# Characterisation of neural activity across the mouse visual cortex during virtual navigation

by

Efthymia (Mika) Diamanti

A dissertation submitted in partial fulfilment of the requirements for the degree of

**Doctor of Philosophy**

Supervisor: Matteo Carandini

Institute of Ophthalmology

University College London

I, Efthymia (Mika) Diamanti, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

*To Andreas and Eleni, my parents*

# Abstract

The brain's visual and navigational systems are thought to be involved in distinct neural processes. Yet, it is known that neurons in areas involved in the formation of spatial representations, such as the hippocampus, are also influenced by visual signals. In this Thesis I asked whether a similar influence exists in the opposite direction, namely whether navigational signals influence processing in primary visual cortex (V1) and in six higher visual areas. In parallel, given that little is known about the role of higher visual areas, especially during behaviour, I will seek to characterise their functional properties and differences across conditions of increased behavioural complexity, from passive viewing of drifting gratings all the way to virtual navigation.

In the first Results chapter, Chapter 3, I will demonstrate that during running through a virtual reality environment, visual responses as early as in V1 are strongly influenced by spatial position. From Chapter 4 onward, together with V1 I will also focus on 6 higher visual areas (LM, AL, RL, A, AM and PM). Specifically, I will attempt to probe activity in these areas across a wide spectrum of conditions: passive viewing of drifting gratings (Chapter 4); active engagement in virtual reality (Chapter 5) and passive viewing in virtual reality (Chapter 6). The results presented in Chapters 5 and 6 will suggest that spatial modulation is present across visual areas specifically during active behaviour. Finally, in Chapter 7 I will ask whether activity in V1, AL and the posterior parietal cortex (PPC) depends on yet another navigational variable, distance run, and how is this dependence different between areas.

In summary, by combining ideas and approaches from research in vision and navigation, I will seek to provide new, intriguing evidence about how neurons across the visual cortex combine visual with navigation-related signals to inform behaviour.

# Impact Statement

How the brain works is a fundamental question, that has always fascinated the scientific community as well as the general audience. Today, much of our insight on the brain's mechanisms originates from studies on the visual system and the navigational system, which are two of the most studied pathways in neuroscience. Yet we do not know how they work together.

In this Thesis I addressed this question and provided several lines of evidence indicating that signals from the brain's navigational system interact with visual signals much earlier than expected. I believe these findings will interest a wide range of neuroscientists, including and beyond the large communities who study vision or navigation. In fact, the importance of some of the findings presented here was recently appreciated by the Nature scientific journal, and I am glad to mention that our manuscript 'Coherent encoding of subjective spatial position in visual cortex and hippocampus' has been accepted for publication. Having our manuscript accepted by such a prestigious, multidisciplinary scientific journal is already a proof of the potential impact of our findings to the scientific community.

In addition, I believe that these findings will also be beneficial beyond the scientific community. I hope that my results obtained here from experiments in mice, will eventually be mapped on our brains and perhaps will provide some new insight on the origins of neuropathological disorders. Indeed, in light of a new, fascinating era in neuroscience, we are now starting to recognise that many of the brain's areas, traditionally thought to function independently, are actually strongly interacting with each other; dysfunction of these interactions may be the cause of several disorders. Along these lines, in this Thesis I provide evidence suggesting that the early visual system is influenced by the brain's navigational system. I am certain that these new ideas will be of interest to the broad audience and will contribute towards a revision of our current knowledge on the fundamental mechanisms underlying cortical processing.

# Acknowledgements

This thesis is the written proof of the amazing, both productive but also very enjoyable, time I had in the Carandini-Harris group and the Saleem lab during my PhD. The work presented here was only made possible thanks to many people who certainly deserve to be acknowledged.

First, I feel very privileged to have been supervised by Matteo Carandini, a scientist with a rare virtue: Matteo's great scientific insight is combined with an excellent, down-to-earth efficiency. These two features of his personality have been inspiring and very crucial in many aspects; Matteo has been a phenomenal mentor when delving into the mysteries of science and also an excellent teacher of impressive precision when having to communicate new ideas; Matteo's decisions and suggestions have been invaluable, mainly because they are always the product of a very specific plan and reasoning (although I still I don't understand why I had to move my desk by exactly 12 cm). For all his support, I am grateful.

Providing a full description of the contribution of Aman Saleem in just a few sentences is rather challenging. Aman has been very present throughout my PhD in all aspects: tackling scientific questions, setting up experiments, analysing data, planning ahead and preparing for posters and talks. Aman's unique scientific ideas but also excellent organisational skills have been vital (although I was not very fond of being asked to tick boxes and make lists, I have to admit it worked!). Aman also has another remarkable quality: on stressful days he is able to calm things down and bring you back to reality (numerous times I knocked on his door saying: 'I've got nothing!'. But after talking to him I felt so much better). For teaching me all these qualities and devoting all this time, I am thankful.

I would also like to thank Kenneth Harris for his outstanding scientific advice and his immediate and very insightful response in all matters, especially with regards to statistics and modelling (I have to admit though, that his insight has sometimes been a bit overwhelming, on occasions when I had to spend a couple of hours to process what Kenneth meant; but this is certainly a sign of his brilliance). Whenever everyone else is

stuck, Kenneth always has the answer. He is also a great teacher; his course on Neuroinformatics is still a major reference for me whenever I need to use more quantitative approaches. For all these reasons, I am grateful.

Many thanks should also be addressed to Julien Fournier, a scientist with an impressive ability to question things 'straight to the point'. His contribution in our recent paper has been invaluable. Here I should also mention that the results of Chapter 7 (which is probably my favourite chapter) are now part of this thesis only thanks to Julien. Just 4 weeks before submitting, he helped me out in doing new analysis and preparing figures. Thanks to him this chapter is now included in the Thesis.

Throughout my PhD I had the chance to focus on science completely hassle-free thanks to Charu Reddy, 'the best lab manager in the world' as I like to call her. Her excellent management skills were noticeable in so many occasions: from animal breeding to all different kinds of admin work. I am particularly grateful to her for helping me out with surgeries on very busy days, but also for her motherly support throughout my PhD.

Michael Krumin and Sylvia Schroeder deserve special thanks for their help with two-photon imaging. Michael's expertise and reasoning on any kind of technical issues has always impressed me (sometimes this was a bit too much. While I wanted to have something done and over, Michael would take his time explaining things from the very first principles; still, everything I know today about optics is thanks to him). Many thanks to Sylvia Schroeder who trained me in two-photon imaging; thanks to her admirable precision and perfection, nothing is broken in the two-photon rig to date.

Many thanks to 'the expert' as I like to call him, Federico Luigi Rossi (I know, his names are in the wrong order); a scientist with outstanding surgical skills (I hope I got a 10% of those) and a brilliant way of scientific thinking. Having started our PhDs almost at the same time, it was a privilege discussing with him all the stresses and shortcomings that both of us came across during this period. Overall, Federico has been an excellent companion through my PhD path.

In addition, many thanks to: Miles Wells for training some of the animals in the circular maze task; I-Chun for her help with modelling and for the fun times we had together; Daisuke Shimaoka for his invaluable help with wide-field imaging during the early days of my PhD; Andy Peters for his help in obtaining retinotopic maps on days when the 'big rig' (now renamed as 'lilrig') was not functional; Sam Failor and Lauren Wool for the excellent

cooperation on those busy days when we all had to use the two-photon rig; Sam Solomon for his (sometimes annoyingly) insightful advice.

Finally, I would like to thank all members in the Carandini-Harris and the Saleem lab for their advice and discussions, and for making this process so enjoyable! I was lucky enough to receive support from exceptional mentors, very bright minds and by now, very good friends. Thank you all!

# Table of contents

# Table of figures

# General introduction

Decades of research have provided valuable insight into the functional and anatomical properties of the brain's visual system. According to a simplified, textbook-like summary of the visual system's properties, visual input impinging on the retina travels through the lateral geniculate nucleus (LGN) of the thalamus into the primary visual cortex (V1). From V1, visual information is then distributed to higher order structures, and, depending on task demands, it is processed by a specialised set of higher visual areas. Indeed, the visual system possesses multiple visual streams, that function in parallel and project to diverse downstream targets (Nassi and Callaway, 2009).

Most of the knowledge on the existence of parallel visual streams stems from studies in the primate cortex (Milner and Goodale, 2008), and in fact the first traces of evidence came from human studies in second world war patients back in the 1960s. In 1969, Newcombe (Newcombe, 1969) observed that subgroups of patients exhibited distinct task performance deficits depending on whether they had gunshot wounds in the parietal or temporal lobe. Patients with parietal lesions could not perform visuospatial tasks, but they were still able to perform visual recognition tasks, whereas patients with temporal lesions performed well in visuospatial tasks, but poorly in visual recognition tasks. After this initial report, decades of research in primates were to follow, until a refined view on the existence of parallel streams was eventually brought forward by Goodale and Milner in 1992 (Goodale and Milner, 1992). According to this view, the visual system possessed two anatomically and functionally distinct pathways, both arising from V1; the 'what' pathway is involved in visual perception and it enables us to perceive features such as shape, orientation, size and faces; the 'where' pathway is involved in vision for action, and it allows us to process spatial context, such as location, distance and motion (de Haan, Jackson, and Schenk, 2018).

Since then, this model proposed by Milner and Goodale has inspired numerous studies to investigate the functional specificity of visual pathways in primates, and continues to be inspiring to date. Nevertheless, based on current knowledge, this model is indubitably an oversimplification. This notion is particularly highlighted in the recent review by (Galletti and Fattori, 2018). Here the authors summarise evidence from primate and human studies suggesting that the 'where' pathway is further divided into two sub-streams. Importantly though, they underline that none of the processing cortical streams should be viewed as an ensemble of interconnected areas dedicated to only one function. Instead, in view of modern theories proposing that the brain's functions arise from networks of broadly interconnected neurons, it is argued that the same neurons participate in more than one processes by dynamically changing their function depending on context.

The ability of neurons to dynamically change their function depending on behavioural context has been emphasised in various ways by yet another considerable set of studies in primates focusing on top-down influences (for a review on top-down influences on visual processing see Gilbert and Li, 2013. Although an exhaustive enumeration of the relevant literature is beyond the scope of this introduction, it should be mentioned that thanks to these top-down influences, which can take many forms, neurons at the earlier processing stages can modulate their activity to reflect specific perceptual demands. Neurons in the visual cortex, for instance, can convey different information about the same visual scene through influences by several factors, varying from attention or expectation to learning and memory (Gilbert and Li, 2012).

If activity in visual cortex can be influenced by such a broad range of cognitive processes, then it is intriguing to hypothesise that it may also be influenced by navigation-related signals during visually-guided navigation.

Navigation, or in other words, one's ability to find her way in an environment is a function of paramount importance, and as such, it is conserved across species. Unlike the origins of our knowledge on visual processing though, which came from studies in the primate and cat, much of the insight on the neural mechanisms underlying navigation originated from studies in the rat. Since the pivotal study of John O'Keefe in the 1970s (O'Keefe, 1976), numerous studies have utilised a plethora of real environments and mazes to dissect the principles of navigation in the rat. Instead, studying navigation in a real environment in humans and primates was less straightforward (although see for eg. (Matsumura et al., 1999) on monkeys performing a real translocation task). Therefore,

pivotal studies seeking for the existence of a spatial cognitive map in primates (Rolls, Robertson, and Georges-François, 1997) and humans (Ekstrom et al., 2003) turned to different experimental designs, such as free-viewing memory tasks (Rolls, Robertson, and Georges-François, 1997) or virtual reality environments (Burgess, Maguire, and O'Keefe, 2002).

The use of virtual reality to probe navigation in mice was introduced a bit later (Dombeck et al., 2007; Harvey, Collman, Dombeck, and Tank, 2009). In this configuration, mice ran through a virtual reality environment while head-fixed. Head-fixation thus eliminated motion artefacts and enabled accurate measurement of neural activity during behaviour. The introduction of head-fixation experiments in mice coincided with the development of powerful genetic tools, which allowed monitoring of large neural populations using optical techniques as well as the precise, but reversible, perturbation of neural activity with the use of optogenetic methods (Zeng and Madisen, 2012). These technical advances marked the era when the mouse would become an increasingly useful model organism for probing neural mechanisms underlying various demanding cognitive processes, such as visual perception or visually-guided navigation. And indeed, it turned out that mice were able to perform behaviours more complex than what was previously appreciated.

By building upon the invaluable insight provided by two parallel streams of research on vision and navigation, in this study we employed the technical advantages of the mouse model, to probe the influence of navigation-related signals across the visual cortex, and we found that these signals powerfully modulate visual responses during active behaviour. It remains to be seen whether similar mechanisms exist in other species. Nonetheless, since visually-guided navigation is a behaviour widely conserved across species, we hope this study will provide new evidence on the fundamental principles underlying sensory processing during this behaviour.

Below I summarise our knowledge on topics presented here, based on research performed in rodents. I will first refer to studies demonstrating the modulation of activity in V1 by non-sensory factors, similar, in the broad sense, to top-down influences found in primates. I will then talk about existing evidence indicating the anatomical and functional organisation of visual areas in the mouse. Next, I will enumerate the few studies that investigated the role of higher visual areas during behaviour, with a specific emphasis on navigation. I will also briefly discuss the little evidence available on the functional

differences between V1 and higher visual areas in the mouse, an important piece of knowledge if we want to continue talking about a potential organisation of the mouse visual areas into hierarchical processing streams. Finally, I will briefly discuss our main hypothesis on the transformation of visual signals into navigational signals and give an overview of the Results Chapters 3 to 7.

## 1.1 Activity in V1 is modulated by non-sensory factors

Activity as early as in V1 of the rodent is rapidly modulated by a broad-range of non-sensory factors. Indeed, V1 neuronal responses can be powerfully shaped by factors such as attention (Zhang et al., 2014), learning (Jurjut, Georgieva, Busse, and Katzner, 2017; Makino and Komiyama, 2015; Poort et al., 2015), reward expectancy (Poort et al., 2015; Shuler and Bear, 2006) or behavioural state (Niell and Stryker, 2010). For instance, learning of a behaviourally-relevant stimulus induces improved sensory processing performed by neurons selective for that stimulus (Jurjut, Georgieva, Busse, and Katzner, 2017; Poort et al., 2015); anticipation of a rewarding (Shuler and Bear, 2006) or aversive event (Makino and Komiyama, 2015) is reflected in the emergent timing activity of many V1 neurons; and locomoting behaviour is associated with enhancement of visually-evoked responses and an increase in the LFP gamma power which is a hallmark of arousal (Lee et al., 2014; Niell and Stryker, 2010). Thus, these studies, together with others, have provided compelling evidence that processing in V1 goes beyond the reliable representation of purely sensory stimuli; and that in fact, higher-level functions, originally thought to occur at a later processing stage, are already present much earlier along the sensory processing pathway.

Much of what we know today about the effect of behavioural context on visual processing in the mouse comes from studies investigating the influence of a simple behaviour: locomotion. Following Niell & Stryker (2010), numerous studies sought to characterise the multifaceted influence of this simple behaviour on V1 neuronal activity. For example, the increase in cortical gain observed during locomotion is associated with subthreshold effects: during locomotion, large-amplitude, low-frequency fluctuations of the membrane potential are suppressed, resulting in reduced membrane potential variability (Bennett, Arroyo, and Hestrin, 2013); furthermore, the membrane potential of excitatory and most inhibitory cell types shifts towards more depolarised potentials (Polack, Friedman, and Golshani, 2013). In addition, the locomotion-induced increase in cortical gain is observed not only in the orientation tuning (Niell and Stryker, 2010) but also in the contrast response functions (Lee et al., 2014). Beyond a uniform increase in cortical gain,

locomotion is also associated with a decrease in the strength of surround suppression encountered by V1 neurons during visual stimulation, thereby affecting spatial integration (Ayaz, Saleem, Schölvinck, and Carandini, 2013; Erisken et al., 2014). It also exerts a stronger influence in the activity of neurons preferring high spatial frequencies, thereby enhancing spatial resolution (Mineault, Tring, Trachtenberg, and Ringach, 2016). Finally, locomotion increases the amount of visual information encoded by the V1 population via two mechanisms: an increase in the firing rate (Dadarlat and Stryker, 2017), and a decrease in noise correlations (Dadarlat and Stryker, 2017; Erisken et al., 2014). Taken together, these findings indicate that the effects of locomotion go beyond a global change in cortical gain. Thus, locomotion cannot be considered the sole epiphenomenon of arousal for a discussion see (Niell, 2015). In fact, two studies that further investigated the effects of locomotion on V1 activity, attempted to disentangle arousal from locomotion (Reimer et al., 2014; Vinck, Batista-Brito, Knoblich, and Cardin, 2015). These studies revealed that although some effects are exclusively related to increased arousal, others are mediated by locomotion per se. These results pointed towards the dual nature of locomotion: it is partly the hallmark of an internal state, but also a physical variable with an important role in visual processing (Wekselblatt and Niell, 2015).

If locomotion plays such a key role in visual processing, then an important question arises: how are locomotor signals combined with visual signals under more natural conditions, such as during navigation? Indeed, during navigation the world around us changes as we move. To investigate the modulation of V1 activity during navigation, several studies have implemented behavioural paradigms in virtual reality (Keller, Bonhoeffer, and Hübener, 2012; Saleem et al., 2013); one advantage of such paradigms is that they allow active engagement with the sensory environment, since the animal's self-motion (speed) controls the update of the visual scenes (Poort et al., 2015); a second advantage, is the precise control of the visual scenes falling on the retina, since the animals are head-fixed. By exploiting the advantages of a virtual reality environment, Saleem et al. 2013 showed that the firing rate of most neurons was well predicted by a continuous model accounting for both position along the virtual reality corridor and speed. To further disentangle modulation by virtual and running speed, they replayed previous sessions to the animal regardless of its running behaviour ('open-loop condition'), and found that many neurons combined running and virtual speed equally. On the other hand, Keller et al. 2012 sought to characterise mismatch signals between self-motion and visual flow in the context

of predictive coding and found that responses of approximately 10% of V1 neurons were time-locked to the brief perturbations (halts) of the visual flow.

Overall, thanks to this plethora of studies on V1 activity during behaviour, not only have we obtained a good grasp of the basic properties of the early visual system in the mouse, but we are also now starting to unravel its neural processes underlying rich behaviours. Yet, much less in known about the functional and anatomical organisation of areas beyond V1, let alone their role in guiding behaviour. In the following sections I first enumerate pivotal studies that attempted to dissect the organization of higher visual areas in the mouse. I then review studies that sought to reveal the functional properties of higher visual areas using classic stimuli. Finally, I outline our current knowledge on the role of higher visual areas in the rodent during complex behaviours.

## 1.2 Functional and anatomical organisation of higher visual areas in the mouse

### 1.2.1 Identification of higher visual areas in the mouse

The first study to precisely parcellate the mouse visual cortex into V1 and 9 higher visual areas based on anatomy and receptive field mapping was the study by Wang and Burkhalter in 2007 (Wang and Burkhalter, 2007). By injecting three anterograde tracers at V1 sites representing distinct parts of the visual field, the authors showed that 9 out of 15 V1 targets, areas POR, P, LM, LI, AL, RL, A, AM and PM, contained complete representations of the visual field (**Figure 1-1**). In addition, all these areas had larger receptive fields than V1.

In the years to follow, two major studies sought to determine the exact number and boarders of higher visual areas using functional methods. Garrett et al. used intrinsic optical imaging to obtain maps of retinotopy together with an automated method which defined boarders between areas based on the reversal of the retinotopic gradient (Garrett, Nauhaus, Marshel, and Callaway, 2014; Juavinett et al., 2017). The authors identified 8 out of 9 areas reported by Burkhalter et al.; specifically, their approach did not classify A as an actual higher visual area. They also listed two new areas for the first time in the mouse, M and LLA. Interestingly, the authors of this study also reported that higher visual areas exhibited biases in visual field coverage. For instance, areas in lateral visual cortex, such as LM, AL or RL have representations of visual space that are more biased towards the centre. Instead, areas of the medial visual cortex, such as PM or M, have representations

that are more biased towards the periphery. This bias in each area's representation of visual space is rather emphasized in a later study by Zhuang et al (2017). Here the authors highlight that none of the higher visual areas represents the whole visual field. In fact, even areas immediately adjacent to V1, represent a distinct portion of the visual field. With respect to the list of retinotopically organised regions in the mouse, the same study listed 4 additional regions of cortex (termed 'patches'). These patches were MMA, MMP, RLL and part of the retrosplenial cortex (RS). Thus, to date the posterior cortex contains 16 retinotopically organised regions (V1, LM, LI, AL, LLA, RL, A, AM, PM, M, P, POR, MMA, MP, RLL, RS), and there is still a possibility that new regions may be added in the future.

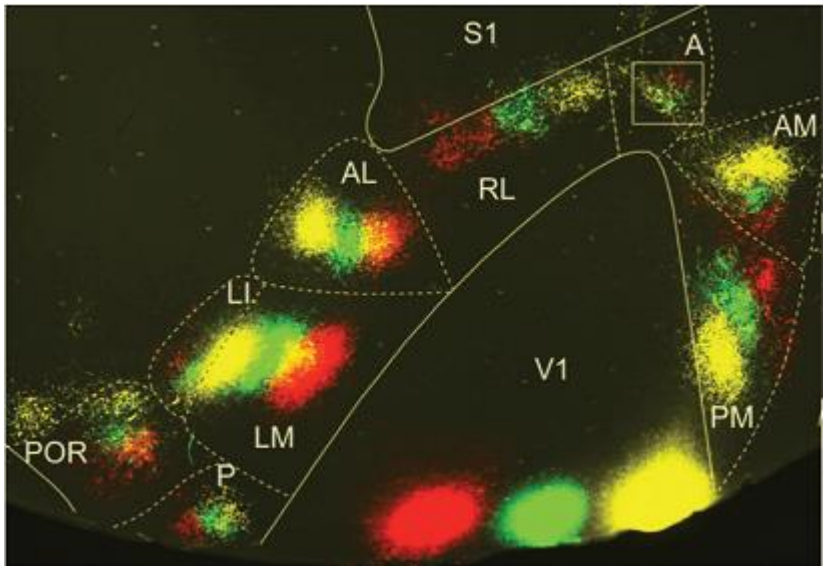


**Figure 1-1**: **Identification of the mouse higher visual areas.**
Projections from V1 to higher visual areas were identified with the use of three anterograde tracers (*yellow, green, red*) injected at different retinotopic positions in V1. Adapted from Wang & Burkhalter (2007). Abbreviations: A, anterior; AL, anterolateral; AM, anteromedial; LI, laterointermediate; LM, lateromedial; P, posterior; PM, posterior medial; POR, postrhinal; RL rostrolateral; V1, primary visual cortex.

Because the set of regions classified as retinotopically organised may not be exhaustive, and also because just the retinotopic organisation of an area does not necessarily imply it is an actual higher visual area, it is important that additional criteria are met for the identification of a set of higher visual areas in the mouse (Glickfeld and Olsen, 2017): these criteria could be functional, based for example on evidence for functional specialisation; alternatively, these criteria could be based on anatomical connectivity, whether for instance higher visual areas receive input from the higher-order visual thalamic nucleus, LP.

### 1.2.2 Functional specialisation of higher visual areas

The functional specialisation of higher visual areas has been the focus of several studies (Andermann et al., 2011; Van den Bergh, Zhang, Arckens, and Chino, 2010; Marshel, Garrett, Nauhaus, and Callaway, 2011; Murakami, Matsui, and Ohki, 2017; Roth, Helmchen, and Kampa, 2012; Smith et al., 2017; Tohmi et al., 2014). All these studies have used similar stimulation protocols to probe the spatiotemporal selectivity of each area. Spatiotemporal selectivity is then used as evidence indicating whether an area belongs to one of two putative processing streams: a stream specializing in object recognition (the 'what' or ventral stream); a stream specializing in motion perception (the 'where' or dorsal stream). For instance, an area belonging to the ventral stream, should have high spatial and low temporal acuity, therefore it should prefer high spatial and low temporal frequencies. Conversely, areas involved in action guidance should have the opposite preference: they should be tuned to high temporal and low spatial frequencies.

The first study to probe spatiotemporal tuning differences between V1 and LM in mouse (termed V2L) found that V1 and LM were very similar in their spatial and temporal frequency tuning (Van den Bergh, Zhang, Arckens, and Chino, 2010). Beyond LM, the first two studies to probe spatiotemporal tuning in more than one higher visual areas in the mouse were Marshel et al. (2011) and Andermann et al. (2011). Using wide-field imaging and 2-photon microscopy in anaesthetised mice, Marshel et al. (2011) analysed the tuning properties of neurons in V1 and 6 HVAs, LM, AL, RL, AM, PM and LI. They found that 5 out 6 HVAs preferred higher temporal frequencies than V1; only PM preferred lower temporal frequencies than V1. On the other hand, PM and LI were found to prefer higher spatial frequencies. The high temporal tuning of AL and the high spatial tuning of PM were also observed in awake mice (Andermann et al., 2011). In this study PM and AL were further discriminated based on peak speed preference, with PM preferring lower speeds than AL. Two following studies in anaesthetised mice confirmed the results reported previously (Roth, Helmchen, and Kampa, 2012; Tohmi et al., 2014). Roth et al. (2012) agreed that neurons in PM were mainly tuned to high spatial and low temporal frequencies (although, in the same study they also report a subpopulation in PM broadly tuned to low spatial frequencies). The study by Tohmi et al. (2014) confirmed the high velocity tuning for AL. It also showed that LM, RL and AM preferred higher stimulus velocities than V1.

More recently, two studies focusing on the development of higher visual areas provided additional evidence on each area's spatiotemporal preferences (Murakami, Matsui, and

Ohki, 2017; Smith et al., 2017). Murakami et al. (2017) considered 3 additional areas: POR, P and A. They found that POR and P, together with LI and PM, preferred high spatial and low temporal frequencies. In contrast, A, together with RL, were tuned to higher temporal frequencies than any other area, as well as to low spatial frequencies. AM and AL had somewhat different properties compared to the aforementioned studies (compare with Marshel et al. (2011) and Tohmi et al. (2014)). AM was the only area in their dataset preferring high temporal but also high spatial frequencies (Marshel et al. (2011) reports preference for low spatial frequencies). AL was not tuned to high temporal frequencies as previously reported. It rather had intermediate properties, but it still preferred higher temporal frequencies than V1. Like AL, LM also had intermediate properties. It preferred slightly higher temporal frequencies than V1, but lower than AL; it also preferred moderate spatial frequencies. Similar were the results on LM's spatial frequency preference based on the developmental experiments by Smith et al. (2017). During development, neurons in LM responded moderately stronger to both low and high spatial frequency stimuli. In contrast, neurons in PM responded more strongly to high spatial stimuli over time.

The broad tuning of LM has also been reported by the few existing studies asking how functional specialisation in higher visual areas arises (Glickfeld, Andermann, Bonin, and Reid, 2013; Matsui and Ohki, 2013). Indeed, axonal projections from V1 to LM have no selectivity for spatial frequencies and exhibit only a modest bias towards lower temporal frequencies. Instead, axonal projections from V1 and LM to areas PM and AL exhibit spatiotemporal frequency biases that match the preference of these areas. Specifically, axonal V1 and LM boutons in AL prefer low spatial and high temporal frequencies; whereas V1 and LM boutons in PM prefer high spatial and low temporal frequencies. These results indicate that spatiotemporal selectivity arises, at least in part, from the functional specificity of axonal projections from V1, and it is most likely further refined through feedforward projections from other higher visual areas such as LM.

Taken together these findings demonstrate that higher visual areas have distinct functional properties. Most studies report that PM, LI, P and POR prefer low temporal and high spatial frequencies, whereas AL, RL, AM and A prefer high temporal and low spatial frequencies. Instead, less clear-cut is LM's preference. Some studies report higher temporal frequency tuning compared to V1; others find that that LM is broadly tuned and more similar to V1 compared to other areas.

Less conclusive have been the attempts of classifying an area as ventral-like or dorsal-like based on its direction tuning properties. For instance, it is expected that an area involved in motion processing (i.e. dorsal-like) should also contain more cells with high direction selectivity compared to V1. According to Marshell et al. (2011) this is indeed the case for AL, RL and AM. High direction selectivity for these areas plus for LM was also demonstrated by Tohmi et al. (2014) and confirmed by Juavinett et al.(Juavinett and Callaway, 2015). for AL, AM and LM (but not for RL). On the other hand, Andermann et al. found lower direction selectivity in AL and PM compared to V1, whereas Roth et al. (2012) found similar direction selectivity between V1 and PM. Similar discrepancies were observed in the degree of orientation selectivity. Marshel et al. (2011) and Tohmi et al. (2014) reported higher orientation selectivity in LM, AL, RL and AM than in V1. LI was the only area with lower orientation selectivity than V1 (Marshel, Garrett, Nauhaus, and Callaway, 2011). In contrast, Andermann et al. found no difference in orientation selectivity between V1 and AL; also, consistent with the later study by Roth et al. (2012), they found no difference in orientation selectivity between V1 and PM. Overall, there has been considerable disagreement between studies about the orientation and direction tuning of higher visual areas. This disagreement can be attributed to the different stimulation protocols which may drive areas differently depending on their spatial and temporal frequency preference (Glickfeld and Olsen, 2017).

A few studies have attempted to identify dorsal-like and ventral-like streams based on other criteria, such as responsiveness to the global motion of plaids (Juavinett and Callaway, 2015) or correlations in response magnitude (Smith et al., 2017). In the study by Juavinett et al., the authors stimulated V1, LM, AL, RL and AM in anaesthetised mice with plaids moving in 12 different directions. Since plaids are composed by two drifting gratings of different orientations, the authors asked whether neurons in these 5 areas were responding to the direction of motion of the two gratings ('component' cells) and/or to the perceived motion of the plaid as a whole ('pattern' cells). They found that all areas contained component cells. The results were less conclusive about the existence of pattern cells in the mouse visual cortex. Nevertheless, based on a small number of pattern cells found in LM and RL, the authors proposed that LM and RL may be part of the motion-processing stream. To reconcile these findings with previous anatomical data (Wang, Gao, and Burkhalter, 2011; Wang, Sporns, and Burkhalter, 2012), which indicated that LM contributes to the ventral stream, the authors propose that LM, similarly to the primate V2, is part of both the dorsal and the ventral pathway. In the second study by Smith et al.(2017)

the authors recorded simultaneously from V1 and areas LM, AL, RL, LI and AM/PM (AM and PM were merged into a single area). Based on the correlations in response magnitudes between areas, two functional clusters were identified: one cluster comprised of LI and LM; a different cluster comprised of AL, RL and AM/PM. Because the clustering of areas into two groups was the same as the one reported previously by Wang et al. 2011 and 2012, the authors concluded that the LM-LI cluster corresponds to a ventral-like subnetwork, whereas the AL-RL-AM/PM cluster corresponds to a dorsal-like subnetwork.

Overall, these studies have elucidated some of the aspects of the functional specificity of the mouse HVAs, based on properties which in classic primate studies have revealed the differential specialization of non-primary cortical areas for either object recognition or motion perception. Nevertheless, additional studies are required before concluding that the distinct properties found in the mouse higher visual areas do actually indicate the existence of areas specialised in either object or motion perception.

### 1.2.3 Anatomical connectivity to and from higher visual areas

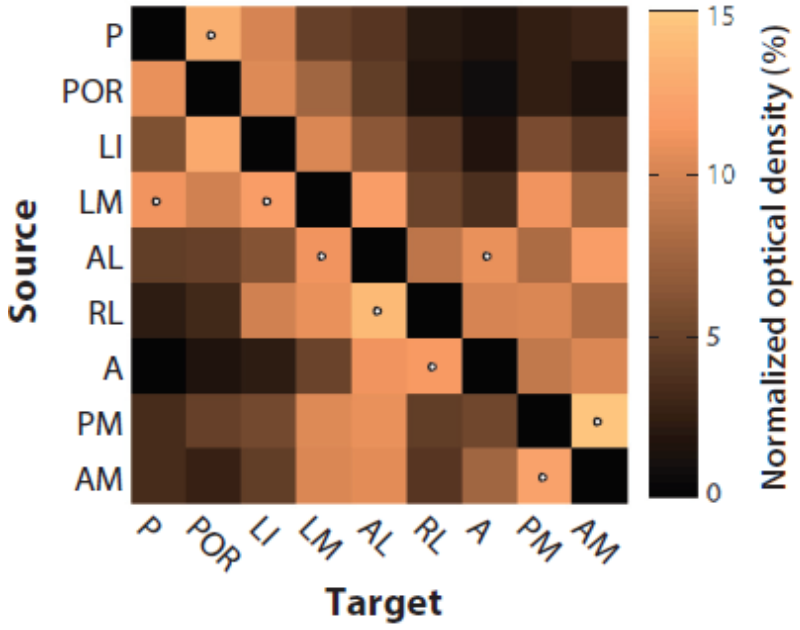


**Figure 1-2: Connectivity between higher visual areas in the mouse**
Summary of the optical density connections between higher visual areas. Adapted by Wang, Sporns, and Burkhalter, 2012. Note that each higher visual area projects to all other higher visual areas. However, based on the strength of their projections, higher visual areas are clustered into the ventral or dorsal substream.

Unlike the primate visual cortex, the mouse V1 has strong projections to at least 9 higher visual areas, LM, AL, RL, A, AM, PM, LI, P and POR (Wang and Burkhalter, 2007; **Figure 1-1**). The fact that V1 projections are distributed across several higher visual areas

could suggest that, unlike primates, higher visual areas in the mouse are not hierarchically organised. Nevertheless, the opposite has been suggested by several studies that sought to investigate connections to and from higher visual areas (D'Souza et al., 2016; Wang and Burkhalter, 2013; Wang, Gao, and Burkhalter, 2011; Wang, Sporns, and Burkhalter, 2012; **Figure 1-2**).

For instance, areas LM and AL, which receive the strongest projections from V1(Wang and Burkhalter, 2007), seem to have a special role along the visual pathway; they are considered the 'gateways' of the ventral and dorsal stream respectively (Wang, Gao, and Burkhalter, 2011; Wang, Sporns, and Burkhalter, 2012). According to these two studies, LM is more strongly connected to V1, LI, P and POR, and this set of visual areas preferentially projects to entorhinal and temporal association regions. Instead, AL is more strongly connected to areas RL, AM, A and PM, and this set of visual areas preferentially projects to the retrosplenial, secondary motor and anterior cingulate cortex. Thus, despite the two areas sharing the same border, they have distinct connectivity profiles and therefore different roles in visual processing.

The same group of investigators has provided additional anatomical evidence for the existence of a dorsal and a ventral pathway in the mouse (Wang and Burkhalter, 2013). Specifically, the ventral cluster of areas LM, LI, P and POR sends non-negligible projections to the superficial layers of the superior colliculus, which are known to be involved in visual processing. Instead the dorsal cluster of areas RL, AM, A and PM sends sparse projections to the superficial layers of the superior colliculus, but also to the deeper layers involved in premotor activity.

A few other anatomical features have served as evidence that some higher visual areas lie at distinct hierarchical levels. LM for example, is thought to be the homologue of primate V2, because it is the only area sharing the vertical meridian with the lateral portion of V1 (Wang and Burkhalter, 2007). In addition, whereas higher visual areas typically receive indirect input through the pathway from the superior colliculus to the lateral posterior nucleus of the thalamus, LM is the only higher visual area receiving direct input from the lateral geniculate nucleus (Oh et al., 2014). Finally, LM has stronger feedback projections to V1 than AL (Wang, Gao, and Burkhalter, 2011), and overall stronger feedforward projections to other higher visual areas (D'Souza et al., 2016), suggesting that LM is at the lower level of the hierarchy after V1. In addition to determining the hierarchical position of LM, the later study by D'Souza et al. also probed the hierarchical

position of PM. The authors used anterograde tracing to label axonal terminals from V1, LM and PM to 9 higher visual areas. They then used the density of axon terminals in layer 1 as a proxy for feedback projections, and the density of axon terminals in layers 2 to 4 as a proxy for feedforward projections. They found that the projections that LM sent out were more than those it received back. On the other hand, PM had similar amounts of feedforward and feedback projections. Given these findings, the authors concluded that LM lies at the bottom of the hierarchy among higher visual areas, followed by PM.

Taken together, numerous anatomical and functional studies have provided valuable evidence on the anatomical and functional properties of higher visual areas in the mouse. Nevertheless, an important question still remains open: what is the type of transformations taking place in the higher visual areas in the mouse? Providing answers to this question is essential before concluding that, similar to the primate, higher visual areas in the mouse are segregated into parallel processing streams. To tackle this question, future effort should be directed towards probing higher visual areas during behavioural tasks. Indeed, only a handful of studies have attempted to probe the role of higher visual areas during behaviour (see for e.g. Funamizu, Kuhn, and Doya, 2016; Itokazu et al., 2018), with the vast majority of those focusing on the posterior parietal cortex (PPC), an area thought to overlap, at least in part, with the higher visual areas, RL, A and AM. Given the proposed overlap between the PPC and the higher visual areas RL, A and AM (Wang, Gao, and Burkhalter, 2011), perhaps additional insight can be drawn from the numerous studies on the rat PPC during complex behaviours (see for e.g. McNaughton et al.; Nitz, 2006; Whitlock et al., 2012). Along these lines, in the following section I will summarise our current knowledge on the role of higher visual areas, including PPC, in the rodent during behaviour. Given the main focus of the current study on virtual navigation, particular emphasis will be given to studies investigating the neural processes underlying navigation in the rodent.

## 1.3 Activity in higher visual areas during behaviour

In contrast to the considerable amount of evidence on the influence of locomotion on V1 activity, few studies have assessed the effects of locomotion on higher visual areas (Andermann et al., 2011; Huh et al., 2018; Lecoq et al., 2014). By comparing epochs of running versus resting, this handful of studies reported that the increase in behavioural gain, previously observed in V1 (Niell and Stryker, 2010), was also present in LM (Lecoq et al., 2014), AL and PM (Andermann et al., 2011), with responses in AL being more strongly modulated by spontaneous running than responses in PM (Huh et al., 2018). To date, there

are no studies on higher visual areas assessing the effects of locomotion beyond the modulation by behavioural state. Therefore, additional studies are required to fully characterise the influence of locomotion in higher visual areas.

Besides locomotion, there is little evidence on the role of higher visual areas during more complex behaviours. If the mouse visual system is indeed composed of different, hierarchical processing levels, then a plausible hypothesis is that during complex behaviours some higher visual areas may be devoted to the multiplexing of sensory with non-sensory signals.

A promising candidate for such an integrative function is the PPC, a cortical region involved in a broad spectrum of behaviours, from decision making processes to navigation see for e.g. Harvey, Coen, and Tank, 2012; Krumin, Harris, and Carandini, 2017. Notably, PPC overlaps, at least in part, with three higher visual areas: RL, AM and A (Glickfeld and Olsen, 2017; Wang, Gao, and Burkhalter, 2011), therefore it is reasonable to think that its function somehow relates to visual processing. Indeed, inactivation of the PPC in rats results in decreased performance during a visually-guided decision-making task (Raposo, Kaufman, and Churchland, 2014). This decrease in visual performance is causally linked to the disruption of visual processing taking place in the PPC, rather than to disrupted accumulation of evidence (Licata et al., 2017). In addition, performance in a memory-guided visual discrimination task is affected when PPC is optogenetically inhibited during stimulus presentation, but it remains unchanged if PPC is inhibited during the delay period (Goard, Pho, Woodson, and Sur, 2016). Instead, performance does not change when inactivating PPC during a decision-making task based on auditory evidence (Raposo, Kaufman, and Churchland, 2014), and in fact PPC is minimally involved in the choice behaviour based on auditory evidence accumulation (Erlich et al., 2015). Taken together, these results demonstrate that the activity in the PPC is more sensitive to visual, rather than auditory, task features, consistent with its anatomical overlap with higher visual areas. In addition, it appears that PPC is not necessary for performance during the delay-working memory epoch of a visual discrimination task.

So far, I outlined recent studies in rodents that shed light on the role of PPC during complex, nonnavigational tasks. What is the role of PPC during navigational tasks? During a virtual navigation, working memory task in mice, PPC neurons encode a specific decision at a specific moment in time, thereby forming sequences of activity at the population level (Harvey, Coen, and Tank, 2012). Therefore, unlike non-navigational tasks, in which PPC is

not involved in working memory (Goard, Pho, Woodson, and Sur, 2016), PPC does encode choice during a memory-guided navigation task. Following a different approach, a later study by Krumin et al (2017), showed that when mice make decisions in a visually-guided virtual navigation task, PPC neurons encode navigational variables, such as position and heading, again highlighting the important role of PPC during navigation.

The involvement of PPC during navigation highlighted by the two aforementioned studies in mice builds upon the long-standing view that the rodent PPC plays an important role during navigation. Evidence supporting this view originate mainly from studies in rats stemming back to the 90s (Save and Poucet, 2009). The original study by McNaughton, 1994 (McNaughton et al., 1994) found that the activity of many PPC cells reflected various self-motion variables, such as forward motion, turning or running. Importantly, neuronal activity did not solely reflect motor responses; some cells encoded combinations of motor variables with spatial variables, suggesting that PPC neurons integrated movement with spatial context. The ability of PPC neurons to code for complex motor behaviours was also demonstrated in a later study (Nitz, 2006), where it was shown that during navigation in labyrinth-like mazes the activity of PPC neurons coding for movement formed sequences that reflected the routes traversed by the animal. Therefore, PPC neurons encode self-motion. These signals can occur up to 500 ms before the execution of movement, when rats forage in an open arena (Whitlock et al., 2012). Instead, when rats are engaged in a goal-directed task in a hairpin maze, PPC neurons change their tuning for self-motion to tuning for spatial position. The anticipatory activity of PPC neurons preceding movement was also demonstrated in the study by Wilber et al. (2014). Cells exhibiting this anticipatory activity performed mixed coding of head direction together with visual cue direction. Notably, modulation by the visual cues was also observed in cells coding for self-motion only, suggesting that the motion-related firing of these cells may be influenced by optic flow (Wilber et al., 2014). Overall, these studies suggest that PPC neurons combine self-motion signals together with visuo-spatial features in the environment and point towards an integrative role of the PPC during navigation.

How could this integrative function in the PPC be useful for navigation? One hypothesis is that the combined, head and cue direction information is then further combined with distance information giving rise to landmark vector cells in the hippocampus (Wilber et al., 2014). To date, it is not clear whether coding for distance information takes place in the PPC. Wilber et al. reported that very few cells in the PPC coded for goal distance during a navigation-to-hidden-goal task. In contrast, in a later

study, goal-distance was reliably decoded by activity in PPC, as well as in PM, during an auditory virtual navigation task (Funamizu, Kuhn, and Doya, 2016). Even in the absence of auditory cues, goal distance could still be decoded from the population activity in the two areas, suggesting predictive activity, but in the presence of auditory cues, decoding errors were reduced, suggesting refinement of encoding activity by sensory input.

Taken together, these findings underline the multifaceted role of PPC during complex behaviours. Although it is still debated which higher visual areas are part of the PPC, existing evidence on the function of the PPC during visually guided behaviours point towards an area involved in the transformations of visual with self-motion signals during navigation; such transformations could well be assigned to a higher visual area. However, drawing conclusions on the functional role of higher visual areas based solely on studies on the PPC is inadequate. Additional effort is required towards deciphering the functional role of the various higher visual areas during complex behaviours.

## 1.4 Differences between V1 and higher visual areas during behaviour

In the previous section I summarised our current knowledge on the role of higher visual areas during behaviour, by mainly extrapolating from studies in the rat PPC. Even fewer are the studies that used behavioural tasks to compare V1 to areas thought to lie higher in the sensory processing stream. Yet, such a comparison is crucial before drawing conclusions on whether hierarchical processing streams do exist in the mouse visual system.

Few recent studies sought to characterise the differences between V1 and PPC (Goard, Pho, Woodson, and Sur, 2016; Pho et al., 2018) or AL/A/AM (Itokazu et al., 2018). During a visual discrimination task, V1 is dedicated to the processing of visual input, whereas PPC coded for both the sensory input and the behavioural choice (Goard, Pho, Woodson, and Sur, 2016). In addition, visual input elicited robust responses in V1 during both conditions of active engagement in the task and passive viewing (Pho et al., 2018). Instead, the majority of PPC neurons responded robustly only during active engagement; PPC neurons also changed their selectivity to match the new target stimulus after reversal of the task contingency, thus suggesting that unlike V1, activity in the PPC reflects the learned task context. In another study, in which mice were trained to perform a visually-guided eye movement task, the authors demonstrated that, unlike V1, areas RL/A/AM

were connected reciprocally to a small portion of the secondary motor cortex involved in eye movements (Itokazu et al., 2018). Correspondingly, RL/A/AM coded for both eye movement and visual signals, whereas V1 population was biased towards purely visual activity.

Overall, it appears that during visually guided, non-navigational tasks V1 activity is restricted to the processing of visual information, whereas higher level functions reflecting task demands arise in a subset of higher visual areas. It remains to be seen whether similar differences also exist between V1 and higher visual areas other than PPC, and whether such differences are still present during visually-guided navigation.

## 1.5 The hypothesis: transformation of visual signals into navigation signals

During visually-guided navigation the brain has to transform visual information encoded by the visual system into spatial information encoded by the navigational system. This transformation involves transitioning across different reference frames. For instance, information processed by the early visual system can be best described in a reference frame centred to the eye. Indeed, the receptive fields of neurons in the early visual system are measured in units of degrees in visual space, because they are invariant to eye orientation. In other words, when the direction of gaze changes, the receptive fields move with the eye. On the other hand, information processed by the brain's navigational system, for example by the hippocampus, can be best described in a reference frame centred to the world. Indeed, neurons in this area called 'place cells' (O'Keefe and Dostrovsky, 1971) represent one's position with respect to the external environment, or equivalently, these representations are fully invariant to one's body orientation. Therefore, for visual information to guide navigation, it is required that visual, eye-centred representations are transformed into spatial, world-centred representations.

Early studies in monkeys sought to dissect the neural circuits involved in coordinate transformations (Cohen and Andersen, 2002). For instance, the original study by Andersen and Mountcastle (1983) demonstrated that neurons in PPC performed coordinate transformations from an eye-centred reference frame to a head-centred reference frame. Specifically, it was hypothesised that for such transformations to exist, visual responses should no longer depend only on the angle $x$ relative to the eye; they should also depend on the angle $y$ relative to the direction of gaze, such that visual responses would become

invariant in a head-referenced frame. To test this hypothesis, the authors measured the receptive fields of the recorded neurons at varying gaze angles and found that visual responses were strongly modulated by the direction of gaze. Because different subsets of neurons coded for different combinations of stimulus position and gaze direction, effectively the whole PPC population coded for all possible stimulus locations relative to eye *and* gaze direction, resulting in a representation of the visual scene that was invariant with respect to the head, rather than just the eye.

What is the mechanism underlying the dependence of PPC visual responses on the gaze angle? One characteristic of this dependence is that it is multiplicative, namely it exerts an effect on the magnitude of the visual response, without affecting the neuron's selectivity (i.e. the neuron's preferred location to fire maximally). This effect is reminiscent of gain modulation phenomena, which allow sensory signals to be combined with modulatory effects non-linearly(Salinas and Abbott, 2001). Consider for instance the influence of visual responses by the direction of gaze described above (Andersen and Mountcastle, 1983). This influence can be well described by the product of two functions: the neuron's response profile $f(x)$, which depends on the angle $x$ relative to the eye and the modulatory factor $g(y)$, termed the 'spatial gain field', which depends on gaze angle $y$. Then the firing rate of the neuron, $r$, is of the form: $r = f(x)g(y)$, with $g(y)$ depending on $y$ linearly (Andersen, Essick, and Siegel, 1985).In addition, this non-linear, multiplicative behaviour can arise from a network with recurrent connections, where neurons with similar tuning properties should excite each other, whereas neurons with different tuning properties should inhibit each other (Salinas and Abbott, 2001).

Overall, the aforementioned studies demonstrated that PPC neurons were able to perform transformations in multiple reference frames and proposed 'gain fields' as the appropriate theoretical framework for describing such operations. Based on existing evidence up until then, such gain fields operated in egocentric (eye or head), rather than allocentric, reference frames. Nevertheless, given the existence of place cells in hippocampus, which encode the animal's position in allocentric coordinates (with respect to the world), it is reasonable to ask which brain area upstream of the hippocampus and downstream of the PPC transforms information from an egocentric to an allocentric reference frame. To isolate allocentric from egocentric gain fields, Snyder, Grieve, Brotchie, and Andersen (1998) asked whether visual responses at identical retinal locations in monkeys were affected by one of two types of transformations: 1. Rotation of body under head, which resulted in rotation of a reference frame with respect to the body, but not to

the world; 2. Rotation of body and head, which resulted in rotation of a reference frame with respect to the world, but not to the body. They found that most responses in area LIP were modulated by rotations of body under head, whereas most responses in area 7a were modulated by rotations of both body and head. These results suggested that body- and world-referenced gain fields were anatomically segregated, with LIP representing visual information in a body-centred reference frame and 7a in a world-centred reference frame. Importantly, the discovery of world-referenced gain fields in area 7a indicated that the brain is able to transform visual information from an egocentric-retinal reference frame into an allocentric-world reference frame, a process particularly useful during navigation.

Overall these studies have provided powerful insight into the neural mechanisms underlying the transformations of information into multiple reference frames. Yet, it remains an open question how visual information transforms into navigation-related signals during visually-guided navigation. Here we hypothesised that similar to the gain fields found in monkeys, visual responses in the mouse may be modulated by spatial context during navigation. Consider for instance a cell that 'likes' the plaids. If the plaids are placed at two different locations in an environment, then this cell should respond similarly at the two locations -purely visual response- (**Figure 1-3** left). Instead, if spatial location plays a role, then the visual response should be different at different locations although the presented stimulus is the same (**Figure 1-3** right).

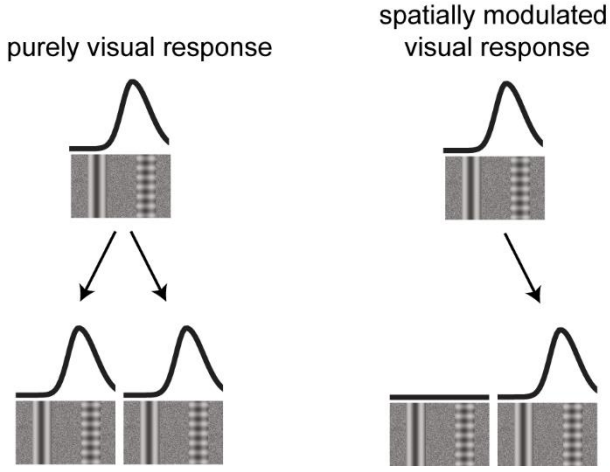


**Figure 1-3: Hypothesis on the modulation of visual responses by spatial location.**
*Left:* A visual cell that 'likes' the plaids exhibits the same response each time a plaid passes through its receptive field. *Right:* Hypothetical modulation of the visual response shown on the left by spatial location. Although the visual stimulus is the same, the cell responds differently at two different locations. *Images with gratings and plaids* represent segments of an environment -linear corridor.

To test this hypothesis, we recorded visual responses in V1 and 6 higher visual areas while mice ran head-fixed through a linear corridor in virtual reality. The corridor had a pair of landmarks, a grating and a plaid, that repeated twice, so that a purely visual neuron would respond similarly in those positions. However, many neurons as early as in V1 preferred to fire more strongly to the landmarks in one position. These results suggest that

visual responses are modulated by spatial context, and that this modulation arises as early as in a primary sensory area in the mouse.

## 1.6 Brief overview

With this review I have hopefully provided the reader with an adequate literature background and hinted on the questions I will address in the next chapters. Following a similar format like the one introduced here, I will first present results on the modulation of V1 by yet another factor, spatial context (Chapters 3). I will next characterise responses in 6 higher visual areas, LM, AL, RL, A, AM and PM, across a wide spectrum of conditions: passive viewing of drifting gratings (Chapter 4); active engagement in virtual reality (Chapter 5) and passive viewing in virtual reality (Chapter 6). In Chapters 5 and 6 I will also ask whether the degree of spatial modulation is the same across areas. Finally, in Chapter 7 I will ask whether activity in V1, AL and the PPC depends on another navigational variable, distance run, and how is this dependence different between areas.

In summary, by combining ideas and approaches from research in vision and navigation, I will seek to provide new, intriguing evidence about how neurons across the visual cortex combine visual with navigation-related signals to inform behaviour. My hope is that these findings will extend beyond the mouse, shedding light on the fundamental principles underlying sensory processing across species during navigation.

# Methods

All experimental procedures were conducted under personal and project licenses issued by the Home Office, in accordance with the UK Animals (Scientific Procedures) Act 1986.

For calcium imaging experiments in the linear virtual corridor and during stimulation with drifting gratings, we used double or triple transgenic mice expressing GCaMP6 in excitatory neurons (5 females, 1 male, implanted at 4-6 weeks). The triple transgenics expressed GCaMP6 fast (Madisen et al., 2015) (Emx1- Cre;Camk2a-tTA;Ai93, 3 mice). The double transgenic expressed GCaMP6 slow (Wekselblatt, Flister, Piscopo, and Niell, 2016) (Camk2a-tTA;tetO-G6s, 3 mice). For calcium imaging experiments in the circular virtual maze we used 4 double transgenics expressing GCaMP6 slow (4 males).

## 2.1 Surgical procedures

4-8-week mice were implanted with an 8 mm circular chamber and a 4 mm craniotomy was performed over the left or right visual cortex under isoflurane anaesthesia (between 1% and 2% depending on animal size and the stage of surgery). Before any operation, animals were injected subcutaneously with Rimadyl (0.5 mg/ml) and intramuscularly with Colvasone (0.2 mg/ml). Bulk of hair was removed with an electric razor, the head was fixed with the use of earbars, the skin was disinfected with iodine and a local anaesthetics (lidocaine) was applied on the scalp. Subsequently, an incision was made exposing the cranium on the left or on the right, such that both lambda and bregma were visible. The periostium was removed and the posterior muscles were gently retracted, so that an 8 mm headplate could be safely attached on the bone instead of the skin or the muscle. The skin in the periphery and the sutures were fixed with Vetbond. Craniotomy was performed by repeatedly rotating a biopsy punch and was shielded with a double coverslip (4 mm inner

diameter; 5 mm outer diameter). Post-surgery, some mice were allowed to recover for 7 days before being water restricted.

## 2.2 Visual stimulation

### 2.2.1 Maps of retinotopy

To measure retinotopy we presented a 14°-wide vertical window containing a vertical grating (spatial frequency 0.15 cycles/°), and swept (Kalatsky and Stryker, 2003; Yang, Heeger, and Seidemann, 2007) the horizontal position of the window over 135° of azimuth angle, at a frequency of 2 Hz. Stimuli lasted 4 s and were repeated 20 times (10 in each direction). We obtained maps for preferred azimuth by combining responses to the 2 stimuli moving in opposite direction, as previously described (Kalatsky and Stryker, 2003).

### 2.2.2 Neuropil receptive fields

To obtain neuropil receptive fields, on each recording session we presented sparse uncorrelated noise for 5min. The screen was divided into a grid of squares of 4 x 4 degrees size. Each square was turned on and off randomly at a 10Hz rate. At each moment in time 2% of the squares were on. To compute the neuropil receptive fields, the field of view was segmented into 5x5 patches (100μm x 100 μm surface per patch). For each patch, we first averaged the raw fluorescence across the patch's pixels. We then computed the stimulus-triggered average of the averaged raw fluorescence trace. The response was further smoothed in space and its peak was defined as the patch's receptive field centre.

### 2.2.3 Drifting gratings

For visual stimulation with drifting sinusoidal gratings two experimental protocols were used: a direction-selectivity protocol, where drifting gratings had varying orientations; a spatial-temporal frequency protocol, where drifting gratings had varying spatial and temporal frequencies. In all protocols, stimulus contrast was 100% and stimulus diameter was 40 deg. Within each session stimulus position was adjusted to match the centre of the receptive fields obtained from the neuropil.

In the direction selectivity protocol, we presented drifting gratings at 12 directions (30 deg step). Each grating had temporal frequency of 2Hz and spatial frequency of 0.04 cpd. Each stimulus presentation lasted 2 s, followed by a 3 s blank period, and was repeated 10 times. On each repeat, the 12 stimuli plus a blank were presented in pseudorandom order.

In the spatial-temporal frequency protocol, we presented a vertical, rightward drifting grating at 4 spatial frequencies (0.02, 0.04, 0.08 and 0.16 cpd) and 5 temporal frequencies (0.5, 1, 2, 4 and 8 Hz). Each stimulus presentation lasted 4 s, followed by a 3 s blank period, and was repeated 10 times. On each repeat, the 20 stimuli plus a blank were presented in pseudorandom order.

## 2.3 Virtual reality set-up

### 2.3.1 Linear corridor

Animals were head-fixed in the centre of 3 LCD monitors placed at a 90deg angle to each other, 19 cm away from each screen, so that visual scenes covered the visual field by 135 deg in azimuth and 42 deg in elevation. Animals ran through the virtual environment by locomoting on a polystyrene wheel which allowed movement along a single dimension (forwards-backwards). The animal's running speed was captured online by a rotary encoder (2400 pulses/rotation, Kübler, Germany).

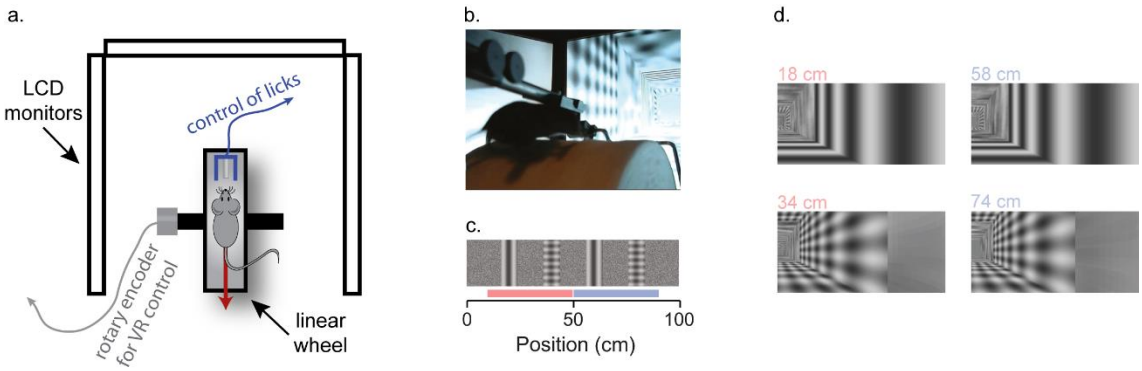


**Figure 2-1: Virtual reality apparatus and design of the virtual corridor**
**a, b.** Animals were head-fixed in the centre of 3 LCD monitors placed at a 90deg angle to each other. Animals ran through the virtual environment by locomoting on a linear, polystyrene wheel. The virtual reality environment was controlled online with a rotary encoder. In rewarded sessions, a drop of water was delivered to the animal through a water spout. Licks were controlled with a lick detector.
**c.** The virtual corridor had two landmarks (a grating and a plaid) that repeated after 40 cm, creating two visually-matching segments (*red* and *blue bars*).
**d.** Example screenshots, showing the visual similarity of the virtual corridor at two pairs of positions spaced 40 cm apart.

In the closed-loop condition, online measurement of running speed was used to control the update of the visual scenes. Upon reaching the 100$^{th}$ cm of the corridor, animals were presented with a grey screen for an inter-trial period of 3 to 5 s, after which they were placed back at the beginning of the corridor for the next trial. The duration of each trial depended on how long it took the animal to reach the end of the corridor. Trials in which

animals did not reach the end of the corridor within 30 s were timed-out and excluded from further analysis. A typical session consisted of more than 50 trials. In the open loop condition, mice were presented with a previous closed-loop session, while still free to run on the wheel.

The VR environment with visually-matching segments was a corridor of 8 cm width and height and a 100 cm length -short corridor-. The background textures of the ceiling, floor and walls were filtered white noise. A grating or a plaid, 8 cm wide each, alternated in the sequence grating-plaid-grating-plaid at 20, 40, 60 and 80 cm from the start of the corridor.

In the visually-identical environment, mice ran the same distance as before (100 cm) and were also placed back at the start of the corridor after 3-5 s presentation of a grey screen. The same four landmarks were also centred in the same positions as before. However, the corridor was extended to 200 cm length, repeating the same sequence of landmarks. The virtual reality software was modified to render only up to 70 cm ahead of the animal, ensuring the visual scenes were strictly identical in the sections between 10 - 50 cm and 50 - 90 cm; the white noise background also repeated with same 40 cm periodicity.

### 2.3.2 Circular maze

Animals were head-fixed in the centre of 3 IPAD LCD screens, 11 cm away from each screen, so that visual scenes covered the visual field by 135 deg in azimuth and 42 deg in elevation. As in the linear maze, animals ran through the virtual track by running on a wheel that allowed movement in the forwards-backwards dimension. Their running speed was captured online by a rotary encoder.

Animals ran over and over again through a virtual circular maze. They were only presented with a grey screen for an inter-trial period of 3 to 5 s if they licked more than 6 times in the wrong place. After the end of the grey screen period animals were placed back at the same position on the maze. A typical session consisted of more than 80 trials.

The virtual reality environment was an 'open' circular maze with 400 cm full length. Visual landmarks were three 'tunnels' of 12 cm perimeter, $L_1$, $L_2$ and $L_3$. Landmark $L_1$ was placed at the reward zone and alternated between a grating or a plaid on each half cycle. Landmarks $L_2$ and $L_3$ (a plaid and a grating) were placed at 40cm and 60 cm away from the reward zone. The background textures were filtered white noise.

## 2.4 Behaviour & training

### 2.4.1 Linear corridor

Mice used for imaging in the linear corridor (short or long) ran freely with no specific task.

Before recordings in the short corridor mice (4 animals) were placed in the virtual environment, typically for 3 days and for up to one week, until they were able to run for at least 80% of the time within a single session. Two of the mice were motivated to run with water rewards, by receiving ~3 μl water with the use of a solenoid valve (161T010; Neptune Research, USA). One animal received rewards at random positions along the corridor. The other received rewards at the end of the corridor. To control for the effect of the reward on V1 responses, no reward was delivered in some sessions (n = 8 sessions; 2 animals).

Prior to recordings in the long corridor, mice (2 animals) were first exposed to 5 sessions in the short corridor, then placed in the long corridor and allowed to habituate to the new environment for another two or three sessions before the start of recordings. One mouse was motivated to run with a reward delivery placed at 100 cm.

### 2.4.2 Circular maze

Mice (n = 4 animals) were trained to lick in a specific location on every half cycle, the reward zone. Trials in which animals did not perform well, i.e. they licked more than six times on a single trial, were timed-out and excluded from further analysis. We also excluded trials in which the animals were not engaged in the task, i.e. when they ran without licking. Animals were rewarded for correct licks with ~3 μl water using a solenoid valve (161T010; Neptune Research, USA), and licks were monitored using a custom device that detected breaks in an infrared beam.

To train mice to perform the task we used a progressive training procedure which lasted approximately 4 to 5 weeks. Initially, animals were allowed to familiarise with running on the wheel while head-fixed. At this stage animals were rewarded if they moved for a few seconds. After mice learned to control the wheel, they were rewarded every time they passed through the reward zone. After this, we introduced trials where the animal was rewarded only when it actively licked in the rewarded region of the corridor without 'punishing' the animal with a time-out if it licked at the wrong place ('bad licks'). In the

next stages we progressively reduced the number of bad licks down to 6. Animals were considered to have learned the task if they were able to perform at least 50 correct trials within a session.

## 2.5 Pupil tracking

We tracked the eye of the animal using an infrared camera (DMK 21BU04.H, Imaging Source) and a zoom lens (MVL7000, Navitar) at 25 Hz. Pupil position and size were calculated by fitting an ellipsoid to the pupil for each frame using a custom software. X and Y positions of the pupil were derived from the centre of mass of the fitted ellipsoid.

## 2.6 Imaging techniques

### 2.6.1 Widefield calcium imaging

For widefield imaging we used a standard epi-illumination imaging system (Carandini et al., 2015; Ratzlaff and Grinvald, 1991) together with an SCMOS camera (pco.edge, PCO AG). A Leica 1.6x Plan APO objective was placed above the imaging window and a custom black cone surrounding the objective was fixed on top of the headplate to prevent contamination from the monitors' light. The excitation light beam emitted by a high-power LED (465 nm LEX2-B, Brain Vision) was directed onto the imaging window by a dichroic mirror designed to reflect blue light. Emitted fluorescence passed through the same dichroic mirror and was then selectively transmitted by an emission filter (FF01-543/50-25, Semrock) before being focused by another objective (Leica 1.0 Plan APO objective) and finally detected by the camera. Images of 200 x 180 pixels, corresponding to an area of 6.0 x 5.4 mm were acquired at 50 Hz.

### 2.6.2 Two-photon imaging

Two-photon imaging was performed with a standard multiphoton imaging system (Bergamo II; Thorlabs) controlled by ScanImage4 (Pologruto, Sabatini, and Svoboda, 2003). A 970 nm laser beam, emitted by a Ti:Sapphire Laser (Chameleon Vision, Coherent), was targeted onto L2/3 neurons through a 16x water-immersion objective (0.8 NA, Nikon). Fluorescence signal was transmitted by a dichroic beamsplitter and amplified by photomultiplier tubes (GaAsP, Hamamatsu). The emission light path between the focal plane and the objective was shielded with a custom-made plastic cone, to prevent contamination from the monitors' light. In each experiment, we imaged 4 planes set apart by 40 µm. Multiple-plane imaging was enabled by a piezo focusing device (P-725.4CA

PIFOC, Physik Instrumente), and an electro-optical modulator (M350-80LA, Conoptics Inc.) which allowed adjustment of the laser power with depth. Images of 512x512 pixels, corresponding to a field of view of 500x500 μm, were acquired at a frame rate of 30 Hz (7.5 Hz per plane).

## 2.7 Pre-processing of imaging data

Pre-processing of raw imaging movies was done using the Suite2p pipeline (Pachitariu et al., 2016) and involved: 1) image registration to correct for brain movement, 2) ROI extraction, i.e. cell detection, 3) correction for neuropil contamination and 4) spike deconvolution.

To manually curate Suite2p's output we used two criteria: one anatomical and one activity-dependent criterion. One of the anatomical criteria in Suite2p is 'area', i.e. mean distance of pixels from ROI centre, normalised to the same measure for a perfect disk. We used this criterion (area < 1.04) to exclude ROIs likely to correspond to dendrites rather than somata. The activity-related criterion is the standard deviation of the cell trace, normalised to the standard deviation of the neuropil trace. We used this criterion to exclude ROIs whose activity was too small relative to the corresponding neuropil signal (typically with std(neuropil corrected trace) / std(neuropil signal) < 2). We finally excluded cells with extremely seldom firing (once or twice within a 20 mins session).

### 2.7.1 Neuropil correction

For neuropil correction, we used an established method (Dipoppa et al., 2018; Peron et al., 2015). We used Suite2p to determine a mask surrounding each cell's soma, the 'neuropil mask'. The inner diameter of the mask was 3 μm and the outer diameter was < 45 μm. For each cell we obtained a correction factor, α, by regressing the binned neuropil signal (20 bins in total) from the 5th percentile of the raw binned cell signal. For a given session, we obtained the average correction factor across cells. This average factor was used to obtain the corrected individual cell traces, from the raw cell traces and the neuropil signal, assuming a linear relationship. All correction factors fell within 0.7 and 0.9.

### 2.7.2 Deconvolution

Deconvolution was performed with Suite2p using an L0-based deconvolution approach (Pachitariu, Stringer, and Harris, 2017). Using this method, neuropil correction and estimation of spike trains were estimated simultaneously, using a model of the form (Pachitariu et al., 2016):

$$\boldsymbol{F_i}(t) = [\boldsymbol{s_i} * \boldsymbol{k_i}](t) + c_i \boldsymbol{N_i}(t) + noise\ ,$$

Where $\boldsymbol{F_i}$ is the raw fluorescence trace of ROI $\boldsymbol{i}$, $\boldsymbol{N_i}$ is the neuropil trace obtained by averaging pixels in a 'donut'-like area around the ROI, $c_i$ is the neuropil scaling coefficient and $\boldsymbol{s_i} * \boldsymbol{k_i}$ is the convolution of the spike train $\boldsymbol{s_i} \geq 0$ with the calcium response kernel $\boldsymbol{k_i}$

$c_i$ and $\boldsymbol{s_i}$ are then estimated iteratively using the cost function:

$Cost(\boldsymbol{s_i}, c_i) = \|\boldsymbol{F_i} - \boldsymbol{s_i} * \boldsymbol{k_i} - c_i \cdot \boldsymbol{N_i}\|^2 + \lambda \cdot L(\boldsymbol{s_i})$ , where $L(\boldsymbol{s_i})$ is the $L_0$ norm.

The iterative approach involved two steps: first, $c_i$ is optimised with simple regression while keeping $\boldsymbol{s_i}$ fixed; second, $\boldsymbol{s_i}$ and $\boldsymbol{k_i}$ are optimized with a spike deconvolution method while keeping $c_i$ fixed.

The spike deconvolution method was non-negative deconvolution using exponential kernels with a kernel decay similar to the decay time of the Calcium indicator GCaMP6s (2s) and an $L_0$ penalty, i.e . a term that penalises non-zero values.

## 2.8 Analysis of responses to drifting gratings

### 2.8.1 Pre-processing

Neuropil-corrected single-cell traces obtained from 4 planes were linearly interpolated to match the sampling rate of the imaging session (30 Hz). Single-cell activity was then baseline-corrected by subtracting the mean activity, and averaged across the time-window when the stimulus was on.

### 2.8.2 Skewness

Since skewness is measure of the asymmetry from the mean of a distribution, we used this measure to determine whether a cell was active or not. For each cell, we computed skewness, s, from its neuropil-corrected or deconvolved fluorescence trace based on the standard formula:

$$s = \frac{E(F-\mu)^3}{\sigma^3}$$

Where $F$ is the neuropil-corrected or deconvolved fluorescence trace, $\mu$ is $F$'s mean, $\sigma$ is $F$'s standard deviation, and $E$ is the expected value of the quantity in brackets.

### 2.8.3 Direction selectivity

Stimulus-evoked responses were obtained by averaging across repeats and by scaling by the maximum response value. Each one of these scaled responses corresponded to the magnitude of a 'response vector', that points towards one of the 12 possible directions (step 30 deg). From these response vectors we obtained the vector sum of the responses, whose normalised length corresponds to the magnitude of direction selectivity, and its direction to the preferred direction of the cell.

To obtain direction tuning curves, for each cell we fit (using least squares) a model function to the average activity for each direction. The model function was the sum of two Gaussians that meet at the peak. We used this fitting approach to determine 4 parameters: the preferred direction; the height of the response to the preferred direction; the height of the response to the non-preferred direction; and the tuning width.

### 2.8.4 Spatial-temporal frequency protocol

To select for responsive cells, we compared the mean activity across repeats during the stimulation period to the mean activity during blank trials. We selected cells whose maximum activity during at least one stimulus type was two standard deviations higher than the mean activity during blank periods. We defined as the cell's preferred spatial and temporal frequency the frequency at which the cell fired maximally, and as the cell's preferred speed the division of its preferred temporal frequency by its preferred spatial frequency.

## 2.9 Analysis of responses in the linear corridor

### 2.9.1 Response profiles as a function of position

To obtain response profiles as a function of position along the corridor, we discretised the position of the animal in 1 cm bins, yielding 100 bins and we calculated the spike count map and occupancy map for each neuron. Both the spike count and occupancy maps were smoothed by convolving them with a fixed Gaussian window (5 cm width for neuropil-corrected fluorescence; 4 cm width for deconvolved traces). Only time points with running speeds greater than 1 cm/s were included in further analyses. We only looked at responses for which the cross-validated reliability (see below) was higher than 0.01. These cells were considered to have activity significantly modulated by position in the corridor. To model single-cell activity under the assumption that responses are identical in the two segments of the corridor, we fit (using least squares) a model function to the response

profile along the visually-matching segment where the cell peaked. The model function was the sum of two Gaussians that meet at the peak. To obtain a prediction along the whole corridor, we then duplicated the fitted response at ± 40 cm away from the maximum.

Spatial modulation ratios and spatial modulation indices were measured by splitting the dataset into odd and even trials. For each cell, the position of the peak response was measured from the response profile averaged across odd trials ('preferred position'). We then obtained the response height at the preferred position and the visually-identical position 40 cm away ('non-preferred position'), using the response profile averaged across even trials. Cells which had a maximal response too close to the start or the end of the corridor (0-15 cm or 85-100 cm) were not considered for analysis of the ratio of responses. Therefore, this excluded cells which responded too close to the start or the end of the corridor, which were outside the visually-matching segments.

For the spatial modulation ratios, we computed the ratio of peak height at the non-preferred position over the peak height at the preferred position:

$$ratio = \frac{peak_{non-preferred}}{peak_{preferred}}$$

Therefore, $ratio = 0$ means response with one peak, whereas $ratio = 1$ means response with two peaks.

For the spatial modulation indices, SMI, we used the standard formula:

$$SMI = \frac{peak_{preferred} - peak_{non-preferred}}{peak_{preferred} + peak_{non-preferred}}$$

Therefore, $SMI = 0$ means response with two peaks, whereas $SMI = 1$ means response with one peak. Note that this is the opposite than the spatial modulation ratio.

To obtain response profile averages across the population with the highest secondary peaks possible, we ignored single-cell activity at the start and end of the corridor (0-15 cm or 85-100 cm).

### 2.9.2 Two-dimensional response profiles and marginals

Two-dimensional response profiles with respect to position and/or virtual and running speed were calculated as previously described (Saleem et al., 2013).

Let $y_j(t)$ be the (deconvolved) firing rate measured at time t for a given trial j smoothed with a 200 ms Gaussian smoothing window, and $p_j(t)$ the independent, continuous variable of interest (position, virtual speed, run speed). $p_j(t)$ was first discretized into n bins for all trials j=1,2,...,m (100 bins for position, 1 cm bin width; 30 bins for speed, excluding stationary periods, i.e. speeds less than 1cm/s). We then computed the spike count map and occupancy map. Each point in the spike count map was defined as the sum of spikes falling in the i-th bin and across all trials m:

$$S_i = \sum_{j=1}^{m} \sum_{t:(p(t)=p_i)} y(t), \;\; p_i = p_1 \ldots p_n,$$

each point in the occupancy map was the total number of values, $k_{ij}$, of the independent variable falling within the i-th bin in the j-th trial:

$$\Phi_i = \sum_{j=1}^{m} k_{ij}$$

Both $S_i$ and $\Phi_i$ were further convolved with an optimal Gaussian window, $\eta_p(0,\sigma)$, of the same width, σ, to give rise to the smoothed discretized signal and occupancy maps. The optimal smoothing window was the one that maximised the cross-validated prediction quality (see below).

We used the two smoothed quantities to obtain an estimate of each neuron's activity as a function of the variable p, $r_p$

$$r_p = \frac{S * \eta_p(0,\sigma)}{\Phi * \eta_p(0,\sigma)}$$

were '*' denotes convolution.

We computed 2-dimensional maps in a similar manner, using two independent Gaussians for smoothing each independent variable, p and q, and 2-dimensional convolution for the estimation of $r_{p,q}$.

To obtain 'separable' 2-dimensional maps we applied singular value decomposition on $r_{p,q}$.

The separable model was defined as:

$$R_{sep} = \boldsymbol{U_1} \sigma \boldsymbol{V_1}$$

where $\boldsymbol{U_1}, \boldsymbol{V_1}$ are the left-singular and right-singular vectors respectively, corresponding to the largest singular value $\sigma$.

### 2.9.3 Response profile prediction quality

Response profiles were estimated based on 80% of the data (shuffled trials) and tested on the remaining 20%. (5-fold cross-validation). For each estimate, a prediction quality, Q, was calculated as the fraction of variance in firing rate explained by the response profile:

$$Q = 1 - \frac{\sum_t (y(t) - y'(t))^2}{\sum_t (y(t) - \mu)^2}$$

where $y(t)$ is the actual, smoothed firing rate of the neuron at time *t*, $y'(t)$ was the estimated firing rate for the same time bin and $\mu$ is the mean firing rate of the training data. Only neurons with a reliability greater than 0.01 were considered for further analysis

## 2.10 Data analysis in the circular maze

### 2.10.1 Response profiles

By inspecting the deconvolved traces obtained from Suite2p, we realised that these had the shape of an exponential function. Thus, to obtain spike counts we took the logarithm of all peaks in the deconvolved trace and rounded them to the closest integer.

Response profiles were estimated as described in **2.9.2**. The total number of position bins was 100 and the bin size 2 cm. Firing rate was smoothed in time with a 250 ms Gaussian window. Smoothing in space of the spike count and occupancy map was performed with a fixed, 4 cm, Gaussian window. The number of cross-validation folds was increased to 20. This enabled estimation of the s.e.m. using Jacknife resampling.

We defined as responsive, cells whose maximum firing was higher than 2 s.e.m. over the mean firing rate in the medium gain condition. To ensure that cells exhibited a similar degree of responsiveness across conditions (low, medium and high gain), we calculated the Pearson cross-correlation coefficient between responses in different conditions and selected response profiles with a correlation coefficient higher than 0.75. The cross-correlation function was further used to determine the shift (mean and s.e.m.) in the response profile at low or high gain relative to the medium gain condition based on the position of the maximum peak in the cross-correlogram in each fold of the cross-validation procedure.

### 2.10.2 Lick distributions

To obtain distributions of licks across correct trials, we considered the position of the first correct lick and computed 'lick maps' as a function of position on the maze as previously described in **2.9.2**. (smoothing in space with a 4cm Gaussian). The average position of first licks was estimated by computing the circular average of the lick distribution. The average position in low or high gain was then compared to the average position in medium gain, to obtain the shift in the lick distribution as a function of gain change.

## 2.11 General linear model analysis

### 2.11.1 Response kernels to stimuli of different orientations

To predict responses to drifting gratings of different orientations (12 orientations) we used ridge regression with a fixed ridge regression coefficient, λ = 0.01. For each orientation we defined a sparse column vector in time, whose values were zero except from the timepoints when the corresponding stimulus appeared on the screen. This vector was then shifted backward in time by half a second and forward in time by 3.5 s in steps of 0.03 s, resulting in a stimulus matrix with 120 columns. Stimulus matrices for each orientation were pulled together, yielding the design matrix, with dimensions (# of frames) X (120 shifts x 12 orientations). We used this matrix to predict absolute fluorescence using 5fold cross-validation.

### 2.11.2 Modelling response profiles in virtual reality

To assess the spatial modulation of V1 neurons while jointly accounting for all other visual and behavioural factors, we fitted the V1 responses to four different general linear (or multilinear ridge regression) models of the form: $\hat{y} = X\hat{\beta}$, where $X$ is an T-by-M matrix with T time points and M predictors, $\hat{y}$ is the predicted calcium trace (T-by-1 array). Optimal coefficient estimates $\hat{\beta}$ (M-by-1 array) that minimise the sum squared error were obtained using: $\hat{\beta} = (X^T X + \lambda I)^{-1} X^T y$, where $\lambda$ is the ridge-regression coefficient.

In the full model, the predictor matrix $X$ contains several sets of columns: a set of spatial basis functions $I_i(x_t)$; pupil position $ex_t, ey_t$; the speed $s_t$ at 5 time lags; pupil diameter $p_t$ again at 5 time lags; and a step function $r_t$ indicating reward with 4 time lags. A model using all these predictors has the form:

$$y_t = \beta_0 + \sum_{i=1}^{12} \beta_i I_i(x_t) + \beta_{ex} ex_t + \beta_{ey} ey_t + \sum_{j=1}^{5} \beta_{s_j} s_{t+\tau_j} + \sum_{j=1}^{5} \beta_{p_j} p_{t+\tau_j} + \sum_{k=1}^{4} \beta_{r_k} r_{t+\tau_k}$$

where $\beta_0$ is a constant.

Predictors were defined similarly to Driscoll et al (Driscoll et al., 2017). The first spatial basis functions corresponded to regions prior to the visually identical segments:

$$I_1(x) = \begin{cases} 1 \; if \; x \in [0,5] \\ 0 \; otherwise \end{cases}$$

$$I_2(x) = \begin{cases} 1 \; if \; x \in [5,10] \\ 0 \; otherwise \end{cases}$$

The basis functions corresponding to visually matching segments consisted of double step functions with weights $a$ and $b$ (defined below):

$$I_3(x) = \begin{cases} a \; if \; x \in [10,15] \\ b \; if \; x \in [50,55] \\ 0 \; otherwise \end{cases}$$

Basis functions $I_4 \ldots I_{10}$ were defined similarly, to cover the x range 15… 50 and 55…90. The final two functions $I_{11}$ and $I_{12}$ covered the non-repeating x ranges [90,95] and [95,100].

The pupil diameter and eye position predictors were all scaled to lie in a range [-1,1], while speed was scaled to the range [0,1]. The pupil diameter and speed predictors were lagged with 5 time shifts: $\tau_j \in \{-1000\ ms\ , -500\ ms\ , 0\ ms, 500\ ms, 1000\ ms\}$ . The reward predictor $r_t$ was defined to be 1 within a 500 ms window of the reward, 0 otherwise, and was lagged by $\tau_k \in \{-1000\ ms\ , -500\ ms\ , 0\ ms, 500\ ms\}$.

The four prediction models used different combinations of these predictors. The first model, *the visual model*, relied on just visual predictors $I_i$, with the constraint that $a = b = \sqrt{2}$ such that the basis functions have unit norm. Therefore, this model resulted in responses that repeated in the visually identical segments perfectly.

The second model, *speed* model, added the influence of running speed $(s_{t+\tau_k})$.

The third model, *the non-spatial model*, added the influence of all other behavioural factors we measured: running speed ($s_{t+\tau_k}$), reward events ($r_{t+\tau_k}$), pupil size ($p_{t+\tau_k}$), and the horizontal ($ex_t$) & vertical ($ey_t$) pupil position.

Finally, the fourth model, *spatial model*, allowed for an independent scaling of the two identical sections of the room. This model allowed $a \neq b$, subject to the constraint that the spatial basis functions had unit norm. To achieve this, we used exhaustive search over a parameter $\alpha \in [0,1]$, with $a = \frac{\alpha}{\sqrt{\alpha^2+(1-\alpha)^2}}$ and $b = \frac{1-\alpha}{\sqrt{\alpha^2+(1-\alpha)^2}}$. Note that $\alpha = 0.5$ would correspond to a purely visual representation with spatial modulation ratio close to 1, while $\alpha = 1$ or $\alpha = 0$ would correspond to a response only in the first or second segment, and a spatial modulation ratio close to 0.

To fit the parameters, we used the ridge regression coefficient, λ that maximized the percentage of variance explained using five-fold cross-validation, searching the values $\lambda$ =0.01, 0.05, 0.1, 0.5 or 1. In the spatial model (where $a \neq b$), we performed multiple ridge regression fits, searching for the optimal value of $\alpha$ using a step size of 0.1, for each λ.

The single cell responses predicted by these models were then processed similarly to the original recorded responses to obtain the response profiles and spatial modulation ratio predicted by the three models. The deviation of the model predictions from the original data were evaluated by fitting an ellipse to the distribution and quantified using the angle of its major axis (the first eigenvector of the covariance matrix).

## 2.12 Decoding population activity

Population activity was decoded using an independent Bayes decoder (Zhang, Ginzburg, McNaughton, and Sejnowski, 1998) under the Poisson assumption. Specifically, for every time bin, we estimated the probability of being at position $x$ given responses across the population $R$:

$$P(x|R) = \frac{1}{Z}\, P(x) \left( \prod_{i=1}^{M} f_i(x)^{r_i} \right) exp\left( -t \sum_{i=1}^{M} f_i(x) \right)$$

Where $f_i(x)$ is the response profile, $r_i$ is the spike count of the $i^{th} (i = 1 \ldots M)$ neuron in a time bin. $Z$ is a normalising constant, which makes the probabilities across positions sum to one.

To account for speed differences as a function of position, the data in medium gain were divided into 3 speed ranges based on the medium speed at each position. Decoding was then performed independently for each speed range. The decoded posteriors obtained

from different speed ranges were pulled together, yielding the decoded posterior as a function of time. To avoid overfitting, we used 20-fold cross-validation.

To obtain the average posterior as a function of position we used a similar approach as in **2.9.2**. The 'posterior' map and occupancy map were smoothed in space with a 4cm Gaussian filter.

The mean and s.e.m. of the decoding error was obtained by aligning the average posterior to the actual position of the mouse. The s.e.m. was obtained from the estimates of the 20-fold cross-validation using Jacknife resampling.

To obtain the mean decoding error for each session, the error map was first averaged across positions. We then excluded values < 1 and took the circular average of the resulting distribution.

# Chapter 3

The mouse primary visual cortex (V1) codes not just visual information, but also carries non-visual signals that reflect the animal's behaviour. The influence of non-visual signals on V1 activity has been demonstrated in a variety of active behaviours, from simple running to more complex learning tasks (see General introduction). Yet, we do not know whether V1 responses are also influenced by signals related to one's position in the environment. The most relevant studies so far have provided valuable insight into the comodulation of visual cortex and hippocampus during memory replay or spatial behaviour, and have suggested that spatial position may be influencing responses as early as in V1 (Ji and Wilson, 2007). However, these studies lack precise control of the visual environment and therefore cannot draw firm conclusions on whether V1 indeed carries non-visual signals related to spatial position.

In this chapter I will ask whether neural activity in V1 is influenced by position-related signals. To address this question, we carefully designed a virtual environment with two visually-matching segments. Given the visually-matching segments, we hypothesised that a purely visual neuron which responds in the first segment should have a similar response in the second segment. Instead, we find that the majority of V1 neurons respond more strongly, or even exclusively, at a specific location in the virtual environment. To establish whether the observed modulation is related to the virtual position of the animal rather than any other behavioural or visual factor, we performed a series of control experiments. These control experiments demonstrated that the modulation of visual responses cannot be explained by other task-related features, such as running speed or reward. It also could not be explained by visual features, such as pupil size or eye position. To assess the relative contribution of spatial position together with these task-related and visual features, we developed a ridge-regression model that takes all these factors into account simultaneously.

We found that the difference between responses to the visually-matching segments can be best explained by spatial position only.

Given these findings, I conclude that modulation by spatial position is present as early as in V1, and propose that cognitive representations of space may well shape early visual processing.

## 3.1 Results

### 3.1.1 Spatial modulation of V1 responses

To assess the influence of spatial position on V1 responses, we used 2-photon calcium imaging to record neural activity in V1, while head-restrained mice ran along a corridor in virtual reality (**Figure 3-1a**). The corridor was 100 cm long and contained two prominent landmarks repeated twice (**Figure 3-1b**); a vertical grating placed at 20 cm was followed by a plaid at 40 cm, then a second vertical grating was placed at 60 cm and was followed by a second plaid at 80 cm. Thus, each landmark was repeated after 40 cm. This repetition resulted in two visually-matched segments of the corridor, one from 10 cm to 50 cm and another one from 50 cm to 90 cm (**Figure 3-1b, c**). Despite the corridor having two visually-matching segments, visual scenes analysed by neurons with central receptive fields would still be different between the first and second segment. This is mainly due to the end wall of the corridor, which appears as an expanding grey square as the animal advances forward. Instead, visual scenes analysed by neurons with receptive fields in the periphery are similar (**Figure 3-1c**). Therefore, we focused our analysis on those portions of medial V1 with receptive field centres >40 deg azimuth. To map these portions of V1 we performed wide-field imaging and obtained maps of retinotopy in transgenic mice expressing the Calcium indicator GCaMP6 in all excitatory neurons (**Figure 3-1d**). We then placed these mice in virtual reality and used their maps of retinotopy to target specific patches of medial V1 across multiple days using two-photon Calcium microscopy (**Figure 3-1e**). Thus, this chronic imaging method allowed us to measure the activity of 8,610 neurons in medial V1 in 4 mice across 18 sessions.

At this point, we set off to characterise visual response profiles in the virtual corridor. Given the repetition of the visual scenes in the two corridor segments, we found, as expected, cells responding to their preferred stimulus, grating or plaid, equally (**Figure 3-2, 1st column**). These responses matched exactly our predictions based on the repetition of

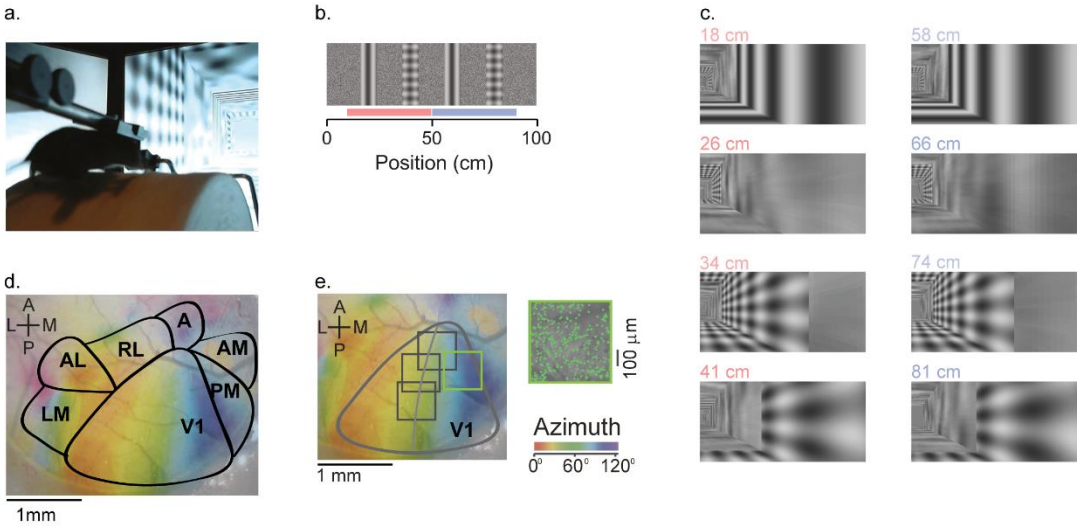


**Figure 3-1: Imaging in Virtual Reality**
**a.** Head-fixed mice ran on a cylindrical treadmill to navigate a virtual corridor presented on three visual displays.
**b.** The virtual corridor had two landmarks (a grating and a plaid) that repeated after 40 cm, creating two visually-matching segments (*red* and *blue bars*).
**c.** Example screenshots, showing the visual similarity of the virtual corridor at two pairs of positions spaced 40 cm apart.
**d.** Example retinotopic map of the cortical surface acquired with widefield calcium imaging. The border of V1 and higher visual areas is defined by inversion of the retinotopic map.
**e.** Same as in **d.** with *squares* denoting the field of view in two-photon imaging sessions targeted to medial V1 in an animal (field of view with *green frame* is shown in the *inset*). Within these sessions we analysed responses from neurons with receptive field centre > 40° azimuth (*curve*).

visual scenes 40 cm away (**Figure 3-2a, dotted line**). We also found cells with weaker responses in the second segment than the first segment, which is what one might expect due to adaptation phenomena (**Figure 3-2a, 2nd column**). However, we also found responses that could not be explained based on pure vision; some cells responded more strongly in the second segment than the first segment (**Figure 3-2a, 3rd column**); even more strikingly, other cells responded only once in the virtual corridor (**Figure 3-2a, 4th column**). This initial observation indicates that the visual stimulation induced in virtual reality gives rise to diverse response profiles, and that notably, the modulation observed in some response profiles cannot be explained based on pure vision.

How evident is the observed modulation across the neural population? To answer this question, we selected response profiles for which variance explained was at least 1% (n = 4,958) and ordered them as a function of the position where each neuron fired maximally. We found that the responses across the population tiled up the whole virtual corridor (**Figure 3-2b**). In addition, most neurons preferred to fire more strongly in one position than in the visually matching position 40 cm away. To assess the robustness of this

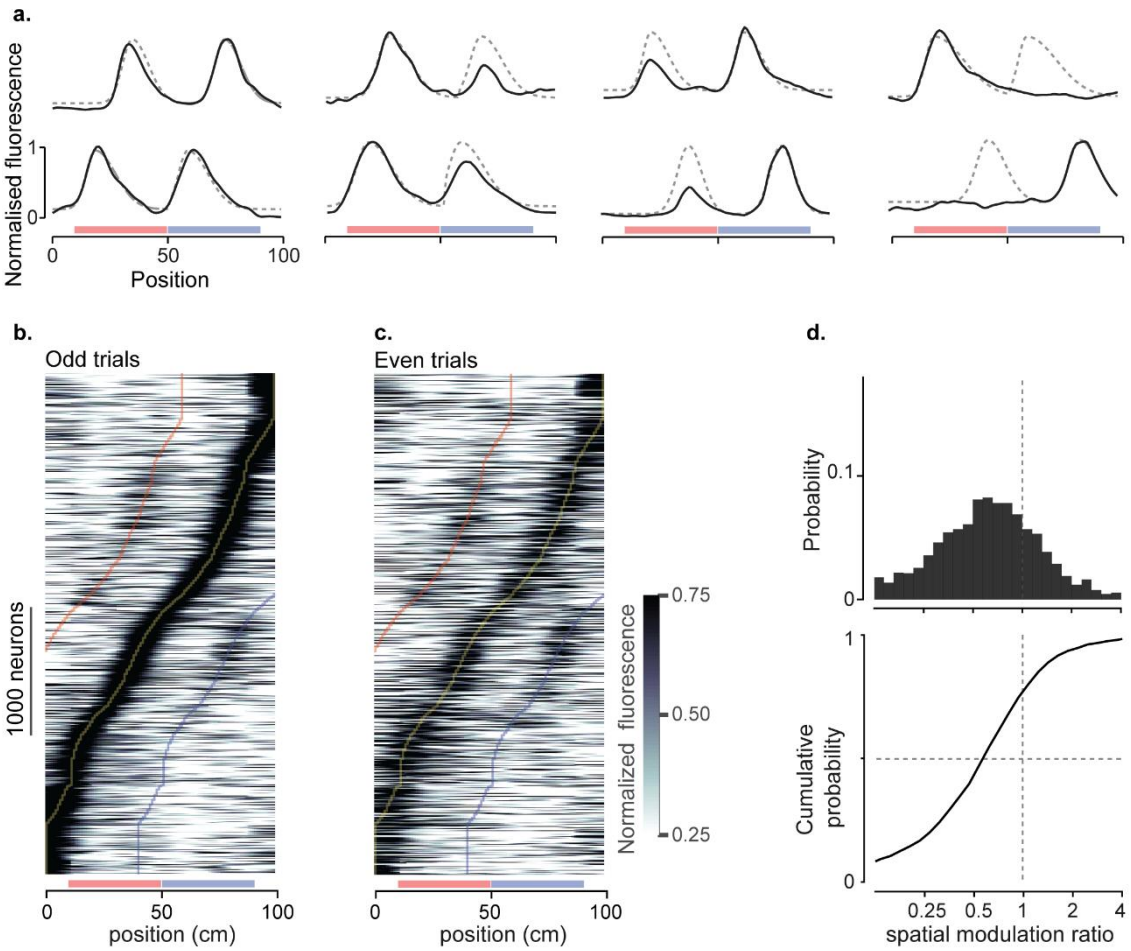


**Figure 3-2: Responses in the primary visual cortex (V1) are modulated by spatial position.**
**a.** Normalised response as a function of position in the virtual environment for eight example V1 neurons. *Dotted lines* show predictions if the responses were identical in the two segments of the corridor (*red* and *blue bars*).
**b.** Normalised response as a function of distance in the virtual environment obtained from odd trials only, for V1 neurons with receptive field centres >40° azimuth and activity significantly modulated by position in the corridor (4,958 of 8,610 neurons; 4 mice). Neurons are ordered based on the position of their maximum response. *Yellow*, *red* and *blue lines* indicate position of maximum +/- 40 cm. *Red* and *blue bars* are as in **a**.
**c.** Same as **b**, for the half of the data (even trials) that were not used to order the responses. Sequence and scaling of the responses are the same as in **b**.
**d.** Histogram (*top*) and cumulative distribution (*bottom*) of the spatial modulation ratio (the ratio of secondary response 40 cm away from peak response, *red* or *blue line* in **b** and **c**, divided by peak response, *yellow line* in **b** and **c**), derived from even trials.

preference for one position, we divided the data between odd and even trials. We used the odd trials to order the response profiles and asked whether the same ordering was preserved in even trials (**Figure 3-2b**, **c**). We found that most neurons exhibited a reliable preference for one position along the corridor in both odd and even trials. To further quantify this preference, we focused on cells that fired within the two visually-matching segments (n = 2,422). Their responses to the landmarks 40 cm from the preferred position

were substantially smaller than the responses at the preferred position (**Figure 3-3**). For these cells we also estimated the 'spatial modulation ratio' i.e. the ratio of the activity 40 cm from the preferred position over the activity at the preferred position. Thus, a ratio close to 1 indicates that a cell responds similarly to the landmarks, whereas a ratio close to 0 indicates strong preference for one position in the corridor. We found that the corresponding distribution of ratios was significantly biased towards lower values and that the median ratio was 0.61 ± 0.31 (median ± m.a.d., in the cross-validated trials). This value was significantly < 1 ($p < 10^{-104}$, Wilcoxon signed rank test). (**Figure 3-2d**). These results indicate a widespread preference of V1 neurons for one location in the virtual environment, thereby suggesting that V1 responses are influenced by spatial position.

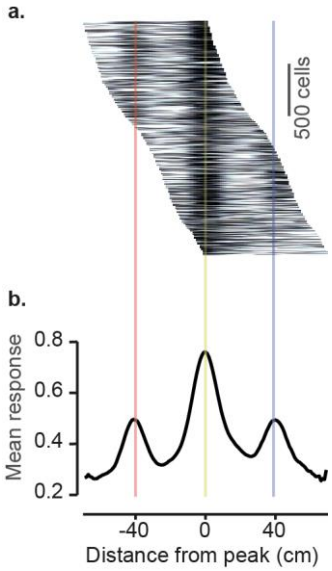


**Figure 3-3: Spatial averaging of visual cortical activity confirms the difference in response between visually matching locations.**
**a.** Mean response of V1 neurons as a function of the distance from the peak response location (2,422 cells with peak response between 15 and 85 cm along the corridor). To ensure that the average captured reliable spatially-specific responses, the peak response location for each cell was estimated from only odd trials, while the mean response was computed from even trials.
**b.** Population average of responses shown in **a.** Lower values of the side peaks compared to central peak indicates strong preference of V1 neurons for one segment of the corridor over the other visually-matching segment (40 cm away from peak response).

## 3.1.2 Modulation is not explained by visual or non-visual factors

Before concluding that V1 responses are indeed influenced by spatial position, it is necessary that we rule out any effects by other visual or task-related variables that are present in our behavioural paradigm. These variables are either related to the animal's behaviour, such as eye position, pupil size, running speed and reward, or to the design of the virtual environment itself, such as visual image, trial onset and trial offset. We first focused on variables related to the animal's behaviour and attempted to control for each one of those individually.

We first focused on the influence exerted by speed because we reasoned that the presumable spatial modulation may be merely due to differences in the speed profile as a function of position in the corridor. Such differences were particularly evident in rewarded

sessions for instance, when the animal had to decelerate to receive a reward (**Figure 3-4a right**). To quantify the effect of speed, we stratified the data into three speed ranges, low, medium and high and estimated responses as a function of position for each speed range (**Figure 3-4b, c**). From this analysis it became evident that some responses were not affected by speed (**Figure 3-4b, c** left panel), whereas others exhibited differences in their shape or height or in both (**Figure 3-4b, c** right panel). Therefore, consistent with previous studies, speed is influencing visual responses. But does speed also affect the ratio of peaks? To answer this question, we followed the same approach as previously, when we used the odd trials as the training set for sorting the responses and the even trials as the test set. This time the training set consisted of responses at medium speed ranges and the test set of

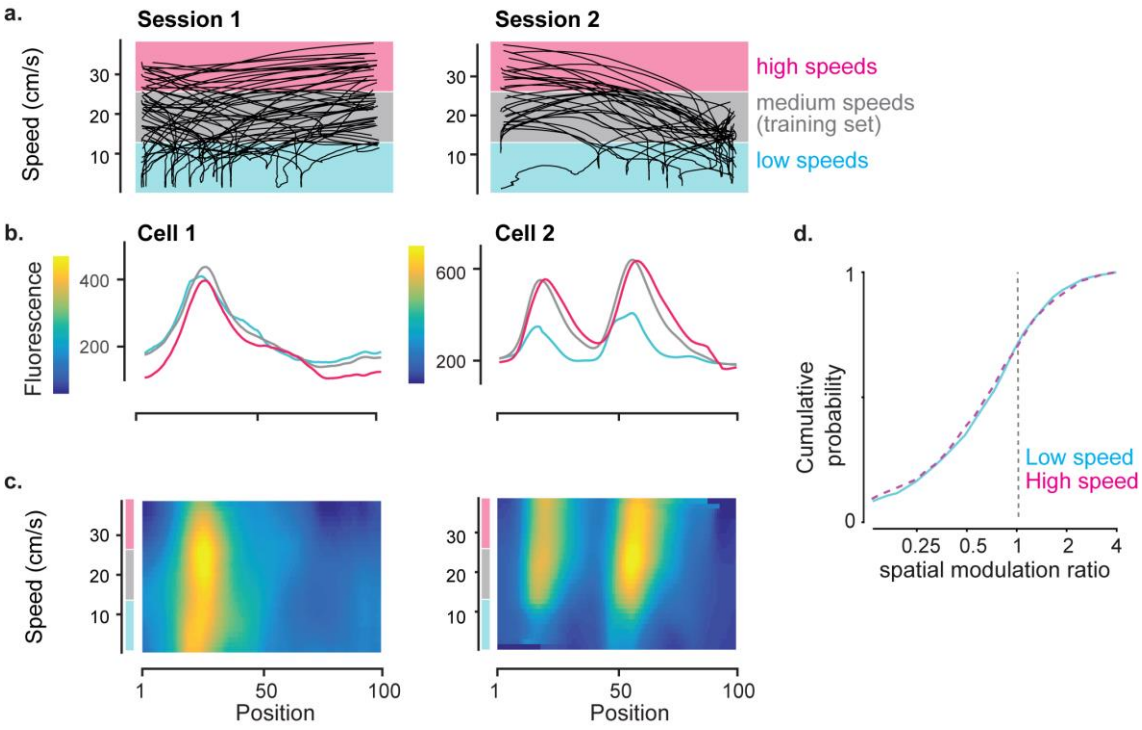


**Figure 3-4: The spatial modulation of V1 responses cannot be explained by speed.**
**a.** Speed-position plots for all single-trial trajectories in two example recording sessions.
**b.** Response profile of example V1 cells in each session as a function of position in the corridor, stratified for three speed ranges corresponding to the shading bands in **a**.
**c.** Two-dimensional response profiles of the same example neurons showing activity as a function of position and running speed for speeds higher than 1 cm/s.
**d.** Cumulative distribution of the spatial modulation ratio after stratifying the data by running speed. The two curves corresponding to low (*cyan*) and high (*purple*) speeds overlap and appear as a single dashed curve.

responses at low and high speed ranges. We first determined the position at which cells fired maximally based on the training set, and then obtained ratios of peaks for the two test sets. This analysis showed that there was no difference in the cumulative distributions of

ratios between low and high speeds (**Figure 3-4d**). In addition, even within a single speed category (medium speed), the spatial modulation ratio was substantially < 1 (0.47 ± 0.22; $p < 10^{-33}$). Therefore, spatial modulation of responses in V1 cannot be explained by speed.

Another factor that may contribute to the modulation of V1 responses is reward. Indeed, it has been previously shown that reward does modulate responses as early as in V1 (Poort et al., 2015; Shuler and Bear, 2006). Also, as it was mentioned earlier, the speed profile depends profoundly on reward locations, because the animal decelerates every time it receives a reward. To account for this factor, we took sessions in which animals ran freely along the corridor in the absence of reward (**Figure 3-5a**, 2 animals, 8 sessions). Again, we found that the distribution of spatial modulation ratios was strongly biased towards values <1 (0.57 ± 0.37, $p < 10^{-14}$). Thus, we concluded that the spatial modulation of V1 responses could not be explained by reward either.

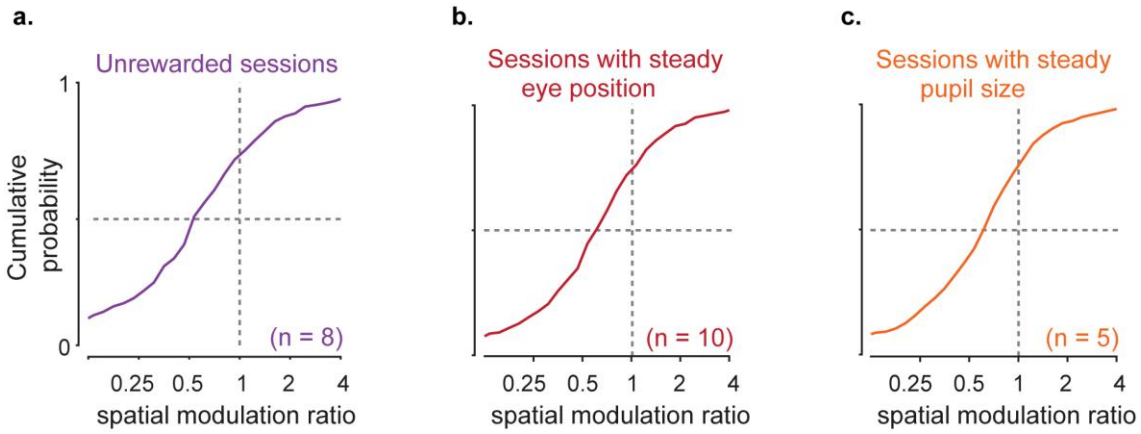


**Figure 3-5: The spatial modulation of V1 responses cannot be explained by reward, differences in eye position or in pupil size.**
**a.** Distribution of spatial modulation ratio for unrewarded sessions (8 sessions; median ± m.a.d.: 0.57 ± 0.37)
**b.** Distribution of spatial modulation ratio for sessions with steady eye position (10 sessions; median ± m.a.d.: 0.63 ± 0.33)
**c.** Distribution of spatial modulation ratio for sessions with steady pupil size (5 sessions; median ± m.a.d.: 0.63 ± 0.33)

Having established that the observed modulation could not be explained by task-related factors, we next asked whether visual factors, such as deviations in eye position or pupil size played a role. We estimated the ratio of peaks for sessions in which eye position (10 sessions) or pupil size (5 sessions) were on average constant along the corridor and found that the distributions were still biased towards lower values (stable size: 0.63 ± 0.33, $p < 10^{-45}$; stable position: 0.63 ± 0.33, $p < 10^{-27}$; **Figure 3-5b, c**). These controls indicate

that displacement of the eye from its central position or pupil size cannot account for the spatial modulation of V1 responses.

### 3.1.3 Relative contribution of visual or non-visual features

So far, I described control analysis that accounted for various confounding factors separately. However, this approach does not rule out the possibility that the modulation of visual responses is mediated by the joint contribution of all these variables, rather than by spatial position per se. In addition, this approach does not rule out the effect of confounds intrinsically linked to the virtual environment, such as the effect of trial onset and offset. To assess the relative contribution of these variables together with spatial modulation, we developed a set of ridge-regression models; these were encoding models of single-cell activity where each variable contributed at least one predictor.

The simplest model, the *visual model* took into account only responses to the visual scenes (which repeat twice), as well as the onset and offset of the trial (which introduce visual transients). Importantly, it enforces equal height of the responses to the two identical segments, thus it is more appropriate for capturing responses with two equal peaks. The second, *speed model*, additionally took into account running speed. The third, *non-spatial model*, additionally took into account reward times, pupil size, and eye position. Finally, the fourth, *spatial model*, additionally allowed visual responses to the two matching segments to differ in amplitude. Thus, it is more appropriate for capturing cells with unequal peaks.

We fit each model to the activity of each neuron and obtained predictions as a function of time (**Figure 3-6a**). Responses with two unequal peaks could be best predicted only by the *spatial model*. We then asked how well the models could predict the spatial modulation ratio. To answer this question, we did the same analysis as we did with the data previously: we transformed predictions as a function of time into smoothed predictions as a function of position (**Figure 3-6b**); we sorted these predictions based on the ordered preferred position of responses obtained from odd trials and we estimated the spatial modulation ratio, i.e. the difference between predicted activity at the preferred position versus predicted activity 40 cm away (**Figure 3-6c**). We found that the most accurate predictions of the ordered activity -and consequently of the spatial modulation ratio- was provided by the *spatial model* only. Indeed, when we plotted the predicted spatial modulation ratio against the measured ratio, we found that the correlation between the data and the model

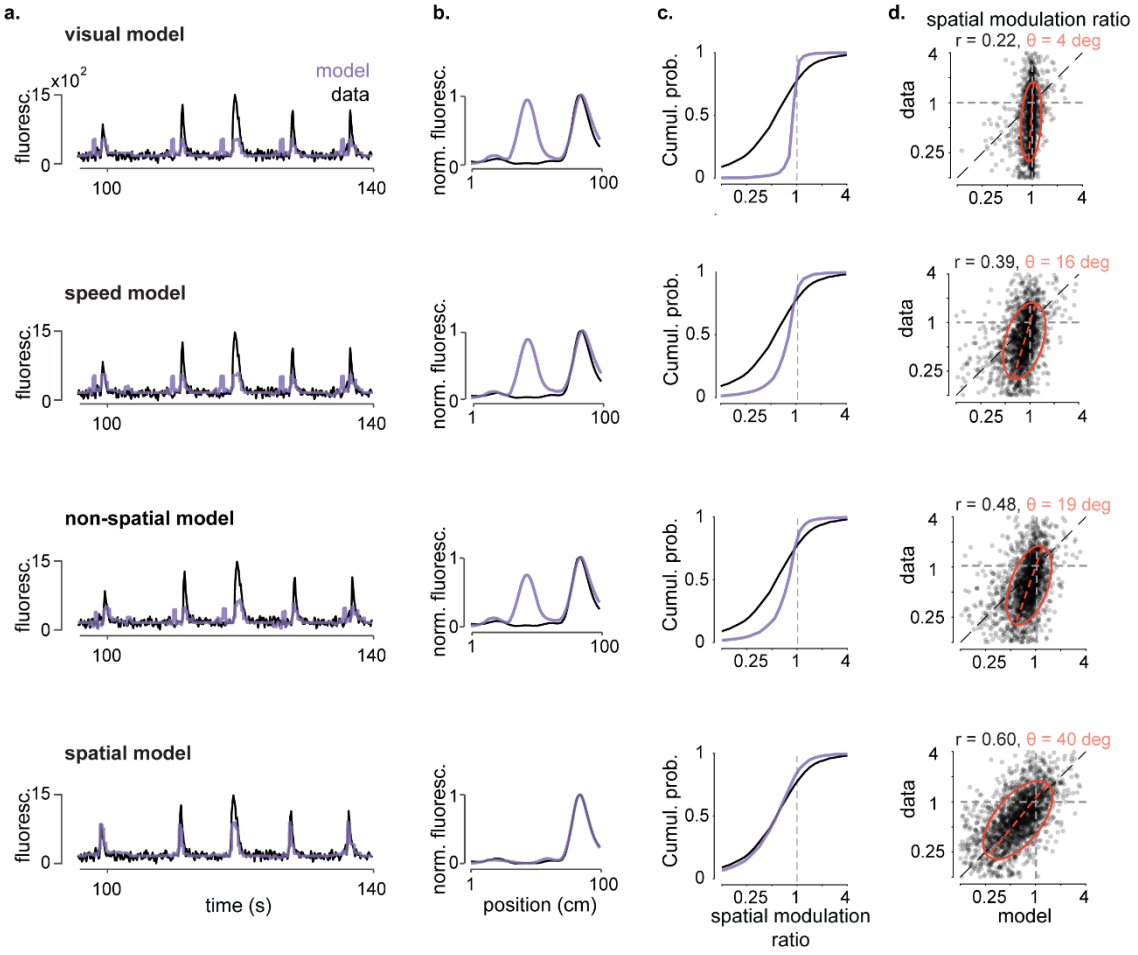


**Figure 3-6: Observed values of spatial modulation ratio can only be modelled using spatial position.**

**a-b.** We constructed four models to predict the activity of individual V1 neurons from successively larger sets of predictor variables. In the simplest, the 'visual' model, activity is required to depend only on the visual scene visible from the mouse's current location, and is thus constrained to be a function of space that repeats in the visually matching section of the corridor. The second 'speed' model also includes modulation by speed. Because speed can differ between the first and second halves of the track, modelled responses need no longer be exactly symmetrical; however, this model does not explicitly use space as a predictor. The third 'non-spatial' model also includes modulation by additional behavioural factors that can differ within and across trials: reward times, pupil size, and eye position. These variables can also differ between the first and second halves of the track; like the previous two models, this model does not explicitly use space as a predictor. The final 'spatial' model extends the previous model by also allowing responses to the two matching segments to vary in amplitude, thereby explicitly including space as a predictor. Example single-trial predictions are shown as a function of time in (a), together with measured fluorescence. Spatial profiles derived from these predictions are shown in (b).

**c.** Cumulative distributions of spatial modulation ratio for the four models (*purple curves*). For comparison, the *black curve* shows ratio of peaks derived from the data (even trials) (median ± m.a.d: visual model: 0.99 ± 0.03; Wilcoxon rank sum test: $p < 10^{-40}$, speed model: 0.87 ± 0.19; $p < 10^{-40}$, non-spatial model: 0.83 ± 0.18; $p < 10^{-40}$; spatial model: 0.60 ± 0.27; p = 0.09).

**d.** Measured spatial modulation ratio versus prediction of the 3 models. Each point represents a cell; *red ellipse* represents best fit Gaussian, *dotted line* measures its slope. The purely visual model (*top*) does poorly, and is only slightly improved by adding predictions from speed, reward, pupil size, and eye position (*middle*). Adding an explicit prediction from space provides a much better match to the data (*bottom*). r: correlation coefficient; θ: orientation of the major axis of the fitted ellipsoid.

increased as we went from the simpler*, visual model* towards the *non-spatial model* which took all measured variables into account (**Figure 3-6d;** Pearson correlation coefficient r, visual model: 0.22, speed model: 0.39, non-spatial model: 0.48, spatial model: 0.60). Yet, the model with the highest correlation to the data, was the *spatial model*. Fits of an ellipsoid to the data showed that the orientation of the ellipsoid approached the diagonal only for the *spatial model*, confirming that this is the model that best matches the data (orientation of the major axis θ, visual model: 4 deg, speed model: 16 deg, non-spatial model: 19 deg, spatial model: 40 deg). Overall, this modelling analysis reinforces the main finding that responses in V1 are modulated by spatial context, beyond the influence of other visual and non-visual factors.

### 3.1.4 Pixel-by-pixel identical visual environment

In the previous section I demonstrated that task-related or visual factors are not sufficient to explain the spatial modulation of V1 responses. However, this analysis does not account for one last confound which is related to the design of the virtual environment itself; that is the end wall of the corridor which is constantly expanding as the animal advances forward. The rationale behind this specific design was to create the sense of being inside a room of finite dimensions. But this choice came at a cost: that the visual scenes processed by cells with central receptive fields were not identical at 40 cm apart. As mentioned previously, due to this limitation we chose to analyse responses of cells with receptive fields >40 deg. However, although images 40 cm apart are identical 'within the receptive fields' of these neurons, images outside their receptive fields are not pixel-by-pixel identical.

To control for this limitation, we ran additional experiments in a visually-identical virtual environment. The environment was a corridor with twice the length of the original one. The landmark textures, a grating and a plaid, were now repeated four times in total (**Figure 3-7a**). In addition, the distance along the corridor that was visible to the animal was reduced to 70 cm, therefore keeping the first and second quarter of the room (every pixel) completely identical (**Figure 3-7b**). Animals ran from 0 cm to 100 cm of this corridor and, as before, after an inter-trial interval they started again at position 0 cm.

Recordings in the visually-identical environment were performed in 2 mice (7 sessions in total). As in the original experiments, we targeted portions of V1 that analysed the peripheral visual field. Even in the extended 'identical corridor' V1 responses were modulated by spatial context in both animals (**Figure 3-7c**). The spatial modulation ratio

was again substantially different to 1 (0.62 ± 0.26, $p < 10^{-81}$, Wilcoxon signed rank test; n = 1044 cells), confirming that spatial modulation of V1 responses could not be explained by visual cues even far outside the classical receptive field.

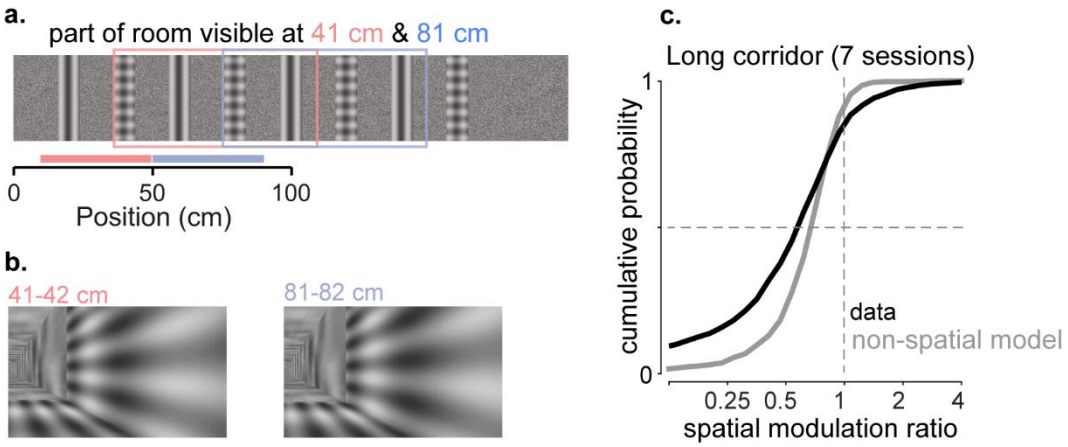


**Figure 3-7: The spatial modulation of V1 responses is not due to end-of-corridor visual cues**
In a virtual corridor of length 100 cm (**Figure 3-1b**) the scenes from the two repeated sections will differ slightly at the perspective vanishing point, as the grey wall at the end of the corridor will occupy more pixels in the second section. Even though this point is far from the receptive field of the measured neurons, it is important to exclude the possibility that this small visual difference could give rise to the position signals measured in V1. We therefore imaged two additional mice in a control experiment, where the on-screen view was absolutely identical in the two sections. This was achieved by creating a 200 cm corridor, "teleporting" the animal back to the start after traversing the first 100 cm, and setting the virtual reality software to provide visibility only up to 70 cm in the distance.
**a.** Diagram of the 200 cm virtual corridor, containing the same grating and plaid as the regular corridor, repeated 4 times instead of twice.
**b.** Visual scenes from locations within the first 100 cm of the extended corridor, separated by 40 cm, are visually identical.
**c.** Cumulative distribution of the spatial modulation ratio across the two mice that were placed in the long corridor (7 sessions, 2 mice; median ± m.a.d: 0.62 ± 0.26; *black line*). *Grey line* shows the spatial modulation ratio predicted by the "non-spatial" model (which predicts activity from the visual scene, trial onset and offset, speed, reward, pupil size and displacement from the central position of the eye, see Supplementary Figure 8 below, non-spatial model). The two distributions are significantly different (Wilcoxon rank sum test; $p<10^{-14}$).

## 3.1.5 Does spatial modulation depend on experience?

Taken together, these results support the hypothesis that responses as early as in V1 are modulated by spatial context. That being the case, it is intriguing to ask how such signals emerge. Do these signals become stronger as the animal becomes more familiar with the environment?

To answer this question, we trained 2 animals to run head-fixed in the absence of any visual stimulus, and sought to measure V1 activity the very first time they experienced the virtual environment. We monitored V1 activity in the periphery, in the same field of view,

and we found that responses on the first day were different from responses on the second day (**Figure 3-8**). Specifically, the secondary peaks 40 cm away were on average higher on the first day, and consequently the ratio of peaks was significantly higher than on the second day (**Figure 3-8b, c**; median ± m.a.d.: first day: 0.71 ± 0.28, second day: 0.56 ± 0.27; Wilcoxon rank sum test; p = 0.005). Importantly, the difference between the first and second day could not be attributed to differences in speed; one animal ran at 13 cm/s on average on the first day and at 15 cm/s on the second day; the second animal ran at 16 cm/s on the first day and at 20 cm/s on the second day. Overall, these preliminary results suggest that the observed spatial modulation in V1 becomes stronger over days. Yet, additional experiments are required to reveal how this spatial modulation is being inherited.

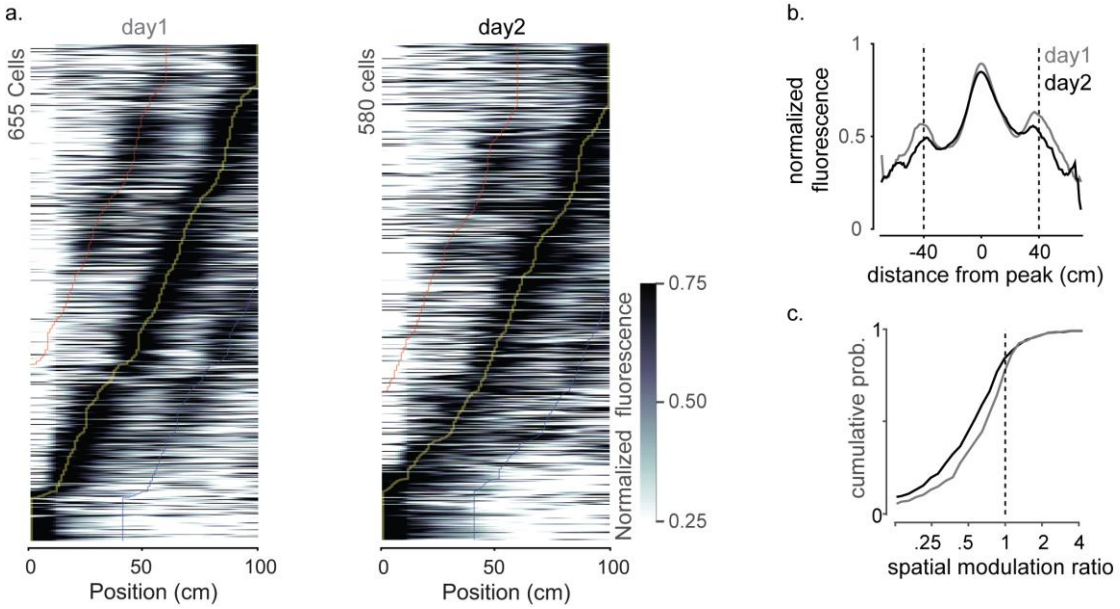


**Figure 3-8: Lower degree of spatial modulation on the first day of exposure to the virtual environment.**

**a.** Normalised response as a function of distance in the virtual environment obtained from even trials only on the first (*left*) and second (*right*) day in virtual reality, for V1 neurons in the same field of view in the periphery (receptive field centres >60° azimuth) and activity significantly modulated by position in the corridor (2 mice; first day: 655 of 755, second day: 580 of 662). Neurons are ordered based on the position of their maximum response. *Yellow*, *red* and *blue lines* indicate position of maximum +/- 40 cm.

**b.** Population average of responses shown in **a** for cells responding to the visually-matching segments obtained on the first day (*grey*) and second day (*black*) in virtual reality. Lower values of the side peaks on day 2 compared to day 1 indicate stronger preference of V1 neurons for one segment of the corridor on day 2.

**c.** Cumulative distribution of the spatial modulation ratio computed on day 1(*grey*) and day 2 (*black*; median ± m.a.d: day 1: 0.71 ± 0.28, day 2: 0.56 ± 0.27. The two distributions are significantly different (Wilcoxon rank sum test; p = 0.005).

## 3.2 Discussion

Modulation of V1 activity in the context of running behaviours has been extensively investigated, both in the absence (Saleem et al., 2013) and presence of visual stimuli (Niell and Stryker, 2010), or even during running through visually-rich virtual reality environments (Saleem et al., 2013). Running through virtual reality environments promotes engagement with the sensory world giving rise to a more naturalistic visual experience (Poort et al., 2015). Yet, how other covariates of this naturalistic experience, beyond running, influence V1 activity, has not been systematically investigated.

Here we asked whether spatial signals, such as the position along the virtual reality corridor, influence V1 responses to the corridor's visual landmarks. Critically, because the corridor contained two visually-matching segments, we hypothesised that purely visual responses between the visually-matching segments should be similar. Instead, we found that many neurons preferred to fire more at a single virtual position, thereby giving rise to a response pattern at the population level that tiled up the whole virtual corridor. The tiling up of the virtual corridor also revealed that many cells responded more strongly in the second half of the virtual environment, suggesting that the observed modulation could not be attributed solely to adaptation influences. Nevertheless, other behavioural variables, such as running or reward, may have potentially influenced the observed modulation (Poort et al., 2015). To assess the influence of these behaviour-related variables, together with visual factors (pupil size, eye movements), we modelled these factors as predictors of the measured responses. We found that speed exerted the strongest effect. Yet, the observed modulation could not be fully explained by speed, neither by the joint contribution of all factors together. Finally, the observed modulation could not be explained by small differences between the two halves of the corridor, which mostly originated from the end grey wall. Taken together, these results point strongly towards the influence of visual responses in V1 by spatial signals, and more specifically by the position of the animal in the environment.

Pivotal studies on the changes in sensory responses induced by behavioural and task-relevant factors have highlighted the underlying role of task engagement or learning (Makino and Komiyama, 2015; Otazu, Tai, Yang, and Zador, 2009; Poort et al., 2015). For instance, the learning-induced improvement in stimulus selectivity is much more pronounced during task engagement than during passive presentation of the same stimuli under anaesthesia (Poort et al., 2015). Thus, it seems that task-engagement accentuates

non-sensory influences on sensory responses. However, we found that in the absence of task, when animals simply ran through the environment without reward, spatial modulation was still present, suggesting that even incidental learning of the spatial features of the environment, in the absence of a task, is sufficient to modulate V1 responses. To which extent learning of the spatial features shapes the spatial modulation of V1 responses remains an open question. Here I attempted to provide some insight into this question based on preliminary data obtained from learning experiments in two animals. These experiments showed that spatial modulation was already present on the first day of exposure to the virtual reality environment. However, the degree of spatial modulation was lower on the first day compared to the second day. Perhaps the apparent modulation observed on the first day can be explained by all other task-related and behavioural factors. This is a testable hypothesis I would like to assess using the model presented in Section **3.1.3**. In addition, it would be useful to assess within-session experience-dependent changes, by computing response profiles from the minimum number of trials possible. Clearly, further experiments and analysis are needed, before concluding on the effects of learning on the modulation of visual responses.

Another important question concerns how similar are the spatial signals observed in V1 to the spatial signals represented in areas traditionally thought to be involved in spatial information processing. In a recent study we showed that V1 and the hippocampus encode similar and consistent estimates of the animal's position and that these encoded positions reflect the animal's subjective estimate of position in the environment (Saleem*, Diamanti* et al., 2018). Yet, these results do not establish whether the similarity of signals in V1 and CA1 is due to feedforward signals from vision or to feedback signals from navigational systems. Given that V1 and CA1 are not connected monosynaptically, we speculate that multiple areas may play a role in conveying both feedforward-visual and feedback-spatial signals between the two regions. For instance, feedforward signals may be conveyed through higher visual areas, whereas feedback signals may be carried through cortical areas involved in processing of spatial information, such as the retrosplenial, prefrontal and parietal cortex (Wang, Gao, and Burkhalter, 2011). Among these areas the PPC is of particular interest, because it is thought to overlap, at least partly, with three higher visual areas, RL, A and AM (Wang, Gao, and Burkhalter, 2011). Therefore, it would be interesting to ask whether the degree of spatial modulation in these higher visual areas is the same as in V1. I will tackle this question in Chapter 5.

Our results demonstrate that spatial signals are present as early as in primary sensory cortex. To better understand how these signals influence sensory responses and inform behaviour under more naturalistic conditions, it would be useful to introduce 2-dimensional virtual environments (Aronov and Tank, 2014; Chen et al., 2018; Cushman et al., 2013), or ultimately record from freely-moving animals in real environments. In such experiments, though, precise control of the visual stimuli would have been challenging. In this study, instead, precise control of the visual stimulation was made possible thanks to head-fixation together with the introduction of a one-dimensional virtual reality environment. This experimental design has proved crucial to the detection of spatial signals in V1.

# Chapter 4

In the previous chapter I sought to characterise V1 responses in virtual reality and demonstrated for the first time, that V1 activity is influenced by yet another non-sensory factor: spatial position. These results add to the broad spectrum of behavioural and task-related factors known to influence V1 responses. Yet, little is known about the impact of non-sensory factors on responses in higher visual areas in the mouse, and even more so, about the functional role of higher visual areas during behaviour. In the next chapters (Chapters 3 to 5) I will address these issues using several experimental protocols of increased behavioural complexity.

In this Chapter I aim to establish a link with previous studies which probed the functional properties of higher visual areas with simple visual stimuli. I will thus present my own results on these functional properties obtained with the use of similar approaches (spatiotemporal frequency and orientation selectivity tuning) (**Section 4.1.1**). Transitioning towards behaviour in virtual reality, I will also compare measured activity in V1 and higher visual areas between viewing of simple visual stimuli versus active behaviour in virtual reality (**Section 4.1.2**).

Overall, this Chapter sets the stage for assessing the functional role of higher visual areas during behaviour, which will be discussed in the next chapters.

## 4.1 Results

### 4.1.1 Visual stimulation with drifting gratings

#### *4.1.1.1 Orientation and direction selectivity across areas*

To characterise the orientation and direction selectivity across visual areas, we presented drifting gratings at 12 different directions of motion (6 orientations, step 30 deg). The

spatial and temporal frequency of the stimuli was fixed at 2 Hz and 0.04 cpd respectively. Many cells exhibited a clear preference for specific orientations or directions across all areas (**Figure 4-1a**; example cells, *black* circles). This preference was confirmed by fitting the sum of two gaussians to the maximum activity during each stimulus presentation (**Figure 4-1a**, example cells, *grey* curve).

We first asked whether the degree of orientation selectivity was the same across areas. We obtained distributions of the orientation vector length for each cell and estimated the median across the population for each area (**Figure 4-1b**). The median vector length for V1 was higher than the rest of the areas, except from LM (vector length > 0.40 for V1 and LM as opposed to vector length ≤ 0.35 for AL, RL, A, AM and PM; K-S test, $p < 10^{-10}$. The stronger orientation selectivity for the V1 and LM populations was also reflected in the percentage of cells classified as orientation selective, based on a threshold of orientation vector length higher than 0.3 (in descending order: LM: 781/1149 = 68%; V1: 6966/10948 = 64%; RL: 2289/4090 = 56%; A: 766/1357 = 56%; PM: 757/1364 = 55%; AL: 599/1152 = 52%; AM: 1005/1927 = 52%). We also asked whether orientation selectivity in any of these areas was biased towards a specific orientation. We obtained distributions of preferred orientations for neurons with orientation vector length higher than 0.3, and we found a bias for cardinal versus oblique orientations in V1 (**Figure 4-1c** ;1-way ANOVA across 9 orientations; $p<10^{-10}$). A similar bias was also observed in AM (1-way ANOVA across 9 orientations; $p<10^{-5}$). No statistically significant bias was observed in the rest of the areas ($p > 0.01$).

We next focused on the direction selectivity of each area by addressing similar questions as for the orientation selectivity. To assess the degree of direction selectivity in each area we obtained distributions of the direction vector length and estimated the median across the population (**Figure 4-1d**). As previously, the median vector length for V1 and LM was higher than the rest of the areas (vector length = 0.25 for V1 and LM as opposed to vector length ≤ 0.23 for AL, RL, A, AM and PM; K-S test, $p << 0.001$). Also, the percentage of direction selective cells (direction vector length > 0.3) was higher for V1 and LM than the rest of the areas (in descending order: LM: 504/1149 = 44%; V1: 4615/10948 = 42%; A: 766/1357 = 36%; AM: 1005/1927 = 35%; PM: 757/1364 = 34%; AL: 391/1152 = 34%; RL: 2289/4090 = 31%). Is direction selectivity biased towards a specific direction in any of the areas? To address this question, we obtained distributions of preferred direction for neurons with direction vector length higher than 0.3 (**Figure 4-1e**). We observed a bias for vertical drifting gratings moving away from the observer in area A, but this bias was not

statistically significant. Similarly, we found no significant bias towards a certain direction in the rest of the areas.

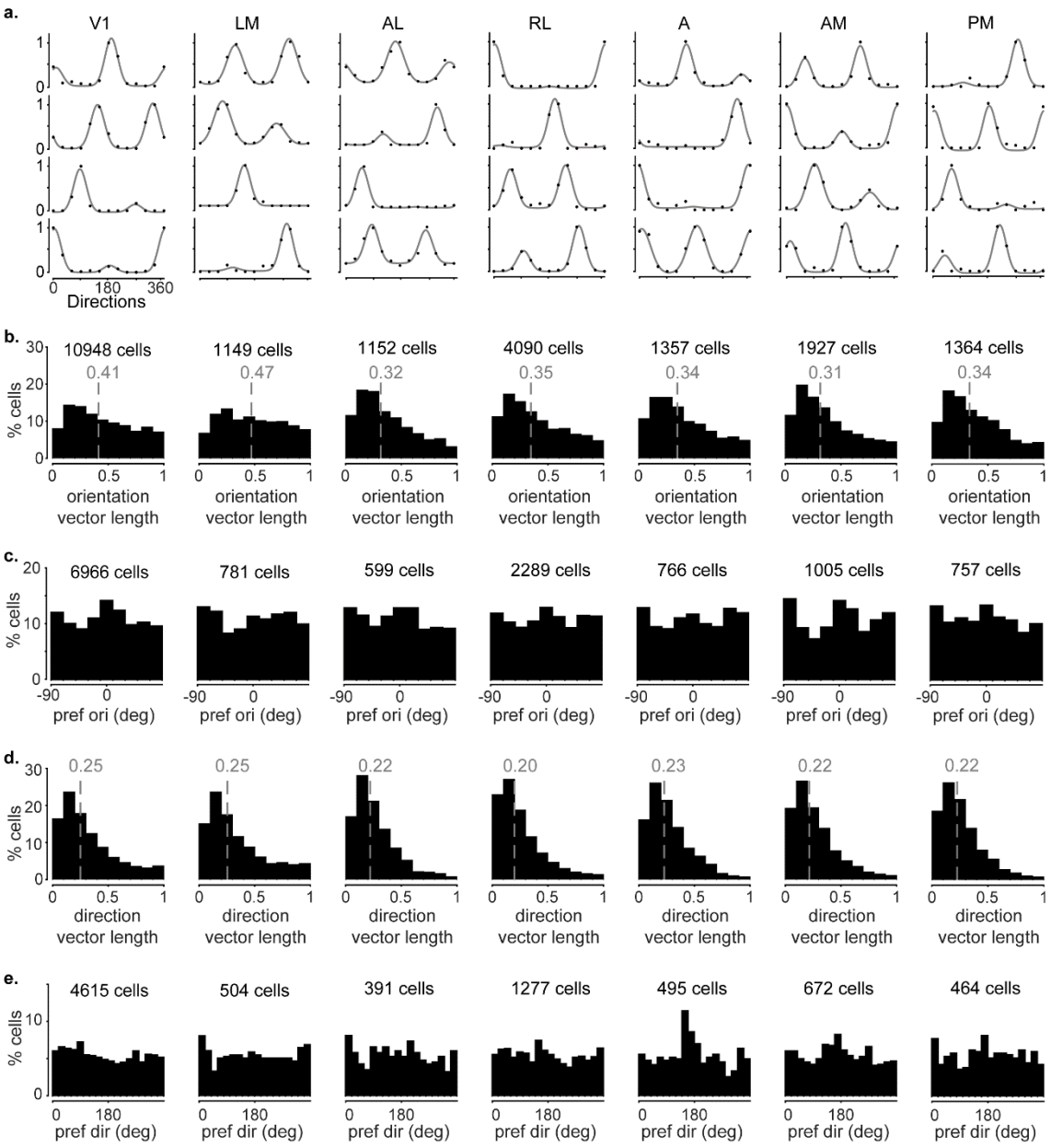


**Figure 4-1: Orientation and direction selectivity across visual areas.**
**a.** Example direction tuning curves across areas (*grey* trace) obtained by fitting the sum of two Gaussians; *black* data points: average response to a given orientation
**b.** Distributions of single-cell orientation vector lengths across areas (*grey dotted* line: median across the population).
**c**. Distributions of preferred orientations across areas for neurons with orientation vector length higher than 0.3.
**d.** Distributions of single-cell direction vector lengths across areas (*grey dotted* line: median across the population) **e.** distributions of preferred directions across areas for neurons with orientation vector length higher than 0.3.

Overall, these results replicate previous findings (Andermann et al., 2011; Marshel, Garrett, Nauhaus, and Callaway, 2011; Roth, Helmchen, and Kampa, 2012) demonstrating that direction and orientation tuning is a feature present in all areas. Notably, we found a bias in V1 for cardinal orientations consistent with Roth et al. (2012). A similar, significant bias was also found in AM, but not in the rest of higher visual areas. Our findings also demonstrated increased orientation and direction selectivity for V1 and LM, perhaps suggesting that these two areas are more similar, and therefore hierarchically closer to each other.

#### *4.1.1.2 Spatial and temporal frequency tuning*

To characterise the spatial and temporal frequency tuning preference across areas, we presented vertical drifting gratings at four spatial and five temporal frequencies. We found cells with robust responses to a specific range of spatial and/or temporal frequencies across all areas, albeit very few of those in area A (**Figure 4-2a**, **Figure 4-3a**). To further assess the spatial and temporal frequency tuning of individual cells, we obtained maps of average stimulus-evoked response at different combinations of spatial and temporal frequencies (Figure 4-2**b**). We then extracted the slice in these maps that corresponded to the maximum response in either the spatial or the temporal frequency and sorted them by preferred temporal or spatial frequency value (**Figure 4-3a, c**). Sorting by preferred spatial or temporal frequency value revealed the frequencies preferred by the majority of cells. To test for within-area biases towards low or high frequencies we also estimated the mean response for each frequency value (**Figure 4-3b, c**)

In the spatial frequency domain (**Figure 4-3a, b**), V1 and LM exhibited a bias towards the lowest spatial frequency (V1: 923/2956, 31% of cells; LM: 99/336, 29%; mean response at lowest frequency significantly different from the mean at all other frequencies, 1-way ANOVA: $p \leq 0.001$). This bias was even more pronounced in areas AL and RL (AL: 75/174, 43%; RL: 325/882, 37%; mean response at lowest frequency significantly different from the mean at all other frequencies, $p << 0.001$). Areas A and AM contained more cells tuned for the lowest frequency (AM: 107/238, 45%; A: 38/95, 40%), but the mean response at lowest frequency was different only from the highest frequency ($p<10^{-10}$). Finally, the majority of cells in PM exhibited an equal preference for both the lowest and the highest spatial frequency, with 43/142 cells preferring the lowest spatial frequency (30%) and 45/142 cells (32%) preferring the highest spatial frequency (no statistically significant difference across frequencies).

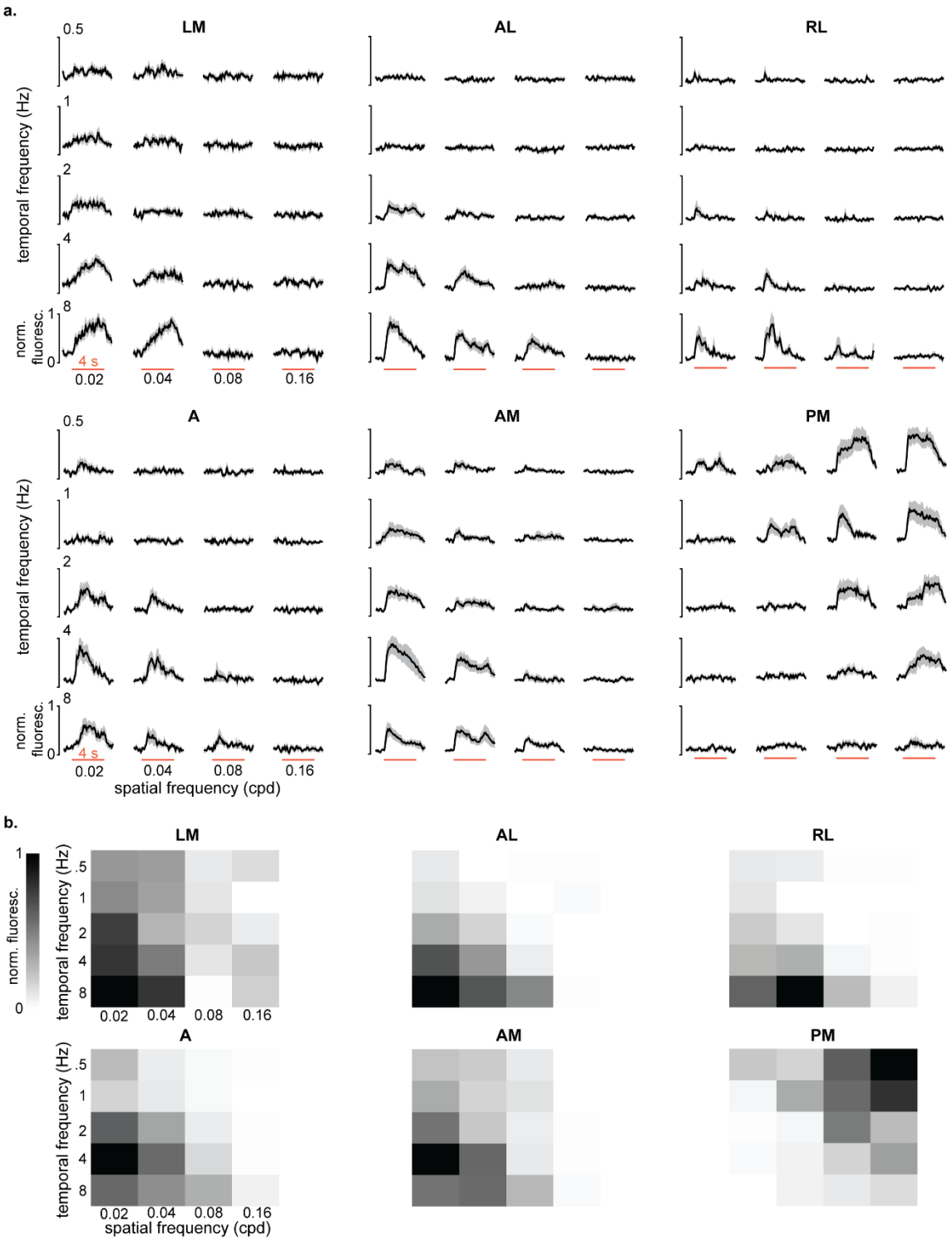


**Figure 4-2: Spatiotemporal frequency tuning across areas**
**a.** Example average responses to drifting gratings of various spatial and temporal frequency across areas. Shaded regions are ±s.e.m.
**b.** Maps across areas of average stimulus-evoked response in the spatial-by-temporal frequency plane.

In the temporal frequency domain (**Figure 4-3c, d**), we found a bias in V1 for the lowest temporal frequency (782/2956, 26% of cells; mean response at lowest frequency

significantly different from the mean at two highest frequencies, $p<10^{-10}$). Instead, most higher visual areas except PM showed a bias towards the highest temporal frequency (in descending order: AL: 61/174, 35%; A: 33/95, 35%; LM: 106/336, 31%; RL: 277/882, 31%; AM: 74/238; 31%; mean response at highest frequency significantly different from the mean at lowest frequency, $p<10^{-10}$). Conversely, PM exhibited a significant bias towards low temporal frequencies (45/142 cells, 32%; mean response at lower frequency significantly different from the mean at the highest frequency, $p<10^{-10}$).

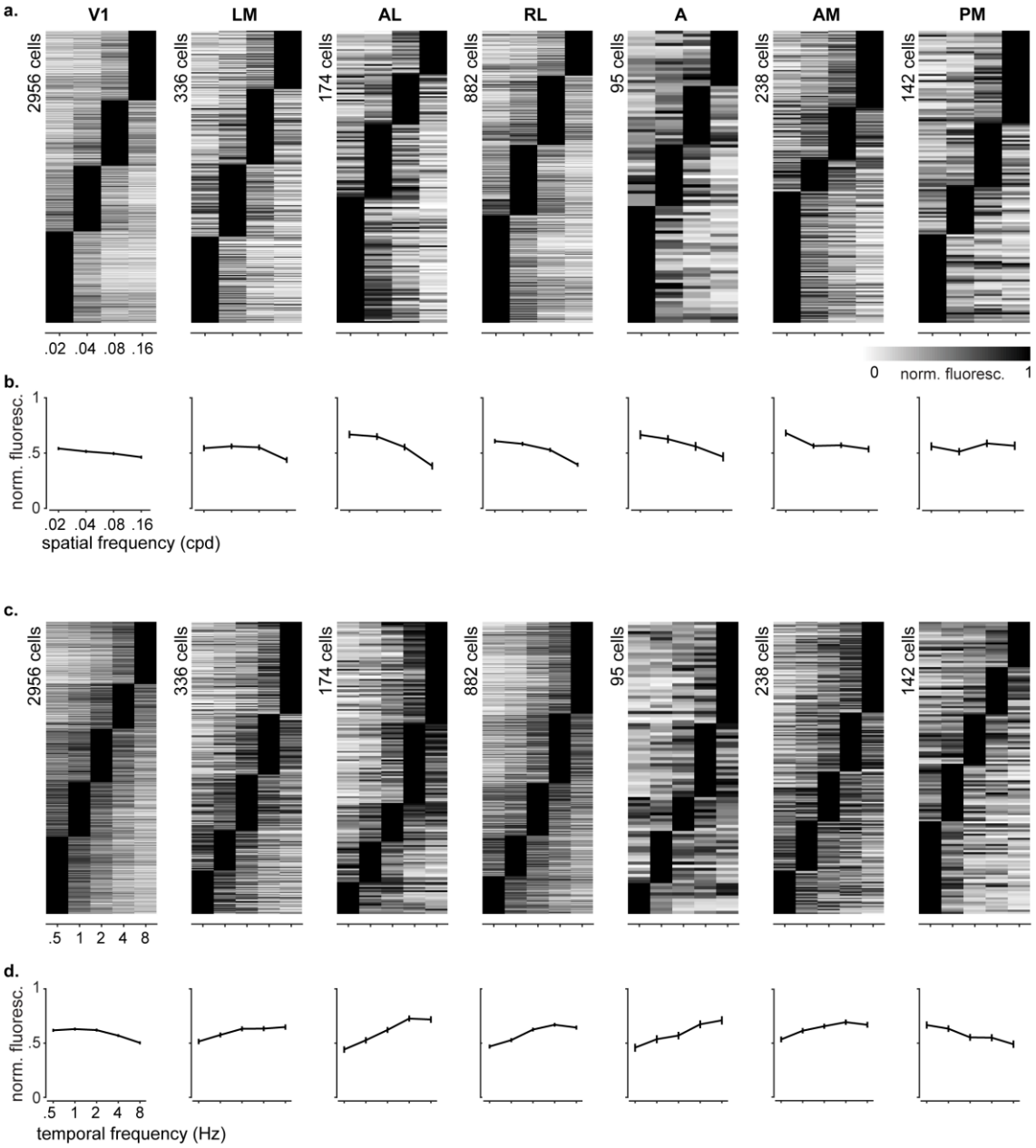


**Figure 4-3 Neurons in visual areas prefer different spatial and temporal frequencies**
**a.** Average responses to different spatial frequencies at the preferred temporal frequency, sorted based on the preferred spatial frequency. (*legend continues on next page*)

**b.** Average response to different spatial frequencies obtained from the population in **a.** Error bars are ±s.e.m.
**c.** Average responses to different temporal frequencies at the preferred spatial frequency, sorted based on the preferred temporal frequency.
**d.** Average response to different temporal frequencies obtained from the population in **c.** Error bars are ±s.e.m.

Up to now we considered spatial or temporal frequency tuning separately. We next sought to determine the number of neurons preferring a specific pair of spatial and temporal frequency, by estimating population density maps in the spatial and temporal frequency domain (**Figure 4-4a**). Neurons in V1 had preferences that spanned the whole spectrum of spatiotemporal frequencies homogeneously. Instead, in the rest of the areas we found biases for a narrow range a spatiotemporal frequency pairs. For instance, most cells in LM preferred high temporal and spatial frequencies. Instead, most neurons in AL, A and AM preferred high temporal but very low spatial frequencies. Cells in RL were tuned towards intermediate to high spatial and temporal frequency values. Finally, most cells in PM preferred the lowest temporal frequency and either very low or very high spatial frequency.

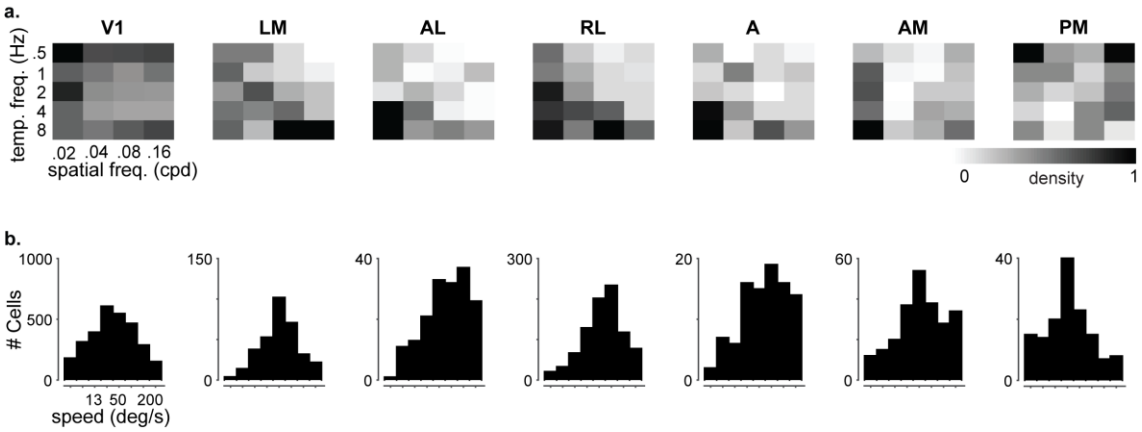


**Figure 4-4 Neurons in visual areas prefer different combinations of spatial and temporal frequencies**
**a.** Population density maps in the spatial and temporal frequency domain across areas
**b.** Ratio of preferred temporal frequency over preferred spatial frequency ('peak speed'). V1, LM and PM prefer lower 'peak speeds' compared to AL, RL, A and AM.

To further assess the preferred spatiotemporal preferences of each cell, we estimated the 'peak speed', i.e. the ratio of preferred temporal frequency over preferred spatial frequency **(Figure 4-4b**). The 'peak speed' tuning across the V1 population was broader compared to the higher visual areas. Instead, LM and PM demonstrated sharp tuning towards medium or lower 'peak speed' ranges respectively. Finally, for areas AL, RL, A and AM, the 'peak

speed' distribution was skewed towards higher speed values. Among all areas, V1 and PM had the lowest median values for 'peak speed' (25 deg/sec), then LM, RL and AM followed with 50 deg/sec, whereas AL and A had the highest median values at 100 deg/sec (most areas except PM had higher median peak speeds than V1; K-S test, $p << 0.0001$).

Overall these results confirm previous findings and suggest that areas AL, RL, A and AM are functionally distinct from LM and PM.

### 4.1.2 Distinct subpopulations respond reliably to drifting gratings or in virtual reality

To obtain a first estimate of the similarity of responses to drifting gratings and during active behaviour (closed loop in virtual reality), we used skewness (the third moment, s) as a measure of whether a cell is active or not; cells with skewness, $s>1$ were considered active (**Figure 4-5a**). Although many cells responded in both conditions, there were also distinct subpopulations that responded to one condition only, with more cells responding to drifting gratings alone than in virtual reality alone. This was probably the case because many cells were selective for spatial and temporal frequencies that were not present in virtual reality. (**Figure 4-5e**; ($s_{gratings} \geq 1$ & $s_{behaviour} < 1$) versus ($s_{gratings} < 1$ & $s_{behaviour} \geq 1$): V1: 26.6% vs 9.1% , LM: 37.1% vs 14.5%, AL: 33.8% vs 9.7%, RL: 30.9% vs 6.2%, A: 34.3% vs 12.3%, AM: 22.8% vs 11.4%, PM: 17.3% vs 14.8%). The existence of distinct subpopulations was also reflected in the relatively low correlations between the two conditions. (Pearson correlation coefficient, r: V1: 0.30, LM: 0.10, AL: 0.25, RL: 0.32, A: 0.32, AM: 0.22, PM: 0.19, $p<10^{-10}$ in all areas).

Because these estimates were derived from all cells recorded, perhaps cells active during presentation of drifting gratings had stimulus preferences different from the landmarks presented in virtual reality. Thus, given that one of the prominent landmarks in virtual reality was a vertical grating, we next investigated the skewness relationship between the two conditions for cells tuned to this specific type of stimulus. Again, we found that cells preferring the vertical grating did not necessarily respond in virtual reality (**Figure 4-5e**, data in *red*, Pearson correlation coefficient, r: V1: 0.35, $p<10^{-4}$, LM: 0.26, $p = 0.07$, AL: 0.58, $p = 0.0001$, RL: 0.36, $p<10^{-4}$, A: 0.34, $p=0.007$, AM: 0.12, $p=0.28$, PM: 0.11, $p=0.44$). In fact, we found that most cells preferring the vertical grating were active during presentation of drifting gratings only and not during active behaviour (% difference

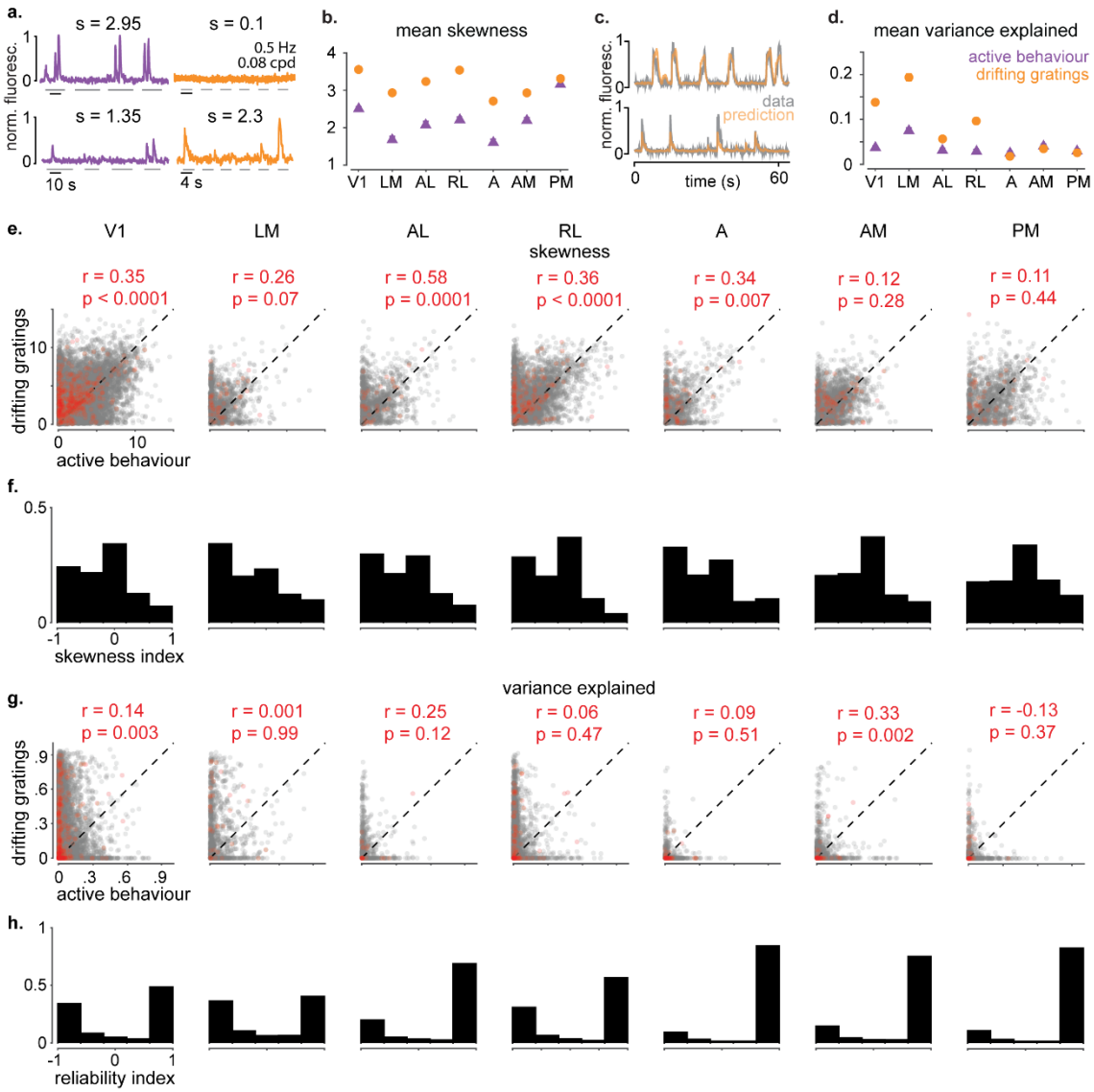


**Figure 4-5: Distinct subpopulations respond reliably to drifting gratings or in virtual reality**
**a.** Normalised fluorescence traces from two example cells during active behaviour (*magenta*; *left*) and during presentation of a vertical drifting grating at 0.5 Hz and 0.08 cpd (*orange; right*). *Grey* bars indicate trial duration. s is skewness.
**b.** Population average of skewness for each visual area during active behaviour (*magenta*) or during presentation of drifting gratings (*orange*). Error bars are ± s.e.m.
**c.** Normalised fluorescence traces (*grey*) and predicted responses (*orange*) for two example cells during presentation of drifting gratings at different orientations.
**d.** Population average of variance explained for each visual area during active behaviour (*magenta*) or during presentation of drifting gratings at different orientations (*orange*). Error bars are ± s.e.m.
**e.** Skewness of single-cell activity across areas during active behaviour (closed-loop) against skewness of single-cell activity during presentation of drifting gratings, for the whole population (*grey*) and for cells preferring the vertical orientation *(red*). r and p is the Pearson correlation coefficient and p-value respectively for cells preferring the vertical orientation.
**f.** Skewness index, s, across areas. s close to 1 means high skewness in virtual reality, whereas s close to -1 means high skewness during presentation of drifting gratings.
**g.** Same as in **a.** for variance explained. The L-shaped distribution in all areas indicates that most cells respond reliably during active behaviour alone or during presentation of drifting gratings alone
**h**. Same as in **f.** for variance explained ('reliability index'). The reliability index gradually becomes more biased away from V1 and LM, towards PM.

between cells active during presentation of drifting gratings only versus cells active during behaviour only: V1: 25%, LM: 26%, AL: 29% RL: 28% A: 26% AM: 12% PM: 15%).

The higher responsiveness of cells during presentation of drifting gratings than during active behaviour was also reflected in the mean skewness across the population (**Figure 4-5b**). Typically, the mean skewness in virtual reality was lower than the mean skewness during stimulation with drifting gratings, with the only exception of PM ((mean ± s.e.m.)$_{behaviour}$ vs (mean ± s.e.m.)$_{gratings}$ , V1: 2.53 ± 0.02 vs 3.56 ± 0.03, LM: 1.71 ± 0.05 vs 2.93 ± 0.06, AL: 2.10 ± 0.06 vs 3.23 ± 0.07, RL: 2.24 ± 0.04 vs 3.54 ± 0.04, A: 1.64 ± 0.05 vs 2.71 ± 0.06, AM 2.23 ± 0.04 vs 2.93 ± 0.05, PM: 3.19 ± 0.07 vs 3.31 ± 0.07). To further classify cells as responsive to drifting gratings and/or in virtual reality, we estimated a skewness index (($s_{behaviour} - s_{gratings}$ )/( $s_{behavour} + s_{gratings}$)). Cells with zero skewness index responded in both conditions, cells with index of -1 responded to drifting gratings alone, whereas cells with index +1 responded in virtual reality alone (**Figure 4-5f**). The mean index across the population was less than zero in all areas, consistent with the observation that cells are more responsive to drifting gratings (mean index ± s.e.m.: V1: -0.20 ± 0.01, LM: -0.13 ± 0.13, AL: -0.27 ± 0.02, RL: -0.26 ± 0.01, A: -0.23 ± 0.01, AM: -0.14 ± 0.02, PM: -0.02 ± 0.01).

We next assessed the reliability of responses using a measure of prediction quality (variance explained, Q): in virtual reality, we estimated the quality of predicting each neuron's activity as a function of position in the virtual corridor ($Q_{behaviour}$). For the stimulation with drifting gratings, we focused on responses to different stimulus orientations, and estimated the quality of predicting each neuron's activity as a function of time ($Q_{Ori}$; **Figure 4-5c**). We then used the relationship between the two measures to ask how different the populations of neurons were responding in virtual reality or to drifting gratings of different orientations.

We found two fairly distinct populations: typically cells with reliable responses in virtual reality were not responding reliably during the orientation selectivity protocol and vice versa (**Figure 4-5g**). Notably, even cells that were tuned to vertical gratings, a stimulus that is also present in virtual reality, exhibited reliable responses to vertical gratings alone for most areas, except A and AM (**Figure 4-5g,** data in *red*; (% difference between reliable responses during presentation of drifting gratings only versus reliable responses during behaviour only: V1: 33, LM: 37, AL: 10, RL: 28, A: -8, AM: -1, PM: 0.14). In addition, the reliability of responses between the two conditions was not correlated (Pearson correlation

coefficient, r; p: V1: 0.14, 0.003; LM: 0.001, 0.99; AL: 0.25, 0.12; RL: 0.06, 0.47; A: 0.09, 0.51; AM: 0.33, 0.002; PM: -0.13, 0.37).

To obtain an additional estimate of how the reliability of responses between the two conditions varied within and between areas, in each area we estimated the mean prediction quality across the population (**Figure 4-5d**). We found that the mean prediction quality for the orientation selectivity protocol was highest in V1 and LM compared to the rest of the areas; within V1 or LM, mean prediction quality for orientation selectivity was also much higher than the same measure in virtual reality (V1: $Q_{ORI} = 0.138 \pm 0.002$ vs $Q_{behaviour} = 0.040 \pm 0.001$; LM: $Q_{ORI} = 0.193 \pm 0.007$ vs $Q_{behaviour} = 0.077 \pm 0.003$). Besides V1 and LM, AL and RL exhibited lower prediction quality for orientation selectivity, but still higher than the prediction quality in virtual reality (AL: $Q_{ORI} = 0.056 \pm 0.004$ vs $Q_{behaviour} = 0.031 \pm 0.001$; RL: $Q_{behaviour} = 0.096 \pm 0.003$ vs $Q_{behaviour} = 0.031 \pm 0.001$). Finally, in areas A, AM and PM the difference in prediction quality between the two conditions was reversed, with responses in virtual reality being on average slightly more reliable than for orientation selectivity (A: $Q_{ORI} = 0.017 \pm 0.002$ vs $Q_{behaviour} = 0.027 \pm 0.001$, AM: $Q_{ORI} = 0.034 \pm 0.002$ vs $Q_{behaviour} = 0.043 \pm 0.001$, PM: $Q_{ORI} = 0.025 \pm 0.002$ vs $Q_{behaviour} = 0.032 \pm 0.002$).

To further classify single-cell responses as more reliable in virtual reality or during the orientation selectivity protocol, we estimated a 'reliability index' $((Q_{behaviour} - Q_{ORI})/(Q_{behaviour} + Q_{ORI}))$. Distributions of the reliability index revealed a trend across areas that was not evident previously (**Figure 4-5h**): V1 and LM had very similar distributions, with almost equal numbers of cells having reliable responses during the orientation selectivity protocol (reliability index = -1) or during behaviour (reliability index = +1). Beyond V1 and LM, we found a gradual bias towards reliable responses in virtual reality. This bias was strongest in A. Consistently, the mean 'reliability index' increased from V1 and LM towards A (mean 'reliability' index ± s.e.m. in ascending order: LM: 0.06 ± 0.02, V1: 0.15 ± 0.01, RL: 0.27 ± 0.01, AL: 0.50 ± 0.02, AM: 0.61 ± 0.02, PM: 0.71± 0.02, A: 0.75 ± 0.02). Notably, unlike previous findings obtained using skewness, we found few cells responding reliably in both conditions (index = 0; V1: 5%, LM: 6%, AL: 3%, RL: 4%, A: 2%, AM: 3%, PM: 2%). Moreover, although based on skewness, cells were more active during presentation of drifting gratings, based on prediction quality their responses were more reliable in virtual reality in most areas, except V1 and LM.

Overall, these findings suggest that reliable responses during active behaviour are distinct from reliable responses during presentation of drifting gratings. This dichotomy

cannot be attributed solely to the fact that cells are active only in one condition but not the other. In addition, although cells in all areas are typically more active during presentation of drifting gratings than during behaviour, they may still respond more reliably during active behaviour. This was particularly evident in areas A, AM and PM, suggesting that these areas may be more sensitive to task engagement.

## 4.2 Discussion

Over the past decade an increasing body of evidence has revealed that the mouse visual cortex comprises of numerous retinotopically organised higher visual areas. At least 6 of these areas (LM, AL, RL, A, AM and PM) can be distinguished from one another based on classic measures of stimulus tuning, such as spatial and temporal frequency preference.

In this Chapter, I used well-established stimulation protocols to probe the orientation selectivity and spatiotemporal tuning of higher visual areas (**Section 4.1.1**). By replicating previous findings on the preferred spatiotemporal tuning of higher visual areas, I confirmed I could use two-photon imaging to target distinct higher visual areas. I also asked how reliable were the responses across conditions of different behavioural complexity: passive viewing of drifting gratings versus active behaviour (closed loop) (**Section 4.1.2**). I found that most responses were reliable only in one of the two conditions. Interestingly, this analysis revealed a progressive reliability bias towards active behaviour when moving away from V1 and LM towards areas A and AM.

I found that V1 and LM had the highest degree of direction and orientation selectivity than any other higher visual area. In addition, similar to previous reports, I observed an orientation bias in V1 towards cardinal orientations (Roth, Helmchen, and Kampa, 2012). In the spatiotemporal frequency domain, I found that V1 was not tuned to a specific spatiotemporal frequency range. In contrast, AL was one of the most sharply tuned areas, preferring high temporal and low spatial frequencies, as previously reported by Andermann et al.(2011). Similar to AL was the spatiotemporal preference of A, consistent with the findings by Murakami et al. (2017). Nevertheless, there were also some discrepancies between our data and previously published studies. Specifically, V1's bias in the nasal-temporal direction was not as clear as previously reported (Hillier et al., 2017; Marques, Nguyen, Fioreze, and Petreanu, 2018). In addition, I did not observe a clear bias for PM towards high spatial frequencies (Andermann et al., 2011). Despite these discrepancies, our findings are consistent to a large extent with previous reports.

### 4.2.1 Consistencies and discrepancies in orientation and direction selectivity

To probe orientation and direction selectivity in V1 and higher visual areas we used drifting gratings at a temporal frequency of 2 Hz and a spatial frequency of 0.04 cpd. These fixed temporal and spatial frequency values have been one of the limitations in our study, because this particular set may not correspond to the range of frequencies preferred by a certain area. Indeed, orientation selectivity is tightly linked to the preferred spatial frequency and direction selectivity to the preferred temporal frequency, therefore it is important that these dependencies are taken into account (Glickfeld and Olsen, 2017). In fact, various studies have emphasised that neglect of these dependencies is the main reason of disagreement between their findings and others (Andermann et al., 2011; Marshel, Garrett, Nauhaus, and Callaway, 2011). Likewise, our result on higher degree of orientation and direction selectivity in V1 and LM disagrees with previous studies: Marshel et al. (2011) reported better direction selectivity in areas AL, RL, AM; Roth et al. (2012) on the other hand reported similar orientation and direction preference in V1 and PM; finally, Andermann et al. (2011) saw no difference in orientation selectivity between V1, AL and PM, but he did see higher direction selectivity in V1 compared to AL and PM; in fact, direction selectivity in AL was minimal. Therefore, our results in direction selectivity are consistent with the study by Andermann et al. (2011), which was also done in awake mice. Instead, the studies by Marshell et al. (2011) and Roth et al. (2012) were performed in anaesthetised mice, which may be another reason for the observed discrepancies.

We also looked at biases towards a specific orientation or direction of motion. Once more, our finding that V1 prefers cardinal directions is consistent with some of the previous findings (Roth, Helmchen, and Kampa, 2012), but it is not in agreement with Andermann et al. (2011) who reports no explicit bias in V1 or PM. In the rest of the areas, we found an overall tendency towards preference for cardinal directions. This tendency was statistically significant for AM. However, in contrast to Roth et al. (2012), preference for cardinal directions was not statistically significant for PM. Conversely, our result of lack of bias for a specific direction in all areas is in disagreement with Hillier et al. (Hillier et al., 2017), who reports preference of V1 for nasal-temporal direction of motion, and with Andermann et al. (2011) who found a bias for upward-downward direction of motion in AL, although in the later study the sample size was relatively small.

### 4.2.2 Consistencies and discrepancies in spatiotemporal selectivity

Consistent with previous anatomical and functional studies, our results suggest that higher visual areas have distinct spatiotemporal tuning properties. Specifically, we find that areas AL, RL and A are tuned to low spatial and high temporal frequencies. Thus, these areas may form a functionally distinct cluster involved in motion processing, as previously suggested by Wang et al. (Wang, Sporns, and Burkhalter, 2012) and followed up by the functional study of Murakami et al. (Murakami, Matsui, and Ohki, 2017). On the other hand, PM is the only higher visual area tuned to low temporal frequencies (Marshel, Garrett, Nauhaus, and Callaway, 2011). In addition, we found that PM prefers both high and low spatial frequencies. This preference of PM is in agreement with Roth et al. (2012), but not with Andermann et al. (2011) and Marshell et al. (2011), who found only high spatial frequency tuning. Finally, LM is tuned to intermediate spatial frequencies but high temporal frequencies (Marshel, Garrett, Nauhaus, and Callaway, 2011). Therefore, higher visual areas can be distinguished from V1 in that they exhibit spatiotemporal selectivity. Indeed, apart from a modest bias towards low temporal frequency (see also Marshel et al., 2011), V1 ‘s preferences spanned the whole spectrum of spatiotemporal frequencies.

### 4.2.3 Passive viewing of drifting gratings versus active behaviour

In this subsection I demonstrated that the subpopulations with reliable responses to drifting gratings of different orientations or during active behaviour were distinct and that this dichotomy could not be attributed to the fact that cells became silent when switching experimental protocol. Notably, very few cells responded reliably to both conditions even in V1.

The main finding here was the clustering of higher visual areas into distinct groups depending on their relative preference to respond reliably to drifting gratings or during active behaviour (**Figure 4-5h**). Specifically, V1 and LM appeared to cluster together, with almost equal amounts of reliable responses in both conditions. Beyond V1 and LM, reliable responses gradually decreased in AL and RL and eventually reached a minimum in area A. Correspondingly, responses during active behaviour became more reliable in areas A, AM and PM, indicating that these areas formed a separate group that responded predominantly during active behaviour. Thus, as we move away from V1 we see two gradual changes: a decrease in the reliability of responses to drifting gratings; an increase in reliability during active behaviour.

It has been reported that areas A, AM and PM are biased towards the periphery (Zhuang et al., 2017), where the optic flow is higher in virtual reality. Therefore, can the increase in reliability during active behaviour for A, AM and PM be due to the fact that these areas experience higher optic flow? Indeed, I cannot exclude the possibility that retinotopic preference may influence the reliability of responses during active behaviour. Nevertheless, the retinotopic bias alone is not sufficient to explain the current result; AL for instance, a lateral visual area that is biased towards the central visual field, contains more reliable responses during active behaviour than its neighbouring area, LM. In addition, the degree of reliability during active behaviour in AL, an area biased towards the central visual field, is equal to the degree of reliability in AM.

Can the decrease in reliability during presentation of drifting gratings be explained by differences in spatiotemporal tuning? I believe this is unlikely for the following reasons. First, to probe orientation selectivity I used a stimulus with intermediate spatial and temporal frequency values (0.04 cpd and 2Hz). These values were outside the range of preferred spatial and temporal frequencies in all areas. Second, areas with different spatiotemporal frequency tuning, such as A and PM exhibited similar degree of reliable responses to drifting gratings. Conversely, areas with similar spatiotemporal frequency preference, such as AL and A exhibited different degree of reliable responses, with A containing the least reliable responses to drifting gratings.

Overall, these results indicate the existence of three groups: 1. V1 and LM contain almost equal subpopulations with reliable responses to the drifting gratings or during active behaviour; 2. AL and RL contain more cells with reliable responses during active behaviour than during presentation of drifting gratings; 3. A, AM, and PM exhibit strong bias towards reliable responses during active behaviour only.

Having characterised responses during active behaviour, in terms of skewness and prediction quality, in the next chapter I focus on the responses profiles across higher visual areas obtained during active behaviour.

# Chapter 5

In Chapter 3 I sought to characterise V1 responses in virtual reality and demonstrated for the first time, that V1 activity is influenced by yet another non-sensory factor: spatial position. These results add to the broad spectrum of behavioural and task-related factors known to influence V1 responses. Yet, little is known about the impact of non-sensory factors on responses in higher visual areas in the mouse, and even more so, about the functional role of higher visual areas during behaviour.

To this end, in this chapter, I set off to characterise response profiles during active behaviour (closed-loop) as a function of position in the virtual corridor across higher visual areas. Specifically, I will address the following questions: is spatial modulation of visual responses also present in higher visual areas during active behaviour? If responses in higher visual areas are indeed influenced by spatial position, is this modulation stronger or weaker compared to V1?

In the next chapter, I will then focus on responses during play-back of virtual reality sessions (open-loop) and ask whether responses during open-loop are different from responses in closed-loop.

By tackling these questions in the current and next chapter, I aim to contribute new piece of evidence on the functional role of higher visual areas in the mouse, and ultimately provide new insight into the mechanisms underlying higher-level visual processing during behaviour.

## 5.1 Results

### 5.1.1 Spatial modulation present across higher visual areas in closed-loop

Before estimating response profiles, we deconvolved the measured fluorescence signal to obtain firing rates for each neuron (**2.7.22.8.1** in Methods). For each area, we then obtained response profiles as a function of position, as previously described (Section **3.1.1**). To exclude very noisy response profiles, we selected cells whose response profiles were predicted with at least 5% prediction quality (LM: 638 out of 1,503 cells (42%); AL: 441 out of 1,741 (25%); RL: 893 out of 5,192 (17%); A: 998 out of 4,126 cells (24%); AM: 1,007 out of 3,278 cells (31%) ; PM: 508 out of 2,512 cells (20%)). We then divided between odd and even trials and sorted response profiles as previously (**Figure 5-1a, b**). In each area the response profiles gave rise to a response pattern across the population that tiled the whole virtual corridor. However, the robustness of the response pattern was not the same across areas. This became more evident when we estimated the mean response of neurons responding to the visually matching segments as a function of the distance from the peak response location (**Figure 5-1c**; LM: 385 cells AL: 227 cells RL: 500 cells A: 473 cells AM: 528 cells PM: 267 cells). In the ideal scenario where all response profiles were exactly the same between odd and even trials, the mean normalised response should have peaked at one. Instead, we found that the average peak height was less than 80%. Among all areas, peak height was lowest in A, suggesting that responses in this area exhibited the highest variability (0.57 ± 0.01). Next to follow were areas PM with average peak height of 0.67 ± 0.02 and AM with 0.70 ± 0.01. AL and RL had a slightly higher peak height of 0.72 (AL: 0.72 ± 0.02, RL: 0.72 ± 0.01). Finally, LM had the highest peak height (V1: 0.77 ± 0.01, LM: 0.80 ± 0.01).

To summarise the degree of spatial modulation in each area, we defined the 'spatial modulation ratio' as the ratio of each cell's responses at the two visually matching locations (±40 cm away from preferred position / preferred position). The median spatial modulation ratio was significantly lower than 1 in all areas (**Figure 5-1d**; median ± m.a.d.: LM: 0.43 ± 0.27, AL: 0.50 ± 0.32, RL: 0.35 ± 0.26, A: 0.45 ± 0.36, AM: 0.57 ± 0.32, PM: 0.52 ± 0.36; Wilcoxon signed rank test: $p < 10^{-13}$ in all areas). Notably, the fact that response profiles tiled the whole corridor indicates that there were many cells that preferentially fired maximally in the second visually-matching segment. Therefore, the bias towards lower values could not be attributed solely to adaptation effects. The percentage of

cells firing in the first half of the corridor ranged between 53% and 58% in most areas, except A where the percentage of cells with preferred position in the first half was lower (A: 48%, LM: 55%, AL: 58%, RL: 56%, AM: 56%, PM: 57%).

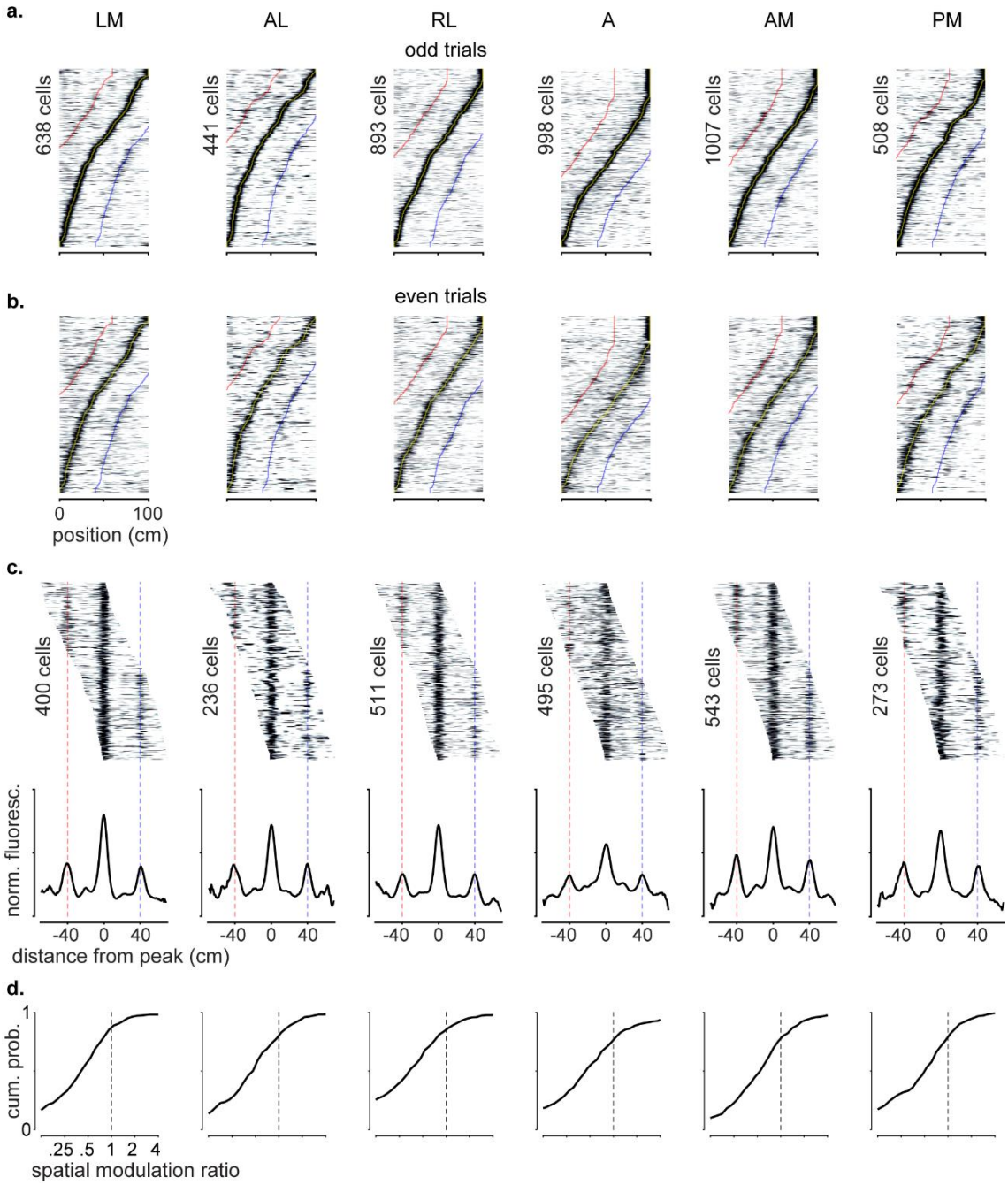


**Figure 5-1: Spatial modulation present across higher visual areas in closed-loop**
**a.** Normalised response in closed-loop as a function of distance in the virtual environment obtained from odd trials only, for neurons across higher visual areas with activity significantly modulated by position in the corridor. Neurons are ordered based on the position of their maximum response. Red and blue lines indicate position at +/- 40 cm from maximum response.
**b.** Same as **a**, for the half of the data (even trials) that were not used to order the responses. Sequence and scaling of the responses are the same as in **a.** *(legend continues on next page)*

**c**. Normalised response of neurons (with peak response between 15 and 85 cm along the corridor) across higher visual areas as a function of the distance from the peak response, obtained by aligning responses in **b.** (top), and population average across these responses (bottom). Red and blue lines same as in **a.** Lower values of the side peaks compared to central peak indicates strong preference of neurons in all areas for one segment of the corridor over the other visually-matching segment (40 cm away from peak response).
**d.** Cumulative distribution for the ratio of secondary response (40 cm away from peak response, red or blue line in a) divided by peak response across areas, derived from even trials.

### 5.1.2 Comparison between V1 and higher visual areas in closed-loop

Having established that responses in visual cortex are modulated by spatial position during active behaviour, I next sought to compare the degree of spatial modulation between higher visual areas and V1. Before tackling this question though, one must be cautious of the potential differences in the modulation ratio caused by the differences in visual scenes between the centre and the periphery. Specifically, the virtual reality environment was by design optimised to drive responses in the periphery, in terms of the chosen spatial frequency of the patterns. In addition, although the virtual reality scenes were very similar in the periphery, this was not the case in the centre particularly due to the grey wall at the end of the corridor. Given these differences between the central and peripheral part of the visual field it was also expected that the response profile patterns might be different depending on the receptive field positions of individual neurons.

To test for this possibility, we first focused on V1, because it was more straightforward to distinguish between V1 portions 'looking' at the centre (lateral V1) or the periphery (medial V1) based on the retinotopic map (**Figure 5-2**). The response profile patterns revealed that responses at the 'non-preferred' position+/- 40cm away from the main peak, were much more pronounced in medial V1 than lateral V1(**Figure 5-2a**). Correspondingly, we found marked differences in the height of the secondary peaks (**Figure 5-2b**) and in the modulation ratios (**Figure 5-2c**; median ± m.a.d.: lateral V1: 0.26 ± 0.21; medial V1: 0.53 ± 0.31). From these findings it became evident that an accurate comparison between areas was only possible if individual cells had similar receptive field positions.

To obtain portions of V1 and higher visual areas with similar receptive fields we implemented a seed-pixel correlation approach. We used the retinotopic map to choose one pixel ('the seed pixel') within the borders of each area. We then correlated the RGB values of this pixel with the RGB values of all other pixels on the retinotopic map. The

correlations of the RGB values gave rise to correlation maps that were clearly distinguished from one another based on preferred azimuth (correlation coefficient > 0.6). We used these maps together with the borders of individual areas to identify cells with similar receptive fields in V1 and each higher visual area. Due to this approach, the receptive fields of the chosen neurons were confined within a certain range of azimuth, similar to previous reports (Zhuang et al., 2017). This range varied form area to area: the chosen portions in LM and AL were biased towards the centre of the visual field, A, AM and PM were biased towards the periphery whereas RL lay between. The portions of V1 used to compare responses in each area also varied accordingly (**Figure 5-3a**).

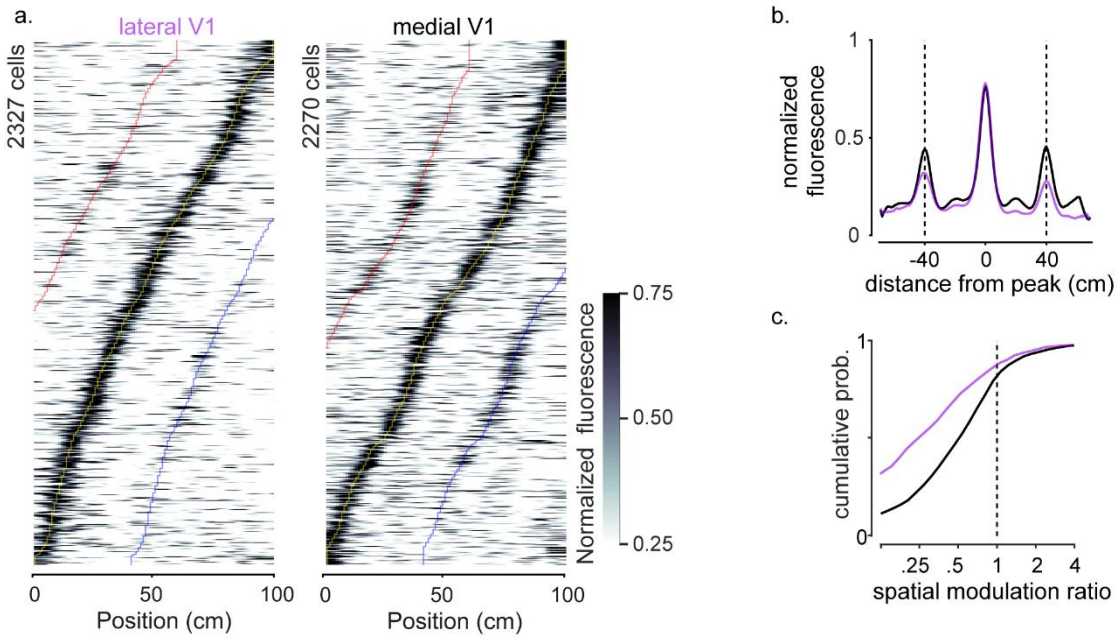


**Figure 5-2: Retinotopic differences in the degree of spatial modulation in V1**
**a.** Normalised response as a function of distance in the virtual environment obtained from even trials only, for neurons in lateral V1 (receptive field centres < 40 deg; *left*) or neurons in medial V1 (receptive field centres > 40 deg).
**b.** Population average across the normalised responses of neurons (with peak response between 15 and 85 cm along the corridor) as a function of the distance from the peak response; *magenta* trace: lateral V1; *black* trace: medial V1.
**c.** Cumulative distribution for the ratio of secondary response divided by peak response, derived from the same response profiles as in **c.;** *magenta* trace: lateral V1; *black* trace: medial V1.

We next asked whether the degree of spatial modulation varied between V1 and higher visual areas, by performing the same analysis as previously described. We arranged response profiles obtained from even trials according to the 'preferred position' of response profiles in odd trials (**Figure 5-3b**). We next obtained mean response profiles as a function of distance from the preferred position (**Figure 5-3c**). The mean response

profiles replicated our earlier finding that area A contained responses with the highest variability. We then divided the activity at ±40 cm away from the preferred position by the activity at the preferred position ('spatial modulation ratio'; **Figure 5-3d**). Consistent with our previous observation in V1, the median modulation ratio of V1 gradually increased away from the centre and towards the periphery (**c** *grey* curve; V1 median ± m.a.d. modulation ratio when comparing with: LM: 0.30 ± 0.28; AL: 0.31 ± 0.30; RL: 0.43 ± 0.29; A: 0.57 ± 0.26; AM: 0.56 ± 0.28; PM: 0.58 ± 0.27). The median modulation ratio in higher visual areas also increased in a similar manner (**Figure 5-3e** *black* curve; LM: 0.38 ± 0.28; AL: 0.35 ± 0.27; RL: 0.40 ± 0.32; A: 0.47 ± 0.37; AM: 0.57 ± 0.27; PM: 0.50 ± 0.33). Nevertheless, we reasoned that judging differences between areas only from the median was not the right measure. Specifically, our cross-validation approach (odd versus even trials) could give rise to ratios higher than one and therefore pull the median towards higher values. This was particularly evident in A, an area with high variability in the responses. We therefore sought to compare differences between distributions instead.

To assess any differences in the spatial modulation distributions between V1 and higher visual areas, we defined the 'spatial modulation index (SMI)' ((preferred position – non-preferred position)/(preferred position + non-preferred position)). Consistent with our earlier findings in V1, the degree of spatial modulation in V1 gradually decreased away from the centre and towards the periphery (*grey* trace; **Figure 5-3e**) Among higher visual areas, we found a pronounced bias for SMI = 1 in A, suggesting that this area contained an abundance of cells with single-peaked responses. When we compared the V1 distribution to the distributions of higher visual areas, we found that the SMI distribution in A and PM was statistically higher than portions of V1 'looking' at the same part of the visual field (1-sided two-sample K-S test, A: $p = 0.0005$; PM: $p = 0.003$). In the opposite direction, we also found that the SMI distribution in LM was statistically lower than V1 (1-sided two-sample K-S test, $p = 0.002$). It should be noted though that cells in these portions of LM and V1 have receptive fields towards the centre, where the virtual reality scenes comprise mostly of high spatial frequency stimuli and the end grey wall. We found no statistically significant differences in the rest of the areas (1-sided two-sample KS test of ratios: LM: AL: $p = 0.95$, RL: $p = 0.17$, AM: $p = 0.89$).

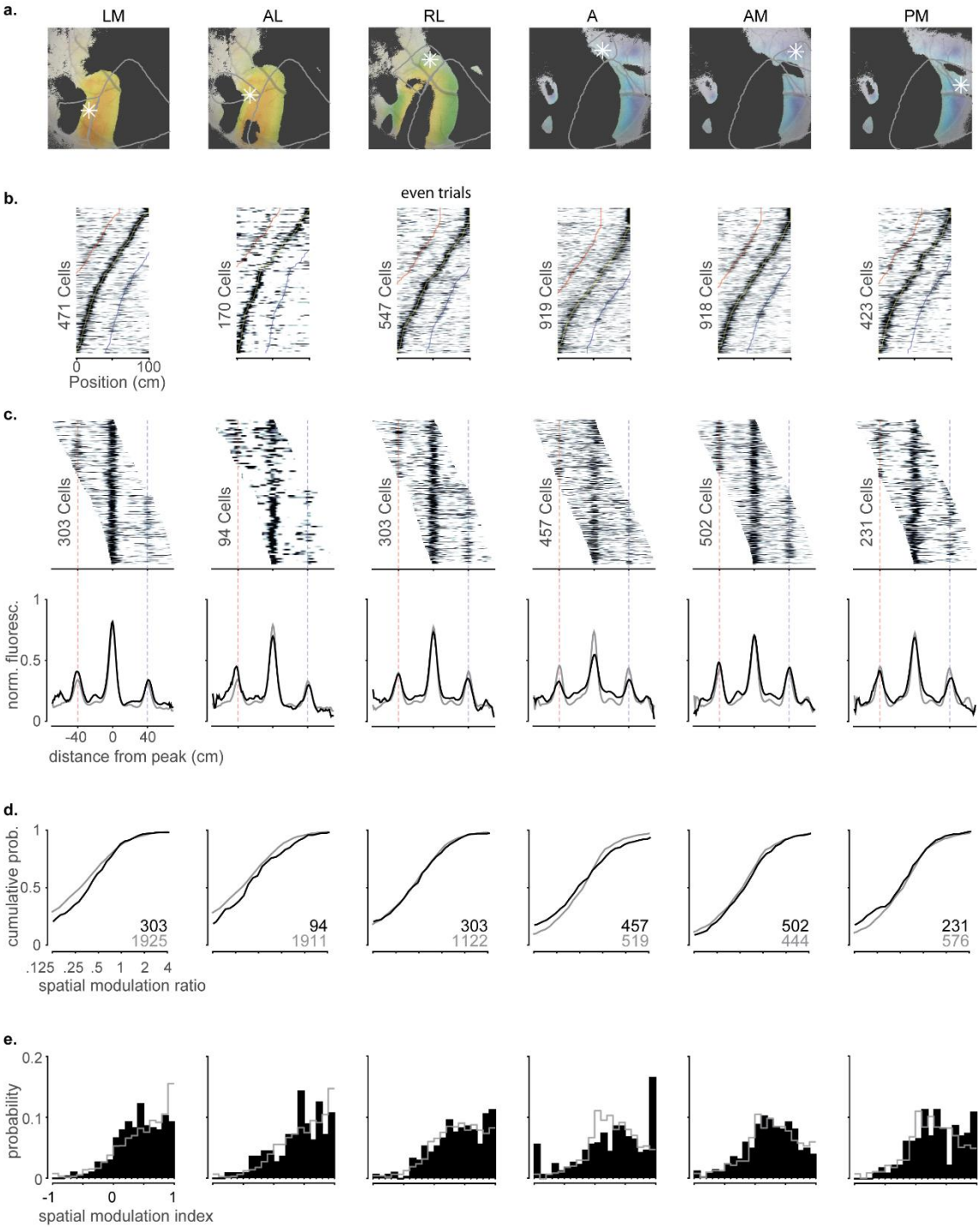


**Figure 5-3: Comparison between V1 and higher visual areas in closed-loop**
**a.** Correlation maps superimposed on an example retinotopic map. Correlation maps were obtained by correlating the azimuth (indicated by an RGB triplet) of a single pixel ('seed pixel'; *white* star) with the azimuth of all other pixels. Correlation maps for each area were obtained by placing the 'seed' pixel within its border (*grey* trace together with the border of V1). Correlation coefficients lower that 0.6 were set to zero (*black* regions).
**b.** Normalised response as a function of distance in the virtual environment obtained from even trials only, for neurons that lie within the 'transparent' regions in **a** and with activity significantly modulated by position in the corridor (variance explained >0.05). (*legend continues on next page)*

Neurons are ordered based on the position of their maximum response obtained from odd trials (not shown). Red and blue lines indicate position at +/- 40 cm from maximum response.
**c.** Normalised response of neurons (with peak response between 15 and 85 cm along the corridor) across higher visual areas as a function of the distance from the peak response, obtained by aligning responses in **b.** (top), and population average across these responses (bottom; *black* trace). For comparison, the population average of V1 responses (with peak response between 15 and 85 cm along the corridor) is superimposed in *grey*. The population average across V1 responses differs between areas, because it was obtained from neurons that lie within the corresponding transparent V1 regions in **a**. Red and blue lines same as in **a.**
**d.** Cumulative distribution for the ratio of secondary response divided by peak response across areas, derived from the same response profiles as in **c.** (*black* trace)**.** Cumulative distribution of the same ratio for V1 cells falling within the transparent regions in a. is superimposed in *grey*.
**e.** Distribution of the spatial modulation index, s, across areas *(black*) obtained from cells in c. s = 1 means response profile with a single peak, s = -1 means response profile with two equal peaks. The distribution of the spatial modulation index for the V1 cells falling within the transparent regions in **a.** is superimposed in *grey*. Note how the spatial modulation is shifted towards lower values in the direction from the centre (area LM) towards the periphery (area PM).

Overall, the analysis presented in this section has been crucial for accurately comparing the degree of spatial modulation between areas. The only spatial modulation distribution that was clearly different with high significance ($p < 0.001$) was in A. We also found that V1 was more strongly modulated than LM. However, given that these cells 'look' straight in the centre of the visual scenes which contain the grey texture of the end wall, this apparent difference could be well explained by the fact that LM has broader receptive fields than V1.

## 5.2 Discussion

In this Chapter, I demonstrated that responses in higher visual areas during active behaviour (closed-loop), are modulated by spatial position. To compare the degree of spatial modulation between areas, I focused on subsets of responses with similar preferred azimuth. This was crucial given the differences in the degree of spatial modulation observed already between medial and lateral V1. This analysis revealed that most of the areas exhibited similar degree of spatial modulation to V1. I also found that the only area with a weaker degree of spatial modulation than V1 was LM, but this difference can be explained by the dependence of spatial modulation by retinotopic preference (see 5.2.1).

This analysis also revealed that the only area with stronger spatial modulation than V1 was A. A is anatomically positioned in the antero-medial portion of the visual cortex, a cortical region thought to coincide with the PPC. PPC has been shown to encode visuo-

spatial features of the environment (Krumin, Harris, and Carandini, 2017; Whitlock et al., 2012) and to integrate these signals together with movement related signals (Whitlock et al., 2012). Also, PPC sends dense projections to the retrosplenial cortex, which further projects to the entorhinal cortex (Whitlock et al., 2012). I propose that this particular function and anatomical connectivity of the PPC is reflected in the stronger spatial modulation found here in A.

### 5.2.1 Effects of retinotopic preference on response profiles

Consistent with the findings by Zhuang et al. (2017), our retinotopic maps obtained with wide-field imaging revealed a bias of higher visual areas towards distinct portions of the visual field. Based on these biases, I used a seed-pixel approach to select the maximum possible number of cells with similar preferred azimuths. For instance, because LM is biased towards the centre (0 to 20 deg azimuth), I placed a seed-pixel at approximately 10 deg azimuth. Instead, for PM, which is biased towards the periphery, I placed a seed pixel at approximately 100 deg azimuth. With this approach also V1 was divided into sub-regions depending on the higher visual area in focus. Specifically, by transitioning from lateral visual areas such as LM towards medial visual areas, such as AM and PM, the correlated portions of V1 also shifted gradually from regions analysing the central part of the visual field towards regions analysing the periphery. Notably, we found a concomitant change in the effect of spatial position, with the degree of spatial modulation decreasing the further away from the centre, in both V1 and higher visual areas.

Can the lower degree of spatial modulation found in LM compared to V1 be explained by LM's bias towards the central visual field? Consider for instance a cell with a receptive field at the centre of expansion. If this is a V1 cell, i.e. a cell with a narrow receptive field, it follows that this cell can only 'see' the end grey wall. Instead, an LM cell, which has a broader receptive field, may also pick up the visual cues outside the grey wall. Therefore, the difference in receptive field size between V1 and LM could explain why LM appears to be more influenced by vision compared to the portion of V1 analysing the centre.

The concomitant change in the degree of spatial modulation as a function of preferred azimuth poses an interesting question: what could be the reason for this shift? Providing definite answers to this question has proven challenging, due to the specific design of the virtual environment. Indeed, although I can claim that the peripheral part of the visual scenes is visually-matching at 40 cm away, this is not necessarily the case for its central portion due to the end grey wall. In addition, to render the virtual corridor with a 3d

perspective, landmarks appearing close to the centre of expansion had to be small and of high spatial frequency. Therefore, it is conceivable that mice cannot even see these visual stimuli appearing far away. To better understand the difference in the degree of spatial modulation between the centre and the periphery, additional experiments are required aiming at precisely mapping the receptive fields of neurons in virtual reality.

The results presented so far from experiments in virtual reality focused on closed-loop. But how different could the responses be during play-back of virtual reality sessions (open-loop)? This will be the focus of the next Chapter.

# Chapter 6

Having established that spatial modulation is present across visual areas, and in particular more pronounced in A during closed-loop, I will now focus on open-loop. In this Chapter I will address two main questions: does the reliability of responses vary between active behaviour in virtual reality ('closed-loop') versus play-back of virtual reality sessions ('open loop')? And, is the degree of spatial modulation different in open loop compared to closed-loop?

Here I find that responses in open-loop are less reliable, weaker and exhibit a lower degree of spatial modulation. These findings suggest that active engagement in virtual reality can shape visual responses to the same virtual reality scenes, yielding more reliable responses during the more naturalistic, closed-loop condition. Thus, these results are reminiscent of studies demonstrating that responses during active engagement are qualitatively different compared to passive viewing (Otazu, Tai, Yang, and Zador, 2009; Pho et al., 2018).

One feature of the open-loop condition, is that the animal's running speed and the update speed of the visual scenes ('virtual speed') are independent from one another. Thus, in the last section of this chapter I will also ask whether the two speeds are weighted equally across areas. I will show that different areas weigh the two speeds differently, and consequently, I will provide additional evidence supporting the functional segregation of higher visual areas.

## 6.1 Results

### 6.1.1 Responses in open loop are less reliable than in closed loop

In section **4.1.2**, I demonstrated that the populations of neurons driven reliably by drifting gratings and by behaviour are distinct. Nevertheless, the visual scenes in virtual reality are marginally similar to the drifting gratings we presented. Even the vertical grating landmark can be effectively different from its drifting equivalent, depending on the running speed of the animal or the neurons' preferred position on the visual scene. But how different are responses during active behaviour in virtual reality ('closed-loop') versus passive viewing of the same virtual reality scenes ('open-loop')?

To assess this question, we compared the reliability of responses between closed loop and open loop. Specifically, we estimated the quality of predicting each neuron's activity as a function of position in the virtual corridor (Q; see Methods). When we plotted prediction quality in closed-loop against prediction quality in open-loop, we found that position in closed-loop could explain a larger fraction of the cell's activity than in open-loop (Figure 6-1**a**). Indeed, despite prediction quality being correlated between the two conditions (Pearson correlation coefficient r > 0.4 with $p<10^{-50}$ in all areas), the mean prediction quality in closed-loop was significantly higher than the mean prediction quality in open loop ($EV_{closed\text{-}loop}$-$EV_{open\text{-}loop}$ , mean ± s.e.m.: V1: 0.028 ± 0.001; LM: 0.047 ± 0.003; AL: 0.016 ± 0.001: RL: 0.016 ± 0.001; A:0.018 ± 0.001; AM: 0.024 ± 0.001; PM: 0.026 ± 0.001; two-sample *t*-test: $p<10^{-10}$ in all areas). We further focused on cells with prediction quality higher that 5% and asked how many of these cells responded reliably in one condition but not the other. We found that at least 12% of cells responded reliably in closed-loop but not in open loop ($Q_{closed\ loop} \geq 0.05$ & $Q_{open\ loop} < 0.05$ ; V1: 23%, LM: 31%, AL: 16%, RL: 12%, A: 18%, AM: 19%, PM: 15%). Instead, a smaller fraction of cells had reliable responses in open-loop but not in closed loop ($Q_{open\ loop} \geq 0.05$ & $Q_{closed\ loop} < 0.05$ ; V1: 3%, LM: 3%, AL: 3%, RL: 3%, A: 4%, AM: 4%, PM: 2%). These findings suggest that responses in open-loop are less reliable than responses in closed-loop. In addition, more cells respond reliably in closed-loop alone than in open-loop alone.

Can the difference in reliability be explained by the fact that cells cease to respond in open-loop? This could be particularly likely, because in our experimental protocol open-loop typically followed closed-loop. To test for this possibility, we investigated the

relationship between skewness in closed-loop and skewness in open-loop (Figure 6-1**b**; note that skewness is much higher here compared to Section **4.1.2**, because here I compute skewness from deconvolved spikes, rather than fluorescence). We found that skewness between the two conditions was highly correlated (Pearson correlation coefficient $r > 0.55$, $p < 10^{-10}$ in all areas). More importantly, skewness in closed-loop was not significantly higher than skewness in open-loop ($mean(Skew_{closed\text{-}loop}) - mean(Skew_{open\text{-}loop})$: V1: -0.36; LM: -0.15; AL: -0.14; RL: -0.74; A: -0.24; AM: -0.10; PM: -0.23; two-sample *t*-test: $p>0.95$ in all areas). Therefore, responses to the visual landmarks during passive viewing are less reliable than responses during active behaviour, even though neurons are active during both conditions.

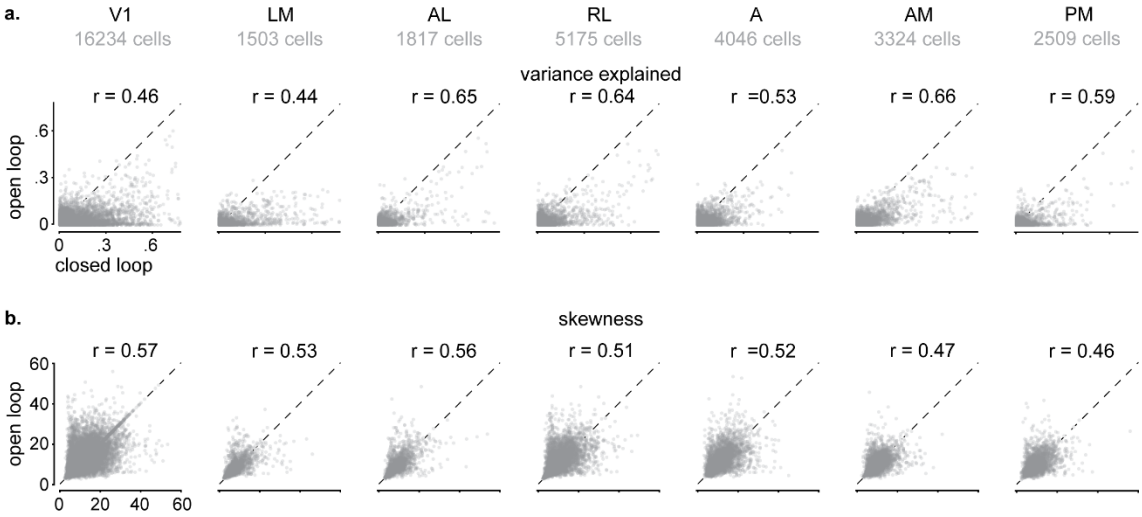


**Figure 6-1: Responses in open loop are less reliable than in closed loop**
**a.** Prediction quality **(**variance explained) of single-cell activity across areas during active behaviour (closed-loop) against skewness of single-cell activity during play-back of virtual reality sessions (open loop). Data are clustered below the diagonal (dotted line), indicating that responses in open loop have lower variance explained than in closed-loop.
**b.** Same as in **a.** for skewness.

Another possibility is that the low reliability in open-loop is due to the fact that visual and self-motion factors are put into conflict. To assess this question, I estimated the reliability of responses (cross-validated variance explained) in open-loop for periods when the speed of the animal was similar to closed-loop. Specifically, I selected segments in open-loop trials in which the running speed of the animal for a given virtual position fell within a 5 cm/s range from the mean closed-loop speed at that virtual position (1 cm bin size; Figure). Comparison of the mean variance between all open-loop trials and trials with similar speed to closed-loop revealed that the reliability of responses increased for similar

speeds in all areas (**Figure 6-4a**; two-sample *t*-test: $p < 10^{-4}$ in all areas). Nevertheless, the reliability was still lower than in closed-loop. This could be either because by comparing to the mean speed in closed-loop we are not excluding the case that on a trial-by-trial basis running speed and virtual speed were not matching. Another possibility would be that behavioural engagement is required for driving responses in virtual reality.

### 6.1.2 Noisier and weaker responses in open loop

Previously I showed that responses of neurons in open-loop are less reliable than responses in closed-loop. Here we asked whether this lower reliability was also reflected in the response profiles as a function of position? For this purpose, we focused on open-loop sessions where the animal was running at least half of the time and estimated response profiles with variance explained higher than 5%. As expected, we found that the percentage of response profiles in open loop with at least 5% variance explained was smaller than in closed-loop (V1: 1,237/14,255 (9% in open loop vs 28% in closed-loop); LM: 227/1,503 (15% vs 42%); AL: 212/1,774 (12% vs 25%); RL: 450/5,087 (9% vs 17%); A: 300/3,105 (10% vs 24%) ; AM: 459/2,923 (16% vs 31%); PM: 145/2,272 (6% vs 20%)). We cross-validated the robustness of the response patterns by arranging response profiles estimated from even trials according to each cell's 'preferred position' of maximal firing obtained from odd trials (**Figure 6-2a**). When we estimated the mean response across cells as a function of distance from the 'preferred' position, we found a striking difference between areas with receptive fields biased towards the central visual field (LM, AL and RL) and areas with receptive fields biased towards the periphery (A, AM and PM). Areas LM, AL and RL exhibited noisier mean normalised response than in closed loop: responses to the visually matching segments peaked at lower values (mean ± s.e.m.: LM: 0.66 ± 0.02 in open loop vs 0.80 ± 0.01 in closed-loop, AL: 0.63 ± 0.03 vs 0.72 ± 0.02, RL: 0.60 ± 0.02 vs 0.72 ± 0.01); the mean response trace typically lacked well defined secondary peaks at 40 cm away (**Figure 6-2b**; compare cyan-open loop with black-closed loop).

Instead, in areas A, AM and PM the mean response to the visually matching segments peaked at approximately similar values as in closed-loop (A: 0.61 ± 0.02 in closed loop vs 0.57 ± 0.01 in open loop, AM: 0.71 ± 0.02 vs 0.70 ± 0.01, PM: 0.66 ± 0.03 vs 0.67 ± 0.02). Especially in areas AM and PM, which in our dataset were exclusively 'looking' in the periphery, the mean response exhibited pronounced secondary peaks. PM in particular had higher secondary peaks in open-loop than in closed-loop, suggesting that the population response profile in PM had become more visual than during active behaviour.

Overall, these results confirm that during open-loop there are fewer cells eliciting reliable responses. Nevertheless, responses of cells with receptive fields in the periphery where the optic flow is higher, are almost as reliable as in closed-loop.

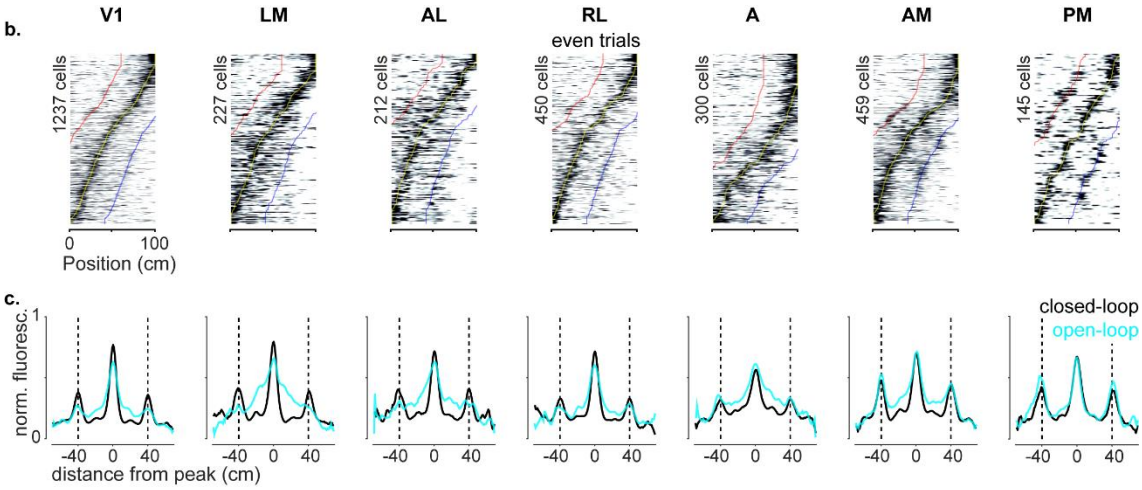


**Figure 6-2: Noisier responses in open loop in the centre but not in the periphery**
**a.** Normalised response in open loop as a function of distance in the virtual environment, obtained from half of the data (even trials) that were not used to order the responses, for neurons across higher visual areas with activity significantly modulated by position in the corridor (variance explained > 0.05). Neurons are ordered based on the position of their maximum response. Red and blue lines indicate position at +/- 40 cm from maximum response. Responses obtained from areas biased towards the central visual field (LM, AL, RL) are noisier compared to areas biased towards the periphery (AM, PM).
**b**. Population average across responses aligned to maximum peak obtained from even trials, for neurons with peak response between 15 and 85 cm along the corridor (*cyan*). For comparison, the population average obtained from closed-loop is superimposed in *black* (same data as in 5.1.1).

### 6.1.3 Lower degree of spatial modulation in open loop

Having shown that responses in V1 and higher visual areas are strongly modulated by spatial position during active behaviour (Chapters 3 and 5), we now asked whether responses of the same cells were different during open-loop. To this end, we arranged response profiles in open-loop (test set) according to the preferred position of cells in closed-loop (train set). Some cells exhibited distinct, albeit weaker, responses to the visually matching segments (**Figure 6-3a**), but overall, the response profile pattern across the population was typically weaker in open-loop than in closed-loop, indicating that many cells ceased to respond to the visually matching segments during open loop (**Figure 6-3b**). How different were the spatial modulation ratios between the two conditions? The median modulation ratio in open loop was significantly higher than in closed loop in all areas, suggesting that response patterns were more visual in open loop (**Figure 6-3c** *cyan*: open

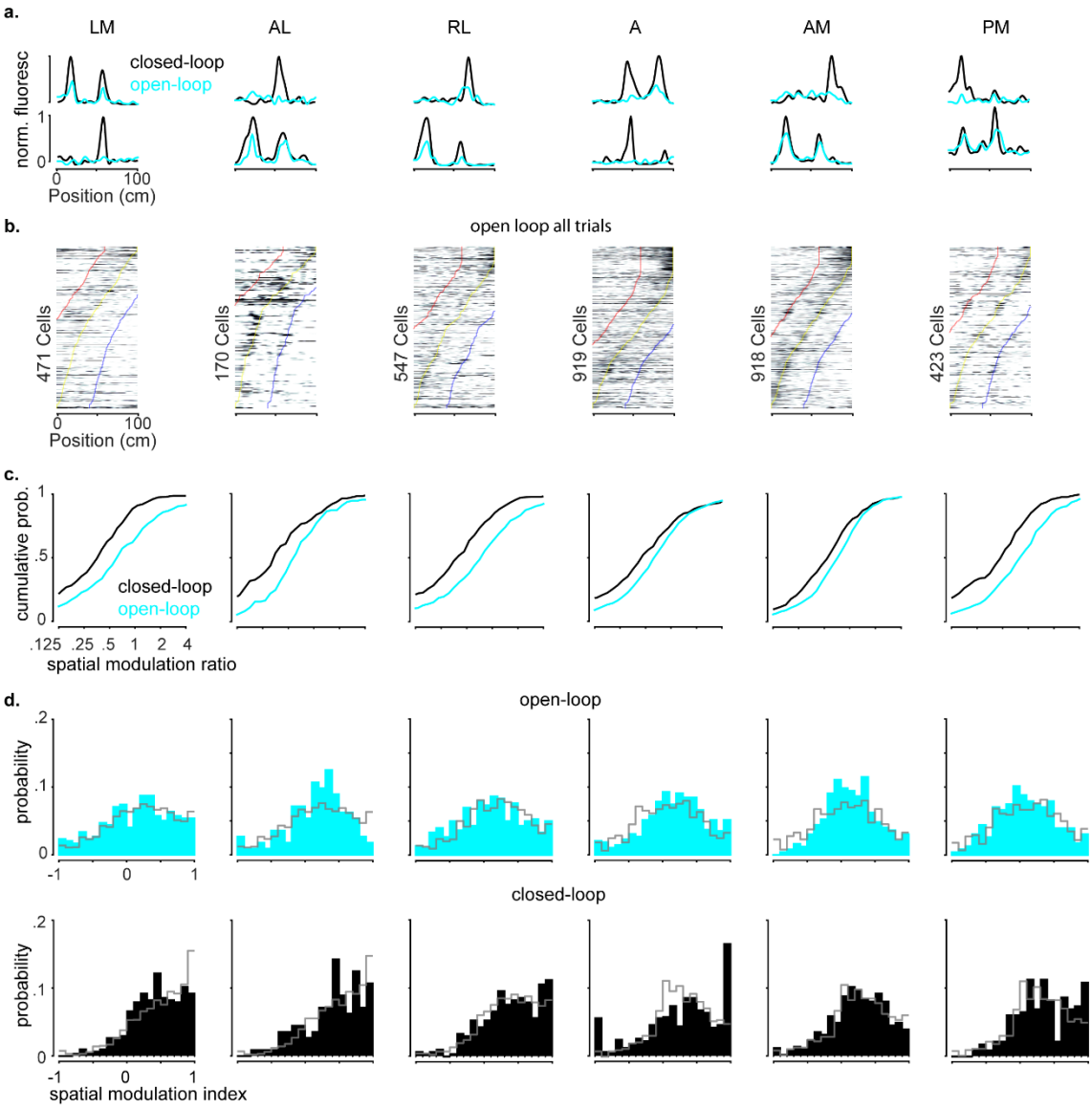


**Figure 6-3: Lower degree of spatial modulation in open loop**
**a.** Single-cell normalised response profiles in open loop (*cyan*) and closed-loop (*black*) across higher visual areas.
**b.** Response as a function of distance in the virtual environment obtained from all trials in open loop, for neurons that lie within the 'transparent' regions of the seed-pixel cross-correlation maps (**Figure 5-3**) and with activity significantly modulated by position in the corridor (variance explained >0.05). Responses were normalised by the maximum peak in closed-loop. Neurons are ordered based on the position of their maximum response obtained from all trials in closed-loop (not shown). Red and blue lines indicate position at +/- 40 cm from maximum response.
**c.** Cumulative distribution for the ratio of secondary response divided by peak response across areas, derived from response profiles in open loop with peak response between 15 and 85 cm along the corridor (*cyan* trace)**.** Cumulative distribution of the same ratio for response profiles of the same cells in closed-loop is superimposed in *black*.
**d**. Distribution of the spatial modulation index, s, across areas *(cyan*) obtained from cells in **c** during open loop. s = 1 means response profile with a single peak, s = -1 means response profile with two equal peaks. The distribution of the spatial modulation index for the V1 cells falling within the transparent regions of the seed-pixel correlation maps is superimposed in *grey* (top). For comparison, the corresponding distributions in closed-loop are also shown (bottom, same data as in **Figure 5-3**).

loop, *black*: closed-loop; $\text{median}_{\text{open-loop}} - \text{median}_{\text{closed-loop}}$ : V1: 0.29; LM: 0.24; AL: 0.17; RL: 0.36; A: 0.23; AM: 0.20; PM: 0.35; two-sample KS test: $p < 10^{-4}$ in all areas).

We next used the same approach as in the previous section. Specifically, we obtained distributions of the spatial modulation index in open loop for cells 'looking' at the same portion of the visual field (**Figure 6-3d top**). From these distributions it became evident that responses with spatial modulation index close to 1 (single-peaked responses) were markedly reduced, both in higher visual areas (*cyan*) and in V1 (*grey* trace) (**Figure 6-3** compare **d top** to **d bottom**). In addition, the difference in distributions between V1 and higher visual areas was less evident in open loop than in closed-loop, albeit still statistically significant for areas A and AM (two-sample K-S test, LM: p = 0.43, AL: p = 0.18, RL: p = 0.23, A: p = 0.003, AM: p = 0.0004, PM: p = 0.67).

Overall, spatial modulation is weaker in open-loop regardless of the animal's running behaviour, but how much weaker is it during periods when the animal was stationary. To assess this question, I selected periods in open-loop when the animal was stationary (running speed lower than 5 cm/s). For each area, I estimated the median spatial modulation ratio obtained from stationary periods in open-loop and found that it was significantly higher than the median spatial modulation ratio obtained from all data in open-loop (Figure; KS-test, $p>10^{-5}$). In fact, for most areas the median spatial modulation was close to unity, suggesting that the overall spatial effect across the population vanished.

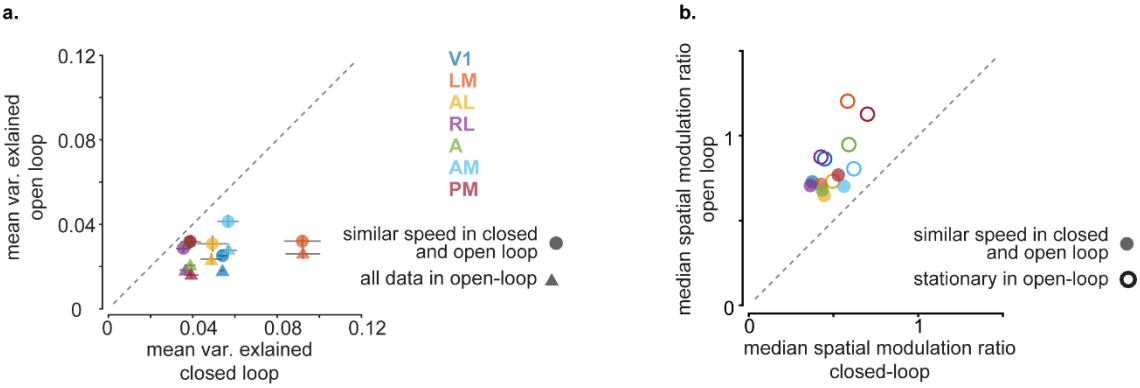


**Figure 6-4: Contribution of self-motion and visual inputs to the main effects in open-loop**
**a.** The reliability of responses in open-loop is more similar to the reliability of responses in closed-loop in all areas when considering similar speeds between the two conditions (*filled circles*).
**b.** Responses in open loop are much more visual (median spatial modulation ratio ~ 1) when the animal is stationary (*open circles*).

### 6.1.4 Relative influence of virtual and running speed

One feature of the open-loop condition is that the update speed of the visual scenes ('virtual speed') does not match the running speed of the animal. Thus, open loop gives the ability to disentangle the influence of the two speed types, thereby differentiating the effects of visual (virtual speed) from non-visual (run speed) factors (see Saleem et al., 2013).

Similarly here, we sought to evaluate the influence of virtual and run speed, by turning to the open-loop condition and asking whether virtual and run speed were able to predict the neurons' activity. For each neuron we obtained cross-validated speed profiles as a function of both speeds (**Figure 6-5a**), and found that many speed profiles could explain at least 5% of the neurons' activity in all areas (V1: 3,432/14,255 (24%); LM: 474/1,503 (31%); AL: 640/1817 (35%); RL: 1457/4753 (31%); A: 1043/3105 (34%); AM: 1094/2920 (37%); PM: 797/2002 (40%)). We next asked whether the influence of visual and run speed could be captured by a multiplicative model; that is a model where the effects of the two speeds are separate from one another ('separable model'). The separable model performed as well as the non-separable model, exhibiting prediction qualities that were very similar to the prediction qualities of the non-separable model (**Figure 6-5b**; Pearson correlation coefficient $r > 0.8$ for most areas except PM, $r = 0.74$ for PM, $p < 10^{-10}$). Accordingly, the mean prediction quality across the population was very similar for the two models (difference in mean prediction quality between separable and joint model: less than 0.002 in all areas; two-sample *t*-test: $p > 0.2$ in all areas), suggesting that the speed influence was well captured by a multiplicative model of virtual and running speed.

By verifying that the separable model provided accurate predictions, we established that the influence of run and virtual speed was independent from one another. Therefore, we could now ask whether the contribution of run speed alone and virtual speed alone was the same across areas. To assess each speed's contribution, we estimated how much of the prediction quality of the separable model could be attributed to one speed type ($Q_{\text{virtual or run}}/Q_{\text{separable}}$, 'relative prediction quality'). The relative prediction quality of run speed and virtual speed varied across areas (**Figure 6-5c**). The relative prediction quality of virtual speed was higher in V1, LM and RL, but lower in areas A and AM. Similar levels of relative prediction quality for the two speeds were found in areas AL and PM.

To further assess the relative weighing of virtual and running speed in each area, we estimated a prediction quality index $((Q_{\text{virtual}} - Q_{\text{run}})/(Q_{\text{virtual}} + Q_{\text{run}}))$. A prediction quality index close to zero means equal contribution of virtual and running speed.

At the two extremes, an index close to -1 means almost exclusive influence by running speed, whereas an index close to 1 means that a cell's activity is influenced almost exclusively by virtual speed. Distributions of the prediction quality index across the population revealed a bias towards indices equal to 1 in most areas (**Figure 6-5d,** V1, LM, AL, RL, PM). Only areas A and AM did not exhibit such bias, consistent with the relative

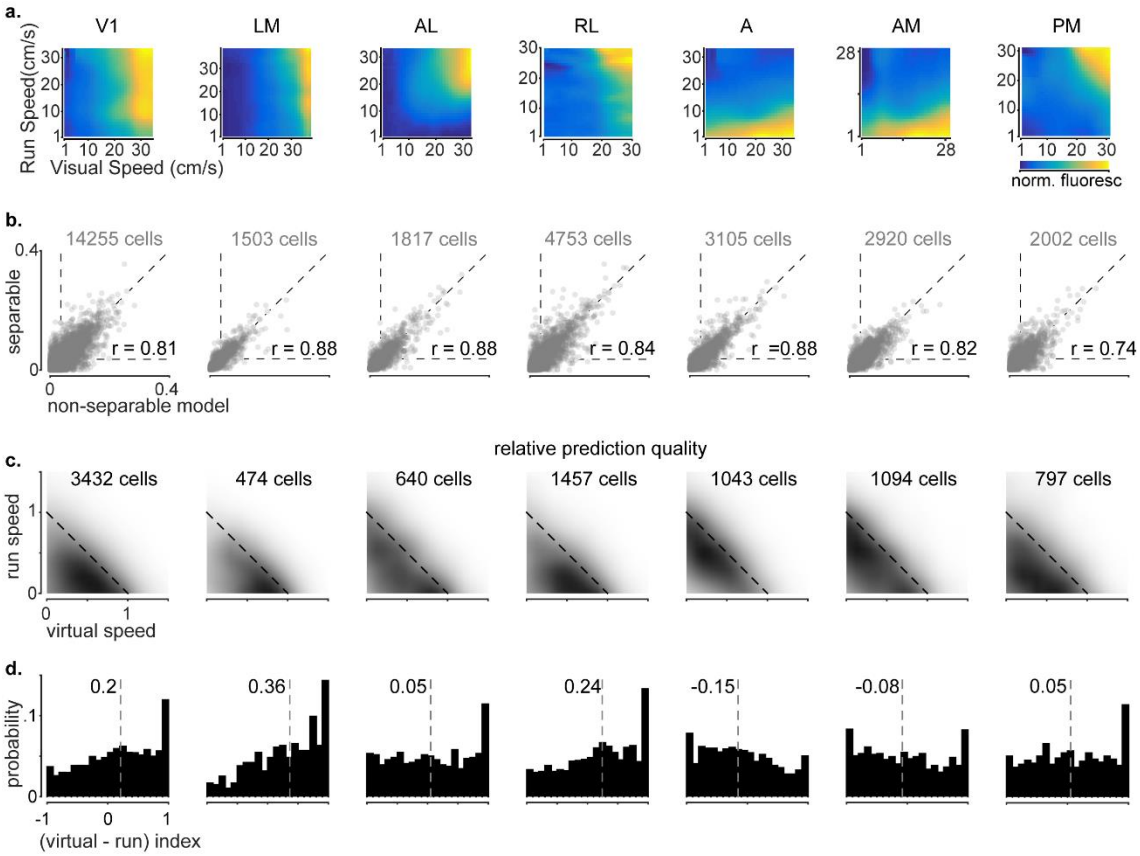


**Figure 6-5: Relative weighing of virtual and running speed varies across areas**
**a.** Examples of 2-dimensional response maps as a function of virtual and running speed across areas.
**b.** Prediction quality (Q) across areas obtained from a joint model of running and virtual speed against prediction quality obtained from a separable model of the two speeds. r: Pearson correlation coefficient.
**c.** Relative prediction quality of virtual speed ($Q_{virtual}/Q_{separable}$) across areas against relative prediction quality of running speed ($Q_{running}/Q_{separable}$). Different areas weigh virtual and running speed differently.
**d**. Prediction quality index, p, (Qvirtual – Qrun )/( Qvirtual + Qrun) across areas. p = -1 means exclusive influence by running speed, p = 1 means exclusive influence by virtual speed. Dotted line indicates the mean across the distribution. Only areas A and AM have negative mean prediction quality index.

prediction quality for running speed being higher in these two areas. The mean prediction quality index also varied accordingly. A and AM had mean prediction quality index less than zero (mean ± s.e.m.; Wilcoxon signed rank test, A: -0.15 ± 0.02; $p < 10^{-10}$, AM: -0.08 ± 0.02; p = 0.0082). V1, LM and RL had the highest mean (V1: 0.20 ± 0.01, LM: 0.36 ± 0.02, RL: 0.24 ± 0.01; $p < 10^{-10}$ in all three areas). AL and PM had mean prediction quality index that was closest to zero (yet, the mean index was statistically higher than zero for PM: 0.05 ± 0.02; p = 0.0021, AL: 0.05 ± 0.03, p = 0.016).

These results suggest that visual areas weigh the visual and non-visual component of speed differently. The highest bias towards virtual speed is found in LM, whereas on the other side of the spectrum, the highest bias towards run speed is found in A. The rest of the areas lie somewhere in-between, with AL for instance, weighing the two speeds equally, and AM being slightly more biased towards running speed.

Taken together, the results presented in Chapters 3, 5 and 6 indicate that responses in V1 and higher visual areas are modulated by spatial position predominantly during active behaviour. Instead during play-back of the virtual reality scenes, responses are weaker, less reliable and significantly less modulated by spatial context.

## 6.2 Discussion

In this chapter I showed that response profiles in open loop differ in various aspects from responses in closed-loop.

First, prediction quality in open loop was lower. This could not be explained by different levels of arousal; first because the animals ran most of the time during both closed- and open loop, and second because I excluded the few sessions during which animals were immobile for at least half of the time. The only occasional difference between the two conditions was the absence of reward in open loop. I find it unlikely that the lack of reward can explain fully the difference in reliability between closed-loop and open loop, especially because some animals were not rewarded in any of the two conditions.

Second, responses in open loop were weaker than in closed loop. This was particularly evident when we scaled response profiles in open loop by the maximum response in closed-loop. These results are reminiscent of two previous studies which assessed the effects of active engagement in a non-navigational task (Otazu, Tai, Yang, and Zador, 2009;

Pho et al., 2018). Unlike passive viewing, engagement in an auditory task suppressed responses in the rat auditory cortex (Otazu, Tai, Yang, and Zador, 2009). Another study also showed that responses in the PPC were much weaker during passive viewing compared to active engagement in a visual discrimination task, whereas responses in V1 remained the same (Pho et al., 2018). In contrast, I found that responses already in V1 were suppressed during replay of the virtual reality. This discrepancy between the current study and Pho et al.(2018) could be due to the difference in behavioural context and the fact that virtual reality tasks create a more naturalistic visual experience.

Third, the degree of spatial modulation was substantially reduced in open loop if animals were not stationary. Perhaps spatial modulation was still present to some extent in open loop because just the replay of the virtual environment is sufficient to create a sense of space. In fact, a recent study probed the effect of passive viewing on the hippocampal place fields and demonstrated that 25% of identified place cells maintained their firing fields (Chen, King, Burgess, and O'Keefe, 2013; see also Terrazas et al., 2005). In addition, learning-related enhancement of visual responses, acquired during engagement in a virtual reality task, is still present, albeit reduced, during replay of previous sessions to fully trained, anaesthetised animals (Poort et al., 2015).

Taken together, two were the main results in the open-loop condition: 1. spatial modulation of visual responses was weaker and 2. visual responses were less reliable. Yet, the reasons for these differences compared to closed-loop are not clear. To provide further insight into these topics I performed additional analysis that focuses on isolating the effects of visual inputs from the physical inputs.

With respect to result 1. I asked whether influence by spatial context can arise in the presence of visual inputs alone, i.e. during periods when the animal is stationary. I found that median spatial modulation was close to 1 (where 1 means two equal peaks), suggesting that the overall spatial effect across the population vanished. These results indicate that visual inputs alone are not sufficient for giving rise to the spatial modulation of visual responses observed in closed-loop, suggesting that visual and self-motion factors need to be combined to give rise to a single modulatory estimate of position.

With respect to result 2. I assessed to which extent the low reliability of responses in open-loop can be explained by the cue conflict between visual and physical inputs. I focused on open-loop trials in which running speed was similar to the mean closed-loop speed and found that the reliability of responses increased for similar speeds in all areas.

This result indicated that the lower reliability in open-loop can be, at least partly, explained by the mismatch between running speed and virtual speed.

Although reliability in open-loop increased after accounting for similar speeds to closed-loop, it was still lower than in closed-loop in many areas. This mismatch could be due to methodological reasons. For instance, instead of comparing to the mean speed in closed-loop, speeds between closed- and open-loop should be compared on a trial-by-trial basis. Also, because even small differences between virtual and physical distance may affect visual responses (see Chapter 7), ideally comparison should be done only for speeds in open-loop that exactly match speeds in closed-loop. Unfortunately, such a comparison was not possible due to the remaining dataset being too small.

### 6.2.1 Effects of retinotopic preference on response profiles

Similar to the closed-loop condition (section **5.1.2**), the effect of retinotopic preference was also reflected in the response profiles of cells responding reliably in open loop. The response profile pattern and population average of areas biased towards the centre (LM, AL, RL) were much noisier compared to areas biased towards the periphery (AM, PM). This difference could be due to the differences in optic flow, which is strongest in the periphery. Consistent with this hypothesis, PM, the area with the highest bias towards the periphery, exhibited the most pronounced responses. The strong influence of these responses by optic flow in open loop could also explain why PM's response profile pattern featured higher responses to the non-preferred position in open loop than in closed-loop, suggesting a stronger influence by vision in this area during replay of the visual scenes.

### 6.2.2 Diverse tuning to virtual or run speed in higher visual areas

Because in the open loop condition running speed is decoupled from the update speed of the visual scenes, I was able to assess the influence of the two types of speed separately. To investigate possible differences between areas, I estimated the quality of predicting single-cell responses from running speed alone or virtual speed alone. This analysis revealed a progressive shift from responses being more influenced by virtual speed in V1 and LM towards strongest influence by running speed in A and AM.

These findings are in agreement with several lines of evidence suggesting a distinct functional role of higher visual areas. First, we found that AL was more strongly modulated by running speed compared to LM. This result is consistent with previous findings suggesting that despite the two areas sharing the same border, their anatomical and

functional connectivity, and consequently their functional role is distinct (Glickfeld, Andermann, Bonin, and Reid, 2013; Wang, Gao, and Burkhalter, 2011). The functional difference between LM and AL is also supported by a recent study showing that spontaneous running modulates more strongly AL than PM (Huh et al., 2018). Another intriguing segregation arising from these results is the stronger modulation by running speed found only in A and AM. Given the current view that A and AM comprise part of the PPC, the stronger modulation by running speed found in these areas is consistent with the strong connectivity of the PPC to the secondary motor cortex (Wilber et al., 2014), and also with the proposed role of the PPC as an area encoding movement-related signals during navigation (McNaughton et al., 1994; Nitz, 2006).

The stronger influence by running speed found in A and AM gives rise to an interesting question: how does the distance travelled influence responses in these areas? Indeed, given that running speed is actually distance travelled per unit time, it is possible that the spatial modulation found in these areas (which was particularly stronger in A compared to V1) does not reflect encoding of spatial position but encoding of distance travelled instead. In a broader sense, perhaps the spatial modulation observed in all areas, including V1, may in fact arise from the influence by distance travelled. This question is explored in detail in the next chapter.

# Chapter 7

Estimates of distance travelled, through integration of self-motion, are known to contribute to the representation of self-location. In hippocampus, movement-related signals are combined with visual information about the environment and shape the spatial representation (Gothard, Skaggs, and Mcnaughton, 1996). If navigational signals are also present in visual cortex (Chapters 3 and 5), we may thus expect that the distance travelled in the environment similarly affects the spatial representation in visual cortex.

To assess the impact of distance travelled on V1 responses, a recent study took advantage of a virtual reality environment to perform gain manipulation experiments in mice (Fournier*, Saleem*, Diamanti* et al., in preparation). Specifically, by changing the relationship between running speed and virtual speed, animals had to run 20% more (gain decrease) or 20% less (gain increase) to reach the reward location (**Figure 7-1a**). While mice performed the task, the authors recorded simultaneously from neurons in CA1 of the hippocampus and in V1. As expected, CA1 place fields shifted along the track depending on the distance travelled (Chen, King, Burgess, and O'Keefe, 2013; Gothard, Skaggs, and Mcnaughton, 1996): place cells fired earlier on the track when mice had to cover longer distances to the reward zone (low gain; **Figure 7-1b**; *cyan*); conversely, they fired later on the track when mice had to cover shorter distances to the reward (high gain; **Figure 7-1b**; *pink*). Surprisingly, V1 cells showed the same behaviour as CA1 place cells; they fired earlier or later on the track when the distance was changed, although visual landmarks remained in the same position, thus indicating that V1 neurons are sensitive to distance travelled (**Figure 7-1c**).

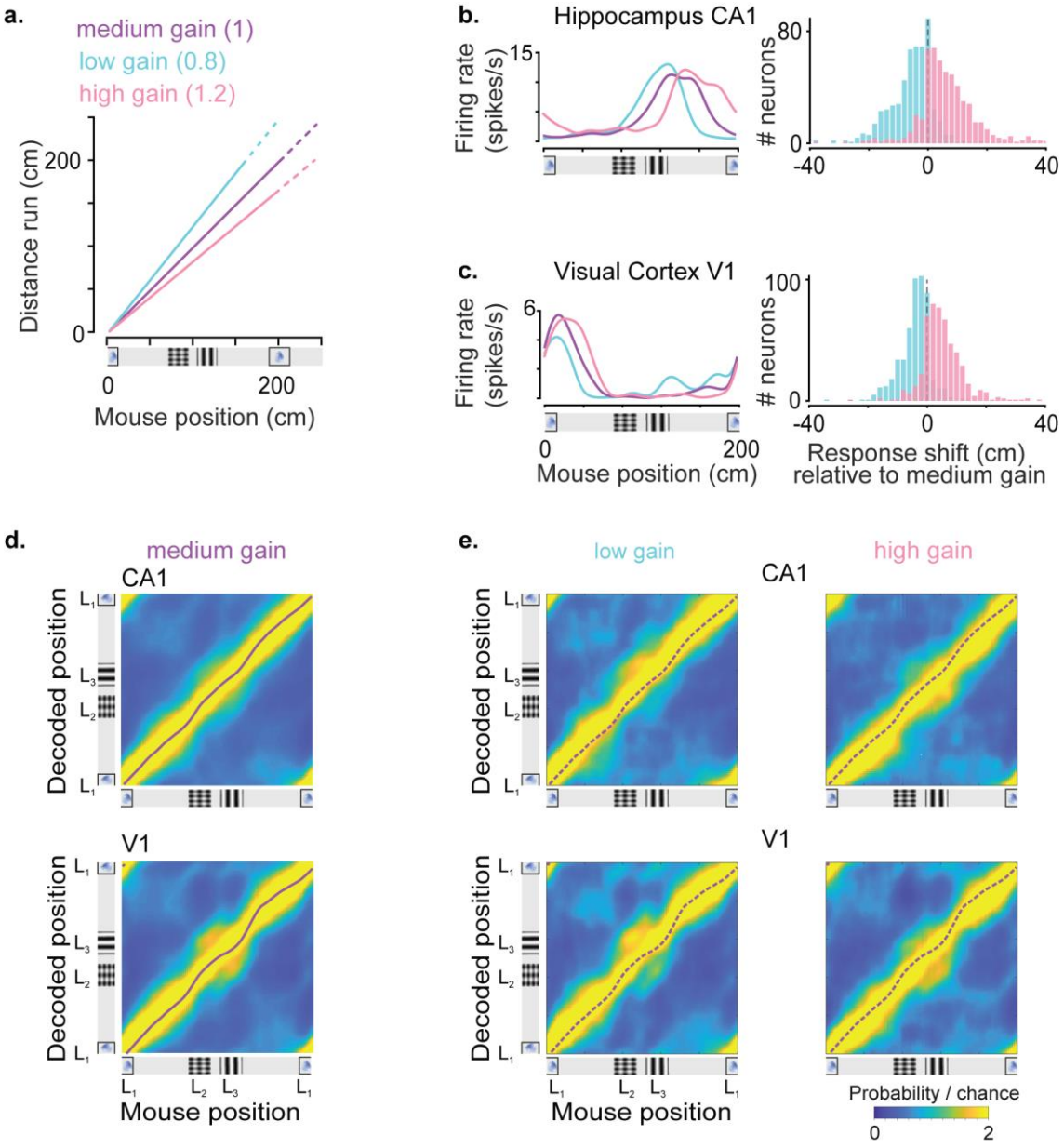


**Figure 7-1 *Courtesy of Julien Fournier*: Influence of distance signals on primary visual cortex and hippocampus during navigation**

**a.** Schematics of the gain manipulation experiments performed in a virtual circular maze. At medium gain the running speed of the animal matched the update speed of the visual scenes in virtual reality. At low gain, animals had to run 20% more (gain decrease) to cover the same distance as in medium gain. At high gain, animals had to run 20% less (gain increase) to cover the same distance as in medium gain.

**b.** Example place field recorded in hippocampus across gain conditions (*left*) and distribution of the response shift across the CA1 population relative to medium gain (*right*).

**c**. Same as in **b** for visual responses recorded in V1.

**d.** Decoding the animal's position from CA1 (*top*) or V1 (*bottom*) population activity at medium gain. The actual position of the animal could be decoded accurately by both populations.

**e.** Changing the gain of the VR shifted the position decoded from the CA1 or V1 population in a more or less uniform manner. At low gain, CA1 and V1 populations represented a position ahead of the animal (*left*); at high gain, CA1 and V1 populations represented a position behind the animal.

To further characterise the influence of the distance run on V1 and CA1 spatial representations at the population scale, the authors used a Bayesian decoder ((Zhang, Ginzburg, McNaughton, and Sejnowski, 1998)) to predict the animal's position from either CA1 or V1 population activity (**Figure 7-1d**, **e**). Interestingly, decoding accuracy was similar between CA1 and V1. At medium gain, the actual position of the animal could be decoded accurately not only by the CA1 population, but also from the V1 population (**Figure 7-1d**), consistent with the presence of position correlates in V1 activity (Chapter 3). Changing the gain of the VR shifted the position decoded from the CA1 or V1 population in a more or less uniform manner. When the distance increased (low-gain condition), CA1 and V1 populations represented a position ahead of the animal (**Figure** 7-1**e**; *left*); conversely, when the distance decreased (high-gain condition), CA1 and V1 populations represented a position behind the animal (**Figure 7-1e;** *right*). Importantly, this shift was smaller than what would be expected if distance run was the only influencing factor, indicating that positions encoded by V1 and CA1 neurons resulted from the integration of both self-motion and visual information.

Overall, these results are strikingly consistent with the findings described in Chapter 3 and further indicate that activity in V1, like in CA1, can be used to decode position based on information provided by visual cues, but also, to some extent, by distance travelled.

Here we asked how the distance travelled affected the spatial representation across three different visual areas: V1, AL and PPC. Among higher visual areas, we chose AL because it has been classified as an area belonging to the 'dorsal' stream (Wang, Gao, and Burkhalter, 2011); we also chose PPC because of multiple evidence from experiments in rats indicating that PPC processes and integrates movement-related signals (McNaughton et al., 1994; Wilber et al., 2014). Our results show that the distance run by the animal affects the representation in all three areas, independently of the visual cues. This effect was markedly stronger in PPC than in V1 or AL: in the low-gain condition, the average position decoded from PPC neurons was further ahead compared to the position decoded from V1 or AL neurons; in the high-gain condition position decoded from PPC was further behind.

To the best of our knowledge, this is the first time that gain manipulation experiments are performed in PPC. Our results suggest that PPC neurons are more sensitive to distance run compared to neurons in V1, consistent with the prominent role of PPC in processing self-motion signals.

## 7.1 Results

### 7.1.1 Mice successfully perform a spatial task in a virtual corridor

To assess the influence of distance run on visual responses during virtual navigation, we trained mice to run through a 400 cm circular corridor and actively lick for water reward at two positions 200 cm apart (**Figure 7-2a**). The corridor contained three types of landmarks, $L_1$, $L_2$ and $L_3$. The doublet $L_2$ and $L_3$ (a plaid and a grating) were placed half way through the reward zone and were fixed throughout. Landmark $L_1$ was placed at the reward location and alternated between a grating or plaid every half cycle, ensuring that mice were not performing a cue detection task. Mice learned to selectively lick at the reward location in most trials (**Figure 7-2b**). The few trials during which they made mistakes were excluded from further analysis.

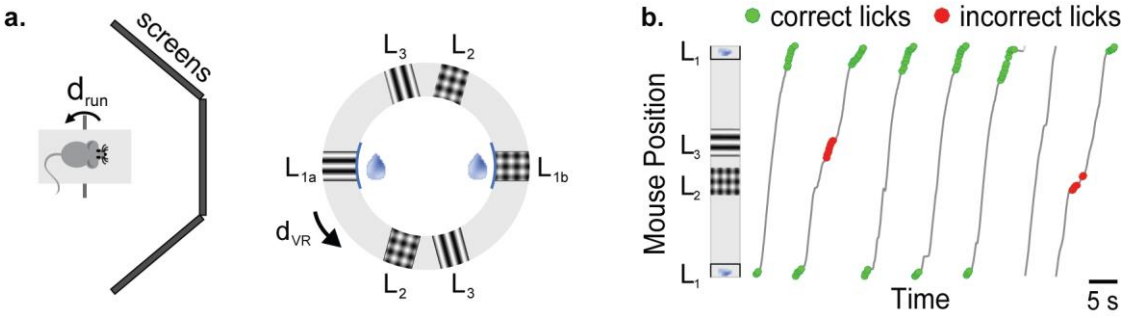


**Figure 7-2** ***Courtesy of Julien Fournier*****: Mice successfully perform a spatial task in a virtual corridor.**
**a.** Mice are head-fixed in front of 3 LCD screens while free to run on a polystyrene wheel which allowed movement along a single dimension (forwards-backwards). The animal's running speed, $d_{run}$, was captured online by a rotary encoder (*left*). The virtual reality environment is a circular maze containing three types of landmarks, $L_1$, $L_2$ and $L_3$. The doublet $L_2$ and $L_3$ (a plaid and a grating) were placed half way through the reward zone and were fixed throughout. Landmark $L_1$ was placed at the reward location and alternated between a grating or plaid every half cycle (*right*).
**b.** Mice had to selectively lick at the reward location. They performed the task successfully in most trials (*green* circles: correct licks; *red* circles: incorrect licks).

### 7.1.2 Distance run influences response profiles

To ensure we were able to obtain robust response profile patterns in this task like those obtained in the linear corridor, we first focused on the medium gain condition. For each area (V1, AL or PPC) we selected cells whose maximum response in the medium gain was at least two s.e.m. higher than the mean and sorted their response profiles as a function of the position at which each cell fired maximally. Notably, we confirmed the observation by Fournier et al. (in preparation) that the response profile pattern was the same between the two semi-circles of the maze, even though the visual cues at the reward locations were

different (data not shown), presumably because the reward and behavioural modulation of responses masked the expected differences due to the difference in visual images. Thus, for all subsequent analysis response profiles in the two semi-circles were pulled together, yielding responses as a function of position along a 200-cm-long virtual corridor. Consistent with previous findings in the linear corridor, neurons responded robustly to the visual landmarks, but they also exhibited a preferred position at which they fired maximally, thereby tiling the whole virtual maze. (**Figure 7-3**).

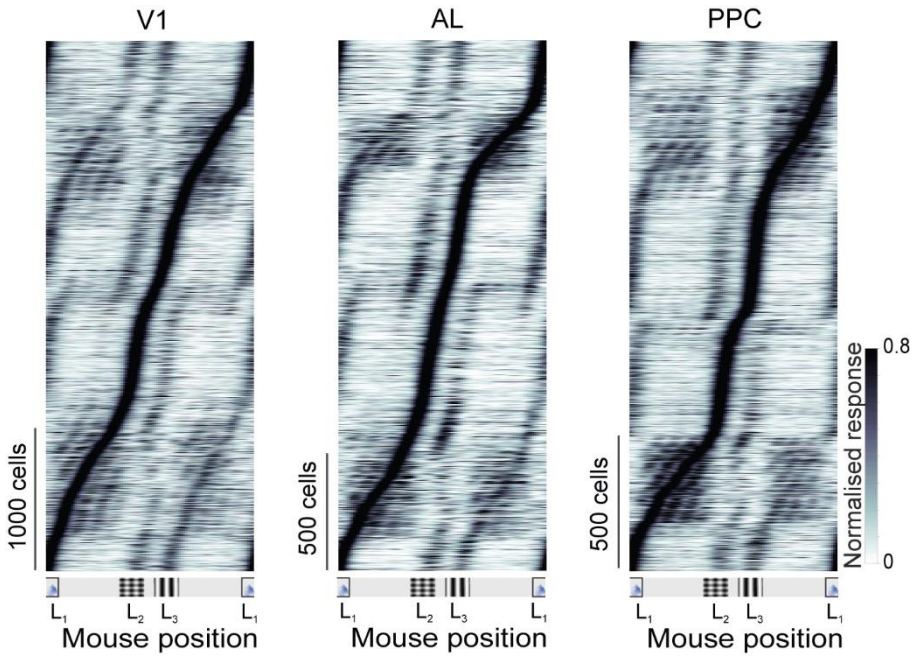


**Figure 7-3: response profiles across the population reveal preference for one position along the virtual corridor in V1, AL and PPC.**
Response profile patterns were obtained from cells whose maximum response in the medium gain was at least two s.e.m. higher than the mean. Response profiles were sorted based on the position at which each cell fired maximally.

We then asked how the response profiles of these cells were influenced by distance run. To assess this question, we focused on cells whose response profiles showed a correlation > 0.75 between low/high and medium gain conditions. We then used the preferred position of maximal firing obtained in the medium gain condition (train set) to arrange response profiles during low and high gain trials (test sets). When we expressed these patterns as a function of distance from the preferred position, we found that in the low gain condition the cells tended to respond before the preferred position at medium gain; in the high gain condition they tended to respond after the preferred position at medium gain (**Figure 7-4a**). Therefore, the maximum response shifted slightly but systematically as a function of gain change.

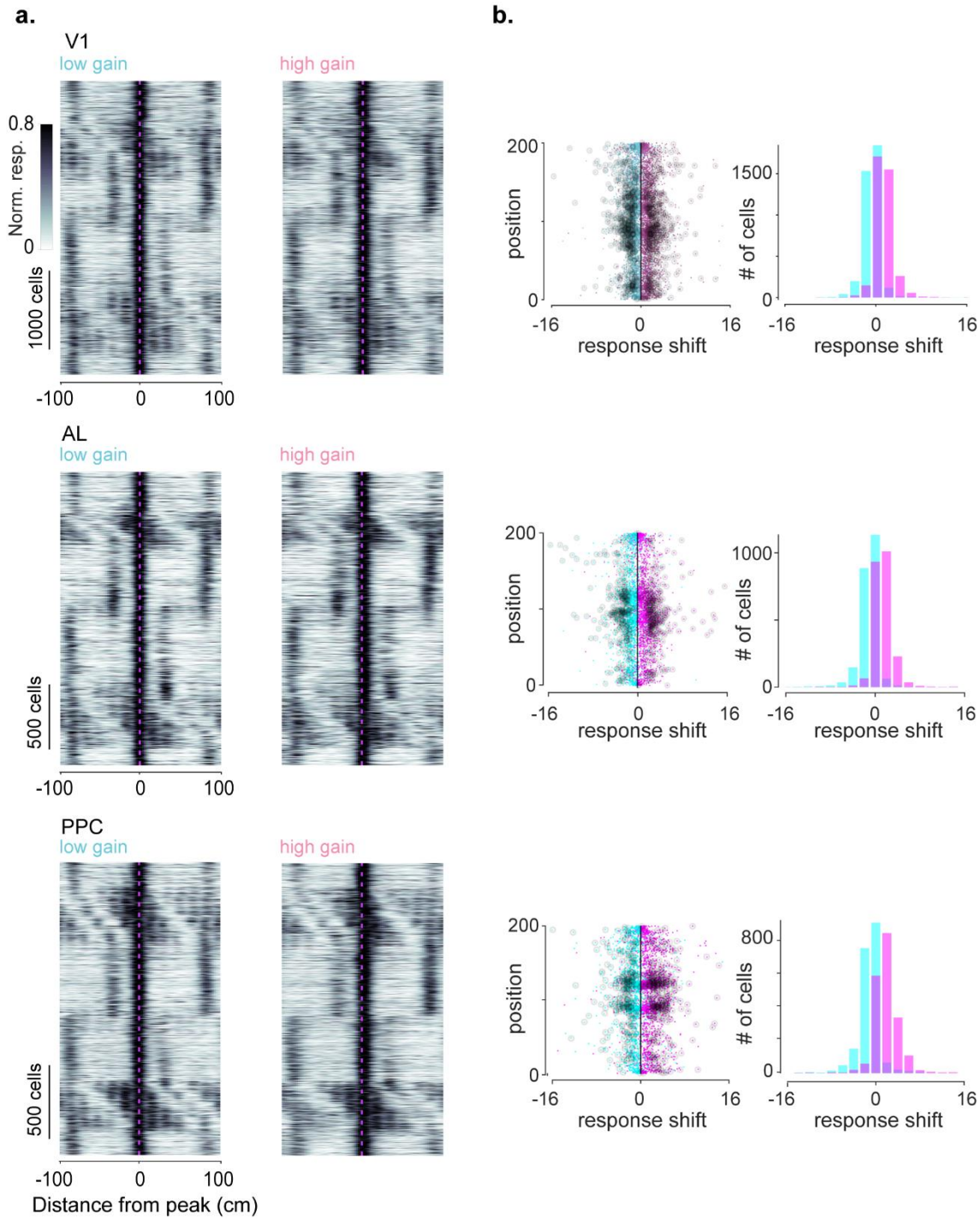


**Figure 7-4: Distance run influences response profiles.**
**a.** Response profiles as a function of distance from the preferred position in V1 (*top*), AL (*middle*) and PPC (*bottom*) at low gain (*left*) and high gain (*right*). The preferred position (*red dotted* line) was obtained from the medium condition for cells whose response profiles showed a correlation > 0.75 between low/high and medium gain conditions.
**b.** Distributions of single-cell response shifts at low gain (*cyan*) and high gain(*pink*), as a function of position (*left*) and distributions of shifts across all positions (*right*) in V1, AL and PPC (*black* circles: cells with a significant shift (p-value < 0.025) in their response profiles).

To further quantify this shift, for each cell we measured the position of the peak in the cross-correlation between the train (medium) and the test (low or high) condition. We next obtained distributions of single-cell response shifts, as a function of position (**Figure 7-4b top**). Cells with a significant shift in their response profiles (**Figure 7-4b** *top*; *black* circles) were identified using a 20-fold cross-validation procedure (p-value < 0.025). We found that for a given condition, at least 14% of cells had significantly shifted response profiles in all areas (low gain, high gain: V1: 16%, 14%; AL: 14%, 15%; PPC: 14%, 20%). We also estimated distributions of shifts across all positions (**Figure 7-4b** *bottom*). These distributions confirmed a systematic shift towards the left for the low gain condition, and towards the right for the high gain condition in all three areas. The mean shift across the whole population was of the order of at least 1 cm in all areas (Mean ± s.e.m.: low gain: V1: -1.2 ± 0.01, AL: -1.2 ± 0.01, PPC: -1.3 ± 0.04; high gain: V1: 1.04 ± 0.02, AL: 1.2 ± 0.02, PPC: 1.8 ± 0.03; one-sample t-test: $p << 10^{-50}$ for all gain conditions and in all areas). Thus, these results extend previous findings by Julien et al. (in preparation) and show that beyond V1, neurons in areas AL and PPC are influenced by self-motion cues, even when visual cues remain in the same place.

### 7.1.3 Gain changes induce shifts in decoding position

To further characterise the effect of distance run at the population scale, we trained a Bayesian decoder on the medium gain condition to estimate the posterior probability of the animal being at a specific location, given the firing activity of neurons as a function of time. We next averaged the posterior for each position bin, to obtain maps of decoded position as a function of the actual position of the mouse (**Figure 7-5a; medium gain**).

Using the same decoder trained in the medium gain condition, we then assessed the effect of distance run on the position decoded from each area (**Figure 7-5a low, high gain**). We found that at low gain, maps obtained from V1, AL and PPC were all shifted upward, reflecting a decoded position that was ahead of the actual position of the animal. Conversely, decoding maps obtained at high gain were shifted downward, reflecting a position behind the animal. The shift in the decoded position, ahead or behind the animal became more evident when we measured the decoding error in each gain condition as a function of position (**Figure 7-5b**); decoding errors for low or high gain were shifted in opposite directions relative to medium gain, thus indicating that the distance run influenced the representation of the environment encoded by V1, AL or PPC neurons.

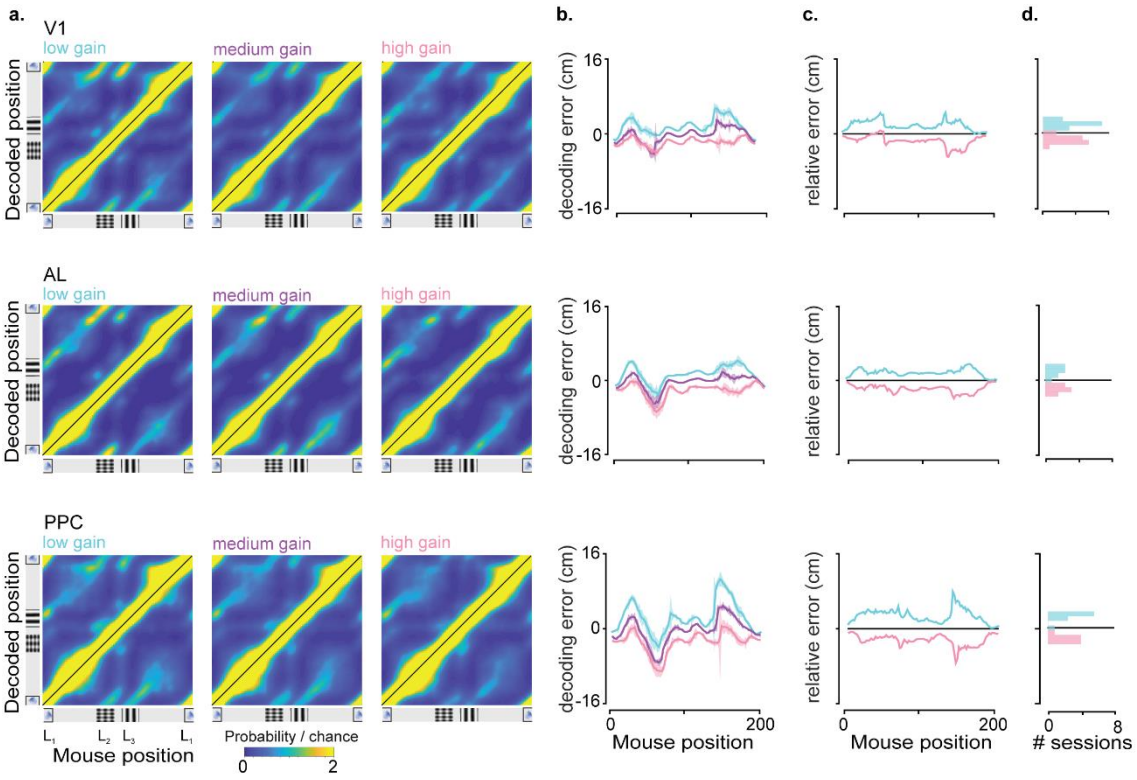


**Figure 7-5: Gain changes induce shifts in decoding position.**
**a.** Position decoded from the population activity inV1 (*left*), AL (*middle*) or PPC (*bottom*) at low (*left*), medium (*middle*) and high (*right*) gain. The decoder was trained in the medium gain condition (*black* line: diagonal).
**b.** Decoding error as a function of position at low (*cyan*), medium (*magenta*) and high (*pink*) gain. Decoding errors for low or high gain are shifted in opposite directions relative to medium gain thus indicating that the distance run influenced the representation of the environment encoded by V1, AL or PPC neurons. Shaded region is s.e.m.
**c.** Relative error at high or low gain relative to the medium gain in V1, AL and PPC (same order and colour code as in **a**). PPC showed larger biases in decoded position compared to V1 or AL when the distance run was changed.
**d.** Distributions of single-session averaged error at high or low gain relative to the medium gain in V1, AL and PPC (same order and colour code as in **a**).

## 7.1.4 Influence by distance run is stronger in PPC

Is the influence by distance run more pronounced in some areas compared to others? When we estimated the error at high or low gain relative to the medium gain (**Figure 7-5c**), we found that PPC showed larger biases in decoded position compared to V1 or AL when the distance run was changed (two-sample t-test: PPC vs V1: $p = 10^{-6}$ at low gain, $p = 10^{-4}$ at high gain; AL vs V1: $p = 0.46$ at low gain, $p = 0.04$ at high gain). Similarly, when we compared the decoding errors averaged across the entire track, the position decoded from PPC was further behind the actual position than V1 at high gain (**Figure 7-5d**; *pink*; V1: -0.73 ± 0.09, AL: -0.91 ± 0.11, PPC: -1.25 ± 0.12; two-sample t-test: AL vs V1: $p_{high\ gain} = 0.22$; PPC vs V1: $p_{high\ gain} = 0.001$); conversely, at low gain, position decoded from PPC was

further ahead than V1 (**Figure 7-5d**; *cyan*; Mean ± s.e.m.: V1: 0.91 ± 0.09, AL: 0.97 ± 0.15, PPC: 1.2 ± 0.13 two-sample t-test: AL vs V1: $p_{low\ gain} = 0.68$; PPC vs V1: $p_{low\ gain} = 0.06$).

Can the higher decoding error observed in PPC be explained by behavioural factors, such as differences in the running speed profile? To answer this question, we compared the speed range across sessions between areas and found that these were very similar (two-sample *t*-test for speeds: AL vs V1: $p_{low\ gain} = 0.74$, $p_{high\ gain} = 0.76$; PPC vs V1: $p_{low\ gain} = 0.63$, $p_{high\ gain} = 0.65$). We also sought to assess the effect of running speed on the magnitude of the decoding error. For each session we expressed the relative decoding error as a function of the mean speed (**Figure 7-6a**).

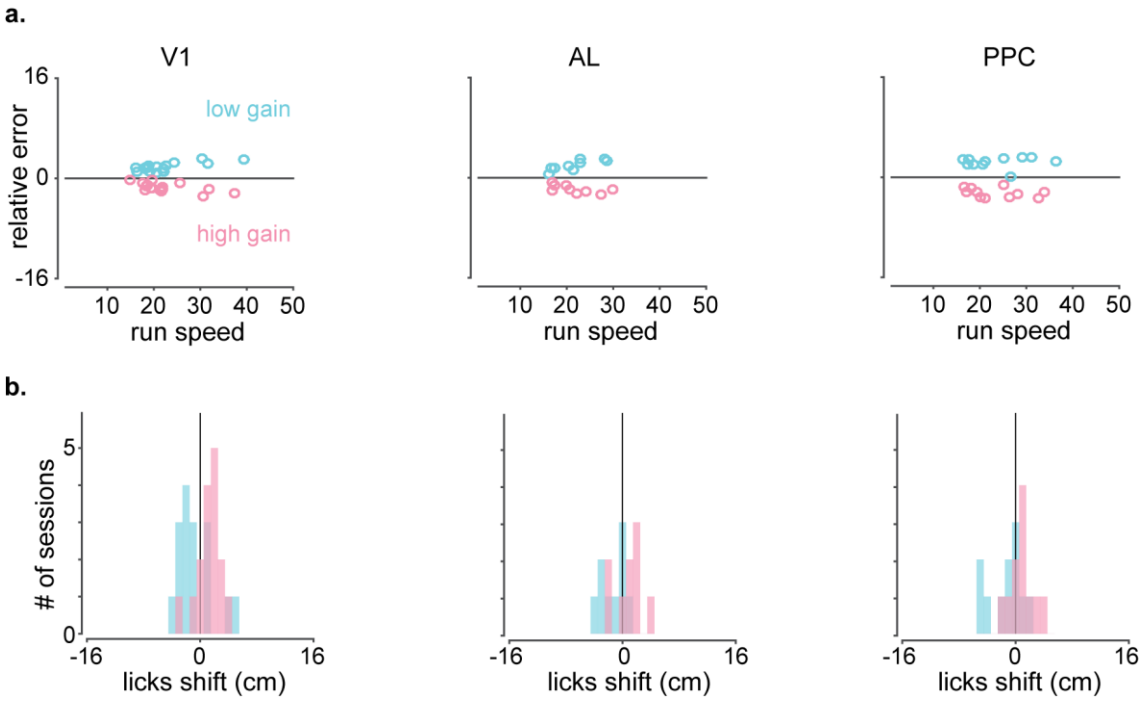


**Figure 7-6: the higher decoding error in PPC cannot be explained by differences in running speed or the animal's licking behaviour.**
**a.** Relative decoding error as a function of the mean speed in V1 (*left*), AL (*middle*) and PPC (*right*). Unlike V1 and AL, the relative decoding error did not correlate with running speed in PPC ($r_{low\ gain} = 0.10$, $p_{low\ gain} = 0.77$, $r_{high\ gain} = -0.36$, $p_{high\ gain} = 0.27$).
**b.** Distributions of of shifts in the licking t low and high gain in V1, AL and PPC (same order as in **a**). These distributions indicate that the position where the animal took the decision to lick shifted according to the gain change (earlier at low gain and later at high gain). The distributions were not different between areas (two-sample *t*-test for shifts in licks: AL vs V1: $p_{low\ gain} = 0.58$, $p_{high\ gain} = 0.66$; PPC vs V1: $p_{low\ gain} = 0.53$, $p_{high\ gain} = 0.76$).

Notably, we found that in V1 and AL the relative decoding error was correlated with running speed especially in the low gain condition (Pearson correlation coefficient (r), p-value (p): V1: $r_{low\ gain} = 0.72$, $p_{low\ gain} = 0.001$, $r_{high\ gain} = -0.60$, $p_{high\ gain} = 0.014$; AL: $r_{low\ gain} =$

0.82, $p_{low\ gain}$ = 0.007, $r_{high\ gain}$ = -0.60, $p_{high\ gain}$ = 0.09). Instead, we found no correlation between the relative decoding error and running speed in the PPC (PPC: $r_{low\ gain}$ = 0.10, $p_{low\ gain}$ = 0.77, $r_{high\ gain}$ = -0.36, $p_{high\ gain}$ = 0.27).

Differences in the decoding error between areas may also be due to differences in the licking behaviour of the animal. Indeed, consistent with neurons in visual cortex, the animal's decision to lick was also shifted when the distance run was changed; the position where the animal took the decision to lick shifted earlier along the track in the low gain condition relative to medium gain, and later along the track in the high gain condition (**Figure 7-6b**). However, we found that these distributions were not different between areas (two-sample *t*-test for shifts in licks: AL vs V1: $p_{low\ gain}$ = 0.58, $p_{high\ gain}$ = 0.66; PPC vs V1: $p_{low\ gain}$ = 0.53, $p_{high\ gain}$ = 0.76). Thus, the larger effects observed in PPC could not be explained by inter-session variability in the animal's strategy to perform the task.

Taken together, these results indicate that responses in V1, AL and PPC are influenced by distance run and that this influence is stronger in PPC regardless of the animal's running or licking behaviour.

## 7.2 Discussion

Throughout this Thesis I have provided several lines of evidence indicating that responses in visual cortex are strongly influenced by navigation-related signals. To better understand the exact content of these signals though, it would be useful to disentangle the effects of actual position from distance run and other covariates of position such as time (see 8.2). Although understanding the influence of all these correlates one by one requires a wide spectrum of experiments that goes beyond the scope of this study, here we attempted to assess the influence of at least one factor: distance run. Indeed, early studies in the hippocampus have shown that distance from the origin together with visual cues shape place fields in a competitive manner (Gothard, Skaggs, and Mcnaughton, 1996). More recently, it was shown that although hippocampal bidirectional cells encoded absolute position in real environments, they switched to a distance code in virtual reality (Ravassard et al., 2013). Finally, another study explicitly dissociated the effect of visual information from self-motion information by reducing the gain of ball-to-virtual movement in half and showing that place field locations shifted accordingly (Chen, King, Burgess, and O'Keefe, 2013). Thus, taken together these studies suggest that coding for distance is a critical factor

that acts independently and competitively with external sensory cues to shape firing fields in the hippocampus, and as such it should be dissociated from signals coding for absolute position.

By implementing a similar approach as Chen et al (2013), I sought to dissociate the influence of distance travelled from visually-guided position signals on responses in V1, AL and PPC during virtual navigation. Specifically, I showed that changing the distance run was associated with a systematic shift in response profiles and a concomitant displacement of the decoded position either ahead of the animal's actual position for longer distances (gain decrease) or behind for shorter distances (gain increase). Therefore, these findings suggest that responses in visual cortex are influenced by distance run, consistent with previous findings in V1 by Fournier et al. (in preparation).

The average shift induced by gain changes was smaller than what would be expected from an influence by distance run only, suggesting that at least in this specific task, in which mice were trained to use visual cues, spatial position based on visual information must be combined with movement-related influences. But how do visual and self-motion signals combine? Some insight into this question can be provided by inspecting the distribution of position decoded by a Bayesian decoder. One possibility is that visual cues and distance run contribute independently. In this case the decoded position should 'jump' between the virtual position determined by the visual landmarks and a shifted position determined by distance run, reflecting variability in firing within each trial. Another possibility is that visual cues and distance run are being weighted together, giving rise to a decoded position that lies between positions determined by visual cues only or distance run only. The distribution of decoded position indicated that position estimates were placed between estimates defined by virtual position alone or distance run alone. But importantly, the combined estimates were not lying along a straight line, but oscillated, suggesting variable, albeit small, differences in the weighting of virtual position or distance run between trials.

Taken together, these results point towards the relative weighting of visual and self-motion cues. Nevertheless, it remains an open question what determines their weighting on a trial-by-trial basis. Answering this question would require knowledge of the point at which the animal starts counting distance run. This knowledge is indispensable for an accurate model of distance run which would take into account the influence of previous moments in time. Because in our behavioural paradigm self-motion and visual cues are

present throughout and the manipulation in gain is small, it has not been possible to determine the starting point from which distance run is estimated. I speculate that in our behavioural paradigm distance run is estimated from a reference point that moves together with the animal. To obtain precise control on that reference point, higher gain manipulations may prove useful. Alternatively, one could use a behavioural paradigm in which visual and self-motion cues are not put into conflict (Chen, King, Burgess, and O'Keefe, 2013). According to this approach, the animal relies on distance run estimated from an initial localisation point previously determined by visual cues.

What is the origin of the observed shift in decoded position during gain changes? Specifically, could this shift be the result of a pure feed-forward response delay? Consider for instance an animal running at 20 cm/s in the medium gain condition. Given the 20% gain change, the animal's speed in the virtual corridor will change by 4cm/s. This means that a 1 cm shift would require a feed-forward delay of 250 ms, way longer than what is expected from visual responses in V1 (<100 ms; (Niell and Stryker, 2008)). Thus, although an influence by feed-forward delay cannot be excluded, I propose that the observed results could only be fully explained by additional delays mediated by feedback signals.

Besides distance travelled, another potential factor which could contribute to the observed shift is running speed. In turn, since running speed is defined as distance travelled per unit time, influences by variable running speed (i.e. same distance travelled at variable time intervals) would indicate that the factor that actually affects neural activity is time, rather than distance. Hence, previous studies have attempted to distinguish distance from time by focusing on trials in which the animal ran at different speeds (see for eg Funamizu, Kuhn, and Doya, 2016; Ravassard et al., 2013). In our case, it has proven difficult to disentangle time from distance mainly due to the stereotyped running speed pattern acquired after animals learned the task. In addition, the stereotyped running speed pattern depended substantially on gain: when approaching the reward zone, animals tended to decelerate earlier in low gain trials and later during high gain trials.

Although our ability to discriminate between distance travelled and time was limited, we sought to control for speed-related effects in two ways: first, decoding was performed separately for different speed ranges; second, we compared the degree of decoding error between sessions during which animals ran at different speeds on average. Notably, we found that the magnitude of the decoding error correlated to some extent with speed in V1 and AL, but additional experiments and further analysis is required to clarify this

dependence. It is possible, for instance, that the observed correlation is not due to effects exerted by time, but rather due to behavioural factors; for example, mice may pay more attention to the visual landmarks when they run slower. On the other hand, we found no dependence between the relative error and running speed in the PPC, implying a potential functional difference between PPC and areas V1 and AL.

The difference between PPC and V1 was further supported by the finding that distance run had a stronger impact on the PPC response profiles compared to V1. Importantly, the difference between areas could not be explained by differences in the licking behaviour of the animal, because the shifts in the licking behaviour induced by gain changes were comparable between V1 and PPC. Therefore, the results presented here suggest that PPC is strongly influenced by distance run, consistent with the proposed role of the PPC in processing and integrating movement-related signals (Wilber et al., 2014) and its stronger connectivity to secondary motor cortex (Oh et al., 2014).

A previous study on the PPC during an auditory, virtual navigation task also suggested that neurons in this area coded for the distance to goal (Funamizu, Kuhn, and Doya, 2016). Instead our gain manipulation experiments, demonstrating a shift earlier along the track in low gain and later along the track in high gain, suggest a code for distance travelled rather than distance to goal. Nevertheless, we did not systematically search for influences by distance to goal and it is a possibility that a subpopulation in PPC may also be modulated by this type of distance. Distance-to-goal signals relate to anticipatory activity which has been previously reported to arise in PPC in the form of the self-motion related signals occurring up to 500 ms before execution of movement (Whitlock et al., 2012). Anticipatory signals reflecting the upcoming reward or punishment have also been reported previously in V1 (Makino and Komiyama, 2015; Shuler and Bear, 2006). It would thus be interesting to assess whether such signals are also present in the response patterns presented here.

Overall, here I demonstrated that distance run, another non-sensory signal known to influence the brain's navigational system, also influences visual responses in V1, AL and PPC. Nevertheless, distance run alone could not explain the relatively small shift in the response profiles. Consequently, decoded position was intermediate between predictions based on visual cues and distance run. Therefore, these findings provide new evidence on the presence of self-motion signals related to the distance travelled in visual areas, thereby reinforcing the conclusion of previous chapters that in visual cortex, responses to visual cues are strongly modulated by navigation-related signals.

# General conclusions

## 8.1 Main findings and limitations

A major focus of this thesis has been the influence of navigational signals on responses across the visual cortex. To this end, I used two-photon imaging to record responses of thousands of cells in V1 and 6 higher visual areas from transgenic mice placed in two virtual reality environments, a linear corridor (Chapters 3,5,6) or a circular maze (Chapter 7). Thorough analysis of the acquired data using several metrics and modelling revealed that correlates of space, known to arise in the brain's navigational system, such as spatial position and distance travelled, are also present in visual cortex; and in fact, these signals occur as early as in V1 (Chapter 3; Saleem*, Diamanti* et al., 2018). Notably, responses recorded in several visual areas during a virtual navigation task encoded the position of the mouse in the environment, with little ambiguity between the visually- matching landmarks, consistent with Saleem*, Diamanti* et al. (2018) (Chapter 7). Moreover, by combining these findings with data acquired under 'passive viewing' conditions, I highlighted the impact of active behaviour on neural activity based on several lines of evidence: 1. Spatial modulation was substantially reduced during play-back of the virtual environment compared to active behaviour (**6.1.3**) 2. during play-back of the virtual environment, responses were weaker and less reliable than during active behaviour. This was the case, even though the animal was running, but its running was not controlling the virtual reality (**6.1.1**) 3. Reliable responses to the visual landmarks in virtual reality did not correlate with reliable responses to drifting gratings, even for cells preferring the vertical drifting gratings, a stimulus that was also a landmark in virtual reality (**4.2.3**). Given these findings I propose that the brain's sensory system is endowed with the ability to enhance or even modulate its activity depending on behavioural demands; this ability is unravelled here specifically in the

context of navigation through evidence demonstrating that visual responses are explicitly influenced by spatial position during active behaviour.

The main limitation in characterising visual responses in virtual reality has been the design of the virtual linear corridor. Specifically, to create a sense of place inside a room, the virtual scenes had to be different between the central and the peripheral visual field. This was first due to the end grey wall which was expanding as the animal advanced forward and second, due to the 3d perspective of the visual environment which imposed that visual stimuli presented in the centre had much higher spatial frequencies compared to the periphery. Thus, it has proven difficult to discern the visual responses of neurons with receptive fields in the central visual field. Correspondingly, the difference in visual scenes between the centre and the periphery has posed a challenge in assessing the role of higher visual areas biased towards the centre (LM, AL, RL) during behaviour in virtual reality, and therefore understanding the underlying reasons for the increased degree of spatial modulation in the centre compared to the periphery (**5.1.2**).

Despite these limitations, this study has elucidated several differences between some of the higher visual areas investigated here, specifically LM, A, AM and PM.

LM, for instance, is an area thought to be the homologue of primate V2 (Wang, Gao, and Burkhalter, 2011) and the only higher visual area receiving inputs from the dLGN, like V1 (Oh et al., 2014). Consistent with these reports, in this study LM appears to be the most similar to V1; visual responses in LM are almost equally reliable during presentation of drifting gratings and in virtual reality, and the corresponding distribution is very similar to V1 (**4.2.3**); the robustness of the response profile pattern in closed-loop, assessed based on the height of the main peak in the population average, is highest in LM together with V1 (**5.1.1**); finally, LM's tuning for speeds in virtual reality is biased towards virtual speed, again similar to V1 (**0**). I also found that the degree of spatial modulation in LM was lowest compared to V1, the only area possessing this feature (although this difference cannot be fully understood for the reasons described in discussion section **5.2**). Taken together, these findings support the proposed contribution of LM to a putative ventral, rather than a dorsal, pathway. In addition, LM's functional similarities to V1 demonstrated here, point to its positioning towards the lower level of a putative hierarchical processing stream, consistent with (Wang and Burkhalter, 2007).

A and AM, on the other hand, appear to be positioned on the other side of the spectrum, as the areas the most different to V1. A and AM are thought to overlap, at least

in part, with PPC (Glickfeld and Olsen, 2017; Wang, Gao, and Burkhalter, 2011), an area associated with higher-order functions particularly during navigation, such as the encoding of spatial signals in the abstract space defined by routes (Nitz, 2009, 2006), the multiplexing of allocentric and egocentric signals (Wilber et al., 2014) decision making (Harvey, Coen, and Tank, 2012). In this study multiple lines of evidence have highlighted the functional differences of A, and sometimes AM, from the other visual areas: the reliability of visual responses in A and AM is strongly biased towards responses in virtual reality, with the reliability of responses to drifting gratings being lowest in A (**4.2.3**); when comparing between cells that 'look' in the same part of the visual field, it appears that A is more strongly modulated by spatial position compared to V1(**5.1.2**); A an AM are the only areas in open-loop with higher degree of spatial modulation than V1; these are also the only areas tuned to running speed, rather than virtual speed, with A exhibiting the strongest bias (**0**); finally, these areas are more strongly influenced by distance travelled (**7.1.4**). Overall these findings provide additional evidence supporting a prominent role of A and AM during navigation and thus, their contribution to a putative dorsal pathway. It is important to note though, that any conclusions regarding spatial representations in PPC, should also take into account the multifaceted and complex nature of the modulatory signals present in this area. To explain the complexity of these signals, work by Andersen and colleagues postulated that in PPC different frames of reference coexist and that the interactions between these frames can be well described by 'gain fields' (see relevant literature in General Introduction). Such frames of reference can be egocentric, or allocentric (Snyder, Grieve, Brotchie, and Andersen, 1998), or even more abstract. For instance, a reference frame can be 'centred' to the route taken (Nitz, 2006): indeed, if the starting and end point of a complex route within the same environment is modified, the spatial firing patterns in the rat PPC shift accordingly, so that they remain invariant with respect to the space defined by the route. Another example of abstract reference frames are the object-centred frames found in the monkey PPC area 7a (Chafee, Averbeck, and Crowe, 2007). Such frames are similar to the route-centred frames found in rats in the sense that they are not defined in the physical space of body or world (Nitz, 2009). Thus, one should be cautious when drawing conclusions on the type of modulation and the reliability of responses particularly in the PPC. The reason being that any unreliability observed in PPC responses may indicate that a different, more abstract frame of reference is more relevant to a specific behaviour.

PM, finally, appears to be positioned in the middle of a putative hierarchical spectrum. In addition, given several anatomical ((Wang, Sporns, and Burkhalter, 2012)) and functional (Funamizu, Kuhn, and Doya, 2016; Roth, Helmchen, and Kampa, 2012) evidence as well as PM's strategic positioning next to retrosplenial cortex (Wang, Sporns, and Burkhalter, 2012), it has been proposed that PM plays an important role during navigation. Here we find several pieces of evidence supporting these views: PM's degree of spatial modulation is slightly but significantly more pronounced than V1, when comparing between cells with receptive fields in the periphery (**5.1.2**); on the other hand, PM has the most robust, purely visual responses during open-loop (**0**); finally, although a subpopulation of cells in PM are tuned for running speed only, on average, PM seems to be influenced by both virtual and running speed equally. Overall, these results support the contribution of PM to a putative dorsal stream, in a manner though different than PPC. Unlike A and AM which seem to vigorously encode self-motion signals, PM seems to be instead also influenced by optic flow. This difference between A/AM and PM is also consistent with the speculation that the putative dorsal stream is split into multiple substreams, with PM belonging to a different substream than A and AM (Murakami, Matsui, and Ohki, 2017).

Overall, here I have highlighted the several differences between higher visual areas which arose during a variety of behavioural conditions. Hopefully these results provide a stepping stone towards the long-sought aim of characterising the role of the mouse higher visual areas during behaviour.

## 8.2 Is modulation of visual responses truly spatial?

To establish that responses in visual cortex were influenced by spatial context, we performed several controls related to either visual, behavioural or task-related factors (see Chapter 3). These involved either the exclusion of each factor individually (see **3.1.2**) or the assessment of all these factors simultaneously (see **3.1.3**). Such controls were crucial in pointing towards the modulation of visual responses by navigation-related signals. Nevertheless, it may also be argued that the nature of this modulation is not spatial per se. For instance, responses to visual stimuli may well be influenced by the familiarity or recency of the stimulus presentation (for a review see Brown and Banks, 2015). Another possibility is that neurons in visual cortex may be modulated by the temporal order of events, or more broadly by time rather than space (for a review see Eichenbaum, 2017). Finally, the observed modulation may originate from more-abstract spatial signals defined by the route taken (Nitz, 2009, 2006; as discussed above).

Here I discuss the potential influence of familiarity or time by referring to the relevant literature. I argue that factors such as familiarity or recency, are not sufficient to explain the observed modulation. Instead the influence by other variables, such as time, cannot be excluded based on the experiments presented here. Future directions in distinguishing potential effects by time are also discussed below.

### 8.2.1 Familiarity, recency and novelty

Familiarity with a stimulus, together with recollection of that stimulus within a previously-experienced context, are the two key processes underlying recognition memory (Eichenbaum, Yonelinas, and Ranganath, 2007). According to the 'dual-process' theories of recognition memory, familiarity and recollection are two distinct processes represented in different subregions of the medial temporal lobes (Eichenbaum, Otto, and Cohen, 1994): hippocampus and the parahippocampal cortex are specifically involved in recollection of spatial, and more generally contextual, information associated to a stimulus; on the other hand, perirhinal cortex, but not hippocampus, encodes familiarity for individual stimuli (Brown and Aggleton, 2001). Therefore, unlike recollection, which involves associations between stimuli and a specific context, familiarity is stimulus-specific, manifested as a reduction in responsiveness to that stimulus (Brown and Banks, 2015).

Reduced responsiveness to a familiar visual stimulus has been previously reported in both rats (Wan, Aggleton, and Brown, 1999) and monkeys (Fahy, Riches, and Brown, 1993; Xiang and Brown, 1998). During a paired viewing procedure, rats were presented with a novel and a familiar stimulus, with each stimulus being visible by one eye only (Wan, Aggleton, and Brown, 1999). Using activation of the immediate early gene *c-fos* as a marker of neuronal activity, it was shown that activity in perirhinal cortex and area TE of the temporal lobe was significantly higher during presentation of novel compared to familiar stimuli. Instead no similar differences were observed in hippocampus. These results were consistent with previous findings in monkeys during a visual recognition task (Xiang and Brown, 1998). In addition, the same study in monkeys dissociated familiarity from recency and novelty: some cells exhibited decreased responses to familiar stimuli, whereas others had decreased responses to any, familiar or novel, stimulus if it had been presented recently. Finally, some cells signalled novelty, by responding more strongly to the first presentation of a novel stimulus compared to the second presentation of both familiar and novel stimuli. Notably, visual responses influenced by familiarity, recency or novelty were found in area TE, perirhinal and entorhinal cortex, but not in hippocampus.

It is important to note that all 'familiarity, recency and novelty' neurons reported by Xiang & Brown (1998) were in fact visually responsive neurons. Similarly, it is conceivable that the visually responsive neurons recorded here may be influenced by familiarity rather than spatial context. Nevertheless, a key feature of familiarity effects is the reduced response to the second presentation relative to the first. Indeed, Xiang and Brown (1998) reported that 98% of neurons responding differently to familiar or recent stimuli, exhibited weaker responses. Reduced responsiveness has also been confirmed by all studies with similar focus (Brown and Banks, 2015). In this study instead, we found that the percentage of cells preferring to fire in the first half of the corridor was almost equal to the percentage of cells preferring to fire in the second half (49% in first half vs 51% in second half). This means that almost half of the cells exhibited stronger responses to the second presentation of the same stimulus, which is the opposite of what is expected from effects of familiarity or recency. Therefore, although an effect of familiarity or recency cannot be excluded, it can only partially explain the influence observed here.

### 8.2.2 Time

Navigation and episodic memory are similar experiences in the sense that they involve the sequential organisation of events, places or memories respectively (Buzsáki and Llinás, 2017). Accordingly, areas in the brain's navigational system, such as the hippocampus, are thought to code not only for the spatial but also for the temporal order of sequential events (Eichenbaum, 2017). Initial evidence on the existence of a temporal code in hippocampus came from studies that controlled location and movement precisely, during either a memory (Manns, Howard, and Eichenbaum, 2007) or a spatial (Pastalkova, Itskov, Amarasingham, and Buzsáki, 2008) task. In correct trials during an odour sequence task, the neural representations of odours presented closer in time were more similar than the representations of odours further apart and importantly this difference was present irrespective of location. Thus, this early study suggested that together with spatial context, hippocampus also represents the temporal order of events (Manns, Howard, and Eichenbaum, 2007). The existence of temporal representations in hippocampus was established shortly after with the discovery of 'time cells' (Pastalkova, Itskov, Amarasingham, and Buzsáki, 2008). In this study rats alternated between the left and right arm of a figure-eight shaped maze. Critically, individual runs were separated by a delay period, during which rats had to ran on a wheel at a constant speed, and thus environmental and movement-related inputs were kept constant. The main finding was that during the delay period, hippocampal neurons were sequentially active at different time

points, such that the population activity covered the whole period during wheel running. Since then, various studies have observed temporal sequences of activity in different behavioural contexts, such as non-spatial tasks (MacDonald, Lepage, Eden, and Eichenbaum, 2011), matching-to-sample tasks paired with head-fixation (MacDonald, Carrow, Place, and Eichenbaum, 2013); wheel running at various speeds (Kraus et al., 2013), classical conditioning (Modi, Dhawale, and Bhalla, 2014); or comparison between tasks with or without working memory demands (Salz et al., 2016).

Given the considerable body evidence that the brain's navigational system codes not only for space but also for elapsed time, how likely is it that in the current study the signals influencing visual responses are temporal rather than navigational? This is a possibility particularly because when running through a linear maze subsequent points in space correspond to subsequent points in time. In addition, given the repetition of visual cues, any modulatory signals may reflect the temporal order of the presented stimuli rather than spatial location per se. Nevertheless, drawing firm conclusions on the contribution of time in the modulation of visual responses observed in this study would require a delay period while keeping environmental cues (visual cues in our case) and self-motion cues (running speed) constant. The grey screen inter-trial intervals could potentially serve as delay periods, but the running speed was found to be highly variable within these periods, thus posing a major limitation towards investigating the potential influence of time. Future work addressing this issue should aim for precise isolation of spatial and movement-related variables.

Even if time is found to play a crucial role in our behavioural paradigm, I speculate that most neurons would still be strongly influenced by spatial context (spatial location or distance run). This speculation is based on studies in hippocampus which aimed at distinguishing the influence of time from spatial location or distance travelled. Of particular interest in this context is the study by Kraus et al. (2013): during the delay period between runs in an eight-figure shaped maze, rats ran on a treadmill while the speed of the treadmill was varied. Thus, manipulation of the treadmill speed allowed discrimination between time and distance. In addition, implementation of generalised linear models enabled quantification of the influence by distance, time but also spatial location. The main finding was that during the delay period most neurons were influenced by both time and distance, with some neurons being tuned to distance only or time only. Importantly, most of these neurons, as well as others, exhibited place fields in the maze, indicating that coding for time does not rule out coding for spatial context. Consistent with these results were the findings

by Salz et al (2016), showing that cells coding for time would sometimes also code for space during a working memory task. Finally, also the earlier study by Pastalkova et al. (2008) reported that, although the sequential patterns of activity when running on a treadmill or through the maze were not related, many hippocampal cells were active during both periods. Based on these studies and others, it has been proposed that populations in hippocampus can code for space along with time and other contextual variables, as long as these are relevant to the task at hand (Eichenbaum, 2014). In fact, in cases where a temporal or spatial code are equally relevant, populations of the same hippocampal neurons can switch between codes depending on behavioural strategy (Cabral et al., 2013). Instead, if locations on the maze are crucial for task performance, spatial context will have a strong influence on the hippocampal populations activity (Kraus et al., 2013). Similarly, in the task presented in Chapter 7, spatial context (reward location and distance run) are very relevant, and therefore are expected to exert a strong influence on the modulation of visual responses.

Overall, given that the brain's navigational system can be considered as a 'general purpose sequence generator' (Buzsáki and Llinás, 2017), with the ability to represent any context relevant to an ongoing experience, one hypothesis is particularly intriguing: if the modulation of visual responses does reflect the interaction between the navigational and visual systems, then the nature of this modulation should be relative, depending on the context represented by the brain's memory systems. Further insight into this hypothesis can be provided by introducing tasks requiring estimates based on place or the memory of ordered events, in parallel or separately.

## 8.3 Future directions

Visual and navigational information are centred in different reference frames, the eye and the world respectively, and therefore our initial hypothesis was that during visually-guided navigation visual information has to be at some point transformed into spatial information (see **1.5**). An attractive theoretical framework supporting this hypothesis was the concept of 'gain fields' and a prominent candidate for participating in transformations across reference frames the PPC (Xing and Andersen, 2000). Nevertheless, here we found that visual responses as early as in V1 are modulated by spatial position, indicating that, at least in the mouse, transformations across reference frames can arise in early sensory cortex. 'Gain fields' still remain an attractive potential mechanism underlying our findings, yet it remains an open question whether vision and spatial signals actually interact in a

multiplicative manner. One way to test if this interaction is multiplicative is by varying the contrast of the visual landmarks. Then a testable prediction is that at 0% contrast any influence by position should disappear. In addition, varying the contrast should have no effect on the ratio of responses, as these should scale with contrast by the same amount.

Another question being raised by our findings is how spatial modulation arises in sensory cortex? Preliminary results indicate that V1 responses on the first day of exposure to virtual reality are more visual compared to subsequent days (**3.1.5**). It would be interesting to assess whether the distribution of ratios acquired on day 1 could be fully explained by influences from behavioural factors using the models described in Section **3.1.3**. Conversely, if spatial modulation is already present on day 1, it is conceivable that this influence has developed within shorter timescales than what we are able to analyse (indeed, to obtain response profiles as a function of position we need to consider at least 20 trials). The possibility that spatial modulation is present already on day 1 is also likely based on findings in hippocampus showing that spatially localised firing fields appear already on the first day of training in virtual reality (Chen, King, Burgess, and O'Keefe, 2013). Therefore, perhaps the best way to go about assessing the emergence of spatial signals in sensory cortex is by lesioning the brain's navigational system prior to exposure to a new environment.

Another question concerns the existence of anticipatory signals. Are these signals more prominent in some areas than others? For instance, it has been shown that PPC contains signals predicting movement up to ~1 sec prior to movement execution (Whitlock et al., 2012). But also, V1 contains anticipatory signals of the upcoming reward (Poort et al., 2015) or aversive event (Makino and Komiyama, 2015). Testing for the existence of anticipatory in our dataset is possible with the model I used to assess the influence of various task-related and behavioural factors (**3.1.3**). These tests will require closer inspection of the weights assigned to the different predictors, particularly for those shifted backward in time. This analysis will provide a better idea on how spatial signals about the actual position intermingle with anticipatory signals of movement or reward.

Finally, having shown that spatial signals are already present in V1, then the next logical step is to assess whether these signals are already present earlier along the visual pathway, such as in the dLGN. Ideally, I would like to be able to measure responses in dLGN simultaneously with V1 and possibly hippocampus. Such large-scale recordings are now made possible with the use of Neuropixel probes (Jun et al., 2017). My next aim is to use

these probes in order to record neural activity across multiple cortical and subcortical areas during virtual navigation.

Clearly more work is needed to understand how these spatial signals are being acquired in visual cortex. To this end, experimental studies targeting the functional connectivity between areas from visual cortex all the way to the hippocampus during navigation would be particularly useful; these studies should go hand-by-hand with computational studies assessing the exact nature of transformations that information undergoes along these pathways. In this Thesis, I presented several lines of evidence on the functional differences between visual areas in the mouse. In parallel, I demonstrated that the mouse cortex does not keep a firm distinction between navigational and sensory systems: rather spatial signals permeate cortical processing. Ideally, these results will pave the way towards a better understanding of the fundamental principles underlying sensory processing during navigation.

# References


Andermann, M.L., Kerlin, A.M., Roumis, D.K., Glickfeld, L.L., and Reid, R.C. (2011). Functional specialization of mouse higher visual cortical areas. Neuron *72*, 1025–1039.

Andersen, R.A., and Mountcastle, V.B. (1983). The influence of the angle of gaze upon the excitability of the light-sensitive neurons of the posterior parietal cortex. J. Neurosci. *3*, 532–548.

Andersen, R.A., Essick, G.K., and Siegel, R.M. (1985). Encoding of Spatial Location by Posterior Parietal Neurons. Science (80-. ). *230*.

Aronov, D., and Tank, D.W. (2014). Engagement of Neural Circuits Underlying 2D Spatial Navigation in a Rodent Virtual Reality System. Neuron *84*, 442–456.

Ayaz, A., Saleem, A.B., Schölvinck, M.L., and Carandini, M. (2013). Locomotion controls spatial integration in mouse visual cortex. Curr. Biol. *23*, 890–894.

Bennett, C., Arroyo, S., and Hestrin, S. (2013). Subthreshold mechanisms underlying state-dependent modulation of visual responses. Neuron *80*, 350–357.

Van den Bergh, G., Zhang, B., Arckens, L., and Chino, Y.M. (2010). Receptive-field properties of V1 and V2 neurons in mice and macaque monkeys. J. Comp. Neurol. *518*, 2051–2070.

Brown, M.W., and Aggleton, J.P. (2001). Recognition memory: what are the roles of the perirhinal cortex and hippocampus? Nat. Rev. Neurosci. *2*, 51–61.

Brown, M.W., and Banks, P.J. (2015). Neuroscience and Biobehavioral Reviews In search of a recognition memory engram. Neurosci. Biobehav. Rev. *50*, 12–28.

Burgess, N., Maguire, E. a, and O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory 1. Neuron *35*, 625–641.

Cabral, H.O., Vinck, M., Fouquet, C., Pennartz, C.M.A., and Rondi-reig, L. (2013). Oscillatory Dynamics and Place Field Maps Reflect Hippocampal Ensemble Processing of Sequence and Place Memory under NMDA Receptor Control. Neuron *81*, 402–415.

Carandini, M., Shimaoka, D., Rossi, L.F., Sato, T.K., Benucci, A., and Knopfel, T. (2015). Imaging the Awake Visual Cortex with a Genetically Encoded Voltage Indicator. J. Neurosci. *35*, 53–63.

Chafee, M. V., Averbeck, B.B., and Crowe, D.A. (2007). Representing Spatial Relationships in Posterior Parietal Cortex: Single Neurons Code Object-Referenced Position. Cereb. Cortex *17*, 2914–2932.

Chen, G., King, J.A., Burgess, N., and O'Keefe, J. (2013). How vision and movement combine in the hippocampal place code. Proc Natl Acad Sci U S A *110*, 378–383.

Chen, G., King, J.A., Lu, Y., Cacucci, F., and Burgess, N. (2018). Spatial cell firing

during virtual navigation of open arenas by head-restrained mice. bioRxiv 246744.

Cohen, Y.E., and Andersen, R.A. (2002). A common reference frame for movement plans in the posterior parietal cortex. Nat. Rev. Neurosci. *3*.

Cushman, J.D., Aharoni, D.B., Willers, B., Ravassard, P., Kees, A., Vuong, C., Popeney, B., Arisaka, K., and Mehta, M.R. (2013). Multisensory control of multimodal behavior: Do the legs know what the tongue is doing? PLoS One *8*.

D'Souza, R.D., Meier, A.M., Bista, P., Wang, Q., and Burkhalter, A. (2016). Recruitment of inhibition and excitation across mouse visual cortex depends on the hierarchy of interconnecting areas. Elife *5*, 1–19.

Dadarlat, M.C., and Stryker, M.P. (2017). Locomotion Enhances Neural Encoding of Visual Stimuli in Mouse V1. J. Neurosci. *37*, 3764–3775.

Dipoppa, M., Ranson, A., Krumin, M., Pachitariu, M., Carandini, M., and Harris, K.D. (2018). Vision and Locomotion Shape the Interactions between Neuron Types in Mouse Visual Cortex. Neuron *98*, 602–615.e8.

Dombeck, D.A., Khabbaz, A.N., Collman, F., Adelman, T.L., and Tank, D.W. (2007). Imaging Large-Scale Neural Activity with Cellular Resolution in Awake, Mobile Mice. Neuron *56*, 43–57.

Driscoll, L.N., Pettit, N.L., Minderer, M., Chettih, S.N., and Harvey, C.D. (2017). Dynamic Reorganization of Neuronal Activity Patterns in Parietal Cortex. Cell *170*, 986–999.e16.

Eichenbaum, H. (2014). Time cells in the hippocampus: A new dimension for mapping memories. Nat. Rev. Neurosci. *15*, 732–744.

Eichenbaum, H. (2017). Time (and space) in the hippocampus. Curr. Opin. Behav. Sci. *17*, 65–70.

Eichenbaum, H., Otto, T., and Cohen, N.J. (1994). Two functional components of the hippocampal memory system. Behav. Brain Sci. *17*, 449–518.

Eichenbaum, H., Yonelinas, A.P., and Ranganath, C. (2007). The Medial Temporal Lobe and Recognition Memory. Annu. Rev. Neurosci. *30*, 123–152.

Ekstrom, a D., Kahana, M.J., Caplan, J.B., Fields, T. a, Isham, E. a, Newman, E.L., and Fried, I. (2003). Cellular networks underlying human spatial navigation. Nature *425*, 184–188.

Erisken, S., Vaiceliunaite, A., Jurjut, O., Fiorini, M., Katzner, S., and Busse, L. (2014). Effects of locomotion extend throughout the mouse early visual system. Curr. Biol. *24*, 2899–2907.

Erlich, J.C., Brunton, B.W., Duan, C.A., Hanks, T.D., and Brody, C.D. (2015). Distinct effects of prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat. Elife *4*, 1–28.

Fahy, F.L., Riches, I.P., and Brown, M.W. (1993). Neuronal activity related to visual recognition memory: long-term memory and the encoding of recency and familiarity information in the primate anterior and medial inferior temporal and rhinal cortex. Exp. Brain Res. *96*, 457–472.

Fournier*, J., Saleem*, A.B., Diamanti*, E.M., Wells, M., Harris, K.D., and Carandini,

M. Influence of distance signals on primary visual cortex and hippocampus during navigation. (in preparation).

Funamizu, A., Kuhn, B., and Doya, K. (2016). Neural substrate of dynamic Bayesian inference in the cerebral cortex. Nat. Neurosci. *19*, 1682–1689.

Galletti, C., and Fattori, P. (2018). The dorsal visual stream revisited: Stable circuits or dynamic pathways? Cortex *98*, 203–217.

Garrett, M.E., Nauhaus, I., Marshel, J.H., and Callaway, E.M. (2014). Topography and Areal Organization of Mouse Visual Cortex. J. Neurosci. *34*, 12587–12600.

Gilbert, C.D., and Li, W. (2012). Adult Visual Cortical Plasticity. Neuron *75*, 250–264.

Gilbert, C.D., and Li, W. (2013). Top-down influences on visual processing. Nat. Rev. Neurosci. *14*, 350–363.

Glickfeld, L.L., and Olsen, S.R. (2017). Higher-Order Areas of the Mouse Visual Cortex. Annu. Rev. Vis. Sci. *3*, annurev-vision-102016-061331.

Glickfeld, L.L., Andermann, M.L., Bonin, V., and Reid, R.C. (2013). Cortico-cortical projections in mouse visual cortex are functionally target specific. Nat. Neurosci. *16*, 219–226.

Goard, M.J., Pho, G.N., Woodson, J., and Sur, M. (2016). Distinct roles of visual, parietal, and frontal motor cortices in memory-guided sensorimotor decisions. Elife *5*, 1–30.

Goodale, M.A., and Milner, A.D. (1992). Separate visual pathways for perception and action. Trends Neurosci. *15*, 20–25.

Gothard, K.M., Skaggs, W.E., and Mcnaughton, B.L. (1996). Dynamics of Mismatch Correction in the Hippocampal Ensemble Code for Space: Interaction between Path Integration and Environmental Cues. J. Neurosci. *16*, 8027–8040.

György Buzsáki; Rodolfo Llinás (2017). Space and time in the brain. Science (80-. ). *485*, 482–485.

de Haan, E.H.F., Jackson, S.R., and Schenk, T. (2018). Where are we now with "What" and "How"? Cortex *98*, 1–7.

Harvey, C.D., Collman, F., Dombeck, D.A., and Tank, D.W. (2009). Intracellular dynamics of hippocampal place cells during virtual navigation. Nature *461*, 941–946.

Harvey, C.D., Coen, P., and Tank, D.W. (2012). Choice-specific sequences in parietal cortex during a virtual-navigation decision task. Nature *484*, 62–68.

Hillier, D., Fiscella, M., Drinnenberg, A., Trenholm, S., Rompani, S.B., Raics, Z., Katona, G., Juettner, J., Hierlemann, A., Rozsa, B., et al. (2017). Causal evidence for retina-dependent and -independent visual motion computations in mouse cortex. Nat. Neurosci. *20*, 960–968.

Huh, C.Y.L., Peach, J.P., Bennett, C., Vega, R.M., and Hestrin, S. (2018). Feature-Specific Organization of Feedback Pathways in Mouse Visual Cortex. Curr. Biol. *28*, 114–120.e5.

Itokazu, T., Hasegawa, M., Kimura, R., Osaki, H., Albrecht, U.R., Sohya, K., Chakrabarti, S., Itoh, H., Ito, T., Sato, T.K., et al. (2018). Streamlined sensory motor

communication through cortical reciprocal connectivity in a visually guided eye movement task. Nat. Commun. *9*.

Ji, D., and Wilson, M.A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. Nat. Neurosci. *10*, 100–107.

Juavinett, A.L., and Callaway, E.M. (2015). Pattern and Component Motion Responses in Mouse Visual Cortical Areas. Curr. Biol. *25*, 1759–1764.

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. Nat. Protoc. *12*, 32–43.

Jun, J.J., Steinmetz, N.A., Siegle, J.H., Denman, D.J., Bauza, M., Barbarits, B., Lee, A.K., Anastassiou, C.A., Andrei, A., Aydin, Ç., et al. (2017). Fully integrated silicon probes for high-density recording of neural activity. Nature *551*, 232–236.

Jurjut, O., Georgieva, P., Busse, L., and Katzner, S. (2017). Learning Enhances Sensory Processing in Mouse V1 before Improving Behavior. J. Neurosci. *37*, 6460–6474.

Kalatsky, V.A., and Stryker, M.P. (2003). New paradigm for optical imaging: temporally encoded maps of intrinsic signal. Neuron *38*, 529–545.

Keller, G.B., Bonhoeffer, T., and Hübener, M. (2012). Sensorimotor Mismatch Signals in Primary Visual Cortex of the Behaving Mouse. Neuron *74*, 809–815.

Kraus, B.J., Ii, R.J.R., White, J.A., Eichenbaum, H., and Hasselmo, M.E. (2013). Hippocampal "Time Cells": Time versus Path Integration. Neuron *78*, 1090–1101.

Krumin, M., Harris, K.D., and Carandini, M. (2017). Decision and navigation in mouse parietal cortex. bioRxiv 166413.

Lecoq, J., Savall, J., Vučinić, D., Grewe, B.F., Kim, H., Li, J.Z., Kitch, L.J., and Schnitzer, M.J. (2014). Visualizing mammalian brain area interactions by dual-axis two-photon calcium imaging. Nat. Neurosci. *17*, 1825–1829.

Lee, A.M., Hoy, J.L., Bonci, A., Wilbrecht, L., Stryker, M.P., and Niell, C.M. (2014). Identification of a brainstem circuit regulating visual cortical state in parallel with locomotion. Neuron *83*, 455–466.

Licata, A.M., Kaufman, M.T., Raposo, D., Ryan, M.B., Sheppard, J.P., and Churchland, A.K. (2017). Posterior Parietal Cortex Guides Visual Decisions in Rats. J. Neurosci. *37*, 4954–4966.

MacDonald, C.J., Lepage, K.Q., Eden, U.T., and Eichenbaum, H. (2011). Hippocampal "Time Cells" Bridge the Gap in Memory for Discontiguous Events. Neuron *71*, 737–749.

MacDonald, C.J., Carrow, S., Place, R., and Eichenbaum, H. (2013). Distinct Hippocampal Time Cell Sequences Represent Odor Memories in Immobilized Rats. J. Neurosci. *33*, 14607–14616.

Madisen, L., Garner, A.R., Shimaoka, D., Chuong, A.S., Klapoetke, N.C., Li, L., van der Bourg, A., Niino, Y., Egolf, L., Monetti, C., et al. (2015). Transgenic Mice for Intersectional Targeting of Neural Sensors and Effectors with High Specificity and Performance. Neuron *85*, 942–958.

Makino, H., and Komiyama, T. (2015). Learning enhances the relative impact of top-down processing in the visual cortex. Nat. Neurosci. *18*, 1116–1122.

Manns, J.R., Howard, M.W., and Eichenbaum, H. (2007). Gradual Changes in Hippocampal Activity Support Remembering the Order of Events. Neuron *56*, 530–540.

Marques, T., Nguyen, J., Fioreze, G., and Petreanu, L. (2018). The functional organization of cortical feedback inputs to primary visual cortex. Nat. Neurosci. *21*, 757–764.

Marshel, J.H., Garrett, M.E., Nauhaus, I., and Callaway, E.M. (2011). Functional specialization of seven mouse visual cortical areas. Neuron *72*, 1040–1054.

Matsui, T., and Ohki, K. (2013). Target dependence of orientation and direction selectivity of corticocortical projection neurons in the mouse V1. Front. Neural Circuits *7*, 1–9.

Matsumura, N., Nishijo, H., Tamura, R., Eifuku, S., Endo, S., and Ono, T. (1999). Spatial- and Task-Dependent Neuronal Responses during Real and Virtual Translocation in the Monkey Hippocampal Formation. J. Neurosci. *19*, 2381–2393.

McNaughton, B.L., Mizumori, S.J., Barnes, C.A., Leonard, B.J., Marquis, M., and Green, E.J. (1994). Cortical representation of motion during unrestrained spatial navigation in the rat. Cereb. Cortex *4*, 27–39.

Milner, A.D., and Goodale, M.A. (2008). Two visual systems re-viewed. Neuropsychologia *46*, 774–785.

Mineault, P.J., Tring, E., Trachtenberg, J.T., and Ringach, D.L. (2016). Enhanced Spatial Resolution During Locomotion and Heightened Attention in Mouse Primary Visual Cortex. J. Neurosci. *36*, 6382–6392.

Modi, M.N., Dhawale, A.K., and Bhalla, U.S. (2014). CA1 cell activity sequences emerge after reorganization of network correlation structure during associative learning. Elife *3*.

Murakami, T., Matsui, T., and Ohki, K. (2017). Functional Segregation and Development of Mouse Higher Visual Areas. J. Neurosci. *37*, 0731–17.

Nassi, J.J., and Callaway, E.M. (2009). Parallel processing strategies of the primate visual system. Nat. Rev. Neurosci. *10*, 360–372.

Newcombe, F.G. (1969). Missile wounds of the brain: A study of psychological deficits. London Oxford Univ. Press.

Niell, C.M. (2015). Cell Types, Circuits, and Receptive Fields in the Mouse Visual Cortex. Annu. Rev. Neurosci. *38*, 413–431.

Niell, C.M., and Stryker, M.P. (2008). Highly Selective Receptive Fields in Mouse Visual Cortex. J. Neurosci. *28*, 7520–7536.

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

Nitz, D. (2009). Neurobiology of Learning and Memory Parietal cortex , navigation , and the construction of arbitrary reference frames for spatial information. Neurobiol. Learn. Mem. *91*, 179–185.

Nitz, D.A. (2006). Tracking route progression in the posterior parietal cortex. Neuron *49*, 747–756.

O'Keefe, J. (1976). Place units in the hippocampus of the freely moving rat. Exp.

Neurol. *51*, 78–109.

O'Keefe, J., and Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Res. *34*, 171–175.

Oh, S.W., Harris, J.A., Ng, L., Winslow, B., Cain, N., Mihalas, S., Wang, Q., Lau, C., Kuan, L., Henry, A.M., et al. (2014). A mesoscale connectome of the mouse brain. Nature *508*, 207–214.

Otazu, G.H., Tai, L.H., Yang, Y., and Zador, A.M. (2009). Engaging in an auditory task suppresses responses in auditory cortex. Nat. Neurosci. *12*, 646–654.

Pachitariu, M., Stringer, C., Schröder, S., Dipoppa, M., Rossi, L.F., Carandini, M., and Harris, K.D. (2016). Suite2p: beyond 10,000 neurons with standard two-photon microscopy. bioRxiv 61507.

Pachitariu, M., Stringer, C., and Harris, K.D. (2017). Robustness of spike deconvolution for calcium imaging of neural spiking. bioRxiv 156786.

Pastalkova, E., Itskov, V., Amarasingham, A., and Buzsáki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. Science *321*, 1322–1327.

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*, 783–799.

Pho, G., Goard, M., Woodson, J., Crawford, B., and Sur, M. (2017). Task-Dependent Representations Of Stimulus And Choice In Mouse Parietal Cortex. bioRxiv 144592.

Pho, G.N., Goard, M.J., Woodson, J., Crawford, B., and Sur, M. (2018). Task-dependent representations of stimulus and choice in mouse parietal cortex. Nat. Commun. *9*, 2596.

Polack, P.O., Friedman, J., and Golshani, P. (2013). Cellular mechanisms of brain state-dependent gain modulation in visual cortex. Nat. Neurosci. *16*, 1331–1339.

Pologruto, T.A., Sabatini, B.L., and Svoboda, K. (2003). ScanImage: Flexible software for operating laser scanning microscopes. Biomed. Eng. Online *2*, 13.

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

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

Ratzlaff, E.H., and Grinvald, A. (1991). A tandem-lens epifluorescence macroscope: Hundred-fold brightness advantage for wide-field imaging. J. Neurosci. Methods *36*, 127–137.

Ravassard, P., Kees, A., Willers, B., Ho, D., Aharoni, D., Cushman, J., Aghajan, Z.M., and Mehta, M.R. (2013). Multisensory Control of Hippocampal Spatiotemporal Selectivity. Science (80-. ). *340*, 1342–1346.

Reimer, J., Froudarakis, E., Cadwell, C.R., Yatsenko, D., Denfield, G.H., and Tolias, A.S. (2014). Pupil Fluctuations Track Fast Switching of Cortical States during Quiet Wakefulness. Neuron *84*, 355–362.

Rolls, E.T., Robertson, R.G., and Georges-François, P. (1997). Spatial View Cells in the

Primate Hippocampus. Eur. J. Neurosci. *9*, 1789–1794.

Roth, M.M., Helmchen, F., and Kampa, B.M. (2012). Distinct Functional Properties of Primary and Posteromedial Visual Area of Mouse Neocortex. J. Neurosci. *32*, 9716–9726.

Saleem, A.B., Ayaz, A., Jeffery, K.J., Harris, K.D., and Carandini, M. (2013). Integration of visual motion and locomotion in mouse visual cortex. Nat. Neurosci. *16*, 1864–1869.

Saleem*, A.B., Diamanti*, E.M., Fournier, J., Harris, K.D., and Carandini, M. (2018). Coherent encoding of subjective spatial position in visual cortex and hippocampus. Nature *562*, 124–127.

Salinas, E., and Abbott, L.F. (2001). Coordinate transformations in the visual system: how to generate gain fields and what to compute with them. Prog. Brain Res. *130*.

Salz, X.D.M., Tiganj, X.Z., Khasnabish, X.S., Kohley, X.A., Sheehan, X.D., Howard, M.W., and Eichenbaum, X.H. (2016). Time Cells in Hippocampal Area CA3. J. Neurophysiol. *36*, 7476–7484.

Save, E., and Poucet, B. (2009). Role of the parietal cortex in long-term representation of spatial information in the rat. Neurobiol. Learn. Mem. *91*, 172–178.

Shuler, M.G., and Bear, M.F. (2006). Reward Timing in the Primary Visual Cortex. Science (80-. ). *311*, 1606–1609.

Smith, I.T., Townsend, L.B., Huh, R., Zhu, H., and Smith, S.L. (2017). Stream-dependent development of higher visual cortical areas. Nat. Neurosci. *20*, 200–208.

Snyder, L.H., Grieve, K.L., Brotchie, P., and Andersen, R.A. (1998). Separate body- and world-referenced representations of visual space in parietal cortex. Nature *394*, 887–891.

Terrazas, A., Krause, M., Lipa, P., Gothard, K.M., Barnes, C.A., and McNaughton, B.L. (2005). Self-Motion and the Hippocampal Spatial Metric. J. Neurosci. *25*, 8085–8096.

Tohmi, M., Meguro, R., Tsukano, H., Hishida, R., and Shibuki, K. (2014). The extrageniculate visual pathway generates distinct response properties in the higher visual areas of mice. Curr. Biol. *24*, 587–597.

Vinck, M., Batista-Brito, R., Knoblich, U., and Cardin, J.A. (2015). Arousal and Locomotion Make Distinct Contributions to Cortical Activity Patterns and Visual Encoding. Neuron *86*, 740–754.

Wan, H., Aggleton, J.P., and Brown, M.W. (1999). Different Contributions of the Hippocampus and Perirhinal Cortex to Recognition Memory. J. Neurosci. *19*, 1142–1148.

Wang, Q., and Burkhalter, A. (2007). Area map of mouse visual cortex. J. Comp. Neurol. *502*, 339–357.

Wang, Q., and Burkhalter, A. (2013). Stream-Related Preferences of Inputs to the Superior Colliculus from Areas of Dorsal and Ventral Streams of Mouse Visual Cortex. J. Neurosci. *33*, 1696–1705.

Wang, Q., Gao, E., and Burkhalter, A. (2011). Gateways of Ventral and Dorsal Streams in Mouse Visual Cortex. J. Neurosci. *31*, 1905–1918.

Wang, Q., Sporns, O., and Burkhalter, A. (2012). Network Analysis of Corticocortical Connections Reveals Ventral and Dorsal Processing Streams in Mouse Visual Cortex. J. Neurosci. *32*, 4386–4399.

Wekselblatt, J.B., and Niell, C.M. (2015). Behavioral State-Getting "In The Zone." Neuron *87*, 7–9.

Wekselblatt, J.B., Flister, E.D., Piscopo, D.M., and Niell, C.M. (2016). Large-scale imaging of cortical dynamics during sensory perception and behavior. J. Neurophysiol. *115*, 2852–2866.

Whitlock, J.R., Pfuhl, G., Dagslott, N., Moser, M.B., and Moser, E.I. (2012). Functional Split between Parietal and Entorhinal Cortices in the Rat. Neuron *73*, 789–802.

Wilber, A.A., Clark, B.J., Forster, T.C., Tatsuno, M., and McNaughton, B.L. (2014). Interaction of Egocentric and World-Centered Reference Frames in the Rat Posterior Parietal Cortex. J. Neurosci. *34*, 5431–5446.

Xiang, J.-Z., and Brown, M.W. (1998). Differential neuronal encoding of novelty , familiarity and recency in regions of the anterior temporal lobe. Neuropharmacology *37*, 657–676.

Xing, J., and Andersen, R.A. (2000). Models of the posterior parietal cortex which perform multimodal integration and represent space in several coordinate frames. J. Cogn. Neurosci. *12*, 601–614.

Yang, Z., Heeger, D.J., and Seidemann, E. (2007). Rapid and Precise Retinotopic Mapping of the Visual Cortex Obtained by Voltage-Sensitive Dye Imaging in the Behaving Monkey. J. Neurophysiol. *98*, 1002–1014.

Zeng, H., and Madisen, L. (2012). Mouse transgenic approaches in optogenetics (Elsevier B.V.).

Zhang, K., Ginzburg, I., McNaughton, B.L., and Sejnowski, T.J. (1998). Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. J. Neurophysiol. *79*, 1017–1044.

Zhang, S., Xu, M., Kamigaki, T., Hoang Do, J.P., Chang, W.-C., Jenvay, S., Miyamichi, K., Luo, L., and Dan, Y. (2014). Selective attention. Long-range and local circuits for top-down modulation of visual cortex processing. Science *345*, 660–665.

Zhuang, J., Ng, L., Williams, D., Valley, M., Li, Y., Garrett, M., and Waters, J. (2017). An extended retinotopic map of mouse cortex. Elife *6*, 1–29.