 $\mu$L water reward. During the training phase, a FA trial was followed by a 3 s time-out. In order to avoid mice developing a strategy of continuous licking, the time-out restarted if the mouse licked during the 3 s. The behavioural analysis methods were the same as described in Section **2.2.11**.

### 3.2.4 Imaging

I imaged mice following the procedures described in Section **2.2.12**, in short: neuronal calcium transients in vS1 were recorded using a Sutter MOM two-photon microscope controlled by ScanImage v.5.3 software. A Mai-Tai laser (SpectraPhysics; 70 fs pulsewidth, 80 MHz repetition rate) emitted a light tuned to 900 nm to excite GCaMP during imaging. I used the resonant scanner in bidirectionnal mode at a

resolution of 512 × 512 pixels. The multiplexed image acquisition was obtained by means of an acousto-optic deflector (DTSX-400-980, AA Optoelectronics), which allows switching between two axially separated fields of view at kHz rates, limited only by the frequency of the resonant scanner (8 kHz) and the image resolution (512 × 512) – allowing for quasi-simultaneous imaging of two fields of view. Images were acquired from two planes separated by 120 $\mu$m quasi-simultaneously giving an acquisition rate per imaged plane of ~15 Hz) [Chong et al., 2019]. I used a 16X/0.80W LWD immersion objective (Nikon) and GaAsP photomultipliers (Hamamatsu) to detect emitted photons. The field of view was 650 × 650 $\mu$m and was imaged down to 400 $\mu$m below the cortical surface. Imaging analysis methods were the same as described in Section **2.2.13**.

### 3.2.5 Longitudinal tracking of neurons across imaging sessions

As described in Section **2.3.7**, mice took variable amounts of time to reach performance criterion. In order to compare changes in the proportion of responsive neurons or in the responses of neurons themselves over the course of training across mice, it was necessary to match the relative timing of learning across animals. Therefore, for each mouse, six steps of training and imaging were identified: pre-training passive imaging, habituation, early training (when the mouse performed <60% correct per session), acquisition (when the mouse performed at 60-70% correct), expert performance (when the mouse performed >70% correct) and post-training passive imaging. Therefore, regardless of the time taken to reach the performance criterion, each mouse generated six neural data sets that were matched according to training and performance stages.

On each day of imaging, the angle and position of the microscope objective were manually adjusted to image the same areas and track changes in responses from the same neurons across days. The objective angle and position from the first day of imaging was used as a guide. The cortical surface was identified daily, imaged briefly with two-photon imaging and compared to previous imaging sessions to ensure consistent imaging depth. In order to ensure correct tracking of neurons and reduce human error, I used the registers2p functionality in the Suite2P package [Pachitariu et al., 2016] to identify vascular and neuronal features (neurites and cell bodies) and register the images offline across days. Overlapping pre-labelled ROIs were identified in registered fields as the same neurons and a final manual correction (i.e. morphological comparisons) ensured that the same neurons were imaged across the experiment (i.e. 'tracked').

Despite my best efforts, not all neurons in the fields of view could be reliably tracked across all imaging sessions. I was able to track 45±28 (mean ± std across all imaged ROIs) neurons across all 6 steps of the imaging and training protocol. On any given day of training, these tracked neurons made up on average between 9.9-11.4% of all imaged neurons (see Table **3.1**). The number of neurons imaged in each step of training did not significantly vary, nor did the proportion of tracked neurons out of a given day's imaged cells ($F_{(5,30)}$=0.437, $p = 0.819$ and $F_{(2.98,17.85)}$=0.393, $p = 0.758$ respectively; repeated-measures ANOVA with Greenhouse-Geisser correction) (Table **3.1** and Figure **3.1**).

| | Training progress | | | | | |
|---|---|---|---|---|---|---|
| | Passive | Hab. | <60% | 60-70% | >70% | Post |
| Total neurons | 453.4 ±171.2 | 387.3 ±125.8 | 418.6 ±166.6 | 401.4 ±146.8 | 415.1 ±139.8 | 438.9 ±153.9 |
| Tracked neurons (%) | 10.0 ±4.2 | 11.4 ±5.1 | 10.4 ±2.8 | 10.8 ±4.1 | 10.1 ±3.0 | 9.9 ±3.5 |

**Table 3.1:** Number of neurons (mean ± std) at different stages of training and number of tracked neurons expressed as proportion of the total number of neurons on a given day (average proportion per mouse, $N = 7$ mice). The different stages of imaging were: pre-training passive imaging, habituation, <60% correct (early training), 60-70% correct (acquisition), >70% correct (expert) and post-training passive imaging. By definition, the number of tracked neurons remained constant across days for each mouse.

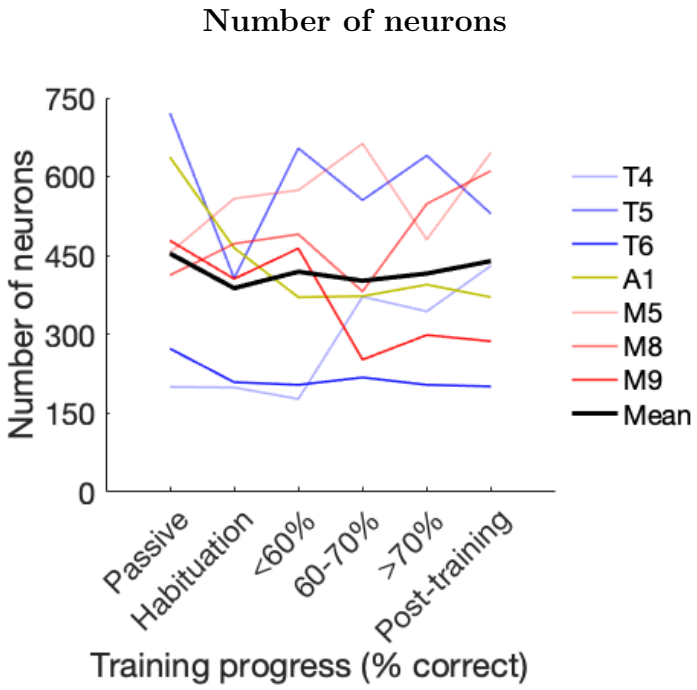



**Figure 3.1:** Number of neurons imaged at different stages of training across mice (tactile trained = blue; multisensory trained = red/orange; auditory trained = yellow). The number of neurons imaged daily per mouse is plotted as a separate coloured line. The average number of imaged neurons is plotted as the thick black line. The different stages of imaging were: pre-training passive imaging, habituation, <60% correct (early training), 60-70% correct (acquisition), >70% correct (expert) and post-training passive imaging. The proportion that these neurons make up of a given day's total neurons is plotted as a separate coloured line for each mouse.

### 3.2.6 Barrel cortex targeting with optical signal imaging

After the final post-training passive imaging session, I performed intrinsic optical signal imaging (IOSI) in all mice that were trained and imaged ($N = 9$). The animals were anaesthetised with chlorprothixene hydrochloride (0.2 mg/mL at 1 mg/kg) and isoflurane (2% in 1.5 L/min $O_2$), and were head-fixed by a clamp on their headplates on a heated mat. A single whisker was threaded into a capillary tube attached to a ceramic piezoelectric stimulator (PB4NB2W Piezoelectric Bimorph Bending Actuator with Wires, Thorlabs). The whisker identity was visually determined during threading based on its position on the whisker pad. IOSI was performed with a Retiga R1 camera set up with a 50 mm and 135 mm lenses (Nikon) attached in a tandem configuration [Juavinett et al., 2017]. The stimulation protocol consisted of 1 s stimulation at 10 Hz with 20 s ITI, repeated 40 times, for a total of 400 deflections. This stimulation protocol was repeated 3-4 times per mouse on different whiskers in order to map the barrel fields in vS1. For every mouse, a map of vS1 barrels was superimposed on an image of the surface of the brain acquired during two-photon imaging sessions (Figure **3.2**). This allowed me to locate the position of the imaging fields within vS1.

In all nine mice that trained on a version of the task, I was able to confirm that the areas imaged were within vS1. Although the spatial resolution of IOSI is limited, responses to whisker stimulation were typically observed between whisker rows B and E, and columns 3 and 6 (Figure **3.2**). While the imaged areas included both barrels and septa, these findings suggest that the majority of neural data obtained originated in primary barrels corresponding to macrovibrissae in vS1.

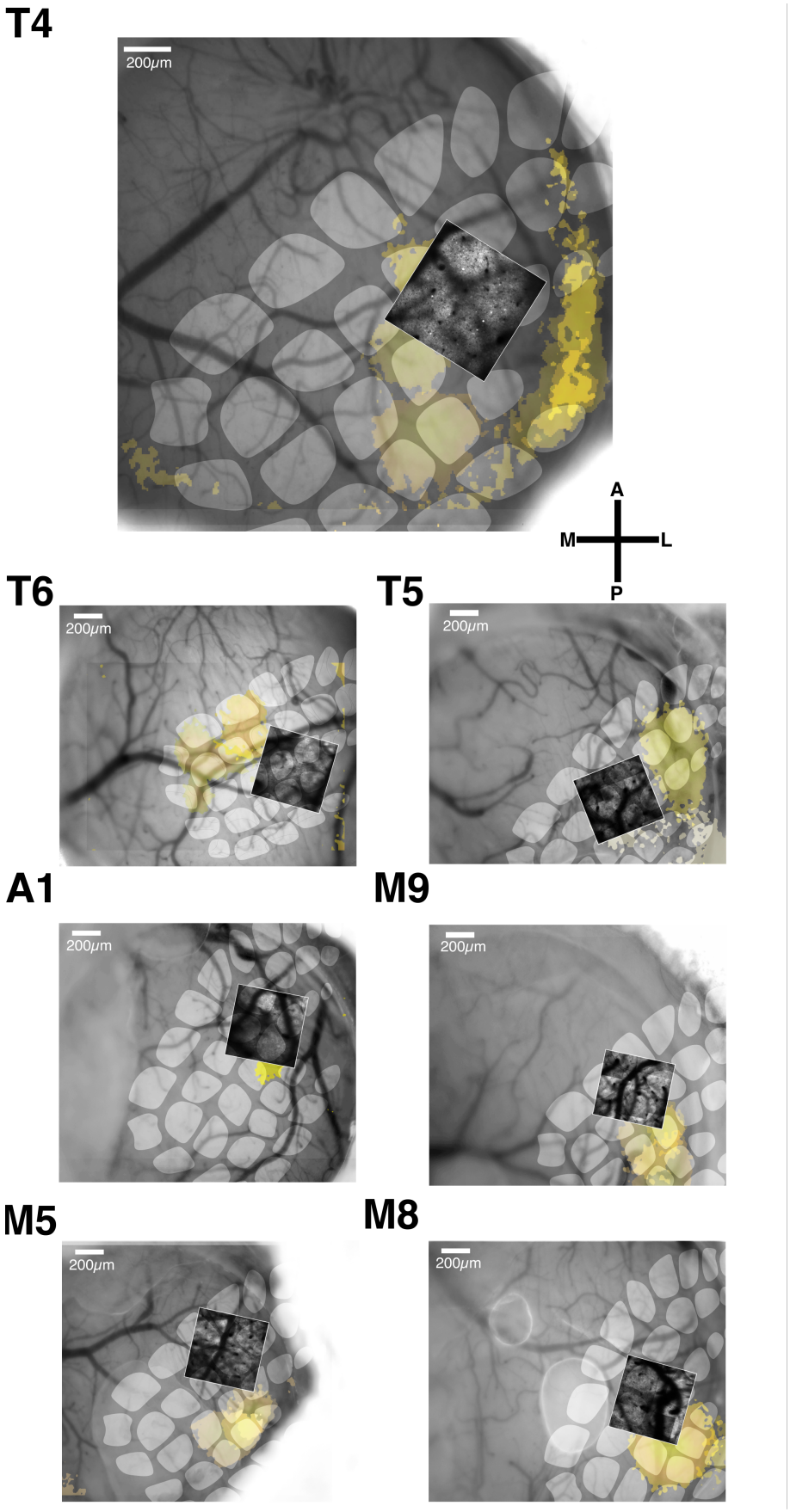

T4
200μm
A
M
L
P
T6
200μm
T5
200μm
A1
200μm
M9
200μm
M5
200μm
M8
200μm

**Figure 3.2:** ***Figure on previous page.*** Terminal intrinsic optical signal imaging (IOSI) for mice trained on the tactile (**T4, T5 & T6**), auditory (**A1**) and multisensory (**M5, M8 & M9**) tasks. For each mouse, whiskers were deflected individually and the corresponding barrels identified (yellow areas: maximal reflectance changes corresponding to increased neural activity). Several barrels were identified in each mouse. These were then mapped to barrels (overlayed in gray). Areas targeted with two-photon imaging are shown with white frames. All panels are presented with the same orientation shown in panel **T4** (A: anterior, P:posterior, M: medial, L: lateral).

### 3.2.7 Support vector machine classification

I implemented a linear support vector machine (SVM) using the sklearn package [Varoquaux et al., 2015] on Python to assess whether the neural responses to stimuli were informative about the frequency of unisensory (tactile or auditory) stimuli, or the tactile, auditory or audio-vibrotactile components of multisensory stimuli. For each mouse, on each day of training, I trained an SVM on trial-by-trial neuronal traces with stimulus identity labels (i.e. the frequency for unisensory stimuli and audio-vibrotactile combination for multisensory stimuli). The classifier was trained on 80% of trials and tested on the remaining 20% on the given day, with a five-fold cross-validation.

I ran the classifier separately on different subpopulations of neurons imaged on a given day – namely all neurons, excited neurons (as described in Section **3.2.4**), suppressed neurons and non-responsive neurons. I also trained and tested the classifier separately on neurons imaged in the shallow plane of multiplexed imaging, corresponding to L2, and neurons imaged in the deep plane, corresponding to L3. In mice trained on the multisensory task, I implemented the SVM independently to classify either the auditory or vibrotactile components of stimulus pairs, as well as the audio-vibrotactile identity of stimuli. The probability of correctly identifying stimuli

by chance varied between unisensory and multisensory across sessions, depending on the number of stimuli (two unisensory and four multisensory), as well as the proportions of stimuli presented in a given session. The classifier performance on stimulus identity categorisation was therefore defined as the difference between the proportion of correctly categorised trials and the proportion of the most frequently presented trial time (i.e. the maximum performance if the classifier simply predicted the same stimulus on every trial), defined as % correct – chance.

## 3.3 Results

### 3.3.1 Neural responses during passive stimulation in naïve mice

Prior to assessing potential effects of training on responses, I first sought to characterise neural responses in naïve animals exposed to auditory, vibrotactile and multisensory stimuli in passive conditions. In passive imaging sessions, mice were head-fixed on top of a wheel in the imaging chamber and the activity of neurons in vS1 was measured with 2-photon calcium imaging while stimuli were presented. Animals were not water-regulated during this imaging session, and they did not receive any reward upon stimulus presentation. I imaged eight mice in total, and each mouse was presented with the two auditory, two tactile and four multisensory stimuli used in later training (described in Sections **2.2.8**, **2.2.7** and **2.2.9**). Responses from individual neurons were analysed offline.

I first assessed the proportion of neurons responsive to auditory, vibrotactile and multisensory stimuli. As defined in Section 2.3.2, the responsiveness of each neuron for each stimulation type (auditory, vibrotactile, or audio-vibrotactile) was

assessed by comparing the maximum fluorescence across two 3-second-long time windows: one positioned immediately before the stimulus (pre-stimulus period), and the second starting from stimulus onset (stimulus period). Pre/post stimulus onset comparisons were done by using a *t*-test (alpha = 0.05) to determine if the response of a given neuron differed across these two time windows, for each stimulus presented. Neurons were considered 'responsive' to auditory stimulation if they responded significantly to at least one of the two sounds presented (paired *t*-test , $p < 0.05$), and responsiveness was similarly determined across the two vibrotactile and four multisensory stimuli. Sample traces from four neurons are shown in Figure **3.3**.

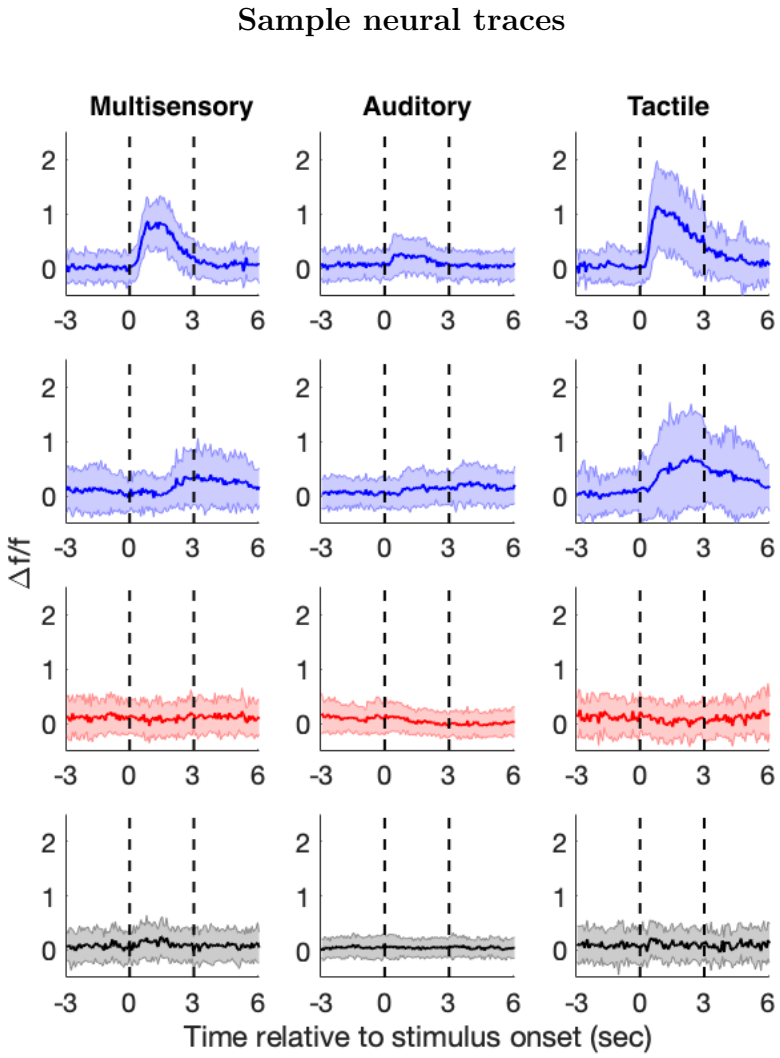



**Figure 3.3:** Average fluorescence traces from four representative neurons (rows) to multisensory (left column), auditory (middle column) and tactile (right column) stimuli presented in a single imaging session. Two tactile stimuli (93 and 129 Hz, delivered with a 250$\mu$m amplitude), two auditory stimuli (6.3 and 12.5 kHz, at a loudness of 70 dB SPL-C) or combinations of the four were used. Blue traces show neurons significantly excited by at least one of the stimuli presented, red traces show neurons significantly suppressed by at least one stimulus, black traces show non-responsive neurons (paired $t$-test , $p$ <0.05; thick line = mean response ± std across trials). Dotted lines: stimulus window.

I performed a 1-way ANOVA to determine if the proportion of responsive neurons differed across the three stimulus modalities: auditory, vibrotactile and audio-vibrotactile (multisensory). As expected in vS1, I found a significant effect

of sensory modality ($F_{(2,21)}$=21.54, $p = 8.18\times10^{-6}$), in which a larger percentage of neurons were responsive to the tactile (Fisher's least significant difference; $p = 1.37\times10^{-5}$) and multisensory ($p = 1.08\times10^{-5}$) stimuli than the auditory stimuli (Figure **3.4** A).

Next, I examined how stimulus-specific the responsive neurons were to the different frequencies of stimulation presented. For each modality, I calculated the percentage of responsive neurons that responded to only one of the presented stimulus values (Figure **3.4** B) or to more than one stimulus values (Figure **3.4 C**). I performed a 1-way ANOVA on the proportion of responsive neurons that responded to only one stimulus value with the three stimulus modalities as predictors, and found that neurons were most selective for the auditory stimuli ($F_{(2,21)}$=48.56, $p = 1.33\times10^{-8}$, post-hoc pairwise comparisons auditory vs. tactile, $p = 1.68\times10^{-6}$; auditory vs. multisensory, $p = 3.61\times10^{-9}$).

Finally, I examined if neurons that responded to only one stimulus frequency were more responsive to low or high frequency stimulation in the auditory and tactile modalities. If age-related hearing loss was present in these mice for these frequency ranges, we might expect more auditory-responsive neurons to respond preferentially to the low frequency stimulus. The proportion of responsive neurons that were significantly responsive to low and high frequency stimuli are plotted in Figure **3.4**. There were no significant differences between the percentage of neurons responding specifically to the low frequency stimulus for auditory and tactile stimulation ($t$-test , t = 0.19, $p = 0.85$). The proportions of responsive neurons plotted in Figure **3.4** are also given in Table **3.2**.

**Responsiveness and selectivity of neurons in pre-training passive conditions**

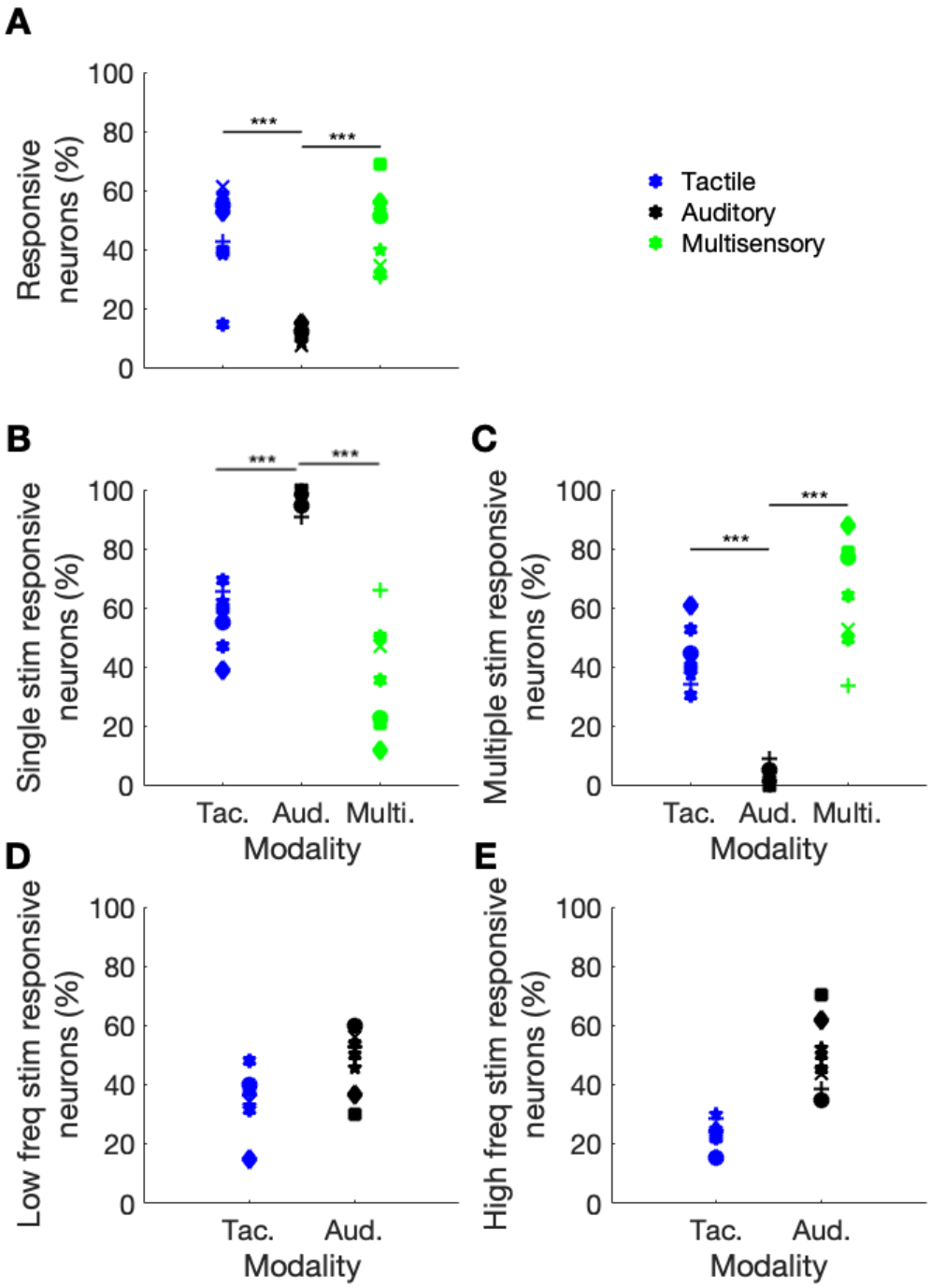


**Figure 3.4:** ***Figure on previous page.*** **A.** Proportion of neurons responsive to tactile, auditory, and multisensory stimuli ($N = 8$ mice). Neurons were considered responsive if they were responsive to at least one stimulus frequency (or a single combination) (paired $t$-test , $p < 0.05$). **B.** Proportion of neurons found to be responsive only to a single auditory or tactile frequency (or only two pairs containing the same tactile component in multisensory conditions). **C.** Proportion of neurons responsive to both frequencies (or to at least 3 combinations in multisensory conditions). **D.** Of all neurons that responded to only one stimulus value (as in **B**) the proportion responsive to only the low frequency stimulus is plotted. **E.** Proportion of selective neurons responsive to only the higher frequency stimulus (defined similarly to **D**. In all panels, asterisks represent the results of pairwise comparisons, resulting from post-hoc LSD tests (**A**-**C**) or $t$-test (**D**-**E**). *** = $p < 0.001$.

| | Proportion of neurons (% ± std) | | |
|---|---|---|---|
| | Tactile | Auditory | Multisensory |
| All responsive | 45.49±15.11 | 11.98±2.89 | 46.09±13.89 |
| | Proportion of responsive neurons (% ± std) | | |
| Single stimulus responsive | 57.13±9.98 | 97.52±3.23 | 38.14±18.56 |
| Multiple stimuli responsive | 42.87±9.98 | 2.48±3.23 | 61.86±18.56 |
| Low frequency responsive | 34.49±9.41 | 48.01±10.25 | Not applicable |
| High frequency responsive | 22.64±5.31 | 49.51±11.89 | Not applicable |

**Table 3.2:** The mean and standard deviations ($N = 8$ mice) of values plotted in Figure 3.4 are provided. The first row shows the proportion of neurons (% of all imaged neurons ± std) responsive to at least one auditory, tactile or multisensory stimulus ('All responsive'). The next row of the table shows the proportion of responsive neurons that responded to only 1 auditory or tactile frequency (or multisensory stimuli with the same tactile component) ('Single stimulus responsive'). The third row shows the % of responsive neurons that responded to more than one presented frequency ('Multiple stimuli responsive'). The fourth and fifth rows give the % of selectively responsive neurons that responded to only the low or high frequency stimulus, respectively.

Hence, while there were significantly fewer auditory-responsive neurons in vS1 of naïve mice imaged in passive conditions, these neurons were significantly more specific in their frequency responses and usually driven by a single frequency only.

### 3.3.2 Neural responses during passive stimulation in trained mice

After measuring the responsiveness of vS1 neurons in untrained mice, I then sought to address how these responses may change after training. I compared how responses to stimuli changed in the same passive stimulation condition after mice had trained on either unisensory or multisensory versions of the discrimination task (Section **2.2.5**).

Two mice trained on the tactile task, one mouse trained on the auditory task and three mice trained on the multisensory task were imaged both before after training during passive exposure to the 2 auditory, 2 tactile and 4 multisensory stimuli. The third mouse trained on the tactile task was presented with tactile stimuli only during the final passive imaging session.

I first aimed to compare how the responsiveness of neurons changed in these mice after training, using the measures defined in Section **3.3.1**. I performed a one-way ANOVA to determine if the proportion of responsive neurons differed across the 3 stimulus modalities: auditory, vibrotactile and audio-vibrotactile. I found a significant effect when comparing the means for each sensory modality ($F_{(2,15)}$=11.95, $p = 7.87\times10^{-4}$). Post-hoc pairwise comparisons revealed that there were significantly more neurons responsive to tactile and multisensory stimuli than auditory stimuli (Fisher's least significant difference; $p = 0.0002$ and $p = 0.0047$ respectively) (Figure **3.5** A).

I then assessed how specific the responses were to stimulus frequencies. I performed a one-way ANOVA on the proportion of responsive neurons that responded to only one stimulus value with the three stimulus modalities as predictors and found

highly significant proportions of specifically responsive neurons for the group means of all modalities ($F_{(2,15)}$=25.80, $p = 1.40\times10^{-5}$). There were more specifically responsive neurons to auditory stimuli than both tactile and multisensory ($p = 0.0035$ and $p = 3.17\times10^{-6}$ respectively) (Figure **3.5** B-C).

Finally, I examined if neurons that responded to only one stimulus frequency were more responsive to low or high frequency stimulation in the auditory and tactile modalities. If age-related hearing loss was present in these mice for these frequency ranges, we might expect more auditory-responsive neurons to respond preferentially to the low frequency stimulus. The proportion of responsive neurons that were significantly responsive to low and high frequency stimuli are plotted in Figure **3.5 D-E**. There were no significant differences between the percentage of neurons responding specifically to the low frequency stimulus for auditory and tactile stimulation ($t$-test , $t$ =-1.1, $p = 0.29$). Detailed proportions of different responsive neurons can be found in Table **3.3**.

**Responsiveness and selectivity of neurons in post-training passive conditions**

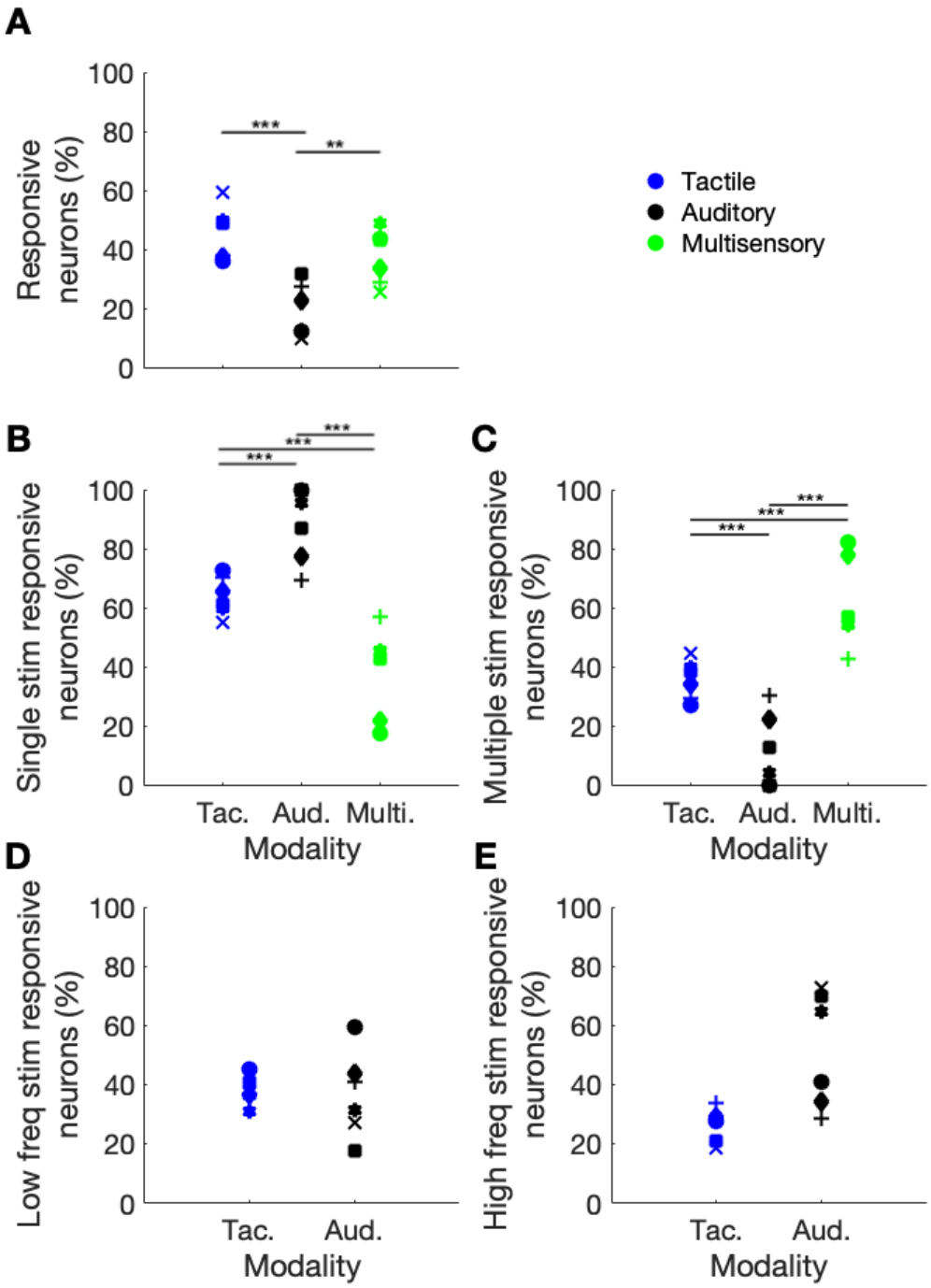


**Figure 3.5:** ***Figure on previous page.*** **A.** Proportion of neurons responsive to tactile, auditory and multisensory stimuli ($N = 6$ mice) respectively after training (a final imaging session for mice that reached criterion, removed from water regulation and passively exposed to stimuli). Neurons were considered responsive if they were responsive to at least one stimulus type (comparison of maximum fluorescence values in a 3 second window before stimulus onset and during stimuli, paired $t$-test , $p < 0.05$). Such examples can be found in Figures **3.3**. **B.** Proportion of neurons found to be responsive only to a single auditory or tactile frequency or a single audio-vibrotactile pair. Responsiveness measures were defined as in **A** and neurons were considered responsive only if found responsive to a single stimulus. **C.** Proportion of neurons responsive to both frequencies (or to at least 2 combinations in multisensory conditions). **D.** Proportion of neurons (out of all responsive neurons) responsive to low frequency stimuli. **E.** Proportion of neurons responsive to higher frequency stimuli, defined similarly to **D**. In all panels, asterisks represent the results of pairwise comparisons, resulting from post-hoc LSD tests (**A**-**C**) or $t$-test (**D**-**E**). *** = $p < 0.001$, ** = $p < 0.01$

| | Proportion of neurons (% ± std) | | |
|---|---|---|---|
| | Tactile | Auditory | Multisensory |
| All responsive | 45.15±9.28 | 19.68±9.19 | 37.40±9.28 |
| | Proportion of responsive neurons (% ± std) | | |
| Single stimulus responsive | 64.26±6.51 | 88.36±12.69 | 38.27±15.31 |
| Multiple stimuli responsive | 35.74±6.51 | 11.64±12.69 | 61.73±15.31 |
| Low frequency responsive | 37.88±4.75 | 36.61±14.60 | Not applicable |
| High frequency responsive | 26.37±5.69 | 51.74±19.54 | Not applicable |

**Table 3.3:** The first part of the table highlights the proportion of neurons (% of all imaged neurons ± std) responsive to any or all auditory, tactile or multisensory stimuli (All responsive) $N = 8$ mice. The second part of the table shows the proportion of neurons (% of all responsive neurons) responsive to only the low frequency stimuli (or multisensory stimuli with the low frequency tactile component) (Low frequency responsive), responsive to only the high frequency stimuli (High frequency responsive), responsive to a single auditory or tactile frequency (Single stimulus responsive) and finally responsive to both frequencies or at least 2 multisensory combinations (Multiple stimuli responsive). $N = 8$ mice.

### 3.3.3 Changes to passive responses pre- and post-training

I sought to assess differences in neural responses to sound between pre-training and post-training passive stimulation sessions. I used a repeated-measure ANOVA

to determine potential differences in the proportion of neurons that were responsive to auditory, vibrotactile and audio-vibrotactile stimuli, with the stage of training (pre- or post-training) set as a predictor variable. There was no effect of training stage ($F_{(2,20)}$=3.08, $p$ = 0.121) or significant interaction between the training stage and stimulus modality ($F_{(2,20)}$=2.92, $p$ = 0.078) on the proportion of responsive neurons (Figure **3.6 A**). Specifically, post-hoc pairwise comparisons did not highlight any significant differences between the proportion of responsive neurons in pre- and post-training passive sessions in tactile ($p$ = 0.23), auditory ($p$ = 0.088) or multisensory sessions ($p$ = 0.11). Similarly, I found no significant effect of training on the proportions of stimulus selectivity ($F_{(2,20)}$=1.67, $p$ = 0.214; Figure **3.6 B-C**) or on the proportion of neurons preferring low or high frequencies ($F_{(2,20)}$=2.77, $p$ = 0.087; Figure **3.6** D-E).

**Comparing the responsiveness and selectivity of neurons in pre- and post-training passive conditions**

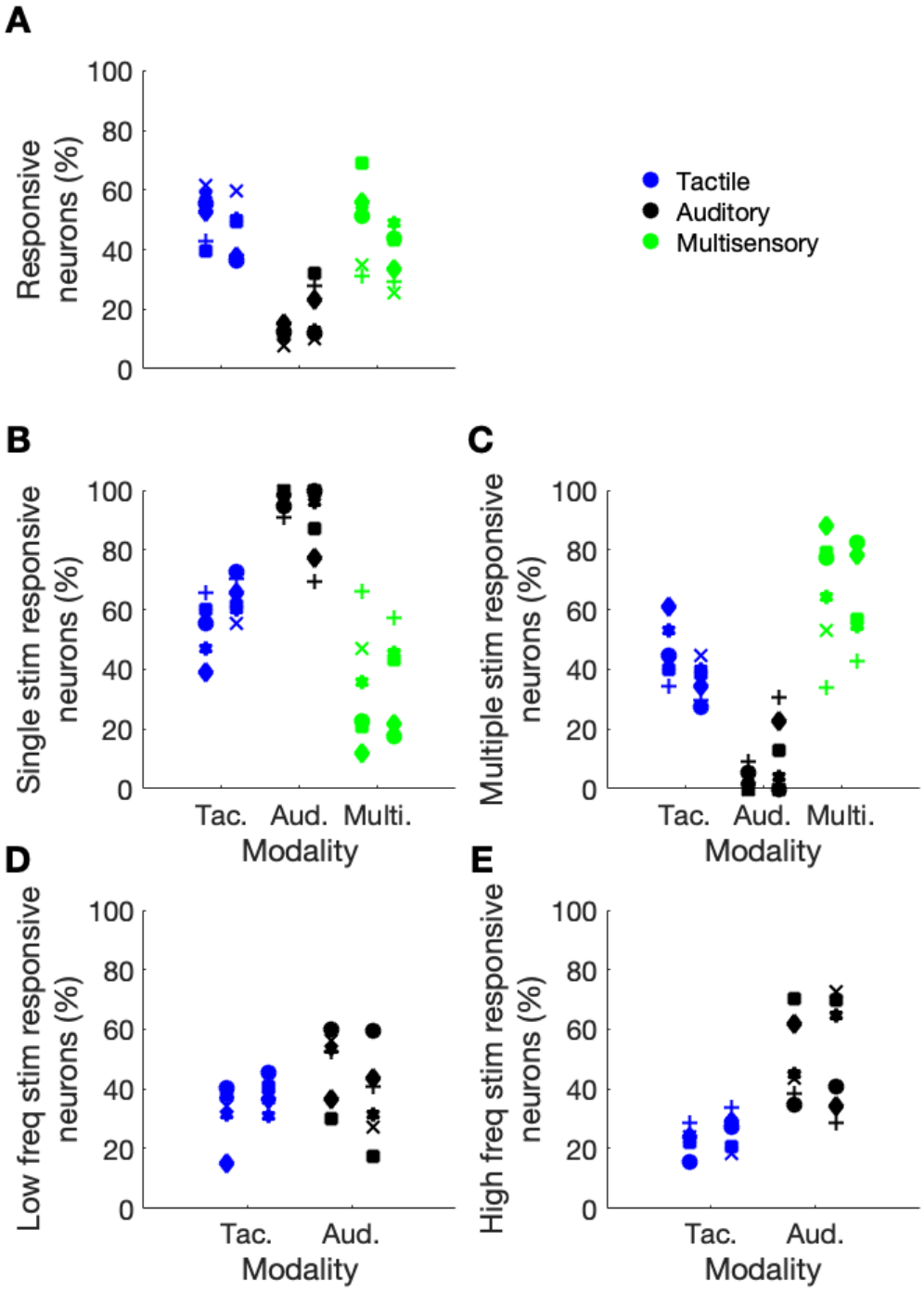


**Figure 3.6:** ***Figure on previous page.*** **A.** Proportion of neurons responsive to tactile, auditory and multisensory stimuli respectively prior to training (left) and after reaching the training criterion (right) ($N = 6$ mice). Neurons were considered responsive if they were responsive to at least one stimulus type (comparison of maximum fluorescence values in a 3 second window before stimulus onset and during stimuli, paired $t$-test , $p < 0.05$). Such examples can be found in Figure **3.3**. **B.** Proportion of neurons found to be responsive only to a single auditory or tactile frequency (or only two pairs containing the same tactile component in multisensory conditions). Responsiveness measures were defined as in **A** and neurons were considered responsive only if found responsive to a single frequency or combination. **C.** Proportion of neurons responsive to both frequencies (or to at least 2 combinations in multisensory conditions). **D.** Proportion of neurons (out of all responsive neurons) responsive to low frequency stimuli. **E.** Proportion of neurons responsive to higher frequency stimuli, defined similarly to **D**.

To summarise, the effects of training on responses to passive stimulation were modest. I did not observe significant training effects on the proportions of responsive neurons or even in the relative proportions of stimulus-selective neurons in unisensory and multisensory conditions. However, I found subtle differences in the patterns of stimuli that elicited responses in trained mice that I had not observed in pre-training conditions. Most notably, I found that a majority of neurons responded specifically to one frequency of vibrotactile stimulation after training.

### 3.3.4 Changes in neural selectivity for stimuli throughout multisensory training

The effects of multisensory learning on neural responses to stimuli may be more apparent if responses are measured while the animal performs the behavioural task. Therefore, I next assessed whether the sensitivity of neurons to a particular unisensory component of multisenory stimuli (i.e. auditory or tactile frequency) changed over the course of training in the three mice trained on the multisensory task. These analyses include the passive stimulation sessions at the start and end

of training (described in the above sections), as well as imaging performed during an habituation session, and while the animal performed behavioural sessions on the multisensory task at each of three different learning stages: early training (<60% correct); mid-training (60-70% correct)); and an 'expert' stage (>70% correct).

For every neuron, I ran a two-way ANOVA on the responses on each trial, with the auditory and tactile stimulus frequencies as predictor variables. If a significant effect of tactile frequency was found, this neuron was identified as a 'tactile-sensitive neuron'. In other words, such a neuron responded differently to multisensory trials when the vibrotactile frequency differed, but was not modulated by the sound frequency. Similarly, if a significant effect of the auditory frequency on a neuron's response was found, this neuron was considered to be 'auditory-sensitive'. Finally, if a significant interaction was found between the auditory and tactile frequencies, the neuron was considered to be a 'multisensory-sensitive' neuron. The proportions of neurons sensitive to each component of the multisensory stimulus throughout training are summarised in Table **3.4** and Figure **3.7**.

| **Proportion of unisensory- and multisensory-sensitive neurons across the course of training (%)** | | | | | |
|---|---|---|---|---|---|
| Tactile | | | | | |
| Pre training | Habituation | <60% correct | 60-70% correct | >70% correct | Post training |
| 6.35% ±3.19 | 2.11% ±0.32 | 1.14% ±0.10 | 1.00% ±0.72 | 4.00% ±0.54 | 4.38% ±1.29 |
| Auditory | | | | | |
| Pre training | Habituation | <60% correct | 60-70% correct | >70% correct | Post training |
| 2.74% ±2.01 | 1.85% ±0.94 | 2.06% ±1.22 | 4.51% ±2.40 | 8.46% ±9.71 | 4.73% ±0.63 |
| Multisensory | | | | | |
| Pre training | Habituation | <60% correct | 60-70% correct | >70% correct | Post training |
| 3.04% ±2.19 | 1.10% ±0.98 | 1.13% ±1.13 | 1.03% ±0.66 | 4.50% ±5.26 | 5.27% ±0.99 |

**Table 3.4:** Mean +/- standard deviation of the proportion of imaged neurons sensitive to the tactile, auditory or multisensory identity of stimuli across multisensory training periods ($N = 3$ mice).

The proportion of neurons sensitive to the tactile, auditory or multisensory components of the stimuli varied significantly across the course of training ($F_{(5,30)}$=3.12, $p = 0.022$, two-way ANOVA). There was no significant difference in the proportion of sensitive neurons between the pre-training and post-training passive sessions ($p = 0.39$, Fischer's least significant difference). However, the proportion of sensitive neurons on the final passive session significantly differed from all 3 stages of active training (<60% correct: $p = 0.0014$; 60-70% correct: $p = 0.00076$; >70% correct: $p = 0.003$). In contrast, there was no difference in the proportion of sensitive neurons between the pre-training passive session and any of the 3 active training stages (p>0.05).

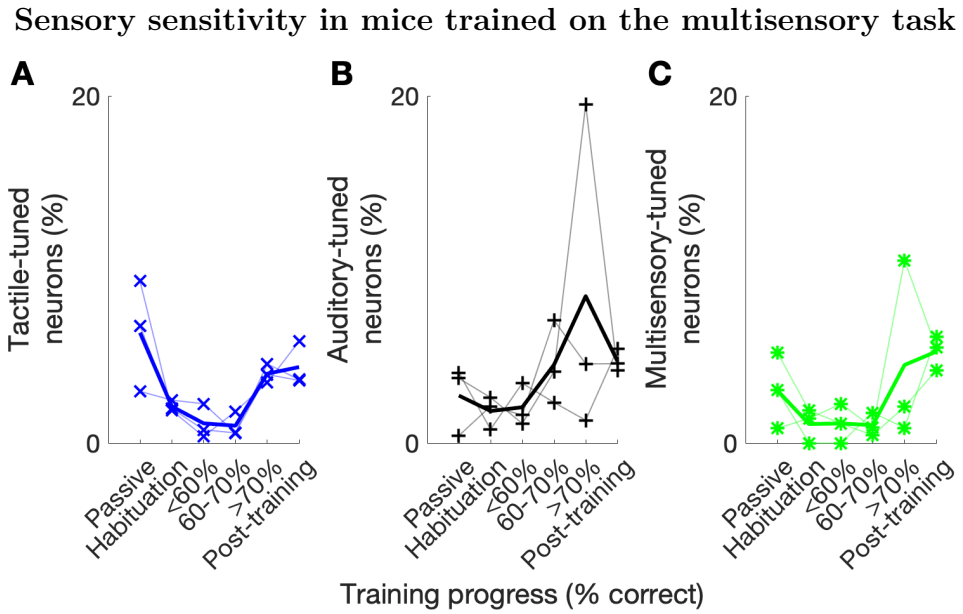



**Figure 3.7: A.** Proportion of neurons sensitive to the tactile identity of multisensory stimuli ($N = 3$ mice). **B.** Proportion of neurons sensitive to the auditory identity of multisensory stimuli ($N = 3$ mice). **C.** Proportion of neurons sensitive to the multisensory identity of multisensory stimuli ($N = 3$ mice).

Hence, while the proportion of tuned neurons remained low across the course of training, a significant post-training increase in the proportion of tuned neurons was observed across all three modalities. The fact that there were no significant differences in the proportion of neurons tuned to the auditory, tactile or multisensory identities of stimuli suggests that training increased neural tuning across all modalities non-specifically. In particular, the increase in auditory-tuned and multisensory-tuned neurons after training suggests the existence of multisensory responses which can be enhanced by training on a multisensory task.

### 3.3.5 Response profile changes throughout training

The above analyses only point out whether neurons respond significantly to stimuli throughout the course of training, but these may miss more subtle changes

in spiking responses. For example, a neuron may change the manner or timing of its response to a stimulus throughout training. I therefore examine potential changes in the response profiles of neurons in the following section.

In the above analyses (sections **3.3.1** onward), every neuron's response to a stimulus presentation was determined by testing if the maximum $\Delta f/f$ value during a 3 s window during stimulus presentation was different compared to a 3 s pre-stimulus window ($t$-test ). A neuron was considered responsive if it passed the $t$-test for at least one of the presented stimulus types (two auditory and tactile stimuli, four multisensory combinations).

In order to compare how responses may vary across cells or the course of training, I further defined the 'responsive' neurons ($t$-test , $p$ <0.05) as exhibiting one of four possible response types. Neurons whose responses reached a maximum $\Delta f/f$ value in the first second of stimulus presentation (i.e. before reaching full stimulus amplitude) were categorised as exhibiting 'sharp onset' responses. Neurons with a peak response during the next second of stimulus presentation, wherein the stimulus reaches its maximum amplitude, were characterised as 'amplitude-locked'. Neurons with peak responses in the 1 s window after the period of maximum stimulus amplitude (i.e. starting 2 s after stimulus onset) were labelled 'delayed onset' responsive neurons. Finally, neurons with a lower peak $\Delta f/f$ value during the 3 s post-stimulus-onset compared to the 3 s pre-stimulus window were labelled 'suppressed' neurons. Average fluorescence traces across all neurons in each of the five response categories (imaged in one behavioural session) are shown in Figure **3.8**.

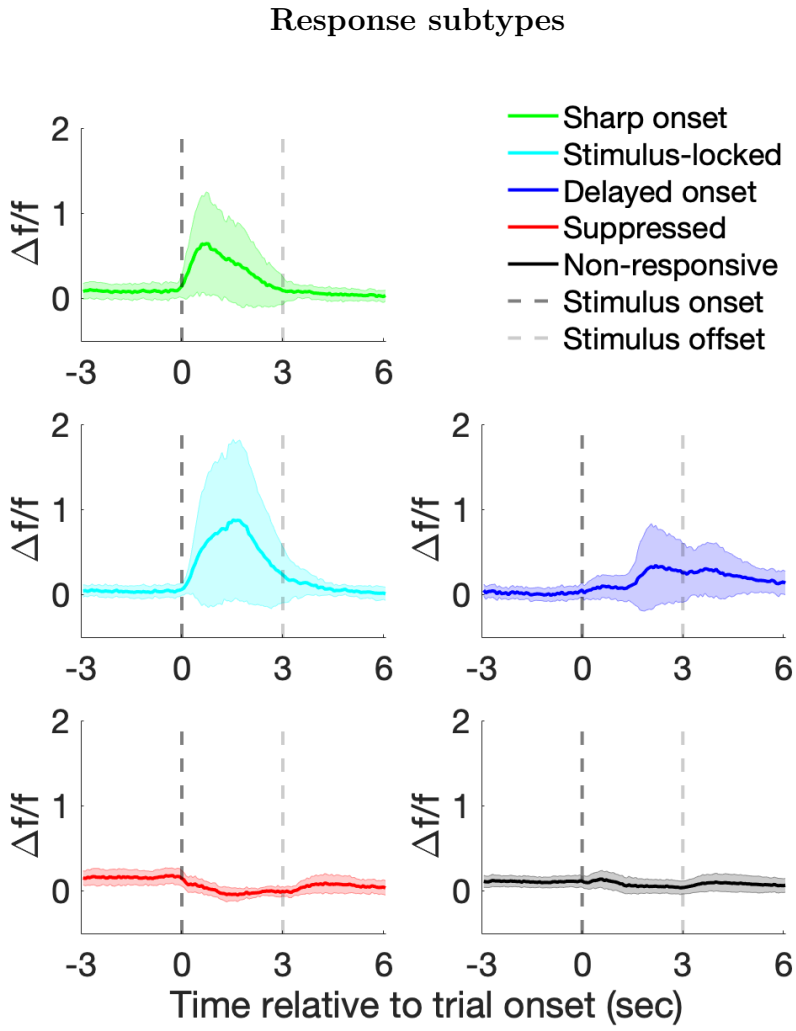



**Figure 3.8:** Response to the high frequency vibrotactile stimulus one imaging field (mouse **T4**), averaged (+/- standard deviation) across all neurons belonging to each of the 5 response types: 'Sharp onset' (green), 'Stimulus-locked' (cyan), 'Delayed onset' (blue), 'Suppressed' (red), and 'Non-responsive' (black) neurons. The vertical dark grey dashed line indicates stimulus onset, and the light grey line indicates stimulus offset.

Unisensory training might be sufficient to trigger changes in these response types in vS1. For example, we might expect to see more sharp onset responses to tactile stimuli following tactile discrimination training. I assessed whether the proportion of different response types varied over the course of training in mice trained on the tactile task ($N = 3$ mice). I used a repeated-measures 2-way

ANOVA with response profile (i.e. the 5 response types) and tactile stimulus frequency ('Go' or 'No-Go') as predictor variables and the proportion of neurons with different response profiles in each of the 6 training stages as the predicted variable. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated ($\chi^2(14) = 35.86$, $p = 0.001$) and I therefore applied a Greenhouse-Geisser correction. I observed no significant difference between proportion of responses to 'Go' and 'No-Go' frequencies ($F_{(3.12,62.42)}$=0.002, $p = 0.999$). However, I found a significant interaction between response profile and training stage ($F_{(12.48,62.42)}$=5.50, $p = 2.11\times10^{-6}$). More specifically, post-hoc pairwise comparisons revealed changes over time in the proportions of differently responsive neurons, which were consistent across both tactile frequencies (Figure **3.9**). For instance, there were significantly more amplitude-locked neurons at later stages of training (>70% correct) than early sessions (<60% correct)($p = 0.042$). In addition, there were significantly more delayed onset and amplitude-locked neurons at the end of training than in the initial habituation sessions ($p = 0.006$ and $p = 0.0175$, respectively). While there were no significant variations in the proportions of sharp onset and suppressed neurons across training stages, there were significantly fewer non-responsive neurons on the last day of training (>70% correct) than on the pre-training passive, habituation and early training sessions ($p = 9.50\times10^{-4}$, $p = 8.45\times10^{-4}$ and $p = 0.011$, respectively). There were significantly more non-responsive neurons on post-training passive sessions than the later stages of training (60-70%: $p = 0.026$ and >70%: $p = 3.25\times10^{-5}$). Across the different types of responsive neurons, only the proportion of delayed onset neurons significantly decreased between the last day of training and the post-training passive session ($p = 8.48\times10^{-4}$). Detailed proportions can be found in Table **3.5**.

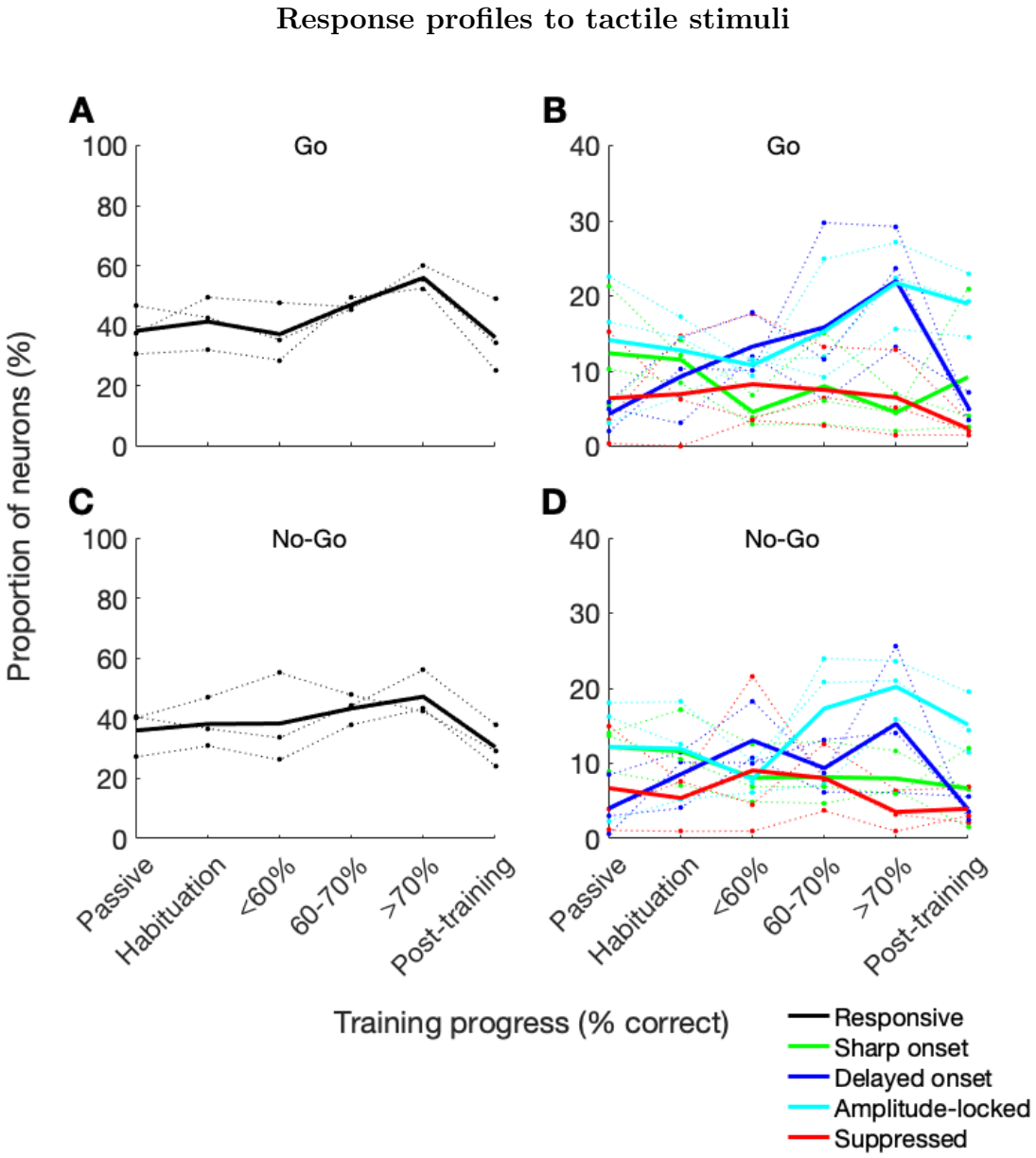



**Figure 3.9:** Proportions of responsive neurons (black) and sub-categories of response profiles: sharp onset (green; peak response before reaching peak stimulus amplitude); amplitude-locked (cyan; peak response during stimulus peak amplitude); delayed onset (blue; peak response after the end of peak stimulus amplitude); and suppressed (red; $\Delta f/f$ value significantly lower during stimulus presentation than pre-stimulus period; $p$ <0.05, left-tailed $t$-test ). Bold lines are averaged across animals ($N = 3$ mice). Thin lines and dots are proportions from individual mice. Plots are shown for 'Go' (**A** & **B**) and 'No-Go' (**C** & **D**) tactile frequencies.

| **Proportion of responses to tactile stimuli (% of imaged neurons)** | | | | | | |
|---|---|---|---|---|---|---|
| **'Go' frequency** | | | | | | |
| | Pre training | Habituation | <60% correct | 60-70% correct | >70% correct | Post training |
| Sharp onset | 12.37 ±8.11 | 11.51 ±2.92 | 4.55 ±2.03 | 7.99 ±6.35 | 4.49 ±2.48 | 9.20 ±10.24 |
| Stimulus locked | 14.12 ±9.94 | 12.75 ±5.46 | 10.78 ±1.24 | 15.31 ±8.40 | 21.72 ±5.77 | 18.95 ±4.23 |
| Delayed onset | 4.28 ±2.07 | 9.27 ±5.77 | 13.25 ±3.99 | 15.82 ±12.34 | 21.99 ±8.18 | 5.21 ±1.87 |
| Suppressed | 6.38 ±7.85 | 6.93 ±7.35 | 8.24 ±8.12 | 7.49 ±5.29 | 6.49 ±5.79 | 2.36 ±1.02 |
| Non responsive | 61.66 ±8.04 | 58.57 ±8.82 | 62.74 ±9.69 | 52.88 ±1.98 | 44.00 ±3.84 | 63.83 ±11.96 |
| **'No-Go' frequency** | | | | | | |
| | Pre training | Habituation | <60% correct | 60-70% correct | >70% correct | Post training |
| Sharp onset | 12.18 ±2.87 | 11.63 ±5.10 | 8.10 ±3.94 | 8.18 ±4.27 | 7.99 ±3.19 | 6.67 ±5.25 |
| Stimulus locked | 12.21 ±8.58 | 11.95 ±6.52 | 7.91 ±2.10 | 17.31 ±8.89 | 20.19 ±3.92 | 15.15 ±4.03 |
| Delayed onset | 4.01 ±4.04 | 8.56 ±3.98 | 13.04 ±4.51 | 9.37 ±3.52 | 15.27 ±9.80 | 3.89 ±1.56 |
| Suppressed | 6.70 ±7.32 | 5.39 ±3.84 | 9.05 ±11.01 | 8.04 ±4.47 | 3.53 ±2.73 | 3.97 ±2.50 |
| Non responsive | 64.06 ±7.59 | 61.90 ±8.20 | 61.75 ±15.06 | 56.74 ±5.01 | 52.81 ±7.78 | 69.58 ±7.00 |

**Table 3.5:** Proportion (mean ± std, $N = 3$ mice) of responsive neurons and response profile subgroups (sharp onset, amplitude-locked, delayed onset, and suppressed) in animals trained on the tactile discrimination task. Proportions are given for responses to 'Go' and 'No-Go' tactile frequencies respectively.

Having established that tactile discrimination training can change the response profiles of neurons in vS1, I next investigated if multisensory training would evoke similar response changes. Here, responses to multisensory stimuli were categorised into the 5 response types, as described above, on responses of

neurons imaged in 3 mice trained on the multsensory categorization task. I used a 2-way repeated-measures ANOVA to investigate how the proportion of different types of responsive neurons varied over the course of training, and across the 4 audio-vibrotactile stimuli. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated ($\chi^2(14) = 142.52$, $p = 2.23\times10^{-23}$) and I therefore applied a Greenhouse-Geisser correction.

I found a significant interaction between response profile and training stage ($F_{(9.49,94.91)}$=8.17, $p = 3.78\times10^{-9}$). There were significantly fewer responsive neurons during post-training passive imaging than in all other sessions (Figure **3.10**; post-hoc pairwise comparisons, pre-training: $p = 2.09\times10^{-8}$; habituation: $p = 2.07\times10^{-8}$; <60%: $p = 3.04\times10^{-7}$; 60-70%: $p = 1.85\times10^{-6}$; >70%: $p = 2.07\times10^{-8}$). As the training stage advanced, the proportion of responsive neurons increased (<60% vs. 60-70%: $p = 0.012$; <60% vs. >70%: $p = 0.0023$). Post-hoc pairwise comparisons also revealed changes over time in the proportions of differently responsive neurons (Figure **3.11**). In contrast with mice trained on the tactile task, pairwise comparisons revealed a significant decrease in the proportion of amplitude-locked and sharp onset neurons between the pre-training and post-training passive sessions ($p = 3.84\times10^{-6}$ and $p = 0.002$ respectively). Similarly, there were significantly fewer delayed onset and amplitude-locked neurons in the post-training passive sessions than in the pre-training habituation sessions ($p = 1.78\times10^{-6}$ and $p = 6.65\times10^{-7}$ respectively). Although not significant overall ($p = 0.051$), a transient increase in the proportion of suppressed neurons was observed between the habituation stage and final training session in one mouse. Finally, I observed a significant decrease in the proportion of sharp onset responsive neurons from last day of training to the post-training passive imaging session ($p = 0.003$). More detailed proportions can be found in Table **3.6**.

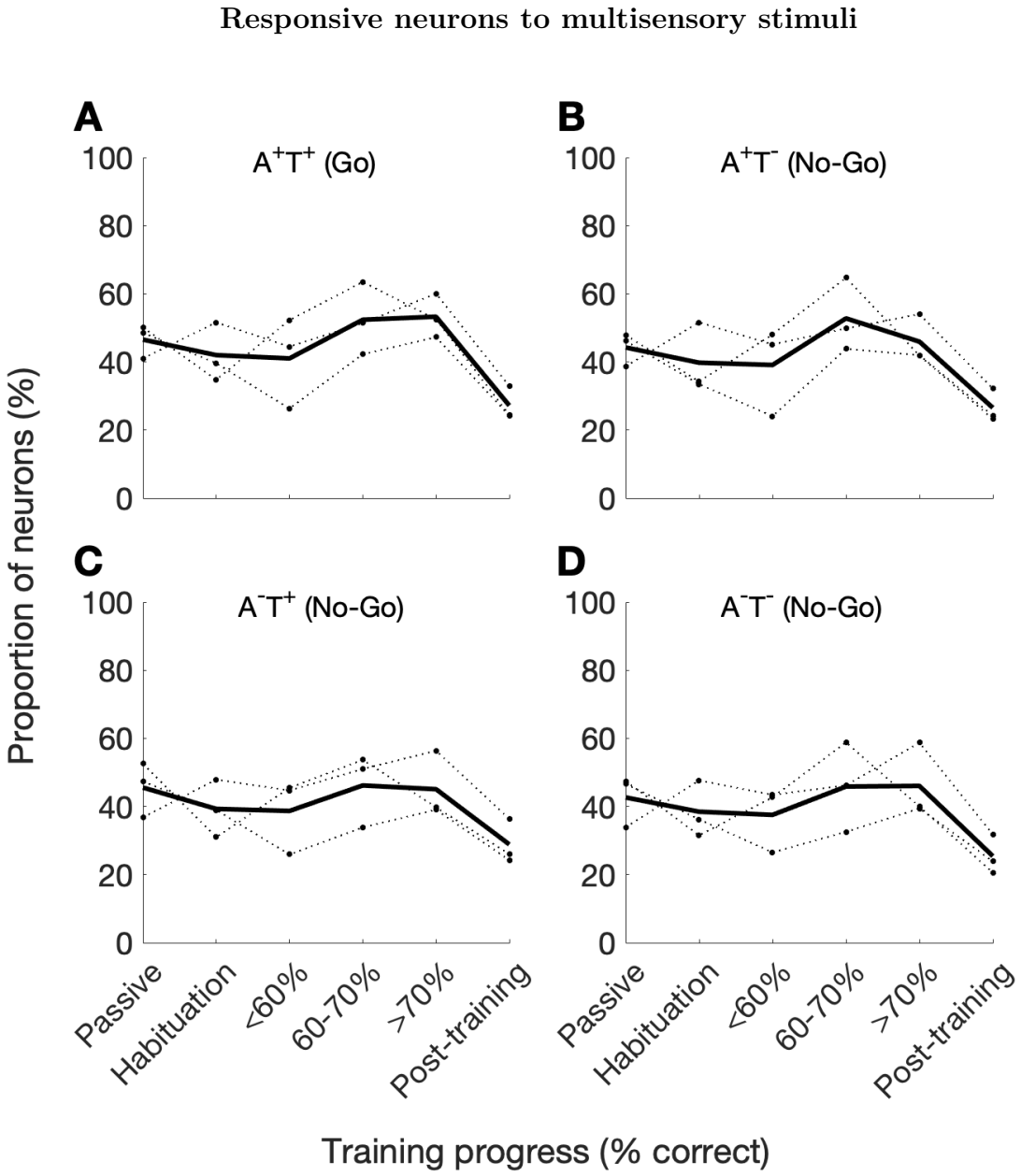



**Figure 3.10:** Proportion of responsive neurons (black) across days of training. Bold lines are averaged across animals ($N = 3$ mice). Thin dashed lines and dots are proportions from individual mice. Plots are shown for 'Go' (**A**) and 'No-Go' (**B, C** & **D**) audio-vibrotactile combinations. More detailed plots of the proportions of different types of responsive neurons can be found in Figure **3.11**.

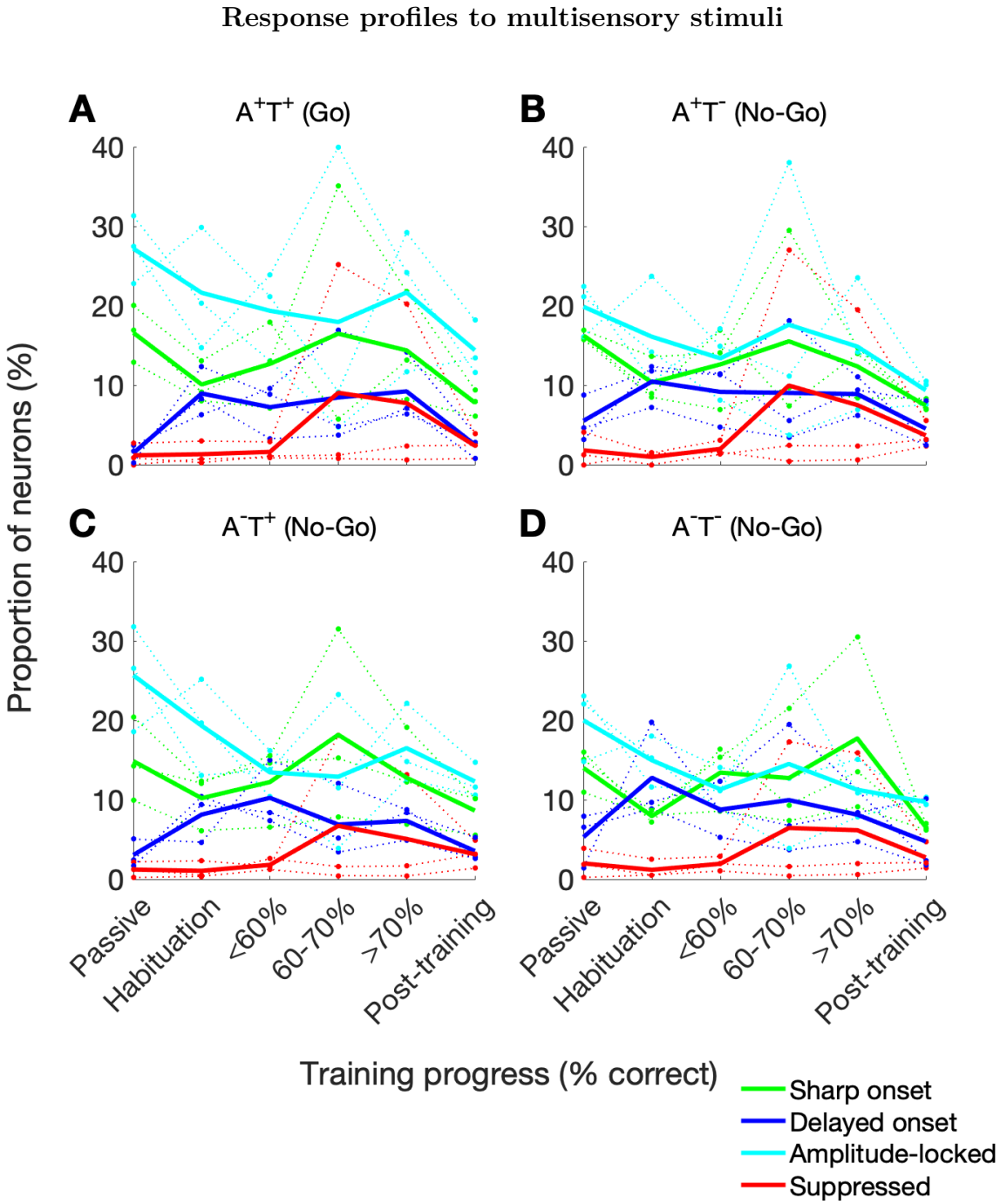



**Figure 3.11:** Details of plots in Figure **3.10**. Proportion of different types of responsive neurons split in separate groups according to their response profile: sharp onset (green), amplitude-locked (cyan), delayed onset (blue), and suppressed (red). Bold lines are averaged across animals ($N = 3$ mice). Thin dashed lines and dots are proportions from individual mice. Plots are shown for 'Go' (**A**) and 'No-Go' (**B, C** & **D**) audio-vibrotactile combinations.

| **Proportion of responses to multisensory stimuli (% of imaged neurons)** | | | | | | |
|---|---|---|---|---|---|---|
| **Go** ($A^+T^+$) | | | | | | |
| | Pre training | Habituation | <60% correct | 60-70% correct | >70% correct | Post training |
| Sharp onset | 16.60 ±3.58 | 10.09 ±2.65 | 12.68 ±5.41 | 16.48 ±16.15 | 14.38 ±6.89 | 7.80 ±1.69 |
| Stimulus locked | 27.20 ±4.26 | 21.64 ±7.67 | 19.37 ±5.62 | 17.95 ±19.20 | 21.68 ±9.02 | 14.40 ±3.39 |
| Delayed onset | 1.44 ±1.13 | 8.96 ±3.10 | 7.23 ±3.46 | 8.45 ±7.33 | 9.22 ±4.30 | 2.50 ±1.55 |
| Suppressed | 1.17 ±1.37 | 1.31 ±1.45 | 1.60 ±1.09 | 9.05 ±13.99 | 7.74 ±10.87 | 2.38 ±1.58 |
| Non responsive | 53.45 ±4.87 | 58.00 ±8.67 | 58.97 ±13.31 | 47.62 ±10.66 | 46.72 ±6.44 | 72.82 ±4.92 |
| **No-Go** ($A^-T^+$) | | | | | | |
| | Pre training | Habituation | <60% correct | 60-70% correct | >70% correct | Post training |
| Sharp onset | 14.87 ±5.27 | 10.20 ±3.52 | 12.24 ±4.96 | 18.19 ±12.08 | 12.78 ±6.11 | 8.64 ±2.66 |
| Stimulus locked | 25.67 ±6.64 | 19.32 ±6.06 | 13.46 ±2.91 | 12.91 ±9.72 | 16.51 ±5.00 | 12.31 ±2.11 |
| Delayed onset | 3.05 ±1.78 | 8.14 ±3.06 | 10.25 ±4.13 | 6.89 ±4.57 | 7.36 ±2.11 | 3.55 ±1.47 |
| Suppressed | 1.22 ±0.98 | 1.06 ±1.10 | 1.81 ±0.70 | 6.72 ±9.88 | 5.08 ±7.01 | 3.15 ±1.76 |
| Non responsive | 54.44 ±8.00 | 60.70 ±8.45 | 61.30 ±11.07 | 53.82 ±10.76 | 54.93 ±9.80 | 71.14 ±6.56 |

| **No-Go** ($A^+T^-$) | | | | | | |
|---|---|---|---|---|---|---|
| | Pre training | Habituation | <60% correct | 60-70% correct | >70% correct | Post training |
| Sharp onset | 16.26 ±0.61 | 10.34 ±2.82 | 12.63 ±5.13 | 15.53 ±12.14 | 12.37 ±3.47 | 7.44 ±0.79 |
| Stimulus locked | 19.88 ±3.33 | 16.10 ±6.86 | 13.38 ±4.65 | 17.61 ±18.06 | 14.86 ±8.30 | 9.36 ±1.60 |
| Delayed onset | 5.51 ±2.93 | 10.46 ±2.83 | 9.16 ±3.86 | 9.03 ±7.93 | 8.88 ±2.46 | 4.53 ±3.06 |
| Suppressed | 1.79 ±2.12 | 0.97 ±0.84 | 1.98 ±0.95 | 9.96 ±14.82 | 7.50 ±10.45 | 3.68 ±1.69 |
| Non responsive | 55.76 ±4.88 | 60.23 ±10.26 | 60.90 ±13.09 | 47.22 ±10.75 | 54.04 ±6.99 | 73.46 ±4.90 |
| **No-Go** ($A^-T^-$) | | | | | | |
| | Pre training | Habituation | <60% correct | 60-70% correct | >70% correct | Post training |
| Sharp onset | 14.01 ±2.72 | 7.95 ±0.74 | 13.43 ±4.24 | 12.74 ±7.67 | 17.73 ±11.31 | 6.57 ±0.39 |
| Stimulus locked | 20.01 ±4.49 | 14.98 ±3.20 | 11.32 ±2.64 | 14.51 ±11.56 | 11.26 ±3.65 | 9.73 ±0.50 |
| Delayed onset | 5.31 ±3.40 | 12.78 ±6.05 | 8.77 ±3.53 | 9.97 ±8.36 | 8.13 ±3.26 | 4.75 ±4.67 |
| Suppressed | 2.00 ±1.83 | 1.19 ±1.17 | 1.95 ±0.91 | 6.46 ±9.43 | 6.17 ±8.43 | 2.75 ±1.77 |
| Non responsive | 57.34 ±7.60 | 61.53 ±8.29 | 62.46 ±9.54 | 54.14 ±13.14 | 53.95 ±10.98 | 74.57 ±5.77 |

**Table 3.6:** ***Table on previous and current pages.*** Proportion (mean ± std, $N = 3$ mice) of non-responsive neurons and responsive neurons split in separate groups according to average response profile: sharp onset, amplitude-locked, delayed onset, and suppressed. Proportions are given for responses to 'Go' stimuli ($A^+T^+$), Auditory-ambiguous 'No-Go' stimuli ($A^+T^-$), Tactile-ambiguous 'No-Go' stimuli ($A^-T^+$) and unambiguous, coherent 'No-Go' stimuli ($A^-T^-$).

To summarise, I observed transient changes in the proportions of neurons exhibiting different response profiles across training stages. I found an increase in the proportion of amplitude-locked and delayed onset neurons in mice trained on the tactile task, which decreased back to pre-training levels in the passive imaging session after training. In contrast, in mice trained on the multisensory task, the proportion of sharp onset, amplitude-locked and delayed onset neurons decreased over the course of training. More strikingly, the proportions of sharp onset and amplitude-locked neurons significantly decreased between pre- and post-training passive sessions, suggesting long lasting changes in the responses of vS1 neurons due to training on the multisensory task. In both trained cohorts, the proportion of responsive neuron significantly increased during training. Finally, a key difference between mice trained on the multisensory and vibrotactile tasks was that there were significantly fewer responsive neurons after the end of training (i.e. when mice were not engaging in the task) in animals trained on the multisensory task, which was not observed in mice trained on the vibrotactile task. This suggests that training on a multisensory task may recruit different subsets of neurons than on a unisensory task in vS1, warranting a more in-depth examination of population activity.

### 3.3.6 Stimulus decoding from vS1 responses

Having observed and characterised changes in the responses of neurons between pre-training and post-training imaging sessions as well changes in response profiles during training, I next aimed to uncover if these changes may improve the discrimination of the trained stimuli in vS1. To answer this question, I implemented a linear SVM classifier in order to assess how well the identity of stimuli could be determined from neural population responses throughout training.

I implemented and tested a linear SVM classifier (described in Section **3.2.7**) in order to determine whether the identity of a given stimulus could be determined based on the respones of vS1 neurons. The classifier was trained and tested independently on 7 different subgroups of neurons: 1) all neurons 2) responsive neurons ($t$-test , 2-tailed, $p$ <0.05); 3) non-responsive neurons ($t$-test , 2-tailed, $p$ >= 0.05); 4) excited (response during the stimulation window was significantly greater than during the pre-stimulus window; $t$-test , 1-tailed, $p$ <0.05); 5) suppressed neurons (response during the stimulation window was significantly less than during the pre-stimulus window; $t$-test , 1-tailed, $p$ <0.05); 6) L2 neurons (imaged at 200 $\mu$m below the cortical surface); 7) and L3 neurons (imaged at 320 $\mu$m below the cortical surface). In any given session, the classifier's performance was expressed as the difference between the percentage of stimulus presentations whose identity was correctly predicted and the chance level determined by the most frequently occurring stimulus type (see Section **3.2.7**).

In mice trained on the tactile version of the task, I trained the classifier to identify the frequency of vibrotactile stimulation presented on each trial based on the subgroup of neurons imaged in a single imaging field of vS1 during a single session. I then used a repeated-measures ANOVA to assess changes in classifier performance over the course of training (6 imaging sessions), across mouse identity ($N$ = 3 mice), and across the 6 different subgroups of neurons (i.e. responsive, unresponsive, excited, suppressed, L2 and L3). The neuronal groups and mouse identity were used as predictor variables and the average classifier performance on stimulus identity categorisation (% correct – chance) as the predicted variable (Figure **3.12**). Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated ($\chi^2(14) = 24.73$, $p = 2.23\times10^{-23}$) and I therefore applied a

Greenhouse-Geisser correction.

I found a significant effect of training stage on classifier performance ($F_{(2.39,34.67)}$=21.97, $p = 5.01\times10^{-8}$). I also found a significant interaction between mouse identity and training stage ($F_{(5.78,34.67)}$=24.09, $p = 6.91\times10^{-11}$), indicating individual differences across mice. Using post-hoc pairwise comparisons, I found a significant increase in classifier performance between the pre-training passive session and middle (60-70% correct) and end of training (>70% correct) ($p = 9.11\times10^{-5}$ and $p = 2.8\times10^{-4}$, respectively). Further, classifier performance increased between early training sessions (<60% correct) and both middle and late training sessions ($p = 2.6\times10^{-4}$ and $p = 0.022$ respectively). However, classifier performance did not significantly differ between pre- and post-training passive imaging sessions ($p = 0.222$). More detailed classification performance breakdowns can be found in Table **3.7**. It is also worth noting that I did not find any significant difference in classification performance between different neuronal subgroups ($F_{(17.33,34.67)}$=0.58, $p = 0.95$). Overall, this suggests that neural responses during the task became more informative about vibrotactile frequencies throughout training.

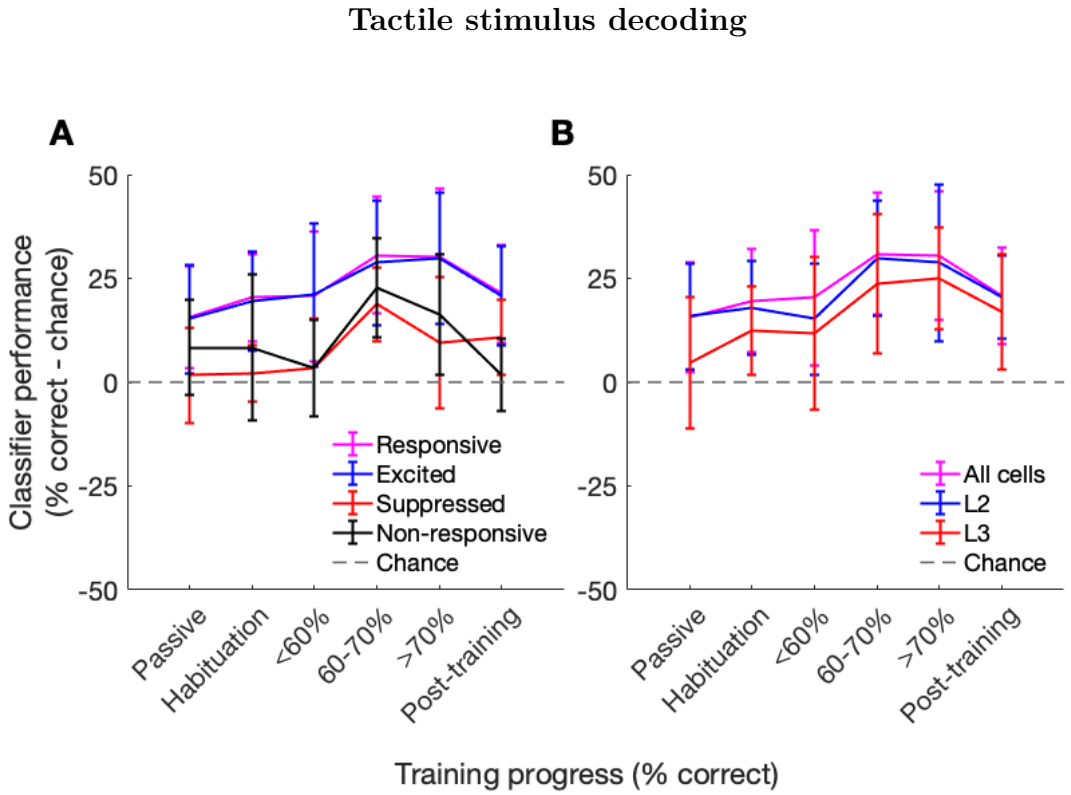



**Figure 3.12:** Classifier performance (mean ± std) over the course of training averaged across mice trained on the tactile task ($N = 3$ mice).

| **Classifier performance (% correct – chance)** | | | | | | |
|---|---|---|---|---|---|---|
| | Pre training | Habituation | <60% correct | 60-70% correct | >70% correct | Post training |
| All neurons | 15.60 ±13.49 | 19.65 ±7.97 | 20.34 ±18.12 | 30.96 ±17.18 | 30.51 ±14.89 | 20.80 ±7.31 |
| Responsive neurons | 15.53 ±13.75 | 20.33 ±4.03 | 20.73 ±16.47 | 30.57 ±16.01 | 30.27 ±15.50 | 21.30 ±7.56 |
| Excited neurons | 15.17 ±15.24 | 19.47 ±3.47 | 21.01 ±18.95 | 28.79 ±17.26 | 29.78 ±15.19 | 20.79 ±8.31 |
| Suppressed neurons | 2.56 ±14.08 | 3.07 ±8.10 | 3.46 ±11.71 | 18.86 ±8.87 | 14.35 ±17.70 | 10.74 ±9.04 |
| Non-responsive neurons | 8.32 ±4.04 | 8.36 ±17.27 | 3.31 ±9.27 | 22.70 ±10.75 | 16.31 ±13.75 | 1.69 ±4.05 |
| L2 neurons | 15.93 ±13.23 | 17.92 ±8.65 | 15.18 ±12.50 | 29.68 ±16.31 | 28.71 ±21.46 | 20.50 ±5.72 |
| L3 neurons | 4.75 ±15.72 | 12.35 ±3.34 | 11.71 ±18.48 | 23.61 ±18.72 | 25.08 ±7.00 | 16.91 ±12.13 |

**Table 3.7:** Classifier performance (% correct – chance) on vibrotactile stimulus identity during training across mice trained on the vibrotactile task ($N = 3$ mice).

I used a repeated-measures ANOVA to assess changes in classifier performance on different neuron sub-populations over the course of training, in mice trained on the multisensory task ($N = 3$ mice). The classifier was trained and tested separately using seven different classes of neurons as input: all neurons, all responsive neurons, excited neurons, suppressed neurons, non-responsive neurons, L2 neurons, and L3 neurons. I set the neuron subgroups, predicted modality (auditory, tactile or audio-vibrotactile pairing) and mouse identity as predictor variables and the average classifier performance on stimulus identity categorisation (% correct – chance) as the predicted variable. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated ($\chi^2(14) = 76$, $p = 2.74\times10^{-20}$) and I therefore applied a Greenhouse-Geisser correction.

I found a significant effect of training stage on classifier performance

($F_{(3.12,162.41)}$=40.97, $p = 2.74\times10^{-20}$). As in the tactile task, I also found a significant interaction between mouse identity and training stage ($F_{(6.25,162.41)}$=67.43, $p = 2.14\times10^{-66}$). Similarly to classification performance in mice trained on the tactile task, I did not find a significant difference between performance on different neuronal subgroups ($F_{(18.74,162.41)}$=1.18, $p = 0.25$).

These findings highlight the fact that classifier performance on different neuron subpopulations did not vary within mice. However, I again observed that the classifier performance varied between individual animals (Figures **3.13**, **3.14** and **3.15**, panels **A-B**, **C-D** and **E-F** respectively). Similarly to mice trained on the tactile task, there was also a significant effect of training stage on classifier performance. However, in contrast to the tactile task, classifier performance significantly improved between the pre-training passive session and post-training imaging ($p = 0.0156$) on the multisensory task. Detailed average classifier performances for each neuron subpopulation across training can be found in Table **3.8**.

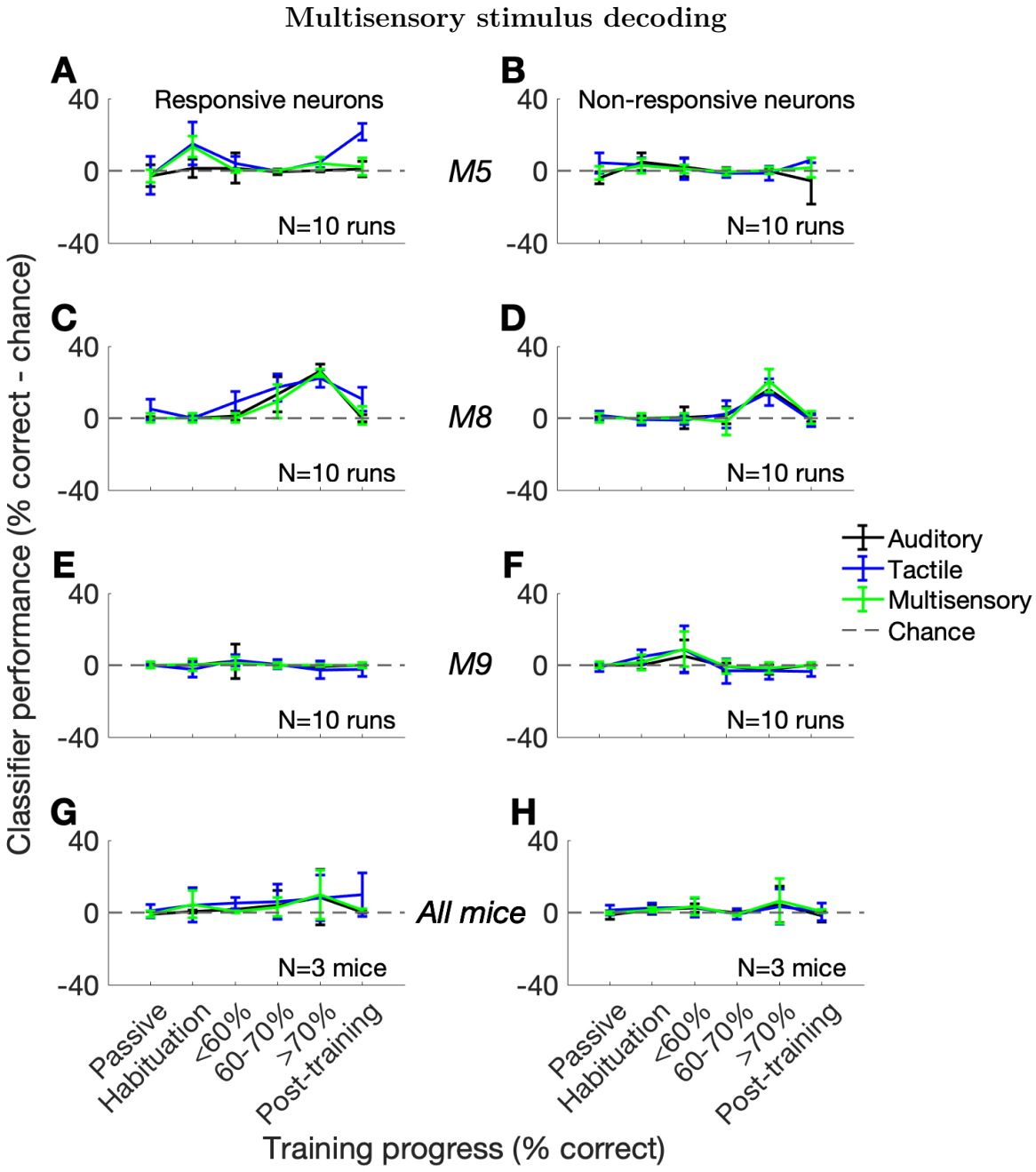



**Figure 3.13:** Performance in stimulus decoding of responsive neurons (**A**, **C**, **E** & **G**) and non-responsive neurons (**B**, **D**, **F** & **H**). The classifier was trained and tested on neural data to determine the auditory, tactile or multisensory identity of trials based on neural calcium transients (black: auditory, blue: tactile, green: multisensory combination, mean±std across 10 train-test repeats). Panels **A** and **B** correspond to the classifier performance in mouse M5, panels **C** and **D** correspond to mouse M8, panels **E** and **F** correspond to mouse M9 and panels **G** and **H** correspond to the average performance across all three mice (mean±std of the means for each mouse).

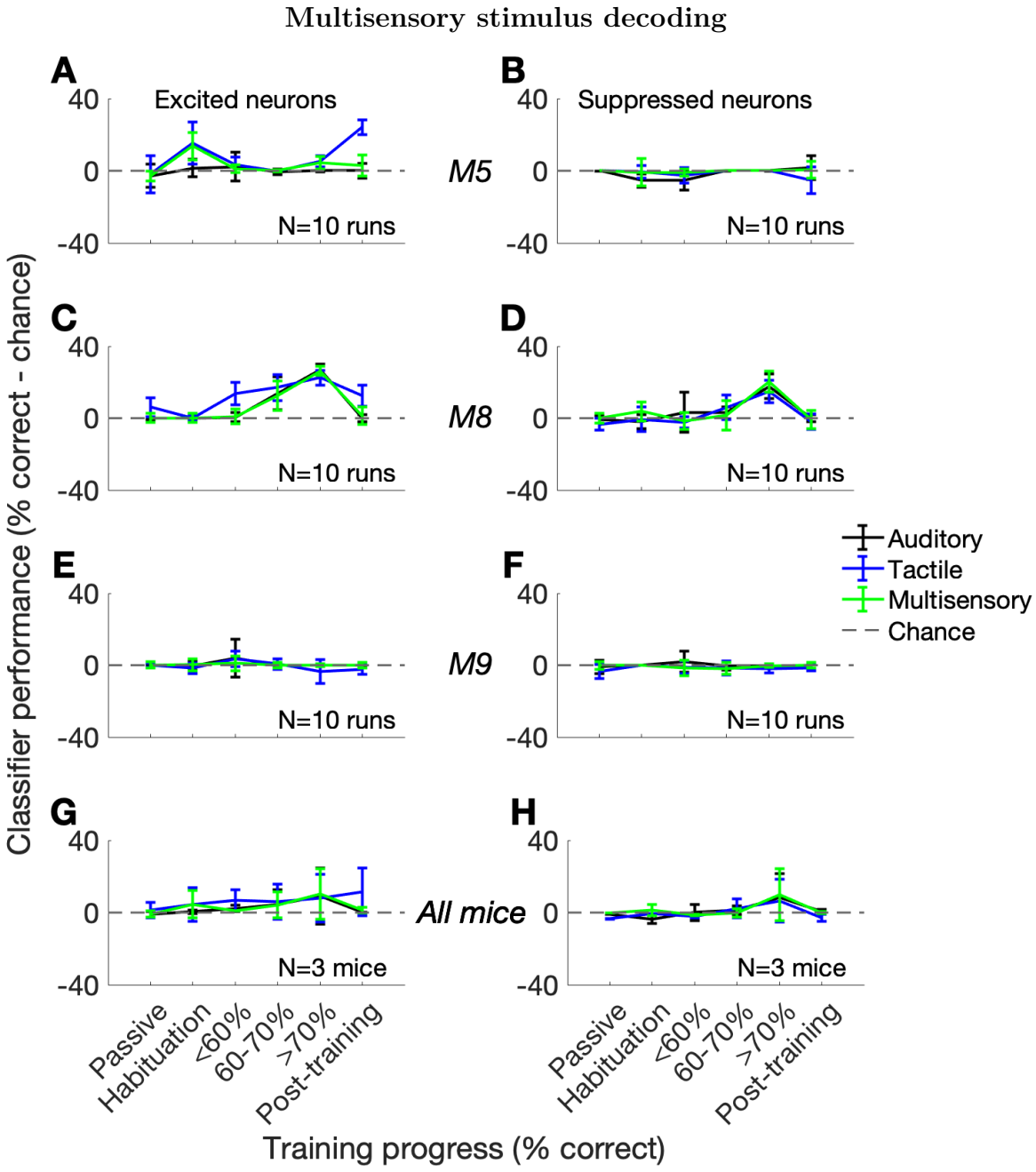



**Figure 3.14:** Performance in stimulus decoding of excited responsive neurons (**A**, **C**, **E** & **G**) and suppressed responsive neurons (**B**, **D**, **F** & **H**). The classifier was trained and tested on neural data to determine the auditory, tactile or multisensory identity of trials based on neural calcium transients (black: auditory, blue: tactile, green: multisensory combination, mean±std across 10 train-test repeats). Panels **A** and **B** correspond to the classifier performance in mouse M5, panels **C** and **D** correspond to mouse M8, panels **E** and **F** correspond to mouse M9 and panels **G** and **H** correspond to the average performance across all three mice (mean±std of the means for each mouse).

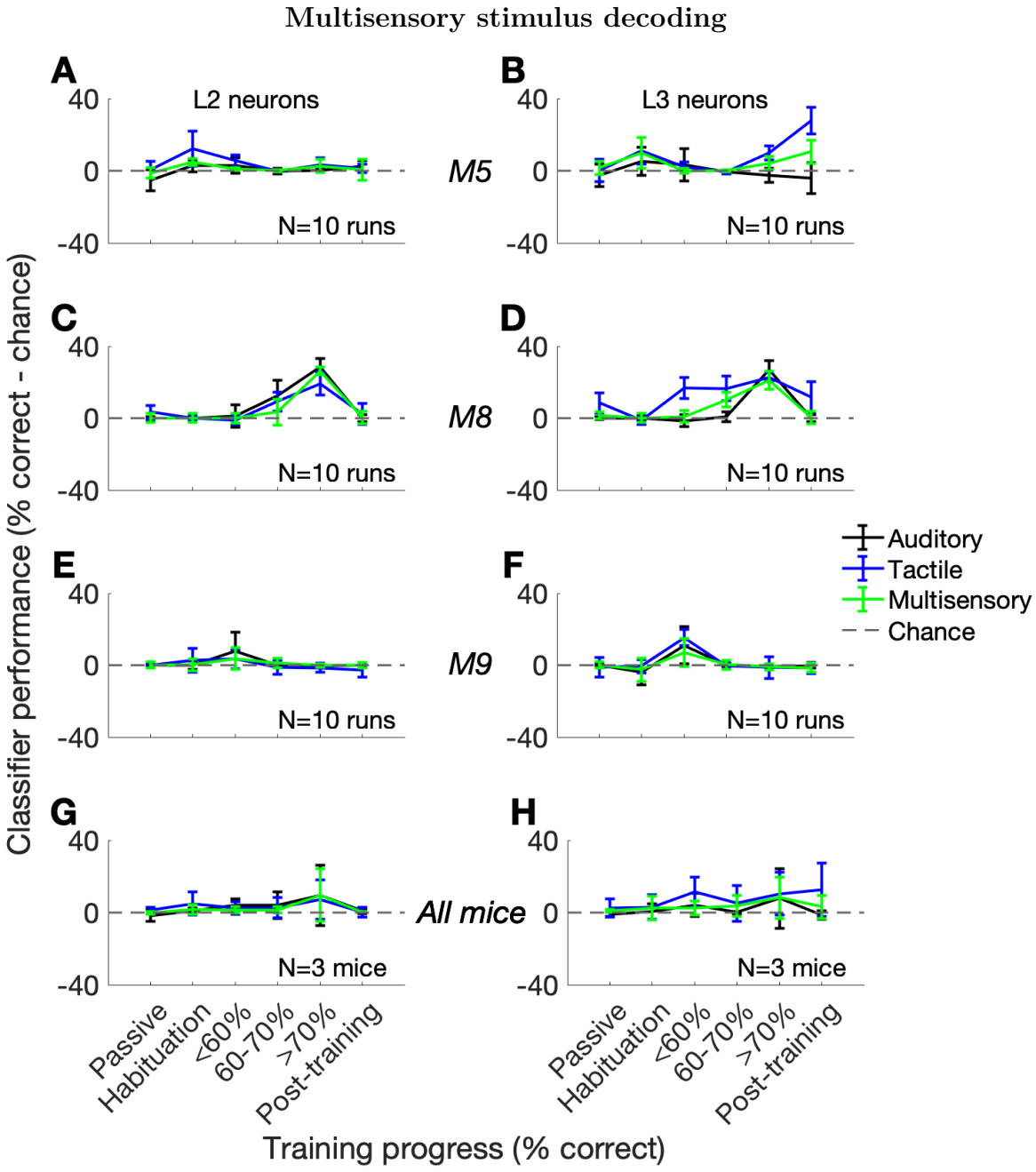



**Figure 3.15:** Performance in stimulus decoding of L2 neurons (**A**, **C**, **E** & **G**) and L3 neurons (**B**, **D**, **F** & **H**). The classifier was trained and tested on neural data to determine the auditory, tactile or audio-vibrotactile multisensory identity of trials based on neural calcium transients (black: auditory, blue: tactile, green: audio-vibrotactile multisensory combination, mean±std across 10 train-test repeats). Panels **A** and **B** correspond to the classifier performance in mouse M5, panels **C** and **D** correspond to mouse M8, panels **E** and **F** correspond to mouse M9 and panels **G** and **H** correspond to the average performance across all three mice (mean±std of the means for each mouse).

| **Classifier performance on categorisation of multisensory stimuli (% correct – chance)** | | | | | | |
|---|---|---|---|---|---|---|
| | Pre training | Habituation | <60% correct | 60-70% correct | >70% correct | Post training |
| All neurons | -0.94 ±1.84 | 4.57 ±7.61 | 1.61 ±2.62 | 1.79 ±2.96 | 9.82 ±13.55 | 0.82 ±1.17 |
| Responsive neurons | -0.89 ±1.76 | 4.63 ±7.70 | 0.46 ±0.63 | 3.09 ±5.37 | 9.83 ±13.52 | 1.20 ±1.07 |
| Excited neurons | -0.87 ±1.72 | 4.66 ±7.75 | 1.01 ±0.08 | 4.19 ±7.19 | 10.23 ±14.05 | 1.41 ±1.43 |
| Suppressed neurons | -0.13 ±0.06 | 1.46 ±3.19 | -1.33 ±0.37 | -0.16 ±2.38 | 9.99 ±14.40 | -0.02 ±0.63 |
| Non-responsive neurons | -0.29 ±0.72 | 1.55 ±1.36 | 3.36 ±4.80 | -1.09 ±0.79 | 6.51 ±12.43 | 0.68 ±0.94 |
| L2 neurons | -0.35 ±0.81 | 1.78 ±2.77 | 1.51 ±1.86 | 1.47 ±1.75 | 9.61 ±14.52 | 0.27 ±0.24 |
| L3 neurons | 1.05 ±0.83 | 2.52 ±6.53 | 2.64 ±3.80 | 3.56 ±5.84 | 8.17 ±11.53 | 3.16 ±6.52 |

| **Classifier performance on tactile categorisation of multisensory stimuli (% correct – chance)** | | | | | | |
|---|---|---|---|---|---|---|
| | Pre training | Habituation | <60% correct | 60-70% correct | >70% correct | Post training |
| All neurons | -0.18 ±3.82 | 5.87 ±9.62 | 5.52 ±4.38 | 4.52 ±8.63 | 7.44 ±10.95 | 8.14 ±13.03 |
| Responsive neurons | 0.87 ±3.77 | 4.22 ±9.49 | 5.21 ±3.24 | 5.85 ±9.70 | 8.07 ±12.65 | 9.85 ±12.05 |
| Excited neurons | 1.40 ±4.26 | 4.57 ±9.37 | 6.82 ±5.89 | 5.88 ±9.68 | 8.05 ±13.23 | 11.47 ±13.19 |
| Suppressed neurons | -3.49 ±0.08 | -0.68 ±0.07 | -1.99 ±0.81 | 2.22 ±5.28 | 6.44 ±11.82 | -2.83 ±2.04 |
| Non-responsive neurons | 1.48 ±2.78 | 2.35 ±2.86 | 2.82 ±5.21 | -0.78 ±2.83 | 3.31 ±9.63 | 0.49 ±4.89 |
| L2 neurons | 1.26 ±1.80 | 4.96 ±6.34 | 2.55 ±3.41 | 2.66 ±5.71 | 7.04 ±10.87 | 0.33 ±2.72 |
| L3 neurons | 2.51 ±5.16 | 3.06 ±6.85 | 11.49 ±8.07 | 5.14 ±9.91 | 10.46 ±11.97 | 12.73 ±14.69 |

| **Classifier performance on auditory categorisation of multisensory stimuli (% correct – chance)** | | | | | | |
|---|---|---|---|---|---|---|
| | Pre training | Habituation | <60% correct | 60-70% correct | >70% correct | Post training |
| All neurons | -1.14 ±1.98 | 1.29 ±2.14 | 2.14 ±3.35 | 3.64 ±6.87 | 9.17 ±16.28 | -0.49 ±0.91 |
| Responsive neurons | -0.93 ±1.62 | 0.48 ±0.86 | 1.65 ±0.50 | 4.22 ±7.87 | 8.50 ±15.36 | 0.34 ±0.53 |
| Excited neurons | -0.93 ±1.62 | 0.43 ±0.89 | 2.21 ±1.72 | 4.37 ±8.13 | 9.01 ±15.59 | 0.03 ±0.02 |
| Suppressed neurons | -0.88 ±0.10 | -3.75 ±2.38 | 0.05 ±4.53 | 1.32 ±2.47 | 8.54 ±12.86 | 0.41 ±1.23 |
| Non-responsive neurons | -1.30 ±2.27 | 1.71 ±2.87 | 2.52 ±2.42 | -0.05 ±1.49 | 4.48 ±10.10 | -1.84 ±3.25 |
| L2 neurons | -1.71 ±2.96 | 1.00 ±1.64 | 4.03 ±3.44 | 4.11 ±7.34 | 9.40 ±16.69 | 0.87 ±1.45 |
| L3 neurons | -0.81 ±1.41 | 0.43 ±4.57 | 4.25 ±6.34 | 0.08 ±0.66 | 7.85 ±16.47 | -1.42 ±2.29 |

**Table 3.8:** Classifier performance (% correct – chance) on vibrotactile stimulus identity during training across mice trained on the vibrotactile task ($N = 3$ mice).

In summary, I trained a classifier to categorise vibrotactile stimuli in mice trained on the unisensory tactile task as well as to categorise auditory, tactile, and audio-vibrotactile stimuli in mice trained on the multisensory task. I found that the amount of information about stimulus identity varied across training. I originally expected to find that, as mice trained, the decoding performance of the classifier would reliably improve. However, while the classifier performance significantly varied over the course of training in both cohorts, I only found a significant improvement in classification performance between pre-training and post-training passive imaging sessions in mice trained on the multisensory task. In addition, there were no significant differences in decoding performance between the auditory and tactile features of multisensory stimuli and audio-vibrotactile pairs. Finally, while there were no significant differences in performance based on different neuronal subgroups (based

on responsiveness or cortical depth), I found significant differences in the decoding performance of neurons imaged across individual mice, in both multisensory-trained and vibrotactile-trained mice.

### 3.3.7 Training alters the responses of most neurons to the trained stimuli

These above findings highlight transient changes within neural populations as well as broader population-wide modifications of responsiveness within vS1. Having observed such effects on whole populations, I next aimed to put the ability to track individual cells across days to profit in order to investigate how a single neuron's responses to stimuli varied throughout training.

Here, I focused on the neurons I was able to track throughout the course of training (an average of 45 neurons (±29 per mouse). For every tracked neuron, I first determined whether it was responsive or not for both passive sessions as well as during training. I then defined each neuron's response stability. Neurons were classified as: always responsive; never responsive; responsive pre-training; responsive post-training; responsive during training; or non-responsive during training. These categories are further described in Table **3.9**.

| **Responsiveness stability of tracked neurons** | |
|---|---|
| **Always responsive** | Either excited or suppressed in all imaging sessions |
| **Never responsive** | Neither excited nor suppressed at any stage of training |
| **Responsive post** | Non-responsive before training (i.e. pre-training passive session) and responsive after training (post-training passive session). |
| **Responsive pre** | Responsive before training (i.e. pre-training passive session) and non-responsive after training (post-training passive session). |
| **Responsive during training** | Non-responsive in both pre- and post-training passive sessions and responsive in at least one session during training. |
| **Non-responsive during training** | Responsive in both pre- and post-training passive sessions and non-responsive in at least one session during training. |

**Table 3.9:** Categories of response stability of tracked neurons. Across mice, each tracked neuron was categorised depending on how stable its responses were across the course of training. Neurons were classified as: always responsive, never responsive, responsive post-training, responsive pre-training, responsive during training, or non-responsive during training. Each neuron could only fall within one category and all tracked neurons were categorised.

I performed a two-way ANOVA with the proportions of neurons exhibiting each of the response stability patterns in each mouse as the predicted variable, and the task the mice were trained on (auditory ($N = 1$), tactile ($N = 3$) or multisensory ($N = 3$)) and stability patterns as the predictors. I found no significant effect of task type on response stability ($F_{(2,24)} = 0,\ p = 1$)). In contrast, I did find significant differences between the proportions of response stability types ($F_{(5,24)} = 12.16,\ p = 6.07\times10^{-6}$). More specifically, post-hoc pairwise comparisons revealed that there were no differences in the proportion of neurons responsive post-training (i.e. were responsive in the post-training but not during the pre-training sessions) compared to those responsive pre-training ($p = 0.5789$). Finally, more neurons were responsive during training than in both passive sessions before and after the training ($p = 3.41\times10^{-5}$) (Figure **3.16**). Detailed proportions can be found in Table **3.10**.

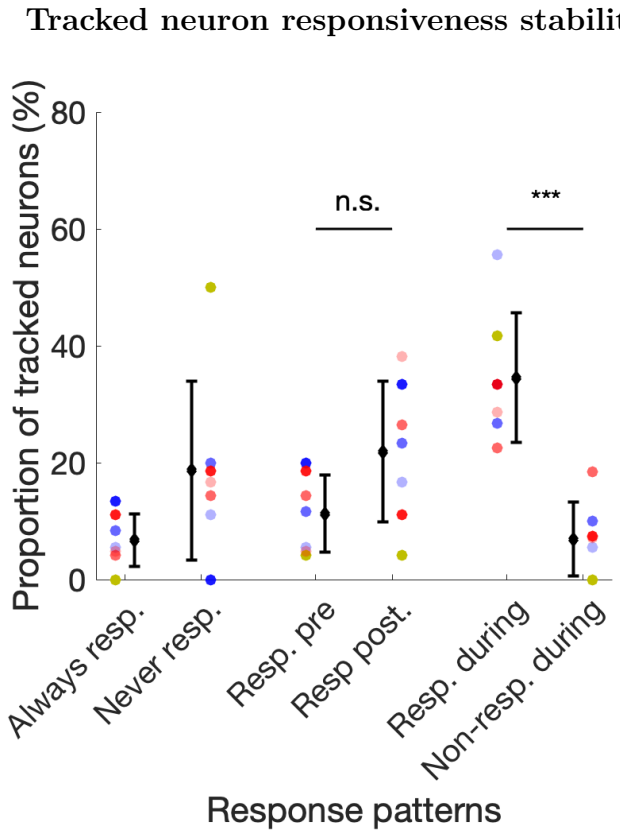



**Figure 3.16:** Proportions of tracked neurons with various response patterns: always resp.=always responsive; never resp.=never responsive; resp. pre=responsive pre-training; resp. post=responsive post-training, resp. during=responsive during training, non-resp. during=non-responsive during training. Each coloured dot corresponds to a single mouse. Mice trained on the tactile (blue), multisensory (red/orange), and auditory (green/yellow) are indicated by their colour. n.s. = non-significant, *** $p$ <0.001, post-hoc pairwise comparisons.

| **Proportions of response profiles (% of tracked neurons ± std)** | | | | | |
|---|---|---|---|---|---|
| **Always resp.** | **Never resp.** | **Resp. pre** | **Resp. post** | **Resp. during** | **Non-resp. during** |
| 6.74 ±4.53 | 18.65 ±15.34 | 11.28 ±6.63 | 21.89 ±12.07 | 34.51 ±11.10 | 6.92 ±6.30 |

**Table 3.10:** Proportions of responsiveness profile changes of tracked neurons. Across mice, each tracked neuron was categorised depending on how stable its responses were across the course of training. Neurons were classified as always responsive, never responsive, responsive post-training, responsive pre-training, responsive during training or non-responsive during training. Each neuron could only fall within one category and all neurons were categorised. The proportions shown are out of all tracked neurons.

To summarise, I categorised changes in individual neuron's responses across training and found that there were significantly more neurons responsive during training than non-responsive during training. Overall, these findings show that a minority of tracked neurons were constant in their responsiveness (or lack thereof), as fewer than 30% of neurons were either always responsive or never responsive throughout all training stages. While the small sample size of tracked neurons prevents wider interpretations about population-wide responses, this finding highlights a broad distinction between neurons that remain responsive or non-responsive over the whole course of training, those who start unresponsive but become responsive by the post-training passive session and, finally, neurons exhibiting either responsiveness or a loss of responsiveness during training while pre- and post-training responsiveness measures remained unchanged. Hence, while the classifier may have failed to decode stimulus information from neural population activity during training, and even post-training in mice trained on the tactile task, stimulus-evoked responses of a subset of neurons varied significantly throughout training, suggesting single-neuron changes in responsivity as a consequence of

training.

### 3.3.8 Neurons in vS1 are modulated by licking

One possible interpretation for the changes in responses and proportions of responsive neurons described throughout this chapter could be that these changes are linked to an animal's behaviour. For instance, except for the pre- and post-training passive sessions during which the animals were not water-regulated or even had a spout within reach, animals could lick to obtain a reward throughout the training protocol. Therefore, one could argue that transient changes in responses during training are actually due to 1) movement of the imaging plane caused by locomotion associated with licking; 2) motor signals within vS1; or 3) neural activity linked to goal-oriented behaviour observed during whisker-guided tasks but not during passive whisker stimulation.

To rule out possible artifacts due to animal locomotion or licking action, I assessed whether neurons exhibited stronger responses when the animal was licking during a stimulus compared to presenting a sensory stimulus alone. I compared neural responses across: 1) 'Lick trials' (i.e. either 'Hit' or FA); 2) 'No-lick trials' ('Miss' or CR); 3) licks during inter-trial periods, at which point there were no sensory stimuli ('Lick'); and 4) times of inter-trial intervals at which no licks or sensory stimulus occurred ('No-lick'). These 4 conditions are summarised in Figure **3.17**.

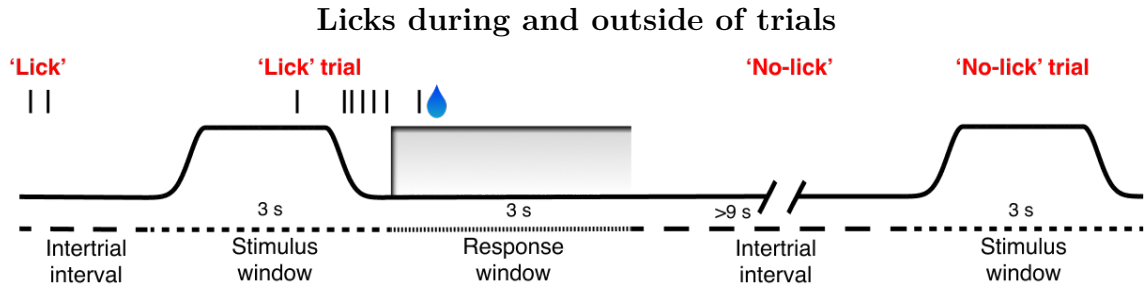



**Figure 3.17:** Summary of the timing of different neural responses for the analysis of licking artifacts. The thick black line depicts the amplitude of stimulus across two trials, including a lick window and intertrial interval between trials and preceding the first one. Licks outside of a stimulus window were defined as 'Lick'. Licks within a stimulus window were called 'Lick trials'. Conversely, an absence of lick during an intertrial interval was defined as 'No-lick' whereas a stimulus window without licks was categorised as a 'No-lick trial'.

At each stage of training, I measured neural responses during licks in trials (6s of recording, centred on the onset of licking instead of the stimulus onset, but during stimulus presentation), during licks outside of trials (6s relative to licking onset), during trials without licks (6s centred around stimulus onset) and finally outside of trials without licks (6s within a period of at least 9s without any licks).

I used a repeated-measures ANOVA with the average neural response amplitude within a given 6 s window as the predicted variable and whether a trial (i.e. stimulus) was occurring, whether a mouse licked or not, the task type (auditory, tactile, multisensory) and the response profile of the neurons (as defined in Section **3.3.5**) as predictors. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated ($\chi^2(5) = 32.74$, $p = 4.24\times10^{-6}$ and I therefore applied a Greenhouse-Geisser correction. I found a significant effect of training stage on responses ($F_{(2.53,313.88)}$=7.47, $p = 2.11\times10^{-4}$), as described in previous sections. More importantly, while there was a significant effect of stimulus (i.e. was a trial occurring?) ($F_{(2.53,313.88)}$=5.29, $p = 2.69\times10^{-3}$), there was no significant effect of

licking ($F_{(2.53,313.88)}$=0.98, $p$ = 0.394). It is however worth noting that there was a significant interaction between licking behaviour and stimulus ($F_{(2.53,313.88)}$=7.72, $p$ = 1.57×10$^{-4}$). In contrast, there was no interaction between the response profile of neurons and licking behaviour ($F_{(10.13,313.88)}$=0.62, $p$ = 0.802).

**In mice trained on the tactile task.**

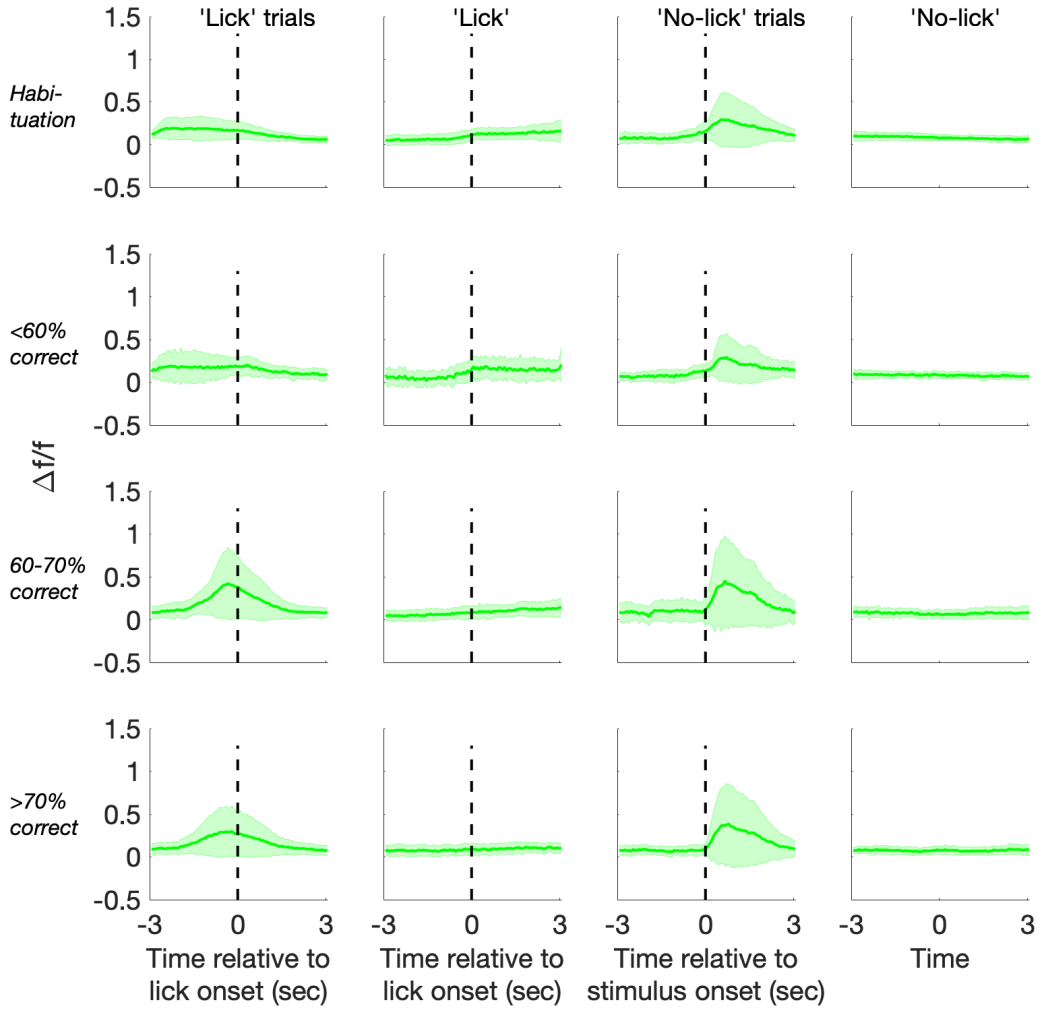



**Figure 3.18:** Responses of sharp onset neurons in 'Lick' trials, around licks (i.e. 'Lick' behaviour but no stimulus), 'No-lick' trials and 'No-lick' behaviour outside of trials. Data shown is mean ± std, $N$ = 3 mice.

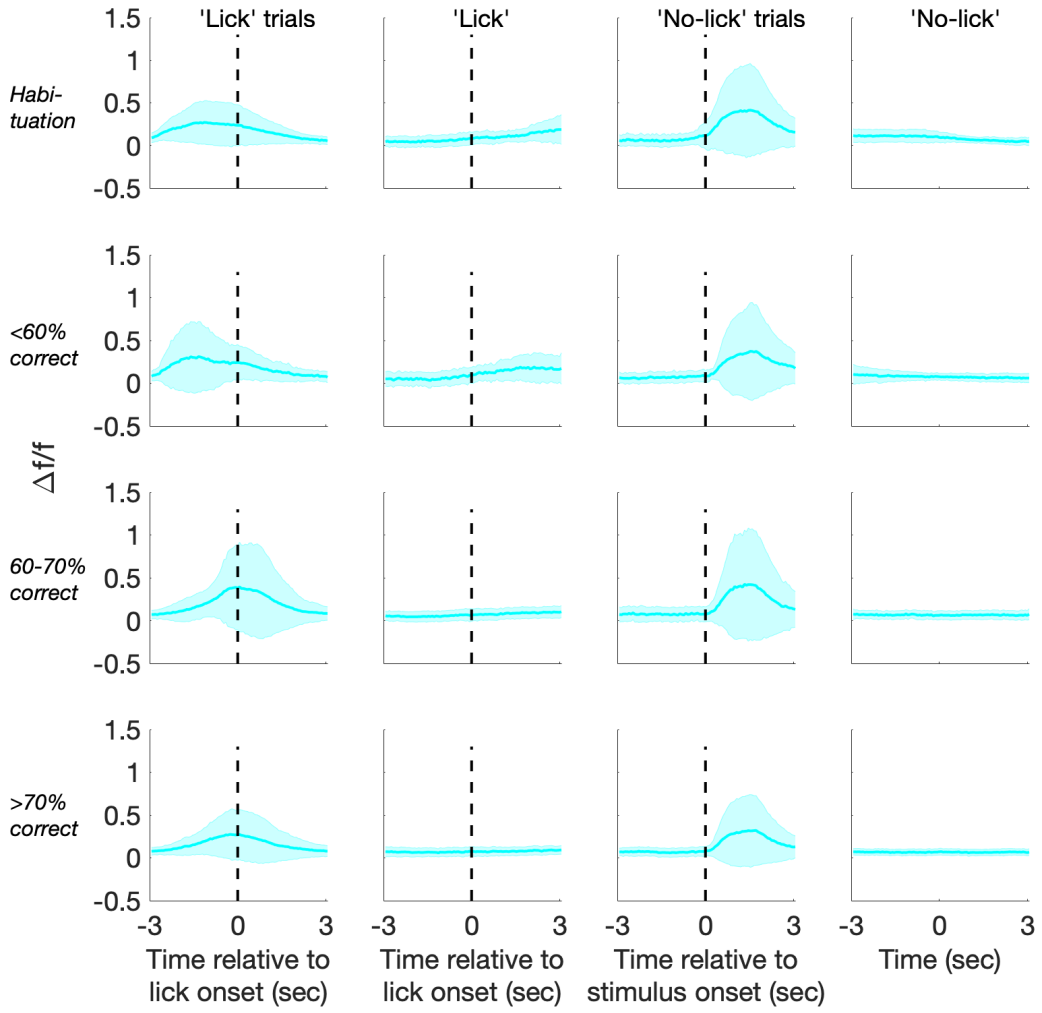



**Figure 3.19:** Responses of amplitude-locked neurons in mice trained on the tactile task in 'Lick' trials, around licks (i.e. 'Lick' behaviour but no stimulus), 'No-lick' trials and 'No-lick' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

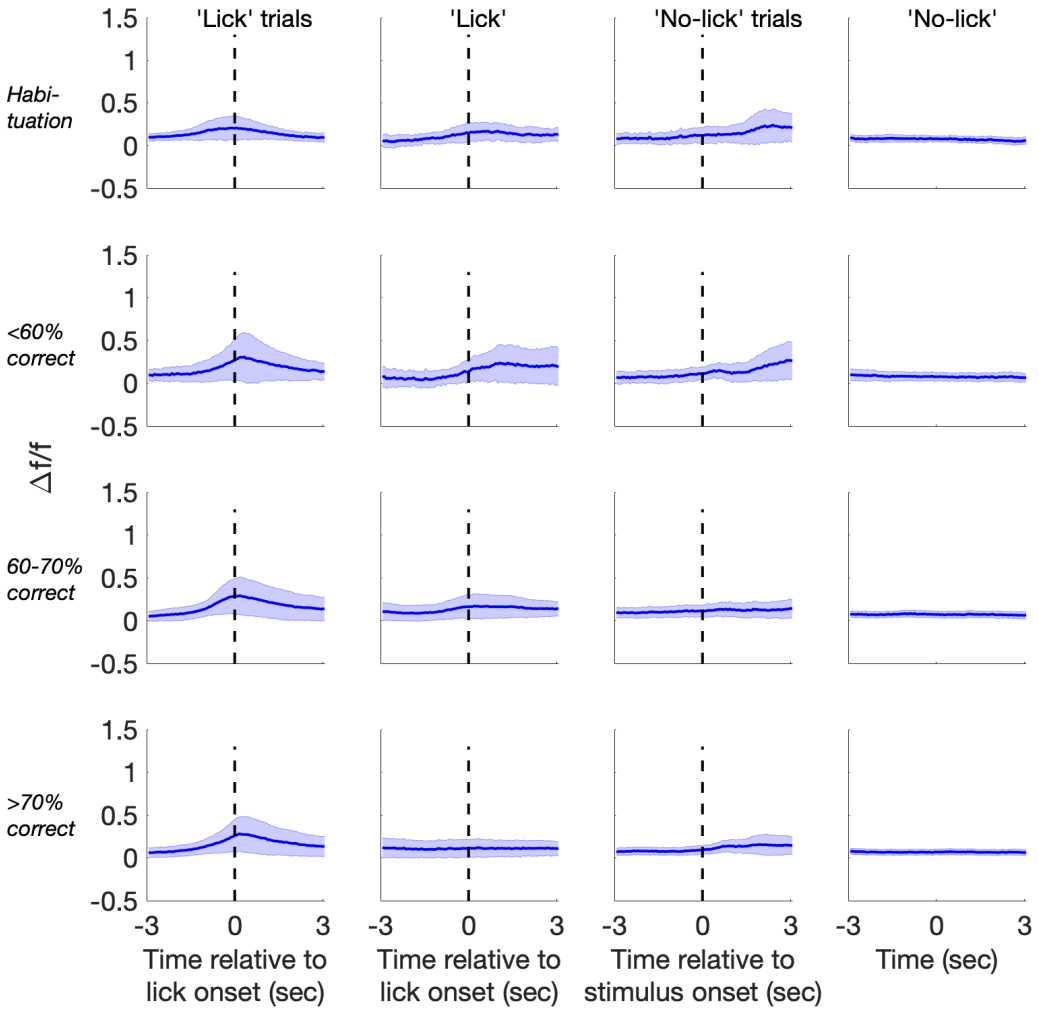



**Figure 3.20:** Responses of delayed onset neurons in mice trained on the tactile task in 'Lick' trials, around licks (i.e. 'Lick' behaviour but no stimulus), 'No-lick' trials and 'No-lick' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

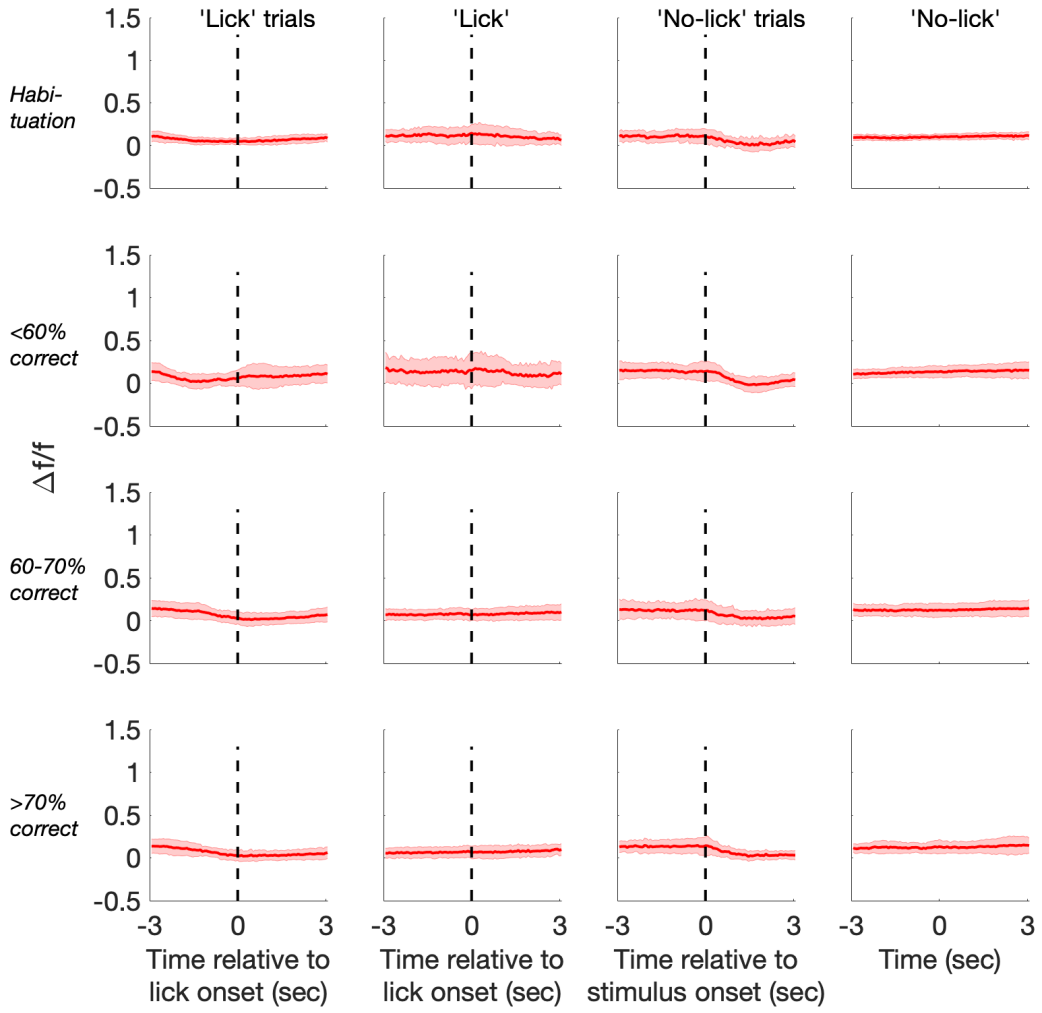



**Figure 3.21:** Responses of suppressed neurons in mice trained on the tactile task in 'Lick' trials, around licks (i.e. 'Lick' behaviour but no stimulus), 'No-lick' trials and 'No-lick' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

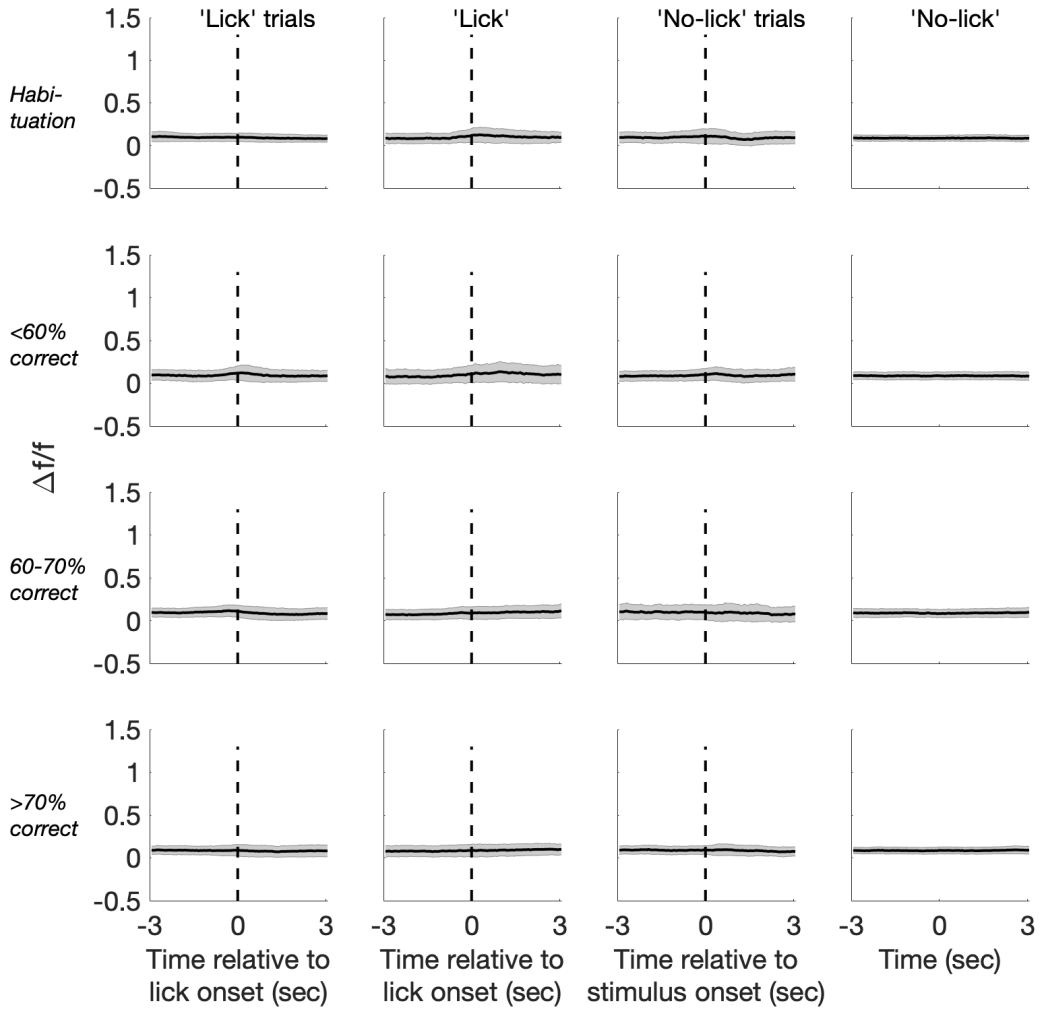



**Figure 3.22:** Responses of non-responsive neurons in mice trained on the tactile task in 'Lick' trials, around licks (i.e. 'Lick' behaviour but no stimulus), 'No-lick' trials and 'No-lick' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

**In mice trained on the multisensory task.**

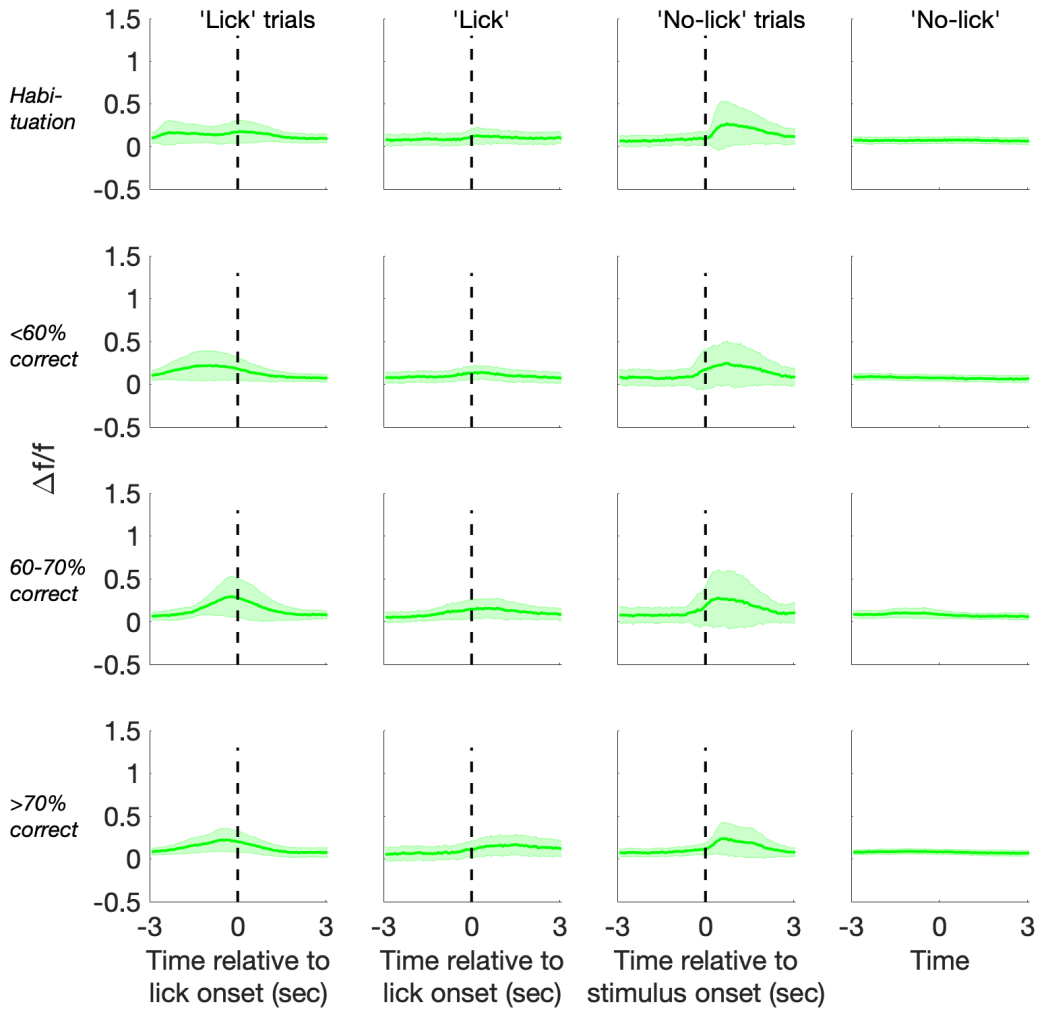



**Figure 3.23:** Responses of sharp onset neurons in mice trained on the multisensory task in 'Lick' trials, around licks (i.e. 'Lick' behaviour but no stimulus), 'No-lick' trials and 'No-lick' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

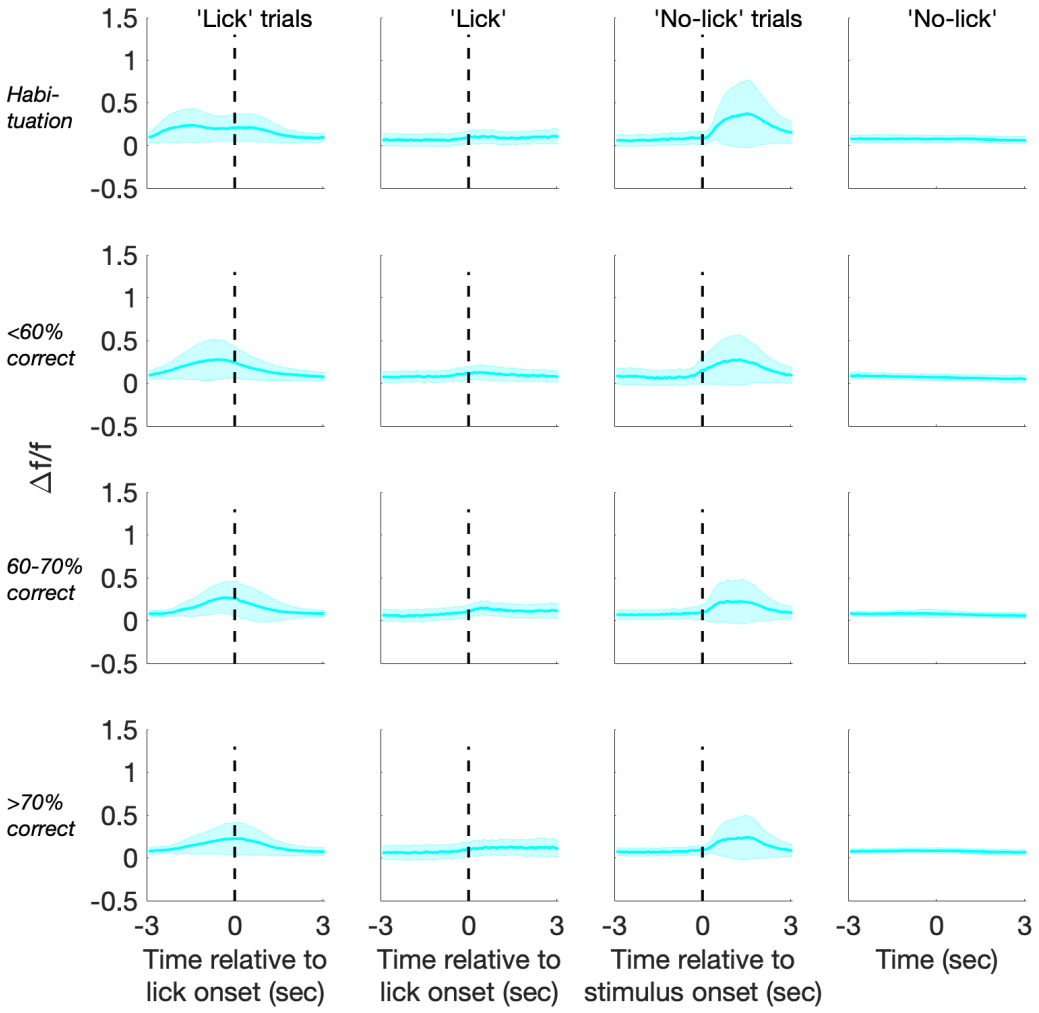



**Figure 3.24:** Responses of amplitude-locked neurons in mice trained on the multisensory task in 'Lick' trials, around licks (i.e. 'Lick' behaviour but no stimulus), 'No-lick' trials and 'No-lick' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

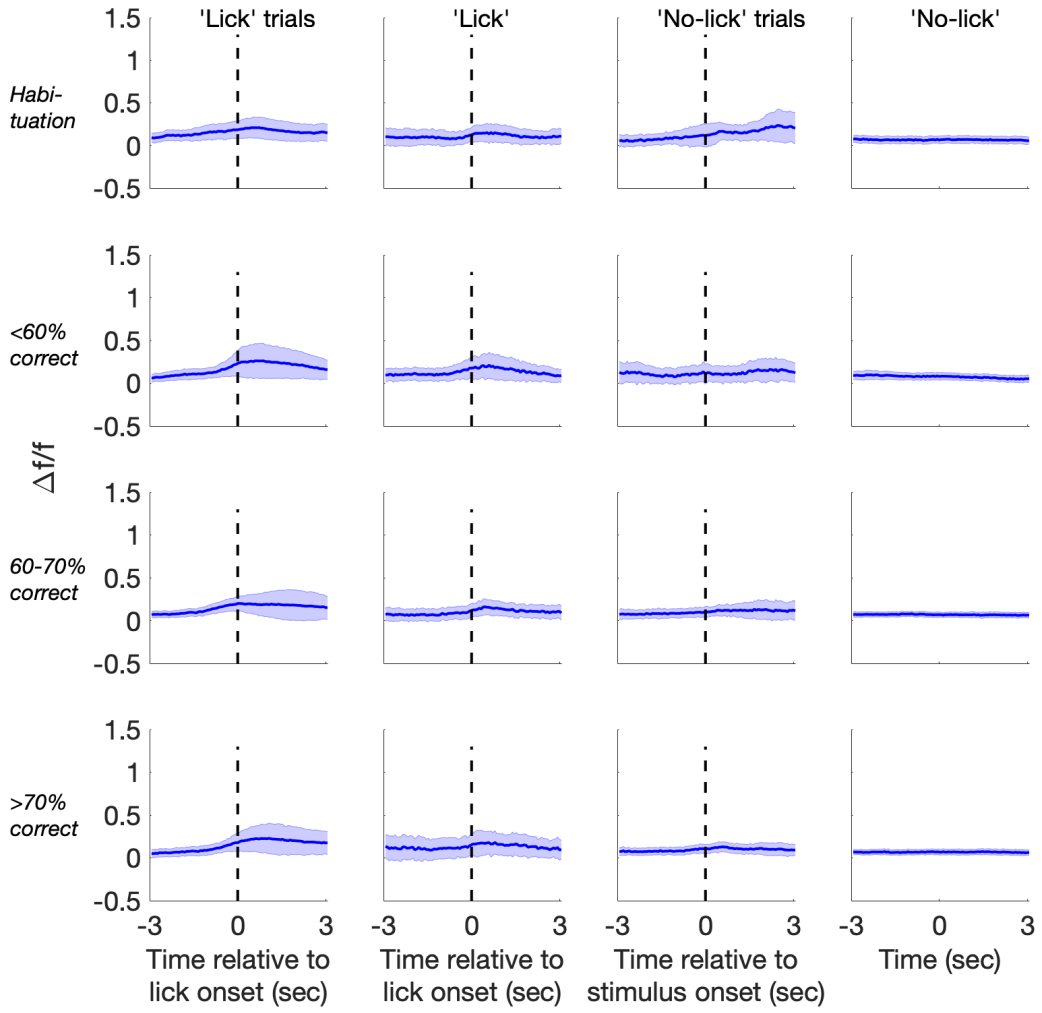



**Figure 3.25:** Responses of delayed onset neurons in mice trained on the multisensory task in 'Lick' trials, around licks (i.e. 'Lick' behaviour but no stimulus), 'No-lick' trials and 'No-lick' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

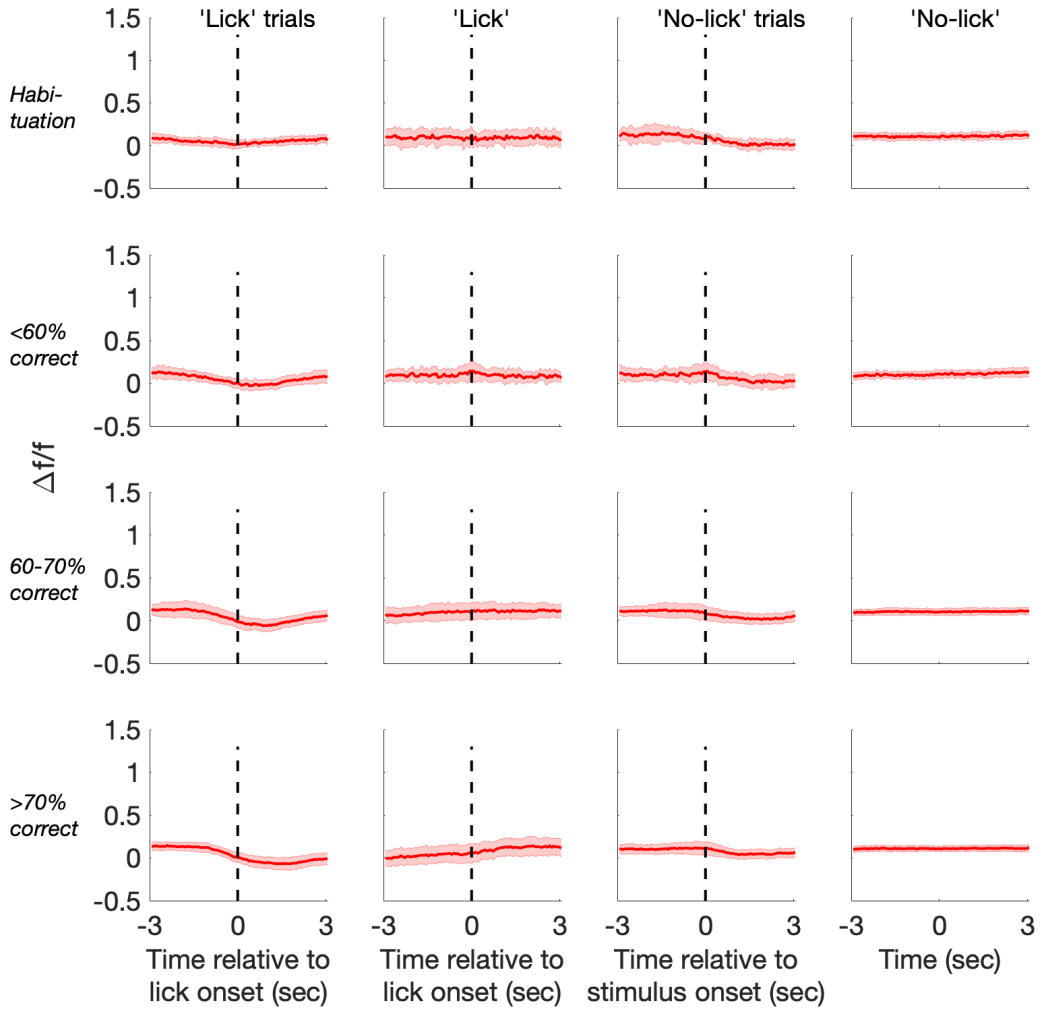



**Figure 3.26:** Responses of suppressed neurons in mice trained on the multisensory task in 'Lick' trials, around licks (i.e. 'Lick' behaviour but no stimulus), 'No-lick' trials and 'No-lick' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

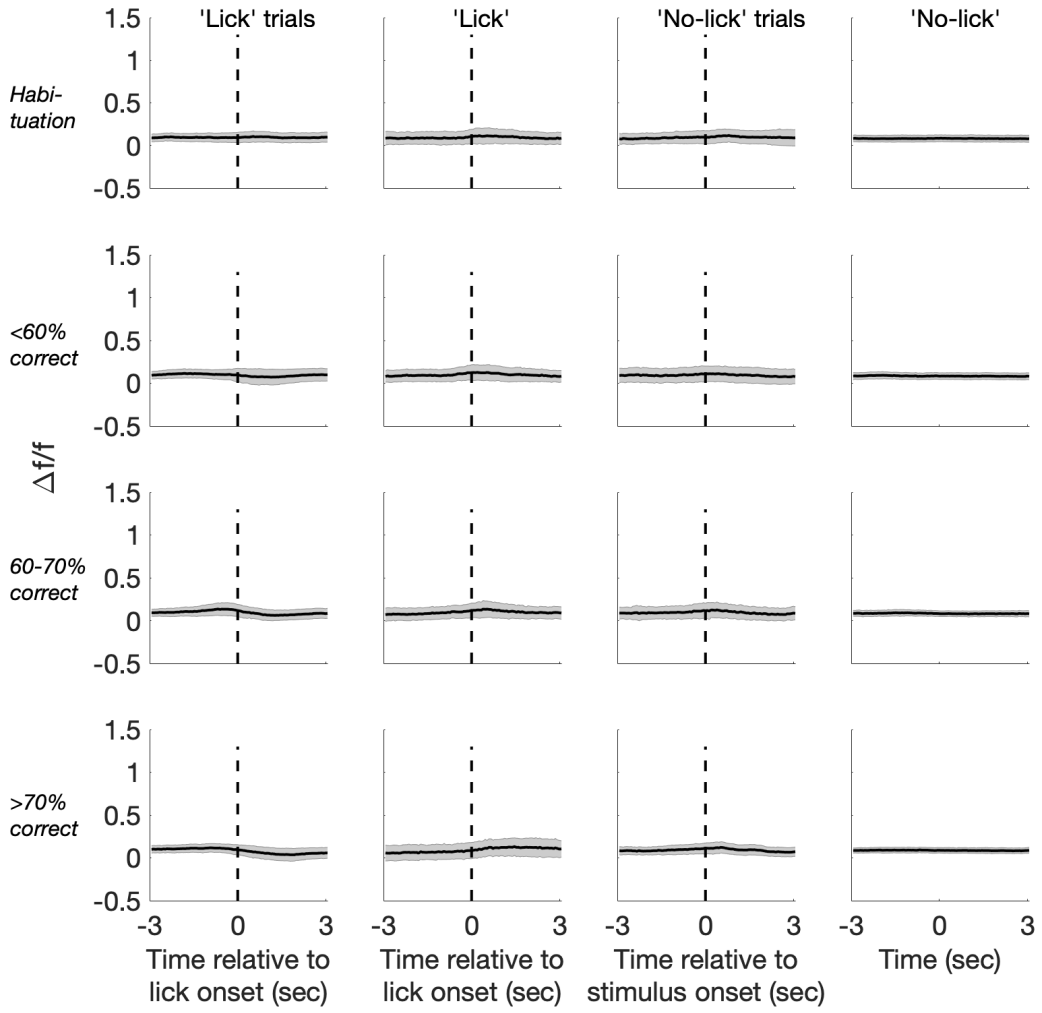



**Figure 3.27:** Responses of non-responsive neurons in mice trained on the multisensory task in 'Lick' trials, around licks (i.e. 'Lick' behaviour but no stimulus), 'No-lick' trials and 'No-lick' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

To summarise, regardless of their response profile, neurons were not responsive to licks alone. There were no significant differences between neural responses during intertrial intervals with or without licking activity. Hence, the changes observed in neural responses cannot simply be attributed to behavioural artifacts. However, there was a significant interaction between licking behaviour and stimulus presentation, suggesting that while licks alone do not affect neural responses, trial outcome and/or

the goal oriented behaviour underlying licking during stimulus presentation modulate neural responses to stimuli.

### 3.3.9 Neurons in vS1 are modulated by locomotion.

Another possible reason explaining changes in responses throughout training could be locomotion, through either 1) movement of the imaging plane or 2) motor signals linked to either locomotion or whisking. The behavioural recording apparatus did not allow me to track whiskers and whisking behaviour, but it did allow me to track the mouse's running.

I examined whether neurons exhibited stronger responses to a sensory stimulus while it was running by comparing responses in 'Run trials' (i.e. either 'Hit' or FA), 'No-run trials' ('Miss' or CR), bouts of running during inter-trial periods, at which point there were no sensory stimuli ('Run'), and times of inter-trial intervals at which no running or sensory stimulus occurred ('No-run'). These conditions are summarised in Figure **3.28**. Their categorisation and the analyses of them are similar that that described for licking above.

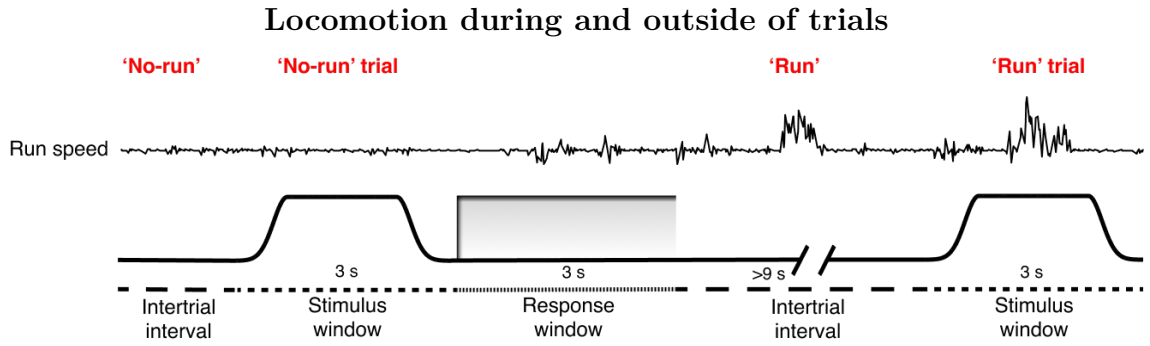



**Figure 3.28:** Summary of the timing of different neural responses for the analysis of locomotion artifacts. The thick black line depicts the amplitude of stimulus across two trials, including a run window and intertrial interval between trials and preceding the first one. The run speed line depicts typical locomotion traces during a session. Locomotion outside of a stimulus window was defined as 'Run'. Runs within a stimulus window were called 'Run trials'. Conversely, an absence of run during an intertrial interval was defined as 'No-run' whereas a stimulus window without locomotion was categorised as a 'No-run trial'.

During both passive sessions and at each stage of training, I measured neural responses around bouts of locomotion in trials (6 s of recording, centred on the onset of running instead of the stimulus onset), around runs outside of trials (relative to running onset), around trials without runs (6s centred around stimulus onset) and finally outside of trials without locomotion (6 s within at least 9 s without any running).

I used a repeated-measures ANOVA with the average neural response amplitude within a 6 s repsonse windows as the predicted variable and whether a trial (i.e. stimulus) was occurring, whether a mouse ran or not, the task type mice were trained on and the response profile of the neurons (as defined in Section **3.3.5**) as predictors. Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated ($\chi^2(14) = 178.7$, $p = 1.21\times10^{-30}$ and I therefore applied a Greenhouse-Geisser correction.

I found a significant effect of training progress on neural responses ($F_{(3.40,428)}$=5.71, $p = 3.62\times10^{-5}$). I found no effect of stimulus presentation (i.e. was a trial occurring?) ($F_{(3.40,428)}$=1.04, $p = 0.38$). This seems unexpected, but one should bear in mind that traces from 'Run-trials' are collected relative to the running onset and not the stimulus onset itself. Similarly, there was no significant effect of running on neural responses ($F_{(3.40,428)}$=1.65, $p = 0.17$). Similar to licking behaviour, I found an interaction between locomotion and stimulus ($F_{(3.40,428)}$=3.07, $p = 0.022$).

**In mice trained on the tactile task.**

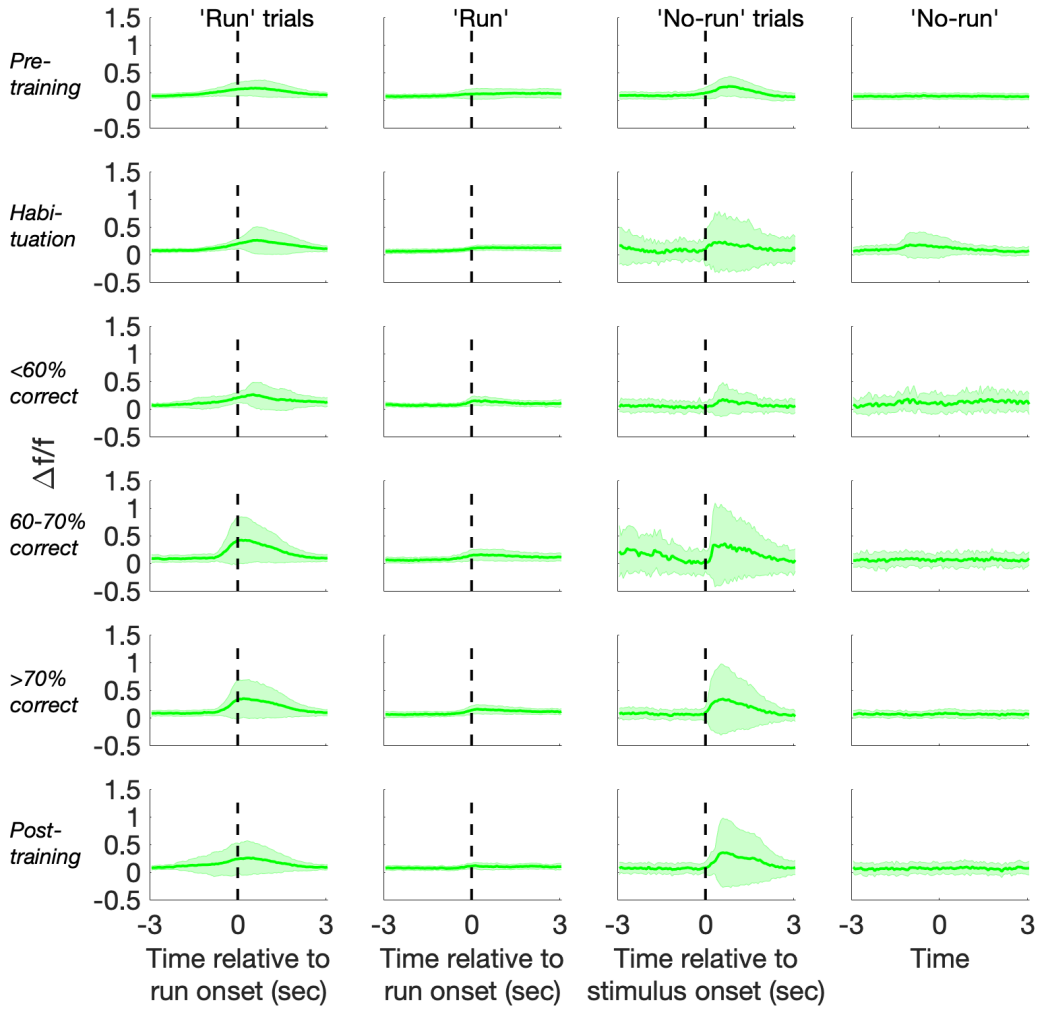



**Figure 3.29:** Responses of sharp onset neurons in mice trained on the tactile task in 'Run trials', around locomotion (i.e. 'Run' behaviour but no stimulus), 'No-run trials' and 'No-run' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

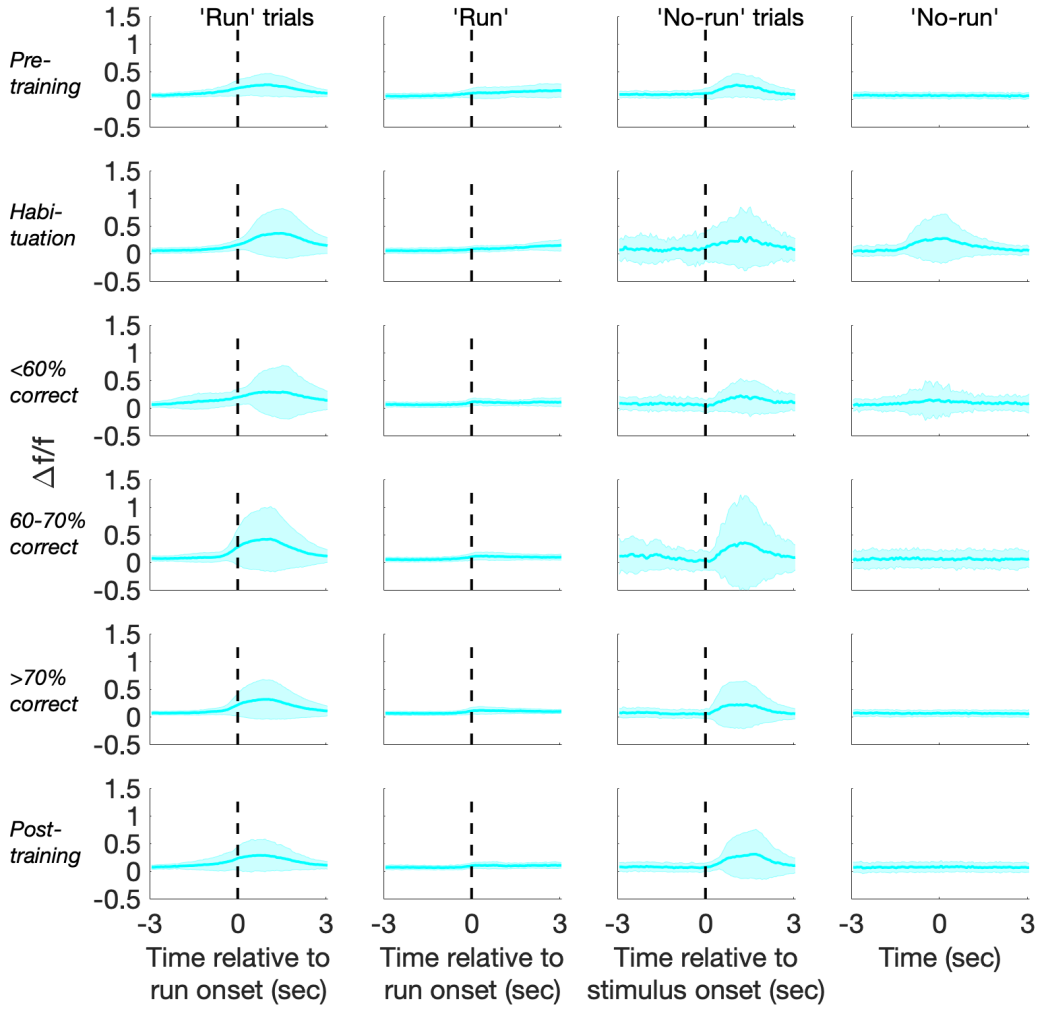



**Figure 3.30:** Responses of amplitude-locked neurons in mice trained on the tactile task in 'Run trials', around locomotion (i.e. 'Run' behaviour but no stimulus), 'No-run trials' and 'No-run' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

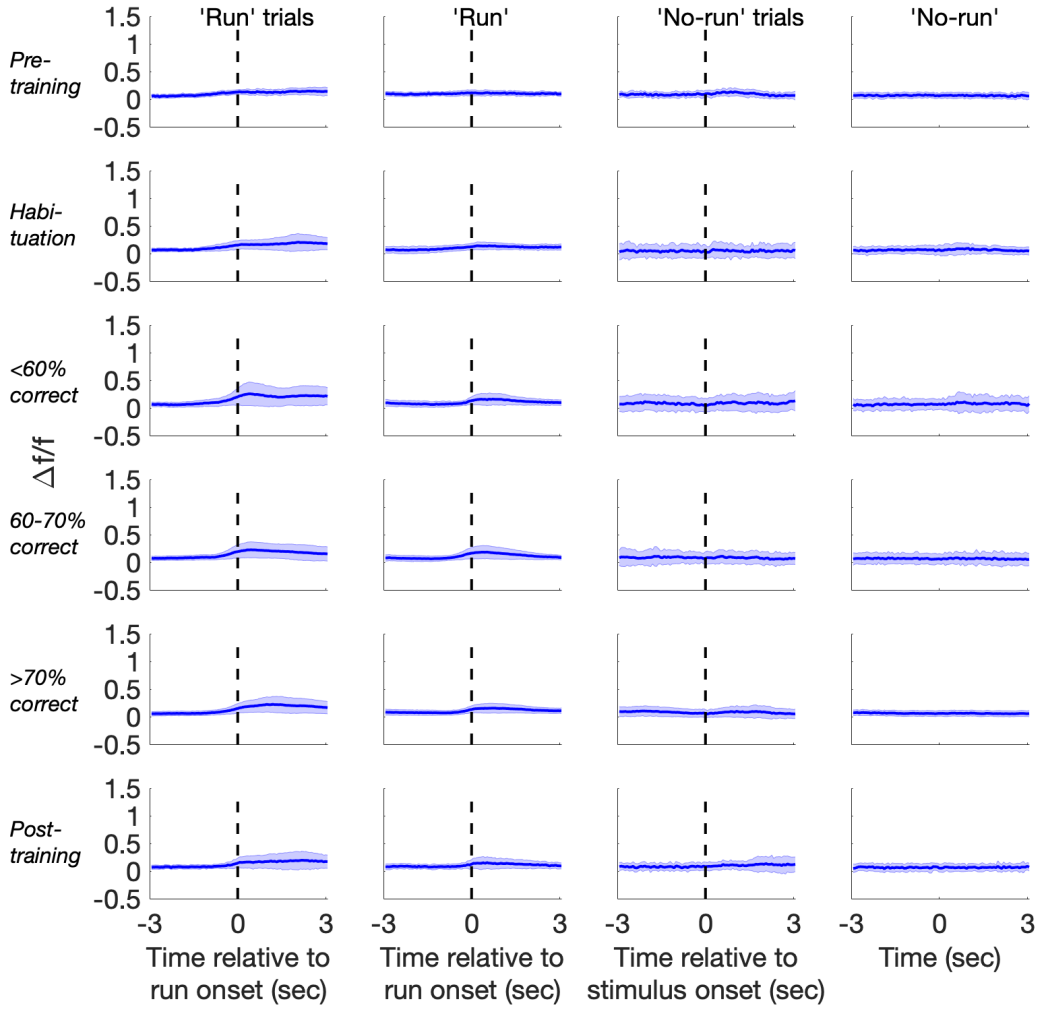



**Figure 3.31:** Responses of delayed onset neurons in mice trained on the tactile task in 'Run trials', around locomotion (i.e. 'Run' behaviour but no stimulus), 'No-run trials' and 'No-run' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

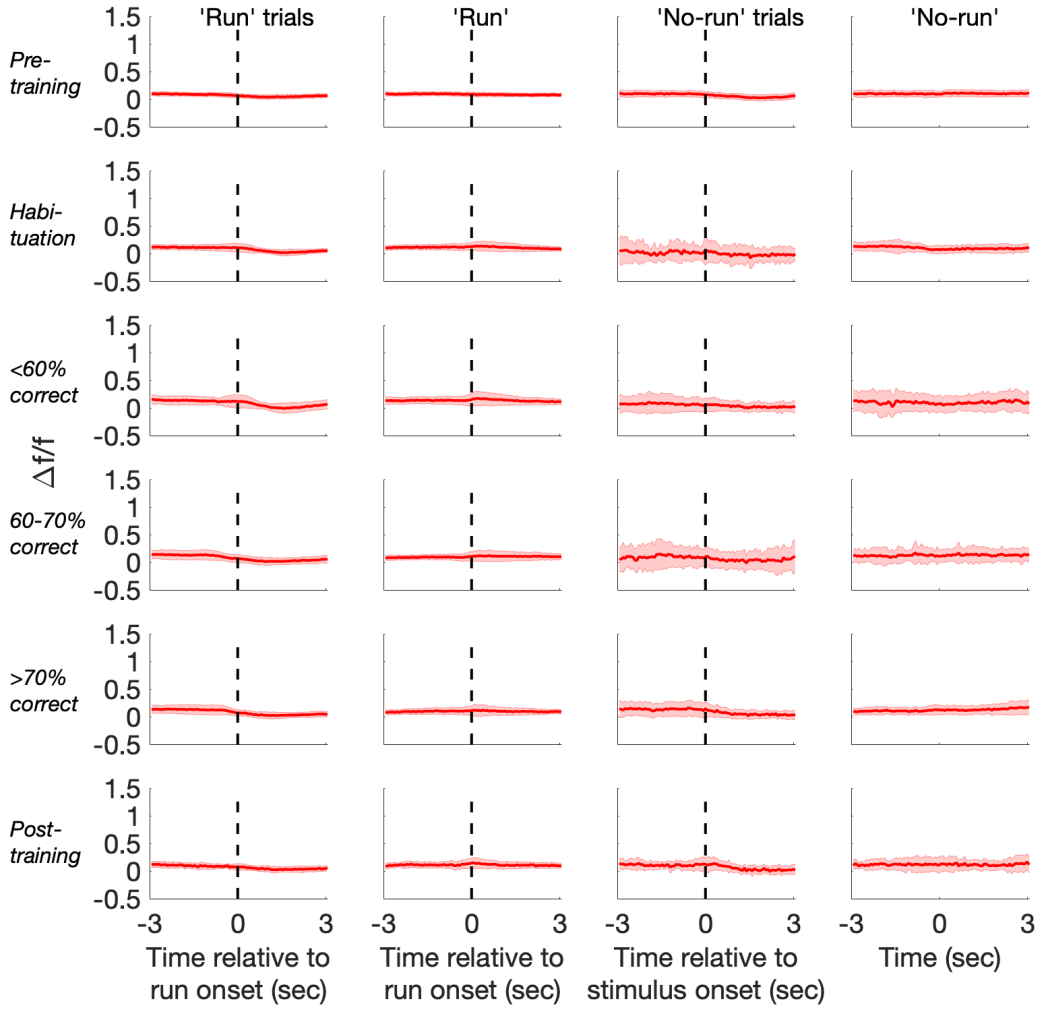



**Figure 3.32:** Responses of suppressed neurons in mice trained on the tactile task in 'Run trials', around locomotion (i.e. 'Run' behaviour but no stimulus), 'No-run trials' and 'No-run' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

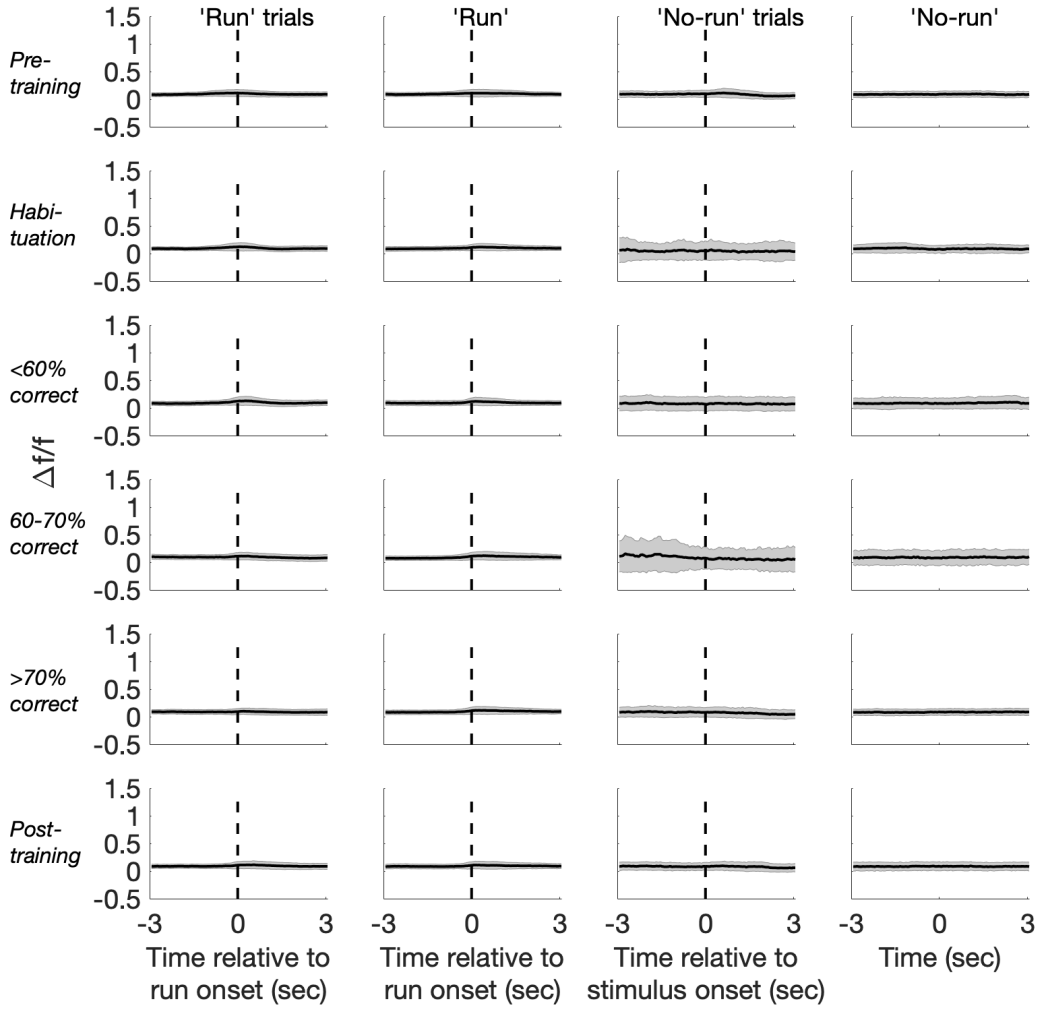



**Figure 3.33:** Responses of non-responsive neurons in mice trained on the tactile task in 'Run trials', around locomotion (i.e. 'Run' behaviour but no stimulus), 'No-run trials' and 'No-run' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

**In mice trained on the multisensory task.**

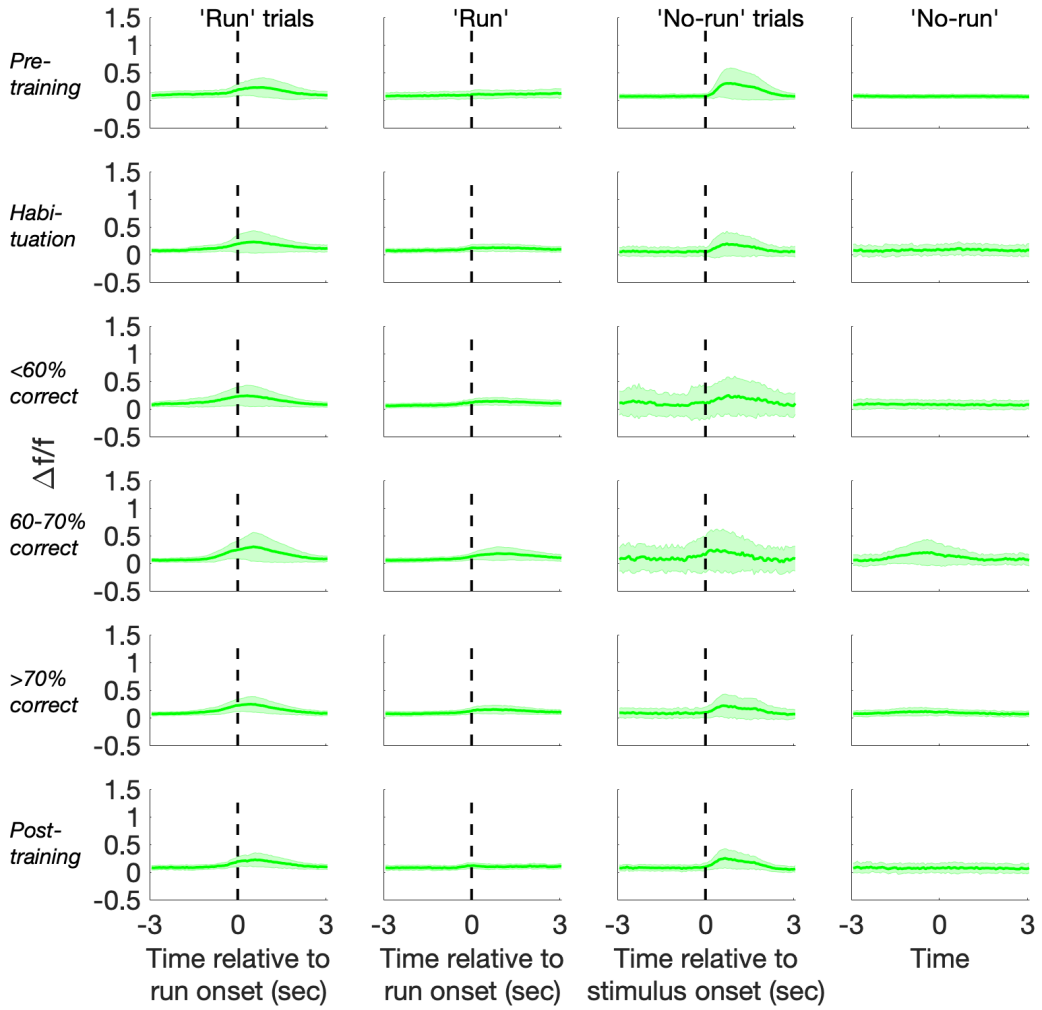



**Figure 3.34:** Responses of sharp onset neurons in mice trained on the multisensory task in 'Run trials', around locomotion (i.e. 'Run' behaviour but no stimulus), 'No-run trials' and 'No-run' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

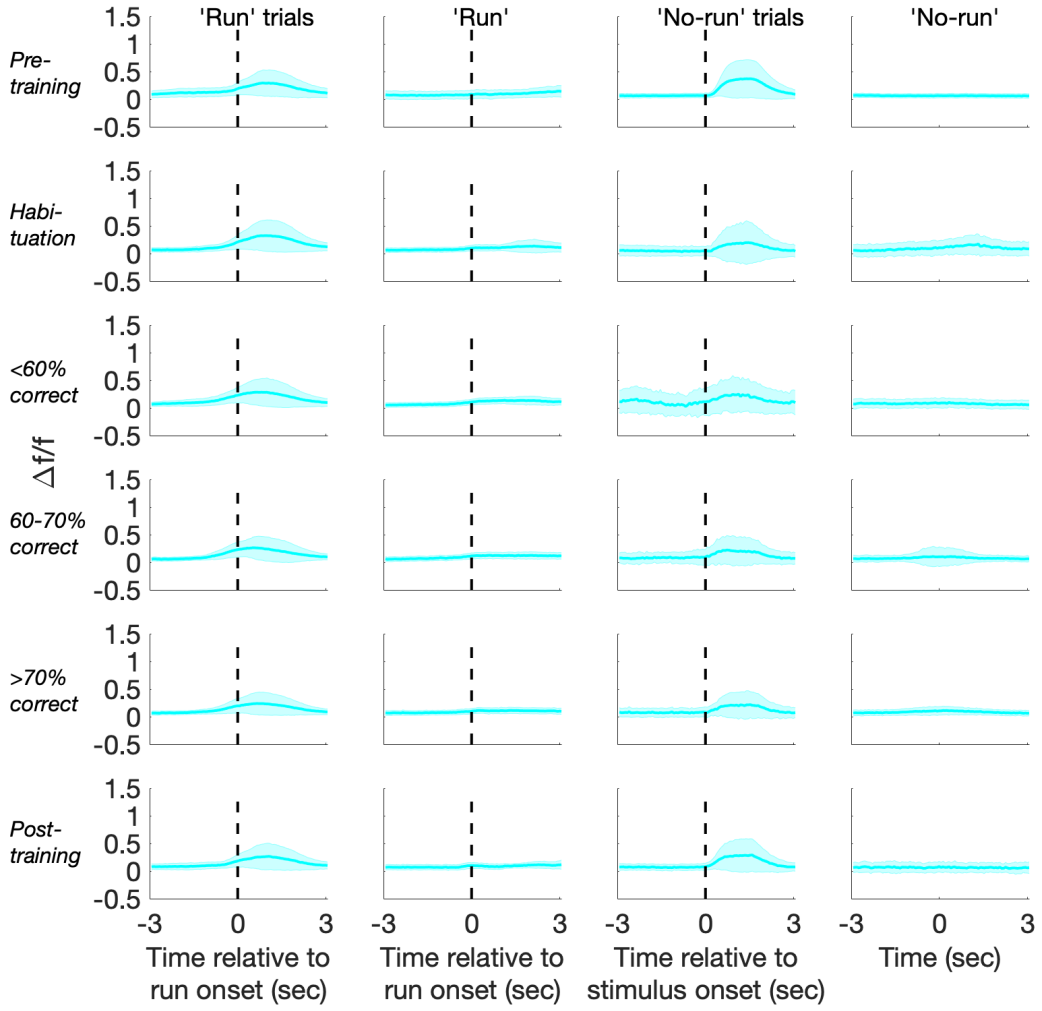



**Figure 3.35:** Responses of amplitude-locked neurons in mice trained on the multisensory task in 'Run trials', around locomotion (i.e. 'Run' behaviour but no stimulus), 'No-run trials' and 'No-run' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

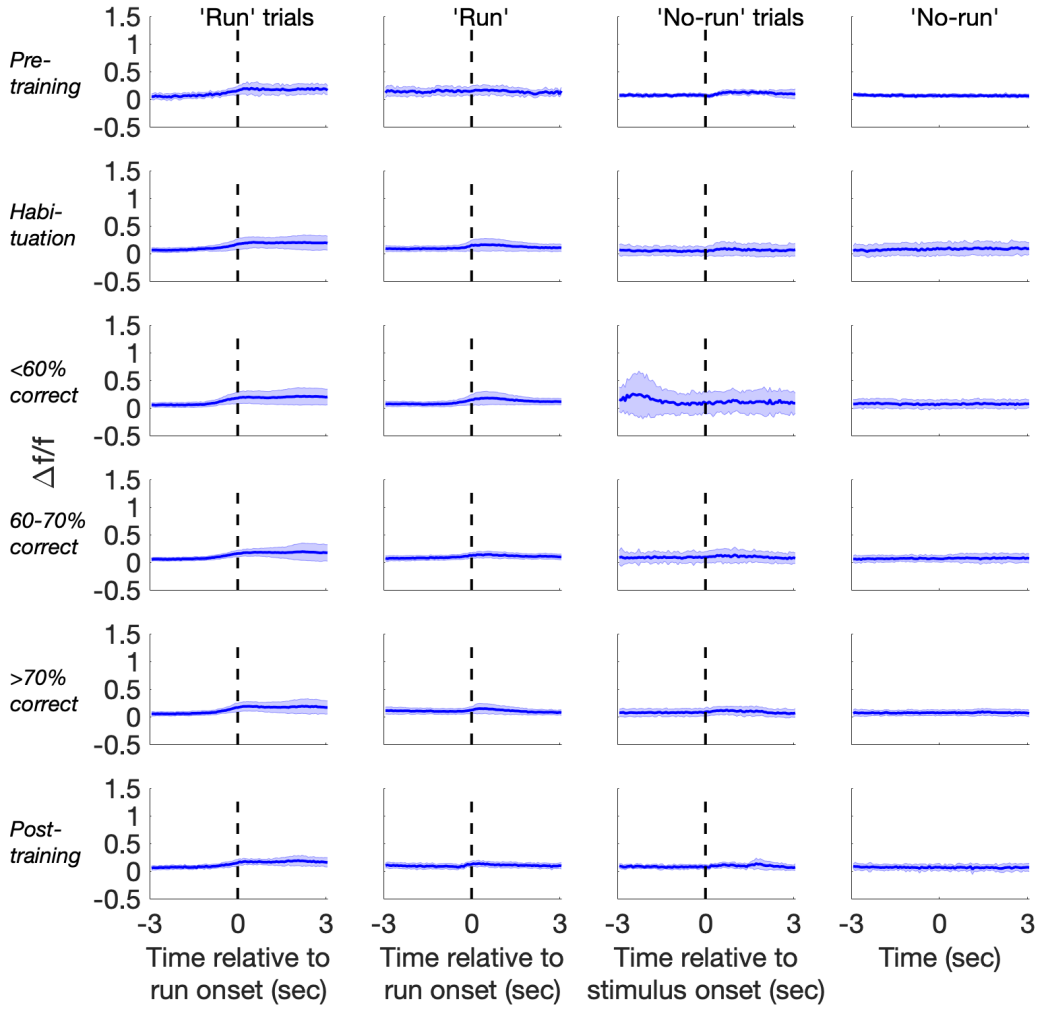



**Figure 3.36:** Responses of delayed onset neurons in mice trained on the multisensory task in 'Run trials', around locomotion (i.e. 'Run' behaviour but no stimulus), 'No-run trials' and 'No-run' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

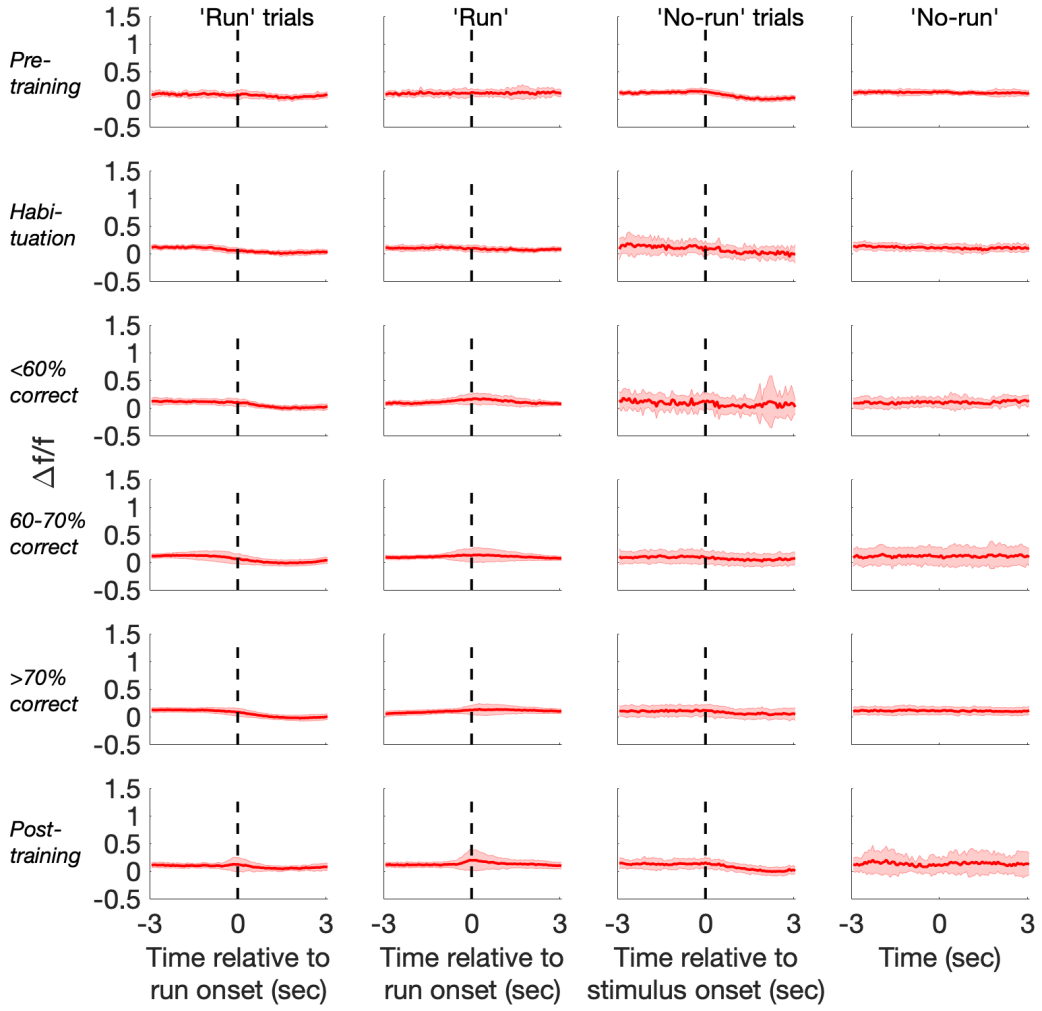



**Figure 3.37:** Responses of suppressed neurons in mice trained on the multisensory task in 'Run trials', around locomotion (i.e. 'Run' behaviour but no stimulus), 'No-run trials' and 'No-run' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

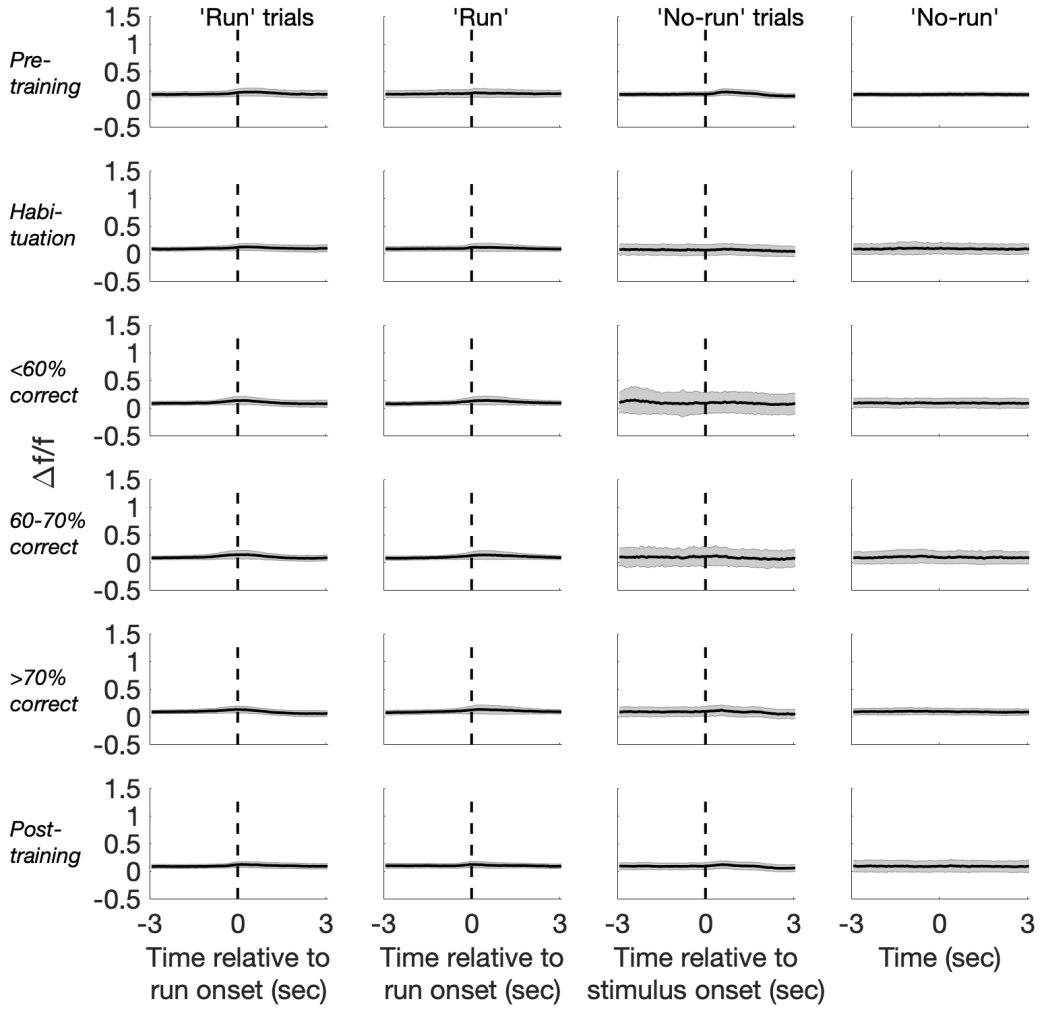



**Figure 3.38:** Responses of non-responsive neurons in mice trained on the multisensory task in 'Run trials', around locomotion (i.e. 'Run' behaviour but no stimulus), 'No-run trials' and 'No-run' behaviour outside of trials. Data shown is mean ± std, $N = 3$ mice.

To sum up, neurons were not significantly responsive to locomotion, reinforcing the point made in Section **3.3.8** that changes in neural responses due to training certainly cannot be fully accounted for by behavioural or motor artifacts. Similarly to licking, I found a significant interaction between running and stimulus presentation, suggesting a modulatory effect of locomotion on sensory responses in vS1.

## 3.4 Discussion

### 3.4.1 Summary of findings

In this Chapter, I described the responses of neurons in vS1 to vibrotactile, auditory and audio-vibrotactile stimuli. I found that in naïve animals, around half of all neurons imaged in vS1 responded to vibrotactile and multisensory stimuli and a tenth responded to auditory stimuli. After tactile training, I observed an increase in the number of neurons responding to tactile stimuli. In contrast, I found a reduction in the number of neurons responding to audio-vibrotactile stimuli, as well as a reduction in the specificity of the responses (i.e. a majority of neurons responded to several audio-vibrotactile stimuli).

In parallel, the timing of neural responses to tactile stimuli within a trial changed throughout training, with early responses to tactile stimuli prior to learning and more neurons showing either amplitude-locked or delayed-responses as mice improved on the task (up to 40% of neurons on the last day of training). In contrast, in mice trained on the multisensory task, a relative majority of neurons responded earlier to the audio-vibrotactile stimulus, exhibiting either sharp onset or amplitude-locked responses although the proportions of both of these response profiles decreased with training (relative majority of 40% of all imaged neurons prior to training and 20% after).

When decoding stimulus identity from neural data using a linear SVM, I found a trend toward an increase in decoding performance in mice trained on the tactile or multisensory task, although the variability in decoding performance was high across individual mice. Nonetheless, the classifier performed similarly in decoding the auditory or tactile components as well as the combined pairs of multisensory

stimuli, suggesting that multisensory information is present as early as vS1 and that its representation is affected by training. Moreover, a majority of tracked neurons responded to stimuli inconsistently across training, being non-responsive to stimuli on some days and responsive on others. This suggests that training may lead to the recruitment of different subpopulations of neurons within vS1.

Finally, I measured responses of neurons during and between stimulus presentations in order to assess whether they may be responsive to other behavioural variables related to the task, including locomotion and licking. While no imaged neurons were responsive to locomotion or licking alone, these behaviours still had a significant effect on neural responses to stimuli. In short, I found that non-sensory behavioural factors may modulate responses to stimuli in vS1.

### 3.4.2 Characterising responsiveness to stimuli

I first aimed to characterise neural responses to stimuli presented in passive, untrained conditions. First, I found that roughly half of the imaged neurons were responsive to vibrotactile stimuli (Figure **3.4**). This finding is perhaps on the higher end of the proportions reported by Peron and colleagues [Peron et al., 2015] but is consistent with other studies [Chen et al., 2015, Ayaz et al., 2019, O'Connor et al., 2010a]. This population of responsive neurons was approximately equally split between selective cells, responding to one stimulus only, and non-selective cells that responded to both presented frequencies.

Interestingly, I found a small but significant proportion of neurons responsive to auditory stimuli. As shown in Figure **3.3**, these responses were typically characterised by low-amplitude fluorescence signals, but they were reliably observed

in around 10% of all imaged neurons. In contrast to tactile responses, a majority of sound-responsive neurons responded selectively to only one of the two auditory frequencies presented, although there were no marked differences between the proportions of cells preferring the low and high auditory frequencies. Although there are reports of auditory stimuli modulating responses in S1 [Zhang et al., 2020, Zhou and Fuster, 2004] robust evidence for auditory-evoked spiking in S1 is scarce [Maruyama and Komai, 2018] and a population-wide analysis of the relative proportions of cross-modal spiking activity has been lacking. The proportions reported in this thesis are consistent with reports of cross-modal responses in S1 [Takagaki et al., 2008], primary auditory [Bizley et al., 2007] and primary visual cortex [Meijer et al., 2017].

I report for the first time detailed, population-wide analyses of auditory responses in vS1. In their investigation of multimodal responses in S1, Zhang and colleagues did not observe neurons responsive to auditory stimuli presented alone [Zhang et al., 2020] but suggested that auditory inputs modulate responses to tactile stimuli. This is in agreement with Maruyama and Komai [Maruyama and Komai, 2018] who found that auditory stimuli produce subthreshold modulation in S1. Why might I have found auditory-responsive neurons which were not reported previously? While the anaesthetised preparation of Maruyama and Komai likely explains the difference, the methods used by Zhang and colleagues were similar to my experiments. For instance, they used mice bred on a C57BL6/J background under 22 weeks of age, expressing the indicator GCaMP6s. It is also worth noting that the proportion of tactile-responsive neurons that they reported (16%) was substantially lower than what I found (45.5%). This discrepancy in findings may in part arise from different definitions of neural responsiveness to stimuli.

Throughout this chapter, I have used a previously defined threshold for a neural response [Rothschild et al., 2010, Panniello et al., 2018], determined by the result of a paired $t$-test comparing trial-by-trial $\Delta f/f$ values in pre- and post-stimulus onset windows. A neuron was considered responsive if passed this $t$-test for at least one of the stimuli I presented. In contrast, Zhang and colleagues [Zhang et al., 2020] combined three measures to define a responsiveness metric, characterised by a one-way ANOVA to determine an effect of stimulus frequency on neural response, the area under tuning curves obtained from a wide range of presented stimuli, and the ratio of activity during and in between stimuli. Their definition is more stringent than mine, and therefore may lead to greater type II error, consistent with the lower number of tactile-responsive neurons they report. This may have caused them to overlook weak responses to sounds. As will be discussed later (Section **3.4.3**), even neurons considered 'non-responsive' in my analyses could be used to classify stimulus identity (Figures **3.12** and **3.13**). This suggests that they were in fact not 'non-responsive' after all, but perhaps that they did not respond reliably enough to pass the statistical threshold of my $t$-test . This highlights the limitations of classifying neurons as 'responsive' using a cut-off criterion.

I found a similar proportion of neurons responsive to multisensory and unisensory stimuli. These multisensory neurons may be the same neurons responding to tactile stimuli and may, in effect, respond to the tactile component of multisensory stimuli. However, while a majority of multisensory neurons responded non-specifically to two multisensory pairs or more (61.9%) (Figure **3.4**), a considerable proportion responded preferentially to only one multisensory combination (38%). Therefore, despite being presented with at least two stimulus pairs containing the same tactile component, their responses were dependent on the auditory component of the pair.

This finding suggests that, beyond identifying a subpopulation of neurons responsive to auditory stimuli alone, I have also found neurons for which responses to tactile stimuli are at the very least modulated by concurrent auditory stimuli. While it may seem trivial to test if the same neurons in my data-set were responsive to both tactile and multisensory stimuli, this is not straightforward. I imaged neurons in passive conditions in separate blocks for each modality. While it may be possible to track most neurons across the imaging session, this passive session was typically three times as long as a training session on the same day. Differences in attention or habituation to sensory stimulation could affect responses. Additionally, while the order of auditory, vibrotactile and audio-vibrotactile experimental blocks were randomised between mice, a single cell may also exhibit signs of habituation over the course of the experimental blocks being separated by up to many weeks. Changes in the animal's attention over the course of these experimental blocks may also impact neuronal activity. Therefore, future studies may consider presenting interleaved unisensory and multisensory stimuli within a single imaging session instead of grouping them into experimental blocks in order to compare responses to each modality for every imaged neuron and, in the process, avoid confounds from habituation.

Therefore, my findings are consistent with those reported in the ferret A1 [Bizley et al., 2007] which identified visually-responsive and visually-modulated neurons in a predominantly auditory region (i.e. the auditory cortex). They further suggest a modulatory effect of sounds on tactile information processing in vS1, consistent with studies in vS1 [Zhang et al., 2020, Maruyama and Komai, 2018] and V1 [Ibrahim et al., 2016]. Ibrahim and colleagues found that concurrent auditory stimuli sharpened the tuning of V1 neuron responses to visual stimuli. In this present study, I only presented two different auditory and tactile stimuli, which makes it

difficult to assess possible changes to tuning curves. Therefore, a follow-up study could present a wider range of tactile and auditory stimuli in order to characterise neural tuning to multisensory stimuli in vS1.

### 3.4.3 Effects of training on responses

A major aim of this work was to assess how training on a unisensory or multisensory discrimination task may alter vS1 responses to stimuli. I found that while the overall proportion of responsive neurons remained constant throughout training, there were more selective than non-selective neurons responding to tactile stimuli at later training stages (Figure **3.5**). Therefore, neurons responded to tactile stimuli with increased specificity throughout training. In contrast, there were more neurons responsive to auditory stimuli following training.

Focusing specifically on mice trained on the same task, I found that the proportion of neurons responsive to tactile stimuli in mice trained on the tactile task did not vary between pre- and post-training passive sessions. In comparison, the proportion of neurons responsive to multisensory stimuli in mice trained on the multisensory task significantly dropped. Across each stage of passive imaging and imaging during training, I identified excited neurons with specific patterns of responsiveness (Figure **3.8**). In mice trained on the tactile task, I found an increase in the proportions of amplitude-locked and delayed onset neurons. In contrast, in mice trained on the multisensory task, I observed a decrease in the proportions of amplitude-locked, sharp onset and delayed onset neurons. I only found differences between pre- and post-training sessions proportions in mice trained on the multisensory task, in which the proportion of sharp onset and amplitude-locked neurons significantly dropped.

A key feature of the changes occurring during training is the fact that these changes were strongest at later stages of training. I found an increase in the proportion of responsive neurons, largest in all mice when they reached the performance criterion. Such changes are consistent with those observed in mice trained on a texture discrimination task [Chen et al., 2015]. Similarly, the increases in amplitude-locked and delayed onset neurons in mice trained on the tactile task were most pronounced at later stages of training. These changes could in fact be related to task performance. One interpretation could be that neural responses are more strongly affected by task-related sensorimotor artifacts such as locomotion or licking. My findings suggest that, while they may have a modulatory effect, these artifacts alone are not enough to explain changes in neural responses. This will be further discussed in Section **3.4.4**. Further, I found no effect of stimulus frequency on these patterns, meaning that the changes in proportions were similar for responses to 'Go' and 'No-Go' stimuli, even at expert stages of training. On the other hand, these neural response changes could reflect processes associated with learning, requiring a level of task engagement that would not be found in post-training passive conditions during which mice were not required to respond to stimuli and were not water-regulated. As such, they may not persist past operant behaviour and return to pre-training levels in post-training imaging sessions, but may underlie population-wide changes in vS1 during task acquisition. One way to verify this theory would consist in presenting mice with these stimuli in passive conditions, without the need for operant behaviour, for several consecutive days and assess whether similar patterns of response changes are observed or whether they are directly dependent on task engagement.

It is interesting here to note that these changes in mice trained on the tactile task were not consistently observed in mice trained on the audio-vibrotactile

task. Indeed, I report instead a drop in the proportion of responsive neurons between pre-training and post-training sessions. There were more sharp onset neurons in mice trained on the multisensory task than those trained on the tactile task. Why might there be a difference between mice trained on the unisensory and multisensory tasks? Based on evidence acquired in passive conditions in this project and in previous literature [Zhang et al., 2020, Ibrahim et al., 2016, Maruyama and Komai, 2018], these findings could be related to the modulatory effect of added auditory stimuli, found to translate as either increased inhibition [Zhang et al., 2020] or subthreshold postsynaptic responses which may change the probability of a single neuron's response to a stimulus [Maruyama and Komai, 2018]. With regards to the differences in proportions of sharp onset, amplitude-locked and delayed onset neurons between mice trained on the tactile and multisensory tasks, it remains unclear why these differences exist.

A potential key to the differences in results across the tactile and multisensory training could reside in the different subpopulations of neurons that remain responsive during training. In the tactile task, the neurons that were responsive in pre-training sessions could remain responsive over the course of training, with some additional neurons becoming responsive during training. In contrast, the proportion of neurons responsive to multisensory stimuli barely increased throughout training before significantly dropping at the end of training. This could suggest that exposure to multisensory stimuli recruits a different subset of neurons, which are modulated by auditory stimuli, replacing those responsive solely to the tactile components. However, when focusing primarily on those neurons that could be tracked across the whole course of training, I did not find an effect of task type on the proportion of neurons that remained reliably responsive, or that developed or lost

responsiveness during training (Figure **3.16**). In order to assess which subpopulation of neurons respond to stimuli of different modalities, follow-up studies could interleave non-informative trials from other modalities. This would allow the experimenter to track neural responses to different modalities while training on a specific version of the task. Further, it would provide an added opportunity to control for the excessive licking reported in the previous Chapter (Section **2.3.8**).

In fact, I found across all mice that a majority of neurons showed changes in their responses, either during training or in post-training imaging (i.e. responsive pre-training and non-responsive after training, or vice-versa). Across mice, more neurons responded during training but not in passive sessions than the other way round, consistent with the observed transient increases in the proportions of responsive neurons as mice trained, which could be explained by either an effect of training or task engagement (Figures **3.9**, **3.10** and **3.11**). Since similar proportions of neurons were found to be only responsive either pre- or post-training, it is unclear whether these are different subpopulations that changed during training or the same neurons. This being said, my analysis of tracked neurons confirmed that some neurons changed responsiveness profiles and therefore that a different subset of cells may have been recruited due to training. Interleaving non-informative trials during a task, as mentioned above, may help both assess effects of trial engagement on responses as well as potential changes in responses to different modalities while training.

This apparent instability in neural responses throughout training is somewhat in disagreement with Peron and colleagues' finding of increased stability of tactile information representation [Peron et al., 2015] in vS1 of mice during training on a pole position discrimination task. An interesting point made by Peron and colleagues can be found in their report of constantly high neurometric stimulus identity decoding,

wherein the neural response of as little as ten neurons to a particular pole position could be reliably inferred from whisker angle and curvature at any stage of training. This reports fits with previous ones of robust responses to whisker positions and object contact throughout training [Crochet and Petersen, 2006, O'Connor et al., 2010a]. This reported robustness matches observed reliable spiking responses during texture perception, characterised by typical 'stick-slip' motion of whiskers across surfaces, wherein vS1 neurons respond to sudden changes in whisker speed of motion [Jadhav and Feldman, 2010, Isett et al., 2018]. Here, I found that the decoding performance of a linear SVM varied widely across animals as well as with the training stage. I did not categorise a neuron's response to whisker curvature and position in my study, but I instead decoded vibrotactile frequency based on neural data in freely whisking mice. In fact, these findings may easily be reconciled. It is apparent from previous studies that vS1 neurons respond reliably to specific whisker motions, positions or curvatures [Jadhav and Feldman, 2010], especially toward the center of barrels [Clancy et al., 2015]. I report findings in mice with intact whisker pads, which differs from studies in which all but one whiskers are threaded. Indeed, the Velcro pad fitted to the vibrotactile apparatus allowed for the non-selective deflection of all macrovibrissae and could be reliably positioned relative to the mice's heads on a day-to-day basis. This is particularly relevant to the overall responsiveness of a population, as Clancy and colleagues highlighted the 'salt-and-pepper' layout of responses in individual barrels of vS1. In contrast, I applied an SVM on neural data obtained from actively whisking mice, without whisker tracing data and, crucially, from neurons within vS1 but potentially located on the edge of barrels or within septa.

Finally, as mentioned previously (Section **3.4.2**), categorising neurons as

'responsive' or 'non-responsive' may not paint a clear picture of the processes underlying sensory responses in vS1. The SVM performed similarly on 'responsive' neurons and 'non-responsive' cells (Figure **3.13**). In fact, the report of there being sufficient information in 'non-responsive' cells for the SVM to perform as well as in 'responsive' ones reinforces the notion that there may be broader, population-wide effects of training at play rather than a small subset of neurons supporting the sensory response altogether.

### 3.4.4 Sensorimotor integration

The barrel cortex is an inherently sensorimotor area, which has even been found to drive whisker retraction in mice [Matyas et al., 2010]. Some neurons in vS1 have been reported to respond only to whisking, irrespective of whether mice made contact with an object through their whiskers [Crochet and Petersen, 2006]. Similarly, Peron and colleagues identified 'touch' and 'whisking' neurons [Peron et al., 2015], consistent with the interleaved distribution of neurons projecting to M1 and S2 [Sato and Svoboda, 2010], while a later study identified neurons whose responses were modulated by locomotion and whisking [Ayaz et al., 2019].

Here, I found that vS1 neurons did not respond to either licking (Section **3.3.8**) or locomotion (Section **3.3.9**) alone. Instead, I observed that these behaviours modulated responses to stimuli. The fact that these neurons did not respond independently to an animal's motor behaviour may be surprising, considering previous findings of neurons encoding whisking or running [Ayaz et al., 2019], or even driving whisking [Matyas et al., 2010]. However, it is worth noting here that I did not track mice whiskers. Whisking and locomotion are tightly controlled, to the extent that mice whisk whenever they run [Sofroniew et al., 2015], although the opposite (i.e.

when mice do not run they do not whisk) is not true. Mice can actively whisk even when not running [Mayrhofer et al., 2013]. Hence, a follow-up study should use a high-speed camera to track whisking during the task in order to assess whether, independent of running or licking, whisking may affect responses to stimuli, which previous literature suggests it would [Crochet and Petersen, 2006, Peron et al., 2015, Ayaz et al., 2019].

The notion that primary sensory cortices respond only to information from their own modality and nothing else is now far outdated. Beyond crossmodal responses, behavioural responses have been observed in other primary sensory cortices, in manners that can be specific to each cortical region. For instance, while locomotion may drive neural activity in V1 [Saleem et al., 2013] or modulate responses in vS1, it has been found to inhibit excitatory pyramidal cells in A1 [Schneider et al., 2014]. Further, spontaneous behaviour has been found to reliably predict hundreds of dimensions of 'spontaneous' neural activity as early as primary sensory cortices, for example V1 [Stringer et al., 2019], sometimes even stronger than sensory events themselves. These findings have been reported across the whole mouse neocortex, finding for example effects of pupil dilation (a common marker for arousal), which may in fact explain the trial-by-trial variability observed in neural responses to stimuli in behaving animals [Musall et al., 2019].

Therefore, while the non-sensory events I recorded in this Chapter did not elicit neural responses on their own, it is not all surprising that they modulated neural responses to sensory stimuli. However, we cannot reliably establish that there are no non-sensory responses in vS1, as I did not record all possible behaviours exhibited by the mice. Most notably, future studies could implement whisker and pupil-tracking, as they appear to be two major modulators of barrel cortex

activity. A growing number of studies are using these techniques to decouple sensory and motor responses [Clack et al., 2012, Peron et al., 2015, Ayaz et al., 2019, Musall et al., 2019, Stringer et al., 2019].

### 3.4.5 Conclusion

To summarise, throughout this Chapter I have reported and characterised responses of vS1 neurons to vibrotactile, auditory and multisensory stimuli. I further assessed how training on the vibrotactile or multisensory tasks may affect these responses. Specifically, I addressed the aims set out in this Chapter:

First, I compared how responsive neurons are to each stimulus type and notably found neurons responsive to auditory stimuli. Second, I confirmed the modulatory effect of auditory stimuli on responses to vibrotactile cues. Finally, I characterised the effects of training on neural responses. I found that training on the tactile and multisensory tasks both elicited changes in neural response profiles and temporary increases in the proportions of responsive neurons. In addition, I found that training on the multisensory task reduced the proportion of responsive neurons after training, a feature not observed in unisensory training.

Combined with the apparent variability in decoding performance of a linear SVM classifier across mice and training stages, and changes in responses to stimuli during and after training in tracked neurons, these findings suggest that training induces population-wide changes in vS1 activity. At each stage of training, different overlapping subsets of neurons gain or lose their responses to stimuli (both to auditory and tactile stimuli), supporting broad changes in vS1 responses to sensory inputs.

While I found high levels of variability between mice, it is not impossible

that these could be explained by non-measured behavioural events such as licking or locomotion, which I found to have a modulatory effect on responses. Further, other events such as whisking were not measured and are likely to further affect neural responses.

To conclude, the findings described in this Chapter are in agreement with previous descriptions of tactile responses as well as multisensory responses in other primary sensory cortices. For the first time, I describe the effects of multisensory training on responses in vS1. These findings support the need for extensive follow-up studies in order to further characterise the modulatory effects of auditory cues on neural responses to tactile stimuli in vS1 and, more generally, determine the role of neural activity not related to sensory input, such as locomotion and attention.

# Chapter 4

# Discussion

## 4.1 Summary

This Thesis aimed to examine the nature of putative multisensory responses in vS1 of mice, and how such responses may vary with experience. As I developed the task upon which to base the project, it became clear that commonly used sensory stimuli came with their shares of caveats. I therefore designed a novel vibrotactile stimulus and adapted it further to the needs of my study.

First, I ensured that the mice could not hear nor see the stimulus and that it was constrained within the tactile modality (Sections **2.2.10** and **2.3.4**). Second, I designed the stimulus to be congruent with the auditory stimuli used for auditory and audio-tactile training (Section **2.3.3**), a feature found to support enhanced discrimination ability and neural responses to multisensory stimuli [Spence and Squire, 2003, Kim et al., 2008, Atilgan et al., 2018, Bizley et al., 2016]. Finally, I ensured that mice could learn to discriminate between auditory, vibrotactile and audio-tactile stimuli reliably (Sections **2.3.4**, **2.3.5**, and **2.3.6**).

In doing so, I first found that mice trained on the multisensory task took longer to train than those trained on unisensory tasks (Section **2.3.7**), although a

follow-up study with age-matched mice trained on the multisensory and auditory tasks should confirm this finding. More importantly, I found patterns of performance improvement during individual sessions (Section **2.3.9**), consistent with a reduction in compulsive licking – a potential shortfall of Go/No-Go tasks discussed in Section **4.2** – and revealing the need for a new method to measure behavioural output from mice (Section **2.3.8**) based on lick-rates and adapted to individual mice and their progress within training sessions.

Having established that mice could train on the task, I then investigated vS1 neural responses to vibrotactile, auditory and audio-tactile stimuli. I found that roughly half of all neurons were responsive to tactile and multisensory stimuli and I report responses to auditory stimuli in vS1 (Section **3.3.1**). Unsurprisingly, fewer neurons responded to auditory stimuli than tactile cues, although these responses were significantly more specific.

Training on the tasks had an impact on the proportion of responsive neurons (Sections **3.3.1** and **3.3.3**). Specifically, training increased stimulus response selectivity to vibrotactile stimulus frequencies – although the fact that only two stimuli per modality were used means we cannot infer about the shape of neurons' tuning curves, as has been done for instance with orientation tuning in passive conditions in V1 of mice [Ibrahim et al., 2016].

While the proportion of neurons responsive to tactile stimuli increased during training before returning to pre-training levels, training reduced the proportion of neurons responsive to multisensory stimuli. In Section **3.4.3**, I suggested that different subnetworks of neurons were recruited during training on the multisensory task than on the tactile task, consistent with reports of experience-induced reinforcement of 'multisensory' local networks in V1 [Knöpfel et al., 2019]. This fits with the observed

changes in response profiles in responsive neurons during training (Section **3.3.5**). In mice trained on the tactile task, training recruited more responsive neurons, matching with an increase in delayed onset and amplitude-locked neurons. In mice trained on the multisensory task, the proportion of responsive neurons decreased although a majority of these neurons were either sharp onset or amplitude-locked neurons, responding to earlier than to tactile stimuli. The fact that a majority of individually tracked neurons exhibited changes in responsiveness over the course of training further supports this dynamic population remodelling claim (Section **3.3.7**). While this varied greatly between animals, the use of a SVM classifier revealed that information about the auditory component of multisensory stimuli was encoded as reliably in vS1 as that about the tactile information contained in the area (Section **3.3.6**).

Finally, while I was able to confirm that neurons were not responsive to locomotion or licking per se (Sections **3.3.8** and **3.3.9**), I still found a modulatory effect of these behavioural variables, suggesting that more complex processes occur as early as vS1 than pure somatosensory responses, consistent with the sensorimotor nature of the area and previous reports of neural correlates of behavioural states in primary sensory areas [Matyas et al., 2010, Ayaz et al., 2019, Stringer et al., 2019, Musall et al., 2019].

## 4.2 The choice of a behavioural paradigm

An important choice made early in the project regarded the distinction between the use of a two alternative forced choice (2-AFC) task, in which animals must respond appropriately to distinct stimuli, or a Go/No-Go task structure, in which animals must respond to a stimulus and refrain from responding to

others. The same task in which animals must discriminate between two pole positions against their whiskers can involve them learning to withhold licking upon presentation of a No-Go tone to avoid a negative outcome in a Go/No-Go version (see [Guo et al., 2014, Poort et al., 2015]), or lick one of two spouts to obtain a reward in a 2-AFC task (see [Mayrhofer et al., 2013, Miyashita and Feldman, 2013a, Busse et al., 2011, Siemann et al., 2015]).

There are several factors to consider in choosing between a Go/No-Go and a 2-AFC task. While both tasks can involve the same stimuli, they are unlikely to involve completely similar cognitive processes. Indeed, teaching an animal to repress its urge to drink is an altogether different task than teaching it to direct that impulse toward different locations. In fact, mice may at first be unable to repress licking in a Go/No-Go task and fail to elicit a response toward the end of an hour-long session [Berditchevskaia et al., 2016]. This is consistent with the findings I report on within-session changes in performance and licking patterns (Section **2.3.9**). This does not necessarily reflect an inability to distinguish between stimuli but rather different levels of thirst or motivation.

Another interpretation for the increase in Miss and CR trials toward the end of tasks could reside in the fact that, in a Go/No-Go task, an experimenter cannot determine whether a lack of lick is reflective of a deliberate withholding or simply a lapse in attention [Stüttgen et al., 2011]. This would further be supported by the observations of within-session variations made in Section **2.3.9**: lapses in attention or decreases in motivation could explain the decrease in Hit and FA rates toward the end of sessions. In fact, combining an overall drop in attention – equating to an equal decrease in Hit and FA rates – with a better control of inhibition in No-Go trials – corresponding to a decrease in thirst – may explain the faster drop in FA rate

than Hit rate and overall performance improvement during sessions. In comparison, in a 2-AFC task, a trial without a lick is simply a missed trial and does not reflect a potential rejection response. Both of these interpretations would suggest that a 2-AFC task may in fact be preferable to a Go/No-Go task for the headfixed imaging of neural responses to sensory stimuli during training.

There is however a crucial advantage in using a Go/No-Go task over a 2-AFC task. In the multisensory task, this structure would have required designing a 4-AFC task. While it is clear that mice can train on 2-AFC tasks [Busse et al., 2011, Mayrhofer et al., 2013, Siemann et al., 2015], whether they can train on a 4-AFC task is to my knowledge unverified. It might simply be the case that 4-AFC tasks are too difficult for mice to train on, especially tasks in which the behavioural output is a targeted lick, requiring four spouts in very close positions and generating numerous possible confounds. For example, requiring 4 spouts next to each other may cause many instances of 'cross-licking' between two or more spouts simply because of their proximity. In addition, using a 4-AFC task combining audio-vibrotactile stimuli would not be the same paradigm as a unisensory 2-AFC task, and performance comparison between unisensory and multisensory tasks may not prove as straightforward.

The use of the Go/No-Go paradigm allows for the presentation of more than two stimuli, which makes it particularly suited for the multisensory version of the task. Indeed, all three No-Go stimuli were pooled in the same 'No-Go' category and generated the same FA or CR outcomes, although knowing their identities and the individual unimodal components that made them up allowed for the subsequent characterisation of an animal's sensory bias on a given day, as described in Section **2.3.10**. More importantly, the overall task structure and required responses from

the mice remained identical to unisensory versions, allowing for the comparison of performance and behaviours between unisensory and multisensory versions of the task in Section **2.3.7**.

## 4.3 Defining a measure of responsiveness

As I mentioned in Section **3.4.2**, the definition of a 'responsive' neuron is not consistent across all studies [Rothschild et al., 2010, Zhang et al., 2020]. This may explain some discrepancies in the proportions of responsive neurons reported across vS1 literature [Peron et al., 2015, Chen et al., 2015, Zhang et al., 2020]. However, the changes observed during training (Sections **3.3.3** and **3.3.5**) were observed with the same responsiveness criterion and therefore do reliably reflect experience-induced changes.

Still, the argument can be made that responsiveness measures only provide partial information on the processes that support responses to sensory stimuli in vS1. This is supported by the finding that non-responsive neurons still informed on the nature of stimuli during training (Section **3.3.6**). Using a responsiveness threshold remains useful, for instance in investigating the sparseness of neural responses in vS1, and should not as such be simply discarded. Rather, this type of analysis should be combined with different methods that do not segregate between responsive or non-responsive cells. For example, Peron and colleagues tracked responses from individual cells and found that a single cell's responses to stimuli predicted the next day of training's responses more accurately [Peron et al., 2015]. In this example, such a finding may suggest a reinforcement and stabilisation of neural responses to tactile stimuli as a result of experience. While my findings may in fact suggest otherwise

that training and experience leads to local network remodelling, this is a striking example of an observation that does not require a 'responsiveness' threshold to reveal processes underlying perceptual training.

## 4.4 Limitations of two-photon imaging

Since there are also reports of thalamocortical and cortico-cortical inputs into L5 and L6 [Ibrahim et al., 2016, Bureau et al., 2006, Henschke et al., 2015], the challenge of depth of imaging decidedly spans the entire thickness of the neocortex. By using two-photon imaging, we were able to image up to L3, at roughly 320 μm below the cortical surface and, as such, probably missed out on potential findings deeper in L5 and L6.

Throughout this Thesis, I measured neural responses by using two-photon imaging. While it enables the simultaneous imaging of hundreds of neurons, one must bear in mind the inherent limitations associated to the biophysical properties of calcium indicators. In particular, indicators may have different rise and decay times, corresponding to the time taken for the fluorescence to increase and decrease back to baseline, and different sensitivities, corresponding to the levels of intracellular calcium at which half the indicator molecules are bound to calcium [Grienberger and Konnerth, 2012]. Here, I had the choice to use either the GCaMP6f indicator or GCaMP6s. While GCaMP6f has faster dynamics than GCaMP6s, it is also less sensitive and may therefore not be suited to reflecting smaller amplitude, subthreshold fluctuations [Chen et al., 2013b]. Here, I used GCaMP6s as it also has a higher signal-to-noise ratio, which allowed me to reliably image into deeper L3 (320 μm below cortical surface).

On the note of imaging depth, while the original theoretical imaging depth limit in two-photon microscopy was reported to be around 600 μm, reports of imaging depths beyond 1mm suggest that it may be possible to include more cortical layers within reach of two-photon imaging [Kobat et al., 2011]. These reports and the development of three-photon imaging [Ouzounov et al., 2017] suggest that imaging into deeper cortical layers is within reach of multi-photon imaging, although these developments are done at the expense of temporal resolution. Hence, it may be preferable to combine two-photon imaging and electrophysiological methods such as multiunit recording to help inform on the laminar structure of multisensory responses in vS1 by revealing the fine-scale temporal structure of responses to multisensory stimuli across the different layers within a column.

## 4.5 What networks support multisensory integration in vS1?

In this Thesis, I reported the observation of cross-modal and multisensory responses in L2/3 of vS1 of mice. I further highlighted changes in responsiveness during training that were specific to multisensory training and were not observed in mice trained on the tactile task (Sections **3.3.3** and **3.3.5**). These findings beg two major questions: 1) what are the inputs into vS1 that support these responses? and 2) how does experience shape local networks within L2/3 of vS1?

According to the canonical model of cortical laminar structure, most thalamic inputs arrive in L4, L5 and L1 of the neocortex [Douglas and Martin, 2004, Bureau et al., 2006]. Decades of research have illustrated the fact that multisensory integration occurs in subcortical areas

[Drager and Hubel, 1975, King and Palmer, 1985, Meredith and Stein, 1986]. In monkeys, some thalamic nuclei project to several primary sensory cortices [Cappe et al., 2009], suggesting that multisensory responses in cortical regions are at least in part due to feed-forward processes. As discussed in Chapter **1** anatomical tracing studies have also revealed the existence of robust direct cortico-cortical connections between A1, V1 and S1 in monkeys [Cappe and Barone, 2005, Cappe et al., 2009, Meredith and Lomber, 2017], ferrets and rodents [Henschke et al., 2015, Teichert and Bolz, 2018, Iurilli et al., 2012].

The layers in which I recorded neural responses do not appear to receive direct thalamic or cortico-cortical inputs. In fact, they receive inputs from other layers within vS1, namely L1, L4 and L5 [Bureau et al., 2006]. I therefore investigated the nature of multisensory responses in vS1 in neurons that did not directly receive inputs from other areas. In fact, Ibrahim and colleagues' study into the nature of audio-visual integration in V1 of mice suggests that some direct cortico-cortical inputs may terminate onto L1 inhibitory interneuron which themselves in turn project into L2/3 and modulate sensory responses, in effect sharpening tuning curves of V1 neurons to visual gratings [Ibrahim et al., 2016].

In this study, I did not investigate intra-vS1 networks that may support multisensory responses in L2/3 although the aforementioned evidence suggests that a combination of thalamocortical inputs and direct cortico-cortical inputs from A1 may be at play. A follow-up study should specifically focus on L1 interneurons, similarly to [Ibrahim et al., 2016], and assess the contribution of cortico-cortical inputs to multisensory responses in L2/3. Similarly, while recording from L4 with two-photon imaging may be challenging due to the depth of these neurons, it still remains feasible and would allow for the comparison of relative contributions of thalamocortical and

cortico-cortical inputs to multisensory integration in vS1.

With regards to the local networks supporting multisensory integration, my findings suggest that, regardless of whether it is unisensory or multisensory, training and experience may induce changes into the responsiveness of individual neurons (Sections **3.3.5** and **3.3.7**). I suggested in Section **3.4.3** that this may relate to the reorganisation of local networks of neurons and population-wide changes in responses to stimuli. In particular, the differences observed between multisensory and unisensory training-induced changes could relate to the reinforcement of different local networks. Recent evidence collected in V1 of mice [Knöpfel et al., 2019] seems to support such a claim. Indeed, Knöpfel and colleagues found increased signal correlations between neurons identified as multimodal, suggesting that specialised networks support multisensory integration in primary sensory cortices.

While assessing the relative contributions of L1 and L4 inputs into L2/3, an adequate follow-up study could aim to 1) measure correlations within L2/3 between neurons, and 2) assess whether separate local networks receive equal inputs from putative cortico-cortical L1 neurons and thalamocortical L4 neurons.

## 4.6 Is the primary somatosensory cortex only sensory?

Given the numerous findings of multisensory responses in primary sensory cortices [Bizley et al., 2007, Ibrahim et al., 2016, Meijer et al., 2017, Knöpfel et al., 2019], it is perhaps unsurprising that I found responses to auditory stimuli in vS1. These findings are consistent with a recurring challenge of the canonical model for cortical parcellation [Wallace et al., 2004], which has by now

showed that different streams of information processing are not as segregated in early cortical stages as could have been initially thought. Of course, this does not mean that different cortical areas do not have modality-specific functional specialisations as primary sensory cortices still predominantly respond to one modality over others. While some findings [Ibrahim et al., 2016, Zhang et al., 2020] may have hinted toward a modulatory role for multisensory integration, in which neurons from an area do not respond alone to another modality but rather present cross-modal modulated responses to their preferred modality, I still report on the presence of a non-negligible proportion of sound-responsive neurons.

However, while my findings only briefly hinted to responses to other non-sensory variables (Sections **3.3.8** and **3.3.9**), a large body of evidence has highlighted the existence of high-dimensional patterns of neural activity that can be explained or explain most observed behaviours [Musall et al., 2019, Stringer et al., 2019]. It makes little doubt that complex behaviours are to some extent represented in vS1.

What might be the purpose of such early cross-talk between sensory modalities and behaviour? A first observation that one should be aware of is the now accepted notion that subcortical multisensory integration is extensive. Why should multisensory integration occur below the cortex and in higher sensory cortices but not at all stages of sensory processing? Modulation of responses at every stage of sensory processing may in fact be more efficient and integrate information across the whole course of processing. The use of the word 'sensory' instead of 'perceptual' may be debated here, as I report effects of non-sensory behavioural variables on neural responses. For now, whether this modulation is bottom-up or top-down remains unclear [Cappe et al., 2009] and the reported effects of training can be interpreted as

having an effect on sensory responses and not necessarily reflect perceptual changes. What is clear though is the fact that vS1 is not a purely unisensory cortex, nor is it a purely sensory area.

## 4.7 General conclusion

In this Thesis, I have described a novel task suited to the study of audio-vibrotactile integration in mice and used it to investigate both the nature of responses to multisensory stimuli and the effects of training on these responses.

I have found vS1 neurons responsive to sound and reliable encoding of auditory information and multisensory information in the area. While my findings require further investigation, they suggest a dual response process to multisensory stimuli in which some neurons exhibit response modulation and others are plainly responsive to cross-modal stimuli.

I furthermore found an effect of training on these responses, once again different depending on whether the training was unisensory or multisensory, supporting a model for experience-induced local circuit remodelling in vS1 and more precisely the further suggestion of multisensory-specific remodelling. To my knowledge, these are the first reports of auditory responses in vS1 and the observed effects of multisensory training on sensory responses in vS1.

These findings undoubtedly warrant further examination of cross-modal and behavioural encoding in primary sensory cortices. Combining the high spatial sampling of multi-photon imaging and the more precise temporal recordings of electrophysiological recordings will help shed a new light on the relative contributions

of separate pathways to multisensory encoding and to the high dimension encoding of non-sensory behavioural features.

# Bibliography

[Adams, 1979] Adams, J. C. (1979). Ascending projections to the inferior colliculus. *Journal of Comparative Neurology*, 183(3):519–538.

[Adibi et al., 2012] Adibi, M., Diamond, M. E., and Arabzadeh, E. (2012). Behavioral study of whisker-mediated vibration sensation in rats. *Proceedings of the National Academy of Sciences*, 109(3):971–976.

[Alais and Burr, 2004] Alais, D. and Burr, D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current Biology*, 14(3):257–262.

[Andermann et al., 2010] Andermann, M. L., Kerlin, A. M., and Reid, R. C. (2010). Chronic cellular imaging of mouse visual cortex during operant behavior and passive viewing. *Frontiers in Cellullar Neuroscience*, 4(March):1–16.

[Angelaki et al., 2009] Angelaki, D. E., Gu, Y., and DeAngelis, G. C. (2009). Multisensory integration: psychophysics, neurophysiology, and computation. *Current Opinion in Neurobiology*, 19(4):452–458.

[Arabzadeh et al., 2006] Arabzadeh, E., Panzeri, S., and Diamond, M. E. (2006). Deciphering the spike train of a sensory neuron: Counts and temporal patterns in the rat whisker pathway. *Journal of Neuroscience*, 26(36):9216–9226.

[Arabzadeh et al., 2003] Arabzadeh, E., Petersen, R. S., and Diamond, M. E. (2003). Encoding of whisker vibration by rat barrel cortex neurons: implications for texture discrimination. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 23(27):9146–54.

[Arabzadeh et al., 2005] Arabzadeh, E., Zorzin, E., and Diamond, M. E. (2005). Neuronal encoding of texture in the whisker sensory pathway. *PLoS Biology*, 3(1).

[Aronoff et al., 2010] Aronoff, R., Matyas, F., Mateo, C., Ciron, C., Schneider, B., and Petersen, C. C. (2010). Long-range connectivity of mouse primary somatosensory barrel cortex. *European Journal of Neuroscience*, 31(12):2221–2233.

[Atilgan et al., 2018] Atilgan, H., Town, S. M., Wood, K. C., Jones, G. P., Maddox, R. K., Lee, A. K., and Bizley, J. K. (2018). Integration of Visual Information in Auditory Cortex Promotes Auditory Scene Analysis through Multisensory Binding. *Neuron*, 97(3):640–655.e4.

[Ayaz et al., 2019] Ayaz, A., Stäuble, A., Hamada, M., Wulf, M.-A., Saleem, A. B., and Helmchen, F. (2019). Layer-specific integration of locomotion and sensory information in mouse barrel cortex. *Nature Communications*, 10(1):2585.

[Barth and Poulet, 2012] Barth, A. L. and Poulet, J. F. A. (2012). Experimental evidence for sparse firing in the neocortex.

[Berditchevskaia et al., 2016] Berditchevskaia, A., Cazé, R. D., and Schultz, S. R. (2016). Performance in a GO/NOGO perceptual task reflects a balance between impulsive and instrumental components of behaviour. *Scientific Reports*, 6(June):1–15.

[Bieler et al., 2018] Bieler, M., Xu, X., Marquardt, A., and Hanganu-Opatz, I. L. (2018). Multisensory integration in rodent tactile but not visual thalamus. *Scientific Reports*, 8(1):1–18.

[Bizley et al., 2016] Bizley, J. K., Maddox, R. K., and Lee, A. K. (2016). Defining Auditory-Visual Objects: Behavioral Tests and Physiological Mechanisms. *Trends in Neurosciences*, 39(2):74–85.

[Bizley et al., 2007] Bizley, J. K., Nodal, F. R., Bajo, V. M., Nelken, I., and King, A. J. (2007). Physiological and Anatomical Evidence for Multisensory Interactions in Auditory Cortex. *Cerebral Cortex*, 17(9):2172–2189.

[Bizley et al., 2013] Bizley, J. K., Walker, K. M. M., King, A. J., and Schnupp, J. W. H. (2013). Spectral timbre perception in ferrets: Discrimination of artificial vowels under different listening conditions. *The Journal of the Acoustical Society of America*, 133(1):365–376.

[Brecht et al., 1997] Brecht, M., Preilowski, B., Merzenich, M. M., and Keck, W. M. (1997). Functional architecture of the mystacial vibrissae. *Behavioural Brain Research*, 84:81–97.

[Brecht and Sakmann, 2002] Brecht, M. and Sakmann, B. (2002). Whisker maps of neuronal subclasses of the rat ventral posterior medial thalamus, identified by whole-cell voltage recording and morphological reconstruction. *Journal of Physiology*, 538(2):495–515.

[Brecht et al., 2003] Brecht, M., Schneider, M., Sakmann, B., and Margrie, T. (2003). Movements evoked by intracellular stimulation of single pyramidal cells in layer 5 and 6 of rat motor cortex. *Nature*, 427(6976):704–710.

[Bregman and McAdams, 1994] Bregman, A. S. and McAdams, S. (1994). Auditory Scene Analysis: The Perceptual Organization of Sound. *The Journal of the Acoustical Society of America*, 95(2):1177–1178.

[Bureau et al., 2006] Bureau, I., Von Paul, F. S., and Svoboda, K. (2006). Interdigitated paralemniscal and lemniscal pathways in the mouse barrel cortex. *PLoS Biology*, 4(12):2361–2371.

[Busse et al., 2011] Busse, L., Ayaz, A., Dhruv, N. T., Katzner, S., Saleem, A. B., Schölvinck, M. L., Zaharia, A. D., and Carandini, M. (2011). The detection of visual contrast in the behaving mouse. *Journal of Neuroscience*, 31(31):11351–11361.

[Cappe and Barone, 2005] Cappe, C. and Barone, P. (2005). Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *European Journal of Neuroscience*, 22(11):2886–2902.

[Cappe et al., 2009] Cappe, C., Rouiller, E. M., and Barone, P. (2009). Multisensory anatomical pathways. *Hearing Research*, 258(1-2):28–36.

[Caves et al., 2018] Caves, E. M., Brandley, N. C., and Johnsen, S. (2018). Visual Acuity and the Evolution of Signals. *Trends in Ecology and Evolution*, 33(5):358–372.

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

[Chen et al., 2013a] Chen, J. L., Carta, S., Soldado-Magraner, J., Schneider, B. L., and Helmchen, F. (2013a). Behaviour-dependent recruitment of long-range projection neurons in somatosensory cortex. *Nature*, 499(7458):336–340.

[Chen et al., 2015] Chen, J. L., Margolis, D. J., Stankov, A., Sumanovski, L. T., Schneider, B. L., and Helmchen, F. (2015). Pathway-specific reorganization of projection neurons in somatosensory cortex during learning. *Nature Neuroscience*, 18(8):1101–1108.

[Chen et al., 2016] Chen, J. L., Voigt, F. F., Javadzadeh, M., Krueppel, R., and Helmchen, F. (2016). Long-range population dynamics of anatomically defined neocortical networks. *eLife*, 5(MAY2016):1–26.

[Chen et al., 2012] Chen, Q., Cichon, J., Wang, W., Qiu, L., Lee, S. J. R., Campbell, N. R., DeStefino, N., Goard, M. J., Fu, Z., Yasuda, R., Looger, L. L., Arenkiel, B. R., Gan, W. B., and Feng, G. (2012). Imaging Neural Activity Using Thy1-GCaMP Transgenic Mice. *Neuron*, 76(2):297–308.

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

[Chong et al., 2019] Chong, E. Z., Panniello, M., Barreiros, I., Kohl, M. M., and Booth, M. J. (2019). Quasi-simultaneous multiplane calcium imaging of neuronal circuits. *Biomedical Optics Express*, 10(1):267.

[Clack et al., 2012] Clack, N. G., O'Connor, D. H., Huber, D., Petreanu, L., Hires, A., Peron, S., Svoboda, K., and Myers, E. W. (2012). Automated tracking of whiskers in videos of head fixed rodents. *PLoS Computational Biology*, 8(7).

[Clancy et al., 2015] Clancy, K. B., Schnepel, P., Rao, A. T., and Feldman, D. E. (2015). Structure of a single whisker representation in layer 2 of mouse somatosensory cortex. *Journal of Neuroscience*, 35(9):3946–3958.

[Colavita, 1974] Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics*, 16(2):409–412.

[Crochet and Petersen, 2006] Crochet, S. and Petersen, C. C. (2006). Correlating whisker behavior with membrane potential in barrel cortex of awake mice. *Nature Neuroscience*, 9(5):608–610.

[De Hoz and Nelken, 2014] De Hoz, L. and Nelken, I. (2014). Frequency tuning in the behaving mouse: Different bandwidths for discrimination and generalization. *PLoS ONE*, 9(3).

[De Kock et al., 2007] De Kock, C. P., Bruno, R. M., Spors, H., and Sakmann, B. (2007). Layer- and cell-type-specific suprathreshold stimulus representation in rat primary somatosensory cortex. *Journal of Physiology*, 581(1):139–154.

[De Kock and Sakmann, 2009] De Kock, C. P. and Sakmann, B. (2009). Spiking in primary somatosensory cortex during natural whisking in awake head-restrained rats is cell-type specific. *Proceedings of the National Academy of Sciences of the United States of America*, 106(38):16446–16450.

[Deneux et al., 2019] Deneux, T., Harrell, E. R., Kempf, A., Ceballo, S., Filipchuk, A., and Bathellier, B. (2019). Context-dependent signaling of coincident auditory and visual events in primary visual cortex. *eLife*, 8:1–23.

[Denk and Svoboda, 1997] Denk, W. and Svoboda, K. (1997). Why multiphoton is more than a gimmick. *Neuron*, 18(Figure 1):351–357.

[Deroy et al., 2016] Deroy, O., Faivre, N., Lunghi, C., Spence, C., Aller, M., and Noppeney, U. (2016). The Complex Interplay between Multisensory Integration and Perceptual Awareness. *Multisensory Research*, 29(6):585–606.

[Diamond and Arabzadeh, 2013] Diamond, M. E. and Arabzadeh, E. (2013). Whisker sensory system - From receptor to decision. *Progress in Neurobiology*, 103:28–40.

[Diamond et al., 2008] Diamond, M. E., Von Heimendahl, M., Knutsen, P. M., Kleinfeld, D., and Ahissar, E. (2008). 'Where' and 'what' in the whisker sensorimotor system. *Nature Reviews Neuroscience*, 9(8):601–612.

[Douglas and Martin, 2004] Douglas, R. J. and Martin, K. A. (2004). Neuronal circuits of the neocortex. *Annual Review of Neuroscience*, 27:419–451.

[Drager and Hubel, 1975] Drager, U. C. and Hubel, D. H. (1975). Responses to visual stimulation and relationship between visual, auditory, and somatosensory inputs in mouse superior colliculus. *Journal of Neurophysiology*, 38(3):690–713.

[Ebara et al., 2002] Ebara, S., Kumamoto, K., Matsuura, T., Mazurkiewicz, J. E., and Rice, F. L. (2002). Similarities and differences in the innervation of mystacial vibrissal follicle-sinus complexes in the rat and cat: A confocal microscopic study. *Journal of Comparative Neurology*, 449(2):103–119.

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

[Ferezou et al., 2007] Ferezou, I., Haiss, F., Gentet, L. J., Aronoff, R., Weber, B., and Petersen, C. C. (2007). Spatiotemporal Dynamics of Cortical Sensorimotor Integration in Behaving Mice. *Neuron*, 56(5):907–923.

[Foxe, 2009] Foxe, J. J. (2009). Multisensory Integration: Frequency Tuning of Audio-Tactile Integration. *Current Biology*, 19(9):R373–R375.

[Foxe et al., 2000] Foxe, J. J., Morocz, I. A., Murray, M. M., Higgins, B. A., Javitt, D. C., and Schroeder, C. E. (2000). Multisensory auditory-somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Cognitive Brain Research*, 10(1-2):77–83.

[Foxe and Schroeder, 2005] Foxe, J. J. and Schroeder, C. E. (2005). The case for feedforward multisensory convergence during early cortical processing. *NeuroReport*, 16(5):419–423.

[Gibson, 1962] Gibson, J. J. (1962). Observations on Active Touch. *Psychological Review*, 69(6):477–491.

[Grienberger and Konnerth, 2012] Grienberger, C. and Konnerth, A. (2012). Imaging Calcium in Neurons. *Neuron*, 73(5):862–885.

[Grinevich et al., 2005] Grinevich, V., Brecht, M., and Osten, P. (2005). Monosynaptic pathway from rat vibrissa motor cortex to facial motor neurons revealed by lentivirus-based axonal tracing. *Journal of Neuroscience*, 25(36):8250–8258.

[Guest et al., 2002] Guest, S., Catmur, C., Lloyd, D., and Spence, C. (2002). Audiotactile interactions in roughness perception. *Experimental Brain Research*, 146(2):161–171.

[Guo et al., 2014] Guo, Z. V., Hires, S. A., Li, N., O'Connor, D. H., Komiyama, T., Ophir, E., Huber, D., Bonardi, C., Morandell, K., Gutnisky, D., Peron, S., Xu,

N.-l., Cox, J., and Svoboda, K. (2014). Procedures for Behavioral Experiments in Head-Fixed Mice. *PLoS ONE*, 9(2):e88678.

[Heffner and Heffner, 2007] Heffner, H. E. and Heffner, R. S. (2007). Hearing ranges of laboratory animals. *Journal of the American Association for Laboratory Animal Science : JAALAS*, 46(1):20–2.

[Helmchen et al., 2018] Helmchen, F., Gilad, A., and Chen, J. L. (2018). Neocortical dynamics during whisker-based sensory discrimination in head-restrained mice. *Neuroscience*, 368:57–69.

[Henschke et al., 2015] Henschke, J. U., Noesselt, T., Scheich, H., and Budinger, E. (2015). Possible anatomical pathways for short-latency multisensory integration processes in primary sensory cortices. *Brain Structure and Function*, 220(2):955–977.

[Hirsch and Curcio, 1989] Hirsch, J. and Curcio, C. A. (1989). The spatial resolution capacity of human foveal retina. *Vision Research*, 29(9):1095–1101.

[Houweling and Brecht, 2008] Houweling, A. R. and Brecht, M. (2008). Behavioural report of single neuron stimulation in somatosensory cortex. *Nature*, 451(7174):65–68.

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

[Isett et al., 2018] Isett, B. R., Feasel, S. H., Lane, M. A., and Feldman, D. E. (2018). Slip-Based Coding of Local Shape and Texture in Mouse S1. *Neuron*, 97(2):418–433.e5.

[Ison et al., 2007] Ison, J. R., Allen, P. D., and O'Neill, W. E. (2007). Age-related hearing loss in C57BL/6J mice has both frequency-specific and non-frequency-specific components that produce a hyperacusis-like exaggeration of the acoustic startle reflex. *JARO - Journal of the Association for Research in Otolaryngology*, 8(4):539–550.

[Iurilli et al., 2012] Iurilli, G., Ghezzi, D., Olcese, U., Lassi, G., Nazzaro, C., Tonini, R., Tucci, V., Benfenati, F., and Medini, P. (2012). Sound-Driven Synaptic Inhibition in Primary Visual Cortex. *Neuron*, 73(4):814–828.

[Jadhav and Feldman, 2010] Jadhav, S. P. and Feldman, D. E. (2010). Texture coding in the whisker system. *Current Opinion in Neurobiology*, 20(3):313–318.

[Jaramillo et al., 2014] Jaramillo, S., Zador, A. M., and Jaramillo, S. (2014). Mice and rats achieve similar levels of performance in an adaptive decision-making task. *Frontiers in Systems Neuroscience*, 8(September):1–11.

[Jousmäki and Hari, 1998] Jousmäki, V. and Hari, R. (1998). Parchment-skin illusion: sound-biased touch.

[Juavinett et al., 2018] Juavinett, A. L., Erlich, J. C., and Churchland, A. K. (2018). Decision-making behaviors: weighing ethology, complexity, and sensorimotor compatibility. *Current Opinion in Neurobiology*, 49:42–50.

[Juavinett et al., 2017] Juavinett, A. L., Nauhaus, I., Garrett, M. E., Zhuang, J., and Callaway, E. M. (2017). Automated identification of mouse visual areas with intrinsic signal imaging. *Nature Protocols*, 12(1):32–43.

[Kandler et al., 2018] Kandler, S., Mao, D., McNaughtun, B. L., and Bonin, V. (2018). Encoding of Tactile Context in the Mouse Visual Cortex. *bioRxiv*.

[Kane et al., 2012] Kane, K. L., Longo-Guess, C. M., Gagnon, L. H., Ding, D., Salvi, R. J., and Johnson, K. R. (2012). Genetic background effects on age-related hearing loss associated with Cdh23 variants in mice. *Hearing Research*, 283(1-2):80–88.

[Kanold et al., 2014] Kanold, P. O., Nelken, I., and Polley, D. B. (2014). Local versus global scales of organization in auditory cortex. *Trends in Neurosciences*, 37(9):502–510.

[Kerr et al., 2007] Kerr, J. N., De Kock, C. P., Greenberg, D. S., Bruno, R. M., Sakmann, B., and Helmchen, F. (2007). Spatial organization of neuronal population responses in layer 2/3 of rat barrel cortex. *Journal of Neuroscience*, 27(48):13316–13328.

[Kim et al., 2008] Kim, R. S., Seitz, A. R., and Shams, L. (2008). Benefits of stimulus congruency for multisensory facilitation of visual learning. *PLoS ONE*, 3(1).

[King and Palmer, 1985] King, A. J. and Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Experimental Brain Research*, 60(3):492–500.

[Knöpfel et al., 2019] Knöpfel, T., Sweeney, Y., Radulescu, C. I., Zabouri, N., Doostdar, N., Clopath, C., and Barnes, S. J. (2019). Audio-visual experience

strengthens multisensory assemblies in adult mouse visual cortex. *Nature Communications*, 10(1).

[Kobat et al., 2011] Kobat, D., Horton, N. G., and Xu, C. (2011). In vivo two-photon microscopy to 1.6-mm depth in mouse cortex. *Journal of Biomedical Optics*, 16(10):106014.

[Krupa et al., 2004] Krupa, D. J., Wiest, M. C., Shuler, M. G., Laubach, M., and Nicolelis, M. A. (2004). Layer-specific somatosensory cortical activation during active tactile discrimination. *Science*, 304(5679):1989–1992.

[Lakatos et al., 2007] Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A., and Schroeder, C. E. (2007). Neuronal Oscillations and Multisensory Interaction in Primary Auditory Cortex. *Neuron*, 53(2):279–292.

[Lee et al., 2008] Lee, S. H., Carvell, G. E., and Simons, D. J. (2008). Motor modulation of afferent somatosensory circuits. *Nature Neuroscience*, 11(12):1430–1438.

[Li et al., 2016] Li, N., Daie, K., Svoboda, K., and Druckmann, S. (2016). Robust neuronal dynamics in premotor cortex during motor planning. *Nature*, 532(7600):459–464.

[Liang et al., 2002] Liang, L., Lu, T., and Wang, X. (2002). Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates. *Journal of Neurophysiology*, 87(5):2237–2261.

[Lomber and Malhotra, 2008] Lomber, S. G. and Malhotra, S. (2008). Double dissociation of 'what' and 'where' processing in auditory cortex. *Nature Neuroscience*, 11(5):609–616.

[Long et al., 2015] Long, M., Jiang, W., Liu, D., and Yao, H. (2015). Contrast-dependent orientation discrimination in the mouse. *Scientific Reports*, 5:1–14.

[Maddox et al., 2015] Maddox, R. K., Atilgan, H., Bizley, J. K., and Lee, A. K. (2015). Auditory selective attention is enhanced by a task-irrelevant temporally coherent visual stimulus in human listeners. *eLife*, 4:1–11.

[Margolis et al., 2014] Margolis, D. J., Lütcke, H., Helmchen, F., Weber, B., and Haiss, F. (2014). *Chronic Two-Photon Imaging of Neural Activity in the Anesthetized and Awake Behaving Rodent*, pages 151–173. Humana Press, Totowa, NJ.

[Maruyama and Komai, 2018] Maruyama, A. T. and Komai, S. (2018). Auditory-induced response in the primary sensory cortex of rodents. *PLoS ONE*, 13(12):1–18.

[Matyas et al., 2010] Matyas, F., Sreenivasan, V., Marbach, F., Wacongne, C., Barsy, B., Mateo, C., Aronoff, R., and Petersen, C. C. H. (2010). Motor Control by Sensory Cortex. *Science*, 330(6008):1240–1243.

[Mayrhofer et al., 2015] Mayrhofer, J. M., Haiss, F., Helmchen, F., and Weber, B. (2015). Sparse, reliable, and long-term stable representation of periodic whisker deflections in the mouse barrel cortex. *NeuroImage*, 115:52–63.

[Mayrhofer et al., 2013] Mayrhofer, J. M., Skreb, V., Von der Behrens, W., Musall, S., Weber, B., and Haiss, F. (2013). Novel two-alternative forced choice paradigm for bilateral vibrotactile whisker frequency discrimination in head-fixed mice and rats. *Journal of Neurophysiology*, 109(1):273–284.

[McGurk and Macdonald, 1976] McGurk, H. and Macdonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264:691–811.

[Meijer et al., 2017] Meijer, G. T., Montijn, J. S., Pennartz, C. M., and Lansink, C. S. (2017). Audio-visual modulation in mouse V1 depends on cross-modal stimulus configuration and congruency. *The Journal of Neuroscience*, 37(36):0468–17.

[Meijer et al., 2018] Meijer, G. T., Pie, J. L., Dolman, T. L., Pennartz, C. M., and Lansink, C. S. (2018). Audiovisual integration enhances stimulus detection performance in mice. *Frontiers in Behavioral Neuroscience*, 12(October):1–12.

[Menzel and Barth, 2005] Menzel, R. R. and Barth, D. S. (2005). Multisensory and secondary somatosensory cortex in the rat. *Cerebral Cortex*, 15(11):1690–1696.

[Meredith and Lomber, 2017] Meredith, M. A. and Lomber, S. G. (2017). Species-dependent role of crossmodal connectivity among the primary sensory cortices. *Hearing Research*, 343(3):83–91.

[Meredith et al., 1987] Meredith, M. A., Nemitz, J. W., and Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *The Journal of neuroscience*, 7(10):3215–29.

[Meredith and Stein, 1986] Meredith, M. A. and Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, 56(3):640–662.

[Minderer et al., 2012] Minderer, M., Liu, W., Sumanovski, L. T., Kügler, S., Helmchen, F., and Margolis, D. J. (2012). Chronic imaging of cortical sensory map dynamics using a genetically encoded calcium indicator. *Journal of Physiology*, 590(1):99–107.

[Miyashita and Feldman, 2013a] Miyashita, T. and Feldman, D. E. (2013a). Behavioral detection of passive whisker stimuli requires somatosensory cortex. *Cerebral Cortex*, 23(7):1655–1662.

[Miyashita and Feldman, 2013b] Miyashita, T. and Feldman, D. E. (2013b). Behavioral detection of passive whisker stimuli requires somatosensory cortex. *Cerebral Cortex*, 23(7):1655–1662.

[Mo and Sherman, 2019] Mo, C. and Sherman, S. M. (2019). A sensorimotor pathway via higher-order thalamus. *Journal of Neuroscience*, 39(4):692–704.

[Morel and Imig, 1987] Morel, A. and Imig, T. J. (1987). Thalamic projections to fields A, AI, P, and VP in the cat auditory cortex. *Journal of Comparative Neurology*, 265(1):119–144.

[Murakami, 1977] Murakami, H. (1977). Rhythmometry on licking rate of the mouse. *Physiology and Behavior*, 19(6):735–738.

[Musall et al., 2019] Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S., and Churchland, A. K. (2019). Single-trial neural dynamics are dominated by richly varied movements. *Nature Neuroscience*, 22(10):1677–1686.

[Necker, 1832] Necker, L. (1832). LXI. Observations on some remarkable optical phænomena seen in Switzerland; and on an optical phænomenon which occurs on viewing a figure of a crystal or geometrical solid . *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 1(5):329–337.

[O'Connor et al., 2010a] O'Connor, D. H., Clack, N. G., Huber, D., Komiyama, T., Eugene, W., and Svoboda, K. (2010a). Supplementary figure and movie legends for

: Vibrissa-based object localization in head-fixed mice. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 30(5):1947–67.

[O'Connor et al., 2010b] O'Connor, D. H., Clack, N. G., Huber, D., Komiyama, T., Myers, E. W., and Svoboda, K. (2010b). Vibrissa-Based Object Localization in Head-Fixed Mice. *Journal of Neuroscience*, 30(5):1947–1967.

[O'Connor et al., 2010c] O'Connor, D. H., Peron, S. P., Huber, D., and Svoboda, K. (2010c). Neural activity in barrel cortex underlying vibrissa-based object localization in mice. *Neuron*, 67(6):1048–1061.

[Olcese et al., 2013] Olcese, U., Iurilli, G., and Medini, P. (2013). Cellular and Synaptic Architecture of Multisensory Integration in the Mouse Neocortex. *Neuron*, 79(3):579–593.

[Ouzounov et al., 2017] Ouzounov, D. G., Wang, T., Wang, M., Feng, D. D., Horton, N. G., Cruz-Hernández, J. C., Cheng, Y. T., Reimer, J., Tolias, A. S., Nishimura, N., and Xu, C. (2017). In vivo three-photon imaging of activity of GcamP6-labeled neurons deep in intact mouse brain. *Nature Methods*, 14(4):388–390.

[Pachitariu et al., 2016] Pachitariu, M., Stringer, C., Dipoppa, M., Schröder, S., Rossi, L. F., Dipoppa, M., Rossi, L. F., Carandini, M., and Harris, K. D. (2016). Suite2p : beyond 10 , 000 neurons with standard two-photon microscopy. *bioRxiv*, pages 1–30.

[Panniello et al., 2018] Panniello, M., King, A. J., Dahmen, J. C., and Walker, K. M. (2018). Local and global spatial organization of interaural level difference and frequency preferences in auditory cortex. *Cerebral Cortex*, 28(1):350–369.

[Pannunzi et al., 2011] Pannunzi, M., Pereda Banos, A., Perez Bellido, A., Soto-Faraco, S., and Deco, G. (2011). Neural mechanisms of audio tactile integration in the flutter range. *BMC Neuroscience*, 12(S1):1–2.

[Peng et al., 2015] Peng, Y., Gillis-Smith, S., Jin, H., Tränkner, D., Ryba, N. J. P., and Zuker, C. S. (2015). Sweet and bitter taste in the brain of awake behaving animals. *Nature*, 527(7579):512–515.

[Peron et al., 2015] Peron, S. P., Freeman, J., Iyer, V., Guo, C., and Svoboda, K. (2015). A Cellular Resolution Map of Barrel Cortex Activity during Tactile Behavior. *Neuron*, 86(3):783–799.

[Petersen, 2007] Petersen, C. C. (2007). The functional organization of the barrel cortex. *Neuron*, 56(2):339–355.

[Petreanu et al., 2012] Petreanu, L., Gutnisky, D. a., Huber, D., Xu, N.-l., O'Connor, D. H., Tian, L., Looger, L., and Svoboda, K. (2012). Activity in motor-sensory projections reveals distributed coding in somatosensation. *Nature*, 489(7415):299–303.

[Poort et al., 2015] Poort, J., Khan, A. G., Pachitariu, M., Nemri, A., Orsolic, I., Krupic, J., Bauza, M., Sahani, M., Keller, G. B., Mrsic-Flogel, T. D., and Hofer, S. B. (2015). Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex. *Neuron*, 86(6):1478–1490.

[Prescott et al., 2011] Prescott, T. J., Diamond, M. E., and Wing, A. M. (2011). Active touch sensing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1581):2989–2995.

[Prönneke et al., 2015] Prönneke, A., Scheuer, B., Wagener, R. J., Möck, M., Witte, M., and Staiger, J. F. (2015). Characterizing VIP neurons in the barrel cortex of VIPcre/tdTomato mice reveals layer-specific differences. *Cerebral Cortex*, 25(12):4854–4868.

[Rock and Victor, 1964] Rock, I. and Victor, J. (1964). Vision and Touch: An Experimentally Created Conflict between the Two Senses. *Science*, 143(3606):594–596.

[Rothschild et al., 2010] Rothschild, G., Nelken, I., and Mizrahi, A. (2010). Functional organization and population dynamics in the mouse primary auditory cortex. *Nature Neuroscience*, 13(3):353–360.

[Rutkowski and Weinberger, 2005] Rutkowski, R. G. and Weinberger, N. M. (2005). Encoding of learned importance of sound by magnitude of representational area in primary auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 102(38):13664–13669.

[Saleem et al., 2013] Saleem, A. B., Ayaz, A. I., Jeffery, K. J., Harris, K. D., and Carandini, M. (2013). Integration of visual motion and locomotion in mouse visual cortex. *Nature Neuroscience*, 16(12):1864–1869.

[Sales, 1972] Sales, G. (1972). Ultrasound and mating behaviour in rodents with some observations on other behavioural situations. *Journal of Zoology*, 168:149–164.

[Sato and Svoboda, 2010] Sato, T. R. and Svoboda, K. (2010). The functional properties of barrel cortex neurons projecting to the primary motor cortex. *Journal of Neuroscience*, 30(12):4256–4260.

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

[Schoups et al., 2001] Schoups, A., R, O., N, Q., and GA, O. (2001). Practising orientation identification improves orientation coding in V1 neurons. *Nature*, 412(6846):549–553.

[Siemann et al., 2015] Siemann, J. K., Muller, C. L., Bamberger, G., Allison, J. D., Veenstra-VanderWeele, J., and Wallace, M. T. (2015). A novel behavioral paradigm to assess multisensory processing in mice. *Frontiers in behavioral neuroscience*, 8(January):456.

[Sofroniew et al., 2015] Sofroniew, N. J., Vlasov, Y. A., Hires, S. A., Freeman, J., and Svoboda, K. (2015). Neural coding in barrel cortex during whisker-guided locomotion. *eLife*, 4.

[Song et al., 2017] Song, Y. H., Kim, J. H., Jeong, H. W., Choi, I., Jeong, D., Kim, K., and Lee, S. H. (2017). A Neural Circuit for Auditory Dominance over Visual Perception. *Neuron*, 93(4):940–954.e6.

[Spence and Squire, 2003] Spence, C. and Squire, S. (2003). Multisensory integration: Maintaining the perception of synchrony. *Current Biology*, 13(13):519–521.

[Stein and Meredith, 1993] Stein, B. E. and Meredith, M. A. (1993). *The merging of the senses.* The MIT Press.

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

[Stiebler et al., 1997] Stiebler, I., Neulist, R., Fichtel, I., and Ehret, G. (1997). The auditory cortex of the house mouse: Left-right differences, tonotopic organization and quantitative analysis of frequency representation. *Journal of Comparative Physiology - A Sensory, Neural, and Behavioral Physiology*, 181(6):559–571.

[Stringer et al., 2019] Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C. B., Carandini, M., and Harris, K. D. (2019). Spontaneous behaviors drive multidimensional, brainwide activity. *Science*, 364(6437).

[Stüttgen et al., 2011] Stüttgen, M. C., Schwarz, C., and Jäkel, F. (2011). Mapping spikes to sensations. *Frontiers in Neuroscience*, 5(NOV):1–17.

[Takagaki et al., 2008] Takagaki, K., Zhang, C., Wu, J. Y., and Lippert, M. T. (2008). Crossmodal propagation of sensory-evoked and spontaneous activity in the rat neocortex. *Neuroscience Letters*, 431(3):191–196.

[Teichert and Bolz, 2018] Teichert, M. and Bolz, J. (2018). How Senses Work Together: Cross-Modal Interactions between Primary Sensory Cortices. *Neural Plasticity*, 2018:1–11.

[Tolias et al., 2007] Tolias, A. S., Ecker, A. S., Siapas, A. G., Hoenselaar, A., Keliris, G. A., and Logothetis, N. K. (2007). Recording chronically from the same neurons in awake, behaving primates. *Journal of Neurophysiology*, 98(6):3780–3790.

[Tollin, 2003] Tollin, D. J. (2003). The lateral superior olive: A functional role in sound source localization. *Neuroscientist*, 9(2):127–143.

[Tollin, 2005] Tollin, D. J. (2005). Interaural Phase and Level Difference Sensitivity in Low-Frequency Neurons in the Lateral Superior Olive. *Journal of Neuroscience*, 25(46):10648–10657.

[Treviño et al., 2013] Treviño, M., Oviedo, T., Jendritza, P., Li, S. B., Köhr, G., and De Marco, R. J. (2013). Controlled variations in stimulus similarity during learning determine visual discrimination capacity in freely moving mice. *Scientific Reports*, 3:1–13.

[Tsukano et al., 2011] Tsukano, H., Hishida, R., and Shibuki, K. (2011). Detection of virtual pitch up to 5kHz by mice. *Neuroscience Research*, 71(2):140–144.

[Tyll et al., 2011] Tyll, S., Budinger, E., and Noesselt, T. (2011). Thalamic influences on multisensory integration. *Communicative and Integrative Biology*, 4(4):378–381.

[Varoquaux et al., 2015] Varoquaux, G., Buitinck, L., Louppe, G., Grisel, O., Pedregosa, F., and Mueller, A. (2015). Scikit-learn. *GetMobile: Mobile Computing and Communications*, 19(1):29–33.

[Veinante and Deschênes, 1999] Veinante, P. and Deschênes, M. (1999). Single- and multi-whisker channels in the ascending projections from the principal trigeminal nucleus in the rat. *Journal of Neuroscience*, 19(12):5085–5095.

[Wachowiak et al., 2013] Wachowiak, M., Economo, M. N., Díaz-Quesada, M., Brunert, D., Wesson, D. W., White, J. A., and Rothermel, M. (2013). Optical dissection of odor information processing in vivo using GCaMPs expressed in specified cell types of the olfactory bulb. *Annals of Internal Medicine*, 158(6):5285–8300.

[Wallace et al., 2004] Wallace, M. T., Ramachandran, R., and Stein, B. E. (2004). A revised view of sensory cortical parcellation. *Proceedings of the National Academy of Sciences*, 101(7):2167–2172.

[Wimmer et al., 2015] Wimmer, R. D., Schmitt, L. I., Davidson, T. J., Nakajima, M., Deisseroth, K., and Halassa, M. M. (2015). Thalamic control of sensory selection in divided attention. *Nature*, 526(7575):705–709.

[Winer and Schreiner, 2005] Winer, J. and Schreiner, C. (2005). *The Central Auditory System: A Functional Analysis*. Springer Science+.

[Woolsey and Van der Loos, 1970] Woolsey, T. A. and Van der Loos, H. (1970). The structural organization of layer IV in the somatosensory region (S I) of mouse cerebral cortex. The description of a cortical field composed of discrete cytoarchitectonic units. *Brain Research*, 17(2):205–242.

[Wu et al., 2013] Wu, H. P. P., Ioffe, J. C., Iverson, M. M., Boon, J. M., and Dyck, R. H. (2013). Novel, whisker-dependent texture discrimination task for mice. *Behavioural Brain Research*, 237(1):238–242.

[Yin et al., 2011] Yin, P., Johnson, J. S., O'Connor, K. N., and Sutter, M. L. (2011). Coding of amplitude modulation in primary auditory cortex. *Journal of Neurophysiology*, 105(2):582–600.

[Young et al., 2017] Young, G. W., Murphy, D., and Weeter, J. (2017). Haptics in music: The effects of vibrotactile stimulus in low frequency auditory difference detection tasks. *IEEE Transactions on Haptics*, 10(1):135–139.

[Zhang et al., 2020] Zhang, M., Kwon, S. E., Ben-Johny, M., O'Connor, D. H., and Issa, J. B. (2020). Spectral hallmark of auditory-tactile interactions in the mouse somatosensory cortex. *Communications Biology*, 3(1):1–17.

[Zhou and Fuster, 2004] Zhou, Y. D. and Fuster, J. M. (2004). Somatosensory cell response to an auditory cue in a haptic memory task. *Behavioural Brain Research*, 153(2):573–578.