Signatures of cortical multisensory integration in mice performing a novel visuotactile evidence accumulation task

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Gerion Nabbefeld, M.Sc.

aus

Berlin, Deutschland

Berichter: Universitätsprofessor Dr. Björn Kampa

Universitätsprofessor Dr. Simon Musall

Tag der mündlichen Prüfung: 15.09.2022

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek verfügbar.

„The difference between screwing around and science is writing it down. “

- Adam Savage

Abstract

Much effort has been focused on studying how the brain processes information from our individual senses. However, the neural mechanisms, that allow the effortless integration of unisensory inputs into multisensory percepts, are largely unknown. To study how neural circuits integrate visual and tactile information, we developed a multisensory discrimination task for head-fixed mice. Here, two sequences of visual, tactile or combined visuotactile stimuli are presented on both sides of the mouse, which has to indicate the higher-rate target-side to obtain a water reward. To ensure integration of sensory information over the entire stimulus period, a short delay was added before the response. Mice achieved high accuracy in all conditions, with improved performance in the multisensory condition. This behavioral task gave us the opportunity to investigate the neural circuits that allow mice to synergistically use both the visual and tactile sensory information to solve the behavioral task.

We then used widefield imaging to measure cortex-wide activity in transgenic mice expressing the Ca^{2+}-indicator GCaMP6s in all cortical excitatory neurons. Here, we found that multisensory stimuli evoked higher neuronal activity compared to unisensory stimulation. This was most evident in the rostrolateral association area RL and parts of medial frontal cortex (mFC), which reliably responded to both visual and tactile stimuli. To better isolate sensory responses from co-occurring task- or behavior-related activity, we used a linear encoding model. Including a multisensory interaction-term significantly improved the predictions of cortical activity. With this approach we identified two key features of sensory evoked responses, depending on the stimulus condition. First, in unisensory trials mice display cross-modal inhibition. Here, in addition to the main sensory responses in the corresponding sensory cortex, robust inhibition of activity in the non-matching sensory cortex was found. Second, we found additional superadditive responses in multisensory trials, likely representing the absence of cross-modal inhibition as well as increased activity in areas RL and mFC.

To understand how sensory information is used to guide behavioral decisions, we first investigated which brain areas displayed activity that reliably reflected the target stimulus side. Here, the medial motor cortex more faithfully reflected the target-side in tactile trials, while secondary visual areas were more reliable in visual trials. In multisensory trials, both

regions accurately reflected the target-side, likely resulting in higher certainty and improved performance in multisensory trials. Finally, using a choice-decoder we identified choice-related neural activity in the anterolateral motor cortex (ALM), as well as in licking-related regions of the primary motor and somatosensory cortex. With this approach, we found no clear modality-specific differences, suggesting that the same neural circuits form decisions in all stimulus conditions.

Our results demonstrate that multisensory stimulation cause widespread cortical activation in mice, which leads to improved task performance. Here, cross-modal inhibition in unisensory trials and superadditive multisensory integration especially in RL and mFC were found in multisensory trials, likely aiding mice in performing the individual task condition. Sensory information is then accumulated over the stimulus period in secondary visual areas and medial motor cortex and this information converges in the secondary motor cortex to form modality-unspecific decisions. These findings give us a much deeper understanding of how the brain processes and generalizes sensory information in order to guide behavioral decisions.

Zusammenfassung

Eine Vielzahl an wissenschaftlichen Arbeitsgruppen konzentriert sich darauf zu erforschen, wie das Gehirn Informationen verarbeitet, die von unseren individuellen Sinnen stammen. Allerdings sind die neuronalen Mechanismen, die die mühelose Integration unisensorischer Eingangssignale in multisensorische Wahrnehmungen ermöglichen, nach wie vor nicht gut verstanden. Um zu untersuchen, wie neuronale Schaltkreise visuelle und taktile Informationen integrieren, haben wir ein multisensorisches Diskriminationsparadigma für kopffixierte Mäuse entwickelt. In selbigem werden zwei Sequenzen von visuellen, taktilen oder kombinierten visuell-taktilen Reizen auf beiden Seiten der Maus präsentiert. Anschließend müssen die Mäuse die Zielseite mit der höheren Rate an Reizen anzeigen, um eine Wasserbelohnung zu erhalten. Um die Integration der sensorischen Informationen über das gesamte Stimulationsintervall zu gewährleisten, wurde vor der Möglichkeit zu antworten eine sogenannte „Delay"-Periode, also eine kurze Verzögerung, eingefügt. Alle getesteten Mäuse erreichten in sämtlichen Konditionen eine hohe Erfolgsrate, mit einer verbesserten Leistung in der multisensorischen Kondition. Dieser Verhaltensversuch gab uns die Möglichkeit die neuronalen Schaltkreise zu studieren, die es den Mäusen erlauben visuelle und taktile Informationen synergetisch zu nutzen um den Verhaltensversuch zu lösen.

Im Anschluss nutzten wir Weitfeldbildgebung, um die kortikale Aktivität während des Verhaltensversuches zu messen. Hierzu wurden transgene Mäuse verwendet, die den Ca^{2+}-Indikator GCaMP6s in allen kortikalen, exzitatorischen Zellen exprimieren. Dadurch haben wir festgestellt, dass multisensorische Reize eine stärkere neuronale Aktivität hervorrufen als unisensorische Reize. Am deutlichsten wurde dies in dem rostrolateralen Assoziationsareal RL und in Teilen des medialen Frontalkortex (mFC), die zuverlässig sowohl auf visuelle als auch auf taktile Reize reagierten. Um die sensorischen Antworten besser von zeitgleich stattfindender versuchs- oder verhaltensbezogener Aktivität zu isolieren, verwendeten wir ein lineares Kodierungsmodell. Die Einbeziehung eines multisensorischen Interaktionsterms in diesem Modell verbesserte die Prognosen der Hirnaktivität signifikant. Mit diesem Ansatz konnten wir zwei Hauptmerkmale der hervorgerufenen sensorischen Antworten in Abhängigkeit der Modalität des Reizes identifizieren. Zum einen fanden wir inter-modale Inhibition, wenn Mäuse unisensorische Versuchsdurchläufe absolvierten. Hier fanden wir, neben den exzitatorischen Reaktionen im entsprechend stimulierten sensorischen Kortex,

auch eine Inhibition der Aktivität im jeweils anderen sensorischen Kortex. Zum anderen, fanden wir zusätzliche, superadditive multisensorische Antworten, die über die lineare Kombination der unisensorischen Antworten hinaus gingen. Dies ist vermutlich begründet in der Abwesenheit der inter-modalen Inhibition und einer erhöhten Aktivierung der Areale RL und mFC.

Um zu verstehen, wie sensorische Informationen die Formation von Verhaltensentscheidungen lenken, haben wir zunächst untersucht, welche Hirnareale Aktivität aufweisen, die die Zielseite zuverlässig widerspiegelt. Hierbei zeigte sich, dass der mediale Motorkortex die Zielseite in taktilen Versuchen zuverlässiger widerspiegelte, während die höheren visuellen Areale diese in visuellen Versuchen verlässlich reflektierten. Bei multisensorischen Versuchen spiegelten beide Regionen die Zielseite zuverlässig wider, was wahrscheinlich zu einer höheren Sicherheit der Tiere und einer besseren Erfolgsrate in multisensorischen Versuchen führte. Schließlich identifizierten wir neuronale Aktivität, die im Zusammenhang mit den Entscheidungen der Tiere stand, mit Hilfe eines entsprechenden Decoders. Solche Aktivitätsmuster fanden wir im anterolateralen Motorkortex (ALM) sowie weiteren Regionen des primären motorischen und somatosensorischen Kortex, welche in Verbindung mit Leckbewegungen der Tiere stehen. Mit diesem Ansatz konnten wir keine eindeutigen modalitätsspezifischen Unterschiede feststellen, was darauf hindeutet, dass dieselben neuronalen Schaltkreise verantwortlich sind die Entscheidungen in allen Modalitätskonditionen zu fällen.

Unsere Ergebnisse zeigen, dass multisensorische Stimulation weitreichende kortikale Aktivierung verursacht, welche zu einer verbesserten Erfolgsrate der Mäuse führt. Hierbei fanden wir inter-modale Inhibition in unisensorischen Versuchen, sowie superadditive multisensorische Integration in multisensorischen Versuchen, vor allem in den Arealen RL und mFC, was die Mäuse wahrscheinlich bei der Durchführung der jeweiligen Konditionen unterstützt. Sensorische Informationen werden dann über die Dauer der Stimulation in den sekundären visuellen Arealen und im medialen Motorkortex akkumuliert und anschließend im sekundären Motorkortex zusammengeführt, wo modalitätsunspezifische Entscheidungen getroffen werden. Diese Ergebnisse geben uns ein viel besseres Verständnis darüber wie das Gehirn sensorische Informationen verarbeitet und generalisiert um das Verhalten zu leiten.

Contributions

The presented multisensory evidence-accumulation task was designed by Björn Kampa and myself, with advice by Simon Musall. This task is based on an earlier multisensory task developed in our Lab by Alexander Bexter and Björn Kampa. The choice of sensory stimuli, the inclusion of a running wheel and the movable lick-spouts to compensate for side-biases of the animals were inspired by this previous study.

While establishing this new setup and task I had assistance from two Master students under my supervision: Severin Graff, who trained the first cohort of animals to establish the new setup and task, in combination with the widefield Ca^{2+}-imaging and Anna Ostenrath, who developed the first preliminary analysis of the widefield imaging data, with allowed us to refine the experimental procedure.

Surgeries for the widefield-imaging experiments were performed by myself, with advice by Simon Musall. Further surgeries, contributing to Figure 4, B were performed by Simon Musall.

Training of the widefield cohort was performed by myself, with assistance from Sandra Brill towards the end of the experimental run. Subsequent cohorts contributing to Figure 4, B were trained by Mira Ritter, Lena Kricsfalussy-Hrabár, Sacha Abou Rachid, Irene Lenzi, Emma Cravo and myself.

The continuous multi-camera acquisition software, as well as the microcontroller-based behavioral software was developed by myself.

Data analysis and visualization was conducted by myself, with advice from Björn Kampa and Simon Musall, concerning analytical approaches and the development of the story.

Acknowledgements

First of all, I want to thank my doctor father Prof. Björn Kampa for giving me to opportunity to conduct my research in his lab. It has been nearly eight years now and it is exciting to see how far the lab has come in this time. I deeply appreciated all the freedom I was given over the years to pursue my projects as I saw fit. This allowed me to develop into an independent researcher capable of pursuing my own research projects in the future. As such I'm looking forward to conducting my next project also in this groups and keep this going for a little bit longer.

I also deeply appreciated the collaboration with my second supervisor Prof. Simon Musall. From helping me establish the surgical procedure in our lab, giving me advice in the design of the behavioral task, to advice on the data analysis. It is hard to imagine how far this project would have come without your help. And I'm looking forward to continuing working together.

I also want to thank all the members of both the AG Kampa and the AG Musall. I have enjoyed my time over the last years with all of you. And I especially want to thank all of the people that assisted me in training the subsequent cohorts of mice and that continue to explore the nature of cortical multisensory integration using this behavioral paradigm.

As a member of the RTG2416 Multisenses-Multiscales I want to thank the German Research Foundation (DFG) for enabling me to conduct this research, as well as allowing me to attend seminars, workshops and conferences that undoubtedly aided my development as a researcher.

I am immensely grateful to my family for all their support and for enduring me being so occupied by my work often for weeks and months at a time. I looking forward to finally visiting all of you again soon!

Finally, I want to thank my amazing fiancée Jenice Linde. I doubt that I would have made it through those last year of the PhD without you. All the long hours and weekends we had to work through, just to keep our projects afloat. You made this possible and you made it all worthwhile throughout!

Funding

I received financial support from the German Research Foundation DFG (RTG2416 MultiSenses-MultiScales).

List of Figures

List of Abbreviations

2AFC	Two-alternative forced choice
ACC or ACAd	Anterior cingulate cortex
AL	Anterolateral visual association area
AMM	Anteromedial motor cortex
AUC	Area under the curve (here: receiver-operator characteristic curve)
AUD	Auditory
AUDp	Primary auditory cortex
Ca^{2+}	Calcium
CaM	Calcium-modulated protein
CI	Confidence interval
CaMK2α	Calcium-calmodulin-dependent kinase II
cpd	Cycles per degree
cv	Cross-validated
DAQ	Data acquisition
GEVIs	Genetically encoded voltage indicators
GECIs	Genetically encoded calcium indicators
GFP	Green fluorescent protein
ITI	Inter trial interval
mFC	Medial frontal cortex
mPFC	Medial prefrontal cortex
MM	Medial motor cortex
MO	Motor
MOp	Primary motor cortex
MOs	Secondary motor cortex
MSI	Multisensory integration

PBS	Phosphate buffered saline
PC	Personal computer
PCIe	Peripheral Component Interconnect Express
PPC	Posterior parietal cortex
R^2	Coefficient of determination, here interpreted as the proportion of explained variance in the context of multiple linear regression models
ΔR^2	Unique explained variance
RL or VISrl	Rostrolateral visual association area
ROI	Regions of interest
RSC	Retrosplenial cortex
s.d.	Standard deviation
s.e.m.	Standard error of mean
SC	Superior colliculus
SNR	Signal-to-noise ratio
SS	Somatosensory
SSp	Primary somatosensory cortex
SVD	Singular value decomposition
TCP	Transmission Control Protocol
tjM1	Tongue and jaw region of the primary motor cortex
tjS1	Tongue and jaw region of the primary somatosensory cortex
tTA	Tetracycline-controlled transactivator protein
TTL	Transistor-transistor logic
V1 or VISp	Primary visual cortex
VIS	Visual
VISa	Anterior visual association area
wS1 or SSp-bfd	Primary barrel cortex

Table of contents

1 Introduction

Every day we are confronted with a complex multisensory environment that we have to make sense out of and interact with. Using specialized sensory organs, we decompose this environment into information of individual sensory modalities. Over the last decades, vast efforts have been devoted to study how our individual sensory systems transduce signals and extract useful features about our environment (Hudspeth & Logothetis, 2000). However, we do not perceive out environment as separate sensory spaces, but rather as being composed of unified objects consisting of all of their different sensory features. The exact mechanisms of how those sensory streams converge and where in the processing hierarchy this takes place to produce these multisensory percepts, is still unresolved (Stein & Stanford, 2008).

Multisensory integration (MSI) is thought to be most widespread in the mammalian cerebral cortex (Choi et al., 2018). Widefield calcium imaging has become a powerful technique to measure such widespread cortical activation patterns during sensory stimulation in transgenic mice (Couto et al., 2021; Ren & Komiyama, 2021; Wekselblatt et al., 2016). In addition, sensory perception has been shown to depend strongly on the behavior of the animal (Carandini & Churchland, 2013; Panzeri et al., 2017; Wekselblatt et al., 2016). Therefore, to investigate how sensory information is integrated across modalities and how this affects perceptual decisions, we use widefield calcium-imaging to measure activity of the dorsal cortex of mice performing a novel multisensory discrimination task. In the following chapters these different aspects will be explain in more detail.

1.1 Multisensory integration

The integration of information from our different senses into a coherent picture of our surrounding is a central feature of our nervous system and the basis of our conscious perception. Deficits in this process have been associated with a variety of neurological disorders. For example, patients on the autism-spectrum have been found to display deficits in MSI, prohibiting the combination of even simple non-social sensory inputs (Ostrolenk et al., 2019). In contrast, schizophrenia patients have been found to display increased MSI compared to healthy controls, potentially leading to inappropriate integration of unrelated

sensory inputs (Stone et al., 2011). This highlights that the process of MSI is a difficult balancing act, with severe consequences when impaired. While the processing of unisensory information in the brain has been a central focus of study since the first neuroscientific studies, the mechanisms of how this information is integrated across sensory modalities is still poorly understood. In ordered, to understand the cause and the mechanisms of such neurological disorders and develop treatments of the symptoms or even the underlying causes, we need a better understanding how the brain integrates multisensory information in healthy subjects.

1.1.1 Evolution of multisensory integration

Sensing the environment and producing appropriate responses, is one of the most fundamental principles of life. As multicellular life became increasingly complex, some aspects of sensing the environment to guide responses of an organism were centralized in the form of a central nervous system. One of the driving forces of our early evolution is thought to be the coevolution of predator and prey (Abrams, 1986). As predators evolved the ability to better locate and hunt their prey, prey in turn had to adapt to detect predators earlier to be able to escape in time. As such, the survival of both these groups depended on reliably and accurately detecting the other. This in part drove the development and refinement of our individual sensory systems. But no system is perfect and likely never will be. As such, sensory systems can't detect signals with perfect accuracy. Each sensory measurement has some amount of uncertainty attached to it. Natural selection can encourage more sensitive and accurate signal detection, improved localization and identification of signal sources, but only up to a certain point after which the energetic cost of further improving a sense will outweigh the benefits. However, by integrating information over different sensory modalities, we can make much more precise estimates of our environment (Körding et al., 2007; Murray et al., 2016). As such, multisensory integration has emerged as a central function and key organizational principle of our brain (Murray et al., 2016; Stein & Stanford, 2008).

1.1.2 Multisensory integration in the superior colliculus

We sense our environment using a wide array of specialized sensory organs, that are each optimized to transduce signals from a specific sensory modality. By integrating information from multiple sensory modalities (cross-modal), the likelihood of detecting an event and reacting to it can be improved (Körding et al., 2007). The term “multisensory integration” (MSI) describes this process of merging information from different senses to be used synergistically and improve our interpretation of the environment (Stein & Stanford, 2008). This process has been studied extensively on the level of single neurons in the superior colliculus (SC; Meredith & Stein, 1983, 1986; Stein & Stanford, 2008; Wallace et al., 1996).

One of the central functions of the SC is to locate sensory stimuli across sensory modalities. As such, this structure contains a high number of multisensory responding neurons (Stein & Stanford, 2008). Activity in SC is related to obvious orientation movements of subjects (Stein et al., 1988). These responses can be enhanced by presenting cross-modal stimuli from a similar direction. This feature of multisensory neurons in SC is facilitated by their modality-specific and spatially overlapping receptive fields (Alvarado et al., 2007). These factors have made the SC a prime target for early studies, revealing general principles of multisensory integration (Wallace et al., 2004).

1.1.3 Neuronal computations underlying multisensory integration

Individual neurons in the SC can display either enhanced or depressed responses to multisensory stimuli compared to the individual unisensory stimuli. This behavior often depends on the features of the stimuli, particularly the congruence of the spatial origin of the individual stimuli and their relative timing (Stein & Stanford, 2008). For example, when signals are registered in two separate sensory modalities, received from similar directions and at roughly the same time, it is very likely that they originate from the same event (Körding et al., 2007). Neurons that react to those stimuli with enhanced multisensory responses, would therefore be more likely to result in a detection of this event and subsequent orienting behaviors (Stein & Stanford, 2008). In addition to an increased likelihood of responses, individuals also display reduced reaction times when presented with matching multisensory stimuli (Diederich & Colonius, 2004; Gleiss & Kayser, 2012), in line with the evolutionary

motivation (see section 1.1.). In contrast, the same stimuli from separate spatial locations, are less likely to relate to the same event and more often result in multisensory depression and would reduce the likelihood of responses. Therefore, both multisensory enhancement and multisensory depression represent important neurological computations.

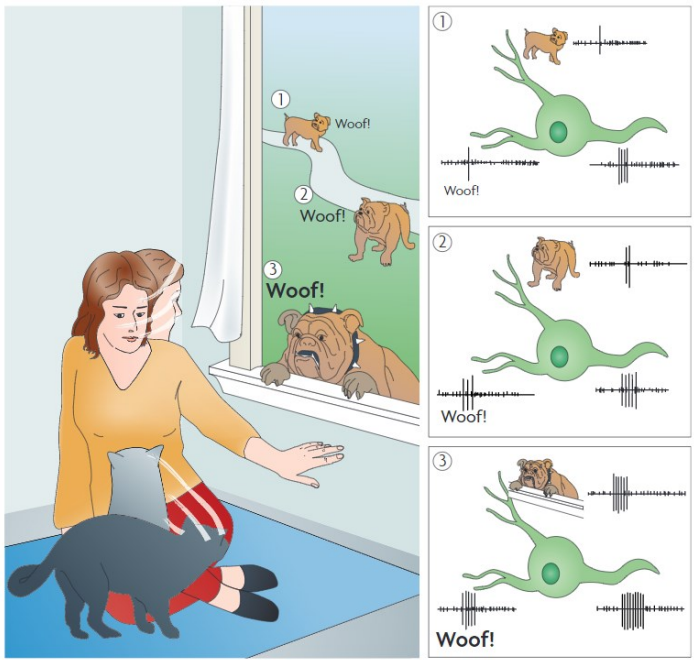

Figure 1: Multisensory processing visualized in an audio-visual context.
Multisensory integration improves our ability to quickly recognize potential threats in our environment and react accordingly. The enhancement effect elicited by multimodal integration is the strongest when the individual unisensory stimuli are weak (scenario 1, superadditive enhancement). This enhancement effect decreases as unisensory signal strengths increase (scenarios 2 and 3, additive and subadditive responses; figure reproduced with permission from Springer Nature: Stein & Stanford, 2008).

Multisensory enhancement can be the result of different underlying computations and is found in three forms: superadditive, additive and subadditive multisensory integration (Stanford & Stein, 2007; Stein & Stanford, 2008).

In the case of additive multisensory integration (Figure 1, scenario 2), individual neurons display responses (measured as evoked firing rates) equivalent to the sum of the evoked responses by the individual unisensory stimulus components. Superadditive multisensory integration (Figure 1, scenario 1) results in responses exceeding the linear addition of the unisensory responses and subadditive integration (Figure 1, scenario 3) results in responses lower than the summed unisensory responses. Both the superadditive and the subadditive responses are forms of non-linear integration, while additive responses represent a linear integration.

This example also illustrates a key feature of multisensory integration termed “inverse-effectiveness”. This principle states that multisensory integration has the greatest benefit, when the unisensory stimulus components are weak or hard to discriminate (Choi et al., 2018; Stanford & Stein, 2007; Stein & Stanford, 2008). Therefore, to investigate multisensory

integration we would ideally study its impact in conditions with different difficulties in the form of varying strength of sensory evidence.

These forms of multisensory integration reflect general computations that are not exclusive to the SC. Similar computations have also been found in macaque prefrontal cortex neurons, when integrating auditory and visual sensory information (Sugihara et al., 2006). Many neurons in the area of study displayed multisensory responses to combined auditory and visual stimuli, to a large extent in the form of multisensory depression. However, when multisensory enhancement was found, the effect was greatest in response to congruent facial imagery (visual) and auditory vocalizations, in contrast to more generic stimuli. This represents an a much more specialized cortical circuit for multisensory integration, when compared to the SC.

While multisensory integration in the SC particularly emphasizes the spatial and temporal congruence of stimuli to locate events and guide orienting behaviors (Wallace et al., 2004), cortical integration appears to be more complex, emphasizing more abstract stimulus features or relationships (Sugihara et al., 2006). To study the role of multisensory integration in the context of higher cognitive functions, we therefore ought to investigate how cross-modal stimuli are processed in the cerebral cortex.

1.1.4 Cortical organization

The cerebral cortex has long been associated with higher cognitive functions, such as working memory and consciousness (Frith & Dolan, 1996). Besides midbrain structures such as the superior colliculus, multisensory integration is thought to be most widespread in the cerebral cortex (Choi et al., 2018). The cerebral cortex also has a central role in perception, which becomes especially obvious when this structure is impaired. For example, individuals that suffered from damage to their visual cortex resulting in partial or complete blindness, can still express rudimentary features of sight (Pöpple et al., 1973), likely based on subcortical visual processing. However, patients consciously report being blind in this region of their visual field. Even when patients registered a stimulus in this blind field of view, they were often unaware that it was in fact a visual stimulus they perceived (Pöpple et al., 1973). Sensory processing is a complex interplay involving many different brain structures, but observations such as the

one described above emphasize the central role that the cerebral cortex plays in our conscious perception of the environment.

Much of our early view of cortical organization was based on lesion studies; observations of damage in a particular region of the cortex coinciding with perceptual and cognitive deficits. Such as the early observation that individuals that became blind as consequence of an injury, often displayed damage to the occipital lobe (Colombo et al., 2002). Similarly, patients that developed a speech impediment after a head injury displayed damage in the same spot of the left lateral frontal lobe, later termed Broca's area (Dronkers et al., 2007). These first studies of cortical function shaped the view of specific perceptual and cognitive functions, associated with particular brain structures, consistent over individuals. In addition, these discoveries of cortical functions closely matched the cortical maps described by Brodmann (Brodmann, 1909), indicating a close link between cortical anatomy and its function.

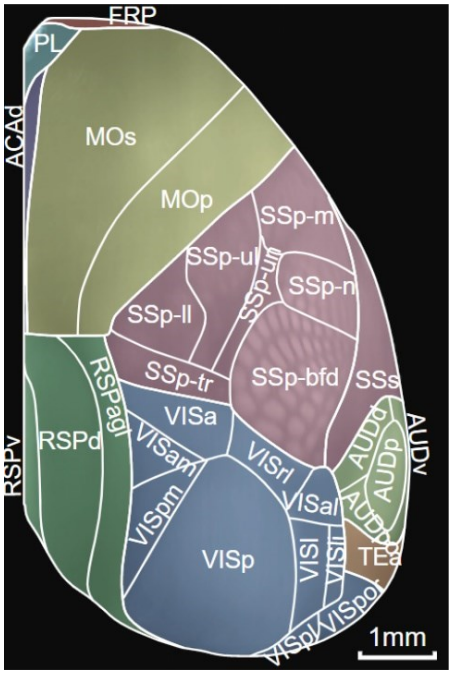

Figure 2: Dorsal view on murine cortical areas outlined according to Allen CCF.
The mice's lisencephalic cerebral cortex is divided into functional areas. Many of these regions were identified to process sensory information (e.g. visual (VIS; blue), somatosensory (SS; red) or auditory (AUD; light green) information) and prepare and execute motoric functions (MO; yellow regions). These functional regions can be further subdivided into primary (VISp, SSp, AUDp, MOp) and individual higher order areas (indicated by white lines).

Reprinted from Cell, Wang et al. 2020. Copyright (2020), with permission from Elsevier.

The mammalian cerebral cortex displays common principles of organization that transform sensory information into motor responses. Sensory information enters the cortex in a modality-specific manner. With the exception of the olfactory sense, information from our sensory organs is relayed over distinct thalamic nuclei to separate cortical regions, the primary sensory cortices (Choi et al., 2018). These primary sensory areas are the first stations of sensory processing in cortex (compare Figure 2, VISp/SSp/AUDp). More specific sensory features are then processed in specialized secondary sensory cortices surrounding the primary sensory areas (see Figure 2; Felleman & van Essen, 1991; Wang & Burkhalter, 2007). These primary and secondary sensory

areas constitute the unisensory domains of the mammalian cortex. The existence of these sensory-specific domains is a highly conserved feature of cortical organization in mammals (Carandini & Churchland, 2013; Krubitzer & Kahn, 2003). Beyond these unisensory regions, association areas are found that link sensory and motor regions (Krubitzer, 2007). These association areas have been assigned a variety of complex functions including long-term memory, working memory, spatial reasoning, abstract thought, emotions, attention and action planning (Frith & Dolan, 1996). As such, these association areas process sensory information across individual modalities and perform more complex and abstract computations on these signals. Association cortices are also connected to premotor regions (see Figure 2; MOs), that relate to the planning of future motor responses (Chen et al., 2017; Esmaeili et al., 2020). Primary motor cortex represents the predominant output region of the cortex that plays a central role in the initiation of voluntary movements (Ebbesen & Brecht, 2017).

This modular organization of the cerebral cortex is thought to be a very flexible schema that allowed mammals to adapt to various environments and promoted the gradual development of more complex cognitive functions (Krubitzer, 2007), particularly by adding functional modules and complexity to the cortical association areas.

1.1.5 Multisensory integration in the cerebral cortex

Multisensory integration is an essential feature of the cortex, especially cortical association areas. A large body of work has characterized unisensory responses in the primary and secondary sensory cortices from humans to rodents (Allen et al., 2017; Ferezou et al., 2007; Grill-Spector & Malach, 2004; Hubel & Wiesel, 1962; Wang & Burkhalter, 2007). However, recently an increasing number of studies have also reported sensory responses to stimuli not matching the preferred sensory modality of these previously considered unisensory regions (Falchier et al., 2002; Iurilli et al., 2012; Laurienti et al., 2002; Molholm et al., 2002; Morrell, 1972). These discrepancies raised concerns about our understanding of the general cortical organization. To address these concerns, efforts have been made to systematically characterize the responsiveness of neurons distributed over vast regions of the rat cerebral

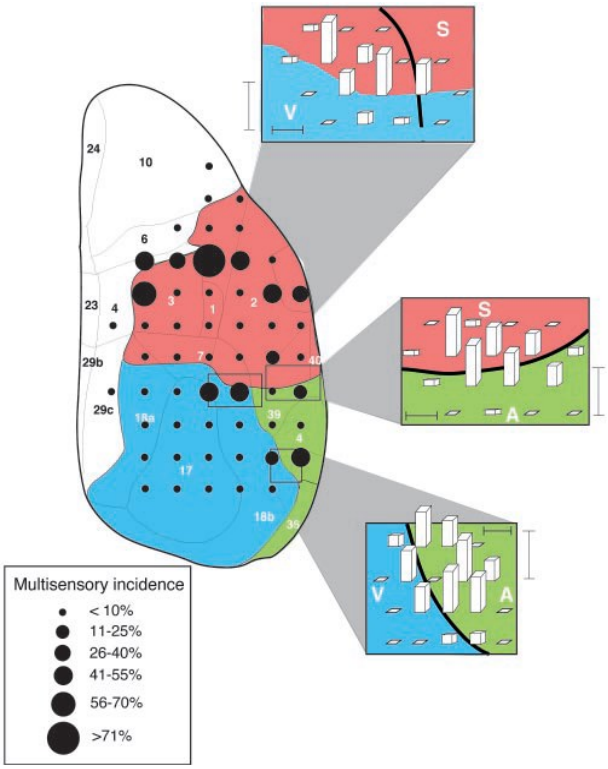

Figure 3: Distribution of multisensory neurons in the rat cerebral cortex
Characterization of neuronal responses using electrophysiological recordings in anesthetized rats. Shading indicates unisensory domains of cortex (blue: visual, red: somatosensory, green: auditory). Circles indicate fraction on multisensory responding neurons in a given recording site. Insets represent higher spatial density measurements along the borders of unisensory regions. Reprinted from PNAS (Wallace 2004), Copyright (2004) National Academy of Sciences, U.S.A.

cortex using electrophysiological recordings (Wallace et al., 2004). The study by Wallace et al., 2004 represents one of the most thorough overviews of the distribution of uni- and multisensory responses in the rodent cortex (Figure 3). In general, this study confirmed the classical view of cortical organization, reporting primarily unisensory responses covering the majority of the sensory dorsal cortex of rats. Considerable fractions of multisensory responding neurons were primarily found at the intersections of unisensory domains at the boarders of the occipital, parietal and temporal lobe (Figure 3; Wallace et al., 2004). In rodents, these regions seem to display both features of secondary sensory areas and multisensory association areas (Olcese et al., 2013; Wang & Burkhalter, 2007). Particularly for the integration of visual and tactile information, the intersecting rostrolateral visual area (area RL; compare Figure 2 “VISrl”) has been investigated in detail in a study in mice (Olcese et al., 2013). In this region a mix of visual and tactile responding cells was found, as well as a high fraction of neurons responding to both modalities. Similar responses have also been found in the anterolateral visual area (area AL) in a study of audiovisual integration in mice performing a behavior task (Meijer et al., 2020). These findings indicate that association areas such as RL and AL play an important role in multisensory integration, especially in a behavioral context.

We want to better understand how primary sensory cortices and higher association areas such as RL work together to integrate multisensory information in a behavioral context and how this information is used to guide behavioral decisions. Therefore, we aimed to study how

activity in these regions is modulated in mice performing a behavioral task involving perception and discrimination of unisensory- and multisensory cues.

1.2 Mouse as a model organism in neuroscience

To investigate fundamental neuroscientific principles, we often turn to animal models. Over the last decades the mouse has become a central model organism in neuroscience (Andermann et al., 2011; Carandini & Churchland, 2013), due to the available genetic tools (O'Connor et al., 2009) and behavioral paradigms allowing for neurological measurements during active behavior (Carandini & Churchland, 2013; Mayrhofer et al., 2013). Particularly this combination of complex behaviors and advanced electrophysiological and imaging techniques, make this a very powerful model system. However, as with any animal model, we can't simply "ask" mice how they perceived a stimulus, we instead use behavioral tasks to probe how sensory information is used to guide behavioral responses.

Even though, association areas contain far larger fractions of multisensory responding neurons (Meijer et al., 2020; Olcese et al., 2013; Wallace et al., 2004), modulation across modalities can also be found in primary sensory areas (Iurilli et al., 2012). It is not known to which extent these two distinct processes contribute to multisensory integration in a behavioral context. It is therefore crucial to study and compare how sensory information is processed in both primary sensory cortices and association areas. Currently, the best approach to simultaneously record cortical activity from many different cortical areas is widefield calcium-imaging (see section 1.4). Crucially, due to the anatomy of the mouse cortex in combination with available preparation techniques (Guo et al., 2014; Silasi et al., 2016), we are only able to reliably record activity in the primary visual and somatosensory cortex of mice (see Figure 13B). Due to these limitations of recording activity from the primary auditor cortex of mice, we were limited to a behavior task that utilizes visual and tactile sensory information.

A visuotactile discrimination task suited for this purpose, has not been available. Therefore, we first had to establish a novel visuotactile discrimination task for head-fixed mice, suitable to study mechanism of cortical multisensory integration. The current behavioral setup and -task are based on an early version of a visuotactile discrimination task, previously presented by Bexter (Bexter, 2022).

1.3 Behavior tasks for rodents

In order to probe sensory perception in animals, we need a readout of how a given stimulus is perceived by an individual. This is commonly achieved by letting animals perform a behavior task. At their essence, all behavior tasks aim to make inferences of how sensory inputs are translated to behavioral responses of animals. The focus of such studies can range from investigating the general psychological principles up to precise signal processing pathways. Many different behavioral paradigms have been established over the years, optimized for the model organism of study and tailored to the precise question at hand. Different behavior task designs for rodents have previously been reviewed (for review, see Carandini & Churchland, 2013). In the following we will focus on behavior tasks for rodents, even though many of the general principles and approaches also apply to other model systems.

1.3.1 Water-restriction

Animals need to be motivated to reliably perform behavior tasks. This is commonly achieved by controlling their water or food intake outside of the behavior tasks and then dispensing the corresponding reward for correct performance during the behavior training (Goltstein et al., 2018). Mice are able to learn behavior tasks under both water- and food restriction and are able to achieve the same performance levels, however mice tend to learn quicker under water restriction (Goltstein et al., 2018). Both of these types of motivation require close monitoring of the physiological state and the well-being of individuals. Notably, there are efforts to simplify these approaches by giving animals more control over their intake, while still maintaining the motivation to perform the behavior tasks. A recent approach has been tested in both rats (Reinagel, 2018) and mice (Urai et al., 2021). Here, animals are given free access to water supplemented with low levels of citric acid, which allowed for free access to water, but results in a self-controlled reduced intake (Urai et al., 2021). As reward in the behavior tasks, either neutral or sweetened water was used to further incentivize the animals. Further, it has been suggested to add non-nutritional sweeteners to the water reward to increase the motivation without affecting the animals' diet (Reinagel, 2018). The behavior task used in this thesis utilizes a standard water-restriction schema, but such alternative

approaches for self-controlled intake are very promising for future research, as they become increasingly established.

1.3.2 Behavioral responses as decision indicators

To judge the animals' decisions, a behavioral readout is required. This can be implemented in different ways. In a freely moving task design, animals commonly have to move to a specific location, indicating their responses to receive a reward (Scott et al., 2015). However, many advanced neurological recording techniques require the head of the animal to remain in a fixed position, referred to as head-fixed behaviors. These tasks necessitate alternative response behaviors. For those purposes, the operation of levers (McKerchar et al., 2005) or tracking devices (Burgess et al., 2017; Sanders & Kepecs, 2012) operated with paws, have previously been used. It has been suggested that coupling the sensory stimulus to the use of a tracking device, giving direct sensory-motor feedback, can help mice to learn a behavior task quicker (Burgess et al., 2017). However, subsequent studies suggested that precisely this direct sensory-motor coupling instead of a delayed response, reduces cortical signals of response preparation linked to decision making, instead indicating subcortical decision processes in such a task design (Steinmetz et al., 2019; Zatka-Haas et al., 2021). Particularly for the purpose of studying cortical perceptual decision-making, this could be counterproductive.

Alternatively, many previous behavior tasks for rodents have used delayed lick-responses (Bexter, 2022; Chen et al., 2017; Esmaeili et al., 2021; Gallero-Salas et al., 2021; Guo et al., 2014; Li et al., 2015; Mayrhofer et al., 2013; Musall et al., 2019). Here, first a sensory stimulus is presented, followed by a delay and only afterwards the animals indicate their responses, separating sensation from the expression of their decisions. This response behavior is well established and activity related to short-term memory (Chen et al., 2017; Gallero-Salas et al., 2021), response preparation (Chen et al., 2017; Esmaeili et al., 2020; Musall et al., 2023; Xu et al., 2022) and movement initiation (Esmaeili et al., 2020; Xu et al., 2022) are reliably reported in the premotor and motor cortex of mice. For those reasons, the behavior task established in this thesis featured delayed lick-responses for the mice to indicate their decisions.

1.3.3 Two-alternative forced choice task

There are different configurations in which sensory stimuli can be presented and decisions can be reported in behavior tasks. The majority of rodent behavior tasks can be described by two main prototypes. In the first type, individuals have to detect and report the mere presence of a sensory cue. In the second type, individuals instead have to discriminate a specific sensory feature between multiple presented options. The first type of task in its simplest form is referred to as a "Go/No-Go" task (Carandini & Churchland, 2013). Here, individuals have to refrain from responding up until they perceived a certain sensory cue, often in the form of the appearance of a stimulus (Meier & Reinagel, 2011), upon which they express a specific behavioral response. This paradigm can also be used to study more complex behaviors, having animals report changes in a specific stimulus feature, such as a change in speed of a visual grating (Orsolic et al., 2021). This is a commonly used task schema, but it comes with drawbacks. On the one hand, responses are linked to the motivation of the subject (Carandini & Churchland, 2013). This makes it hard to distinguish if the absence of a response in a given trial was caused by the individual not perceiving a sensory cue or due to disengagement from the task. On the other hand, choices are intrinsically linked to motor responses, which drive considerable amounts of cortical activity on their own (Musall et al., 2019; Shimaoka et al., 2018), making the distinction of perceptual and movement-related neural responses difficult. This issue can in part be alleviated by giving animals two separate response options: "yes" and "no" (Carandini & Churchland, 2013). This design effectively decouples the indicated decision from the willingness to respond. However, responses still depend on a subjective decision threshold that determines if an individual tends to report "yes" more or less often than others. This can be addressed using signal detection theory (Stanislaw & Todorov, 1999), but it can make it more difficult to compare neural activity relating to perception over individuals.

The second type of behavioral task design, reduces these confounds. In this task design multiple stimuli are presented and an individual has to decide between those options. The more stimuli and response options are presented, the more difficult a task becomes to solve. As such, commonly only two options are presented in rodent studies (Huber et al., 2008; Mayrhofer et al., 2013; Raposo et al., 2012). Since subjects have to choose one stimulus over the other, this type of task is commonly referred to as a two-alternative forced choice (2AFC)

task. By having individuals directly compare stimuli, it is easier to determine how well two stimuli can be distinguished from one another and this can be directly linked to how distinct neuronal responses are to those stimuli. Therefore, a 2AFC task design was best suited for the use in this study to investigate how unisensory and multisensory stimuli are perceived and integrated by mice to form decisions.

1.3.4 Sensory stimuli

Next, we needed to determine the sensory stimuli that are presented in the behavior task. As mentioned previously, the aim of this study is to investigate cortical multisensory integration. For this purpose, widefield calcium-imaging of the dorsal surface of cortex was employed. With this approach, we were limited to studying cortical regions that are optically accessible in mice using available surgical preparations (Guo et al., 2014; see Section 1.4). While the somatosensory- and visual cortical areas can reliably be imaged with this approach, the same is not the case for the auditory cortex. Therefore, in order to to study the relative contributions of primary sensory cortices and related association areas, we were restricted to visual and tactile sensory stimuli. A crucial aspect of the stimuli used in this study was they should not contribute sensory information across the stimulated modalities, potentially convoluting visual and tactile sensory evidence.

Since rodents heavily rely on their whisker system and devote extensive neuronal resources to the interpretation of tactile whisker sensations, many previous studies have investigated this system in detail (Ayaz et al., 2019; Chen et al., 2017; Ferezou et al., 2007; Petersen, 2007). However, usually tactile stimuli are presented by the means of a physical object deflecting the whiskers (Ayaz et al., 2019; Ferezou et al., 2007; Gallero-Salas et al., 2021; Gilad & Helmchen, 2020; Guo et al., 2014; Petersen, 2007). Such physical stimulation apparatuses could have had visible features detectable by the mice, interfering with the stimulation. Therefore, we presented the tactile stimuli in the form of brief airpuffs against the whiskers of mice, similar to previous studies conducted in both rats (Ollerenshaw et al., 2012) and mice (Bernhard et al., 2020). With this it could be ensured that there was no visual component as part of the tactile stimulus presentation.

The presentation of visual stimuli generally has no such constrains, especially when using monitors or similar remote light sources. Conceptually, the closest analog to our tactile airpuffs would be a simple flashing visual stimulus. Such flashing visual cues have successfully been used with rats (Scott et al., 2015). However, prior attempts in our lab both by Bexter (Bexter, 2022) and myself have shown that mice struggled with using such flashing visual cues when performing a behavior task. Alternatively, sinusoidal gratings are commonly used as visual stimuli in mouse behavioral tasks and mice can perform these tasks well (Bexter, 2022; Burgess et al., 2017; Busse et al., 2011; Goldbach et al., 2021). Therefore, we sought to combine the characteristics of a moving grating that mice tend to perform well with in a behavioral task, with the ability to present isolated cues. We realized this by using a single cycle of a drifting sinusoidal grating as an individual cue (for further discussion see section 3.1.2). These cues can be flexibly combined into longer, irregular sequences of cues allowing for more complex tasks designs, such as an evidence accumulation task.

1.3.5 Evidence accumulation task

In our everyday life we often have to make decisions based on noisy, ambiguous sensory information, where the correct choice is not always clear from the current sensory input. Instead, we have to integrate information over time to form an appropriate decision (Brunton et al., 2013; Scott et al., 2015). Such an accumulation of sensory evidence represents a complex cognitive task requiring cortical integration (Pinto et al., 2019). It has been shown that rats and mice are capable of performing evidence accumulation tasks in different sensory modalities (Brunton et al., 2013; Erlich et al., 2015; Morcos & Harvey, 2016; Pinto et al., 2018; Scott et al., 2015). Further, this paradigm also has proven to be suited for the study of multisensory integration in rodents, where animals are capable of improving their choices by integrating visual and auditory information (Pisupati et al., 2021; Raposo et al., 2012; Siemann et al., 2015). In this study, a similar task design was used to investigate visuotactile integration using a 2AFC evidence accumulation task design. Here, we present mice with random sequences of sensory cues on both their left and right side and after a short delay the mice have to indicate the side with the higher number of cues presented by licking the corresponding water-spout. By decoupling the sensory modalities of our visual and tactile cues we were able study mice performing the same behavior tasks, based on either the visual

or the tactile information alone, as well as the combined multisensory cues. In addition, such a task design where mice perceiving visual and tactile information over time, also represents a naturalistic behavior of mice having to navigate narrow tunnels in their natural environments.

To summarize, according to the aspects discussed above, we determined the general features of the required behavior task. For the purpose of subsequent widefield calcium-imaging, we designed a task in which head-fixed mice were presented with random sequences of either visual, tactile or multisensory cues on their left and/or right side. After a delay, mice were presented with two spouts to indicate the side where they perceived the higher number of cues and by licking the corresponding spout received a water reward. This task design allowed us to investigate how mice integrate sensory information over time and study the effects of cross-modal stimulation using widefield calcium-imaging.

1.4 Widefield calcium-imaging

To study how sensory information is integrate by mice to solve the visuotactile evidence accumulation task, we needed to simultaneously record neural activity in many different cortical areas, from primary sensory cortices over association areas to motor cortex. This required homogeneous and stable signals over the entire cortex over the course of the behavioral training. For this purpose, optical imaging of genetically encoded functional indicators was the best suited approach for this study.

There are two main types of genetically encoded functional indicators: genetically encoded voltage indicators (GEVIs) and genetically encoded calcium (Ca^{2+}) indicators (GECIs). These two indicator types both reflect neural activity, but in different ways.

GEVIs are membrane-associated indicators, that change their fluorescence signals in accordance to the instantaneous membrane potential of a cell (Akemann et al., 2012). Therefore, these indicators allow for a very high temporal resolution when measuring neuronal activity. Such indicators have successfully been used in both widefield imaging (Carandini et al., 2015) and even two-photon microscopy (Akemann et al., 2013). However, these indicators give weaker signals with lower signal-to-noise ratios (SNR) compared to many

GECIs. This is in part due to the association of the sensors with the cell membrane, limiting the indicator densities per cell. GEVIs are a powerful tool, particularly for single cell studies, even allowing to study signal propagation within a neuron's dendrites (Akemann et al., 2013). However, for the purpose of comparing neural responses over many brain areas simultaneously, in a behavioral context, GECIs were the better suited indicator.

GECIs are indirect indicators of neural activity. The firing of action potentials of a cell leads to an influx of calcium ions (Ca^{2+}) into a cell's cytoplasm, mediated by voltage-gated Ca^{2+} channels (Chen et al., 2013). Therefore, changes in the intracellular Ca^{2+}-concentration can be used as a measure of neural activity. GECIs respond to these changes in intracellular Ca^{2+}-concentration by changing their conformation, resulting in altered fluorescence intensity. GECIs are therefore indirect indicators of neural activity with slower dynamics than GEVIs. However, this allows for stronger fluorescence signals and higher SNR. This is crucial for the study of neural activity during active behavior.

One of the most commonly used GECIs are indicators of the GCaMP-family (Allen et al., 2017; Andermann et al., 2011; Musall et al., 2019; Wekselblatt et al., 2016). This indicator is composed of three fused functional subunits (Chen et al., 2013): a calcium-modulated protein (CaM), a CaM-interacting peptide (M13; Crivici & Ikura, 1995) and a green fluorescent protein (GFP). Once the intracellular Ca^{2+}-concentration of a cell increases and all four Ca^{2+}-binding sites of CaM are occupied, CaM interacts with M13, in turn resulting in a conformational change of GFP, resulting in an increase in fluorescence (Chen et al., 2013). These changes in fluorescence can then be used to optically measure neural-activity from many distributed regions of the brain.

Indicators of the GCaMP-family (Chen et al., 2013) have been shown to generate strong signals with sufficient SNR to measure neural responses on a single trial level, allowing to analyze responses on the timescale of behavior (Musall et al., 2019; Wekselblatt et al., 2016). This critical advantage made Ca^{2+}-imaging using GCaMP, the best approach to study behavior-related neural activity in the context of this thesis. The emergence of transgenic mouse lines expressing GCaMP and other genetically encoded functional indicators, in combination with advanced optical imaging approaches (Guo et al., 2014; Silasi et al., 2016) now allow to simultaneously record activity over extensive regions of the cortex, over the long periods of behavioral studies (Wekselblatt et al., 2016). Using a well-established

transgenic mouse line expressing GCaMP6s (Wekselblatt et al., 2016), the slow but high affinity version of this generation of GECIs, allowed us to record neural activity from visual- and somatosensory cortex up to the motor and premotor areas of mice performing the visuotactile evidence accumulation task.

1.5 Aims

The human cerebral cortex is the most complex structure known to man. It plays a central role in our higher cognitive functions and is thought to be the neural correlate of our consciousness. Yet, at its essence the cortex serves the purpose of integrating sensory information about the environment and guiding our responses to it. A crucial step in this process is integrating the information from all our senses to construct a unified model of our surrounding. Despite the importance of cortical multisensory integration and its effect on guiding our behavior, many of the mechanisms are still poorly understood. The central aim of this study is to advance our knowledge of how sensory information across modalities is integrated in the mammalian cortex and how this information is used to guide behavioral responses. For this purpose, three objectives will be accomplished:

Objective 1: Establishment of a novel visuotactile evidence accumulation task.

Previous studies in rats using a task design similar to ours, reported robust performance improvements in rat performing multisensory trials (Pisupati et al., 2021; Raposo et al., 2012). The first milestone will be to establish if mice are also able to synergistically used the visual and tactile sensory information to improve their behavioral performance. By focusing on the visual and tactile sensory modalities we are in the best position to study cortical circuits for sensory decision making across modalities, using subsequent widefield calcium-imaging. Further, the evidence accumulation task design allows to compare unisensory and multisensory performance over a wide range of task difficulties.

Objective 2: Characterization of multisensory integration in cortical circuits

In the second part, widefield imaging in transgenic mice expressing the genetically encoded calcium-indicator GCaMP6s will be used to measure activity over the entire dorsal surface of the cortex. Having designed the behavior task to utilize visual and tactile stimuli, we are able

to study sensory responses in both primary sensory cortices as well as association areas, such as the rostrolateral visual area (RL) previously shown to integrate visual and tactile sensory information in both rats and mice (Olcese et al., 2013; Wallace et al., 2004). This allows us to directly compare the relative distribution of multisensory depression and enhancement that can be found in primary sensory cortices and higher association areas during active behavior task performance.

Objective 3: Identification of signatures for sensory evidence and behavioral decisions across modalities

Finally, we investigate how mice integrate sensory information to solve the behavior task. This comprises of two distinct aspects. On the one hand, to solve the evidence accumulation task mice have to integrate sensory information over time to identify the target stimulus side. One the other hand, mice have to make decisions on which side to respond.

To identify cortical circuits that integrate the sensory information, reflecting the target-side, we compare the activity of individual cortical areas depending on the side where the target stimulus was presented. For this purpose, we compute how distinct cortical areas respond depending on the target-side using the area under the receiver-operator characteristic curve (AUC). This allows us to identify cortical areas that accumulate sensory evidence over the course of a trial and investigate if sensory evidence is accumulated separately for the individual modalities or modality-unspecific. By comparing target representations in unisensory and multisensory trial we will also gain a deeper insight into the mechanism that facilitate multisensory perception.

To probe how this sensory information is translated into responses in the decision-making process, we aim to identify cortical areas that reliably reflect the upcoming choices of mice. For this purpose, a recently published choice-decoder will be used (Musall et al., 2023). Subsequently we aim to investigate if there are modality specific differences, as an indication of distinct circuits, forming either visually- or tactilely driven decisions. This approach will allow us to outline a likely cortical circuit that mice use to integrate visual and tactile information to guide their behavioral responses. With this we can address the central question at which stage of the cortical processing, information is integrated across sensory modalities and how this information is used to guide behavior.

2 Results

2.1 Mice performing a visuotactile evidence accumulation task

2.1.1 Behavioral task design

The aim of this thesis was to study how mice integrate information from two different sensory modalities to guide their behavioral responses. For this purpose, we designed a novel visuotactile evidence accumulation task. In this task mice were presented either with visual, tactile or the combination of visual and tactile information in the multisensory condition. In all three stimulus conditions mice had the same objective: to integrate sensory evidence during the stimulus period and report which side they perceived the higher number of sensory cues on. Since the general task design was the same in all stimulus conditions, this allowed us to identify and distinguish modality specific mechanism of sensory perception and general principles of the decision-making process, found in all stimulus conditions.

A schematic of the setup and the time-course off an example multisensory trial are shown in Figure 4A. In this multisensory task subjects are confronted with three trials types: visual, tactile or multisensory. In the visual condition mice were presented with cues in the form of individual cycles of a moving grating, displayed on monitors to the left and right of the mice. In the tactile condition subjects were presented with short airpuffs directed against the whiskers. Finally, in the multisensory condition both visual and tactile cues are presented simultaneously and temporally aligned. In each trial mice were presented with sequences of 0 to 6 cues at their left and/or right side during the 3 s stimulus period. After a short delay period of 0.5 s mice were presented with two water spouts. Here, the subjects had to indicate the side where they perceived the higher number of sensory cues, by licking the corresponding spout to receive a small water reward (typically 2µl).

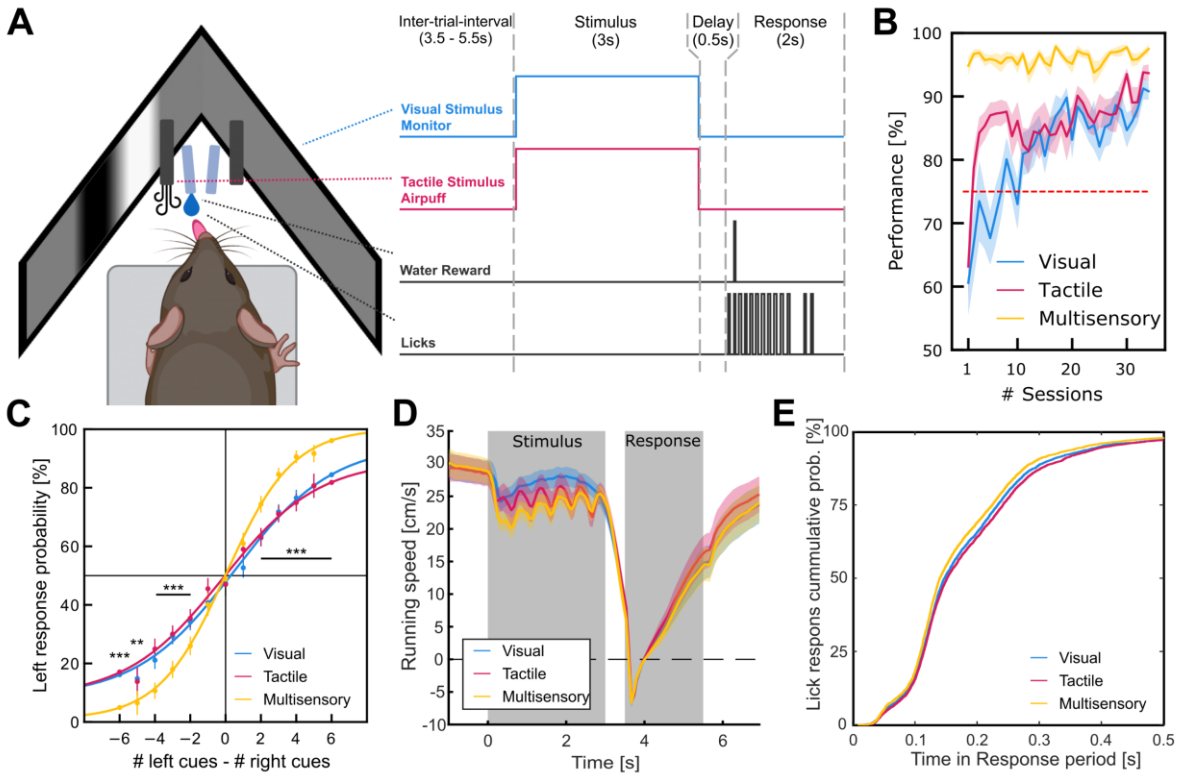

Figure 4: Mice performing the visuotactile evidence accumulation task.
(A) Left: Schematic setup configuration. The mouse is placed on a running wheel with two monitors in front for visual stimulation, airpuff ports for tactile stimulation and movable water spouts for response detection and reward delivery. Right: Schematic of an example correct multisensory trial. (B) Individual learning curves for visual, tactile and multisensory trial performance, aligned to first session, respectively. Error indicates standard error of mean (s.e.m.) over individuals. (C) Psychometric curves for visual, tactile and multisensory task performances, depending on the difference in number of left and right cues. Significance indicators for multisensory compared to best unisensory trial performance (Binomial test, n=67,834 trials from 4 mice). (D) Running speed of mice depending on trial modality. Error shading indicates 95% confidence interval (CI). (E) Cumulative distribution of initial lick response depending on stimulus condition for correct response trials.

After this 2 s response period, an inter-trial-interval (ITI) of 3.5s is employed before the next stimulus onset. However, if a mouse did not respond in a given trial the next trial was delayed by 1 s and if the mouse responded incorrectly, the next trial was delayed by 2 s. This delay-based reinforcement was included for two reasons. First, to give additional positive feedback by presenting the next trial and therefore the next opportunity for a reward sooner. Second, to discourage potential behavioral strategies that maximize trial counts by guessing instead of maximizing fraction of correct trials. Based on this task design mice were trained one modality at a time to keep the task as simple as possible in the early training stages.

2.1.2 Learning

Mice were trained to perform the behavior task first in the visual detection condition. The detection condition represents the easiest task condition with the maximum number of target cues, in the absence of any distractors. Here, mice displayed performances above 75% of correct responses (the criterion for a successfully learned behavioral task) on average after 7 (1 - 9) sessions (median and range, from 12 mice, excluding one outlier requiring 18 session; Figure 4B). After this, mice were trained on the tactile detection condition. Here, mice displayed performances above 75% already after 2 (1 - 5) session (median and range; from 12 mice, excluding one outlier requiring 19 sessions) after introduction. Once the mice were able to perform both the visual and the tactile detection condition, multisensory trials were introduced in addition to the already learned visual and tactile trials. Mice were able to perform this condition above criterion instantly from the first session on; session 1 (1 - 1) (median and range; from 12 mice). Strikingly, with a performance of 94.9% ± 4.9% (mean ± s.d., from 12 mice) mice displayed higher performances in the multisensory condition, compared to even the preferred unisensory condition with 88.7% ± 7.6% (mean ± s.d., from 12 mice), already in the very first session ($P < 0.05$; t-test; from 12 mice). This showed that mice were capable of achieving high performances in all three stimulus conditions of this visuotactile detection task. Furthermore, it demonstrated that subjects were capable of using information of both modalities presented simultaneously to improve their success-rate even in these simplest trial conditions.

2.1.3 Improved discrimination performance in multisensory trials

Once the mice had learned to perform the detection component of the task in all stimulus conditions, discrimination trials were introduced by presenting a random number of target cues and adding distractor cues. In these trials, mice had to indicate the side with the higher number of cues presented over the stimulus period. In all three stimulus conditions, the performances depending on the difference in target and distractor cue number, were well described by sigmoid psychometric curves (number of trials: 67,834 from 4 mice). Furthermore, subjects displayed very similar performances in both unisensory conditions (Figure 4C). In contrast, mice were able to achieve significantly higher performances in all except the most difficult multisensory discrimination trials ($P < 0.001$, for all |#left cues -

#right cues| > 1; with the exception of the condition #left cues - #right cues = -5, with $P < 0.01$, n=67,834 trials from 4 mice). This demonstrates that the mice are able to use visual and tactile information synergistically to make more accurate decisions, resulting in higher performances in multisensory trials over a broad range of task difficulties.

Having seen that mice display more successful behavioral responses in multisensory trials, we next investigated if mice displayed any further differences in their behavior, while performing this task in different stimulus conditions. For this purpose, we compared the running and licking behavior in the different stimulus conditions. These represent the clearest behavioral readouts of the subject's behavior in this task.

2.1.4 Differences in running behavior depending on stimulus condition

Having seen that the mice can achieve different performance levels in uni- and multisensory trials we investigated if mice also displayed modality-specific variations in other behavioral features. For this purpose, first the running speed of the mice was compared over modalities (Figure 4D). Here, a very stereotypical running pattern could be found that was tightly linked to the time-course of trials. In general, mice displayed a high running speed in the inter-trial-interval and a slight reduction in running speed after stimulus onset. Then, during the delay mice drastically reduced their speed and came to a complete halt, and even seemed to move backwards in the attempt of halting on the wheel, within the first 0.5s of the response period. After this, mice started running again.

Clear differences in the running behavior could only be found during the stimulus presentation. Here, in both tactile and multisensory trials clear dips in running speed every 0.5 s were found, corresponding to the individual airpuffs presented. Further, mice displayed the highest average speed during the stimulus period in visual, compared to tactile trials ($P < 0.01$, 164 sessions from 4 mice; Wilcoxon rank-sum test) and multisensory trials ($P < 0.001$, 159 sessions from 4 mice; Wilcoxon rank-sum test). In addition, significantly lower speeds were also found in multisensory trials compared to tactile trials ($P < 0.001$, 159 sessions from 4 mice; Wilcoxon rank-sum test). Strikingly, towards the end of the stimulus period the running speeds in all stimulus conditions converged again.

In summary, subjects displayed strong modulation in their running behavior over the time-course of a trial, with clear modality-specific differences in their running behavior during the stimulus presentation. These differences between stimulus conditions have the potential to cause differences in the cortical activity (Shimaoka 2018) and have to be considered when interpreting the cortical activity of the mice in response to the stimuli.

2.1.5 Faster licking responses in multisensory trials

Having seen that the mice display differences in their running behavior, we next investigated their response behavior. For this purpose, we compared the time of the first lick response in a given trial between modalities (Figure 4E). Here, in all stimulus conditions the majority (>95%) of lick responses were given within the first 0.5 s of the response period. Accordingly, in the subsequent widefield Ca^{2+}-imaging we will focus the analysis of response-related activity on these first 0.5 s of the response period. Only a small difference in the response times of 4 ms (median difference; $P < 0.05$, n=11,705 trials from 4 mice; Wilcoxon rank-sum test) was found between visual and tactile trials. In contrast, median lick responses in multisensory trials were 8ms faster compared to visual trials ($P < 10^{-5}$, n=11,738 trials from 4 mice) and 12ms faster compared to tactile trials ($P < 10^{-12}$, 12,161 trials from 4 mice; Wilcoxon rank-sum test). While, these differences in the millisecond range likely would not affect our recordings of cortical activity using widefield Ca^{2+}-imaging, they might be indicative of more general features, such as the confidence of the mice (Pisupati et al., 2021; for discussion see section 3.1.4). Nonetheless, this shows that mice displayed differences in multiple behavioral features, depending on the stimulus condition of the task, which have to be considered for the interpretation of the cortical activity.

To summarize, we were able to establish a novel visuotactile evidence accumulation task, in which mice were able to perform in two distinct unisensory conditions, as well as a multisensory trial condition. We found that mice consistently displayed higher performances in the multisensory compared to either of the unisensory trial conditions. Also, differences in further behavioral parameters such, as their running and licking behavior were seen. This behavioral task now gave us the opportunity to study the cortical activity of mice performing the task in the individual stimulus conditions to identify modality-specific sensory responses

and decision-making processes, reaching beyond the confounds of individual sensory modalities to guiding behavior.

2.2 Widefield Ca^{2+}-imaging to measure cortical activity of mice performing the multisensory behavioral tasks

The establishment of a multisensory behavior task for head-fixed mice, now allowed us to measure and compare the cortical activity of mice, by employing widefield Ca^{2+}-imaging. For this purpose, transgenic mice were used that express the genetically encoded fluorescence Ca^{2+}-indicator GCaMP6s in cortical pyramidal neurons (Wekselblatt et al., 2016). Using a clear-skull preparation, similar to previous studies (Guo et al., 2014; Musall et al., 2019; Zatka-Haas et al., 2021), optical access to the entire dorsal surface of the cortex was achieved. This preparation allowed to chronically record cortical activity of subjects, performing the multisensory behavior task. This allowed us to records an extensive dataset of ~1000 hours of continuous high-resolution imaging data. To enable the storage and analysis of these vast amounts of data we use a dimensionality-reduction approach (Musall et al., 2019; see section 5.6.1 for details). All analysis presented is based on this preprocessed data. To compare activity over sessions and individuals, recordings were aligned to the Allen Common Coordinate Framework v3 (Allen CCF; Wang et al., 2020) based on anatomical and functional information (see Figure 13).

First, activity in simple detection trials was compared between individual stimulus conditions, to investigate the general patterns of cortical activity in mice performing the task in the individual stimulus conditions (Figure 5). Here, in visual trials (1,031 correct detection trials from 4 mice) activity was primarily found in primary visual cortex (V1) and secondary visual cortices (Andermann et al., 2011; Marshel et al., 2011; Wang & Burkhalter, 2007), as well as the anterior cingulate cortex (ACC, corresponding to posterior ACAd in Figure 2; Murphy et al., 2016) and the fontal pole of the cortex of the mice during the stimulus period.

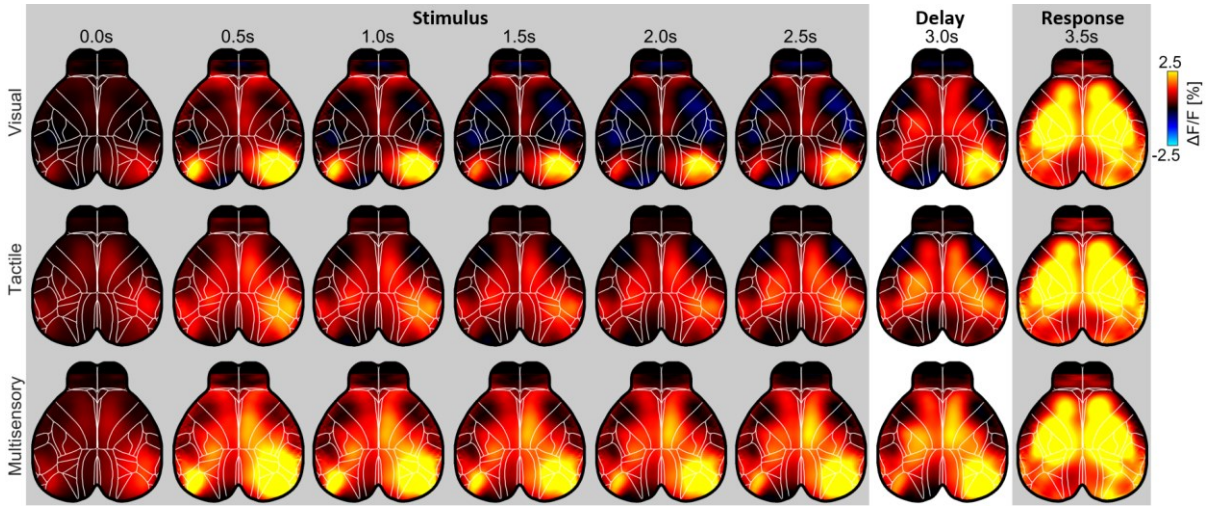

Figure 5: Similar patterns of cortical activity in mice performing visual, tactile and multisensory detection trials.
Average cortical activity in mice performing correct visual (top), tactile (middle) and multisensory (bottom) detection trials. Activity was averaged over left and right trials by inverting hemispheres in right trial, so that the right hemisphere represents activity contralateral to target stimuli presented on the left. Recordings are aligned to the Allen CCF, with dorsal region outlines superimposed. Shadings indicate stimulus, delay and response period. Individual maps represent 0.5s time-bins of cortical activity, starting with the stimulus onset.

In tactile trials (1,159 correct detection trials from 4 mice) activity was found in the barrel cortex (wS1; corresponding to SSP-bfd in Figure 2), surrounding areas including area RL and the medial motor cortex (MM, corresponding to the medial region of MOs in Figure 2; Chen et al., 2017; Guo et al., 2014), also commonly referred to as primary whisker motor cortex (Ferezou et al., 2007) of the mouse. In comparison, in multisensory trials (961 correct detection trials from 4 mice) broader activation of the dorsal cortex was found, but particularly in the occipital and parietal cortex and MM. During the delay period, activity similar to the modality-specific stimulus responses could still be seen. In addition, increased activity in the parietal and medial motor cortex was found in all stimulus conditions. Finally, during the response period, high activity of the entire dorsal cortex was seen in all stimulus conditions, with the strongest activity in the premotor- and motor cortex and parts of the somatosensory cortex.

In summary, during the stimulus period modality-specific, but partially overlapping patterns of cortical activity could be found. Towards the delay and especially the response period, cortical activity became more similar over modalities, suggesting a common cortical circuitry

by which the individual modality specific signals converge to eventually produce the same motor-response. To investigate this overlap in responses over the time-course of a trial, next we analyzed which cortical regions consistently displayed significant activation in all stimulus conditions to identify likely regions for multisensory integration.

2.3 Overlapping visual and tactile responses in PPC and medial frontal cortex.

Having seen that mice display similar patterns of cortical activity in all stimulus conditions over the course of an average trial, we then characterized which cortical regions displayed responses in both visual and tactile trials to identify regions with a high potential for multisensory integration. To study this overlap in cortical responses, we expressed cortical activity during correct detection trials in the form of a z-score (relative to the standard-deviation over sessions and mice), to focus on the most consistent cortical responses in the individual stimulus conditions (Figure 6A).

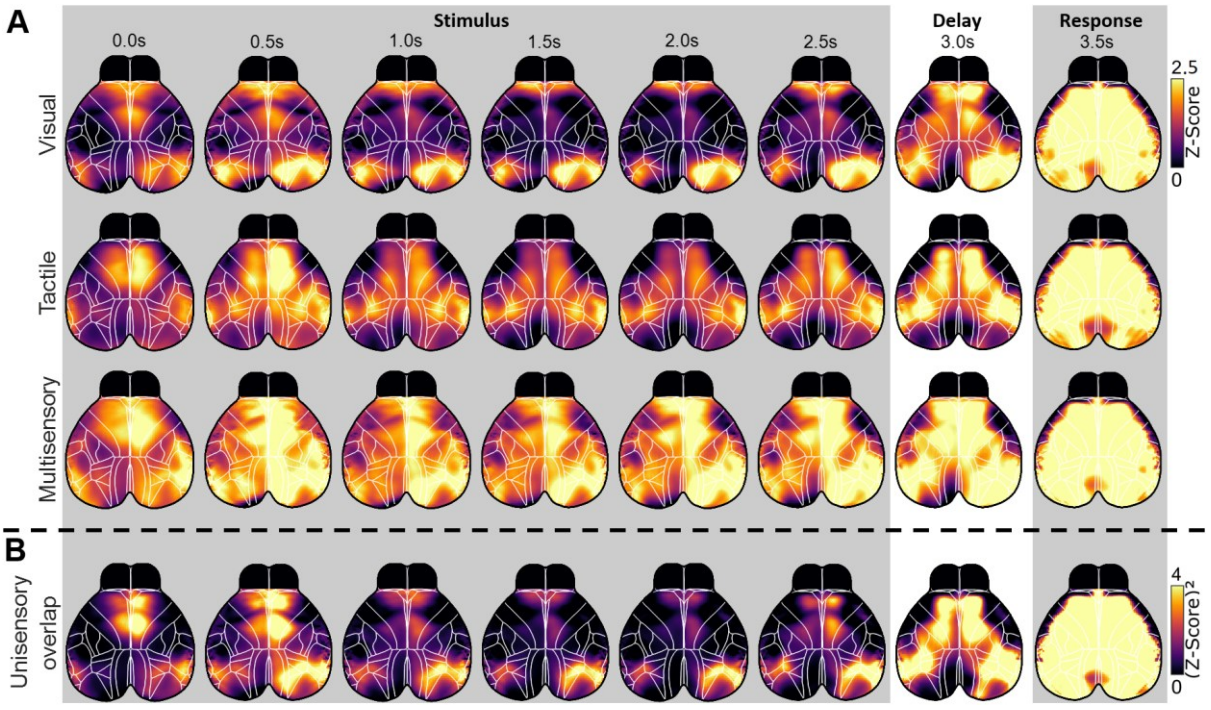

Figure 6: Overlapping visual and tactile responses in PPC and medial frontal cortex.
(A) Z-Score of average cortical activity in mice performing correct visual (top), tactile (middle), multisensory (bottom) detection trials, with stimuli on the left side. Same as in Figure 5, maps represent 0.5s time-bins starting at the stimulus onset. (B) Overlap in visual and tactile responses in A, expressed as product of their Z-Score maps.

Here, during the stimulus and delay period in visual trials (1,031 correct detection trials from 4 mice) three main clusters of activity became visible: a broad activation of the occipital lope including V1, secondary visual areas, as well as the posterior retrosplenial cortex (RSC, compare Figure 2), then activity in ACC and MM and finally a region in the anteromedial motor cortex (AMM) of the mice. Generally, activity was found bilaterally, with stronger and more reliable responses on the hemisphere contralateral to the stimulus presentation. In tactile trials (1,159 correct detection trials from 4 mice), over the entire stimulus and delay periods broad bilateral activation of the parietal cortex and the entire medial frontal cortex (mFC; including AMM and MM) was observed. Here, particularly activity in AMM and MM displayed stronger contralateral responses. In multisensory trials (961 correct detection trials from 4 mice), activity was generally increased, compared to unisensory conditions. Here, the same regions responding in the individual unisensory conditions displayed strong and consistent responses. These included primary- and secondary sensory cortices, as well as RSC and mFC. During the response period, broad activity of almost the entire dorsal cortex was found in all stimulus conditions. To distinguish which components of these multisensory trial responses, correspond to modality-specific responses and which to general activity, found in both visual and tactile trial, we investigated the overlap in unisensory trial responses.

To study the overlap of activity in the unisensory trials, we compute the product of the z-scored activity in visual and tactile trials. This approach highlighted regions in the cortex that reliably responded in both visual and tactile trials (Figure 6B). Overlapping activity in the stimulus period was consistently found in area RL (compare VISrl in Figure 2) between the two primary sensory cortices, as well as mFC. The activity was highest in the time frame from 0.5 s to 1 s after stimulation onset (second time-bin in Figure 6B), corresponding to the second stimulus cue presented, matching the stronger stimulus responses in the individual stimulus condition trials. The overlapping activity increased during the delay period compared to the stimulus, on both the ispi- and contralateral hemisphere with strong bilateral activity in mFC, PPC and portions of the somatosensory cortex. During the response period the entire dorsal cortex displayed strong activation in all stimulus conditions, which was also reflected in the unisensory response overlap.

In summary, over the time-course of a trial overlapping responses in visual and tactile trial were found especially in RL and mFC. Here RL displayed strong overlapping responses over the entire trial, while AMM and MM displayed the strongest responses at the beginning of the stimulus presentation and again at the end of the stimulus- and during the delay period. In the initial response period, broad activation over the entire dorsal cortex was found in every stimulus condition. This suggest, that the regions RL and mFC represent sensory information across modalities, likely in a shared neural circuitry. Toward the delay and response period, cortical activity became more similar, indicating a convergence of the processing streams, as part of the decision-making process and the response preparation.

2.4 Multisensory enhancement in RL and MM

Having seen that there is potential for multisensory integration, particularly in RL and mFC of mice, next we investigated if these regions in particular, displayed higher activity when mice were performing multisensory trials. For this purpose, we first computed the maximum response of the individual unisensory trials (Figure 7A). Here, we found the occipital lobe including V1, secondary visual cortices and RSC, as well as the region at the anterior pole of the cortex, displayed higher activity in visual compared to tactile trials. In comparison, the somatosensory cortex and a large region of the motor cortex, were more active in tactile trials. These patterns of preferred sensory modality were very consistent over the entire time-course of a trial. Next, we compared activity in multisensory trials to these maximum unisensory responses (Figure 7B). This way the multisensory responses were compared to the individually preferred unisensory response for each region separately. Here, it was seen that over the course of the stimulus period, activity in multisensory trials increasingly exceeded the preferred unisensory responses. This was in fact most evident in area RL and MM, as our previous results suggested. This excess activity was strongest on the hemisphere contralateral to the target stimulus, but was also seen in part on the ipsilateral hemisphere. During the delay period, activity in multisensory trials exceeded that of the unisensory trials, however to a lesser extent than at the end of the stimulus period. Finally, during the response period the activity of the unisensory trials exceeded that of the multisensory trials again. This suggests that differences in unisensory and multisensory trials could especially be found during the stimulus and the delay period of the task and most prominently in area RL and MM.

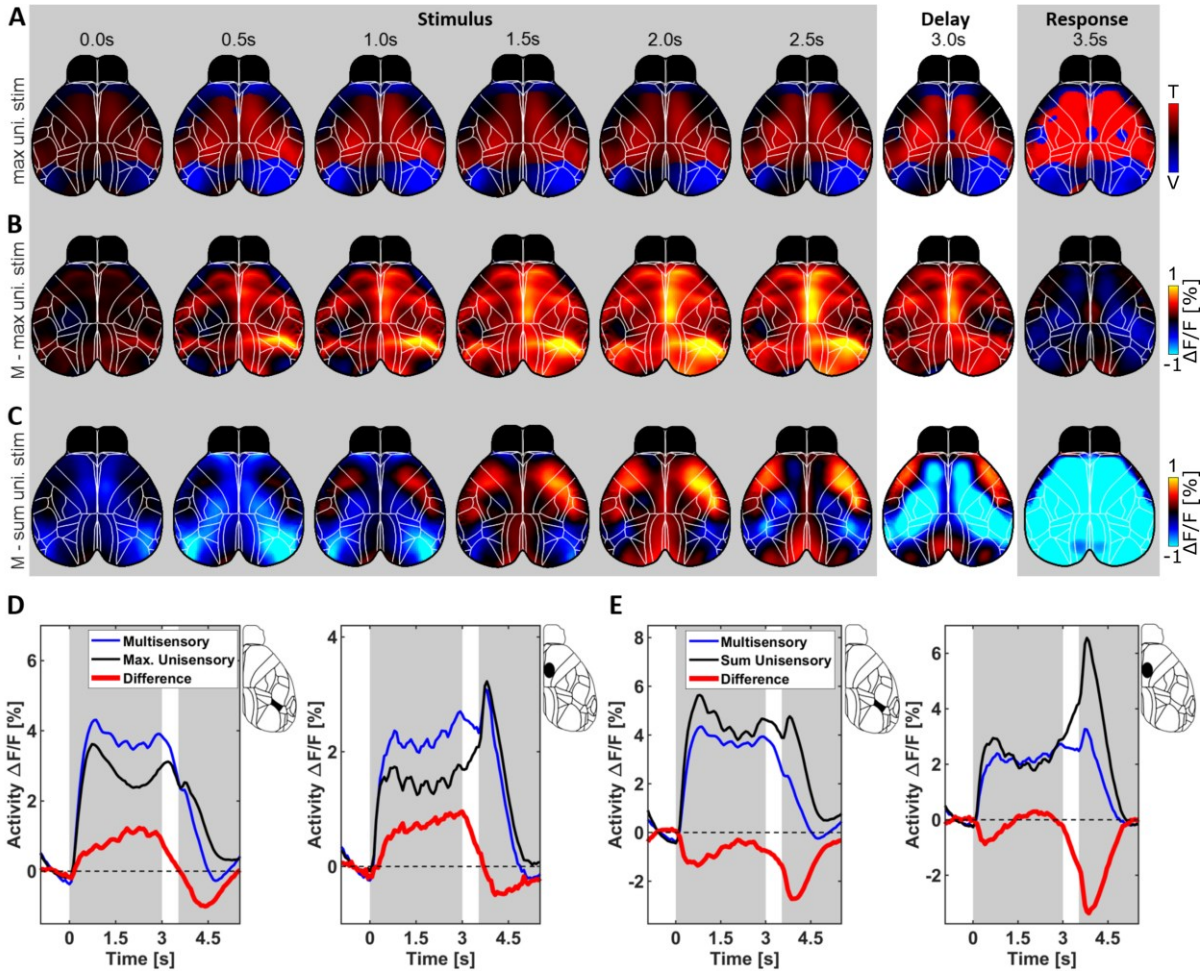

Figure 7: Multisensory enhancement in RL and MM.
(A) Maps of preferred modality in correct unisensory detection trials (see Figure 5). Blue: higher response in visual (V) trials. Red: Higher responses in Tactile (T) trials. Maps scaled between 0% and 2.5% ΔF/F. (B) Difference of Multisensory (M) trial activity and maximum unisensory response (shown in A). (C) Difference of Multisensory activity and the sum of visual and tactile trial activity. Maximum is computed pixel-wise for individual frames over time. (D) Multisensory and maximum unisensory trial activity and their difference of RL and MM; corresponding to B. (E) Same as D, for comparison of multisensory and summed unisensory activity; corresponding to maps in C.

To study these responses in more detail, we investigated the activity in RL and MM on the side contralateral to the stimulus over the course of a trial (Figure 7D). Here, in both regions activity sharply rose from the stimulus onset on in both the preferred unisensory and the multisensory trials. Starting with the stimulus onset, excess multisensory activity was found both in RL and MM and continued to steadily increase over the entire stimulus period. In both uni- and multisensory trials activity in RL was decreasing again after the stimulus, with activity in multisensory trial falling faster. Accordingly, the difference in activity between both trial

types was decreasing as well. Then, at the end of the delay period activity in RL reached the same level in uni- and multisensory trial. A similar trend became apparent for MM, where the difference in activity also decreased to almost the same level in uni- and multisensory trials before the response. However, whereas in multisensory trials this was the result of a decrease in activity, in unisensory trials activity was rising throughout the delay period. This suggests, that activity is not merely decaying during the delay period, as might be expected due to the nature of the calcium-indicator (Chen et al., 2013), but instead seems to approach a common level of activity. This would argue for the hypothesis of converging sensory information on a common circuitry to prepare the responses of the mice. During the response period, activity in RL remained elevated longer in unisensory trials and only converges again toward the end of the response period. In MM activity peaks within the first 500ms of the response period, around the time where the mice reported the majority of their decisions (see Figure 4E). After that, the activity sharply decreased again towards baseline levels. The in comparison decreased multisensory activity seems to result from a delay in the downward slope of the response. To summarize, over the entire stimulus and delay period of the task RL and MM displayed consistently higher multisensory responses compared to the preferred unisensory responses, indicating simultaneous processing of both sensory modalities in multisensory trials, as suggested by the overlapping unisensory responses (see Figure 6).

However, notably this comparison seems to highlight regions in which the preference switches from one unisensory modality to another (Figure 7A, top), particularly between the occipital- and the parietal cortex and again between the motor cortex and the anterior pole of the cortex.

Having seen that there is in fact increased activity in response to multisensory stimuli, compared to the preferred unisensory responses, we next investigated if this activity could be described as a simple linear summation of the individual unisensory responses, potentially indicating spatially overlapping, but separate neuronal populations responding. For this purpose, the activity in multisensory trials was compared with the sum of the unisensory trial responses (Figure 7C and E). Here, primarily reduced responses were found initially. The strongest reduction in activity was present from stimulus onset on in V1 and secondary visual cortices, as well as region of the wS1 and MM. Strikingly, these were the same regions that displayed the highest stimulus responses (Figure 5 and Figure 6). Over the course of the

stimulus period, the difference in activity of multisensory and summed unisensory trials became smaller. Toward the end of the stimulus period regions in the lateral motor- and parietal cortex displayed higher activity compared to the summed unisensory responses. However, this was driven by reduced activity in the unisensory trials compared to baseline, rather than increased activity in the multisensory trials (compare Figure 5). Toward the delay and especially during the response period, almost the entire cortex displayed a very strong reduction in activity compared to the sum of the unisensory responses. This would suggest that areas responsive to the stimuli, primarily display subadditive multisensory responses. However, especially the strong relative reduction in activity during the response, highlighted a crucial shortcoming of this approach of comparing sensory responses. Here, the average trial responses were treated as approximations for sensory evoked responses. This assumption is flawed. Even though this becomes most evident during the response period, during the entire trial, cortical activity is influence by aspects other than the sensory stimuli presented (for discussion see section 3.2.3). To study multisensory responses in the context of a behavioral task, we needed to isolate the individual sensory responses from any other cooccurring task- or behavior-related activity. To achieve this, a linear encoding model was used that allowed us to decompose cortical activity into individual components relating to specific task- or behavioral events (Musall et al., 2019; Orsolic et al., 2021).

2.5 Linear encoding model isolates sensory responses.

The previous section indicated that in many regions of the cortex activity in multisensory trials exceed that found even in the preferred unisensory trials, but did not reach the level of cortical activity that would be predicted by a linear summation of the average visual- and tactile trial activity. However, these comparisons likely suffered from the presence of non-stimulus related cortical activity. This emphasized the need to account for components in the cortical activity that do not represent direct sensory responses. Therefore, a linear encoding model was used as an approach to decompose cortical activity into components that relate to specific task- or behavioral events.

The objective of the encoding model is to produce a prediction of cortical activity at any point in time, using information about the task that the mouse was performing, as well as

information about the behavior of the mouse itself. The aim of this model was to determine if the cortical activity in multisensory trials could be explained as a linear summation of the individual visual and tactile sensory responses, together with the non-sensory related information about the task. The alternative hypothesis to this was that there is predictive power in the cooccurrence of the visual and tactile stimuli; a signature of non-linear multisensory integration in the cortex.

To model cortical activity over time, relating to specific task- and behavioral events, each type of event was modeled as a sequence of average cortical activity patterns depending on the time relative to the corresponding event. In the case of a sensory stimulus cue, individual regressors were used, one for each timepoint in the imaging data relative to the onset of the sensory cue (Figure 8A). Here, each regressor only applied to a single frame relative to the time of the event, but repeatedly for each occurrence of this event in the session; in this case each stimulus presentation. This allowed us to model the average cortical response over time, aligned to the stimulus onset. The same principle applied to behavioral events, e.g. in the form of a lick by the mouse. Since this event is the result of a movement of the animal, which is accompanied by preparatory activity (Li et al., 2015), activity related to this behavior can be expected both preceding and following the movement event (Musall et al., 2019). Therefore, regressors are designed that apply to frames both before and after the occurrence of these events. Finally, analog signals such as the current running speed of the mice were used as well to scale an average pattern of cortical activity at the current time. With these building blocks, a design matrix was created based on all the information used to predict the cortical activity of mice at any point in time during a given session.

To address the question if multisensory stimulus responses could be modeled as a linear sum of the visual and the tactile stimulus responses, the model was designed to use the same regressors to describe the visual stimulus response in both visual and multisensory trials and the same tactile response regressors in tactile and multisensory trials (Figure 8B).

As a result, activity related to the visual stimulus, had to be represented the same way in unisensory and multisensory trials. In turn, multisensory stimuli were represented as the linear summation of both unisensory responses. By then adding an additional multisensory regressor whenever visual and tactile stimuli were presented simultaneously, any remaining sensory activity that deviated from the sum of unisensory responses and could not be

accounted for by other task or behavioral parameters, therefore would have to be modeled using this additional multisensory regressor.

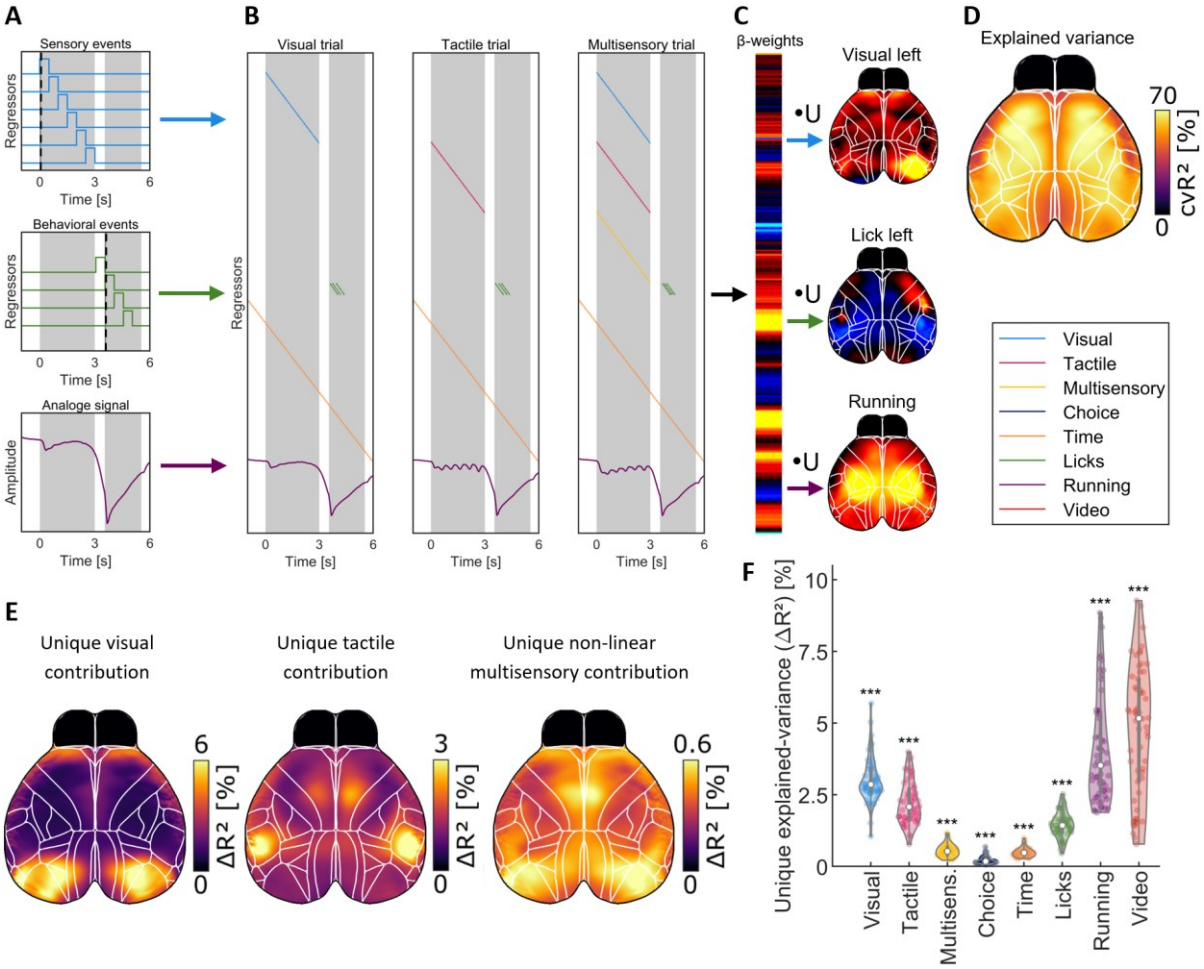

Figure 8: Linear encoding model to isolate sensory responses.
(A) Schematic visualization of sensory and behavior regressors, corresponding to events, symbolized by dashed line. Shading indicates stimulus and response periods, analogous to Figure 5. (B) Schamatic for an example visual (left), tactile (middle) and multisensory (right) trial, as part of the full model design-matrix. (C) example β-weights, here presented corresponing to the first spatial dimension of the imaging data. By convolution β-weights with spatial components (U), weights are projected into original pixel-space for subsequent analysis and visualization. (D) Cross-validated explained variance over trials, averaged over sessions and mice. (E) Maps of unique explained variance for visual (left), tactile (middle) and non-linear multisensory regressors (right), indicating regions in which model predictions are improved by information from the corresponding regressors. (F) Unique explained variance for main groups of regressors, averaged over the entire optically accessable dorsal cortex, and compared over mice and sessions. Each regressor group presented, reliably contributed significant information to model prediction ($P < 10^{-9}$; Wilcoxon sign-rank test).

This can be seen in the schematic design matrices for the example trials of each stimulus condition (Figure 8B). Here, in visual trials only the sensory regressors relating to the visual

stimulus are active. The same applies to tactile regressors in tactile trials. Then, in multisensory trials the same visual and the tactile sensory regressors which are used to model the stimulus responses in the unisensory trials are active simultaneously. In addition, a third sensory regressor is used which is only active when both visual and tactile cues are presented at the same time. This additional multisensory regressor cannot contribute additional information to the prediction of cortical activity if the sum of the visual and tactile stimulus responses is sufficient to describe the multisensory stimulus responses, since this activity is already explained by the other sensory regressors. However, if there are non-linear interactions between the unisensory stimulus responses, these would be described best by this additional multisensory regressor.

Up to this point, cortical activity was studied in simple detection conditions, due to the homogeneity of stimulus presentations, simplifying the interpretation of neuronal responses. However, to be able to separate sensory responses from other task-aligned activity, we needed to decouple these responses from the general time-course of a trial as much as possible. This was achieved by modeling the neuronal responses to the presentation of individual sensory cues, that could occur at different times during the stimulus period. Therefore, the model was trained on both detection and discrimination trial to also investigate sensory responses related to the presence of individual sensory cues at different time points in a trial.

Using this general design of modeling individual sensory-, task- and behavioral events, we trained the model to best predict the imaging data using ridge-regression. However, instead of modeling the individual pixels of the raw imaging data, we predicted the temporal components of the imaging data, containing the same information (after convolution with the corresponding spatial components) in a lower dimensional space, to reduce the computational cost and prevent overfitting (Musall et al., 2019). As a result of the training β-weights were obtained, that link how the presence of each regressor translates to activity of the spatial components of the imaging data, at the respective frame (Figure 8C, left). These weights could then be convolved with the spatial components of the imaging data to project the weights into a cortical map. Exemplary β-weight maps for a visual stimulus presented on the animals' left side, a lick on the left spout and running were obtained and can be seen in

Figure 8C (right). These maps display signals in areas that match known functions of these regions (Wang et al., 2020).

To evaluate how reliable this model could predict cortical activity in mice performing the behavioral task a tenfold cross-validated proportion of explained variance (cvR^2) was computed (Figure 8D). This revealed that the model could explain 55.03% ± 0.74% (mean ± s.e.m., n=58 sessions) of the single trial variability over the entire dorsal cortex. This showed that the model design is well suited to describe the cortical activity of mice performing the multisensory evidence accumulation task. This now allowed us to study if the encoding model required non-linear multisensory responses to improve the predictions of cortical activity in multisensory trials.

2.5.1 Non-linear multisensory activity

To answer the question if there are non-linear multisensory responses in the mice performing the behavior task, we evaluated if the predictions of cortical activity by the encoding model could be improved by including an additional multisensory regressor. To test to which extent each regressor contributes to the model predictions the unique explained variance (ΔR^2) for selected groups of regressors was determined. This was achieved by shuffling every regressor over time, for example related to the visual stimulus, to destroy any correlation with the actual stimulus presentations. This reduced model was then trained again to predict the cortical activity as well as possible. By comparing the explained variance (cvR^2) of the full model with that of the reduce model, any activity that could solely be explained by the visual stimulus was revealed (Figure 8E, left). In this case, activity that is uniquely explained by the visual regressors was found in the occipital cortex of the mice, particularly in V1, secondary visual cortices and the posterior RSC, as well as in the anterior pole region of the cortex. Similarly, unique explained variance for the tactile stimulus regressors was found in wS1 and MM (Figure 8E, middle). Finally, we found that the multisensory interaction regressors in fact also had unique information that allowed to explain cortical activity, which was not explained by the visual and tactile stimulus regressors or any other task- or behavior regressor (Figure 8E, right). However, in contrast to the visual and tactile stimulus regressors, the unique explained variance of the multisensory regressors was found broadly distributed over the

cortex. However, the highest contributions of the non-linear multisensory information were found in V1, RL and MM of the mice.

To compare how reliably groups of regressors contributed unique information to the model predictions, we computed the ΔR^2 over trials within individual sessions and then compared those over sessions and mice (Figure 8F; 58 sessions from 4 mice). This revealed that the behavior of mice as well as non-stimulus-related task information explain a high proportion of the variance in the cortical activity. Particularly videography, capturing movements of the individuals (excluding running and licking) with $\Delta R^2 = 4.73\% \pm 0.30\%$ (mean ± s.e.m., n=58 sessions; $P < 10^{-9}$, Wilcoxon sign-rank test) and running speed with $\Delta R^2 = 4.07\% \pm 0.25\%$ (mean ± s.e.m., n=58 sessions; $P < 10^{-9}$, Wilcoxon sign-rank test) contributed the highest fraction of unique information to the model in order to explain the cortical activity. This highlights the necessity to account for the behavior of the subjects when studying cortical activity and is in line with previous the reports by Musall et al., 2019 and Orsolic et al., 2021.

While explaining less activity than videography and running, information about the lick-responses also provided significant unique information (Figure 8Figure 6F). Here, regressors relating to the lick-responses contributed $\Delta R^2 = 1.46\% \pm 0.06\%$ (mean ± s.e.m., n=58 sessions; $P < 10^{-9}$, Wilcoxon sign-rank test) and the choices expressed by these licks, modelled over the entire trial duration contributed $\Delta R^2 = 0.22\% \pm 0.02\%$ (mean ± s.e.m., n=58 sessions; $P < 10^{-9}$, Wilcoxon sign-rank test). Furthermore, the time regressors, predicting a general task-aligned time-course of cortical activity in all trial types, also contributed further unique information that is not explained by other task- and behavior related features with $\Delta R^2 = 0.49\% \pm 0.02\%$ (mean ± s.e.m., n=58 sessions; $P < 10^{-9}$, Wilcoxon sign-rank test). With this we could now account for the contribution of behavior- and non-stimulus-related task information to the cortical activity. This allowed us to address the question if stimulus responses in multisensory trials could be described as a linear integration of visual and tactile sensory responses.

We found that information about the visual stimulus contributed $\Delta R^2 = 3.01\% \pm 0.10$ (mean ± s.e.m., n=58 sessions; $P < 10^{-9}$, Wilcoxon sign-rank test) to the model predictions. Similarly, tactile stimulus information provided $\Delta R^2 = 2.22\% \pm 0.10$ (mean ± s.e.m., n=58 sessions; $P < 10^{-9}$, Wilcoxon sign-rank test), averaged over the entire dorsal cortex. In comparison to the visual and tactile regressors, multisensory regressors provided less but still

highly significant unique information to the model predictions with $\Delta R^2 = 0.56\% \pm 0.03$ (mean ± s.e.m., n=58 sessions; $P < 10^{-9}$, Wilcoxon sign-rank test). However, it has to be considered here, that this information does not represent stimulus responses themselves, but rather deviations from an assumed summation off visual and tactile responses.

This finding demonstrates that there are in fact non-linear multisensory stimulus responses in mice performing the visuotactile evidence accumulation task, which cannot be explained by other factors relating to the task or the behavior of the mice. To investigate how this non-linear multisensory information modulates cortical activity, next we examined the individual β-weight kernels of the sensory regressors used to model the cortical activity.

2.6 Superadditive multisensory activity revealed using a linear encoding model

For our model we chose to represent the first sensory cue in a sequence on either the left or the right side, using different regressors than for the subsequent cues. The main reason for this was the commonly observed phenomenon of sensory adaptation, where sensory responses are strongest to the onset of a stimulus, but then adapt and display lower levels of sustained activity (Adibi & Lampl, 2021; Keller et al., 2017). Therefore, two distinct sets of regressors were used to describe the sensory responses in each stimulus condition. First, a stimulus onset kernel was used, which was aligned to the presentation of the first sensory cue in the sequence and ranges over the remaining stimulus and delay period. Second, subsequent cue kernels were used that captured stereotypical responses around the individual subsequent cue presentations.

To study how responses to multisensory stimuli were modeled by the encoding model, we first investigated the individual β-weight kernels of the stimulus onset regressors (Figure 9A). Here, the visual stimulus kernel displayed the highest positive weights in the contralateral V1 ($P < 0.001$, n=57 sessions from 4 mice, individually tested in both the first and last second of stimulus period, Wilcoxon sign-rank test). Further positive weights were found in the secondary visual cortices, ACC and the anterior pole of the cortex; the same regions displaying ΔR^2 for visual regressors (Figure 8). These positive weights mean that the encoding model uses increased activity in these regions in order to predict the cortical activity, whenever a

visual stimulus was presented. As such, these β-weights over time can be treated as an approximation of the cortical activity corresponding to the individual events. However, particularly towards the end of the stimulus period, also negative β-weights were found over broad regions of the somatosensory cortex and parts of the motor cortex. Here, in the last second of the stimulus period both ipsilateral wS1 ($P < 0.01$, n=57 sessions from 4 mice, Wilcoxon sign-rank test) and contralateral wS1 ($P < 0.001$, n=57 sessions from 4 mice, Wilcoxon sign-rank test) displayed significant negative β-weights.

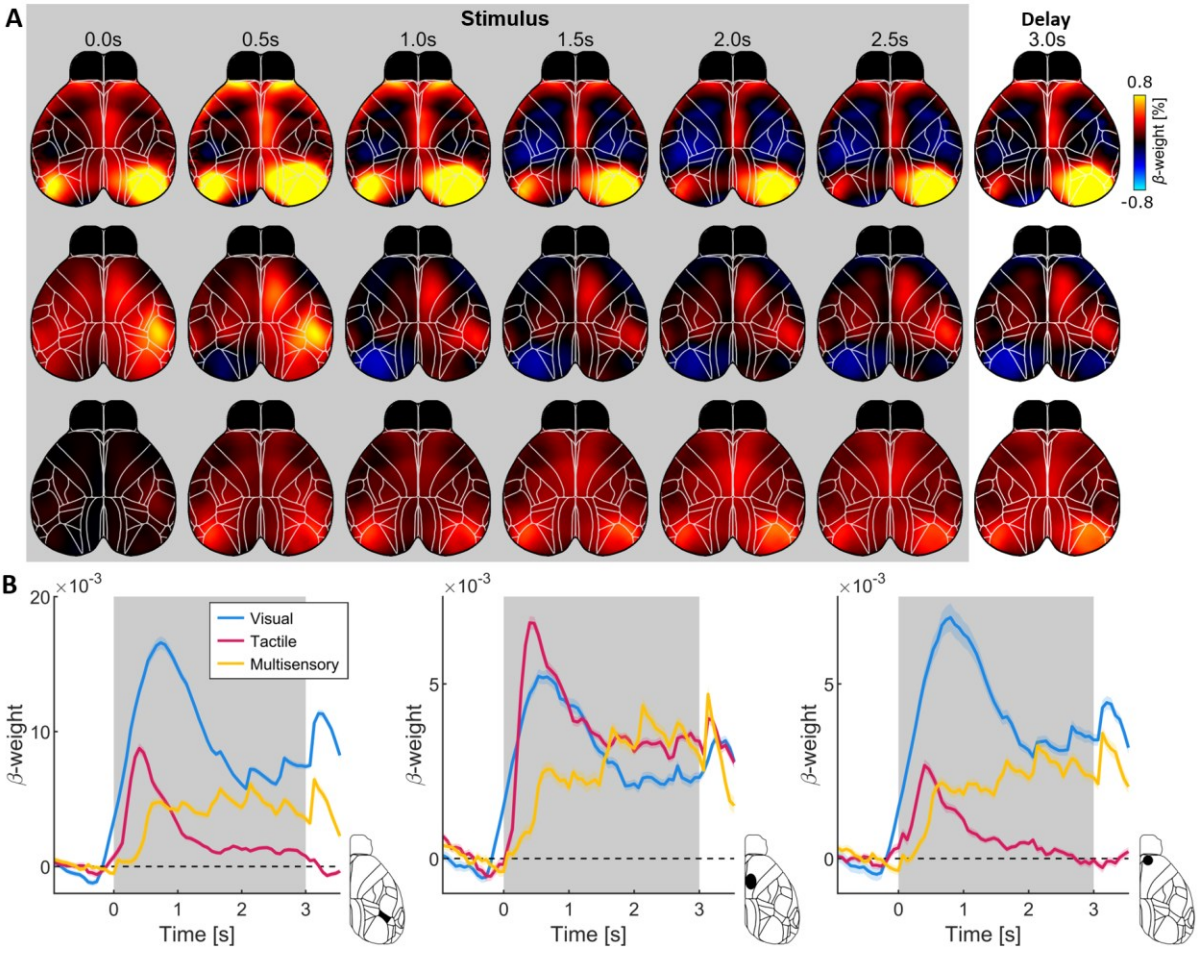

Figure 9: Super-additive multisensory activity revealed using a linear encoding model.
(A) Average β-weight maps of the stimulus onset kernels in the encoding model (see Figure 8). Maps represent 0.5s time-bins starting with the stimulus onset. Times displayed on top of the maps indicates beginning of the 0.5s time-bin over which weights is averaged. Shading indicates stimulus and delay period, applicable for the majority of trials with the first sensory cue presented at the beginning of the stimulus period. Top: average visual kernel over the course of a trial. Middle: average tactile kernel. Bottom: average multisensory interaction kernel. (B) Traces (mean ± s.e.m.) to compare stimulus conditions, for areas RL (left), MM (middle) and AMM (right). Maps indicate corresponding ROI location.

These indicate that the presentation of visual stimuli resulted in an inhibition of activity in these regions. A similar pattern was found for the tactile stimuli. Here, positive tactile β-weights were consistently found in contralateral wS1 ($P < 0.001$, n=57 sessions from 4 mice, individually tested in both the first and last second of stimulus period, Wilcoxon sign-rank test) and MM over the course of a trial. Furthermore, an inhibition of activity during the last second of the stimulus period was found in ipsilateral V1 ($P < 0.001$, n=57 sessions from 4 mice, Wilcoxon sign-rank test), but not contralateral V1 ($P > 0.4$, n=57 sessions from 4 mice, Wilcoxon sign-rank test). This indicates that cross-modal inhibition is a general feature in unisensory trials.

To investigate responses in multisensory trials, we studied the multisensory interaction kernels that model sensory responses deviating from a linear combination of both visual and tactile responses. This multisensory interaction kernel displayed almost exclusively positive β-weights over broad regions of the dorsal cortex, with higher bilateral symmetry than found in the visual and tactile β- kernels. Here, the highest positive weights were found in both ipsi- and contralateral V1 ($P < 0.001$, n=57 sessions from 4 mice, Wilcoxon sign-rank test), RL and MM. The same regions in which these regressors contributed the most unique information for the model predictions. This suggests that multisensory stimuli evoke predominantly superadditive multisensory integration.

To compare the sensory responses between the visual, tactile and the multisensory interaction kernels, the weights in areas RL, MM and AMM were examined in detail (Figure 9B). Generally, very similar response profiles were found within the individual modality regressors over all three areas. The visual and tactile β-weights quickly rose and peaked within the first second after the aligned stimulus onset and after that fell again. However, in all three areas the visual β-weights stayed elevated compared to baseline for the remaining stimulus and delay period. The same characteristic was also found for tactile β-weights, but only in MM. In area RL and AMM the tactile β-weights fell close to baseline levels again after the initial sensory response. Comparing these responses to the multisensory kernel, a different trend could be found. The β-weights of the additional multisensory kernel displayed a delayed rise and showed no clear signs of falling again over the remaining stimulus and delay period. In area MM and AMM the weights even seemed to rise over the remaining stimulus and delay period. These results showed a clear tendency that argued for strictly super-additive

multisensory integration of the visual and tactile information in the cortex. However, a potential issue might arise from the consistent timing of the stimuli. While this is not the case for every trial, in many trials the first sensory cue was presented at the beginning of the stimulus period. Therefore, these regressors had a strong tendency to be correlated with the general timing of the task. This might have resulted in these regressors capturing cortical activity that was related to the general time-course of trials, in addition to the sensory information. To address this concern, we next inspected the subsequent stimulus cue kernels, modelling individual sensory cue responses.

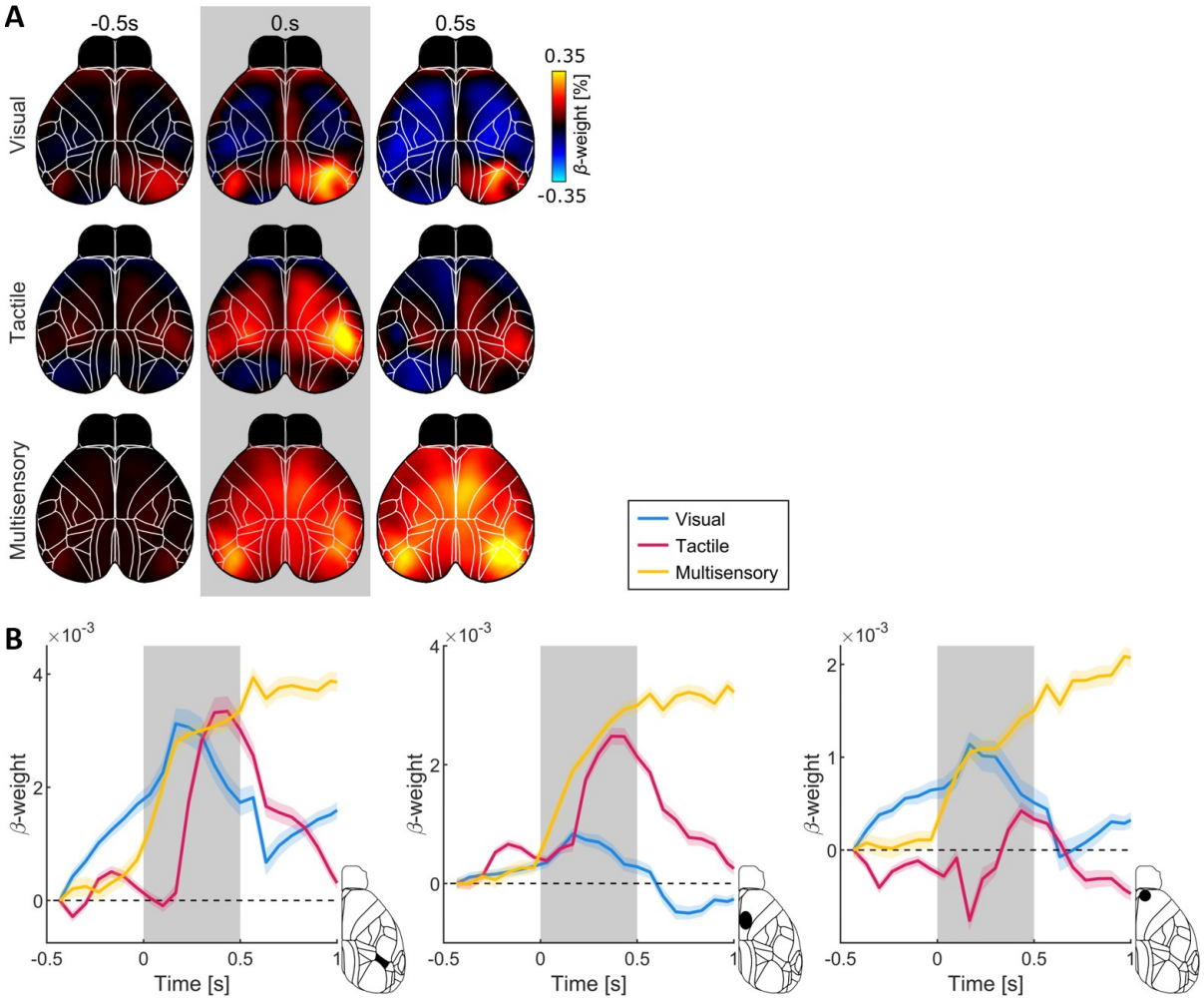

Figure 10: Super-additive multisensory activity in response to subsequent sensory cues. (A) Average β-weight maps of the subsequent sensory cue kernels in the encoding model (see Figure 8). Maps represent 0.5s time-bins around the individual cue onset. Times displayed on top of the maps indicates beginning of the 0.5s time-bin over which activity is averaged. Shading indicates the corresponding stimulus cue. Top: average visual cue kernel. Middle: tactile cue kernel. Bottom: multisensory interaction cue kernel. (B) Traces (mean ± s.e.m.) to compare stimulus conditions, within areas RL (left), MM (middle) and AMM (right). Maps indicate corresponding ROI location.

The subsequence sensory cue kernels, were aligned around the onset of each individual sensory cue, excluding the initial cue which was described by the sequence onset kernel. These regressors displayed far less alignment to the general time-course of a trial, since they only capture activity in a short time-window around individual cues, which occurred at different points in time in the remaining 2.5 s of the stimulus period. Despite that, very similar patterns of activity were found in the subsequent cue kernels compare to the sequence onset kernels. Here, the visual cue kernel displayed positive weights in the visual cortices (contralateral V1, $P < 0.001$, n=57 sessions from 4 mice, Wilcoxon sign-rank test) and to a lesser extent also in ACC and AMM (Figure 10A, top). This activity peaked during the 0.5 s cue presentation window in areas RL and the AMM (Figure 10B). Weights sharply dropped again after the cue period. This general tendency was also evident in MM but with much lower β-weights compared to the tactile cue kernel. Similar to the sequence onset kernel, following the current cue presentation, significant negative β-weights are found in both ipsi- and contralateral wS1 ($P < 0.001$, n=57 sessions from 4 mice, Wilcoxon sign-rank test).

The tactile β-weight kernels displayed positive weights following the airpuff delivery, especially in contralateral wS1 ($P < 0.001$, n=57 sessions from 4 mice, Wilcoxon sign-rank test) and MM (Figure 10A, middle). This response profile also was seen in area RL and MM, but was not clearly visibly in AMM. Following the current cue presentation negative β-weights were found in ipsilateral V1 ($P < 0.001$, n=57 sessions from 4 mice, Wilcoxon sign-rank test), but not in contralateral V1 ($P > 0.06$, n=57 sessions from 4 mice, Wilcoxon sign-rank test).

Finally, we compared these responses again to the multisensory interaction kernels. Here, positive β-weights after the individual cue onsets were found broadly distribute over the cortex. This also applied to area RL and MM after the individual cue presentation, but could also be seen in AMM. In all three regions persistent or even further rising positive β-weights were found after the cue presentation. Combined with the responses of the stimulus onset kernels, this strongly suggested that with every single multisensory cue, activity continued to increase in many distributed cortical areas, with the strongest responses in RL and MM. This behavior of accumulating activity might be the key for the mice to perform multisensory trials more successful that the individual unisensory trials.

These finding gave us a clearer understanding of the individual unisensory responses and even demonstrated superadditive multisensory responses in mice performing the multisensory

evidence accumulation task. To understand how these sensory responses might be used to eventually guide the behavioral decisions of individuals, we next studied if the cortical activity that we measured contained sufficient information for the mice to solve the behavioral tasks. To address this question and identify potential differences in visual and tactile task performances, we investigated which cortical regions display activity that could be used by the mice to solve the behavior task.

2.7 Areas RL and MM represent modality-specific sensory evidence

To understand how mice could solve the behavior task, we investigated which cortical regions displayed activity that would allow individuals to reliably identify the target stimulus side. We consistently observed strong and symmetric co-fluctuations of activity in both hemispheres on a single trial level. These made the interpretation of activity of an individual hemisphere out of the context of the other hemisphere difficult. Furthermore, since the objective of the mice was to integrate sensory information from both sides over time to eventually produce a lick response either towards the left or the right side, we suspected the presence of lateralized cortical activity that coded for the side of the target stimulus. This is of particular interest in more difficult discrimination trials, where the mice are presented with both sensory evidence for and against the target-side. Therefore, to identify cortical regions coding for the target stimulus side, the differential in cortical activity between both hemispheres was computed on a trial-by-trial bases. Then, the area under the receiver-operator-characteristic curve (AUC) was computed. This metric indicates how distinct the distributions of responses were over trials, depending on the target-side. Here, AUC-values close to 0.5 indicate low separability, meaning that a given cortical area did not response differently in left target- and right target trials. Therefore, such an area would likely be less informative for the mice's decision-making process. In contrast, AUC-values below 0.5 indicate decreased responses in the hemisphere contralateral to the target stimulus side, compared to the ipsilateral side. Finally, AUC-values greater than 0.5 indicate increased contralateral activity compared to the hemisphere ipsilateral to the target stimulus side.

In this step of the analysis, we want to identify cortical activity that would allow the mice to determine the target stimulus side based on the sensory information presented. But crucially,

we wanted to distinguish these responses from activity related to the decision-making process, as well as the preparation and execution of the lick-responses. Therefore, we had to minimize the impact of activity related the behavioral choices of the mice. To achieve this, we balanced the number of left and right, as well as correct and incorrect response trials, effectively decoupling the sensory information about the target-side from the responses given by the mice. This now allowed us to identify differences in cortical activity between the hemisphere's indicative of the target stimulus side and compare these between stimulus conditions.

First of all, in visual discrimination trials, AUC-values especially in secondary visual cortices, but also in V1, RSC and ACC were consistently higher on the side contralateral to the target stimulus over the entire stimulus and delay period (Figure 11A). In addition, decreased activity was found over a lateral band reaching from ALM over lateral M1 to the anterolateral somatosensory cortex on the contralateral compared to the ipsilateral hemisphere. During the stimulus presentation in tactile trials the highest AUC-values were found in wS1 and MM, but also to a lesser extent in the visual cortices and RSC. However, during the delay period high, positive AUC-values were primarily found in wS1 and MM. Finally, in multisensory discrimination trials, high positive AUC-values were found in wS1, secondary visual cortices, RSC and MM; the same regions seen in unisensory trials. This activity remained indicative of the target stimulus side over the entire stimulus period. During the delay period the highest AUC-values were still found in MM and the higher visual areas, including RL.

When comparing the regions predictive of the target stimulus side between the visual and tactile trials especially during the delay period, it appeared that the activity in MM was much more predictive of the target in tactile trials, while activity in higher visual cortices more reliably indicated the target in visual trials. In multisensory trials, both of these regions display high AUC values and could therefore both be used to solve the behavioral task.

To further investigate the role of RL and MM in the different stimulus conditions, we compared the AUC-values over the course of a trial (Figure 11B). In both areas the AUC-values rose quickly with the stimulus onset and remained above chance level for the entire stimulus and delay period, only falling again during the response.

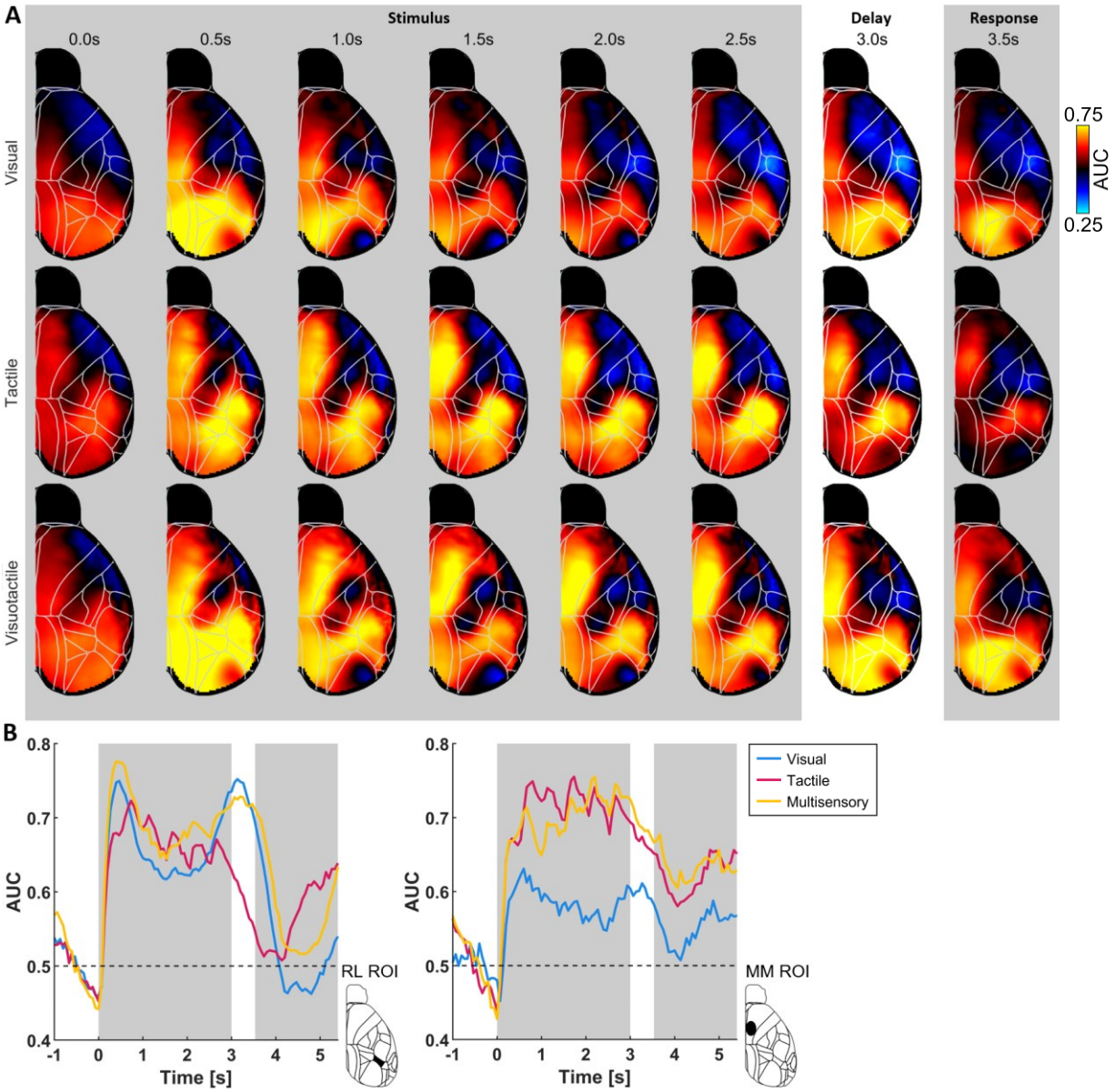

Figure 11: Areas RL and MM represent modality specific sensory evidence
Cortical regions representing sensory evidence were investigated by analyzing how reliably activity in individual regions reflects the target stimulus side. For this, differences in the activity of the left and right hemisphere were analyzed in discrimination trials. To focus on the sensory responses, a balanced number of correct and incorrect as well as left and right trial was used. Here, AUC-values are computed pixel-wise, reflecting how distinct responses are modulated depending on the target-side in individual trials. (A) From top to bottom: Rows display AUC-value maps for visual, tactile and multisensory discrimination trials. Individual maps represent averages over 0.5s time-bins, beginning at the time indicated on top of the maps. Shading highlights the stimulus and response period. (B) Traces to compare AUC-values between stimulus conditions in areas RL and MM. Cortical maps indicate ROI locations.

To compare how informative activity in RL and MM was for the choices of the subjects, we investigated if AUC-values were significantly different during the delay period preceding the response in the individual stimulus condition. In area RL, AUC-values were significantly higher in visual trials compared to tactile trial ($P<0.01$, n=2,304 trials from 4 mice, Wilcoxon rank-sum test). Activity in RL was also more reliable in multisensory compared to tactile trial ($P<0.05$, n=1928 trials from 4 mice, Wilcoxon rank-sum test). However, comparing AUC-values in visual and multisensory trials, no significant difference was found ($P>0.83$, n=2,096 trials from 4 mice, Wilcoxon rank-sum test).

Next, we compared AUC-values in MM. Here, significantly higher AUC-values were found in tactile- compared to visual trials ($P<0.001$, n=2,304 trials from 4 mice, Wilcoxon rank-sum test). Responses in MM also more reliably reflected the target-side in multisensory compared to visual trials ($P<0.001$, n=2096 trials from 4 mice, Wilcoxon rank-sum test). However, no significant difference was found between tactile and multisensory trials ($P>0.92$, n=1928 trials from 4 mice, Wilcoxon rank-sum test).

To summarize, while similar regions could be found in all stimulus conditions with activity predictive of the target stimulus side, we found a more accurate representation of tactile targets in MM, while activity in area RL more accurately reflected the target-side in visual trials, immediately preceding the responses of the mice. In multisensory trials both of these areas predicted the target-side as accurately as in the best unisensory condition. While these regions didn't display more reliable activity in multisensory trials, the redundant representation of sensory evidence in multiple cortical regions likely lead to higher certainty and more successful decisions in multisensory trials.

2.8 Activity in motor- and parietal-cortex are predictive for upcoming choices

Finally, to understand how perception translates to the animal's decisions, a recently established logistic-regression decoder (Musall et al., 2021) was used to determine which cortical regions displayed activity predictive of the animals' choices. This decoder was designed to predict the responses of mice given the activity at a specific time point during a trial. This allowed us to investigate how accurately the animals' decisions could be predicted over the course of individual trials, based solely on the cortical activity in the current trial. To

compare responses related the animals' decision instead of the sensory evidence, a balanced number of left and right as well as correct- and incorrect response trials were used within individual sessions, similar to the previous section. With this decoder, we found that during the stimulus period the accuracy of choice predictions steadily increased to 77.9% ± 2.1% (mean with 95% confidence interval (CI), n=78 sessions from 4 mice) correct choice-predictions at the end of the stimulus period (Figure 12A). Even in absence of further sensory information during the delay period, prediction accuracy continued to increase to 81.0% ± 2.0% (mean with 95% CI, n=78 sessions from 4 mice). The highest accuracy of the decoder with 90.8% ± 2.0% (mean with 95% CI, n=78 sessions from 4 mice) was reached 0.8 s into the response period, after which the accuracy steadily decreased again. Reaching these levels of prediction accuracy even before the mice gave their responses, suggested the presence of robust features in the cortical activity that could be used by the decoder to accurately predict the upcoming responses of the individuals on a single trial basis.

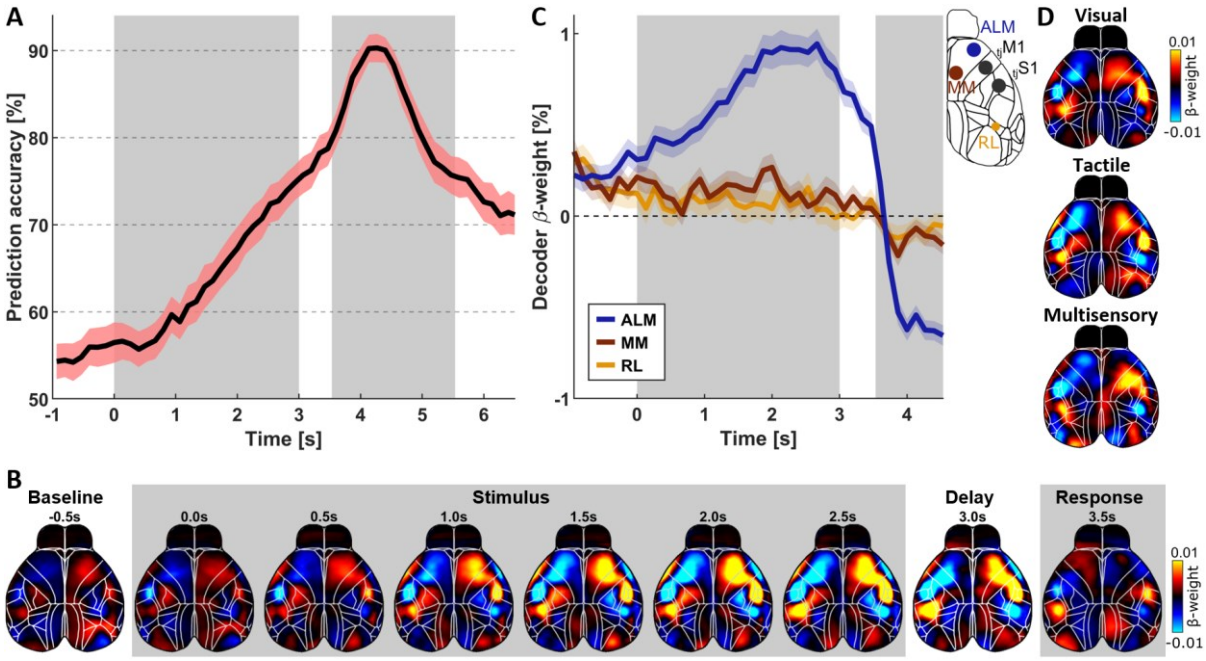

Figure 12: Motor- and Parietal cortex are predictive for upcoming choices.
(A) Prediction accuracy of the logistic regression decoder fit to all stimulus conditions over the course a trial, as mean ± 95% CI. (B) Average β-weight maps for 0.5s time-bins over the course of a trial. Times indicated on top of each map indicate beginning of the time-bin. Positive weights indicate increased activity related to left-choices. (C) Left: β-weights over time for areas RL, MM and ALM, as mean ± s.e.m. Right: indicator of ROI locations. (D) Average weights binned over stimulus and delay period for visual (top), tactile (middle) and multisensory trials (bottom). Based on separate decoders fit only on trial of the corresponding modality.

To study the structure of the cortical activity that was used by the decoder to predict responses, we investigated the average β-weight maps over the course of a trial (Figure 12B). Generally, β-weights of the decoder were found broadly distributed over the entire dorsal cortex. However, especially over the course of the stimulus and the delay period, a robust pattern of β-weights in the anterolateral motor cortex (ALM; compare Figure 12C, right; also referred to as tjM2 see Esmaeili et al., 2020), the tongue and jaw region of the primary motor cortex (tjM1) and the corresponding region of the somatosensory cortex (tjS1) was found.

Next, to investigate if either RL or MM also displayed clear choice-related signals in addition to the sensory responses seen in the previous sections, β-weights in these regions were compared to those found in ALM. Here no clear changes in choice-predictive β-weights were observed during the stimulus presentation and the delay period in either RL or MM. During the delay, immediately preceding the response, β-weights found in RL and MM were not significantly different from zero ($P(RL)>0.15$; $P(MM)>0.11$, n=78 sessions from 4 mice, Wilcoxon sign-rank test). In contrast, β-weights in ALM displayed a strong increase over the course of the stimulus period, meaning that activity in ALM became increasingly predictive for the upcoming choice of the animals upon presentation of sensory information. Here, β-weights found in ALM during the delay period were significantly higher than zero on the side contralateral to the stimulus ($P<10^{-7}$, n=78 sessions from 4 mice, Wilcoxon sign-rank test).

Lastly, to study if there are indications of modality-specific differences in the choice-predictive areas, separate decoders were trained on trials of only one stimulus condition at a time. With this approach, separate β-weight maps for visual (n=3,072 trials from 4 mice), tactile (n=4,668 trials from 4 mice) and multisensory trials (n=476 trials from 4 mice) were computed (Figure 12D). In all stimulus condition, very similar patterns of positive β-weights across ALM, tjM1 and tjS1 were found, with no obvious modality-specific choice areas appearing.

To summarize, in the previous sections we found that both RL and MM displayed sensory responses in the individual stimulus conditions and even supperadditive multisensory responses. Further these regions contained information about the accumulated sensory evidence that mice could use to solve the task, but did not express activity predictive of the responses the mice were going to give at the end of a trial. Instead, this activity was found in ALM, as well as tjM1 and tjS1. With the latter two regions likely relating to the execution of the motor-responses themselves (Xu et al., 2022), while ALM seems to be closer to the

decision-making process itself, as indicated by previous studies in tactile delayed-response task (Chen et al., 2017; Esmaeili et al., 2020).

3 Discussion

Over the course of this thesis we established a novel visuotactile evidence accumulation task for head-fixed mice. Mice successfully learned to perform this task in all stimulus conditions. This then allowed us to study how the mice perceive the sensory stimuli and translate this information into response behaviors, using widefield Ca^{2+}-imaging. With this, we identified regions in the posterior parietal cortex (PPC) and the medial frontal cortex, which displayed spatially overlapping sensory responses in both visual and tactile trials. Using a linear encoding model, we were able to isolate sensory responses from other task-aligned cortical activity. With this, we found cross-modal inhibition in unisensory trials, where activity is inhibited in the non-stimulated sensory cortex. Furthermore, we were able to reveal that mice display superadditive multisensory responses broadly distributed over the cortex, but particularly in area RL and MM. Then, to link how sensory activity is translated to behavioral responses, we identified regions that reliably represented the target stimulus side and could therefore be readout by the mice to form correct decisions. These regions were broadly distributed, but displayed modality specific differences. In visual trials, more robust target-side specific responses were found in the secondary visual cortices, whereas activity reflecting the target in tactile trials was primarily found in MM. Both of these regions reliably represented the target-side in multisensory trials, likely resulting in higher certainty and improved performance of mice in multisensory trials. To study if the same regions also reflect the behavioral responses of subjects, we build a choice-decoder that reliably predicted the upcoming responses of mice. In all stimulus conditions, this decoder primarily used activity in area ALM, tjS1 and tjM1 to predict the behavioral responses. This indicates that mice accumulate the sensory evidence in modality-specific regions and this information is then relayed to area ALM to form the behavioral decisions across modalities. Furthermore, this circuit facilitates mice to use visual and tactile sensory evidence synergistically to increase their certainty and improve their behavioral performance. In the following these results will be discussed in detail.

3.1 Visuotactile evidence accumulation task

3.1.1 Task

To study how mice integrate sensory information over different sensory modalities to guide their behavioral decision, we established a novel visuotactile evidence accumulation task. This task incorporated different components of previously published behavioral paradigms (Bexter, 2022; Musall et al., 2019; Pinto et al., 2018; Raposo et al., 2012; Scott et al., 2015). Key features of this task were the independent presentation of visual and tactile sensory evidence, allowing us to compare performance and neural responses in unisensory and multisensory trials, and the delayed response period that required short-term memory (Chen et al., 2017; Gallero-Salas et al., 2021).

We found similar behavioral choice accuracy to previous reports in rats (Raposo et al., 2012; Scott et al., 2015) and mice (Musall et al., 2019; Pinto et al., 2018). This behavioral task gave us the opportunity to study how mice represent the sensory information in cortex and how this is used to solve the task, by the means of widefield Ca^{2+}-imaging.

3.1.2 Sensory stimuli and task design

In this study we used individual cycles of a full field moving grating as visual cues for the mice. However, this was not our first approach. Initially, we attempted to train a previous cohort of mice on visual stimuli presented in the form of LED flashes, similar to visual stimuli successfully used in rats (Raposo et al., 2012; Scott et al., 2015). However, our mice were unable to learn the behavior task when presented with LED flashes (Bexter, 2022). Previous studies in mice had successfully used visual stimuli in the form of moving gratings or moving virtual objects presented on screens or domes (Burgess et al., 2017; Busse et al., 2011; Goldbach et al., 2021; Pinto et al., 2018). Therefore, we adapted our stimulation protocol to also employ moving gratings. To give the mice the best chances of learning the visual task, we chose a spatial frequency and speed of the gratings that well matched the preferred features of the primary and secondary visual cortices (Marshel et al., 2011). We also ensured that the selected stimuli matched sensory features that are suitable for stimulation of area RL, due to the known role of RL in visuotactile integration (Olcese et al., 2013).

Mice were able to reliably learn this task when presented with drifting gratings, suggesting that the movement of the visual cue might be an essential feature for the mice to perform visual behavior tasks. This is also supported by the finding that mice can perform visual behavioral tasks when stimulated using LED arrays that simulate a moving visual stimulus (Musall et al., 2019).

However, as a result of the low spatial frequency as well as the relatively low speed of the stimuli, the stimulus presentation period became very long, compared to previous studies (Chen et al., 2017; Musall et al., 2019; Raposo et al., 2012; Scott et al., 2015). This reduced the number of trials that could be studies in the same given time. In general, it would be advantageous if this feature of the task could be improved. Subsequent studies could attempt to reduce the duration of the stimulus period, for example by presenting the individual cues drifting with the same speed, but for a shorter amount of time, in quicker succession. However, great care has to be taken when altering such fundamental features of a behavioral task. Any alteration has a high potential to fundamentally change how sensory information is perceived, the way visual and tactile information is integrated and might even prevent the mice from learning the task in the first place.

In addition to the long stimulus period, the task also features a relatively long inter-trial-interval. The reason for this was that our previous experiments, using widefield Ca^{2+}-imaging in combination with this task, revealed cortical activity patterns more than 1s after the end of the response period, that otherwise could have extended into the next trial; affecting the interpretation of measured cortical activity. However, one could attempt to randomize and on average reduce the length of the inter-trial interval, while accounting for such trial by trial differences, for example using a linear encoding model, similar to the one employed in this thesis. However, the resulting uncertainty about the time-course of a session might affect the performance of the mice, by prohibiting task-aligned preparatory activity preceding the stimulus onset. An increase in such task-aligned movements over the course of training has previously been reported (Musall et al., 2019) and might reflect a critical aspect of the learning process.

We presented tactile stimuli in the form of airpuffs, similar to previous approaches (Bexter, 2022; Ollerenshaw et al., 2012), to prevent any visible components in the tactile stimulus, as would be the case using moving rods or rough surfaces which have been employed in previous

studies (Ayaz et al., 2019; Li et al., 2015; Petersen, 2007). It was taken great care that the airpuffs reliably deflected the whiskers of the mice, while avoiding direct airflow towards the face. This was confirmed by the reliable responses in wS1 seen in Figure 6 and even more clearly in the unique explained variance plot in Figure 8E.

3.1.3 Learning

To teach mice the task, they were first trained on visual, then tactile and finally multisensory detection trials. Using this approach, it could be avoided that mice get overwhelmed by the alternating sensory scenery in the early training stages. Mice reliably learned to perform the task in the individual stimulus condition (Figure 4B) and later were able to perform the task also when presenting trials with different sensory modalities in random succession.

3.1.3.1 Unisensory learning

Over the course of training, strong differences could be found in the learning curves of the individual stimulus condition. It took mice multiple sessions to reach reliable performance in the initial visual condition, while mice displayed high performance in tactile trials after, in many cases just one session of learning to adapt to the new stimulus (Figure 4B). This much quicker learning of the subsequent tactile condition, suggests that it was easier for the mice to learn this novel tactile condition, compared to the initial visual condition. A likely cause of this result is that mice initially also had to learn the general principals of the behavioral task. These include habituation to the water-restriction, the setup and the training schedule, the association of lick responses at presented water-spouts to receive the water reward and then finally the association of sensory evidence informing their subsequent, delayed response actions. All of these aspects that mice have to learn, likely contributed to the longer time mice needed to learn the initial visual task condition.

Therefore, the faster learning of the subsequent tactile condition, could be an indication that mice are able to generalize learned principles of performing the behavior task which they acquired during the visual condition. This implies that mice have the ability to generalize the sensory information to such an extent, that the same neural circuit that translate the sensory evidence into response behaviors in a visual condition, could adapt to the new form of sensory

evidence in the tactile condition. Aspects of such generalized circuitry flexibly linking visual and tactile sensory information to the same motor responses should also be evident in the widefield Ca^{2+}-imaging. In fact, while we have found evidence for regions that individually represent either visual or tactile evidence more reliably (Figure 11), no clear modality-specific choice areas could be found (Figure 12).

Nonetheless, there is also the possibility that it is generally easier for mice to learn the tactile detection condition compared to the visual condition. To address this question, future studies could start to train mice first on the tactile condition and only subsequently introduce the visual detection condition. By then comparing the relative amount of time mice need to learn the individual conditions, it could be determined if one stimulus condition is inherently easier for the mice to learn than the other.

3.1.3.2 No additional multisensory learning

After successfully learning to perform the detection task in both the visual and tactile condition, multisensory detection trials were introduced. Strikingly, mice were able to perform this condition from the first session on at higher success-rates than in the familiar unisensory trial conditions (Figure 4B). This shows that mice readily employed the learned principles from the unisensory conditions to the multisensory trials, without first having to adapt to the new stimulus condition, over an extended period of time. A simple explanation for such an effect would be that mice simply focus on sensory information from just one of the two presented modalities and ignore the other. If this were the case, we would expect the mice to perform on par with the individual unisensory conditions or even slightly worse due to the additional effort of having to ignoring the other modality. Instead, we find that mice perform the multisensory condition instantly better than either unisensory condition, demonstrating that they are using both the visual and the tactile information synergistically, even without requiring additional, extensive learning.

3.1.4 Lapse trials

Comparing performances in the more difficult discrimination trials, we found that mice also consistently performed the multisensory condition better than the unisensory discrimination

conditions (Figure 4C). Mice are therefore able to take advantage of the additional sensory information over a broad range of task difficulties, similar to previous reports (Bexter, 2022; Coen et al., 2023; Raposo et al., 2012). Surprisingly, we even found robust improvements in multisensory detection trials compared to unisensory trials. Here, in the detection condition we presented the maximum number of 6 sensory cues on the target-side an no distractor cues, making the sensory evidence unambiguous. Despite that and the extensive training time (>6months) of the mice on detection trials, (where throughout at least a subset (>10%) of trials in a given session were detection trials) mice were not able to reliably perform these conditions at 100% success rate. Such few but persistent error trials independent of training duration and task difficulty have previously been termed “lapse-trials” (Busse et al., 2011; Carandini & Churchland, 2013; Gold & Ding, 2013). A recent study investigated the source of such presumable “lapse-trial” in rats performing a similar audiovisual evidence accumulation task and found that the best explanation for the persistence of these incorrect response trials is a trade-off between behavioral strategies of exploration and exploitation (Pisupati et al., 2021). Under this interpretation, these error trials are caused by animals exploring the consequences of alternative actions, in contrast to the known, rewarded actions. In their experiments they also observed a decrease in the lapse-rate of rats performing multisensory trials, the likely reason for this was lower uncertainty in these trials compared to the unisensory conditions. Our results in mice performing a visuotactile evidence accumulation task are well in line with these findings. In addition to this, we also found significantly earlier lick responses in mice performing multisensory (Figure 4E). Due to the response delay included in the task design, these earlier responses do not represent quicker reaction times as have been found previously in multisensory tasks (Diederich & Colonius, 2004; Hirokawa et al., 2011; Sakata et al., 2004). Instead, these quicker responses could be further evidence for higher certainty of mice performing multisensory trials, compared to unisensory trials. In conclusion, features of this higher certainty in multisensory trials should also manifest themselves in the neural activity of the mice (see section 3.2.6).

3.1.5 Differences in running behavior

Mice displayed very stereotypical running behavior in the task. This might reflect an important behavioral strategy to perform this task in particular. Especially in visual trials this fixed

running pattern might help the mice to interpret the sensory information, potentially using existing circuitry that links visual flow and locomotion (Bexter, 2022; Leinweber et al., 2017). To study the relevance of this running behavior for the general task performance, subsequent studies could prohibit this form of locomotion. This could either be achieved by blocking the wheel in a subset of trials or even entire sessions and compare performances in running and stationary trials. The wheel could also be replaced entirely with a stationary platform from the beginning of the training to test if this hypothetical link of locomotion and visual flow is beneficial or even essential for mice to learn the visual task condition.

We also found that cortical activity related to the running of the mice made up a substantial fraction of the cortical activity that was measured using the widefield Ca^{2+}-imaging (Figure 8D). It is therefore essential to explicitly account for such behavior related responses when interpreting measurements of neural activity (Shimaoka et al., 2018). This is especially true if such behaviors display strong alignment to the task itself (Musall et al., 2019), as was the case for the running behavior of our task.

3.2 Widefield Ca^{2+}-imaging

To study how mice solve the behavior task in the different stimulus conditions we used widefield Ca^{2+}-imaging to record activity over the dorsal cortex of mice. For this purpose, a preparation was used that allowed to record cortical activity from the entire dorsal surface of the cortex through the intact skull (Guo et al., 2014). In mice performing all three stimulus conditions, the same general patterns of cortical activity were found in single trials of all stimulus conditions. First, stimulus specific but partially overlapping responses occurred during the stimulus period. Then, activity patterns became more similar during the delay (Figure 5 and Figure 6). Here, very similar increases in activity are found bilaterally in the medial somatosensory- and primary motor cortex, particularly in the front- and hind limb areas. These responses likely reflect, at least in part the strong reduction in running speed preceding the response (Figure 4D). In addition to this, also bilaterally increased activity could be found in the medial frontal cortex (mFC). Such responses have previously been linked to short-term memory functions in delayed-response tasks (Chen et al., 2017; Gallero-Salas et al., 2021; Guo et al., 2014). Finally, during the response period the entire cortex displayed

strong activity, with no clear indication for modality-specific differences in cortical activity. This suggests a convergence of sensory information from different modalities to drive choice-related activity patterns that are modality independent.

3.2.1 Cortex-wide visual response patterns

The stimuli in visual trials primarily resulted in activity of the primary- and secondary visual cortices (Andermann et al., 2011; Marshel et al., 2011; Murakami et al., 2015; Vanni et al., 2017; Wang & Burkhalter, 2007). Further, visual responses could be found in ACC, which have directly been described previously (Murakami et al., 2015; Vanni et al., 2017) and are also evident in the study by Musall et al., 2019. However, we also found robust visual responses at the anterior medial motor cortex of the mice, which have not been described previously.

The reason for this is in part that only few studies employed cortical preparations that allowed to record activity up to the anterior edge of the cortex in mice performing visual tasks (Musall et al., 2019; Zatka-Haas et al., 2021). Here, both of these studies report consistent responses in the occipital cortex and ACC, but visual responses in the anterior medial motor cortex were only evident in the study by Zatka-Haas et al., 2021. A recent study using high density electrophysiological recordings in mice performing almost the same task as used by Zatka-Haas, has also reported visual responses in both the primary and secondary motor cortex, the mPFC (medial prefrontal cortex; including the prelimbic and infralimbic cortex) and the orbitofrontal cortex of mice (Steinmetz et al., 2019). This raises the question under which conditions these responses in AMM, which were also seen in our study, can be found. Due to the localization of the responses, we initially had concerns that these signals in part might reflect light contamination from the stimulus presentation. However, we found that these signals were very reliable (Figure 6A). Since light contamination would be the consequence of errors in the light shielding in individual sessions, the high reliability of these signals makes it unlikely that they are entirely explained by signal contaminations. A possible alternative explanation might be slight technical differences in the recordings themselves. Our widefield recordings appear to be a bit more defocused than previous studies (Guo et al., 2014; Musall et al., 2019), this is likely caused by a slightly lower imaging plane. As a result, our recordings could be more sensitive to signals from lower regions in the frontal cortex. The reliable visual

responses we observed, might to an extent reflect responses of the frontal pole area, the orbitofrontal and the prelimbic cortex of the mice.

These areas have been linked to visual spatial memory, working memory in general, rule learning and cognitive flexibility in rats (Boulougouris et al., 2007; McDonald & White, 1994; Rich & Shapiro, 2007; Tsutsui et al., 2016). Recent studies in mice also have found central roles of orbitofrontal cortex in associative learning in a visual task (Liu et al., 2020) and linked mPFC (including prelimbic and infralimbic cortex) to learned rules in a visual categorization task (Goltstein et al., 2021). In our task, mice have to rapidly switch their attention between visual, tactile or both sensory modalities. This might require a level of cognitive flexibility in line with the established functions of mPFC and orbitofrontal cortex. The presence of visually evoked responses in AMM might therefore depend on the precise nature of the behavioral tasks employed. However, further studies would have to be conducted, ideally using single cell recordings, to unravel the precise origin of these responses and to which extent they depend on the predictability of the behaviorally-relevant sensory modality.

3.2.2 Response overlap in unisensory trials

To analyze the spatial overlap in the cortical responses to the visual and tactile stimuli, we first identified regions with a high potential for multisensory integration. We chose to investigate this overlap in responses by computing the product of the z-scored unisensory responses. This approach emphasizes regions that, on the one hand reliably respond to the individual sensory stimuli, and on the other hand display strong responses in both visual and tactile trials. With this approach we found particularly strong overlap in visual and tactile responses in area RL, which is in line with previous studies in rats (Wallace et al., 2004) and mice (Olcese et al., 2013). This suggests that our approach is well-suited to identify multisensory responses in specific cortical areas. Further, strong overlap could also be found in the frontal area MM. This area has been shown to represent whisker-related tactile sensory information (Chen et al., 2017) and display features of short-term memory in delayed-response tasks (Chen et al., 2017; Gallero-Salas et al., 2021; Guo et al., 2014). However, the role of MM in processing of visual information is unclear. Particularly strong and lateralized overlapping responses in MM were found during the delay period of the task. This

lateralization of responses especially in multisensory trials, makes this area a very promising region to be used by mice to improve their choices in multisensory trials (further discussed in Section 3.2.6). A third region was found that seems to display strong overlap in unisensory trial responses, a region in the anterior medial motor cortex. However, activity in this region was less pronounced that that found in PPC and MM, therefore later analysis mainly focused on the latter two regions. Next, we studied if these areas in particular displayed increased responses in mice performing multisensory trials.

3.2.3 Multisensory enhancement of cortical responses

Our initial recordings in mice performing detection trials, showed higher cortical activity in multisensory, compared to unisensory trials (Figure 4). In the previous section we identified overlapping unisensory responses in PPC and the medial frontal cortex that display a high potential for multisensory integration (Figure 5). Next, we investigated if multisensory enhancement was found particularly in these regions, when mice were presented with multisensory stimuli. For this purpose, we first computed the maximum responses over the unisensory trials. These maps showed well defined regions with either visual or tactile preference, persistent over the duration of a trial. The sensory preference of the regions found are well in line with the sensory domains outlined by the Allen CCF (Figure 2; Wang et al., 2020) and agree with previous reports of modality specific cortical responses (Allen et al., 2017). By comparing multisensory trial responses with these preferred unisensory responses, we characterized the extent of multisensory enhancement and depression in cortical responses. This revealed multisensory enhancement in many regions of the cortex during the stimulus and delay period of the tasks. Furthermore, we found the highest multisensory enhancement in PPC, primarily around area RL and in the mFC of the mice, well in line with our previous observations. These results for PPC also agrees with the reports by Wallace et al., 2004 that identified the highest proportion of multisensory responding neurons in between the unisensory domains of the cortex. Particularly for area RL a gradual decrease in the proportion of visually responding cells and an increase in tactile neuron with increasing distance from V1 towards S1 has been found, with multisensory neurons at the intersection (Olcese et al., 2013).

However, there is a potential issue with this approach of comparing multisensory and maximum unisensory trial responses using widefield imaging. Even though a gradual decrease in unisensory responses, with increasing distance from the primary sensory cortices, would be in line with the reports by Olcese et al., 2013, they could also be in part the result of light scattering in the tissue. By then comparing multisensory responses with the maximum unisensory response we might overemphasize the boarders of the unisensory domains. This makes is difficult to judge if the increased multisensory responses are the result of increased local activity or an artifact of overlapping scattered light from the neighboring visually and tactilely responding regions at the border of preferred sensory domains.

Subsequently, we compared multisensory trial responses to the sum of the unisensory responses. One the one hand, this circumvented the issue of overemphasizing the boarders between sensory domains, since activity in both unisensory trial conditions is taken into account, including potential light scattering. On the other hand, this approach in principle would have allowed us to characterize multisensory responses as being superadditive, additive or subadditive (Stein & Stanford, 2008). These results indicated that multisensory responses are subadditive during the stimulus presentation. However, by far the largest signals were actually found in the response period of the task, in the absence of sensory information. This highlighted another crucial flaw when using trial averaged activity as an approximation of sensory responses. This approach neglected the fact that mice have temporal expectations about the task (Esmaeili et al., 2021; Orsolic et al., 2021), express response preparation (Chen et al., 2017; Guo et al., 2014; Li et al., 2015; Xu et al., 2022) and execute instructed as well as uninstructed movements (Musall et al., 2019; Orsolic et al., 2021; Shimaoka et al., 2018), all of which are accompanied by considerable cortical activity that is visible in each sensory condition.

When testing if multisensory responses could be described as the sum of visual and tactile evoked activity we implicitly tested if the following relationship holds:

$$M - (V + T) = 0 \quad \text{(I)}$$

Here, M is the sensory response in multisensory trials, V the sensory response evoked by the visual stimulus and T the response evoked by the tactile stimuli. However, by approximating

sensory responses with trial averages, we actually compare both the sensory response and any cooccurring non-sensory activity over conditions:

$$(M + \text{non-sensory}) - ((V + \text{non-sensory}) + (T + \text{non-sensory})) = 0 \Leftrightarrow$$

$$M - (V + T + \text{non-sensory}) = 0 \quad \text{(II)}$$

Here, the non-sensory term refers to any task- or behavior related activity that is not reflected in the sensory responses. With this we see that the non-sensory related activity does not cancel out, which biases the results of this analysis toward subadditive effects with multisensory stimulation. We therefore cannot directly compare the sensory responses without explicitly accounting for all non-sensory cortical activity present. Furthermore, here we assume the simplest case possible, that the non-sensory activity is identical in all stimulus condition. But we have seen already that this is not the case (Figure 4D and E). Mice express different running patterns in the individual stimulus condition. It has been shown that running has diverse effect on cortical activity (Pakan et al., 2016; Polack et al., 2013; Shimaoka et al., 2018). This means that our comparison of multisensory trial activity with preferred unisensory trial activity is flawed, since we are not only comparing the sensory responses, but also the differences in responses related to other behavioral features. In order to study cortical sensory responses in awake mice, especially when performing a behavioral task, we must therefore account for the cortical activity that is linked to the behavior of the subjects as well as other task parameters that are temporally-aligned to the sensory stimuli. To achieve this, we used a linear encoding model to decompose the cortical activity into individual components relating to specific task- or behavioral events (Musall et al., 2019; Orsolic et al., 2021).

3.2.4 Linear encoding model

To study the sensory responses in mice performing an active behavioral task, it is essential to account for any cortical activity that does not reflect a direct sensory response but might occur at the same time as a sensory stimulus. A recently published model was the best approach for this application (Musall et al., 2019). This model was designed to reconstruct the cortical activity measured in individual sessions based on linear combinations of cortical responses related to individual task- or behavior-related events. The approach allows to

flexibly define event types that have repeating occurrences within a given session and the model determines the best average cortical response pattern over time that relates to the individual event type. Even though, this means that for example a visual stimulus can only be described as an average cortical response, we are still able to capture differences in cortical activity between trial, as long as these differences in activity are related to an external feature such as a movement of the mouse that we can observe. By recording the individual lick-responses, the running speed of the mice and videography of further movements, we could use this approach to decompose the sensory responses related to our stimulation and other task-aligned activity. Furthermore, since we were presenting random sequences of sensory cues in discrimination trials, we could distinguish individual sensory cues responses from simultaneously occurring, task-aligned activity.

The model we build based on combining a large array of task- and movement-related variables, was capable of explaining more than half the variability in the cortical activity on a single trial level. This high prediction-accuracy, gave us confidence that the model was well-suited to describe the neural activity measured during the performance of our task.

We found that information related to the behavior of the mice explained a large fraction of the cortical activity. Particularly, information about the running speed of the mice and further movements of the animals captured using videography, explained the highest fraction of cortical responses. This observation is in line with previous reports employing similar modeling approaches (Musall et al., 2019; Orsolic et al., 2021). Even though, the behavioral decisions of the subjects represent the main drive to perform the task, choice regressors explained a comparatively small fraction of the cortical activity, estimated using the unique explained variance metric. However, this was likely caused by us modelling the impact of choices over the entire trial separately from the individual lick responses of the animals during the response period. Therefore, choice-related signals during the response period could likely still be modelled accurately, based solely on the presence of lick-responses. However, this still indicates that choice-related signals did not dominated the cortical activity on mice outside of the response window; or alternatively that choice-related activity could be predicted based on movements of the subjects. Notably, while conducting the experiments, we had the impression that subjects tended to display a directional running behavior, seeming to lean

toward the target stimulus side, even before the response period. However, this observation requires further analysis to be conclusive.

With this modelling approach we can only predict cortical activity that is related to features that are externally expressed by the subjects. While our model has access to behavioral parameters that have previously been used to estimate the behavioral state of mice, such as the running speed and even the pupil diameter as part of the videography (Shimaoka et al., 2018), there are likely fluctuations of the behavioral state or even thought processes, that do not have an external readout and cannot be measured with our current methods. Such internal processes likely account for a large fraction of the variance in neural activity that our model cannot predict at the current time.

Since the model explains the cortical activity in the form of individual, stereotypical event responses, we could now extract the sensory response kernels for different modalities. This allowed us to address if multisensory responses could be described as a linear combination of visual- and the tactile stimulus response or alternatively, if there is a non-linear multisensory response component when presenting both stimuli at the same time. To test this hypothesis, we modelled the sensory responses in multisensory trials as the linear combination of the visual and the tactile responses (as found in the unisensory trials), combined with an additional multisensory interaction regressors. This additional regressor allowed the model to account for neural activity that deviated from the linear combination of the individual unisensory responses. We studied to which extent each of those regressor groups contributed to the model predictions by determining the unique explained variance (ΔR^2). This ΔR^2 represents the most conservative estimate of what each regressor contributes to the model predictions (Musall et al., 2019). The fact that we found unique explained variance for visual and tactile regressors in their respective primary sensory cortices as well as known association areas, demonstrates that this approach recapitulates known features of evoked cortical responses of these modalities. Furthermore, this showed that linking the unisensory regressors and the non-linear multisensory regressors in multisensory trials did not impact these representations in the model. The fact that we also found unique explained variance for the non-linear multisensory regressors, demonstrates that there are non-linear interactions between the visual and the tactile sensory responses in multisensory trials. This means that multisensory responses in the widefield data are not simply explained as the

combination of visual and tactile responses, excluding the possibility that neural populations that respond to either visual or tactile stimulation in overlapping sensory areas are strictly separated. Further, the map of unique explained variance for the additional multisensory information had clear structure, with the highest contribution bilaterally in MM, RL and V1. This makes it unlikely that the additional multisensory information reflects unspecific responses, for example related to the arousal of the animals. Instead, these regressors seem to capture specific patterns of cortical responses, deviating from a linear integration in multisensory trials.

Since in this step of the analysis we were only studying if there is unique information in the non-linear multisensory regressors, we could not distinguish if these reflect additional superadditive activity, subadditive responses or even multisensory depression (Stein & Stanford, 2008). For this, we needed to study the multisensory β-weight kernels used to model the cortical responses, which could have indicated one of two alternatives. On the one hand, there might have been individual neurons responding to both the visual and the tactile stimuli that displayed activity in multisensory trials that was lower than the linear sum of the individual stimulus responses, as measured using the widefield Ca^{2+}-imaging. This would take the form of negative multisensory β-weight kernels, indicating either subadditive multisensory integration or even multisensory depression, depending on the magnitude of the β-weights compared to the unisensory responses. On the other hand, if we would have found additional activity represented by the non-linear multisensory kernel, this would instead indicate either superadditive multisensory responses in individual neurons or a recruitment of additional neurons that only respond to multisensory stimulation. This would be reflected in positive β-weights of the additional multisensory kernels. To answer this question, we inspected the β-weights of the uni- and multisensory kernels.

3.2.5 Sensory responses isolated using the linear encoding model

To understand how the encoding model used the additional multisensory information to improve the predictions of cortical activity, we studied the β-weight kernels that were used to model the neural responses to individual sensory cues. For our model, we chose to represent the very first sensory cue in a sequence on either the left or the right side, using a

different response kernel than for the subsequent cues. The main reason for this is the commonly observed phenomenon of sensory adaptation, where sensory responses are strongest to the onset of a stimulus sequence, but then adapt and display lower levels of sustained activity (Adibi & Lampl, 2021; Keller et al., 2017). Therefore, two distinct sets of regressors were used to describe the sensory responses in each stimulus condition. First, a stimulus onset kernel, which was aligned to the presentation of the first sensory cue in the sequence, ranging over the remaining stimulus and delay period. Second, subsequent cue kernels that capture stereotypical response around the individual cue presentations.

To investigate if neural activity in response to multisensory stimuli could be characterized as either subadditive or superadditive in specific cortical areas, we investigated the multisensory β-weights. Both the onset kernels and the subsequent cue kernels consistently displayed positive multisensory β-weights. Since these multisensory responses represent additional activity after accounting for the individual responses to the visual and tactile stimulus components, this finding demonstrates the existence of widespread superadditive multisensory responses in the cortex. This suggests that in mice performing multisensory trials, either neurons responding to the individual modalities display supperadditive multisensory responses or that additional neurons are recruited in multisensory trials that are unresponsive in unisensory conditions.

To better understand these responses, we also investigated the individual visual and tactile β-weight kernels. These sensory kernels predominantly displayed activity in the respective primary sensory cortices and association areas. However, these responses were also accompanied by inhibition of activity in regions corresponding to the other, non-matching sensory modality.

Inhibition of activity even in primary sensory areas by stimuli of a different modality, have previously been reported in both anesthetized and awake mice (Iurilli et al., 2012; Lohse et al., 2021). However, the underlaying mechanisms can range from direct cortico-cortical projections (Iurilli et al., 2012) to subcortical circuits involving many different brain structures (Lohse et al., 2021), depending on the modalities involved. Cross-modal inhibition is therefore not cause by a single universal mechanism, but rathe comprises a number of mechanisms that each serve specific biological and computations functions in the processing of sensory information. Furthermore, this cross-modal inhibition is not exclusive to mice. Laurienti et al.,

2002 found similar results in humans using functional magnetic resonance imaging when presenting auditory and visual stimuli. Crucially, while cross-modal inhibition was evident under unisensory stimulation, it was absent in multisensory trials. This conditional inhibition has been linked to sensory attention (Laurienti et al., 2002; Macaluso et al., 2000), in specific cases mediated by the multisensory parietal association cortex (Macaluso et al., 2000). Our findings in the context of a behavioral tasks are well in agreement with these reports.

According to these findings, the cross-modal inhibition we found in unisensory trials, could reflect a form of sensory attention; a mechanism to prioritize cortical resources, where responses to one sensory modality suppress activity in other sensory cortices that are currently not informative for the behavior. Therefore, cross-modal inhibition could aid mice in unisensory trials, but might also hinder multisensory trial performance. The superadditive multisensory responses we found, could therefore indicate the absence of this inhibition or a counteracting mechanism that prevents the otherwise mutual inhibition of sensory responses in primary sensory cortices.

This could also explain the high amounts of unique explained variance by the multisensory regressors we found in the primary visual cortex, in addition to RL and mFC. The tactile stimulus kernels displayed a selective inhibition of V1, ipsilateral to the tactile stimulus. These responses most likely could not be sufficiently counteracted in multisensory trials by the activity predicted for the visual stimulus and therefore would have to be accounted for using the additional multisensory response kernel. These findings suggest that mechanisms of sensory attention could be a key difference between unisensory and multisensory trial performance.

Comparing sensory responses in area RL and the mFC, we saw that visual responses were stronger in area RL and tactile responses were stronger in MM. This gives us an indication that these regions might have differential roles in the integration of sensory information of these two modalities. To study the individual roles of these areas and how their activity might contribute to the capability of mice to solve the task, we next investigated which areas displayed activity that reliably reflected the target stimulus side.

3.2.6 Visual and tactile evidence representation in RL and MM

To understand how the sensory information could guide the behavior of the mice, we first determined which cortical regions displayed activity that reliably reflected the target stimulus side. For this purpose, we studied which cortical activity patterns could identify the target-side in discrimination trials. We chose to focus on discrimination trials, because in these trials the current sensory input at any given time in the stimulus period, could reflect the target-side as well as the distractor-side. The eventual target can therefore only be determined by integrating sensory evidence over the entire stimulus period. In contrast, in a detection trial the target-side is evident at any time in the stimulus period and no accumulation of sensory evidence would be required.

Many cortical areas reflected the target stimulus side. These target-selective responses could especially be found in the primary and secondary sensory areas as well as the retrosplenial cortex and mFC. Having seen that the RL and MM display the highest overlap in visual and tactile trial responses and particularly high superadditive multisensory activity, we compared how these two regions represented visual and tactile information in the unisensory trials and how this translates to multisensory trials.

Over the course of a trials, RL and MM displayed AUC-values above chance-level over the entire stimulus and delay period of trials. Surprisingly, we found that the AUC-values in these regions already reach relatively high values shortly after the stimulus onset. Since we investigated the neural activity in discrimination trials, sensory evidence would have to be accumulated over the entire stimulus period to reliably identify the target-side. Therefore, the target should not have been evident for the subjects this early in the stimulus period. The best explanation for this observation is that in the process of balancing the number of left and right as well as correct and incorrect trials, we introduced structure into the sequences of target and distractor cues of the trials included in the analysis. Subsequent studies could attempt to improve this procedure of balancing the number of included trials, in these regards. In the context of this study, we focused the analysis on the delay period of the task immediately preceding the behavioral responses.

In the delay period, activity in area RL reflected the visual target more reliably than the tactile target. The opposite was the case for MM, where activity reflected the target more accurately

in tactile trials over the entire stimulus and delay period. This suggests that, even though both of these regions displayed sensory response in all stimulus conditions, these two regions in particular reflect sensory evidence differently, depending on the sensory modality. This representation of tactile evidence in MM is in line with a previous study on a tactile discrimination task, which included an extensive response delay period (Chen et al., 2017). Here, MM also reflected the sensory evidence acquired during the stimulus presentation, and maintained this information until the response behavior could be executed. Our results indicate that this area, which has also previously been linked to general short-term memory functions (Gallero-Salas et al., 2021) does more faithfully reflects tactile sensory evidence, compared to visual. In visual trial, instead the secondary visual cortices including area RL reflected visual evidence more reliably.

To test how individual cortical areas reflect the target stimulus side, we computed the AUC-values as a normalized measure of target-selectivity for cortical responses within individual sessions and then averaged these over sessions and mice. This, allowed us to identify regions that displayed target-selective activity, pooled over all discrimination difficulties. To further study this relationship, it would be possible to narrow this approach down and compute neurometric curves (Gold et al., 2010), that reflect how well the activity in a given cortical region reflects the target-side at specific task difficulties (in the form of specific differences in the number of target- and distractor cues) for the individual stimulus condition. However, this was not achievable with our current approach, due to the low number of choice-balanced trials that are available within an individual session once trials are further subdivided by modality and target-distractor cue difference. A subsequent study could use the existing dataset and attempt to pool choice-balanced trials for the individual difficulties over sessions and even mice to compute the AUC over the entire compiled dataset. However, due to the computational resources required for such an analysis, it is unlikely that this is achievable on a pixel-by-pixel basis as presented in this thesis, pooled over difficulties. However, an alternative approach could be to focus on specific areas of interest, such as RL and MM and selectively pool activity averaged over such region of interest to simplify the analysis. As a result, one would obtain AUC-values for specific cortical regions over all task difficulties for the individual stimulus conditions. Based on our current findings, we suspect that the neurometric function of RL would match the psychometric curve of task performance

more closely in visual trials, than that in tactile trials. While the opposite would be expected for MM, which would likely reflect tactile trial performance more closely. However, AUC-values found in RL and MM during multisensory trials were not significantly higher than those in the individually preferred unisensory conditions. Therefore, it would be unlikely that the increase performance in these trials could be explained using activity in either region alone.

Since both RL and MM reliably reflected the target-side in multisensory trials, subjects likely combine information from both these areas to improve their decisions in multisensory trials. In this analysis, we investigated the activity of each region in the cortex in isolation, only in reference to the opposite hemisphere. Using information from multiple cortical regions simultaneously, more elaborate analytical approaches could likely improve the accuracy of decoding the target stimulus side from cortical activity. Such an approach could further explore the relationship of RL and MM, and how information from these two areas could be combined to guide the behavioral decision in multisensory trials.

Such studies of the existing dataset would grant even deeper insights into the underlying neural circuitry, and how these sensory responses are used to form decisions and guide the behavior. Lastly, to investigate if the same areas that represent the target-side also display choice-related activity and animal responses, we studied how the cortical activity relates to the responses of the mice.

3.2.7 Choice-related activity in the lateral frontal cortex

Finally, to study how the mice translate sensory evidence into response behaviors, we trained a choice decoder that predicted the upcoming responses of subjects in individual trials, based only on the current cortical activity. We chose this approach of training a decoder with L1 regularization to identify the most informative components of the cortical activity (Musall et al., 2023). The high accuracy of >80% of correct response-predictions at the end of the delay period demonstrated that mice display reliable patterns of cortical activity that are predictive of their upcoming responses. Furthermore, the decoder's choice prediction accuracy continuously increases over the course of the stimulus and delay period up to the animals' responses. This indicates that mice do not form their decision at a specific time in the trial, for example during the delay shortly preceding the response execution, but instead seem to

plan and continuously update their future response from the stimulus onset on. The areas that most reliably predicted the animals' responses were ALM, tjM1 and tjS1. Similar patterns of choice-related cortical activity have been reported previously in an auditory task, using the same analytical approach (Musall et al., 2023). Activity particularly in these areas has been linked to the preparation and execution of directional lick responses, as commonly used in rodent behavior tasks (Esmaeili et al., 2020; Li et al., 2015). Here, ALM seems to represents more latent, contextual features that influence the motor preparation in line with its role as a premotor area (Esmaeili et al., 2020; Guo et al., 2014; Xu et al., 2022). In contrast, tjS1 and tjM1 instead reflect features more immediate to the execution of the lick responses (Esmaeili et al., 2020; Xu et al., 2022). ALM therefore represents the cortical area closest to the decision-making process, as previously indicated in tactile delayed lick-response tasks (Chen et al., 2017; Esmaeili et al., 2020; Li et al., 2015). Accordingly, studies on behavior tasks such as the one presented in this thesis, requiring mice to lick at water spouts to indicate their decisions have commonly found choice-related signals in the secondary motor cortex of mice (Chen et al., 2017; Esmaeili et al., 2021; Guo et al., 2014; Mohan et al., 2022; Musall et al., 2023). In contrast, in a behavior task for mice that features a direct sensory-motor response, instead of a delayed response, only a low fraction of neurons was found in the secondary motor cortex displaying preparatory activity (Steinmetz et al., 2019). Accordingly, it has been speculated that central computations of the decision-making process in such a task could instead take place in subcortical areas (Zatka-Haas et al., 2021). This emphasizes the importance of considering the behavioral context of the decision-making process. Having subjects express their decision in the form of delayed response behaviors in this study, gave us the opportunity to investigate sensory integration and decision-making processes in cortical circuits.

Subsequently, we trained separate decoders to predict responses in the individual stimulus conditions, to investigate if there is evidence for modality-specific choice-related regions. Here, once again we found similar patterns of high β-weights in ALM, tjM1 and tjS1 in all stimulus conditions, indicating that the same regions reflect the choices of the mice across modalities. This suggests that sensory evidence, which we found to be represented in distributed modality-specific cortical regions, converges and is used modality-unspecific in the ALM and further reflected in tjM1 and tjS1. One open question in this context that remains at this point is if the visual and tactile sensory information is still separately represented inside

ALM and then locally translated into the licking decision of the mice, or if these two streams of sensory information that are still separately represented in RL and MM are integrated at another intermediate stage, such as the parietal cortex, and ALM already receives sensory-unspecific information.

4 Conclusions

In this study we investigated how the cerebral cortex integrates multisensory information and how this information guides behavioral decisions. To achieve this, we first established a novel visuotactile evidence accumulation task for head-fixed mice. Mice were able to reliably learn this task in the individual unisensory and the multisensory conditions. Here, we found that mice displayed significantly higher performance in multisensory trials over a broad range of task difficulties. This gave us the opportunity to study the cortical activity of mice performing this task. With this, we were able to identify regions in the posterior parietal cortex and the medial frontal cortex, where sensory responses overlapped in visual and tactile trials, indicating a high potential form multisensory integration. Using a linear encoding model, we were able to isolate the sensory responses from other task-aligned activity, relating for example to movements of the mice. This approach revealed signs of widespread superadditive multisensory integration in the cortex of mice performing multisensory trials.

Further, we found that in unisensory trials increased activity in the corresponding sensory cortices was accompanied by an inhibition of activity in the non-stimulated sensory cortices. Similar cross-modal inhibition has previous been reported in humans presented with unisensory stimuli, but was absent when both sensory modalities were presented simultaneously (Laurienti et al., 2002). Such a mechanism of inhibition only present in unisensory trials, would in part explain the superadditive multisensory responses we found. Such mechanisms of either conditional cross-modal inhibition, as a form of sensory attention or superadditive multisensory integration, likely aid mice in performing the task in the different stimulus conditions.

To investigate how sensory information is integrated by mice to solve the behavior task, we computed how reliably each cortical region reflected the target-side over the course of a trial.

Here, we found that the medial motor cortex more reliably reflected the target-side in tactile trials, while the secondary visual cortices preferably reflect visual targets. In multisensory trials, both regions represented the target-side with a reliability comparable to the prefered unisensory condition. This additional, redundant information likely increased the certainty of mice performing multisensory trials, leading to the improved multisensory performance.

To investigate if these areas also reflect the behavior decisions of the mice, we used a choice-decoder to predict the upcoming responses of mice. This decoder primarily used information in ALM, tjM1 and tjS1 to predict the behavioral responses. These areas are linked to the planning and the execution, particularly of lick responses (Xu et al., 2022). However, no clear modality specific differences in the choice predictive areas were found.

Our findings suggest that cross-modal inhibition found in unisensory trials is absent in individuals performing multisensory trials, allowing to integrate information from both sensory modalities. Visual and tactile information is accumulated preferentially in sensory-specific regions of the cortex, over the course of a trial. This information then converges in the secondary motor cortex to form decisions and plan responses modality-unspecific. These findings outline a general circuitry by which sensory information is integrated across modalities to guide the behavioral decisions and outlines targets for future studies of cortical multisensory integration.

5 Materials and Methods

5.1 Animals subjects

All animal procedures were performed in strict compliance with the EU directives 86/609/EWG and 2007/526/EG guidelines for animal experiments and were approved by the local government (Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein Westfalen, Recklinghausen, Germany). Data presented in this thesis was obtained from four female mice with transgenic background. These mice expressed the genetically encoded calcium indicator GCaMP6s in excitatory neuron of the telencephalon (for reference see Wekselblatt et al., 2016). Mice were obtained by crossing two individual mouse lines. The first mouse line (DBA-Tg(tetO-GCaMP6s)2Niell/J; JAX 024742) resulted in GCaMP6s expression

under control of tetracycline-responsive regulatory element (tetO). The second mouse line (Cg-Tg(CaMK2α-tTA)1Mmay/DboJ; JAX 007004) expressed the tetracycline-controlled transactivator protein (tTA) under the promotor of the forebrain-specific calcium-calmodulin-dependent kinase II (CaMK2α; Mayford et al., 1996). Mice were genotyped to confirm presence of CaMK2α-tTA and tetO-GCaMP6s (heterozygous or homozygous). Primers used for genotyping of tetO-GCaMP6s were 5´-GTG TCA GAG GTT TTC ACC GTC-3´ (forward). Primers for CaMK2α-tTA were 5´-GAC CTG GAT GCT GAC GAA GG-3´ (forward) and 5´-GCA GCT CTA ATG CGC TGT-3´ (reverse). For the learning curves presented in Figure 5B, additional eight male C57BL/6J mice were included.

All mice were housed in fixed groups of 2-3 animals; preferably siblings. In their home cages, mice were provided with environmental enrichment and had access to food *ad libitum*. Cages were kept in a ventilated cage cabinet, which provided artificial illumination for a 12 h:12 h-reversed day-/night cycle. Mice were trained during this simulated night-period.

Surgeries were conducted when the mice reached an age of 8-12 weeks. Behavioral training was started after a one- to two-week recovery from surgical procedures. Training took place only on week days between 8 am and 6 pm. To motivate animals to participation in experiments, mice were kept water-restricted for the days of training. Water deprivation was suspended during weekends. Throughout the water-restriction, mice were monitored and scored on a daily basis. Over the course of a daily training session, mice received unlimited access to water in form of rewards for correctly trial responses. If mice drank less than 1 ml during the experiment, they were later supplemented with the respective missing amount of water absorbed in oats. The stop criterion for experiments, when the weight of a mouse dropped below 20% of its reference body weight was never reached throughout experiments.

5.2 Surgical procedure

Four female transgenic mice and eight male C57BL/6J mice were operated at an age of 8-12 weeks. The surgeries consisted of skull clearing and head-holder attachment. Anesthesia was induced using isoflurane (3% in 100% O_2). Throughout the surgery, anesthesia was maintained at ~1.5% isoflurane in oxygen. Diluted Buprenorphine was injected (subcutaneous 0.1 mg/kg) as a general analgesic agent at the beginning of the surgery and again at the end, according

to the animal license. The hair over the skull was removed using epilation cream and cleaned using phosphate buffered saline (PBS).

Bupivacaine (0.25%, subcutaneous 0.08 ml in total per mouse) was injected under the scalp, as well as the bridge of the nose and the region of the ear bars as an additional local analgesic. The skin over the skull was cut with an incision along the midline, pushed to the sides and attached to the skull using tissue adhesive (Vetbond, 3M). A custom-designed head-attachment was fixed using dental cement either by first applying Metabond (Parkell) and then Ortho-Jet (Lang Dental) or directly using Ortho-Jet dental cement. The cleaned intact skull was cleared by applying a thin layer of cyanoacrylate (Zap-A-Gap CA+, Pacer Technology). Mice were returned to their home cage on a heating pad for recovery. For 3 days following the surgical procedure, the drinking water was supplemented with Enrofloxacin (0.16 ml in 160 ml drinking water, anti-bacterial) and Buprenorphine (5 ml in 160 ml drinking water).

To prevent systemic infection, an antibacterial therapy, starting one day before the surgical intervention and lasting for three days after the surgery, was given by supplementing the drinking-water with Enrofloxacin (1ml/L, Baytril®). As a further analgesic treatment, the drinking-water was supplemented with Buprenorphine (5 ml in 160 ml drinking water) for three days following the surgery, in addition to the Enrofloxacin.

5.3 Behavior

5.3.1 Visuotactile evidence accumulation task

Individual sessions of the behavior task were comprised of interleaved visual, tactile and multisensory trials. Each of those stimulus conditions followed the same general schema (Figure 4A). A trial started with a 3 s stimulus period, in which mice were presented with sequences of sensory cues on their left and/or right side. Up to 6 sensory cues were presented at fixed time-points every 0.5 s throughout the stimulus period. In detection trials the maximum number of 6 cues were presented in absence of any distractors. In discrimination trials the presence of a cue in each time-bin was drawn randomly, with a probability of 0.7 on the target-side and 0.3 on the distractor-side, the same probabilities used in a previous study (Scott et al., 2015). In case the target sequence by chance contained fewer cues than the distractor sequence, these two were switched so that the target-side never had fewer cues

than the distractor-side by definition. After the stimulus period a 0.5 s delay was used. This was followed by the 2 s response period during which mice were presented with two spouts on either side at which mice had to indicate their decisions in the form of licks to receive a small water reward. Between the response period and the next stimulus presentation there was a 3.5 s inter trial interval (ITI). To encourage correct response behaviors, a time-based punishment was included, which increases the ITI, either by 1 s if mice didn't respond in the current trial or by 2 s if mice responded incorrectly. Mice were able to run on a custom 3D-printed wheel. The setup was controlled using a microcontroller running custom code (see section 5.3.5.1), which in turn was controlled using custom python code running on a personal computer (PC; see section 5.3.6).

5.3.2 Visual stimuli

Visual stimuli were presented on two monitors (LG 23MB35PMF, 60 Hz refresh rate) mounted on the left and the right side of the mice. Monitors were in direct contact to fill the visual field of mice as well as possible and were positioned so that the eye of a mouse was closest to the center of the corresponding monitor at a distance of 18 cm. This resulted in an effective angle of ~60° in the horizontal between the two monitors. Visual cues were designed to be similar to an individual cycle of a moving grating, with a spatial frequency of 0.018 cycles per degree (cpd), to fill precisely half the width of the monitor, drifting at a temporal frequency of 2 Hz to match the 0.5 s cues presentation period. Cues consisted of a single cycle of a sine-wave, with its magnitude scaled by a factor of 1.25 and clipped to values within the range [-1, 1]. Here, a value of -1 resulted in black color, a value of 0 results in gray and a value of 1 in white. This preserved the general character of the grating while also giving the stimulus a slightly higher contrast. For the presentation, values between -1 and 1 were mapped into 24-bit color space. Cues were presented on both sides of the mice drifting from nasal to temporal, with the dark component of the cue leading, followed by the bright component, as shown in Figure 4A.

5.3.3 Tactile stimuli

Tactile stimuli were presented in the form of individual airpuffs, delivered via stainless steel tubes (16G gavage needle, straight, 125 mm long, AliExpress) mounted on top and in front of

the animals. Tubes were mounted at an altitude angle of 45° and an azimuth angle of 10° directing air current inwards toward the mice. Spout positions were carefully calibrated to only deflect the whiskers, close to the whisker pad. Individual airpuffs were generated using short (20 ms) digital pulses generated by the microcontroller (see section 5.3.5.1), triggering an opto-coupled relay (AQZ105, Panasonic) switching a supply current (24 VDC, FSP-1243, 72 W, Voltcraft), in turn opening a solenoid valve (VDW22LA, 24 VDC, SMC). This valve controlled the flow of compressed air (0.03 bar), resulting in a brief and subtle pulse of air flow.

5.3.4 Water reward delivery

The reward delivery was controlled with by the microcontroller (see section 5.3.5.1). Here, a digital pulse was generated to trigger an opto-coupled relay (AQZ105, Panasonic) switching a supply current (24 VDC, FSP-1243, 72 W, Voltcraft), which in turn opening a solenoid valve (VDW22LA, 24 VDC, SMC). The length of the digital pule controlling the open-duration of the individual valves was regularly calibrated to ensure a stable and equal reward size for both valves.

5.3.5 Behavior setup control

To control the behavior setup with high precision while allowing appropriate control by the user, a two-layered design was used. The frontend was implemented in python (see section 5.3.6), including a graphical user interface for the user. This program controlled the general task execution, determining the individual trial settings. The backend was built around a microcontroller that managed the presentation of individual trials with extremely high (µs) temporal precision (see section 5.3.5.1).

5.3.5.1 Microcontroller software

The main components of the behavior setup were controlled using custom software running on a microcontroller (Teensy 3.2, PJRC) in the following referred to as teensy. The main functions of this code are the following:

1. Initiate visual stimulus, then align and control tactile stimulus presentation
2. Control the stepper motors for the lick spout presentation
3. Detect lick responses using capacitive touch sensors connected to the spouts

4. Initiate water reward delivery if applicable
5. Report registered responses to the control PC to be saved

With these functions in mind, custom code was written using Arduino IDE in combination with the Teensyduino add-on (PJRC). The teensy code was based on a fixed sequence of 5 states that defined a given trial: Trial initiation, Stimulus presentation, delay, response period and ITI.

During the trial initiation the teensy read the defining parameters of the upcoming trial from the controlling PC. This was implemented using serial USB communication, where a fixed sequence of bytes was read, defining the individual trial parameters. These parameters included, inter alia, the desired lick spout positions, the opening durations of the airpuff- and water valves, and the individual sequences of tactile cues. Once the new trial was defined, the lick spouts were moved to the “out”-position, a position outside the reach of the mice, as preparation for the later presentation in the response period. This was followed by a predefined 2 s baseline interval. At the end of this interval, the teensy code transitions into the stimulus period.

At the beginning of the stimulus period, a signal byte was sent to the control PC, requesting the presentation of the visual stimulus. The PC then presented the visual stimulus accordingly. With the presentation of the first stimulus frame, a photodiode indicator appeared in the corner of the monitors indicating the start of the visual stimulus. This was the case independent of the stimulus condition of the current trial. For example, in a tactile trial the stimulus program still continued presenting a gray image, indistinguishable from the image outside of the stimulus period, with the addition of the photodiode indicator. This signal was read out by a photodiode and allowed to align the visual and tactile stimulus presentation, by compensating for the delay caused by the visual stimulus presentation using monitors. This photodiode signal resets the current trial clock to the beginning of the stimulus period and if desired, tactile cues are presented at their defined time-points.

At the end of the stimulus period, the code transitioned to the delay state. In this state, no further functions were executed other than waiting for the beginning of the response period. At the beginning of the response period, the lick spouts are moved from the “out”-position to the “in”-position, where the mice can reach the spouts with their tongue.

The behavior of the setup during the response period was defined by two parameters, that were set during the trial initiation: The first parameter defines if only the target spout

(referred to as “single spout”) or both spouts are supposed to be presented. The second parameter indicates if the reward was supposed to be automatically delivered (referred to as “auto-reward”) or only upon correct lick responses. For details about the lick-detection see section 5.3.5.2. Depending on the combination of these two parameters the setup behaves as follows:

Case 1: Single-spout was on and auto-reward was on

If mice licked, then the reward was immediately delivered, otherwise it was automatically dispensed 0.5 s into the response period.

Case 2: Single-spout was on and auto-reward was off

Reward was dispensed only if the mouse licked, otherwise no reward was dispensed.

Case 3: Single-spout was off and auto-reward was on

This behavior further depended on the response of the mouse:

- If mouse licked the correct spout: reward was delivered (this case was indistinguishable from a regular self-performed trial, from the perspective of the mouse).
- If the mouse did not respond or lick the incorrect spout, responses were ignored and the reward was automatically delivered on the target-side 0.5 s into the response period.

Case 4: Single-spout was off and auto-reward was off

This case represented the standard self-performed trial, where mice had to lick the correct spout to receive a water reward.

Further, if a mouse responded incorrectly in a self-performed trial, both spouts were immediately moved to the “out”-positions, signaling the incorrect response. However, if the mouse responded correctly or alternatively after the automatic reward delivery, only the incorrect spout was moved out, so the mouse could collect the reward. This behavior of the setup was a helpful tool during the early training period and commonly used during the first few trials of every session to motivate mice.

Once a response of the animal was registered or the auto-reward was delivered, these signals were transmitted to the controlling PC over serial USB to be saved. At the end of the response period, the spout position was recalibrated (see section 5.3.5.3). With this, the inter-trial-interval (ITI) began. In self-performed trials, the duration of the ITI depended on the response

of the mice (see section 5.3.1). Once the ITI was over, the code returned to the trial initiation state to start the next trial.

5.3.5.2 Lick-detection

Throughout the trial, capacitive signals from the two lick spouts were monitored using the "touchRead"-function of the teensy. The sensors were calibrated over a 2 s interval at the beginning of the session in absence of licks, to determine an appropriate signal threshold. This threshold was 5 standard deviations above the signal acquired during the calibration. However, when desired the threshold could also be adjusted manually using the computers graphical user interface, which signaled rather to increase or decrease the individual threshold used by the teensy code by a predefined value. The presence of lick signals above threshold was then converted into a binary signal that was recorded by the DAQ (data acquisition) card of the PC. Responses were registered if a mouse licked a given spout twice out of the first three contacts.

5.3.5.3 Lick spout position control

The spout positions were controlled using stepper motors (NEMA 8, Pololu) in combination with driver boards (DRV8834, Pololu). For this purpose, lick spouts were mounted on top of stepper motors using custom 3D-printed parts and motors in turn were fixed in the setup using another custom 3D-printed part. Motor positions were recalibrated after every trial by moving spouts out until they made an electrical contact to a reference point, which was detected by the teensy to update the current motor position. This was a necessary step due to the nature of stepper-motors, which operate in relative steps but don't have an absolute position readout. With this, it could be ensured that the spouts were always presented precisely at the desired positions.

5.3.6 Python based behavior software

The central behavior software, that controlled the behavior task, was implemented in Python run on the control PC (personal computer). The main functions of this program were to define the sequence of trials that were presented, communicate with the image acquisition software, present the visual stimuli, record setup- and behavior-related signals and save this

information. To be able to control this software, a GUI (graphical user interface) was created implemented in PyQt5, that allowed the user to define parameters of the task and control the behavior of the setup.

When started, the python behavior software first initialized the visual stimulus software using the python subprocess library. This routine continuously communicated with the behavior software over TCP (Transmission Control Protocol) implemented using the socket library and upon request displayed the prerendered visual stimuli using the PsychoPy library. Upon initialization, this software established a TCP connection to the behavior software and opened two fullscreen windows on the setup monitors. These monitors now displayed the neutral gray image with a black photodiode indicator, which was always present outside of the stimulus period. This approach of running the visual stimulus in an independent python instance was necessary due to limitations of PsychoPy, only being able to run on the main thread of a python instance, which otherwise would have interfered with the GUI. Once the visual stimulus software was initialized, the connection to the teensy microcontroller was established. And the startup routine of the setup was initialized. This included the calibration of the photodiode, the calibration of the lick-detection, calibration of the stepper motors and subsequent positioning at the default “in”-positions of the spouts.

After those startup routines ran, the GUI opened for the user. Here, the user could now enter the mouse ID and specify the general session type. Session types are defined by the presented sensory modality conditions and the mode of stimulus presentation (detection vs. discrimination task). Upon selection, default session settings are applied. Individual parameters concerning, for example, the probability of discrimination trials, probability of single-spout trials (assisted trials where only the target spout is presented; excluded from analysis) could then be changed if needed by the user. These and further settings, with the exception of the mouse-ID, could be adjusted flexibly throughout the session.

Once the session was started by the user, the mouse-ID as well as a time-stamp were transmitted to the image acquisition software, to simplify merging the individual recordings after the session. With this, the first trial is initiated by sending the required trials parameters to the microcontroller. Over the course of a trial digital and analog signals about the timing of the behavior task, the synchronization trigger for the image acquisition, as well as the running speed of the mice and the lick responses are recorded using a data-acquiring card

(PCIe-6323, National Instruments). This information, together with the responses of the animals are stored in the h5-file format for later analysis.

5.4 Training

The behavior training was started on the detection condition. In this condition the maximum number of 6 cues is presented on the target-side in the absence of distractor cues. To not overwhelm the mice, they are initially only trained on visual detection trials until the mice reach stable performance. This performance value is calculated as the fraction of correct responses over all responses given. If this performance either exceeded 75% for three consecutive sessions or 90% in a single session, mouse has successfully learned the task in the visual detection condition. Once this was the case the tactile detection condition was introduced. In the following two sessions, 50% of trial remained visual detection trials and the other 50% were tactile. After those first two transition sessions, the fraction of tactile trials is increased to 100%. After the mouse had learned to perform the tactile condition as well, the mouse is presented again a single session with equal amounts of visual and tactile detection trials. After this, the multisensory condition is introduced at an equal fraction to the individual unisensory conditions (33.3%). Since there is nothing inherently new to learn for the mice, they perform this condition very well from the first session on. At this stage the mice have learned the full multisensory task in the detection conditions. In the final stage of training discrimination trials, consisting of a variable number of target- and distractor cues are introduced. Initially, the fraction of discrimination trials is being kept low and slowly increased over time as the mouse displays higher performances also in the discrimination conditions. With this training regime a high number (n=12) of mice could be trained to perform this multisensory accumulation of evidence task.

5.4.1 Bias- correction

Mice often display side-biases in the form of responding more often toward on side compared to the other. To control these biases a number of measures were taken. The water-valves dispensing the rewards were carefully calibrated in regular intervals to ensure equal reward sizes. The sequence of left and right target trials was generated in a pseudo-random manner.

Further, if by chance three consecutive trials were presented with the target on the same side, the probability to present the next target again on this side is reduced to 25%.
Finally, to mitigate any remaining side-biases, the performance in the ten most recent left and right trials was tracked and compared with one another to detect a bias. Based on the extent of the response imbalance, the position of the spouts during the response period was automatically adjusted, within adjustable ranges of possible spout position. This was realized in such a manner that the preferred spout became more cumbersome to reach until the proportion of left and right responses equalized again. With these measures the proportion of left and right responses was kept roughly equal, independent of the trial conditions.

5.5 Behavior videography

The movements of subjects were monitored using two CMOS-cameras (acA1920-155um, Basler), controlled by the Image acquisition software (see section 5.6.2). One camera captured the facial movements of the individuals (front view) and one the body movements (side view). During the session, videos were saved as raw binary field in the TDMS-format and afterward compressed using H.264 encoding with a compression factor of 17, implemented in python using “ffmpeg” (Tomar, 2006).

5.6 Widefield imaging

The widefield Ca^{2+}-imaging was based on an inverted tandem-lens macroscope (Ratzlaff & Grinvald, 1991), with a 50 mm objective (Nikon AF-D 50mm f1,4; Foto Erhardt) at the bottom, above the subjects and a 85 mm objective on top, in front of a sCMOS (scientific complementary metal-oxide semiconductor) camera (Edge 5.5, PCO). A schematic of this macroscope is shown in Figure 13A. Frames were acquired using custom software (see section 5.6.2) at an imaging rate of 30 Hz using 4 x 4 spatial binning resulting in a resolution of 640 x 540 pixels. Frames were alternatingly illuminated using a blue LED (470 nm, M470L3, Thorlabs) and a violet LED (405 nm, M405L3, Thorlabs) with a 405 nm excitation filter (#65-133, Edmund optics). LEDs were collimated using adjustable collimation lenses (SM2E, Thorlabs). Both excitation light paths were merged using a dichroic mirror (no. 87–063, Edmund optics) and reflected onto the brain using a second dichroic mirror (G381323036,

Qioptiq). GCaMP fluorescence signals were acquired using a 525 nm emission filter (MF525-39, Thorlabs) mounted in front of the camera. By alternatingly illuminating frames using either the blue or the violet LED, two separate movies at 15 Hz each are acquired. Here, frames acquired under blue illumination reflect Ca^{2+}-dependent fluorescence, while frames acquired under violet illumination represent Ca^{2+}-independent fluorescence signal (Allen et al., 2017; Lerner et al., 2015). This allowed us to compensate for Ca^{2+}-independent signal by subtracting the rescaled frames acquired under violet illumination, from the preceding blue illuminated frames, later (see section 5.6.1).

5.6.1 Preprocessing of widefield imaging data

First the individual acquired frames were registered to an average reference frame, computed over the entire session using a subpixel image registration routine, based on a discrete Fourier transform (Guizar-Sicairos et al., 2008). This was separately performed for frames acquired under blue and violet illumination, with corresponding reference frames. These reference frames were computed as the average frame over the entire sessions under the corresponding illumination. With this small movements in the image plane could be compensated.

Next, the differential calcium dependent signal was computed. For this, the Ca^{2+}-independent fluorescence signals acquired under violet illumination were rescaled and subtracted from the Ca^{2+}-dependent fluorescence, according to the approach presented by Allen et al., 2017.

The further preprocessing of the widefield imaging data was based on the dimensionality reduction approach established by Musall et al., 2019). Here, the 300 dimensions with the highest variance of the differential Ca^{2+}-signal were computed using a Singular value decomposition (SVD). This returned spatial components U and scaled temporal components V. If not otherwise stated, all analysis was performed on these SVD components. This vastly reduced the computational costs of analysis. In addition, V was high-pass filtered for frequencies greater 0.1 Hz using a zero-phase, second-order Butterworth filter. Maps in the original pixel space were then reconstructed by convolving V with U.

Imaging data was aligned to the Allen common coordinate framework v3 (Wang et al., 2020), similar to the approach by Allen et al., 2017 and Musall et al., 2019, using anatomical markers: the left, center and right location where the cortex contacts the olfactory bulbs and the

medial point on the midline at the base of the retrosplenial cortex. Alignment was confirmed for individual mice using visual field sign mapping (Garrett et al., 2014) to confirm locations of visual areas (Figure 13).

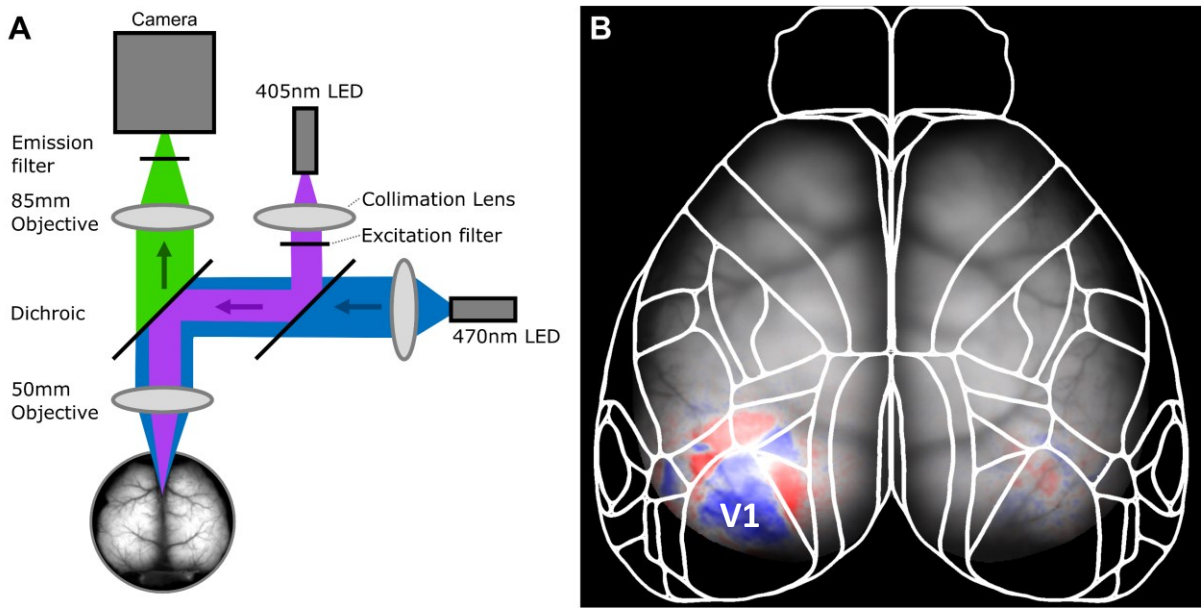

Figure 13: Widefield imaging setup and alignment of recording
(A) Schematic of whole-cortex widefield Ca^{2+}-imaging macroscope. (B) Visual field sign maps acquired according to the method established by Garrett et al., 2014 to identify the location of primary visual cortex (V1) and higher visual areas. Patches in blue indicate visual areas with mirrored representation of visual space, found in V1, as well as the anterolateral- and the anteromedial visual area. Patches in red indicate visual areas with non-mirrored representation of visual space found in the posteromedial-, rostrolateral- and the anterior visual area. Visual field sign maps were used as reference for the alignment of all widefield recordings to the Allen CCF.

5.6.2 Image-acquisition software

The image acquisition was managed by custom software written in LabView (National Instruments). This program controlled the widefield imaging, including the alternating frame illumination, as well as the acquisition of two further video streams of the mice's behavior, using two CMOS-cameras (acA1920-155um, Basler). For this purpose, all three cameras were configured to acquire individual frames upon receiving an external TTL (transistor-transistor logic)-trigger. These triggers were continuously generated using a data-acquisition card (PCIe-6323, National Instruments). With this, the PCO-camera was triggered at a fixed rate of 30 Hz,

while both behavior cameras were each triggered at 90 Hz, resulting in a fixed correspondence of every third behavior frame with the individual widefield frames. Further to switch the illumination the widefield frame, additional signal-lines were coupled to the same trigger routine that alternatingly triggered either the blue or the violet LED-driver (Cyclops LED Driver 3.6, Open ephys). Frames for each of those cameras were acquired and temporarily stored in individual ring buffers until they were saved to individual TDMS-files. For the purpose of later alignment with the behavior tasks, an additional synchronization trigger was recorded and stored together with the video recordings.

5.7 Data analysis

5.7.1 Average response analysis

To compute average responses in the individual conditions (see Figure 5 and Figure 7) session-wise averages over neural responses were computed. For this purpose, the average over the temporal components (V) was computed over all trials within as session, for example corresponding to correct, left visual detection trials and the corresponding time-bin was computed. This average V was then convolved with the spatial components (U) to generate a map of average cortical response in the corresponding condition and time-bin. Further, maps of left and right trials were averaged (independent of the number of trials that contributed to these two conditions) by inverting the hemispheres in left trials to display activity in terms of ipsi- and contralateral responses. Then, a mouse-wise average response was computed by combining all session-wise averages, weighted by the number of trials that contributed to the individual maps. Finally, these mouse-wise responses were averaged to obtain the final average neural response maps. Traces of average activity for individual regions of interest (ROIs) were computed by spatially averaging activity in a circular patch with a radius of ~375 µm.

5.7.2 Analysis of significant unisensory response overlap

To identify cortical regions that reliably response in both visual and tactile trials and therefore have a high potential for multisensory integration, first average cortical responses were computed for each stimulus condition within a session. Then the z-score was calculated over

these session-wise average responses. Then, to compute the overlap in visual and tactile responses the pixel-wise product of the corresponding z-sore movies was calculated.

5.7.3 Encoding model

The linear encoding model used in this thesis was based on the model presented by Musall et al., 2019. Our model was build based on different task- and behavioral signals to predict the observed cortical activity. For this purpose, the continuous recording of widefield- and behavior recordings, were sliced into sequences of individual trials and concatenated. Here, frames from 1 s before the stimulus onset to the end of the response period constituted a single trial. Then, a design matrix was constructed for each individual session, containing all task- and behavior related information for the model to predict the observed cortical activity (see section 5.7.3.1). The model was fit to the temporal SVD components (V) of the Ca^{2+}-imaging data using ridge regression to prevent overfitting and to allow many regressors to contribute to the model predictions. Here, the regularization penalty was estimated for each regressors individually (see Musall et al., 2019). To estimate how well the model generalized to frames within a session it was not trained on, we used ten-fold cross-validation to compute explained variance (cvR^2). Here, for each iteration of the model fit in the cross-validation process the multisensory-shuffle regressors (see section 5.7.3.1) were reshuffled using the “randperm”-function in Matlab. To compute the unique explained variance (ΔR^2), additional regressors groups were shuffle using the same procedure.

5.7.3.1 Design matrix

The design matrix used by the linear encoding model links how specific task- or movement events translate to patterns of cortical activity over time. This design matrix consisted of two part. The first part was comprised by the task related regressors, concerning the sensory stimuli, the timing of a trial, the animal’s responses and the reward delivery. This design matrix had a shape of #frames by #regressors.

Each type of event, with the exception of analog signals (running and videography), was modeled by a sequence of individual regressors. Each regressor was active in a single frame, with consecutive regressor shifted by one frame to create a sequence that could capture an event response kernel, around the time of a corresponding event. This specific sequence of

regressors, capturing the time-course of an event-related response, was copied to each individual occurrence of this event in the session.

The individual regressors groups comprising the task-related part of the design matrix are listed below (Table 1). Sensory events were modeled by two separate types of regressors: onset regressors and subsequent cue regressors. The onset regressors ranged from before the first cue presented on the corresponding side up to the end of the delay period. The subsequent cue regressors covered a narrower time-window around the individual sensory cues, excluding frames in the delay period, if applicable. Each sensory stimulus was modeled using these two types of regressors. Sensory cues on the left and the right side of the mouse were modeled separately. Visual regressors were active in both visual and multisensory trials (whenever a visual stimulus was presented), tactile regressors were active in both tactile and multisensory trials and multisensory regressors were active whenever both visual and tactile regressors were active. This allowed to capture non-linear multisensory dynamics. Furthermore, a copy of the multisensory regressors was included that was shuffled in time to destroy any correlation which the task and the neural activity. These shuffle-regressors were reshuffled with every iteration of the model fit as part of the cross-validation process of the full model (see section 5.7.3.2).

In addition to the sensory information, further regressors were included that captured neural activity over the course of entire trials. These included the relative time-course within a trial ("Time", applying to every single trial); if a mouse chose the left side at the end of the current trial ("Choice"), as well as the choice in the previous trial, to capture potential signatures of side-bias; and if a mouse received a water reward in the current trial ("Reward"), as well as the presence of a reward in the previous trial.

The second part of the model design matrix contained movement related signals. These included lick at the water spouts, the running speed and the video recordings of the mice in the setup. For this purpose, both running speed and licks were resampled to match the framerate of the Ca^{2+}-imaging data (15 Hz). Here, the running speed reflected the average speed during a given frame. The left and right licks indicated, if a mouse contacted the corresponding spout at the time of the current frame.

The movement-related part of the design matrix was comprised of event related regressors, capturing a response around individual events and analog regressors. The event related

regressors were modeled similar to the task-related regressors. Here, individual movement events were modeled for 0.5 s before to 1.5 s after the individual movement-event.

Table 1: Task related regressors in the linear encoding model

Regressor type	Regressor descriptor	Range around event
Sensory - Onset cue	Visual, left	1 s before onset – end of delay
	Visual, right	
	Tactile, left	
	Tactile, right	
	Multisensory, left	
	Multisensory, right	
	Multisensory, left – shuffle control	
	Multisensory, right – shuffle control	
Sensory - Subsequent cue	Visual, left	0.5 s before onset – 1 s after onset (only within stimulus period)
	Visual, right	
	Tactile, left	
	Tactile, right	
	Multisensory, left	
	Multisensory, right	
	Multisensory, left – shuffle control	
	Multisensory, right – shuffle control	
Time	Time within a trial	Entire trial
Choice	Mouse responded “Left” in the current trial	
Choice - previous trial	Mouse responded “Left” in the previous trial	
Reward	Mouse received a reward at the end of trial (correct response)	
Reward - previous trial	Mouse received a reward in the previous trial	

These events included: first left or right lick in a trial, the subsequent licks on the corresponding spout, the reward delivery (independent of the side), the start as well as the end of forward movement and the start and end of backward movement on the wheel (running speed > 1cm/s; Polack et al., 2013).

Additional analog running regressors, relating only to the current frame included: running speed, temporally filtered running speed (gaussian filter with sigma=0.5s; implemented in Matlab using the “filter”- and the “gausswin”-function) and two binary vectors representing the presence of forwards and backwards motion (speed > 1 cm/s) in the corresponding direction.

The last part of the design matrix were the video recordings of the animals’ behavior in the setup. Here, similar to the widefield-preprocessing, the dimensionality of the individual behavior videos was reduced to 200 components using SVD. Subsequently, the two behavior video components were merged using a second SVD, extracting the 200 highest variance components over both video recording. This approach was taken, since both video-stream contained redundant information about the animals’ movements and this allowed to reduce the complexity of the design matrix and circumvent overfitting. Furthermore, signals concerning the licks and the running speed of the mice, are also contained in the video recordings, a QR decomposition was used to remove this information from the SVD components of the video recordings included in the design matrix. With this, it could be ensured that the videos only contribute information not already contain the other movement variables included in the model.

5.7.3.2 Unique explained variance analysis

To determine the unique contribution specific types of information to the overall model predictions, we computed the unique explained variance of the corresponding groups of regressors. For this purpose, the selected regressors were shuffled in time to destroy the correlation with the imaging data. This reduced version of the model was again fit to the same data to identify which features of the Ca^{2+}-imaging data could no longer be predicted by the reduced model. By computing the differential of the cvR^2 of the full model and the cvR^2 of the new reduced model, the unique explained variance (ΔR^2) was determined.

To compute ΔR^2 for specific types of model input, each regressor corresponding to this groups was individually shuffled in time using the “randperm”-function in Matlab, in every iteration

of the cross-validation process. For example, in the case of visual information, this included both the onset and subsequent visual cue regressors, for both the left and right side. To compute ΔR^2 of running information, both the time-varying running events, as well as the analog running regressors were shuffled.

To investigate if specific groups of regressors reliably contributed unique information to the model predictions, the ΔR^2 was computed ten times over for each regressors group and the average computed over the repeated runs to reduce the impact of the shuffling itself. Then the ΔR^2 was averaged over the entire visible surface of the cortex, to obtain an average value for the entire cortex. These average ΔR^2 values are then tested over all sessions included, using a Wilcoxon sign-rank test.

5.7.3.3 Analysis of sensory β-weight kernels

For the analysis of the modeled sensory responses, we inspected the β-weights corresponding to the visual, tactile and non-linear multisensory regressors. The β-weight translate how the presence of a specific event relates to the individual spatial components (U) of the Ca^{2+}-imaging data. To inspect the β-weight we convolved them with U to obtain maps of activity corresponding to the current individual event in the original pixel space. Since the linear encoding model uses individual regressors shifted in time to model the average time-course of a response, relative to the time of an event, we can insect these β-weight maps of individual regressors relating to the same event over time. This sequence of β-weight maps, was then visualized the same way as the original Ca^{2+}-imaging data (see section 5.7.1).

5.7.4 AUC analysis

To identify cortical regions that reliably reflect the target stimulus side the area under the receiver-operator-characteristic curve (AUC) was computed. Here, we selected all discrimination trials within a given session and balanced the numbers of correct and incorrect as well as left and right trials. This was achieved by randomly removing trials until an equal number of trials in all conditions was reached (at least 20 trials per session). This was done to reduce the impact of choice-related signals on this analysis. Individual frames were reconstructed to by convolving V with U and the frames were spatially binned 4 by 4. This reduced signals noise on the individual trial level and also reduced the computational cost of

the later AUC calculation. Then, the differential in activity between the left and right hemisphere was computed. The AUC was calculated pixel by pixel using the “colAUC”-function in Matlab (Tuszynski, 2012), to indicate how distinctly different the activity of a given pixel was over time and trials depending on the target stimulus side. To generate to AUC-maps presented in Figure 11, individual session-wise AUC sequences were averaged, weight by the number of trials that contributed in the individual sessions.

5.7.5 Decoding model

To predict the responses of the mice based on the widefield imaging data a logistic regression decoder was trained on the temporal components of the data and the responses of the mice at the end of the individual trials as the labels. The model was fit using the ‘fitclinear’ function in Matlab for each time point in a trial individually for a given session. To prevent overfitting only the 50 most informative components were used for the decoder. Further, a L1 penalty was used to train the decoder to limit the weights to the most informative of these components. To prevent the decoder from using sensory information or imbalances in left and right responses for the predictions an equal number of correct and error trials, as well as an equal left and right response trials were selected to train the decoder. This was achieved by randomly removing trials until an equal number of trials in all conditions was reached (at least 60 trials). The accuracy of the decoder was determined for each time point individually using 10-fold cross-validation. To create the cortical maps of β-weights the weighted temporal components are convolved with the spatial components of the imaging data.

6 Literature

Abrams, P. A. (1986). Is predator-prey coevolutlon an arms race? *Trends in Ecology & Evolution*, *1*(4), 108–110. https://doi.org/10.1016/0169-5347(86)90037-6

Adibi, M., & Lampl, I. (2021). Sensory Adaptation in the Whisker-Mediated Tactile System: Physiology, Theory, and Function. *Frontiers in Neuroscience*, *15*, 1406. https://doi.org/10.3389/fnins.2021.770011

Akemann, W., Mutoh, H., Perron, A., Park, Y. K., Iwamoto, Y., & Knöpfel, T. (2012). Imaging neural circuit dynamics with a voltage-sensitive fluorescent protein. *Journal of Neurophysiology*, *108*(8), 2323–2337. https://doi.org/10.1152/jn.00452.2012

Akemann, W., Sasaki, M., Mutoh, H., Imamura, T., Honkura, N., & Knöpfel, T. (2013). Two-photon voltage imaging using a genetically encoded voltage indicator. *Scientific Reports*, *3*(1), 2231. https://doi.org/10.1038/srep02231

Allen, W. E., Kauvar, I. V., Chen, M. Z., Richman, E. B., Yang, S. J., Chan, K., Gradinaru, V., Deverman, B. E., Luo, L., & Deisseroth, K. (2017). Global Representations of Goal-Directed Behavior in Distinct Cell Types of Mouse Neocortex. *Neuron*, *94*(4), 891-907.e6. https://doi.org/10.1016/j.neuron.2017.04.017

Alvarado, J. C., Vaughan, J. W., Stanford, T. R., & Stein, B. E. (2007). Multisensory Versus Unisensory Integration: Contrasting Modes in the Superior Colliculus. *Journal of Neurophysiology*, *97*(5), 3193–3205. https://doi.org/10.1152/jn.00018.2007

Andermann, M. L., Kerlin, A. M., Roumis, D. K., Glickfeld, L. L., & Reid, R. C. (2011). Functional specialization of mouse higher visual cortical areas. *Neuron*, *72*(6), 1025–1039. https://doi.org/10.1016/j.neuron.2011.11.013

Ayaz, A., Stäuble, A., Hamada, M., Wulf, M. A., Saleem, A. B., & Helmchen, F. (2019). Layer-specific integration of locomotion and sensory information in mouse barrel cortex. *Nature Communications 2019 10:1*, *10*(1), 1–14. https://doi.org/10.1038/s41467-019-10564-8

Bernhard, S. M., Lee, J., Zhu, M., Hsu, A., Erskine, A., Hires, S. A., & Barth, A. L. (2020). An automated homecage system for multiwhisker detection and discrimination learning in mice. *PLOS ONE*, *15*(12), e0232916. https://doi.org/10.1371/journal.pone.0232916

Bexter, A. (2022). *A Virtual Spatial Navigation Task for Multisensory Discrimination [Doctoral dissertation, RWTH Aachen University], University Library RWTH Aachen*. https://doi.org/10.18154/RWTH-2022-04889

Boulougouris, V., Dalley, J. W., & Robbins, T. W. (2007). Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat. *Behavioural Brain Research*, *179*(2), 219–228. https://doi.org/10.1016/j.bbr.2007.02.005

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Barth J. A.

Brunton, B. W., Botvinick, M. M., & Brody, C. D. (2013). Rats and humans can optimally accumulate evidence for decision-making. *Science*, *340*(6128), 95–98. https://doi.org/10.1126/science.1233912

Burgess, C. P., Lak, A., Steinmetz, N. A., Zatka-Haas, P., Bai Reddy, C., Jacobs, E. A. K., Linden, J. F., Paton, J. J., Ranson, A., Schröder, S., Soares, S., Wells, M. J., Wool, L. E., Harris, K. D., & Carandini, M. (2017). High-Yield Methods for Accurate Two-Alternative Visual Psychophysics in Head-Fixed Mice. *Cell Reports*, *20*(10), 2513–2524. https://doi.org/10.1016/j.celrep.2017.08.047

Busse, L., Ayaz, A., Dhruv, N. T., Katzner, S., Saleem, A. B., Schölvinck, M. L., Zaharia, A. D., & Carandini, M. (2011). The detection of visual contrast in the behaving mouse. *Journal of Neuroscience*, *31*(31), 11351–11361. https://doi.org/10.1523/JNEUROSCI.6689-10.2011

Carandini, M., & Churchland, A. K. (2013). Probing perceptual decisions in rodents. In *Nature Neuroscience* (Vol. 16, Issue 7, pp. 824–831). https://doi.org/10.1038/nn.3410

Carandini, M., Shimaoka, D., Rossi, L. F., Sato, T. K., Benucci, A., & Knöpfel, T. (2015). Imaging the Awake Visual Cortex with a Genetically Encoded Voltage Indicator. *The Journal of Neuroscience*, *35*(1), 53–63. https://doi.org/10.1523/JNEUROSCI.0594-14.2015

Chen, T. W., Li, N., Daie, K., & Svoboda, K. (2017). A Map of Anticipatory Activity in Mouse Motor Cortex. *Neuron*, *94*(4), 866-879.e4. https://doi.org/10.1016/j.neuron.2017.05.005

Chen, T. W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., Schreiter, E. R., Kerr, R. A., Orger, M. B., Jayaraman, V., Looger, L. L., Svoboda, K., & Kim, D. S. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature 2013 499:7458*, *499*(7458), 295–300. https://doi.org/10.1038/nature12354

Choi, I., Lee, J. Y., & Lee, S. H. (2018). Bottom-up and top-down modulation of multisensory integration. In *Current Opinion in Neurobiology* (Vol. 52, pp. 115–122). Elsevier Ltd. https://doi.org/10.1016/j.conb.2018.05.002

Coen, P., Sit, T. P. H., Wells, M. J., Carandini, M., & Harris, K. D. (2023). Mouse frontal cortex mediates additive multisensory decisions. *Neuron*, *111*(15), 2432-2447.e13. https://doi.org/10.1016/j.neuron.2023.05.008

Colombo, M., Colombo, A., & Gross, C. G. (2002). Bartolomeo Panizza's Observations on the optic nerve (1855). *Brain Research Bulletin*, *58*(6), 529–539. https://doi.org/10.1016/S0361-9230(02)00831-6

Couto, J., Musall, S., Sun, X. R., Khanal, A., Gluf, S., Saxena, S., Kinsella, I., Abe, T., Cunningham, J. P., Paninski, L., & Churchland, A. K. (2021). Chronic, cortex-wide imaging of specific cell populations during behavior. *Nature Protocols*, *16*(7), 3241–3263. https://doi.org/10.1038/S41596-021-00527-Z

Crivici, A., & Ikura, M. (1995). Molecular and Structural Basis of Target Recognition by Calmodulin. *Annual Review of Biophysics and Biomolecular Structure*, *24*(1), 85–116. https://doi.org/10.1146/annurev.bb.24.060195.000505

Diederich, A., & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: Effects of stimulus onset and intensity on reaction time. *Perception & Psychophysics*, *66*(8), 1388–1404. https://doi.org/10.3758/BF03195006

Dronkers, N. F., Plaisant, O., Iba-Zizen, M. T., & Cabanis, E. A. (2007). Paul Broca's historic cases: high resolution MR imaging of the brains of Leborgne and Lelong. *Brain*, *130*(5), 1432–1441. https://doi.org/10.1093/brain/awm042

Ebbesen, C. L., & Brecht, M. (2017). Motor cortex — to act or not to act? *Nature Reviews Neuroscience 2017 18:11*, *18*(11), 694–705. https://doi.org/10.1038/nrn.2017.119

Erlich, J. C., Brunton, B. W., Duan, C. A., Hanks, T. D., & Brody, C. D. (2015). Distinct effects of prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat. *ELife*, *4*(4). https://doi.org/10.7554/eLife.05457

Esmaeili, V., Tamura, K., Foustoukos, G., Oryshchuk, A., Crochet, S., & Petersen, C. C. (2020). Cortical circuits for transforming whisker sensation into goal-directed licking. *Current Opinion in Neurobiology*, *65*, 38–48. https://doi.org/10.1016/j.conb.2020.08.003

Esmaeili, V., Tamura, K., Muscinelli, S. P., Modirshanechi, A., Boscaglia, M., Lee, A. B., Oryshchuk, A., Foustoukos, G., Liu, Y., Crochet, S., Gerstner, W., & Petersen, C. C. H. (2021). Rapid suppression and sustained activation of distinct cortical regions for a delayed sensory-triggered motor response. *Neuron*, *109*(13), 2183-2201.e9. https://doi.org/10.1016/j.neuron.2021.05.005

Falchier, A., Clavagnier, S., Barone, P., & Kennedy, H. (2002). Anatomical Evidence of Multimodal Integration in Primate Striate Cortex. *Journal of Neuroscience*, *22*(13), 5749–5759. https://doi.org/10.1523/JNEUROSCI.22-13-05749.2002

Ferezou, I., Haiss, F., Gentet, L. J., Aronoff, R., Weber, B., & Petersen, C. C. H. (2007). Spatiotemporal Dynamics of Cortical Sensorimotor Integration in Behaving Mice. *Neuron*, *56*(5), 907–923. https://doi.org/10.1016/j.neuron.2007.10.007

Frith, C., & Dolan, R. (1996). The role of the prefrontal cortex in higher cognitive functions. *Cognitive Brain Research*, *5*(1–2), 175–181. https://doi.org/10.1016/S0926-6410(96)00054-7

Gallero-Salas, Y., Han, S., Sych, Y., Voigt, F. F., Laurenczy, B., Gilad, A., & Helmchen, F. (2021). Sensory and Behavioral Components of Neocortical Signal Flow in Discrimination Tasks with Short-Term Memory. *Neuron*, *109*(1), 135-148.e6. https://doi.org/10.1016/j.neuron.2020.10.017

Garrett, M. E., Nauhaus, I., Marshel, J. H., Callaway, E. M., Garrett, M. E., Marshel, J. H., Nauhaus, I., & Garrett, M. E. (2014). Topography and areal organization of mouse visual

cortex. *Journal of Neuroscience*, *34*(37), 12587–12600. https://doi.org/10.1523/JNEUROSCI.1124-14.2014

Gilad, A., & Helmchen, F. (2020). Spatiotemporal refinement of signal flow through association cortex during learning. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-15534-z

Gleiss, S., & Kayser, C. (2012). Audio-Visual Detection Benefits in the Rat. *PLOS ONE*, *7*(9), e45677. https://doi.org/10.1371/JOURNAL.PONE.0045677

Goldbach, H. C., Akitake, B., Leedy, C. E., & Histed, M. H. (2021). Performance in even a simple perceptual task depends on mouse secondary visual areas. *ELife*, *10*, 1–39. https://doi.org/10.7554/eLife.62156

Gold, J. I., & Ding, L. (2013). How mechanisms of perceptual decision-making affect the psychometric function. *Progress in Neurobiology*, *103*, 98–114. https://doi.org/10.1016/J.PNEUROBIO.2012.05.008

Gold, J. I., Law, C. T., Connolly, P., & Bennur, S. (2010). Relationships between the threshold and slope of psychometric and neurometric functions during perceptual learning: Implications for neuronal pooling. *Journal of Neurophysiology*, *103*(1), 140–154. https://doi.org/10.1152/jn.00744.2009

Goldman-Rakic, P. S., & Rakic, P. (1991). Distributed Hierarchical Processing in the Primate Cerebral Cortex. *Cerebral Cortex*, *1*(1), 1–47. https://doi.org/10.1093/cercor/1.1.1-a

Goltstein, P. M., Reinert, S., Bonhoeffer, T., & Hübener, M. (2021). Mouse visual cortex areas represent perceptual and semantic features of learned visual categories. *Nature Neuroscience*, *24*(10), 1441–1451. https://doi.org/10.1038/s41593-021-00914-5

Goltstein, P. M., Reinert, S., Glas, A., Bonhoeffer, T., & Hübener, M. (2018). Food and water restriction lead to differential learning behaviors in a head-fixed two-choice visual discrimination task for mice. *PLoS ONE*, *13*(9). https://doi.org/10.1371/journal.pone.0204066

Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, *27*, 649–677. https://doi.org/10.1146/ANNUREV.NEURO.27.070203.144220

Guizar-Sicairos, M., Thurman, S. T., & Fienup, J. R. (2008). Efficient subpixel image registration algorithms. *Optics Letters*, *33*(2), 156. https://doi.org/10.1364/OL.33.000156

Guo, Z. V., Li, N., Huber, D., Ophir, E., Gutnisky, D., Ting, J. T., Feng, G., & Svoboda, K. (2014). Flow of cortical activity underlying a tactile decision in mice. *Neuron*, *81*(1), 179–194. https://doi.org/10.1016/j.neuron.2013.10.020

Hirokawa, J., Sadakane, O., Sakata, S., Bosch, M., Sakurai, Y., & Yamamori, T. (2011). Multisensory Information Facilitates Reaction Speed by Enlarging Activity Difference between Superior Colliculus Hemispheres in Rats. *PLOS ONE*, *6*(9), e25283. https://doi.org/10.1371/JOURNAL.PONE.0025283

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, *160*(1), 106–154. https://doi.org/10.1113/jphysiol.1962.sp006837

Huber, D., Petreanu, L., Ghitani, N., Ranade, S., Hromádka, T., Mainen, Z., & Svoboda, K. (2008). Sparse optical microstimulation in barrel cortex drives learned behaviour in freely moving mice. *Nature*, *451*(7174), 61–64. https://doi.org/10.1038/nature06445

Hudspeth, A., & Logothetis, N. K. (2000). Sensory systems. *Current Opinion in Neurobiology*, *10*(5), 631–641. https://doi.org/10.1016/S0959-4388(00)00133-1

Iurilli, G., Ghezzi, D., Olcese, U., Lassi, G., Nazzaro, C., Tonini, R., Tucci, V., Benfenati, F., & Medini, P. (2012). Sound-Driven Synaptic Inhibition in Primary Visual Cortex. *Neuron*, *73*(4), 814–828. https://doi.org/10.1016/j.neuron.2011.12.026

Keller, A. J., Houlton, R., Kampa, B. M., Lesica, N. A., Mrsic-Flogel, T. D., Keller, G. B., & Helmchen, F. (2017). Stimulus relevance modulates contrast adaptation in visual cortex. *ELife*, *6*. https://doi.org/10.7554/eLife.21589

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS ONE*, *2*(9). https://doi.org/10.1371/journal.pone.0000943

Krubitzer, L. (2007). The magnificent compromise: cortical field evolution in mammals. In *Neuron* (Vol. 56, Issue 2, pp. 201–208). https://doi.org/10.1016/j.neuron.2007.10.002

Krubitzer, L., & Kahn, D. M. (2003). Nature versus nurture revisited: an old idea with a new twist. *Progress in Neurobiology*, *70*(1), 33–52. https://doi.org/10.1016/S0301-0082(03)00088-1

Laurienti, P. J., Burdette, J. H., Wallace, M. T., Yen, Y. F., Field, A. S., & Stein, B. E. (2002). Deactivation of Sensory-Specific Cortex by Cross-Modal Stimuli. *Journal of Cognitive Neuroscience*, *14*(3), 420–429. https://doi.org/10.1162/089892902317361930

Leinweber, M., Ward, D. R., Sobczak, J. M., Attinger, A., & Keller, G. B. (2017). A Sensorimotor Circuit in Mouse Cortex for Visual Flow Predictions. *Neuron*, *95*(6), 1420-1432.e5. https://doi.org/10.1016/j.neuron.2017.08.036

Lerner, T. N., Shilyansky, C., Davidson, T. J., Evans, K. E., Beier, K. T., Zalocusky, K. A., Crow, A. K., Malenka, R. C., Luo, L., Tomer, R., & Deisseroth, K. (2015). Intact-Brain Analyses Reveal Distinct Information Carried by SNc Dopamine Subcircuits. *Cell*, *162*(3), 635–647. https://doi.org/10.1016/j.cell.2015.07.014

Li, N., Chen, T. W., Guo, Z. V., Gerfen, C. R., & Svoboda, K. (2015). A motor cortex circuit for motor planning and movement. *Nature*, *519*(7541), 51–56. https://doi.org/10.1038/nature14178

Liu, D., Deng, J., Zhang, Z., Zhang, Z. Y., Sun, Y. G., Yang, T., & Yao, H. (2020). Orbitofrontal control of visual cortex gain promotes visual associative learning. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-16609-7

Lohse, M., Dahmen, J. C., Bajo, V. M., & King, A. J. (2021). Subcortical circuits mediate communication between primary sensory cortical areas in mice. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-24200-x

Macaluso, E., Frith, C. D., & Driver, J. (2000). Modulation of Human Visual Cortex by Crossmodal Spatial Attention. *Science*, *289*(5482), 1206–1208. https://doi.org/10.1126/science.289.5482.1206

Marshel, J. H., Garrett, M. E., Nauhaus, I., & Callaway, E. M. (2011). Functional specialization of seven mouse visual cortical areas. *Neuron*, *72*(6), 1040–1054. https://doi.org/10.1016/j.neuron.2011.12.004

Mayford, M., Bach, M. E., Huang, Y.-Y., Wang, L., Hawkins, R. D., & Kandel, E. R. (1996). Control of Memory Formation Through Regulated Expression of a CaMKII Transgene. *Science*, *274*(5293), 1678–1683. https://doi.org/10.1126/science.274.5293.1678

Mayrhofer, J. M., Skreb, V., Von der Behrens, W., Musall, S., Weber, B., & Haiss, F. (2013). Novel two-alternative forced choice paradigm for bilateral vibrotactile whisker frequency discrimination in head-fixed mice and rats. *Journal of Neurophysiology*, *109*(1), 273–284. https://doi.org/10.1152/jn.00488.2012

McDonald, R. J., & White, N. M. (1994). Parallel information processing in the water maze: Evidence for independent memory systems involving dorsal striatum and hippocampus. *Behavioral and Neural Biology*, *61*(3), 260–270. https://doi.org/10.1016/S0163-1047(05)80009-3

McKerchar, T. L., Zarcone, T. J., & Fowler, S. C. (2005). Differential acquisition of lever pressing in inbred and outbred mice: comparison of one-lever and two-lever procedures and correlation with differences in locomotor activity. *Journal of the Experimental Analysis of Behavior*, *84*(3), 339–356. https://doi.org/10.1901/jeab.2005.95-04

Meier, P., & Reinagel, P. (2011). Rat performance on visual detection task modeled with divisive normalization and adaptive decision thresholds. *Journal of Vision*, *11*(9), 1–17. https://doi.org/10.1167/11.9.1

Meijer, G. T., Marchesi, P., Mejias, J. F., Montijn, J. S., Lansink, C. S., & Pennartz, C. M. A. (2020). Neural Correlates of Multisensory Detection Behavior: Comparison of Primary and Higher-Order Visual Cortex. *Cell Reports*, *31*(6). https://doi.org/10.1016/j.celrep.2020.107636

Meredith, M. A., & Stein, B. E. (1983). Interactions Among Converging Sensory Inputs in the Superior Colliculus. *Science*, *221*(4608), 389–391. https://doi.org/10.1126/science.6867718

Meredith, M. A., & Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, *56*(3), 640–662. https://doi.org/10.1152/jn.1986.56.3.640

Mohan, H., An, X., Kondo, H., Zhao, S., Matho, K. S., Musall, S., Mitra, P., & Huang, Z. J. (2022). Cortical glutamatergic projection neuron types contribute to distinct functional subnetworks. *BioRxiv*, 2021.12.30.474537. https://doi.org/10.1101/2021.12.30.474537

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cognitive Brain Research*, *14*(1), 115–128. https://doi.org/10.1016/S0926-6410(02)00066-6

Morcos, A. S., & Harvey, C. D. (2016). History-dependent variability in population dynamics during evidence accumulation in cortex. *Nature Neuroscience*, *19*(12), 1672–1681. https://doi.org/10.1038/nn.4403

Morrell, F. (1972). Visual System's View of Acoustic Space. *Nature*, *238*(5358), 44–46. https://doi.org/10.1038/238044a0

Murakami, T., Yoshida, T., Matsui, T., & Ohki, K. (2015). Wide-field Ca2+ imaging reveals visually evoked activity in the retrosplenial area. *Frontiers in Molecular Neuroscience*, *08*(June). https://doi.org/10.3389/fnmol.2015.00020

Murphy, T. H., Boyd, J. D., Bolaños, F., Vanni, M. P., Silasi, G., Haupt, D., & Ledue, J. M. (2016). High-throughput automated home-cage mesoscopic functional imaging of mouse cortex. *Nature Communications*, *7*. https://doi.org/10.1038/ncomms11611

Murray, M. M., Lewkowicz, D. J., Amedi, A., & Wallace, M. T. (2016). Multisensory Processes: A Balancing Act across the Lifespan. *Trends in Neurosciences*, *39*(8), 567–579. https://doi.org/10.1016/j.tins.2016.05.003

Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S., & Churchland, A. K. (2019). Single-trial neural dynamics are dominated by richly varied movements. *Nature Neuroscience*, *22*(10), 1677–1686. https://doi.org/10.1038/s41593-019-0502-4

Musall, S., Sun, X. R., Mohan, H., An, X., Gluf, S., Li, S.-J., Drewes, R., Cravo, E., Lenzi, I., Yin, C., Kampa, B. M., & Churchland, A. K. (2023). Pyramidal cell types drive functionally distinct cortical activity patterns during decision-making. *Nature Neuroscience*, 2021.09.27.461599. https://doi.org/10.1038/s41593-022-01245-9

O'Connor, D. H., Huber, D., & Svoboda, K. (2009). Reverse engineering the mouse brain. *Nature*, *461*(7266), 923–929. https://doi.org/10.1038/nature08539

Olcese, U., Iurilli, G., & Medini, P. (2013). Cellular and synaptic architecture of multisensory integration in the mouse neocortex. *Neuron*, *79*(3), 579–593. https://doi.org/10.1016/j.neuron.2013.06.010

Ollerenshaw, D. R., Bari, B. A., Millard, D. C., Orr, L. E., Wang, Q., & Stanley, G. B. (2012). Detection of tactile inputs in the rat vibrissa pathway. *J Neurophysiol*, *108*, 479–490. https://doi.org/10.1152/jn.00004.2012.-The

Orsolic, I., Rio, M., Mrsic-Flogel, T. D., & Znamenskiy, P. (2021). Mesoscale cortical dynamics reflect the interaction of sensory evidence and temporal expectation during perceptual decision-making. *Neuron*, *109*(11), 1861-1875.e10. https://doi.org/10.1016/j.neuron.2021.03.031

Ostrolenk, A., Bao, V. A., Mottron, L., Collignon, O., & Bertone, A. (2019). Reduced multisensory facilitation in adolescents and adults on the Autism Spectrum. *Scientific Reports 2019 9:1*, *9*(1), 1–9. https://doi.org/10.1038/s41598-019-48413-9

Pakan, J. M., Lowe, S. C., Dylda, E., Keemink, S. W., Currie, S. P., Coutts, C. A., & Rochefort, N. L. (2016). Behavioral-state modulation of inhibition is context-dependent and cell type specific in mouse visual cortex. *ELife*, *5*. https://doi.org/10.7554/eLife.14985

Panzeri, S., Harvey, C. D., Piasini, E., Latham, P. E., & Fellin, T. (2017). Cracking the Neural Code for Sensory Perception by Combining Statistics, Intervention, and Behavior. *Neuron*, *93*(3), 491–507. https://doi.org/10.1016/j.neuron.2016.12.036

Petersen, C. C. H. (2007). The Functional Organization of the Barrel Cortex. *Neuron*, *56*(2), 339–355. https://doi.org/10.1016/J.NEURON.2007.09.017

Pinto, L., Koay, S. A., Engelhard, B., Yoon, A. M., Deverett, B., Thiberge, S. Y., Witten, I. B., Tank, D. W., & Brody, C. D. (2018). An accumulation-of-evidence task using visual pulses for mice navigating in virtual reality. *Frontiers in Behavioral Neuroscience*, *12*. https://doi.org/10.3389/fnbeh.2018.00036

Pinto, L., Rajan, K., DePasquale, B., Thiberge, S. Y., Tank, D. W., & Brody, C. D. (2019). Task-Dependent Changes in the Large-Scale Dynamics and Necessity of Cortical Regions. *Neuron*, *104*(4), 810-824.e9. https://doi.org/10.1016/j.neuron.2019.08.025

Pisupati, S., Chartarifsky-Lynn, L., Khanal, A., & Churchland, A. K. (2021). Lapses in perceptual decisions reflect exploration. *ELife*, *10*, 1–27. https://doi.org/10.7554/ELIFE.55490

Polack, P. O., Friedman, J., & Golshani, P. (2013). Cellular mechanisms of brain state-dependent gain modulation in visual cortex. *Nature Neuroscience*, *16*(9), 1331–1339. https://doi.org/10.1038/nn.3464

Pöpple, E., Held, R., & Frost, D. (1973). Residual Visual Function after Brain Wounds involving the Central Visual Pathways in Man. *Nature*, *243*(5405), 295–296. https://doi.org/10.1038/243295a0

Raposo, D., Sheppard, J. P., Schrater, P. R., & Churchland, A. K. (2012). Multisensory decision-making in rats and humans. *Journal of Neuroscience*, *32*(11), 3726–3735. https://doi.org/10.1523/JNEUROSCI.4998-11.2012

Ratzlaff, E. H., & Grinvald, A. (1991). A tandem-lens epifluorescence macroscope: Hundred-fold brightness advantage for wide-field imaging. *Journal of Neuroscience Methods*, *36*(2–3), 127–137. https://doi.org/10.1016/0165-0270(91)90038-2

Reinagel, P. (2018). Training Rats Using Water Rewards Without Water Restriction. *Frontiers in Behavioral Neuroscience*, *12*. https://doi.org/10.3389/fnbeh.2018.00084

Ren, C., & Komiyama, T. (2021). Characterizing Cortex-Wide Dynamics with Wide-Field Calcium Imaging. *Journal of Neuroscience*, *41*(19), 4160–4168. https://doi.org/10.1523/JNEUROSCI.3003-20.2021

Rich, E. L., & Shapiro, M. L. (2007). Prelimbic/infralimbic inactivation impairs memory for multiple task switches, but not flexible selection of familiar tasks. *Journal of Neuroscience*, *27*(17), 4747–4755. https://doi.org/10.1523/JNEUROSCI.0369-07.2007

Sakata, S., Yamamori, T., & Sakurai, Y. (2004). Behavioral studies of auditory-visual spatial recognition and integration in rats. *Experimental Brain Research*, *159*(4), 409–417. https://doi.org/10.1007/S00221-004-1962-6/TABLES/3

Sanders, J. I., & Kepecs, A. (2012). Choice ball: a response interface for two-choice psychometric discrimination in head-fixed mice. *Journal of Neurophysiology*, *108*(12), 3416–3423. https://doi.org/10.1152/jn.00669.2012

Scott, B. B., Constantinople, C. M., Erlich, J. C., Tank, D. W., & Brody, C. D. (2015). *Sources of noise during accumulation of evidence in unrestrained and voluntarily head-restrained rats*. https://doi.org/10.7554/eLife.11308.001

Shimaoka, D., Harris, K. D., & Carandini, M. (2018). Effects of Arousal on Mouse Sensory Cortex Depend on Modality. *Cell Reports*, *22*(12), 3160–3167. https://doi.org/10.1016/j.celrep.2018.02.092

Siemann, J. K., Muller, C. L., Bamberger, G., Allison, J. D., Veenstra-VanderWeele, J., & Wallace, M. T. (2015). A novel behavioral paradigm to assess multisensory processing in mice. *Frontiers in Behavioral Neuroscience*, *8*(Jan). https://doi.org/10.3389/fnbeh.2014.00456

Silasi, G., Xiao, D., Vanni, M. P., Chen, A. C. N., & Murphy, T. H. (2016). Intact skull chronic windows for mesoscopic wide-field imaging in awake mice. *Journal of Neuroscience Methods*, *267*, 141–149. https://doi.org/10.1016/j.jneumeth.2016.04.012

Stanford, T. R., & Stein, B. E. (2007). Superadditivity in multisensory integration: putting the computation in context. *NeuroReport*, *18*(8), 787–792. https://doi.org/10.1097/WNR.0b013e3280c1e315

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149. https://doi.org/10.3758/BF03207704

Stein, B. E., Scott Huneycutt, W., & Alex Meredith, M. (1988). Neurons and behavior: the same rules of multisensory integration apply. *Brain Research*, *448*(2), 355–358. https://doi.org/10.1016/0006-8993(88)91276-0

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, *9*(4), 255–266. https://doi.org/10.1038/nrn2331

Steinmetz, N. A., Zatka-Haas, P., Carandini, M., & Harris, K. D. (2019). Distributed coding of choice, action and engagement across the mouse brain. *Nature*, *576*(7786), 266–273. https://doi.org/10.1038/s41586-019-1787-x

Stone, D. B., Urrea, L. J., Aine, C. J., Bustillo, J. R., Clark, V. P., & Stephen, J. M. (2011). Unisensory processing and multisensory integration in schizophrenia: A high-density electrical mapping study. *Neuropsychologia*, *49*(12), 3178–3187. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2011.07.017

Sugihara, T., Diltz, M. D., Averbeck, B. B., & Romanski, L. M. (2006). Integration of auditory and visual communication information in the primate ventrolateral prefrontal cortex. *Journal of Neuroscience*, *26*(43), 11138–11147. https://doi.org/10.1523/JNEUROSCI.3550-06.2006

Tomar, S. (2006). Converting video formats with FFmpeg. *Linux Journal*, *2006*(146), 10.

Tsutsui, K. I., Oyama, K., Nakamura, S., & Iijima, T. (2016). Comparative overview of visuospatial working memory in monkeys and rats. In *Frontiers in Systems Neuroscience* (Vol. 10, Issue DEC). Frontiers Research Foundation. https://doi.org/10.3389/fnsys.2016.00099

Tuszynski, J. (2012). *colAUC*. MATLAB Central File Exchange.

Urai, A. E., Aguillon-Rodriguez, V., Laranjeira, I. C., Cazettes, F., Mainen, Z. F., & Churchland, A. K. (2021). Citric acid water as an alternative to water restriction for high-yield mouse behavior. *ENeuro*, *8*(1), 1–8. https://doi.org/10.1523/ENEURO.0230-20.2020

Vanni, M. P., Chan, A. W., Balbi, M., Silasi, G., & Murphy, T. H. (2017). Mesoscale mapping of mouse cortex reveals frequency-dependent cycling between distinct macroscale

functional modules. *Journal of Neuroscience*, *37*(31), 7513–7533. https://doi.org/10.1523/JNEUROSCI.3560-16.2017

Wallace, M. T., Ramachandran, R., & Stein, B. E. (2004). A revised view of sensory cortical parcellation. *Proceedings of the National Academy of Sciences*, *101*(7), 2167–2172. https://doi.org/10.1073/pnas.0305697101

Wallace, M. T., Wilkinson, L. K., & Stein, B. E. (1996). Representation and integration of multiple sensory inputs in primate superior colliculus. *Journal of Neurophysiology*, *76*(2), 1246–1266. https://doi.org/10.1152/jn.1996.76.2.1246

Wang, Q., & Burkhalter, A. (2007). Area map of mouse visual cortex. *Journal of Comparative Neurology*, *502*(3), 339–357. https://doi.org/10.1002/cne.21286

Wang, Q., Ding, S. L., Li, Y., Royall, J., Feng, D., Lesnar, P., Graddis, N., Naeemi, M., Facer, B., Ho, A., Dolbeare, T., Blanchard, B., Dee, N., Wakeman, W., Hirokawa, K. E., Szafer, A., Sunkin, S. M., Oh, S. W., Bernard, A., … Ng, L. (2020). The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. *Cell*, *181*(4), 936-953.e20. https://doi.org/10.1016/j.cell.2020.04.007

Wekselblatt, J. B., Flister, E. D., Piscopo, D. M., & Niell, C. M. (2016). Large-scale imaging of cortical dynamics during sensory perception and behavior. *J Neurophysiol*, *115*, 2852–2866. https://doi.org/10.1152/jn.01056.2015.-Sensory-driven

Xu, D., Dong, M., Chen, Y., Delgado, A. M., Hughes, N. C., Zhang, L., & O'Connor, D. H. (2022). Cortical processing of flexible and context-dependent sensorimotor sequences. *Nature*. https://doi.org/10.1038/s41586-022-04478-7

Zatka-Haas, P., Steinmetz, N. A., Carandini, M., & Harris, K. D. (2021). Sensory coding and the causal impact of mouse cortex in a visual decision. *ELife*, *10*. https://doi.org/10.7554/eLife.63163

Eidesstattliche Erklärung

Ich, Gerion Nabbefeld erklärt hiermit, dass diese Dissertation und die darin dargelegten Inhalte die eigenen sind und selbstständig, als Ergebnis der eigenen originären Forschung, generiert wurden. Hiermit erkläre ich an Eides statt:

1. Diese Arbeit wurde vollständig oder größtenteils in der Phase als Doktorand dieser Fakultät und Universität angefertigt;
2. Sofern irgendein Bestandteil dieser Dissertation zuvor für einen akademischen Abschluss oder eine andere Qualifikation an dieser oder einer anderen Institution verwendet wurde, wurde dies klar angezeigt;
3. Wenn immer andere eigene- oder Veröffentlichungen Dritter herangezogen wurden, wurden diese klar benannt;
4. Wenn aus anderen eigenen- oder Veröffentlichungen Dritter zitiert wurde, wurde stets die Quelle hierfür angegeben. Diese Dissertation ist vollständig meine eigene Arbeit, mit der Ausnahme solcher Zitate;
5. Alle wesentlichen Quellen von Unterstützung wurden benannt;
6. Wenn immer ein Teil dieser Dissertation auf der Zusammenarbeit mit anderen basiert, wurde von mir klar gekennzeichnet, was von anderen und was von mir selbst erarbeitet wurde;
7. Kein Teil dieser Arbeit wurde vor deren Einreichung veröffentlicht.

_____

Ort, Datum, Unterschrift