Single-neuron correlates of visual object representation in the human brain: effects of attention, memory, and choice

Thesis by
Juri Minxha

In Partial Fulfillment of the Requirements
for the degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2018
(Defended December 15, 2017)

ACKNOWLEDGEMENTS

After 6 years, I now have a wife, less hair, and a rekindled appreciation for scientific research. I would like to dedicate the next few paragraphs of this dissertation to acknowledging the people who either directly or indirectly made this happen.

My entry into neuroscience was through the field of brain-machine interfaces. As an undergraduate, I studied electrical engineering, but after spending some time in a neuroscience lab as a research assistant, I naturally veered towards it (and thankfully, away from analog circuit analysis). I came to Caltech to work on brain-machine interfaces in Richard Andersen's lab. In the two years I was in the lab, I met some truly wonderful people there. I want to thank Richard for always being supportive and introducing me to the world of monkey and human electrophysiology. I want to thank Tyson Aflalo, Christian Klaes (my office mate for 2 years), Spencer Kellis, and Boris Revechkis for their mentorship and for the pranks and adventures we shared as the first human recordings were getting off the ground.

In the remainder of my time at Caltech, I worked in the labs of Ralph Adolphs and Ueli Rutishauser. A graduate student could not possibly ask for a better combination of advisers than Ralph and Ueli. It is actually because of a serendipitous meeting with Ueli, while I was doing a two-month internship in Frankfurt in 2010 that I decided to consider graduate school in neuroscience, and specifically at Caltech. Ueli is the technically minded adviser who can offer advice on everything from programming routines and analysis techniques to broad conceptual questions. On a few occasions, we have even sat together to debug faulty code. It is hard to imagine anyone having the patience for that. I am also very grateful to Ralph for creating a sense of community in the lab and making me feel welcome from the start. His incredibly insightful remarks, his unfaltering work ethic, and his accessibility at all times have frequently left me stupefied. I hope that someday I can emulate their seemingly boundless energy and patience for mentorship. I can only hope that this is something that will come with age.

Most of the data used in this thesis comes from human subjects that are implanted with electrodes for the purpose of monitoring and localizing their seizures. This data is incredibly rare and valuable, and it was made possible by a large group of people. I would be remiss if I did not single out Dr. Adam Mamelak, our neurosurgeon. Adam is not only a wonderful doctor but also a very insightful scientist. We have worked on a number of projects together, and I have always looked forward to his unique, outside-the-box perspective on the data that we collect. I would also like to acknowledge Drs. Jeffrey Chung and Chrystal Reed, our neurologists. Jeffrey and Chrystal have made sure that our patients are happy and that our recording sessions run smoothly. Thanks to Adam, Jeffrey, and Chrystal, I have also had the privilege of observing a number of clinical procedures, a unique experience that is not shared by most neuroscientists. While I will not name them individually, I also want to acknowledge the patients and the large group of dedicated staff members at Cedars Sinai who made all of this research possible.

I would also like to acknowledge the two other members of my committee, John O'Doherty and Doris Tsao. I have interacted with John and Doris extensively in the time that I have been at Caltech, and I look to them and the quality of their work as prime examples of what I would like to do and who I would like to be as a scientist.

Science (at least the version that I have been engaged in) is a community endeavor. Over the years, I have received a lot of feedback from my lab mates, in both the Rutishauser and Adolphs Labs. I want to thank the members of the Adolphs lab for the supportive community we have built over the years, the frequent trips to the Catalina Island, and the personal conversations we have shared. I want to single out Tim, my office mate, for being a constant source of refreshing conversation, serving me with the daily reminder of the hobbies and leisure activities someone *could* engage in if they had the time.

I also want to acknowledge the current and former members of the Rutishauser Lab. Of these, I would like to single out Jan Kaminski, whose skepticism has always kept me on my toes during all lab presentations. I would like to thank all the lab members for the academic and personal conversations we have shared throughout the years.

I would be remiss if I did not mention my cohort of fellow CNS students. I specifically want to single out Gidi Nave and Luke Urban. Gidi was my first friend at Caltech, and his friendship over the years is something I truly cherish and hope to cultivate for the rest of my life. Luke and I had very similar trajectories through graduate school; we both started out working with brain-machine interfaces. Our light-hearted conversations over the years have often provided the much-needed break from the day-to-day grind.

I want to recognize all of my friends at Caltech, from whom I feel I have taken so much. Caltech is a small place and my friends are all from diverse fields. Sometimes it feels that my PhD was not just in neuroscience, but also astrophysics, computer vision, mathematics, etc.

I want to thank all the members of my Albanian family, Moza, Halit, Rudi, Edmond, Alba, Linda, and Kevin as well as my Brazilian family for their unwavering support throughout my time here.

My time at Caltech is a very special time, because apart from all the wonderful things I learned and the great people I met, this was the place where I met my wife. I met Gabi in February of 2013, during her CNS interview weekend. She had that infectious enthusiasm for science that sometimes, as a third-year graduate student, you forget. Over the past four years, we have lived together, traveled, published papers, been resident associates, and celebrated our wedding with all of our friends, and soon, we are about to embark on an extended trip around the world. All of this means so much more because of her.

ABSTRACT

My thesis is primarily focused on the single-neuron representations of visual categories, memory, and choice in the medial temporal lobe (MTL) and the medial frontal cortex (MFC) of the human brain. Most of the data presented here comes from single-unit recordings conducted in neurosurgical patients who were implanted with depth electrodes used to monitor their epileptic seizures. Given the significance of the data, it seems only appropriate that Chapter II (the first data chapter) of this thesis is dedicated to detailing the process of acquiring such recordings. The chapter provides extensive detail on the surgical techniques, anatomical targets, data acquisition, and post-processing procedures that produced the data we analyze here.

In Chapter III, we will focus on representations of visual categories in single cells in the amygdala. One very special brand of visual categories is faces. We know that neurons in the primate amygdala respond prominently to socially significant stimuli, such as faces, yet the contribution of these responses to social perception remains poorly understood. In the first part of this thesis, we evaluated the representation of faces in the primate amygdala during naturalistic conditions by recording from both human and macaque amygdala neurons during free viewing of identical arrays of images with concurrent eye tracking. Among the notable differences, we found that in both monkeys and humans, the majority of face-selective neurons preferred faces of conspecifics, a bias also seen behaviorally in their fixation preferences. We found that response latencies, relative to fixation onset, were shortest for conspecific-selective neurons. Our data suggests that overt attention to faces gates amygdala responses, which in turn prioritize species-typical information for further processing.

In Chapter IV, we investigated how visual representations in the medial temporal lobe are subsequently used to make two types of decisions: a recognition memory choice ("Have you seen this image before?"), and a stimulus categorization choice ("Is this a face?"). We show that: (1) there are distinct populations of cells in the medial frontal cortex (including dorsal-anterior cingulate cortex and pre-supplementary motor cortex) encoding recognition memory or categorization-based choices; (2) category-selective cells in the medial

temporal lobe are insensitive to such task conditions; and (3) spike-field coherence between field potentials in the medial temporal lobe and action potentials in the medial frontal cortex are enhanced during recognition memory choices. This suggests that inter-areal communication between these two brain regions may be facilitated selectively in tasks that rely on recognition memory-based information.

Overall, we have been able to show, using concurrent eye-tracking and single-unit recordings, that the visual representations of objects (and specifically faces) in the MTL are gated by fixations, as measured in a free-viewing task. Additionally, we have also observed representations of abstract choice in the medial frontal cortex during a memory – recognition and visual categorization task. Based on our coherence analysis between MFC cells and local field potentials in the MTL, we believe that this representation of choice in MFC is enabled by internal representations of recognition memory in MTL.

PUBLISHED CONTENT AND CONTRIBUTIONS

Minxha, Juri, et al. "Fixations gate species-specific responses to free viewing of faces in the human and macaque amygdala." *Cell reports* 18.4 (2017): 878-891.

Minxha, Juri, et al. "Surgical and Electrophysiological Techniques for Single-Neuron Recordings in Human Epilepsy Patients" In *Extracellular Recording Approaches, Springer, Ed. Sillitoe, Roy, ISBN 978-1-4939-7548-8 (2018)*

Minxha, Juri, and Ueli Rutishauser. "Making decisions based on autobiographical memories." *Neuron* 86.2 (2015): 350-352.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

ACC	anterior cingulate cortexOver
AED	anti-epileptic drug
BOLD	blood-oxygen-level-dependent
CV	coefficient of variation
DOS	depth of selectivity
EEG	electroencephalography
EPSP	excitatory postsynaptic current
ERP	event-related potential
FDR	false discovery rate
fMRI	functional magnetic resonance imaging
iEEG	intracranial electroencephalography
ISI	interspike interval
KS-test	Komogorov-Smirnof goodness-of-fit test
LFP	local field potential
mPFC	medial prefrontal cortex
MTL	medial temporal lobe

MUA	multi-unit activity
OFC	orbitofrontal cortex
PCA	principal component analysis
PFC	prefrontal cortex
pre-SMA	pre supplementary motor area
RT	reaction time
ROC	receiver operator characteristic
SNR	signal-to-noise ratio
SVM	support vector machine
SUA	single-unit activity
STD	standard deviation
SEM	standard error of the mean
TLE	standard temporal lobe epilepsy
TSLS	time since last spike

Chapter I: General Introduction

1.1 Overview

Imagine that you are walking down a crowded New York City street. If you are a local, you are largely ignoring the deluge of information all around you. If you are a tourist, you are eagerly sampling the sights, sounds, and smells the city has to offer as you walk by. You shift your gaze to sample the many different objects in your environment and you may choose to linger on things that pique your interest (ex. an interesting face or an advertisement). This scenario highlights the incredibly dynamic nature of the world we live in. And yet, despite this, our brain has relatively stable representations of the objects in our world. For example, we can effortlessly tell if what we are looking at is a face. If probed, we can also answer different questions about the face: "Does it seem happy?" or "Is that my cousin Alba?" We can extract a lot of visual information from the objects we encounter in the world, and we can parse this information in a task-dependent way. Much of what we know about the visual representation of objects in the human brain (and specifically representations of faces in the amygdala) has been learned from experiments that do not necessarily capture the information-rich and dynamic nature of the world where we might encounter these objects. For example, most of what we know about the processing of faces in the amygdala has been learned from the presentation of static stimuli. While much has been learned from such experiments, it is unclear how the associated neural responses generalize to cases where (1) there is competition for the subject's attention, and (2) the subject is free to look at whatever they prefer (free-view). One of the first objectives in this thesis is to determine what these visually evoked responses to faces in the amygdala look like when we lift some of the experimental restrictions that have been used in past experiments (i.e. enforced fixation, single static image). Understanding how the representation of faces is affected during free viewing is only the first step. The second objective of this thesis is to understand how visual representations in the medial temporal lobe (specifically amygdala and hippocampus) allow us to make decisions about what we are looking at. For example, I could show you an image and ask if it is an image of a human face. This is a trivial task and most people would be able to answer this question very quickly. Separately, I can show you the same exact image, and ask if you have ever seen it before. The second question is a little more difficult because it

requires memory (i.e. a memory-guided decision). These two tasks require that you extract two relatively orthogonal pieces of information about the same stimulus and yet we are able to flexibly address these questions as they are presented to us. Therefore, throughout this thesis, our objectives are twofold: (1) characterize visual representations in the amygdala under natural viewing conditions (specifically for faces), and (2) understand how the visual representations in MTL can allow us to make simple decisions about what we are looking at (e.g. "It's a face" or "I have never seen that before").

Most of the data that is collected to address the objectives outlined above comes from neurosurgical patients who were implanted with depth electrodes used for monitoring their epileptic seizures. Typical targets for these electrodes include bilateral hippocampus and amygdala in the medial temporal lobe and pre-supplementary motor (pre-SMA), dorsal-anterior cingulate cortex (dACC) in the medial frontal cortex. This data is made up of 1973 isolated putative cells. The contributions to this number from all the areas recorded are as follows: 594 in amygdala, 404 in the hippocampus, 399 in pre-SMA, and 576 in dACC. This data was collected over 49 independent sessions, from 19 different patients, and represents a total recording time of 22.3 hours (this number represents only the portions of the recordings that were directly used for this thesis, not the total amount of recording).

1.2 Visually-selective neurons in the human amygdala

There is an extensive body of literature detailing the response properties of visually selective neurons in the human amygdala. About 20% of the human amygdala neurons show high-level visual category selectivity (Fried, MacDonald et al. 1997, Kreiman, Koch et al. 2000, Rutishauser, Ye et al. 2015). The response latency of the visually selective cells in the amygdala is approximately 400ms (Mormann, Kornblith et al. 2008, Rutishauser, Ye et al. 2015) , and it can be both excitatory and inhibitory (Rutishauser, Tudusciuc et al. 2011). Some studies have also noted a difference in the responses between right and left amygdala. Category-specific responses to animals for example seem to be more prevalent in the right amygdala (Mormann, Dubois et al. 2011). Certain categories of images are represented more in the population of visually selective amygdala cells. Up to 50% of human amygdala cells

respond in some manner to faces (Rutishauser, Tudusciuc et al. 2013) and these face responses can show selectivity for face-parts (such as eyes or nose) while others only respond to the presence of the entire face. Some studies have explored responses to emoting faces and have noted that the responses of face-selective cells in the amygdala correlate more with the subject's choice (i.e., subjective perception of the face) than with the sensory evidence (Wang, Tudusciuc et al. 2014). On a follow-up study, it was also found that the face responses in the amygdala encode the intensity of a specific facial emotion (Wang, Yu et al. 2017). Together, this body of literature shows robust visual responses in the human amygdala but it does a poor job of explaining what the computational purpose of these responses might be. The latter studies mentioned above (specifically by Wang et al.) have started to address this by linking the face responses to behavioral measurements (e.g. the subject's perception/choice). It still remains unclear however what the computational purpose of such visually highly-selective neurons (which be found throughout the MTL not just amygdala (Mormann, Ison et al. 2014)) actually is.

1.3 Face processing in the primate amygdala

Faces are a category of images that deserve special consideration when looking at visual responses in the primate amygdala. The purpose of this section is to catalogue what is currently known about face responses in the primate amygdala and it is mostly derived primarily from work that has been done in humans and monkeys.

In the macaque, face responsive cells in the amygdala are predominantly found in the lateral and basolateral nuclei (Hoffman, Gothard et al. 2007). Subsets of these face-selective cells have been shown to respond only to a specific monkey (i.e., identity coding cells) or only to a specific facial expression (expression-coding cells) (Gothard, Battaglia et al. 2007). Using movies of conspecifics, a population of cells in the amygdala was shown to have eye-contact sensitive responses (Mosher, Zimmerman et al. 2014) that were only apparent when the gaze of the monkey in the movie met the gaze of the monkey *viewing* the movie. This suggests that face responses in the amygdala are not just a coarse response but actually contain a lot of structure that might only be apparent by monitoring neural activity while also tracking eye

movements. The diverse tuning properties of face-responsive amygdala cells, which we describe here, also allude to a possible computational role that might distinguish face processing in the amygdala from that seen in areas of the cortex. These responses are not passive (i.e., simply signal the presence of a face), but could be used to make inferences in a social setting. They can possibly be used to infer emotion (recall the study by Wang et al. which showed parametric encoding of the intensity of an emotion), or facilitate social communication through eye gaze. The latter is one of the most powerful instances of non-verbal communication; it can be used to signal (e.g. you can cue someone to look at something with your eyes) or to cue for intimacy (among others).

Evidence of finer structure in the face responses has also been found in humans. In a study using stimuli that only revealed parts of a face, Rutishauser et al. discovered cells that were selective for regions of interest within the face (i.e., mouth, left eye, and right eye) as well as cells that only responded to the presence of the entire face (Rutishauser, Tudusciuc et al. 2011). These kinds of responses can also be seen outside of the amygdala (in face patches for example) so again the question arises, what is the computational advantage of having cells with tuning for face-parts in the amygdala? One possible explanation could be that amygdala responses do not only reflect the perception of faces but are actually used in the production of facial expressions (Livneh, Resnik et al. 2012). This hypothesis is based on anatomical evidence of projections from the amygdala to cingulate motor areas. This is a very exciting hypothesis because it suggests that the amygdala forms a closed control-system suited for adaptive social behavior; on the one end monitoring the environment for socially relevant cues and on the other end using these cues to generate contextually meaningful behaviors.

1.4 Using more naturalistic stimuli in probing amygdala responses

In relating behavior to patterns of neural activity, most electrophysiological experiments (specifically visual experiments in human and non-human primates) tend to use stimuli that have been stripped of all non-essential elements. The reason for this is that the more variables you remove from the experimental paradigm, the easier it becomes to interpret the associated

neural patterns. As an example, most of the literature detailing the properties of face cells in the primate amygdala has been based on the presentation of static images of faces. This experimental scenario is far removed from our own experience as we move around the world. In the real world, we are constantly updating our view by directing our gaze at different objects in our environment. Furthermore, we rarely experience these objects in isolation; they are usually embedded in a sea of other objects (i.e. crowdedness), they are often themselves animate (moving and emoting), and their relationship to us can vary (ex. far or close). Given this constant competition for our attention, on average we sample the items in our environment for brief periods of time (200-500ms) before moving to the next items that attract our attention. Of course, the actual numbers can vary widely depending on what we are doing and what is around us, but the point is that the processes underlying naturalistic viewing behavior usually occur on short timescales and under information-rich conditions.

On the one hand, the importance of using more ethologically valid behavioral tasks cannot be understated. On the other hand, neural data is much easier to interpret if the number of variables is small. Improvements in sensors that allow us to measure behavior, such as eye trackers, and methodological innovations in automatic labeling of behavior using machine vision, have allowed researchers to probe more and more interesting behaviors in the laboratory. In some cases, these new experimental paradigms have expanded our understanding of the associated brain structures involved in the behavior, and in other cases, they have reshaped it (Sheinberg and Logothetis 2001). As an example, the use of more naturalistic stimuli in recent years (e.g. movies) has led to findings that would have otherwise been overlooked. In a recent study, Mosher et al. (Mosher, Zimmerman et al. 2014) annotated instances when the monkey (frame-by-frame annotation, under free-viewing conditions) fixated on the eyes of conspecifics in the video. In doing so, they found eye cells in the macaque amygdala that are modulated by directed versus averted gaze. They discovered these cells only after they replaced static images with videos of conspecifics and allowed the monkey to freely view the movie (Gothard, Mosher et al. 2017).

1.5 Representation of memory and visual category in the human medial temporal lobe

So far, we have focused on face processing in the amygdala, but cells in the medial temporal lobe -- anatomically comprised of the hippocampus, amygdala, fornix, and the surrounding perirhinal, entorhinal, and parahipocampal cortices – are selective for a wide variety of visual categories and concepts (of which faces are an instance). Other cells in the MTL signal whether a stimulus is new or old. This memory signal can emerge even after a single exposure (Rutishauser, Mamelak et al. 2006). Furthermore, it has been shown that this memory signal is not binary, but rather is modulated by memory strength (as measured by subjective confidence ratings). One important thing to note is that the visual category of a particular image, and whether it is novel or familiar, are independent pieces of information and therefore not correlated (this *mostly* true, since we do have a better memory for particular categories of images). Accordingly, the groups of cells that carry these two types of information are largely disjoint. While these two types of signals in the MTL have been well characterized, it is still unknown if these representations of visual category and memory are modulated by task. The memory signal carried by cells in the human MTL for example, was identified in the context of a recognition task. It is unclear however if this signal would still be present if the subjects were not asked a memory related question.

1.6 Interactions between the medial temporal lobe and the medial frontal cortex

While much is known about the representations of declarative memory in the human medial temporal lobe, little is known about how these representations are accessed in order to make a memory-based decision. The retrieval and use of memory representations for decisions is thought to rely on interactions between the hippocampus and the medial frontal cortex, two areas that are correlated in memory demanding tasks in humans (Anderson, Rajagovindan et al. 2009), macaques (Brincat and Miller 2015), and rodents (Siapas, Lubenov et al. 2005). There is anatomical evidence for both direct (Lavenex, Suzuki et al. 2002) and indirect (Ito, Zhang et al. 2015) pathways between the hippocampus and the medial frontal cortex. Furthermore, electrophysiological studies have shown that theta rhythms in the hippocampus entrain cells in the MFC (Jones and Wilson 2005, Siapas, Lubenov et al. 2005). The extent to which MFC cells cohere to hippocampal oscillations is correlated with performance in

spatial navigation and working memory tasks in animals (Hyman, Zilli et al. 2005, Hyman, Zilli et al. 2010). It is unclear however if this inter-area coherence between MFC and hippocampus is needed for accessing representations of declarative memory in the MTL in order to make a memory-based decision.

1.7 Overview of the thesis

In Chapter II, I will detail the methodology we use to collect single-cell recordings from human subjects. The detail is extensive and covers everything from the surgical methods for implantation of the electrodes to the processing pipeline that we use for the collection and analysis of the neural and behavioral data. In Chapter III of this thesis, I will focus on data collected from single cells in the monkey and human amygdala using concurrent eye-tracking during a free-viewing task as well as a covert attention task (the latter was done in humans only). From there, in Chapter IV, we will look at representations of abstract choice as a function a visual stimulus and task demands. This is the only chapter that includes data recorded outside of the medial temporal lobe. Methodologically, it is also a departure from the rest of the chapters because we also look for the first time at evidence of inter-area communication through measures of spike-field interactions. In Chapter V, I will expand on the implications of these results, while also presenting some additional, related data, as well as a few ideas about follow-up studies and future work.

References

Anderson, K. L., et al. (2009). "Theta oscillations mediate interaction between prefrontal cortex and medial temporal lobe in human memory." Cerebral Cortex **20**(7): 1604-1612.

Brincat, S. L. and E. K. Miller (2015). "Frequency-specific hippocampal-prefrontal interactions during associative learning." Nature neuroscience **18**(4): 576.

Fried, I., et al. (1997). "Single neuron activity in human hippocampus and amygdala during recognition of faces and objects." Neuron **18**(5): 753-765.

Gothard, K. M., et al. (2007). "Neural responses to facial expression and face identity in the monkey amygdala." Journal of neurophysiology **97**(2): 1671-1683.

Gothard, K. M., et al. (2017). "New perspectives on the neurophysiology of primate amygdala emerging from the study of naturalistic social behaviors." Wiley Interdisciplinary Reviews: Cognitive Science.

Hoffman, K. L., et al. (2007). "Facial-expression and gaze-selective responses in the monkey amygdala." Current Biology **17**(9): 766-772.

Hyman, J. M., et al. (2005). "Medial prefrontal cortex cells show dynamic modulation with the hippocampal theta rhythm dependent on behavior." Hippocampus **15**(6): 739-749.

Hyman, J. M., et al. (2010). "Working memory performance correlates with prefrontal-hippocampal theta interactions but not with prefrontal neuron firing rates." Frontiers in integrative neuroscience **4**: 2.

Ito, H. T., et al. (2015). "A prefrontal-thalamo-hippocampal circuit for goal-directed spatial navigation." Nature **522**(7554): 50-55.

Jones, M. W. and M. A. Wilson (2005). "Theta rhythms coordinate hippocampal–prefrontal interactions in a spatial memory task." PLoS biology **3**(12): e402.

Kreiman, G., et al. (2000). "Category-specific visual responses of single neurons in the human medial temporal lobe." Nature neuroscience **3**(9): 946-953.

Lavenex, P., et al. (2002). "Perirhinal and parahippocampal cortices of the macaque monkey: projections to the neocortex." Journal of Comparative Neurology **447**(4): 394-420.

Livneh, U., et al. (2012). "Self-monitoring of social facial expressions in the primate amygdala and cingulate cortex." Proceedings of the National Academy of Sciences **109**(46): 18956-18961.

Mormann, F., et al. (2011). "A category-specific response to animals in the right human amygdala." Nature neuroscience **14**(10): 1247-1249.

Mormann, F., et al. (2014). "Visual cognitive adventures of single neurons in the human medial temporal lobe." Single Neuron Studies of the Human Brain: Probing Cognition: 121-150.

Mormann, F., et al. (2008). "Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe." Journal of Neuroscience **28**(36): 8865-8872.

Mosher, C. P., et al. (2014). "Neurons in the monkey amygdala detect eye contact during naturalistic social interactions." Current Biology **24**(20): 2459-2464.

Rutishauser, U., et al. (2006). "Single-trial learning of novel stimuli by individual neurons of the human hippocampus-amygdala complex." Neuron **49**(6): 805-813.

Rutishauser, U., et al. (2011). "Single-unit responses selective for whole faces in the human amygdala." Current Biology **21**(19): 1654-1660.

Rutishauser, U., et al. (2013). "Single-neuron correlates of atypical face processing in autism." Neuron **80**(4): 887-899.

Rutishauser, U., et al. (2015). "Representation of retrieval confidence by single neurons in the human medial temporal lobe." Nature neuroscience **18**(7): 1041-1050.

Sheinberg, D. L. and N. K. Logothetis (2001). "Noticing familiar objects in real world scenes: the role of temporal cortical neurons in natural vision." Journal of Neuroscience **21**(4): 1340-1350.

Siapas, A. G., et al. (2005). "Prefrontal phase locking to hippocampal theta oscillations." Neuron **46**(1): 141-151.

Wang, S., et al. (2014). "Neurons in the human amygdala selective for perceived emotion." Proceedings of the National Academy of Sciences **111**(30): E3110-E3119.

Wang, S., et al. (2017). "The human amygdala parametrically encodes the intensity of specific facial emotions and their categorical ambiguity." Nature Communications **8**.

Chapter II: Surgical and Electrophysiological Techniques for Single-Neuron Recordings in Human Epilepsy Patients

2.1 Overview

Extracellular recordings of single-neuron activity in awake behaving animals are one of the principal techniques used to decipher the neuronal basis of behavior. While only routinely possible in animals, rare clinical procedures make it possible to perform such recordings in awake human beings. Such human single-neuron recordings have started to reveal insights into the neural mechanisms of learning, memory, cognition, attention, and decision-making in humans. Here, we describe in detail the methods we developed to perform such recordings in patients undergoing invasive monitoring for localization of epileptic seizures. We describe three aspects: the neurosurgical procedure to implant depth electrodes with embedded microwires, electrophysiological methods to perform experiments in clinical settings, and data processing steps to isolate single neurons. Together, this chapter provides a comprehensive overview of the methods needed to perform single-neuron recordings in humans during psychophysical tasks.

2.2 Introduction

Invasive intracranial EEG (iEEG) monitoring is routinely performed in patients who are believed to have localization-specific epilepsy, but the exact source of the seizure onset cannot be identified using non-invasive methods such as scalp EEG, MRI scans, PET and SPECT studies, or MEG. For these patients, either surface "grid" electrodes or penetrating "depth" electrodes provide a precise method to better identify seizure onset and spread patterns. "Grid" electrodes are sheets of electrodes imbedded in a thin sheet of silicone. They are typically used to identify the site and spread pattern of neocortical seizures on the brain surfaces, and perform cortical functional mapping via electrical stimulation. In contrast, depth electrodes penetrate the brain surface and pass through both cortical grey and white matter. The distal end of the electrode typically rests in deep cortical or

subcortical locations such as the amygdala, hippocampus, primary visual cortex, or insula. Depth electrodes are typically used to identify seizure onsets in patients suspected of mesial temporal or mesial frontal seizures, although more recently the use of many small stereotactic EEG (SEEG) electrodes has been used as an alternative to grids (Mullin, Shriver et al. 2016). In addition to seizure monitoring, depth-electrode monitoring offers the unique opportunity to address important research questions on the function of the human nervous system (Fried, Rutishauser et al. 2014). For example, our current choice of electrode, called the Behnke-Fried (BF) hybrid depth electrode (Fried, Wilson et al. 1999), has a standard array of 4-8 circular platinum-iridium ECoG electrodes spaced at 5 mm intervals along the electrode shaft, and has a hollow core. Through the hollow core, we thread a bundle of 9 microwires (40 μm diameter, platinum-iridium) (see Fig. 1A). These wires are contained in an insulated covering except on the distal end, where they come out in a "flower spray" configuration. The microwires and associated macro-electrode external sheath are FDA approved, and manufactured by Adtech Instrument Corp (Racine, WI). The wires extend 15 mm from the shaft. Eight of the wires are insulated while an additional single referential ground wire is uninsulated. The assembly also has a green shrink-wrap sheath that sits over the insulation and is used during insertion to protect the wires (see below for details). We use the BF electrodes specifically to record multiunit and single unit extracellular activity at the most medial aspect of the electrode target. At present, we have not found any other electrodes that can reliably record single unit activity along the shaft of the main electrode, although newer technologies are being developed for this purpose. Importantly, no additional risk over standard clinical procedures is incurred by inserting microwires in addition to standard depth electrodes (Hefft, Brandt et al. 2013, Schmidt, Wu et al. 2016). The focus of this chapter is to provide a detailed description of the surgical methodology involved in the insertion of hybrid depth electrodes. In addition, we also briefly summarize subsequent methods to obtain reliable single unit recordings from the microwires in a clinical scenario. Our intention is that this detailed, step-by-step description will prove a useful guide to others interested in performing recordings in humans at the single-neuron level. Of note, there are several alternate techniques for inserting depth electrodes (see (Mehta, Labar et al. 2005, Misra, Burke et al. 2014)). We

describe the method we have successfully employed for the last 12 years, but acknowledge that other methods or modifications may be equally successful.

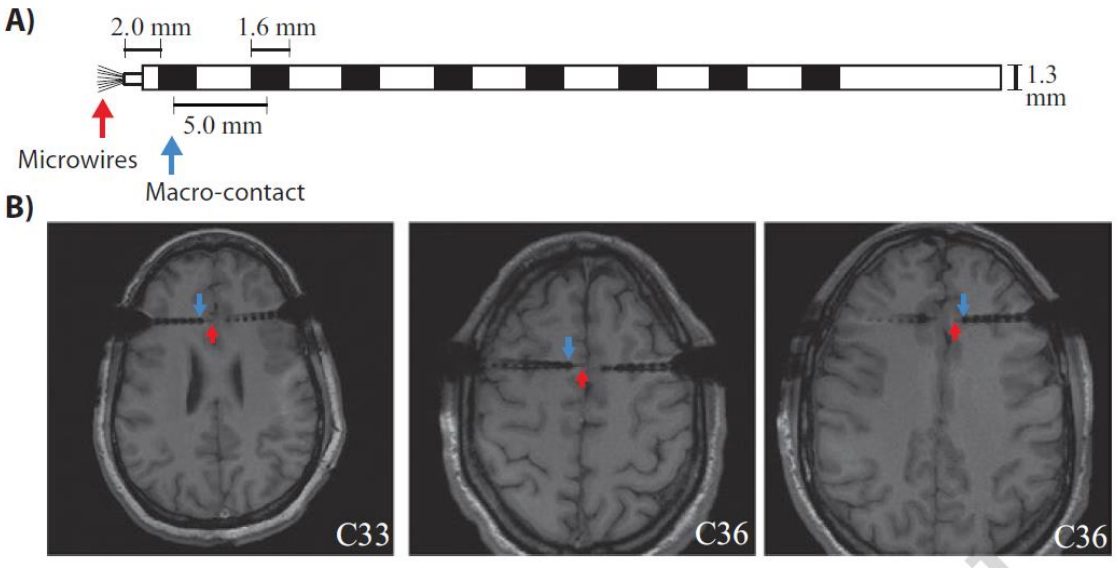


Figure 2.1: Electrodes used and postoperative MRIs. **(a)** Sketch of the hybrid macro-micro depth electrode. **(b)** Example postoperative MRIs illustrating the depth electrode placement.

2.3 Surgical Methods

2.3.1 Target Selection

Placement of depth electrodes must always be dictated primarily by clinical concerns. Patients are undergoing depth electrode monitoring for the primary purpose of identifying a seizure focus. Because insertion of depth electrodes can carry substantial risks such as brain bleeding, stroke, infection, and even death, strict ethical standards must be maintained at all times (Mamelak 2014). Thus, it is unethical and unjustified to insert electrodes in non-clinically relevant areas, or regions used only for research application.

Failure to follow such strict ethical standards is likely to lead to potential harm to patients, which can never be justified.

In general, patients undergoing depth electrode monitoring fall into two categories: (1) seizures are suspected to arise from a medial temporal or limbic structure, but non-invasive monitoring and imaging tests are not sufficient to justify proceeding directly to a surgical intervention. Common examples of this include patients with suspected unilateral onset of seizures in the hippocampus or amygdala, but patients do not meet so called "skip" criteria, so that depth electrode monitoring is used to confirm that all the seizures arise from one mesial temporal lobe versus having bilateral independent seizure onsets, or evidence that the seizures do not arise from the mesial temporal lobe at all. Another common situation is a case where the patient is believed to have localization specific epilepsy, but non-invasive monitoring cannot reliably identify the site. In those cases, depth electrode monitoring is used both to determine lateralization (i.e. what hemisphere does a seizure focus arise from) and localization (i.e. from what lobe of the brain or general region does the seizure arise from). Often in these cases, patients subsequently go on to subdural grid or high density SEEG monitoring to further localize the seizures.

For most typical depth electrode cases, we rely upon orthogonal trajectories, and place bilateral symmetric electrodes. The typical medial targets are: amygdala, mid-body of hippocampus, medial orbito-frontal cortex (OFC), anterior cingulate cortex (ACC), and pre-supplementary motor area (pSMA). In addition, electrodes may often be placed in the parahippocampal gyrus, insula (frontal or temporal opercular regions), parietal cortex, or in any overt structural abnormalities such as cortical dysplasias or regions of gliosis. Thus, in general, our patients are implanted with 5- 8 electrodes in each hemisphere (see Table 1 for typical targets). Note that for all targets, the electrode tip is centered approximately 5 mm more lateral than the desired recording site, to allow room for the microwires to protrude from the end of the macroelectrode into cortical structures.

Table 2.1: Example electrode locations

Target	Approximate anterior–posterior location	Approximate vertical location	Other
Orbitofrontal cortex	25 mm anterior to anterior commissure		
Dorsal anterior cingulate cortex	18–30 mm anterior to anterior commissure	Mid-body of cingulate cortex	
Pre-SMA	5–8 mm anterior to anterior commissure	20–30 mm above the AC	
Amygdala	Anterior to temporal horn of ventricle	3–5 mm lateral to the uncus	Medial aspect of amygdala, basolateral nucleus
Hippocampus	10–15 mm posterior to amygdala target	Mid-body of hippocampus, slightly superior	Tip in CA1–3 rather than dentate gyrus

2.3.2 Stereotactic targeting

Depth electrodes need to be inserted with a high degree of precision. Both the final target position as well as the trajectory from the surface of the brain to the target must be precisely planned to avoid injury to vascular structures such as veins and arteries on the brain surface and deep within the brain. Accidental puncture of veins and arteries is the primary cause of morbidity (brain injury) from depth electrode insertion, and these structures have substantial patient-specific variability. To accomplish this task, we rely on the use of frame-based stereotaxis. We utilize a Codman-Roberts- Wells (CRW) stereotactic frame (Integra) and an attachment to the CRW frame that has been specifically designed for orthogonal depth electrode placements (Cosgrove Depth Electrode Insertion Kit, Adtech Instrument Corp). Unfortunately, we are not aware of a similar insertion kit design to work with other common stereotactic frames l, although the design could easily be modified for those systems. Frame-based stereotaxic methods have a targeting accuracy of less than 1 mm, and can be used with a variety of commercially available stereotactic planning software suites. These stereotactic planning systems are

routinely utilized by the majority of neurosurgical centers around the world. With this method, a metal frame is attached to the patient's head using 4 disposable screws that penetrate the skin and press on but do not go into the skull (Figure 1). We utilize ear bars inserted into the auditory canals of the patient at the time of frame placement to ensure a perfectly centered and orthogonal frame placement.

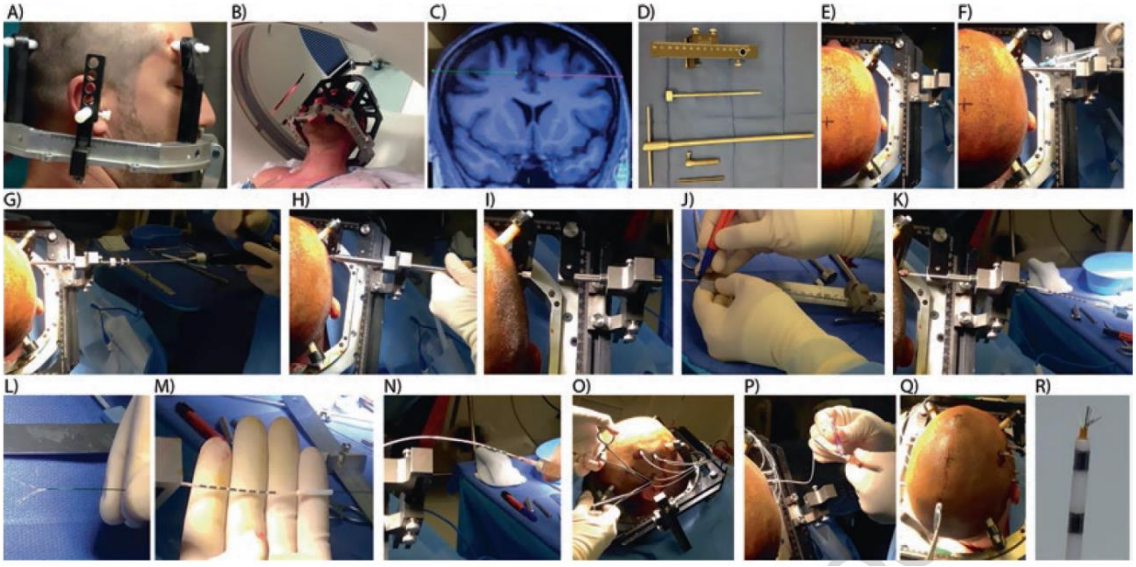


Figure 2.2: Stepwise process of depth electrode insertion (A) Stereotactic frame placement. (B) CT scan with fiducial localizer to create patient specific coordinate system. (C) Screenshot from planning system (Framelink, Medtronic) used to determine stereotactic coordinates and trajectories. (C) Tools used for implantation (Cosgrove Depth guide, coring tool, anchor bolt wrench, reducing tubes for electrodes). (E) CRW base frame with Cosgrove Depth Guide mounted on patient right. (F) Coring tool inserted to skin entry point, and lidocaine injection. (G) Hand held twist drill passed directly through coring tube to drill entry point and anchor bolt site. (H) Bolt insertion using hex wrench. (I) Anchor bolt inserted in skull, Cosgrove guide mounted with reducing tube for electrode insertion. (J) Marking length of electrode from end of reducing tube to target. (K) Inserted hybrid macro-electrode before removal of electrode stylet, inserted to length measured in (J). (L) Microwires (before cutting) together with green sheath. (M) Insertion of microwires through already implanted macroelectrode. (N) Final coupling of macro-and microelectrode. (O) Securing electrode by tightening anchor bolt after microwires are inserted. (P) Application of surgical glue to distal end to secure micro-and macroelectrode coupling. (Q) Insertion of ground/reference strip. (R) Assembled hybrid macroelectrode with microwires extending (after deplant).

2.3.3 Surgical protocol – pre-operative procedures

Pre-operative Area:

1) Full head shave.
2) Intravenous propofol sedation given by anesthesiologist.
3) CRW stereotactic head frame applied using standard neurosurgical methods (Fig. 2A):

a. Local anesthetic injected at each pin site (Albright's solution, a pH balanced mixture of 1% Lidocaine, 0.25% Marcaine, sodium bicarbonate, and epinephrine).
b. Frame positioned with ear bars in bilateral auditory canals and held by technician. Good orthogonal position in all three planes confirmed by direct visualization by surgeon.
c. Skull pins inserted at each of the four posts of the frame and tightened to secure the frame in a good orthogonal trajectory.
d. Apply CT or MRI compatible stereotactic localizer.

CT or MRI scanner:

1) Patient is taken to the CT or MRI suite for scanning once sedation has worn off (typically 5-10 minutes). Our preference is to obtain a high resolution 3T MRI prior to surgery, as this is a routine part of the pre-surgical epilepsy workup, and is therefore almost always available. On the day of the surgery, a CT scan is acquired (0º gantry angle, 1.2 mm slice thickness, axial images from the top of the frame base through the vertex, helical scan mode). Iodine contrast is also given to better visualize vascular structures. The scan is performed with the patient's head secured to a CT fixation holder to assure that the scan is orthogonal (see Fig. 2B). The CT acquisition typically takes under 1 minute. We do not typically acquire a CT angiogram (CTA) sequence, although this can be used to better visualize surface and deep vasculature if desired.
2) The localizer is removed and the patient is transferred to the operating room.

Registration of patient-specific stereotactic space[1]:

1) CT images are transferred to the stereotactic planning computer via a PACS server.

[1] We use the Framelink® Stereotactic Planning Software suite (Stealth Station, Medtronic) for planning at our institution. There are however many alternative solutions that are just as reliable (e.g. Brain lab, Radionics).

2) The previously acquired MRI is also imported. If an MRI is not available, then the CT is not acquired and we acquire a 1.5 T MRI with the stereotactic frame and localizer on, and use these images. MRI should include at least one high-resolution multi-slice (120-180, typically 144) axial or coronal sequence and a similar high-resolution T2 weighted image for best anatomic imaging.
3) Using the planning software (Framelink, Medtronic, Inc), the MRI and CT images are co-registered and aligned (Auto-Merge feature).
4) The fiducial rods from the localizer detected on the CT or MRI scan are then registered, generating a patient-specific stereotactic Cartesian coordinate system.
5) Once this step is complete, a full Cartesian coordinate system is established in which each point in space is associated with a specific Lateral (x), Anterio-Posterio (y), and Vertical (z) coordinate that corresponds to identical x, y, z values on the frame. A full description of stereotactic methods is beyond the scope of this chapter.

Target Calculation (see Fig. 2C):

1) The most medial point for the electrode is chosen on the displayed axial, coronal, and sagittal MRI. Typical targets are listed in Table 1. The monitor cursor is placed on the target point.
2) Clicking the computer mouse on the chosen target generates the x, y, and z coordinate values for that point on the screen in the lower corner of the image panel. These points are recorded on an electrode targeting sheet. The "set target" button on the software is selected.
3) A point is then chosen on the lateral skin surface that is roughly parallel to the targeting the A-P and vertical planes. Once selected, the x, y, and z coordinates are displayed by clicking the mouse similar to how the initial medial target was selected. The point is adjusted with small movements of the cursor to ensure that the A-P (y), and Vert (z) values for the entry point are within 1 mm of the same values for the chosen target point. This assures that the trajectory of the electrode will be orthogonal to the insertion guide and parallel to the base frame. Once this point is established, the "set entry" button is selected, resulting in the display of the electrode trajectory from surface to medial point.

4) Using the trajectory path tool, the trajectory is then followed in multiple planes to ensure that no surface or deep vessels are violated, and that sulci are avoided as much as possible. If need be, the trajectory is then iteratively adjusted until the surgeon is satisfied with the trajectory.
5) This process is repeated for every electrode, generating an average of 10-16 trajectories.
6) The insertion depth of each electrode then recorded on the target-planning sheet as 190 mm minus the absolute value of the lateral coordinate. This measurement determines the stopping point for insertion of the electrode at the lateral edge of the insertion guide (see 2.2.4 for details), as the system is designed so that the back end of the insertion guide with reducing tube in place measures exactly 190 mm from the center of the stereotactic frame.

2.3.4 Surgical protocol – intraoperative procedures

Once all target coordinates and trajectories have been defined, the patient is ready to be implanted. The patient is brought into the operating room. We typically perform these procedures under total general anesthesia with either laryngeal mask or endotracheal tube insertion. However, the procedures can also be performed using only propofol sedation and local anesthetic if desired. In our experience this is not necessary, slows the procedure, and increases patient discomfort. The procedural steps for insertion are as follows.

Patient positioning and prepping:

1) Patient is placed in a semi-sitting "lounge chair" on the operating room table with the CRW frame secured in place using a standard neurosurgical head holder (Mayfield head holder with CRW adaptor plate). The head is position almost upright with the frontal eminence (top of the forehead) uppermost in the field, allowing symmetric access to both lateral sides of the head.
2) The entire head is prepped with an iodine-based antiseptic solution. Care is taken to make sure the prep extends below the zygoma (cheek bone) on both sides and up to the frame rods.

3) The CRW base ring is attached to the head frame and secured with the locking knobs.
4) A sterile "U" drape is draped around the base of the frame but below the base ring and extended around the front of the patient with a full or three-quarter sheet placed to cover the patient's body. This creates a sterile field (also see Note 1).

Equipment Required:

1. Bovie electrocautery unit and cautery pencil.
2. Modified electrocautery stylus (custom-made 225 cm-long insulated rod with 2 cm tip exposed at end, inserts in coring tube).
3. Cosgrove Depth Electrode Insertion Kit (see Fig. 2D), containing:

a. CRW Stereotactic Electrode Insertion Guide.
b. Reducing tube – non-slotted.
c. Reducing tube – slotted.
d. Slotted electrode guide insert (inserted in slotted reducing tube).
e. Coring tool.

4. Handheld neurosurgical twist drill.
5. Disposable drill kit for depth electrode anchor bolts (Adtech DDK2-2.8-30x for standard and BF depths, DDK2-2.4-30X for SEEG anchor bolts and electrodes).
6. Anchor bolts.
7. Appropriate number of BF electrodes.
8. 1x4 or 1x6 contact subdural strip electrode (for ground/reference).
9. Basic surgical instrument tray with forceps, cocker clamps, etc.
10. Small gauge K-wire or Steinman pin.

Anchor bolt placement:

The standard process of inserting BF hybrid depth electrodes using the Cosgrove Depth Insertion Kit is identical for all electrodes and insertion sites. We typically perform all right-side insertions first, followed by all left-sided insertions, as this is far more time efficient.

1. The Cosgrove Depth Guide is placed on the A-P mount side of the CRW base frame and set to the A-P (y) coordinate for the electrode to be inserted (see Fig. 2E) (The CRW uses a Vernier scale to allow for accuracy up to 0.1 mm).
2. The height of the Cosgrove Guide is adjusted to match the Vertical (z) coordinate of the target for the electrode to be inserted.
3. The non-slotted reducing tube is inserted into the guide until flushed with it, and locked in place with the tightening screw.
4. The coring tool is inserted through the reducing tube to the skin, marking the entry point. This point is injected with a 2-3 cm wheel of lidocaine 0.5% with epinephrine (see Fig. 2F).
5. The coring tool is twisted directly through the skin and muscle down to the skull. The sharp edges on the coring tool allow it to cut through the skin and muscle but not penetrate bone. Once making firm contact with bone, it is locked in place with the tightening screw.
6. The modified Bovie tip is inserted directly through the hollow opening in the coring tool down to the bone and coagulation of the deep tissue carried out for several seconds. This prevents bleeding from the muscle that might occur from coring.

a. If no modified Bovie tip is available, the K-wire can be inserted and the distal end of the K-wire touched to the Bovie to transmit current to the deep tissues. If this method is used, care is taken to make sure the K-wire does not touch the side walls of the coring tube, resulting in an electrical short and no tissue coagulation.

7. The handheld twist drill is passed through the coring tube opening to the bone, and a twist drill hole is made in the bone. An adjustable stop on the drill is set to minimize risk of plunging into brain. Ideally, the drill penetrates the skull but stops at the dura. Standard neurosurgical techniques are applied to achieve this depth (see Fig. 2G).
8. The drill is removed and the sharp end of the K-wire is inserted down the coring guide. This allows the surgeon to ensure that the bone has been completely breached and to palpate the dura. The dura is then punctured with the K-Wire. The process can be repeated as needed to ensure complete drilling and dural opening.
9. The coring tool and reducing cannula are removed.

10. An anchor bolt is placed on the distal end of the hex wrench supplied with the Cosgrove Kit (see Fig. 2H). This wrench is designed to match the hexagonal shape of the anchor bolt, with a width that is the same as the Cosgrove Guide. This ensures that the insertion of the anchor bolt will be exactly in line with the drill hole and will remain orthogonal at all times.
11. The hex wrench with anchor bolt is inserted through the insertion guide into the drilled hole through the skin and hand–tightened into the calvarium. This typically requires 15-20 half-turns of the wrench. Care must be taken not to turn too quickly or with too much pressure to avoid fracturing of the anchor bolt or the underlying bone. Typically, the anchor bolt is advanced until the hexagonal aspect of the bolt is touching the skin surface. This ensures excellent purchase in the bone but not too deep a penetration to cause epidural hematoma.
12. Once in place, the hex wrench is removed and the surgeon confirms that the anchor bolt is tightly secured.

Hybrid microwire insertion:

1. The length of the electrode that had previously been measured (Step 6 in "Electrode Trajectory" section above) is noted. The macro-contact portion of the BF hybrid electrode is measured to this length from the distal tip of the electrode on a ruler, with the length marked using a surgical marking pen (see Fig. 2J).
2. The distal end of the microwire bundle is then cut as a single bundle using very sharp tonotomy scissors. We typically cut the wires to be 4-5 mm long for optimal results (see Note 3 and Fig. 3L).
3. The green protective sheath is gently pulled over the microwire bundle to protect the wires during insertion (see Fig. 3L). The sheath should just cover all the wires but not be pulled up too far to avoid it coming off or bending during insertion.
4. The slotted reducing tube and slotted guide cannula are assembled so that the distal ends of both pieces are flush. They are then inserted as a single assembly into the Cosgrove Guide (see Fig. 2I).

5. The K-wire is again passed though the guide cannula and the anchor bolt through the dura to ensure clear passage of the electrode.
6. The BF macroelectrode is inserted through the guide cannula and opening in the anchor bolt and passed until the marked point on the electrode just aligns with the back end of the guide cannula (see Fig. 2K). This is the target depth.
7. With the surgeon carefully holding the end of the electrode that enters the anchor bolt (so that it cannot slip), the electrode stylet is removed. The guide cannula and slotted reducing tube are then unscrewed and gently pulled back and disassembled, leaving only the electrode exiting from the Cosgrove Insertion Guide. The distal electrode is then passed through the Cosgrove Guide opening so that the electrode sits completely outside the Cosgrove assembly. The anchor bolt set screw is tightened one half-turn.

a. A slotted guide cannula is needed because the distal outer dimeter of the BF macroelectrode is larger than the proximal end, and if it were not slotted, it could not be freed from the guide assembly.

b. Placement of the macro-contact and holding it in place while the distal end is removed from the Cosgrove Guide is the most precise portion of the procedure and the easiest place for the surgeon to accidentally move the electrode depth. Care should be taken to ensure that the surgeon has a secure grip on the electrode just as it enters the anchor bolt during this entire process to minimize risk of migration.

8. The surgeon picks up the microwire assembly and hands the distal end of the microelectrode bundle to an assistant to hold and advance, while the tip is inserted in the macroelectrode.
9. The surgeon holds the distal end of the macroelectrode, taking great care to not pull out or advance the electrode (it is not yet secured). He/she then threads the microwire assembly into the hollow opening of the macroelectrode and gently advances it until the distal connecter bushing tightly connects with the distal connector on the macroelectrode (see Fig. 2M). It is advanced until the collar on the back end of the microwire assembly aligns with the blue line present on the macroelectrode, indicating it has been fully inserted with the distal end protruding from the distal end of the macroelectrode by the

amount that was pre-cut in Step 2 (see Fig. 2N). The collar should fit snugly into the distal end of the macrowire assembly.

10. The tightening screw on the anchor bolt is then tightened, first finger tight and then further with a Kocher clamp to lock the entire electrode assembly in place (see Fig. 2O).
11. A small drop of surgical glue (e.g. Dermabond, Indermil) is applied to the electrode-coupling site at the distal end to prevent the assembly from separating (see Fig. 2P).

The process (Steps 1-11) is then repeated for the next electrode, until all electrodes have been inserted.

12. Once all electrodes are inserted, a 1-2 cm incision is made in the midline scalp at the parietal vertex. A hemostat or similar clamp is used to create a small sub-galeal pocket and the 1x4 subdural strip is inserted into the subgaleal space with the contacts pointing outward. This will serve as the ground and reference contacts for the recordings (see Fig. 2O). The incision is closed with a nylon suture, and the electrode tail is secured to the scalp.
13. An A-P and lateral skull X-ray are taken after all electrodes are inserted but prior to completion of the procedure. Review of the X-ray may identify electrodes that are misplaced or accidently pulled out. Any misplaced electrodes are then re-inserted with final placement again confirmed by X-ray.
14. The exact number and/or color scheme for each electrode and its location are double-checked with a technician to ensure correct identification of each wire for subsequent EEG recordings.

Completion of procedures:

Once all electrodes are in place, the head must be properly dressed to prevent infection:

1. The entire head is cleaned as best as possible with several wet lap sponges, removing any dried blood and betadine paint.
2. 1 cm wide x 2-3 cm long strips of Xeroform or similar bandaging material are cut and wrapped around the base of each anchor bolt.
3. The surgeon holds the head while as assistant releases the frame from the Mayfield head holder, unscrews the four skull pins, and removes the CRW frame.

a. Care is taken not to hit any of the protruding anchor bolts or accidentally pull on the electrodes during removal.

4. Several gauze sponges are placed on both sides of the head and a full head dressing is applied. We use two Kerlix rolls for this purpose. The electrode tails are brought out through the top of the dressing with care taken to ensure they are not buried in it. The exiting tails are further covered with additional gauze sponges and secured with silk tape. A Spandage expandable bandage net is also applied.
5. The patient is awakened and taken to the recovery room.

Post-surgery procedures:

1. A non-contrast brain MRI is obtained within 4 hours of insertion. This confirms each electrode tip location and identifies any sites of bleeding.

a. A CT scan is not advisable as there is a tremendous metallic artifact making any interpretation difficult.

2. A-P lateral and submental vertex plain films are taken.
3. The patient is transferred to the Epilepsy Monitoring Unit (EMU) for recovery, and electrodes are hooked up for continuous EEG monitoring, typically the following day.
4. Evaluation of the microwires for detecting single unit and multiunit activity is typically first started 1-2 days after insertion to allow the patient to recover from the procedure.

Removal of electrodes:

Once monitoring is completed, the electrodes and anchor bolts must be removed. This is a straightforward surgical procedure, typically done under propafol sedation. The basic steps are as follows:

1. In the OR, the patient is given propofol sedation. The head dressing is cut off, with care taken not to accidentally cut the electrode wires. The head is placed on a gel donut for support.
2. The head is not cleaned with betadine until after the electrodes are removed.
3. Starting with the right side, the head is turned to the left. The surgeon puts on gown and gloves.

4. Using the Kocher clamp, the tightening screw on each electrode is loosened. We typically loosen all screws at one time.
5. Each electrode is then pulled out from the anchor bolt and inspected. The distal microwires should be visible (see Fig. 2R) and the entire electrode array removed (also see Note 2).
6. Using the hex wrench, unscrew each anchor bolt and remove it.
7. The entire side of the head where the bolts were removed is painted with betadine scrub paint.
8. Each insertion site is closed with a single 2-0 or 3-0 nylon suture. The closing stich must tightly bring together the skin edges and should be inspected to ensure no egress of CSF.
9. Once one side is completed, the head is turned to the other side and the same process (1-8) repeated for the other side,
10. Once all electrodes are removed, the head is cleaned with a moist lap sponge and water. Antibiotic ointment is then applied to each suture site. No other dressing is required.
11. The patient is awakened and returned to his/her room.

2.4 Methods for data acquisition and behavioral testing

A typical intracranial recording setup relies on three separate computers: an acquisition system, a stimulus presentation system, and an eye-tracking system. Below, we briefly outline the configuration of each setup. Together, we have found this to be a very reliable setup for use in the clinical setting.

2.4.1 Data acquisition system

For data acquisition, we use the Atlas system from Neuralynx Inc. All signals from the microwires are pre-amplified on the head with small pre-amplifiers (headstages) that attach directly to the pig-tail connector of the microelectrode. All microwire recordings are performed broadband (0.1 Hz – 9 kHz bandpass filter) and are sampled at 32kHz. In addition, this system allows the monitoring of all signals originating from macroelectrodes (depth electrodes, grids) and to pass these signals on to a clinical system running in parallel.

Together, this configuration allows us to connect only a single system to the patient, which lowers noise and avoids interference (see Notes 4 and 5). While we monitor the broadband recordings throughout the experiment, all processing of the data (i.e. filtering, spike detection, spike sorting) is redone during offline analysis. We typically set the input range to ±2500 μV, resulting in <1 μV resolution. This is especially critical for spike sorting (see below), which relies on the shape of the spike waveforms themselves. Alternative products from other manufacturers (including Blackrock Microsystems Inc. and TDT Inc.) offer similar solutions to the one we described.

2.4.2 Stimulus Presentation and eye tracking

We implement all experimental tasks in Matlab with Psychophysics Toolbox (Brainard 1997, Pelli 1997). This well-tested and extensively utilized toolbox has been utilized by numerous human intracranial experimenters and is well-suited for this purpose. We typically show stimuli on a 19-inch screen with a resolution of 1024 x 768 pixels. The screen is supported by an arm mount and also carries the camera and infrared light source for the eye-tracker. We monitor monocular gaze position with a 500Hz sampling rate with an Eyelink 1000 system (SR Research Inc.). We utilize a 9-point calibration grid to determine the eye-to-screen coordinate transformation. Throughout a typical experiment, we can monitor eye position with an accuracy of 0.42 DVA ± 0.15.

2.4.3 Response boxes and keyboards

To collect responses from the subjects, we primarily use the RB-740 and the RB-844 response pads (Cedrus Inc.) These response pads offer more reliable timing compared to a regular keyboard. Also, they are fully customizable, contain only a few buttons, and can be changed from experiment to experiment. As a result, we find response pads to be easier to use for patients.

2.4.4 Synchronization and data transfer

Since the three systems are independent, it is essential to synchronize behavioral events. We use the stimulus presentation system as the master system. Whenever a significant behavioral event occurs (stimulus onset, stimulus offset, button press), this system sends an event to both the acquisition system as well as the eye-tracking system. This is achieved by utilizing the parallel port to send a signal to the transistor-to-transistor logic (TTL) input port on the data acquisition system. The same events are also sent to the eye-tracking system utilizing an Ethernet IP connection and the EyeLink toolbox (Cornelissen, Peters et al. 2002). This way, the point of time at which each behavioral event occurred is known on all three systems, despite the underlying clocks not being synchronized. Also see (Rutishauser, Kotowicz et al. 2013) for further technical details on how to communicate between the three systems involved and how to utilize these connections for real-time closed-loop experiments.

2.4.5 Methods for data processing

Spike sorting is the process of extracting action potentials from the raw intracranial recording and attributing them to a particular neuron ("unit"). As is the case with all unsupervised clustering problems, one of the main challenges of spike sorting is estimating the number of neurons (i.e. clusters) that a given electrode is "listening" to. While there are many spike-sorting solutions available (see Table 2), they all execute a similar workflow: signal conditioning and filtering of the raw trace, followed by spike detection and alignment, feature extraction, and finally, clustering (Lewicki 1998, Gibson, Judy et al. 2012). The features used for clustering are either the raw spike waveforms or derivatives thereof. To judge the quality of sorting, additional properties of the spike train associated with a given cluster have to be considered, including the distribution of inter-spike intervals, firing rates, and autocorrelations. Together, these pieces of information provide evidence for whether a given cluster can be considered representative of a single neuron or not. While the focus of this chapter is on the surgical aspects of depth electrode implantation, we briefly summarize the standard steps we utilize below. Please see (Rutishauser, Cerf et al. 2014) for further details.

2.4.6 Filtering

The first step in the processing pipeline is to remove low frequency content from the raw trace (Fig. 3a) by band-pass filtering the raw signal in the 300-3000 Hz frequency range (Fig. 3b). In order to preserve the shapes of the spike waveforms, it is important that the filtering process does not introduce phase distortions (Quiroga 2009), which is achieved by using a zero-phase digital filter ((Fig. 7). For real-time applications, such filtering is not possible because it is non-causal. In that case, the alternative is to use a linear phase FIR filter and directly account for the group delay introduced by the filter (delay = (L-1)/2, where L is the filter length).

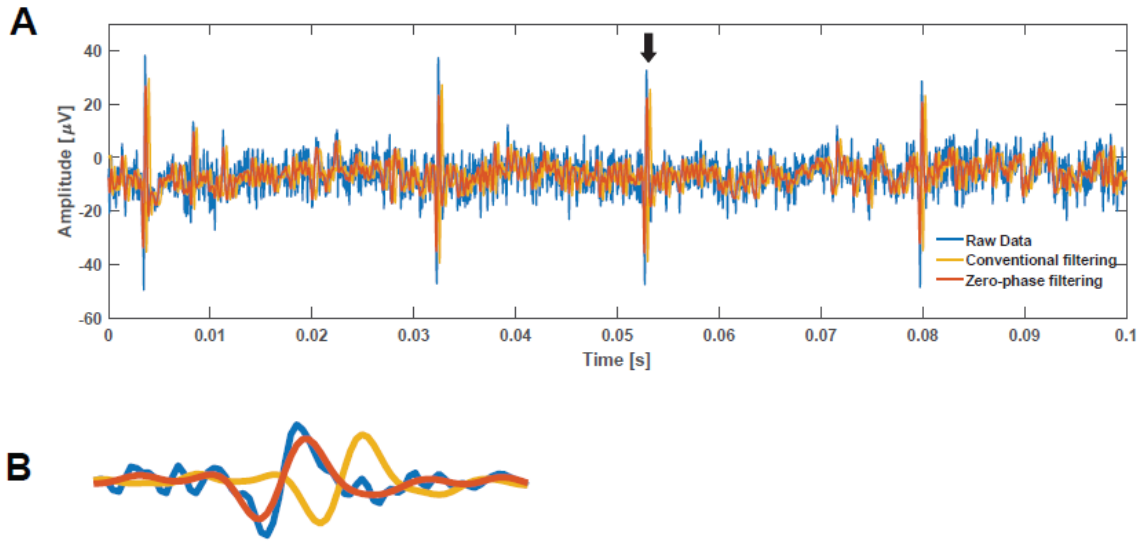


Figure 2.3: Zero-phase filtering (a) Bandpass filtering is a common first step in improving the signal-to-noise ratio of spike waveforms. By retaining spectral information in a specific frequency range (300-3000 Hz for the example shown here), we can improve detection and sorting of spike waveforms. The way in which the filtering is performed however, can greatly change your results. Here we show the results when we filter the raw data (blue trace, Fs = 32000Hz) with no phase distortion (red trace) and when we filter in the conventional way (yellow trace). Zero-phase filtering was implemented with the Matlab function *filtfilt.* (b) While both methods preserve the waveform shape, conventional filtering delays the spike waveform while zero-phase filtering does not.

2.4.7 Spike detection and extraction

In order to extract the action potentials from the filtered signal, we assume that the individual spikes are above the noise floor. Since the noise floor may not necessarily be stationary, we use a time-dependent threshold that is a function of the underlying noise properties of the

signal. Specifically, the threshold is set to be a multiple (typically around 5) of the estimated standard deviation of the filtered trace. While this may work perfectly well for spikes that have large waveforms, it may miss some of the smaller spikes that are much closer to the noise floor. A few simple techniques can help improve the signal-to-noise ratio of spikes and therefore improve detection. One such technique is to use the energy of the signal instead of the raw trace (Bankman, Johnson et al. 1993). The energy operator amplifies small differences between the spike amplitude and the noise floor, making it easier to set a threshold for spike detection. Since the processes that we are interested in (i.e. spikes) unfold over approximately 1ms, we compute the local energy of the signal at that time scale by convolving it with a rectangular kernel of 1ms width. This results in an "energy signal (Fig. 3c), which is then thresholded. The threshold is set to a multiple of the standard deviation of the energy signal (here, 5x s.d.). The threshold parameter (i.e. the multiple of the standard deviation) can vary case by case. For a channel that has very few spikes, we may need to set this parameter higher than usual to avoid picking up noise. On the other hand, in situations with cells of very high firing rates with large amplitude waveforms, the threshold has to be set lower than usual (i.e. 3 or 4x s.d.). For each threshold crossing, we extract a fixed number of samples before and after from the raw signal (typically, 2.5ms total length). This 2.5ms long trace is the waveform of the spike and forms the basis for all processing that follows.

2.4.8 Spike sorting (clustering)

Spike sorting involves two steps: identification of features from each waveform followed by unsupervised clustering of these features. The most commonly used feature for clustering is the spike waveform. The goal is to identify features of the waveform that maximally separate different cells. On one extreme, we could use a single scalar, such as the peak-to-trough amplitude or the spike width. On the other extreme, we could use the entire waveform and clustered in this N-dimensional space (where N is the number of samples that make up the waveform). An alternative approach is to capture as much of the variance in the waveforms as possible using only a few dimensions. This can be achieved by utilizing a dimensionality reduction technique such as principal component analysis (PCA).

Once spike waveform features have been identified, an unsupervised clustering algorithm is used to partition the space. In practice, a variety of algorithms have been used for human single-neuron recordings, including OSort, Wave_clus, and variants of Klustakwik/Klusta (see Table 2). Some are parametric and make assumptions about the underlying distribution of the data (such as Gaussian Mixture Models). Others are nonparametric and rely on heuristics computed directly from the data. We utilize the OSort (Rutishauser, Schuman et al. 2006) algorithm for spike detection and clustering. OSort uses a distance metric between the raw waveforms for clustering, runs spike-by-spike (online), and determines the number of clusters automatically. Other spike sorting approaches used for human single-neuron recordings are example pipelines include Wave_clus (Quiroga, Nadasdy et al. 2004) and Klusta (Rossant, Kadir et al. 2016).

Table 2.2: Spike sorting algorithms commonly used for human single-neuron recordings

Name	Filtering (Hz)	Automatic	Detection	Features used	Sorting algorithm	Notes
OSort [18]	Yes	300–3000	Can be but manual refinement recommended	Signal energy, computed locally with 1 ms kernel	Full, raw, decorrelated waveforms	Discriminative, with manual oversight	Does not generalize to polytrodes
Klusta [20, 39]	No	500—95% of Nyquist	Can be but manual refinement recommended	Dual threshold based on filtered trace; flood-fill algorithm used for detecting events over time and electrodes (i.e., space)	Principal components of spatiotemporal spike signature	Parametric, discriminative, maximum likelihood fitting using GMM with expectation maximization (EM)	Tailored for dense electrode arrays, activity between channels not independent
Wave_Clus [19]	No	300–3000	Yes	Amplitude threshold based on filtered trace; multiple of noise estimate	Wavelet coefficients of waveform	Nonparametric, discriminative (superparamagnetic clustering)	Can deal with polytrode data

2.4.9 Quality metrics

We rely a list of quantitative metrics to assess how likely a given cluster represents a single neuron and to assess whether a given cluster is over- or under-merged.

Stable waveform: A key metric is the peak-to-trough amplitude of the spikes associated with a cluster as a function of time. Ideally, the amplitude of the spike should remain constant throughout the experiment (see Fig. 6B). Large deviations in the shape of the waveform are usually indicative of electrode movements (the electrode changes position with respect to the cell) or of an artifact of the spike-sorting algorithm (ex. two clusters were merged when they should not have been, or a single cluster was split into two when it should not have been). This can often be corrected by manually merging clusters.

Stable firing rate (on long enough time scales): A second metric to tracking the stability of a cell is firing rate as a function of time. While there may be task-dependent modulation at finer time scales, on long enough time scales, the average firing rate should be relatively stable. As in the case of the peak-to-trough amplitude, a large deviation in the average firing rate of a cell is usually indicative of electrode movement or over-splitting in spike sorting (see Fig. 6B for an example). This can often be corrected by manually merging clusters.

Inter-spike interval histogram: One of the unmistakable features of a cell is the distribution of its inter-spike intervals (ISIs). The ISI histogram can be used in two ways: (1) to tell the difference between cells that have similar waveform shapes but might be functionally different, and (2) to verify that there are no violations of the refractory period (i.e. there are few ISIs < 3ms). Violations of the refractory period are indicative of a cluster representing multi-unit activity.

Alignment check: An important check of the quality of a cluster is the distribution of the peak amplitude across all the waveforms. This distribution should be unimodal and tightly clustered around the true mean. A multimodal distribution is indicative of over-clustering (two or more clusters have been mistakenly merged into one) or misalignment of the waveforms (see Fig. 6A).

Cluster quality checks: We typically provide, for each cluster included in a paper, histograms of a number of spike isolation quality metrics to allow an assessment of how well-separated the cells included in a study were (see supplementary Figures in (Rutishauser,

Ross et al. 2010, Kaminski, Sullivan et al. 2017) for examples). These include: projection tests between all possible pairs of clusters on the same wire (Pouzat, Mazor et al. 2002, Rutishauser, Schuman et al. 2006, Rutishauser, Ye et al. 2015), isolation distance (for each cluster versus all other detected spikes on a wire), L-ratio (Schmitzer-Torbert, Jackson et al. 2005, Hill, Mehta et al. 2011), %ISI violations <3ms, and signal-to-noise of the mean waveform of a cluster.

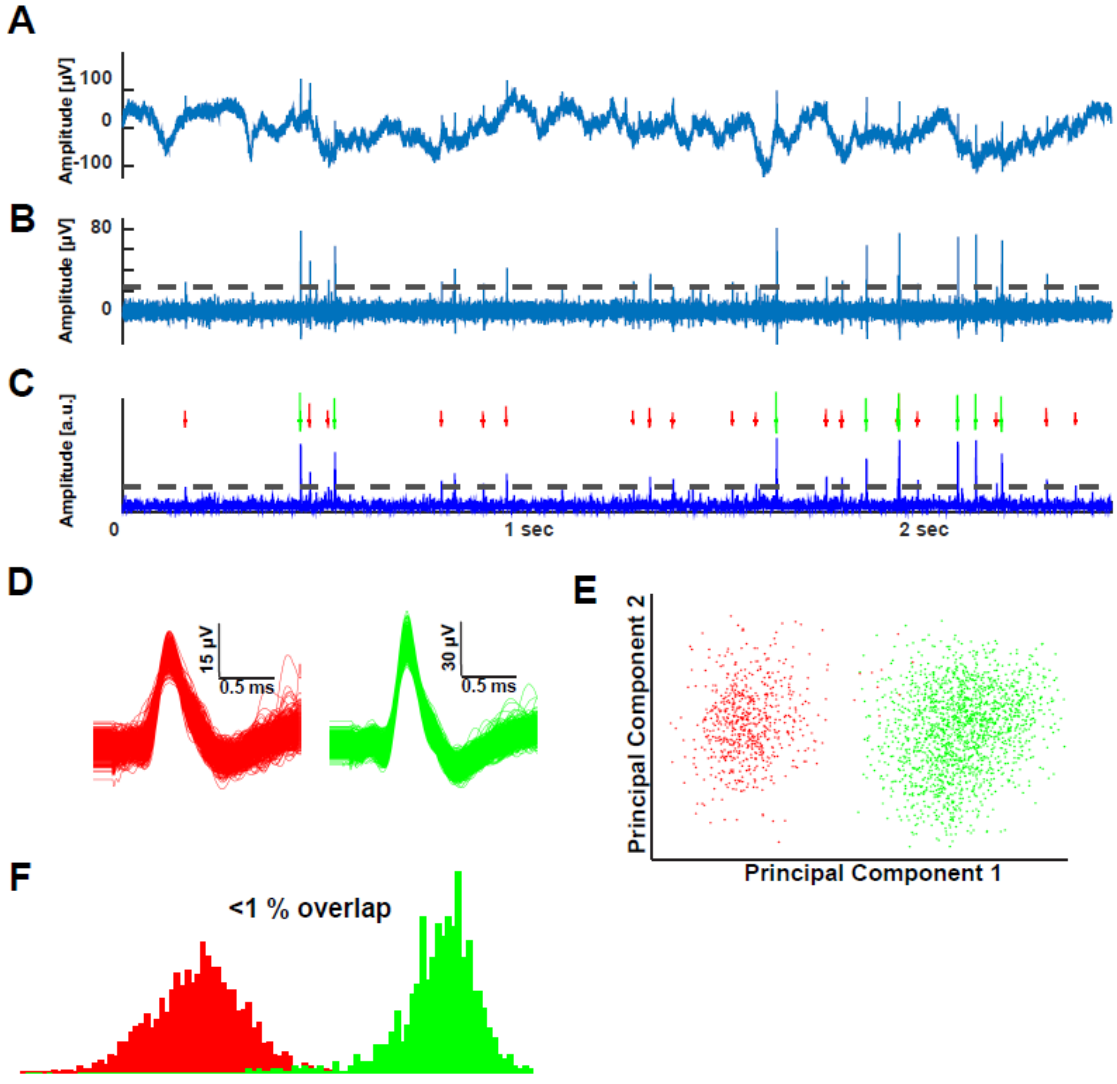


Figure 2.4: **Spike detection and sorting.** (A) Example broadband recording from a microwire in the amygdala (with bipolar referencing). (B) The same trace as in (A) but bandpass filtered in the 300 – 3000 Hz band. (C) Signal used for spike detection (green line is the threshold). Signal shown is the local energy (with a 1ms kernel), which improves the SNR of spikes with respect to the baseline. (D) After spike detection and alignment, two

prominent waveform shapes, each one belonging to a different underlying unit (green and red), were identified on this channel. (E) The individual waveforms (256 samples per electrode) from the two clusters projected into principal component space. The clusters are well-separated. (F) Projection test to validate the separation between the two putative single-units (clusters). Shown are two overlapping histograms, each corresponding to one cluster. There was less than 1% overlap.

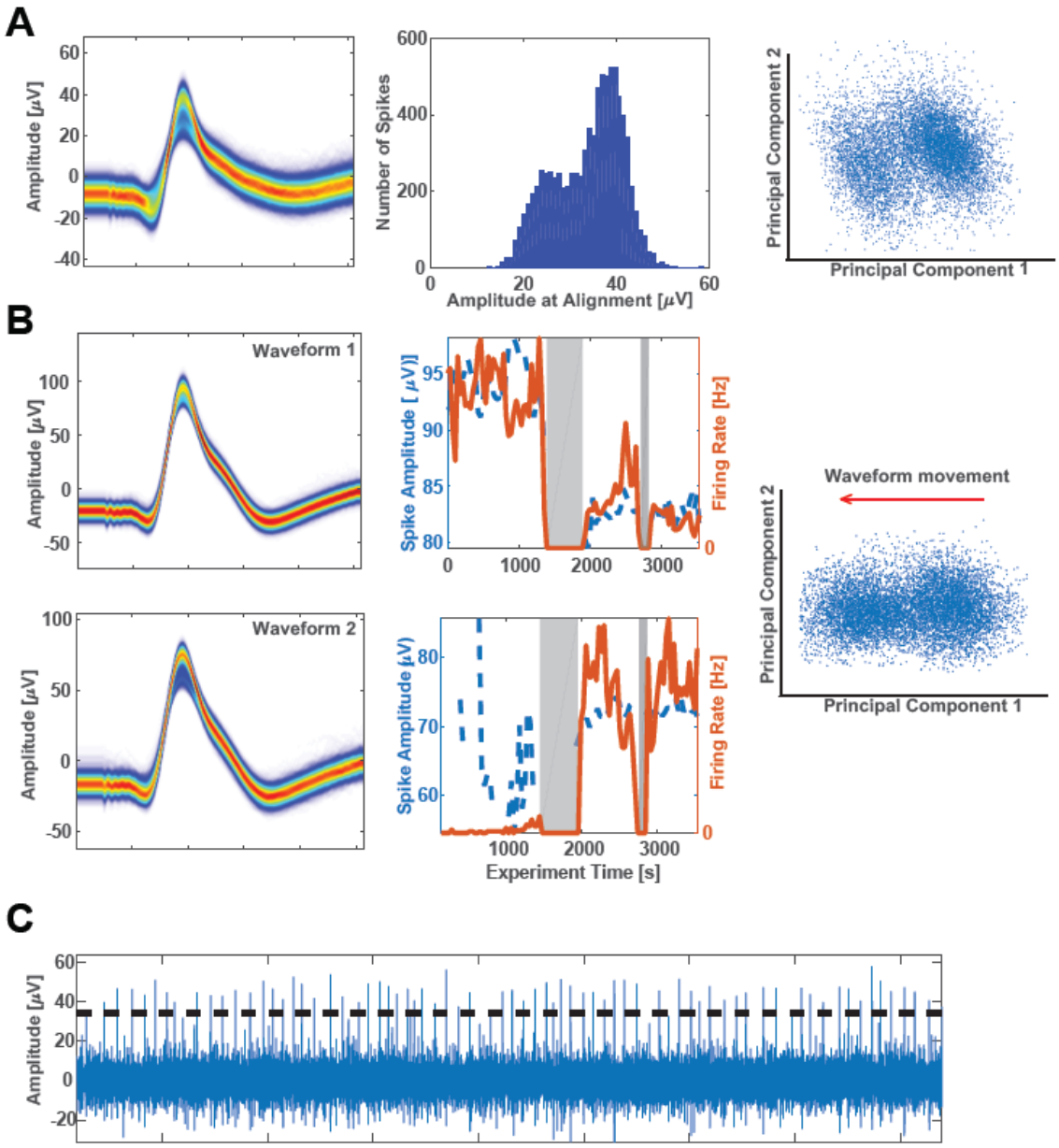


Figure 2.4: Typical problems in spike sorting. (A) One of the most common problems in spike sorting is over-merging of two clusters. This can happen if the within-group differences between spike waveforms are larger than that of across groups. In this case, the algorithm (in this case Osort) will merge the two clusters into a single cluster. There are a few simple ways to detect this kind of phenomenon. Illustrated here is the distribution of amplitudes at the alignment point (middle panel). If bimodal distribution, this is an indication

that this cluster is a composite of two (or more) other clusters. A projection of the individual spikes into PCA space confirms this (notice the two clusters in the third panel). (B) An example of under-merging of two clusters due to non-stationarities through the experiment (usually due to electrode movement). The firing rate and amplitude (second column) of the waveforms as a function of experiment time is a useful tool in diagnosing such problems. Note that during grey periods, recording was off. (C) High firing rates can bias the automatic threshold selection (dashed line), leading to missing spikes with lower amplitudes. For channels with high firing rates of high-amplitude spikes, the threshold has to be lowered manually.

2.5 Notes

Note 1: The CRW frame rods remain exposed during the procedure and are not sterile. Thus, the surgeon must take care not to accidentally touch them during the procedure. If touched, or even if there is a concern of possible touching, the surgeon gloves are changed. We typically change gloves at least 3-4 times during a procedure.

Note 2: During removal, if an electrode does not easily slide out from the anchor bolt and "hangs up," the macro contact may be caught on the dura. Do not pull hard on the electrode, as this may shear the electrode and leave a residual in the brain. Rather, cut the electrode as close to the anchor bolt as possible. Then, when the anchor bolt itself is removed, the electrode will come out with it in the largest majority of cases.

Note 3: Although we have experimented with various lengths ranging from 3-8 mm, 4-5 mm seems to given optimal recordings. Care is taken to cut quickly, and to not crush the distal ends of the microwires. A sharp scissor helps with this. The microwire bundle is examined and can be slightly fashioned with the scissor blades to make sure the wires protrude in a "flower spray" configuration.

Note 4: The choice of grounds and reference electrodes is critical for stable single-neuron recordings with high signal-to-noise ratios and an absence of movement artifacts. Strips implanted below the scalp (see Surgical Methods), with the exposed contacts pointing away from the brain, provide the best ground and reference contacts. For single-neuron recordings alone, the best reference is local, i.e. one of the eight micro-wires serves as a reference for the other 7 micro-wires. Such bi-polar recordings have the highest signal-to-noise ratio because that way, all wires have approximately the same impedance, their tips are located within a few mm of each other and common low-frequency activity cancels out. However, this configuration cancels out much of the local field potential (LFP). Thus, if LFP is important, it is advisable to use either the local reference wire or a remote reference.

Note 5: Common sources of recording noise, including line noise, are devices connected to the patient. Before starting a recording, unplug all devices that are directly connected to the patient or which are being touched by the patient. All such devices should run on battery and be disconnected from the wall. This, in particular, includes IV pumps, leg warmers, remote controllers, cell phones, computers, etc.

References

1. Mullin, J.P., M. Shriver, S. Alomar, I. Najm, J. Bulacio, P. Chauvel, and J. Gonzalez-Martinez, *Is SEEG safe? A systematic review and meta-analysis of stereo-electroencephalography-related complications.* Epilepsia, 2016. **57**(3): p. 386-401.
2. Fried, I., U. Rutishauser, M. Cerf, and G. Kreiman, *Single Neuron Studies of the Human Brain: Probing Cognition*. 2014, Boston: MIT Press.
3. Fried, I., C.L. Wilson, N.T. Maidment, J. Engel, E. Behnke, T.A. Fields, K.A. MacDonald, J.W. Morrow, and L. Ackerson, *Cerebral microdialysis combined with single-neuron and electroencephalographic recording in neurosurgical patients - Technical note.* Journal of Neurosurgery, 1999. **91**(4): p. 697-705.
4. Schmidt, R.F., C. Wu, M.J. Lang, P. Soni, K.A. Williams, Jr., D.W. Boorman, J.J. Evans, M.R. Sperling, and A.D. Sharan, *Complications of subdural and depth electrodes in 269 patients undergoing 317 procedures for invasive monitoring in epilepsy.* Epilepsia, 2016. **57**(10): p. 1697-1708.
5. Hefft, S., A. Brandt, S. Zwick, D. von Elverfeldt, I. Mader, J. Cordeiro, M. Trippel, J. Blumberg, and A. Schulze-Bonhage, *Safety of hybrid electrodes for single-neuron recordings in humans.* Neurosurgery, 2013. **73**(1): p. 78-85; discussion 85.
6. Mehta, A.D., D. Labar, A. Dean, C. Harden, S. Hosain, J. Pak, D. Marks, and T.H. Schwartz, *Frameless stereotactic placement of depth electrodes in epilepsy surgery.* J Neurosurg, 2005. **102**(6): p. 1040-5.
7. Misra, A., J.F. Burke, A.G. Ramayya, J. Jacobs, M.R. Sperling, K.A. Moxon, M.J. Kahana, J.J. Evans, and A.D. Sharan, *Methods for implantation of micro-wire bundles and optimization of single/multi-unit recordings from human mesial temporal lobe.* J Neural Eng, 2014. **11**(2): p. 026013.

8. Mamelak, A., *Ethical and Practical Considerations for Human Microelectrode Recording Studies*, in *Single Neuron Studies of the Human Brain*, I. Fried, et al., Editors. 2014, MIT Press: Boston
9. Pelli, D.G., *The VideoToolbox software for visual psychophysics: Transforming numbers into movies.* Spatial Vision, 1997. **10**: p. 437-442.
10. Brainard, D.H., *The Psychophysics Toolbox.* Spatial Vision, 1997. **10**: p. 433-436.
11. Cornelissen, F.W., E.M. Peters, and J. Palmer, *The Eyelink Toolbox: eye tracking with MATLAB and the Psychophysics Toolbox.* Behav Res Methods Instrum Comput, 2002. **34**(4): p. 613-7.
12. Rutishauser, U., A. Kotowicz, and G. Laurent, *A method for closed-loop presentation of sensory stimuli conditional on the internal brain-state of awake animals.* J Neurosci Methods, 2013.
13. Gibson, S., J.W. Judy, and D. Markovic, *Spike Sorting The first step in decoding the brain.* Ieee Signal Processing Magazine, 2012. **29**(1): p. 124-143.
14. Lewicki, M.S., *A review of methods for spike sorting: the detection and classification of neural action potentials.* Network, 1998. **9**(4): p. R53-78.
15. Rutishauser, U., M. Cerf, and G. Kreiman, *Data Analysis Techniques for Human Microwire Recordings: Spike detection and Sorting, Decoding, Relation between Neurons and Local Field Potential*, in *Single Neuron Studies of the Human Brain*, I. Fried, et al., Editors. 2014, MIT Press: Boston. p. 59-98.
16. Quiroga, R.Q., *What is the real shape of extracellular spikes?* J Neurosci Methods, 2009. **177**(1): p. 194-198.
17. Bankman, I.N., K.O. Johnson, and W. Schneider, *Optimal detection, classification, and superposition resolution in neural waveform recordings.* IEEE Trans Biomed Eng, 1993. **40**(8): p. 836-41.
18. Rutishauser, U., E.M. Schuman, and A.N. Mamelak, *Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, in vivo.* Journal of neuroscience methods, 2006. **154**(1): p. 204-224.

19. Quiroga, R.Q., Z. Nadasdy, and Y. Ben-Shaul, *Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering.* Neural Computation, 2004. **16**(8): p. 1661-1687.
20. Rossant, C., et al., *Spike sorting for large, dense electrode arrays.* Nat Neurosci, 2016. **19**(4): p. 634-41.
21. Rutishauser, U., I.B. Ross, A.N. Mamelak, and E.M. Schuman, *Human memory strength is predicted by theta-frequency phase-locking of single neurons.* Nature, 2010. **464**(7290): p. 903-7.
22. Kaminski, J., S. Sullivan, J.M. Chung, I.B. Ross, A.N. Mamelak, and U. Rutishauser, *Persistently active neurons in human medial frontal and medial temporal lobe support working memory.* Nat Neurosci, 2017.
23. Rutishauser, U., E.M. Schuman, and A.N. Mamelak, *Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, in vivo.* J Neurosci Methods, 2006. **154**(1-2): p. 204-24.
24. Pouzat, C., O. Mazor, and G. Laurent, *Using noise signature to optimize spike-sorting and to assess neuronal classification quality.* Journal of Neuroscience Methods, 2002. **122**(1): p. 43-57.
25. Rutishauser, U., S. Ye, M. Koroma, O. Tudusciuc, I.B. Ross, J.M. Chung, and A.N. Mamelak, *Representation of retrieval confidence by single neurons in the human medial temporal lobe.* Nature Neuroscience, 2015. **18**(7): p. 1041-50.
26. Hill, D.N., S.B. Mehta, and D. Kleinfeld, *Quality metrics to accompany spike sorting of extracellular signals.* Journal of Neuroscience, 2011. **31**(24): p. 8699-8705.
27. Schmitzer-Torbert, N., J. Jackson, D. Henze, K. Harris, and A. Redish, *Quantitative measures of cluster quality for use in extracellular recordings.* Neuroscience, 2005. **131**(1): p. 1-11.
29. Adolphs, R., *What does the amygdala contribute to social cognition?* Annals of the New York Academy of Sciences, 2010. **1191**(1): p. 42-61.
30. Paton, J.J., M.A. Belova, S.E. Morrison, and C.D. Salzman, *The primate amygdala represents the positive and negative value of visual stimuli during learning.* Nature, 2006. **439**(7078): p. 865-870.

31. Tsao, D.Y., W.A. Freiwald, R.B. Tootell, and M.S. Livingstone, *A cortical region consisting entirely of face-selective cells.* Science, 2006. **311**(5761): p. 670-674.
32. Rutishauser, U., O. Tudusciuc, S. Wang, Adam N. Mamelak, Ian B. Ross, and R. Adolphs, *Single-Neuron Correlates of Atypical Face Processing in Autism.* Neuron, 2013. **80**(4): p. 887-899.
33. Rutishauser, U., O. Tudusciuc, D. Neumann, A.N. Mamelak, A.C. Heller, I.B. Ross, L. Philpott, W.W. Sutherling, and R. Adolphs, *Single-unit responses selective for whole faces in the human amygdala.* Current Biology, 2011. **21**(19): p. 1654-60.
34. Rutishauser, U., A.N. Mamelak, and R. Adolphs, *The primate amygdala in social perception - insights from electrophysiological recordings and stimulation.* Trends Neurosci, 2015. **38**(5): p. 295-306.
35. Adolphs, R., F. Gosselin, T.W. Buchanan, D. Tranel, P. Schyns, and A.R. Damasio, *A mechanism for impaired fear recognition after amygdala damage.* Nature, 2005. **433**(7021): p. 68-72.
36. Dal Monte, O., V.D. Costa, P.L. Noble, E.A. Murray, and B.B. Averbeck, *Amygdala lesions in rhesus macaques decrease attention to threat.* Nature communications, 2014. **6**: p. 10161-10161.
37. Adolphs, R., D. Tranel, and A.R. Damasio, *The human amygdala in social judgment.* Nature, 1998. **393**(6684): p. 470-474.
38. Mormann, F., S. Kornblith, R.Q. Quiroga, A. Kraskov, M. Cerf, I. Fried, and C. Koch, *Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe.* Journal of Neuroscience, 2008. **28**(36): p. 8865-8872.
39. Harris, K.D., D.A. Henze, J. Csicsvari, H. Hirase, and G. Buzsaki, *Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements.* J Neurophysiol, 2000. **84**(1): p. 401-14.

Chapter III: Fixations Gate Species-specific Responses to Free Viewing of Faces in the Human and Macaque Amygdala

3.1 Overview

Neurons in the primate amygdala respond prominently to faces. This implicates the amygdala in the processing of socially significant stimuli, but its contribution to social perception remains poorly understood. We performed a comprehensive evaluation of the representation of faces in the primate amygdala during naturalistic conditions by recording from both human and macaque amygdala neurons during free viewing of identical arrays of images with concurrent eye-tracking. Neurons responded to faces only when they were fixated, suggesting that their activity was gated by visual attention. Further experiments in humans utilizing covert attention confirm this hypothesis. The majority of face-selective neurons preferred faces of conspecifics in both species, a pattern mirroring first fixation preferences. Response latencies, relative to fixation onset, were shortest for conspecific-preferring neurons, and were ~100ms shorter in monkeys compared to humans. This argues that attention to faces gates amygdala responses, which in turn prioritizes species-typical information for further processing.

3.2 Introduction

Faces are important stimuli for primate social behavior. Humans and macaques share a homologous set of cortical regions specialized for processing faces (Tsao, Moeller et al. 2008), and in macaques, these "face patches" contain neurons almost entirely selective for faces (Tsao, Freiwald et al. 2006). Together, face patches constitute an interconnected system for constructing face representations from facial features (Moeller, Freiwald et al. 2008, Freiwald, Tsao et al. 2009). A key unanswered question is how this cortical representation of faces guides social behavior. The amygdala is a key structure in such subsequent processing: it is reciprocally connected with the cortical face patches (Grimaldi, Saleem et al. 2016), contains a large proportion of neurons selective for faces (Sanghera, Rolls et al. 1979, Gothard, Battaglia et al. 2007, Rutishauser, Tudusciuc et al. 2011, Mosher, Zimmerman et al. 2014), and is critical for primate social behavior.

The amygdala processes stimuli with ecological significance, including not only social stimuli such as faces, but conditioned and unconditioned rewards and punishments (Paton, Belova et al. 2006, Adolphs 2010). Face-selective responses are prominent in the amygdala of both humans and monkeys (Sanghera, Rolls et al. 1979, Fried, MacDonald et al. 1997, Gothard, Battaglia et al. 2007, Rutishauser, Tudusciuc et al. 2011, Mosher, Zimmerman et al. 2014), as would be expected from the highly processed visual inputs the amygdala receives from the multiple areas where face-selective cells have been discovered (Gross, Rocha-Miranda et al. 1972, Bruce, Desimone et al. 1981, Perrett, Rolls et al. 1982, Rolls 1984, Desimone 1991). This picture suggests a limited contribution of the amygdala to face processing: all its face selectivity might be explained by the inputs from face-selective cortical regions. In addition, it is commonly believed that the large receptive fields of the neurons that provide input to the amygdala would result in visual receptive fields of amygdala neurons that are not spatially restricted (Gross, Bender et al. 1969, Boussaoud, Desimone et al. 1991, Barraclough and Perrett 2011). It has been proposed that the amygdala responds to faces even when they are not attended (Vuilleumier, Armony et al. 2001) or consciously perceived (Tamietto and de Gelder 2010). This view of the amygdala's function fits with a long-standing debate about whether the amygdala mediates rapid automatic and relatively coarse detection of significant stimuli through a route of subcortical inputs (Cauchoix and Crouzet 2013). These views challenge observations that human amygdala neurons show exceedingly long visual response latencies (Mormann, Kornblith et al. 2008, Rutishauser, Mamelak et al. 2015), and that fMRI activation of the amygdala appears to require visual attention (Pessoa, McKenna et al. 2002). In the absence of comparative studies using the same stimuli and the same paradigm in both species, it is nearly impossible to determine whether these are differences between the two species or rather different experimental conditions. Resolving these disparate conclusions thus requires a more comprehensive investigation, which no single study has yet accomplished: assessing amygdala responses to faces at the single-unit level across both monkeys and humans, and investigating selectivity and response latency in relation to fixation onset during free-viewing with concurrent eye-tracking.

Although we know much about how face-selective responses may arise from the geometric and semantic features of faces (Tsao, Freiwald et al. 2006), this knowledge has been derived from studies with static stimuli of single faces displayed on a featureless background and in the absence of eye movements. As such, little is known about face responses during natural vision and their potential modulation by attention. During natural vision, the visual system has to contend with complex and dynamic visual scenes in which multiple items compete for attention. Eye movements select from the multitude of possible fixation targets that are salient or behaviorally significant elements of the scene (such as faces). Under these conditions, the response properties of cortical visual neurons can be modulated dramatically (Sheinberg and Logothetis 2001, Rolls, Aggelopoulos et al. 2003). Indeed, eye movements and the use of naturalistic stimuli change the selectivity and response reliability of neurons even in early visual areas (Gallant, Connor et al. 1998, David, Vinje et al. 2004). Similar modulation of attention-related neural activity has been found in parietal and prefrontal visual areas involved in the planning and elaboration of a sequence of fixations during natural vision (Zirnsak and Moore 2014). Likewise, neurons in the inferotemporal cortex that are selective for an item embedded in a crowded scene respond to their target stimulus only during fixations on that particular item (Sheinberg and Logothetis 2001). Thus, throughout the brain, visual processing is strongly influenced not only by the identity of objects, but also by how fixations select them. However, almost nothing is known about how fixations affect visual processing in the amygdala.

In the context of natural vision, the amygdala is of particular interest because amygdala lesions are known to interfere with the efficient visual exploration of faces (Adolphs, Gosselin et al. 2005), and because amygdala neurons respond to dynamic social signals such as eye contact (Mosher, Zimmerman et al. 2014). Moreover, lesions of the amygdala produce a complex constellation of impairments in social behavior (Adolphs, Tranel et al. 1994, Adolphs, Tranel et al. 1998). The amygdala is thus a prime candidate for mediating between the perceptual representations of faces in the cortical face-patch system and the mediation of social behaviors based on such perception. Elucidating this role, however, requires both a more naturalistic presentation of stimuli and a better quantification of how they are attended. Here we achieved both these imperatives by allowing subjects to freely

view a complex array of images that competed for attention while we monitored eye movements. Our focus was on the category selectivity of amygdala neurons during natural vision, with specific emphasis on the potential category selectivity for conspecific and heterospecific faces. We focused on faces not only because of their patent ecological relevance, but also because they are the visual category of stimuli consistently explaining the largest proportion of variance of the responses of amygdala neurons (Gothard, Battaglia et al. 2007, Rutishauser, Tudusciuc et al. 2011).

In addition to using free viewing and eye-tracking, we sought to find convergent evidence by presenting identical stimuli to both monkeys and humans in an attempt to help generalize findings across species. This allowed us to compare between responses to faces in each species, including differences in neuronal response selectivity and latency. To achieve this, we presented humans and monkeys with the same arrays of images for free viewing. The arrays contained images of monkey and human faces intermixed with images of non-face objects. We addressed three tightly related questions. First, do neural responses in the human and monkey amygdala depend on the visual category of attended stimuli as assessed by fixation location? Second, are face-responsive neurons in the amygdala biased for faces of conspecifics? Third, are the response latencies to faces different in the two species? Together, this comparative study reveals that primate amygdala responses during free viewing are profoundly influenced by attention to socially relevant stimuli during natural vision. Our findings suggest a mechanistic basis for the role of the primate amygdala in attentional selection for social stimuli, by which amygdala responses in turn regulate responses in visual cortex in a top-down manner (Pessoa and Ungerleider 2004).

3.3 Results

3.3.1 Task and Behavior

We tested a total of 12 human epilepsy patients (28 sessions, 50 ± 1 trials per session, ± s.d.) and 3 healthy macaques (16 sessions, 113 ± 13 trials per session, ± s.d.). Subjects freely viewed complex visual stimuli (Fig. 1A). Each stimulus consisted of a circular array of eight images chosen at random from two face categories (human and monkey faces) and two non-

face categories (either flowers and fractals or fruits and cars, depending on the version of the task performed). The non-face categories were later pooled for analysis as "distractors" (Distractor Group #1 contained cars and fractals and Distractor Group #2 contained fruits and flowers). Each image array was displayed for 3-4s, and subjects were free to view any location. Although stimuli were identical for the two species, each necessitated slightly different task conditions (see below for behavioral controls). Macaques received a fixed amount of reward after conclusion of a trial if they maintained their gaze position anywhere on the screen during the entire stimulus period. Trials were aborted and no reward was given if a monkey moved its gaze off the screen within the first 3s of stimulus onset. This achieved attention to the image array while encouraging free exploration. Human subjects were instructed to freely view the stimuli for a later memory test (Fig. 3.1A). To verify that our two tasks produced largely comparable fixation behaviors in the two species, we compared the scan-paths used by humans and monkeys to explore the image arrays (Fig. 3.1B shows an example).

To further ensure that our memory task does not introduce fixation preference biases, we also asked two groups of healthy human control subjects to perform the identical task (Fig. 3.2): one with the same instructions as the patients ("memory controls") and one that did not know that a memory test would follow ("free viewing control"). Patients had good recognition memory (average across n=14 sessions, 70%, p=0.001 vs. chance, one-sided binomial test), showing that they attended to the stimuli. The patients' performance was somewhat lower than that of the memory group but not the free viewing group (Fig. 3.2D, $t(19)=2.845$, $p=0.01$ and $t(19)=1.467$, $p=0.16$ respectively). Crucially, the probability that the first fixation landed on a human face was not influenced by task instruction and did not differ between any of the groups ($t(19)=0.8538$, $p=0.40$ and $t(19)=0.3013$, $p=0.77$, respectively). Subjects in all three groups explored all stimuli extensively regardless of task instructions (Fig. 3.2A,B), confirming that informing subjects of a subsequent memory test without further specific information does not affect fixation preferences for faces.

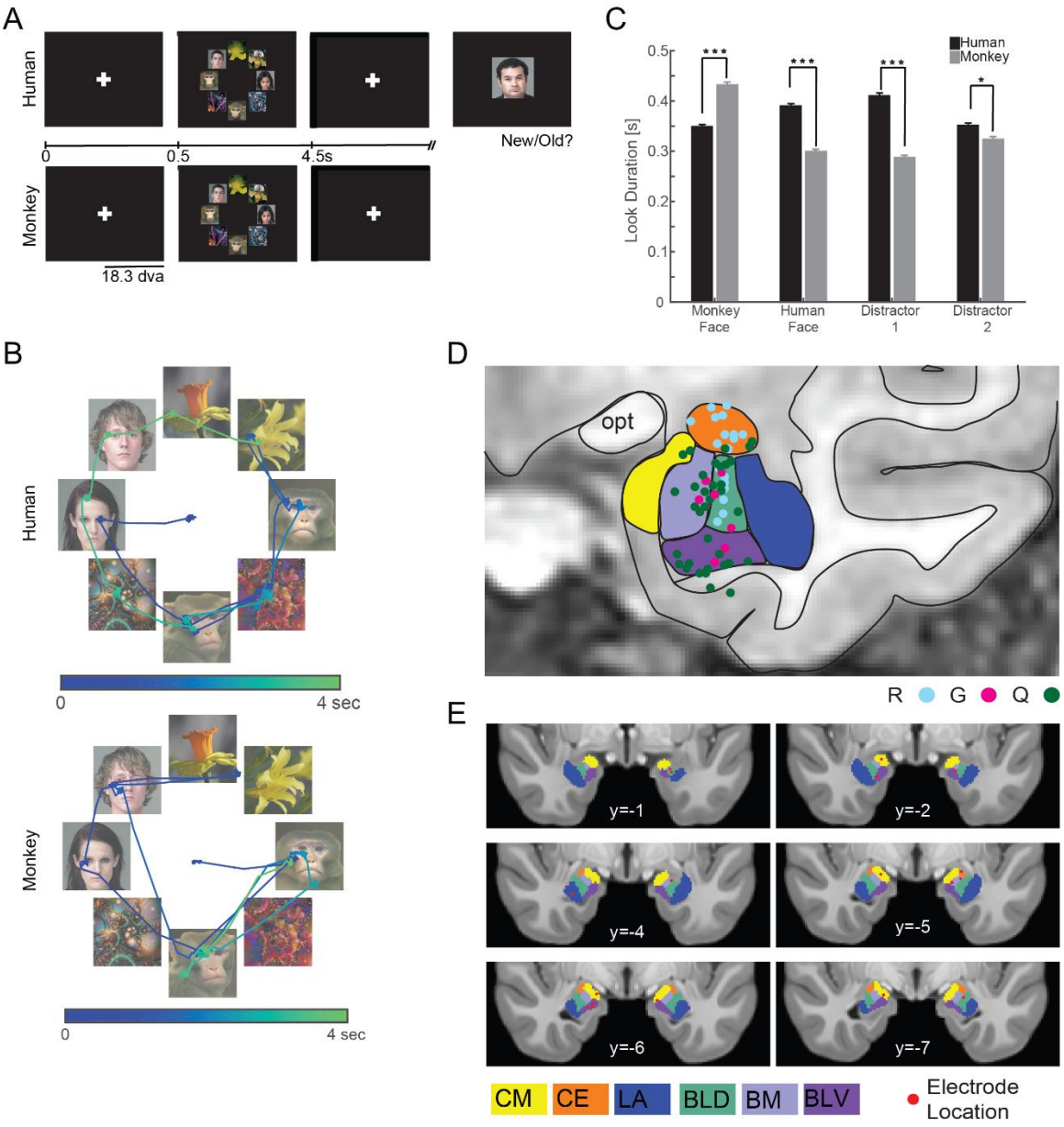


Figure 3.1: **Task, behavior, and recording locations.** (a) Task performed by human and monkey subjects. Task version #1 is shown. (B) Example scan paths from a human (top) and monkey (bottom) viewing the same stimuli. Dots are fixations, and lines are saccades. The first saccade (starting at the center) targets the face of a conspecific. Trial time is encoded by color. (C) Look duration ("dwell time") on each stimulus category for humans (black) and monkeys (gray). Monkeys fixated longer on conspecific faces and on flowers, while humans fixated longer on conspecific faces and on fractals (three stars indicates $p<0.001$ and one star indicates $p<0.05$, two-tailed t-test). (D-E) Recording locations. Amygdala nuclei are indicated in color. (D) Recording sites in the three monkey subjects (R,G,Q in different color dots) collapsed onto a representative coronal section. (E) Human recording sites (red dots) in MNI152 space. Abbreviations: LA = Lateral Nucleus, BLD = Dorsal Basolateral, BM = Basomedial, CE = Central, CM = Cortical and Medial Nuclei, BLV = Ventral Basolateral.

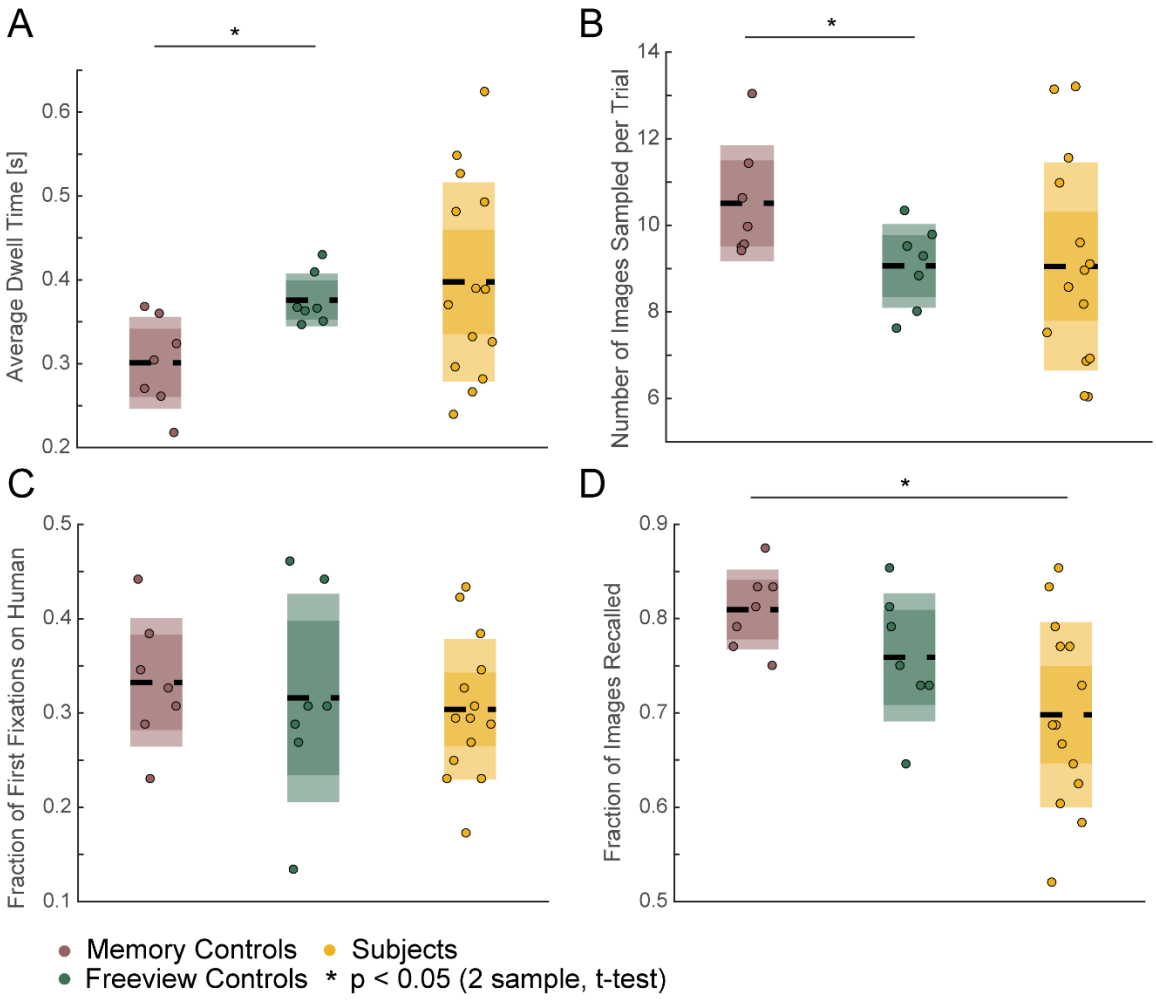


Figure 3.2: Comparison of behavior of human neurosurgical subjects with normal control subjects, related to Figure 3.1. (A) The average look time for each image for three groups of subjects: memory controls (brown), free-view controls (green), and neurosurgical subjects (yellow). Each dot represents the average across one session (n=7, n=7, n=14 for the memory, free-viewing, and subject groups respectively). For the subjects, we only used the sessions that used the same stimuli as the control group. (B) Average number of images visited during each trial. (C) Proportion of all first fixation that landed on a human face was comparable across all subject groups. (D) Memory retrieval performance of the three subject groups. While on average control subjects that knew of the later memory test performed better than the free-view control subjects, it was not statistically significant ($t(12)=1.674$, $p=0.12$). Neurosurgical subjects performed significantly above chance ($p=0.002$, binomial test) but worse than the memory control group ($t(19)=2.845$, $p=0.01$).

The majority of all fixations landed on one of the eight images in both human patients and macaques (84% and 89%, respectively; Fig. S1F) and subjects looked longer and earlier at faces of conspecifics compared to faces of heterospecifics (Fig. 1C, see legend for statistics). A reliable viewing pattern for both species was that the probability of looking first at a conspecific face was higher than that of first looking at a heterospecific face (32% vs. 24%, $t(54) = 3.63$, $p=0.0006$, in humans and 32% vs. 20%, $t(30) = 2.77$, $p=0.01$, in monkeys, paired t-tests, chance is 25%, Fig. S1A-B). Also, in both species some sequences of fixations (i.e. human followed by monkey face) were more likely than others (Fig. S1C-D), which shows that fixation location was influenced by image content throughout. This suggests that even before launching a sequence of exploratory eye movements, conspecific faces competed successfully for fixation priority in both humans and monkeys. Together, this argues that the location of the faces on the screen was attended to and influenced the fixation patterns of both humans and monkeys, a mechanistic hypothesis that we tested further with a covert attention task in the human subjects that is described below.

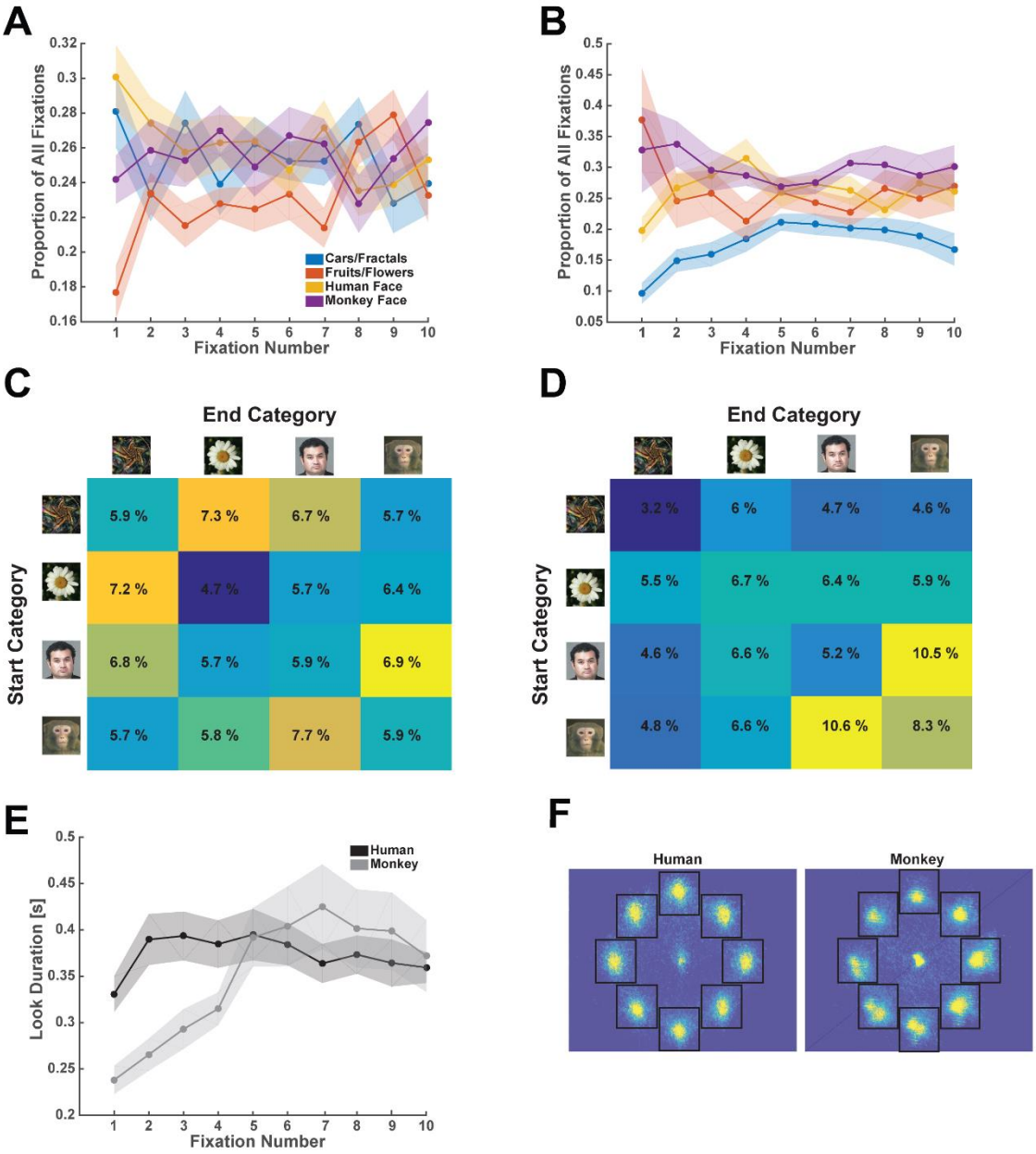


Figure 3.3: Comparison of human and monkey behavior, related to Figure 3.1. (A) Proportion of fixations that landed on a given image category as a function of fixation number and category for human subjects. (B) The same plot as in (A) but for monkeys. (A,B) Note how, for the first fixation, both species were more likely to look at faces of their own species compared to the other species. (C,D) Frequency with which each of the 16 possible category transitions occurred for human (C) and monkey (D) subjects. While all possible transitions occurred, some were more likely (i.e. human face – monkey face, and monkey face – human face). (E) Look duration as a function of fixation number. Both humans and macaques modulated look duration as a function of time in the trial (p=0.01 and P=10-7, respectively, 1x10 ANOVA). However, this difference was entirely accounted for by the difference between the first and second fixation in humans, but not macaques (p=0.09 and P=10-5, respectively, 1x9 ANOVA). (F) Heat maps showing the eye-tracking data for all sessions in humans (left) and monkeys (right). Both humans and monkeys sampled all 8 images.

3.3.2 Electrophysiology

We isolated 422 and 195 putative single units from the human and macaque amygdala, respectively (on average 19 and 12 per session, respectively). We only analyzed cells with a mean firing rate of >0.5 Hz (290 and 148 units, respectively; mean firing rates 2.28 Hz and 5.6 Hz, range 0.5-26.4 Hz and 0.5-72.8 Hz). Throughout the manuscript, we use the terms neuron or cell to refer to a putative single unit, and we only used units satisfying multiple conservative criteria (see Methods for details). To facilitate direct comparison between species, all spike detection, sorting, and data analysis was performed using the same methods for all recordings from both species (see Fig. 3.4 for electrophysiological properties of cells in both species).

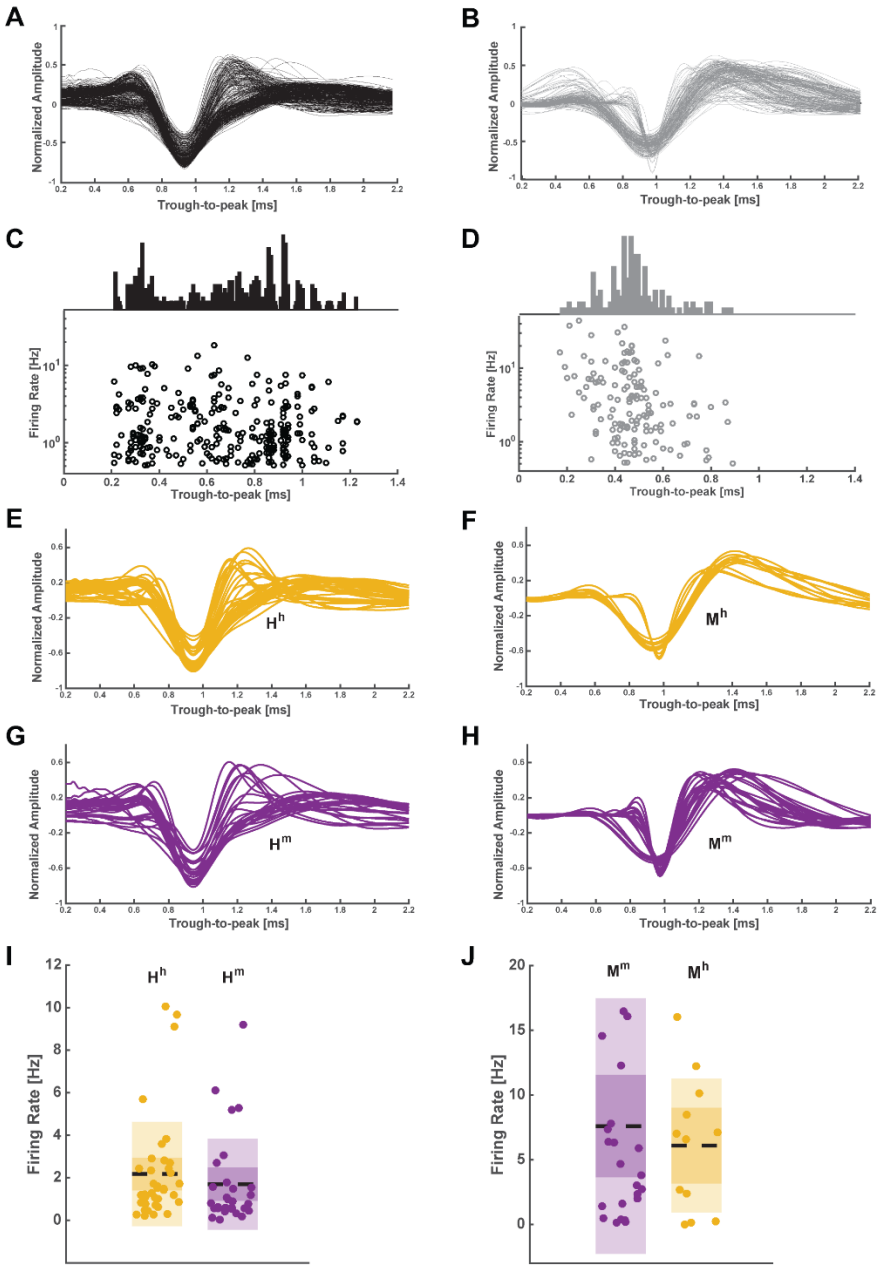


Figure 3.4: Electrophysiological properties of neurons in both species, related to Figure 2. (A) Mean waveform shapes for all tuned cells in the human amygdala (n=85). The waveforms were all peak-normalized, and aligned by their peak (in the case of a positive peak, we flipped the waveform in order facilitate alignment). (B) Mean waveform shape for all tuned cells in the monkey amygdala (n=61). (C) Trough-to-peak distances (in ms) as a function of firing rate for neurons recorded in humans (n=85). Trough-to-peak distances were bimodally distributed ($p<0.0001$, Hartigan dip test). (D) Same as (C), but for all cells recorded from the monkey amygdala. There was no evidence for a bimodal distribution of trough-to-peak distances ($p=0.3$, Hartigan dip test). (E-H) Mean waveforms for al H^h, M^h, H^m, and M^m cells, respectively. (I) Mean firing rates for all H^h (yellow, 2.18 ± 0.38 Hz) and H^m cells (purple, 1.70 ± 0.40 Hz). There was no significant difference between these two populations ($p=0.40$, 2-sample t-test). (J) Mean firing rates for all M^h (yellow, 6.08 ± 1.50 Hz) and M^m cells (purple, 7.60 ± 2.01 Hz). There was no significant difference between these two populations ($p=0.62$, 2-sample t-test).

3.3.3 Fixation-target sensitive neuronal responses

We first determined whether the responses of amygdala neurons were modulated by the identity of the fixated stimuli. For each neuron, we tested whether the firing rate following fixation onset co-varied with the identity of the fixated images (see Methods). When features within an image were successively fixated, time of fixation onset was determined by the first fixation that fell within that image's region ("look onset", see Methods and Fig. S2). We found that 20% (n=85/422) of human and 31% (n=61/195) of macaque neurons significantly modulated their firing rate after fixation onset (Fig. 2). These "fixation-target sensitive" responses appeared transiently, shortly after fixation onset (see below for a latency analysis).

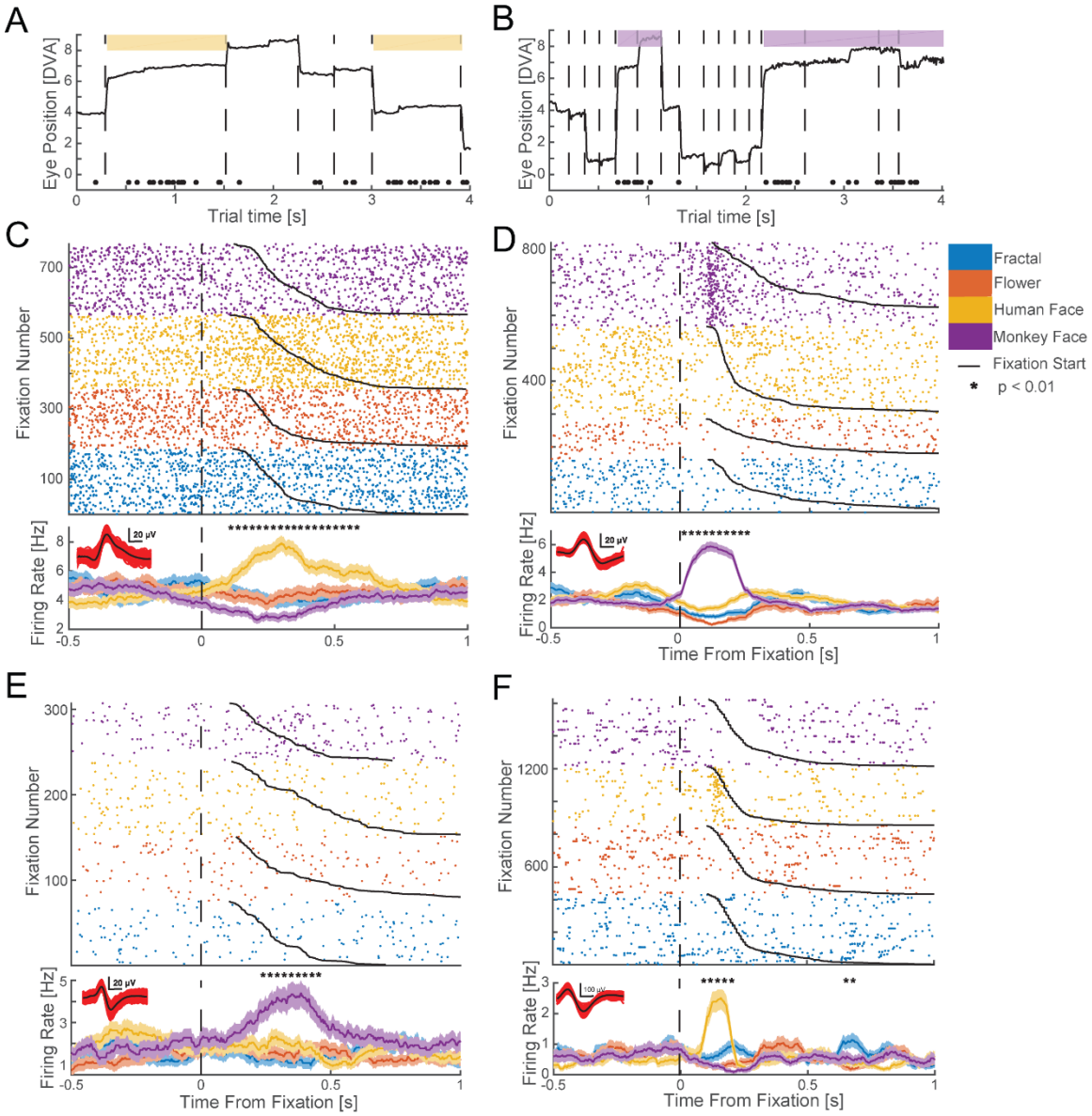


Figure 3.5: Example single neurons with fixation-related activity. (A,B) Example trial from a human (A) and monkey (B) face-selective neuron. Spikes are indicated by black dots. Whenever gaze fell onto a conspecific face (colored patch), the neuron increased its activity. (C-F) Rasters (top) and mean firing rate (PSTH, bottom) for neurons recorded in humans (C,E) and monkeys (D,F). Neurons are selective for conspecific (C,D) and heterospecific (E,F) faces. (C,D) show the activity of the neurons depicted in (A,B). t=0 marks fixation onset. Trials were sorted by category of the fixated image (color code) and fixation duration (black line). Stars above the PSTHs indicate bins of neural activity (of 250ms duration) with a significant (1x4 ANOVA, $p<0.01$) difference in firing rate. Horizontal scalebar for waveforms is 0.2mV. The four neurons are from different subjects.

We further characterized the category selectivity of fixation-target sensitive responses (Fig. 2C,D). We first classified each fixation-target sensitive cell according to the category to which it responds most strongly (highest firing rate) at the point of time at which neurons provided most category information (see Methods and Fig. 3.6).

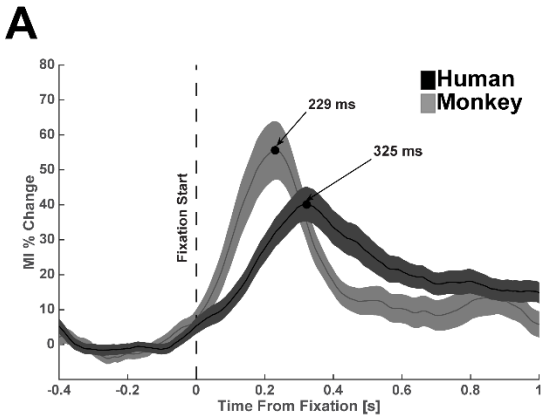


Figure 3.6: Using mutual information to determine the position of the analysis window for selectivity analysis, related to Figure 3. (A) Time-course of information, quantified as mutual information (MI; peak normalized) between the firing rate and visual category for all neurons recorded in monkeys (N=195, light gray) and humans (N=422, black). The point of time at which MI was maximal (t=325ms and 229ms, respectively) was used to place the analysis window for all further analysis.

The majority of fixation-target sensitive neurons preferred faces of conspecifics: 49% (n=42/85) and 39% (n=24/61) in humans and macaques, respectively (Fig. 3A). A smaller proportion preferred faces of heterospecifics (that is, faces of the opposite species): 32% (n=27/85) and 18% (n=11/61) in humans and macaques, respectively. Together, about 71% of all fixation-sensitive neurons preferred faces (Fig. 3A, 81% in humans and 58% in macaques). In contrast, only 19% (n=16/85, in humans) and 42% (n=26/61, in macaques) were sensitive to the non-face categories we used (flowers, fractals, fruits, and cars). Since subjects were free to look at any of the images, we had no way of ensuring that they would sample uniformly from the different image categories. In order to ensure that the tuning of the cells was not confounded by the number of fixations on each category, we carried out a control analysis in which we selected cells after equalizing the number of fixations for each

image category by subsampling. This revealed similar proportions: 50±3% and 40±2.2% of neurons preferred faces of conspecifics, respectively (±s.d. across 100 bootstraps). Thus, most primate amygdala fixation-target sensitive neurons responded to faces rather than nonsocial object categories, and there were two groups of such neurons: those that increase their firing rate whenever fixations are made onto faces of conspecifics, and those that increase their firing rate only when looking at faces of heterospecifics (Fig. 3B-E shows the average response of all four types of face cells). For clarity, we label each type of face-sensitive cell with a capital letter to signify the species in which the cell was recorded (H or M) and with a lowercase letter to signify the tuning of the cell to human or monkey faces (h or m).

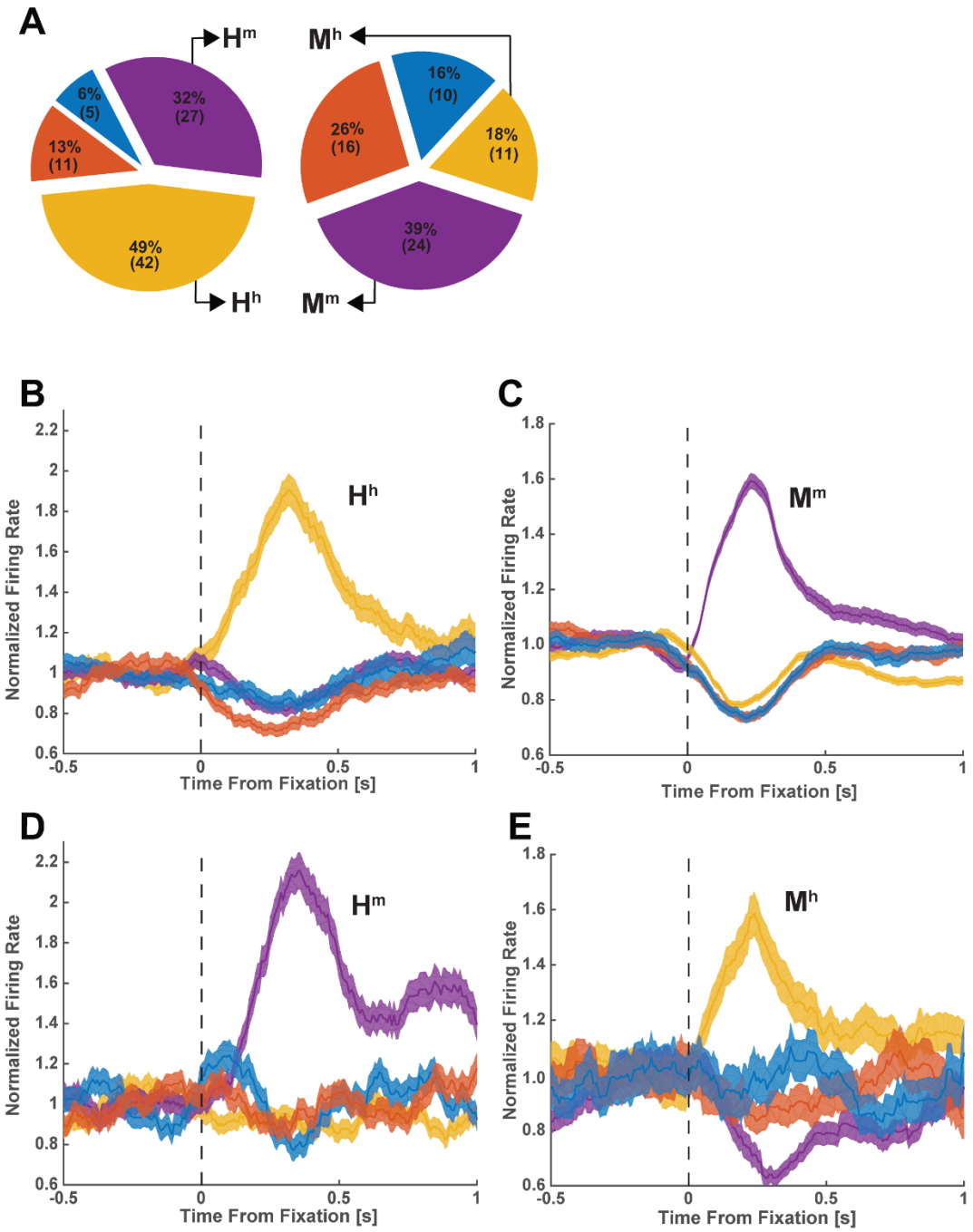


Figure 3.7: Population analysis and cross-species comparison of fixation-related visual category-selectivity. (A) Preferred stimulus of all recorded visually selective cells in human (left) and monkey (right) amygdala. The largest proportion of neurons responded maximally to faces of conspecifics: 49% (N=42) and 39% (N=24) of selective neurons in humans and monkeys, respectively. (B-E) Average normalized PSTHs of the four groups of face cells we identified (Hh, Mm, Hm, Mh). The middle row (B,C) shows neurons selective for conspecifics in humans (left, Hh, N=42) and monkeys (right, Mm, N=24). The bottom row (D,E) shows cells selective for heterospecifics in humans (left, Hm, N=27) and monkeys (right, Mm, N=11). Errorbars are ±s.e.m. across cells. Notation: Hh = human cell selective for human faces; Hm = human cell selective for monkey faces; Mm = monkey cell selective for monkey faces; Mh = monkey cell selective for human faces.

We next determined whether fixation-target sensitive neurons differentiated between multiple categories, i.e., whether they also responded to images from a non-preferred category with firing rates that were different from the baseline (Fig. 2C-D, Fig. 3B-E). Indeed, some neurons showed a pattern of response that appears optimized to differentiate between all of our categories. To quantify this effect, we calculated two metrics for each cell: (1) the number of pairs of image categories discriminated by each neuron (i.e. human faces vs. monkey faces) and (2) the depth of selectivity (DOS) index (Rainer, Asaad et al. 1998) commonly used to determine the extent to which visual neurons differentiate between stimuli.

We found that neurons in the human amygdala differentiated between, on average, 3.47±0.1

pairs of categories (out of 6), while neurons in the monkey amygdala differentiated between 4.15±0.2 pairs (Fig. 4A-B). Thus, neurons in humans differentiated between significantly fewer ($p = 0.002$, 2-sample KS test) pairs of categories compared to neurons in macaques. Similarly, the depth of selectivity (DOS), an index of the narrowness of tuning to a specific category, of all human neurons was larger than that of macaque neurons (0.54±0.02 vs. 0.43±0.03, $p=0.0003$, 2-sample KS test, Fig. 4C), but was at the same time significantly lower than 1 ($p<1e-37$, 2-sample KS test). Note that a DOS value of 1 means exclusive tuning to one stimulus, but no response to all other stimuli; in contrast, a DOS value of 0 implies no preferred tuning. We observed DOS values of 0.18-0.87 in humans and 0.11-0.90 in macaques (Fig. 4C-D). DOS values for neurons recorded in humans were significantly larger than those for neurons recorded in monkeys (Fig. 4E), a result compatible with the sparser response profile over categories as shown in Fig. 4B. While the DOS differed between species (see above), it did not differ significantly between cells tuned to conspecific and heterospecific faces (1x2 ANOVA, $F(1,67)=1.75$, $p=0.19$ and $F(1,33)=0.52$, $p=0.47$ in humans and monkeys, respectively). We also estimated DOS values using fixations (50/50 split) not previously used to select neurons and found that DOS values are highly reliable and significantly larger than those of unselected cells (Fig.

4C,D, see legends for statistics). Taken together, these observations show three important similarities between neurons in the human and monkey amygdalae: both contain fixation-target sensitive neurons; these neurons show category-specific responses; and the largest subset of such neurons responds preferentially to conspecific faces. A difference between the species was that human neurons have a sparser response profile over categories.

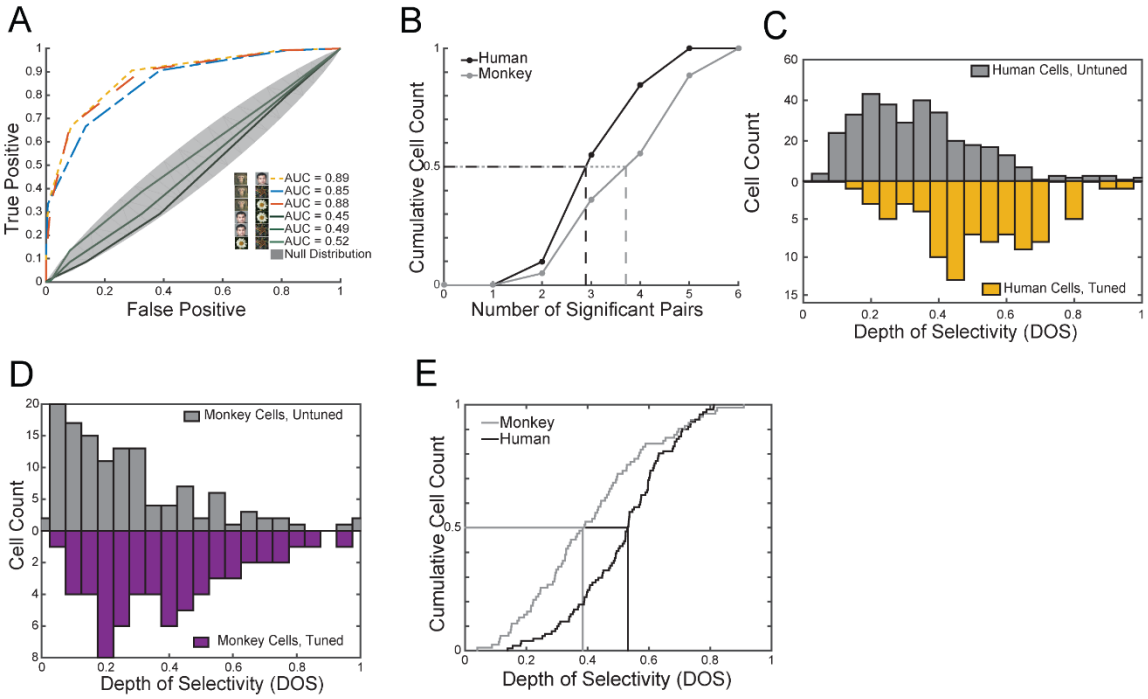


Figure 3.8: **Monkey and human amygdala cells differ in their depth-of-selectivity.** (A) Single-cell ROC analysis example. The monkey cell shown (identical to that in Fig. 2D) responded only to images of conspecifics, allowing it to discriminate 3 pairs of categories (dashed colored lines). (B) Distribution of the number of significant contrasts (see A) for all visually tuned neurons in humans (black) and monkeys (gray). Cells recorded in monkeys differentiated significantly more contrasts (4.15±0.2) than human cells (3.47±0.1, $p<0.002$, 2-sample KS test). (C,D) Population summary. Comparison of depth-of-selectivity (DOS) values for tuned and untuned cells in human (C) and monkey (D). In both species, the DOS values are significantly greater in the tuned population ($p<1e-16$, in humans and $p<5e-5$ in monkeys, 2-sample KS test). For tuned cells, DOS values were calculated for a subset of fixations that were not used in the selection of that cell (i.e. to determine its tuning). (E) Depth-of-selectivity (DOS) for all visually tuned neurons in humans (black) and monkeys (gray). Human cells had significantly larger DOS values (0.51±0.02 vs. 0.43±0.03, $p<0.0003$, 2-sample KS-test).

3.3.4 Interspecies comparison of response latencies of face-selective neurons

We next compared the latency of the fixation-target sensitive neurons between species. We estimated the response latency for each cell to test whether the time at which the modulation of firing rate was first detectable systematically co-varied as a function of species and

stimulus type. We first quantified latency differences using a single-neuron moving-window regression model to estimate the amount of variance in the firing rate that can be explained at any point of time by the visual category of the fixated stimulus (Fig. 5A-D, 1x4 ANOVA, moving window of 250ms, step size = 8ms). We then estimated the effect size ω^2 as a function of time separately for each neuron to determine the point of time, relative to fixation onset, at which ω^2 first became significant (Fig. 5A-D). We found that the onset of the fixation-sensitive neurons in macaques was, on average, 76 ms earlier than in humans (101±7.5 ms versus 177±8.7 ms, p=5e-8, 2-sample KS test, Fig. 5D). Also, the proportion of neurons that became visually selective increased earlier in macaques compared to humans (Fig. 5E), and the point of time at which neurons provided the most information (peak of ω^2) was 113 ms earlier in macaques compared to humans (209±8.9 ms vs. 322±7.5 ms, p<1e-14, 2-sample KS test, Fig. 5F). Together, this shows that, regardless of stimulus selectivity, cells in the human amygdala respond approximately 100ms later relative to fixation onset compared to cells recorded in the macaque amygdala.

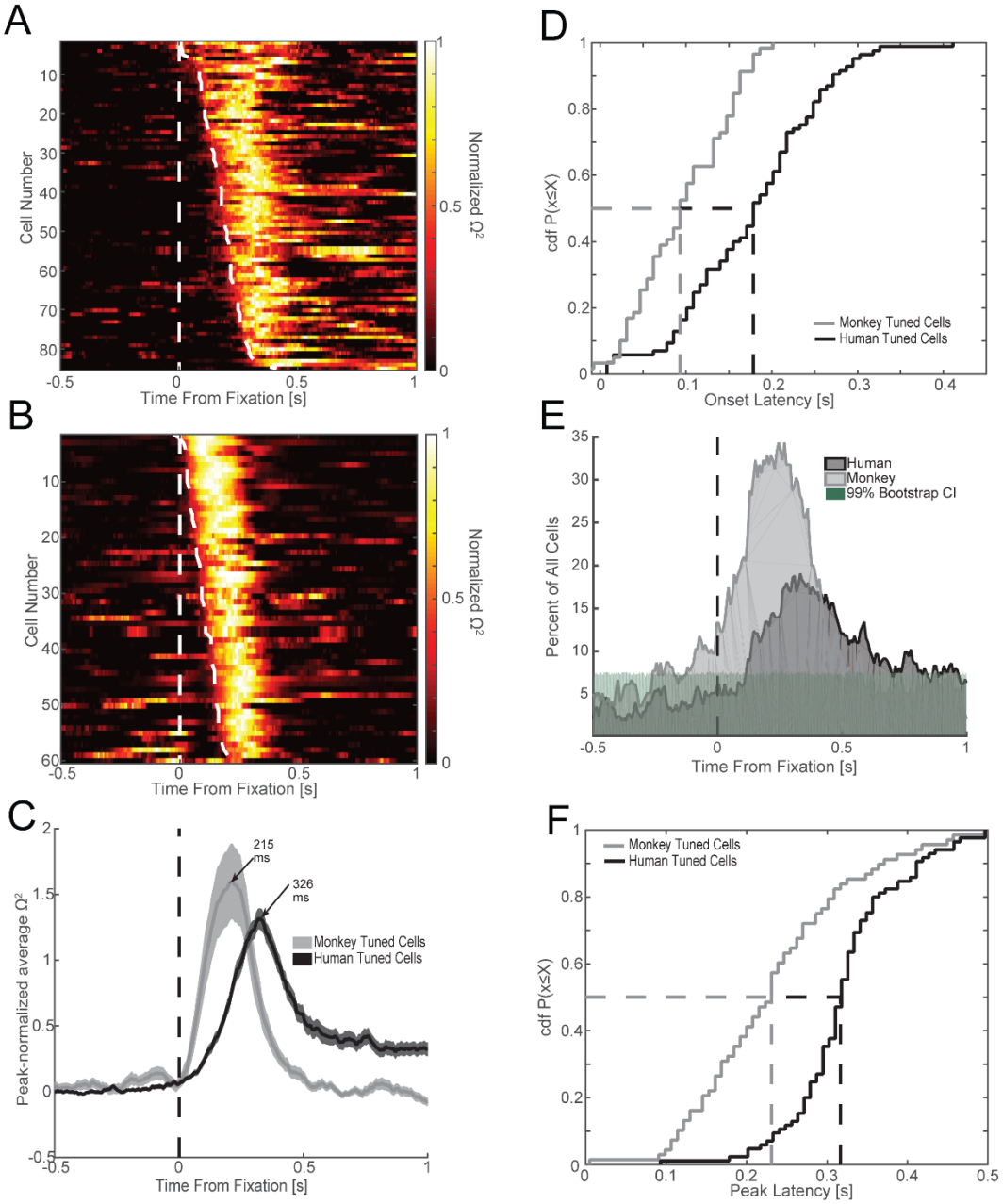


Figure 3.9: Interspecies comparison of response latency relative to fixation onset. (A,B) Effect size (Ω^2) for all visually selective human (A) and monkey cells (B) as a function of time and sorted by earliest point of significance (only cells that are significant at the $p<0.01$ level are shown). Each cell's effect size is normalized to its peak. (C) Mean normalized effect size for all visually selective cells recorded in humans (N=85) and monkeys (N=61). (D) Cumulative distribution of the onset latency computed using the effect size (see A-B). The mean onset latency was significantly earlier in monkeys (101±7.5 ms) compared to humans (177±8.7 ms, $p<5e-8$, two-sample KS test). (E) Proportion of all recorded cells that were sensitive to the identity of the fixated stimulus as a function of time (bin size 250ms, step size 8ms). Shading shows the 99th percentile of the bootstrap distribution. (F) Cumulative distribution of the time from fixation onset until peak effect size. Peak effect size was reached significantly earlier in monkeys compared to humans (209±8.9 ms vs. 322±7.5 ms, $p<1e-14$, 2-sample KS test).

We next compared, within each species, whether there were latency differences between cells tuned for different stimuli. We used two methods (Fig. 6) to measure the response latency difference between the two most prominent cell categories that we found: face cells for conspecific and heterospecific faces (H^h versus H^m in humans and M^m versus. M^h in macaques). Using our selection criteria, we computed the number of cells that would be tuned for each category as we shifted the point of analysis from 500ms before the onset of fixation, until 1000ms after the onset of fixation (step size = 8ms). Our measure of latency was the point in time where the proportion of cells tuned exceeded that expected by chance for the first time (see Methods). Using this approach, we found that cells that were selective for conspecific faces responded significantly earlier than cells that were selective for heterospecific faces in both species (Δ_{human}=70ms, Δ_{monkey}=90ms, Fig. 6 A,B). In addition, we also confirmed this result using a moving-window ROC analysis and found a similar difference (Δ_{human}=62ms, Δ_{monkey}=38ms, Fig. 6 C, D). Together, this shows that information about conspecific faces is available at an earlier point of time relative to information about faces of other species in both humans and monkeys.

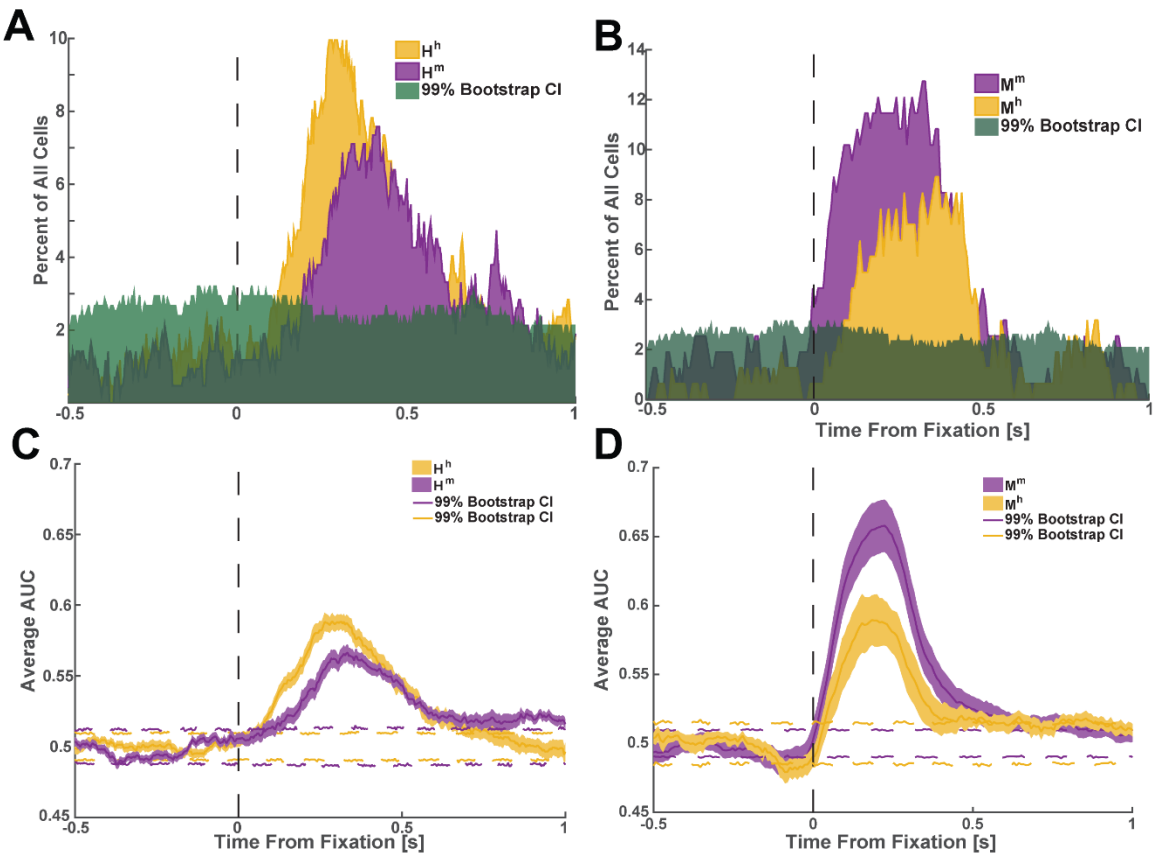


Figure 3.10: Face cells responded earlier and more strongly to conspecific compared to heterospecific faces. (A) Proportion of all recorded cells in humans (out of N=422) selective for fixations on conspecific (Hh, yellow) and heterospecific faces (Hm, purple). Shading indicates the 99% confidence interval. (B) Proportion of all recorded cells in monkeys (out of N=195) selective for fixations on conspecific (Mm, purple) and heterospecific (Mh, yellow) faces. (C,D) Average AUC as a function of time. Dotted colored lines indicate the 99% confidence interval.

3.3.5 Category-preference of fixation-sensitive neurons during covert attention

In this experiment, we asked whether fixation-target sensitive cells retain their tuning for peripherally presented stimuli when these are attended but not fixated. In a separate experiment, we recorded 119 cells (6 human subjects, 8 sessions) during a covert attention task with enforced central fixation. Images were identical to those used in the free-viewing task except that only a single image was shown at one (randomized) array location in isolation. Subjects maintained fixation at the center of the screen while a single stimulus was shown in the periphery (Fig. 4.1A). We found that of all the tuned cells in the covert condition (n=31/119, 26%), 25/31 of neurons were tuned to either human or monkey faces (1x4 ANOVA, n=16, Fig. 4.1B,E; n= 9, Fig. 4.1C,F respectively). For a subset (n=10) of these

face-selective neurons, we also recorded responses during the free-viewing task. Of these 10 cells, all maintained their face selectivity across the two task conditions and a comparison of all cells recorded in both tasks (n=31) revealed a high probability for cells to either be tuned in both tasks or neither tasks ($p=0.004$, Odds Ratio: 22.8, Fisher's test of association). Also, the proportion of face selective cells was not significantly different across the two tasks (25/31 and 69/85 in covert and free viewing task respectively, $\chi^2 = 0.0042$, $p = 0.94$). Notably, cells responded significantly earlier in the free-viewing condition compared to the cells recorded in the covert attention condition (Δ_{peak} = 191ms, Fig. 4.1D). This is expected because during the covert attention condition, the location of the stimulus was unpredictable, and thus deployment of covert attention could only be initiated following stimulus onset. In addition, the depth of selectivity was significantly larger in the covert compared to the free-viewing condition ($p<0.01$ for all n=422 cells in free viewing and n=119 in covert condition, $p<0.05$ for all n=85 tuned cells in free viewing and n=31 cells in covert condition, 2-sample KS test). Together, this data supports the hypothesis that amygdala neurons selective for faces and other complex visual objects are responsive to the currently attended visual stimulus, both during free viewing as well as during covert attention.

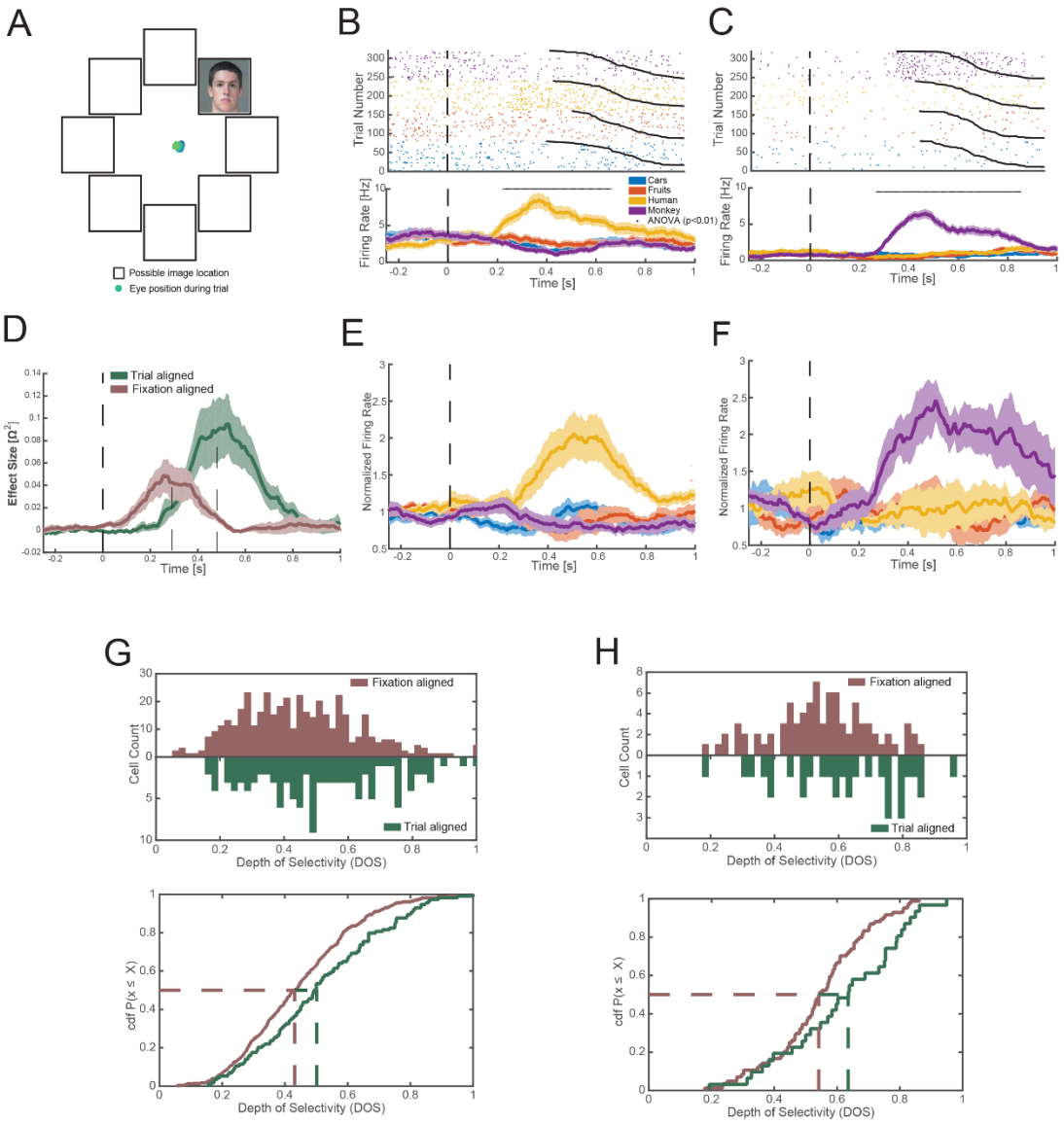


Figure 3.11: Face-selective amygdala neurons recorded in humans during covert tasks respond to covertly attended faces. (A) Subjects fixated at the center of the screen and indicated by button press whether a peripheral image shown showed a car. Shown is a single example trial, with eye-tracking data (blue) indicating that subjects maintained fixation. (B-C) Example face-selective neurons with a response selective to the identity of the peripheral stimulus. t=0 is stimulus onset. (D) Comparison of response of face cells in covert and free-viewing sessions for the subset of cells which were recorded in both tasks (4/7 sessions, n=10). The average effect size is shown fixation-and trial onset aligned. (E) PSTH of all human face-selective neurons (Hh, n=16) during the fixation-enforced covert attention condition. (F) PSTH of all monkey face-selective neurons (Hm, n=9) during the fixation-enforced covert attention condition. (G,H) Population-level comparison between the covert and free-viewing tasks for all recorded (G) and only visually tuned (H) cells. (G) DOS values were significantly larger in the covert attention task compared to the free-viewing task ($p<0.01$, 2-sample KS test). (H) DOS values were selectivity higher in the covert attention task ($p<0.05$, 2-sample KS test).

3.4 Discussion

Our results reveal that amygdala activity during active exploration of complex scenes is strongly modulated by the currently fixated stimulus. In contrast, previous studies in the amygdala of humans (Fried, MacDonald et al. 1997, Kreiman, Koch et al. 2000, Rutishauser, Tudusciuc et al. 2011) and macaques (Rolls 1984, Leonard, Rolls et al. 1985, Desimone 1991, Nakamura, Mikami et al. 1992, Gothard, Battaglia et al. 2007) relied on isolated single objects and were thus unable to investigate whether responses were modulated by gaze or not. Indeed, the assumption so far has been that because inferotemporal cortex neurons have large receptive fields for images shown in isolation (Tanaka 1993, Tovee, Rolls et al. 1994), the response of amygdala neurons should not depend on fixation location. However, here we find that the effective receptive field is relatively small in our task. This finding is similar to the response properties of "eye cells" in the macaque amygdala, which respond only when a monkey fixates on the eyes of another monkey (Mosher, Zimmerman et al. 2014).

Little is known about the effective receptive field sizes and their dependence on stimulus density for human and macaque amygdala neurons. In higher visual cortical areas in macaques, receptive fields encompass the entire hemifield (Gross, Bender et al. 1969, Boussaoud, Desimone et al. 1991, Barraclough and Perrett 2011). At the same time, many such neurons have heightened sensitivity to information present at the fovea (Moran and Desimone 1985, Rolls, Aggelopoulos et al. 2003). Once animals are allowed to actively explore complex visual scenes, however, receptive fields of neurons in macaque TE can shrink considerably (Sheinberg and Logothetis 2001, Rolls, Aggelopoulos et al. 2003). While it is possible that neurons in the amygdala inherit some of their properties from the same higher visual cortical areas (Amaral, Price et al. 1992, Rolls, Aggelopoulos et al. 2003, Barraclough and Perrett 2011), the significantly increased response latencies and complex selectivity changes we show make it unlikely that the responses we document are simply representing cortical input.

Our results show that the fixation-dependent responses were likely an effect of attention. This is because covert attention produced the same conclusions, even in some of the very same

cells. The strength of this result is limited to humans, because we did not perform the same task in monkeys due to the difficulty of training monkeys accustomed to free viewing on a covert attention task. An additional difference between covert and overt attention is that the sharpness of tuning (sparsity) was greater (more sparse) during covert compared to overt attention. A plausible explanation for this result is that the overt attention task still permits some influence from the other concurrently presented (unattended) images. In contrast, this source of competition is removed in the covert task (since only a single stimulus was presented). Indeed, unattended task-irrelevant peripheral faces can impair performance in a variety of settings (Landman, Sharma et al. 2014), and it is possible that the reduction of selectivity we observed here is a reason for this effect.

3.4.1 Role of face cells in social behavior

Our findings underscore the importance of using more naturalistic stimuli with inherent biological significance, in conjunction with behavioral protocols that better approximate natural vision. The finding that face-selective neural responses in the amygdala are strongly related to visual attention is ecologically important, because, in real social situations, directing one's gaze towards or away from faces and parts thereof (in particular the eyes) is a crucial social signal and sets the affective tone of the social interaction (Emery 2000). The amygdala is crucially involved in this process (Adolphs, Gosselin et al. 2005, Rutishauser, Mamelak et al. 2015), and impairments in directing gaze to faces are a prominent deficit in autism that is thought to be partially due to amygdala dysfunction (Baron-Cohen, Ring et al. 2000, Rutishauser, Tudusciuc et al. 2013). While a preference for features, such as the eyes, can be explained by perceptual properties (Ohayon, Freiwald et al. 2012), the conspecific-preference we showed cannot be attributed to low-level stimulus properties. Together, this indicates that face-sensitive cells in the amygdala might report not only the presence, but also the relative salience of stimuli. No such observations have been reported for cortical face cells, making it possible that this species-specific face signal might be computed locally within the amygdala.

3.4.2 Information represented by face cells

Face cells also responded to several other categories, either by a decrease or by a more moderate increase in firing rate relative to baseline. Indeed, both the number of pairs of categories that a cell's response differentiates and depth of selectivity indicated that neurons in both species differentiated between more pairs than would be expected by a sparse and specific response to just one category. Notably, cells in macaques differentiated between more pairs and had lower DOS values, indicating that macaque cells were less specifically tuned. Together, this suggests that primate amygdala neurons, including face cells, carry information about several categories, but that human neurons are more selective. Category selectivity is a prominent feature of visually responsive neurons in several areas of the human (Fried, MacDonald et al. 1997, Kreiman, Koch et al. 2000) and macaque (Bruce, Desimone et al. 1981, Perrett, Rolls et al. 1982, Gothard, Battaglia et al. 2007) temporal lobes. However, the amygdala of both species also contains more specific cells, such as "concept cells" that only respond to the face of a particular individual (Quiroga, Reddy et al. 2005), cells that signal certain emotions or facial expression (Gothard, Battaglia et al. 2007), and cells that signal the familiarity of stimuli (Rutishauser, Ye et al. 2015). It remains to be investigated whether these cells are similarly sensitive to fixation location.

3.4.3 Latency Differences

We found that, in both species, face-cells responded significantly earlier to faces of conspecifics relative to heterospecific faces. Behaviorally, both macaques and humans preferentially process faces of conspecifics more efficiently and are better at differentiating individuals of the same species (Pascalis and Bachevalier 1998, Dufour, Pascalis et al. 2006). In macaques, face-sensitive cells in the inferotemporal cortex differentiate between human and macaque faces (Sigala, Logothetis et al. 2011) and respond earlier to human compared to non-primate animal faces (Kiani, Esteky et al. 2005). The same neurons, however, showed no latency difference when comparing humans versus macaques (Kiani, Esteky et al. 2005). A new hypothesis motivated by our result is that the human versus macaque same-species

latency advantage is first visible in the amygdala as the result of the higher social significance attributed by the amygdala to conspecific faces.

Human single-neuron onset latencies are considerably slower compared to those of macaques (Leonard, Rolls et al. 1985, Mormann, Kornblith et al. 2008, Rutishauser, Mamelak et al. 2015) in many brain areas, including the amygdala (Mormann, Kornblith et al. 2008). However, inter-species comparisons of latencies are challenging because of variable experimental conditions, tasks, and stimuli. In particular, previous work in humans has argued that because receptive fields are large, control for eye movements is not necessary (Mormann, Kornblith et al. 2008). Here, we showed that this assumption is not valid. Instead, we performed a rigorous comparison of response latencies by comparing the fixation-aligned responses of face-cells tuned to conspecifics. This ensured that in both species, we relied on the earliest and strongest known amygdala response. With this approach, we determined that human amygdala neurons had response latencies that were on average ~100ms longer than those in macaques. Thus, our work shows that this frequently observed inter-species difference (Mormann, Kornblith et al. 2008) cannot be explained by methodological differences. This raises the important question of whether this latency difference is already present in higher visual areas or whether it first emerges in areas of the medial temporal lobe. This will require human single-neuron latency estimates in higher cortical visual areas, which have not been performed to date. Notably, recordings from early visual areas V2/V3 in humans indicate that the response latencies in these areas do not differ between monkeys and humans (Self, Peters et al. 2016). This raises the possibility that local processing in higher areas specific to humans is responsible for this substantial increase in response latency.

3.5 Conclusions

Faces are stimuli of high significance for primates, and the brains of several species contain multiple areas connected in a network specialized for face processing (Desimone 1991, Emery 2000, Tsao, Freiwald et al. 2006, Tsao, Moeller et al. 2008). Exploring the division of labor among the different nodes in this network has been a fruitful approach to capturing more general, circuit-level principles of neural computation. Indeed, a detailed analysis of

face cells throughout the brain revealed a distributed but interconnected system of cortical face patches specialized for different components of face processing (Kanwisher and Yovel 2006, Tsao, Moeller et al. 2008). However, most of what is known about this network has been derived exclusively from work in macaques, even though it is often assumed that the properties of this system are the same in humans (Barraclough and Perrett 2011). Here, we present critical, direct evidence for significant differences and commonalities. It is likely that the face-responsive properties of amygdala neurons arise, at least in part, through convergent inputs from several cortical areas where face cells have been identified. However, the face cells in the amygdala do not merely recapitulate the response properties of face cells in cortical areas, but show pronounced effects of species-specific relevance, and of attention. These findings revise our view of the amygdala's contribution to face processing from that of an automatic and broad detector, to that of a highly selective and attention-dependent filter. These effects likely constitute an essential ingredient for guiding processing in downstream regions, and ultimately for generating social behavior in real-world settings where many stimuli constantly compete for attention.

Methods

Human electrophysiology

Human subjects were patients being evaluated for surgical treatment of drug-resistant epilepsy who provided informed consent and volunteered for this study. Monocular gaze position was monitored at 500Hz (EyeLink 1000, SR Research). The institutional review boards of Cedars-Sinai Medical Center and the California Institute of Technology approved all protocols. We recorded bilaterally from the amygdala using microwires embedded in macroelectrodes. From each microwire, we recorded the broadband 0.1-9000Hz continuous extracellular signal with a sampling rate of 32kHz (Neuralynx Inc). One microwire on each macroelectrode served as a local reference (bi-polar recording). All included patients had clearly distinguishable spiking activity on at least one electrode in at least one amygdala.

Monkey electrophysiology

A custom-built 7-channel Eckhorn drive (Thomas Recording, Germany) advanced 7 microelectrodes (1-2 MΩ) to the recording targets in the right amygdala. The reconstructed anatomical location of each neuron recorded relative to a generic coronal section through the mid-amygdala is shown in Figure 1. Single unit activity was pre-amplified via a built-in head stage with 20 gain (Thomas Recording, Giessen, Germany), amplified and filtered (1,000 gain; 600-6,000 Hz filter, Lynx-8, Neuralynx, Bozeman, MT, USA), and sampled continuously at 40 kHz (Power 1401, Cambridge Electronic Design [CED], Cambridge, UK). Eye-position was monitored using ISCAN infrared eye tracker. Monkeys were seated in a primate chair 57 cm from a monitor and before each recording session began, they underwent a 9-point eye-position calibration with $\pm$1 dva resolution. The stimuli were displayed and the monkeys' behavior was monitored with the Presentation software (Neurobehavioral systems, Albany, CA, USA).

Behavioral task: free viewing

Monkeys were trained to fixate on a white cross. If the monkeys maintained gaze on the fixation spot for at least 100ms, a circular array of images subtending 23.4 x 23.4 dva was presented. Monkeys were allowed to freely scan the scene for 3-4 s, but were required to keep their gaze within the boundaries of the array for at least 3 s. Monkeys received a 0.5–1 ml juice reward followed by a 3s inter-trial interval if this condition was met. If the monkey failed to fixate, or looked outside the boundary of the image, the trial was terminated, reward was withheld, and the array was repeated. Humans were instructed to freely observe the arrays for a fixed amount of time (4s). After each array, a blank screen with a fixation cross was displayed for 1s.

Behavioral task: covert attention

This experiment was carried out only in humans. Human subjects were instructed to maintain fixation at the center of the screen. The same stimuli were used. Stimuli were displayed in the periphery (6 DVA) in one of eight possible locations (Fig 7A). Subjects were instructed to maintain fixation and answer a yes/no question about the image ("does the image contain a car, yes or no") with a button press. Images stayed on the screen until an answer was provided (with a time-out of 5s). In each session, subjects viewed 320 images chosen equally from the four stimulus categories (monkey face, human face, fruits, and cars).

Spike sorting and single-neuron analysis

The raw signal was filtered with a zero-phase lag filter in the 300-3000Hz band and spikes were detected and sorted using a semi-automated template-matching algorithm (Rutishauser, Schuman et al. 2006, Rutishauser, Cerf et al. 2014). In humans, channels with interictal epileptic spikes in the LFP were excluded. We used the same processing pipeline to process the monkey and human recordings (see supplementary methods). All PSTH diagrams were computed using a 250ms window with a step-size of 7.8ms. No smoothing was applied.

Localization of electrodes (humans)

Electrodes were localized based on pre-and post-operative T1 structural MRIs and a high-resolution amygdala atlas with identified sub-nuclei (Tyszka and Pauli 2016). Only electrodes that could be localized to the amygdala were included. We used the following processing pipeline to transform the post-operative MRI into the same space as the Atlas. We extracted the brains from the pre-and post-operative T1 scans (Segonne, Dale et al. 2004) and aligned the post-operative to the pre-operative scan with Freesurfer's mri_robust_register (Reuter, Rosas et al. 2010). We then computed a forward mapping of the pre-operative scan to the CIT168 template brain (Tyszka and Pauli 2016) using a concatenation of an affine transformation followed by a symmetric image normalization (SyN) diffeomorphic transform computed by the ANTs suite of programs (Avants, Duda et al. 2008). This resulted in a post-operative scan overlayed on the MNI152-registered version of the CIT168 template brain (Tyszka and Pauli 2016). We then used Freesurfer's freeview program to mark the electrodes as point sets to determine in which amygdala nucleus the tips of the microwires were located.

Processing of human eye-tracking data

Calibration was performed using the built-in 9-point calibration grid and was only used if validation resulted in a measurement error of <1 dva (average validation error was 0.62 dva). We used the Eyelink system which automatically annotates fixations and saccades from the continuous stream of data using a motion, velocity, and acceleration threshold (default thresholds). In order to get reliable estimates of tuning and latency, we used stringent selection criteria to exclude fixations from analysis that were too short or those who landed on the same image as the previous fixation (Fig S3, "conservative criteria").

Using look-onset instead of fixations

In the monkey, fixations and saccades were annotated using Cluster Fix (Konig and Buffalo 2014). In the human subjects, we used the annotation provided by the Eyelink system. However, to bypass potential differences in the two annotation approaches, we used the "look onset" instead of fixation onset for all analysis in this paper. The "look onset" is the point of time at which the first data point fell onto a particular image (Fig 3.11). This way, we pooled all successive fixations that fell on the same stimulus into a single "look" (Fig 3.11B). While the look duration is thus typically longer than fixation duration (Fig 3.11E-F), our analysis depends only on the onset and is thus insensitive to this difference. Aligning with look-onset instead of fixation onset resulted in qualitatively similar neural responses (Fig 3.11C-D). At the same time, using look-onsets instead of fixation onsets has several advantages, including that it (1) is insensitive to idiosyncrasies that might arise from the two separate annotation methods in the two species, (2) provides a more conservative estimate of the neural response latency, and (3) ignores fine structure in the neural response that may be driven by successive fixations on the same stimulus. Throughout the manuscript, fixation onset refers to look onset unless mentioned otherwise (i.e. Fig 3.11).

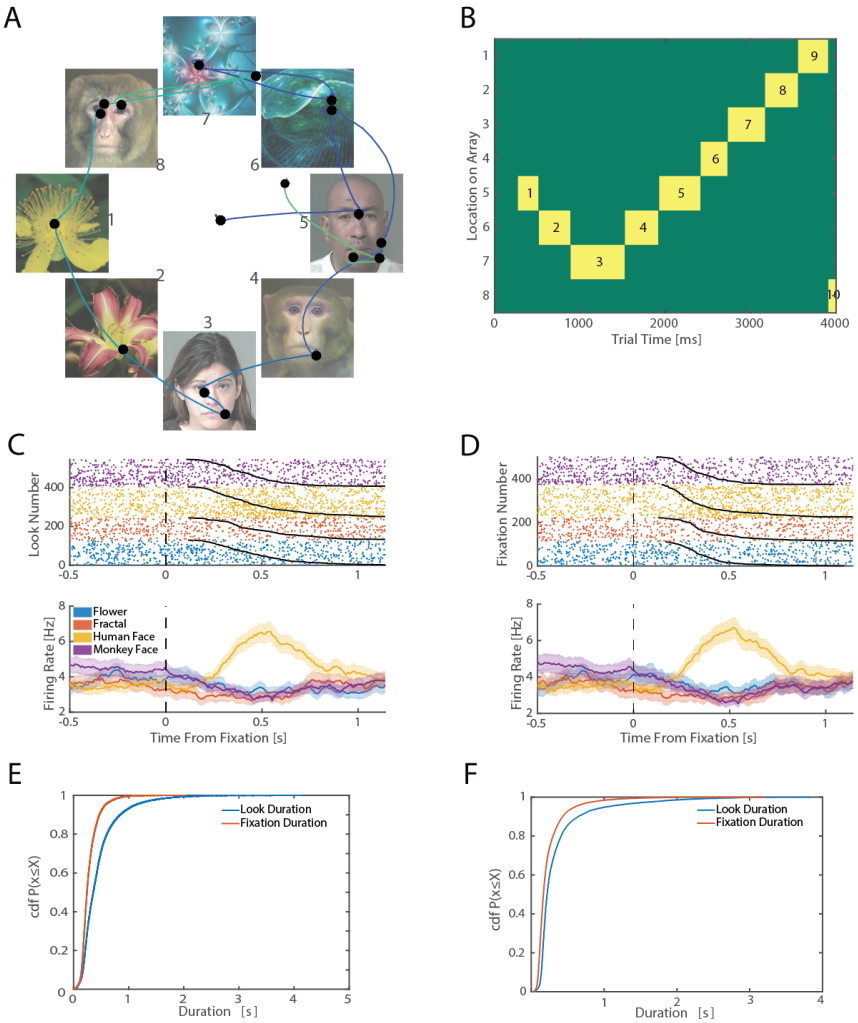


Figure 3.12: Comparison of fixation onset and look onset methods, related to Figure 1. (A) Example scan path from a single trial from a human subject. See Fig 3.1 for notation. (B) Summary of eye-tracking data into discrete periods of "looks" (yellow squares). Successive fixations that fall on the same image are pooled together into a single "look". The y-axis denotes the location of the look in the array as indicated in (A). (C, D) Comparison of a single-cell response, aligned to fixation onset (C) and look onset (D). Note the virtually identical response of the cell using the two criteria. For each, Raster and PSTH are shown. (E) Cumulative distribution for fixation (red) and look (blue) duration. Look duration was longer because of the pooling of several fixations into one look. (F) Same as (E), but for monkeys.

Selection of units

We determined whether a cell's response is sensitive to the identity of fixated stimuli using a 1x4 ANOVA of the spike counts during a 250ms-long time window centered on the point of time at which MI was maximal for each species (t=332ms and t=229ms, respectively). We excluded successive fixations that fall on the same category (Fig S3C, conservative criteria). To achieve this, we included only fixations that were not preceded or succeeded by fixation(s) on an image of the same category for at least 100ms. If the ANOVA was significant ($p<0.05$), we determined the category with the largest mean response in the same time window. This category was used as the preferred category of the cell.

Single-neuron ROC analysis

Neuronal ROCs were constructed based on the spike counts in a 250ms long window, centered at the peak of the mutual information. We varied the detection threshold between the minimal and maximal spike count observed, linearly spaced in steps of 1 spike. The AUC of the ROC was calculated by integrating the area under the ROC curve (trapezoid rule). The AUC value is an unbiased estimate for the sensitivity of an ideal observer that counts spikes and makes a binary decision (present or absent) based on whether the number of spikes is above or below a threshold. We generated a bootstrap distribution of the AUC values by randomly scrambling the image labels and computing the AUC values 1000 times. All statistical tests were based on the 99% confidence interval of this bootstrap distribution.

Mutual information (MI)

In order to determine the post-fixation window of analysis in monkeys and humans, we computed the mutual information between the spike counts (S) and the image category (C) using: $I(S,C) = \sum_{S,C} P(S,C) \cdot log_2 \frac{P(S|C)}{P(S)}$, where C is a discrete variable that can take 1 of 4 possible values, and S is also a discrete variable that can take 1 of N possible values, depending on the maximum firing rate of the cell. The mutual information was computed for each cell and at each point along the PSTH (from -0.5s to 1s around the fixation). The mutual information for each cell was then averaged to produce the mean trace. The location of the center of the fixed window for all follow-up analysis was set to the point of time at which MI was maximal (see Fig 3.6).

Estimation of latency

We relied on three different methods for latency estimation. (1) In the first case (Figure 5A), we compute for each cell the bootstrap distribution of the effect size, by scrambling the labels on each fixation. We then use the point where the cell's effect size (Ω^2) crosses the 99% confidence interval of the bootstrap distribution as our estimate of onset latency. (2) In a second method (Figure 6A, 6B), we systematically move our analysis window from 500ms before fixation to 1000ms after fixation (in increments of 8ms) and use our selection criteria to count the number of tuned cells. We also compute the bootstrap distribution of the number of tuned cells, by scrambling the labels on each fixation. We use the point where the number of tuned cells crosses the 99% confidence interval of the bootstrap distribution as our estimate of latency. Note that given our selection criteria ($p<0.05$), the chance level for tuning is 1/20. The chance level for a particular category (ex. to be tuned for human faces) is smaller, 1/80. (3) The third method (Figure 6C, 6D) relies on the average AUC, computed as the average of the individual cell AUCs for each category. We use the point where the average AUC cross the 99% confidence interval as our estimate of latency.

Behavioral Controls

We conducted a separate control experiment in 14 healthy individuals, in order to address the role that task instructions played in the way that subjects look at the images on the array. These control participants were randomly assigned to either the memory or free-viewing group. The memory group (n=7) was explicitly told to remember the images presented in the arrays. The free-viewing group (n =7) on the other hand, was not told about the memory component of the task and was simply instructed to look at the images on the screen. All subjects inspected the same exact arrays of images (n=52 trials), populated with images of fruits, cars, monkey, and human faces. The instructions were intended to mimic the task "instructions" that our subjects and the monkeys received. We compared the behavior between these two control groups as well as our subjects across 4 different metrics (Figure S5): (1) average looking time on an image in the array; (2) average number of images visited in a trial; (3) the proportion of first fixations that landed on a human face; and (4) performance on the recognition trials.

Assessment of selectivity

We used ROC analysis between all 6 possible pairs of stimulus categories to assess the number of pairwise comparisons that each neuron was able to differentiate. For each of the 6 possible comparisons, we computed the moving window AUC and compared this to the bootstrap distribution, which was generated by shuffling the fixation labels and computing the AUC 1000 times. In addition, we quantified the depth of selectivity DOS of each neuron by $DOS_i = \frac{n-\frac{\sum_{j=1}^{n} R_j}{R_{max}}}{n-1}$, where n is the number of categories (n=4), R_j is the mean response to category j, and R_{max} is the maximal mean response. D varies from 0 to 1, with 0 indicating an equal response to all categories and 1 exclusive response to one but none of the other categories. Thus, a DOS value of 1 is equal to maximal sparseness.

Regression analysis

We used the regression model $S(t) = \alpha_0(t) + C$ to estimate whether the firing rate S was significantly related to the factor category (C, 1-4). Spike counts S(t) were computed for a 200ms window that was moved in steps of 50ms. We quantified the effect size of the factor category using ω^2, which is less biased than percentage variance explained (Olejnik and Algina 2003). Models were fit and effect sizes calculated using the effect size toolbox functions mes1way and mes2way (Hentschke and Stuttgen 2011). The null distribution was estimated by randomly scrambling the fixation labels and fitting the same model 1000 times. Estimates of latency were based on the first time the actual value was located outside of the 99% confidence interval.

References

Adolphs, R. (2010). What does the amygdala contribute to social cognition? Ann N Y Acad Sci 1191, 42-61.

Adolphs, R., Gosselin, F., Buchanan, T.W., Tranel, D., Schyns, P., and Damasio, A.R. (2005). A mechanism for impaired fear recognition after amygdala damage. Nature 433, 68-72.

Adolphs, R., Tranel, D., and Damasio, A.R. (1998). The human amygdala in social judgment. Nature 393, 470-474.

Adolphs, R., Tranel, D., Damasio, H., and Damasio, A. (1994). Impaired Recognition of Emotion in Facial Expressions Following Bilateral Damage to the Human Amygdala. Nature 372, 669-672.

Amaral, D.G., Price, J.L., Pitkanen, A., and Carmichael, S.T. (1992). Anatomical organization of the primate amygdaloid complex. In The Amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction, J.P. Aggleton, ed. (New York: Wiley-Liss), pp. 1-66.

Baron-Cohen, S., Ring, H.A., Bullmore, E.T., Wheelwright, S., Ashwin, C., and Williams, S.C.R. (2000). The amygdala theory of autism. Neurosci Biobehav Rev 24, 355-364.

Barraclough, N.E., and Perrett, D.I. (2011). From single cells to social perception. Philos Trans R Soc Lond B Biol Sci 366, 1739-1752.

Boussaoud, D., Desimone, R., and Ungerleider, L.G. (1991). Visual topography of area TEO in the macaque. J Comp Neurol 306, 554-575.

Bruce, C., Desimone, R., and Gross, C.G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. J Neurophysiol 46, 369-384.

Cauchoix, M., and Crouzet, S.M. (2013). How plausible is a subcortical account of rapid visual recognition? Front Hum Neurosci 7, 39.

David, S.V., Vinje, W.E., and Gallant, J.L. (2004). Natural stimulus statistics alter the receptive field structure of v1 neurons. J Neurosci 24, 6991-7006.

Desimone, R. (1991). Face-selective cells in the temporal cortex of monkeys. J Cogn Neurosci 3, 1-8.

Dufour, V., Pascalis, O., and Petit, O. (2006). Face processing limitation to own species in primates: a comparative study in brown capuchins, Tonkean macaques and humans. Behav Processes 73, 107-113.

Emery, N.J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. Neurosci Biobehav Rev 24, 581-604.

Freiwald, W.A., Tsao, D.Y., and Livingstone, M.S. (2009). A face feature space in the macaque temporal lobe. Nat Neurosci 12, 1187-1196.

Fried, I., MacDonald, K.A., and Wilson, C.L. (1997). Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. Neuron 18, 753-765.

Gallant, J.L., Connor, C.E., and Van Essen, D.C. (1998). Neural activity in areas V1, V2 and V4 during free viewing of natural scenes compared to controlled viewing. Neuroreport 9, 2153-2158.

Gothard, K.M., Battaglia, F.P., Erickson, C.A., Spitler, K.M., and Amaral, D.G. (2007). Neural responses to facial expression and face identity in the monkey amygdala. J Neurophysiol 97, 1671-1683.

Grimaldi, P., Saleem, K.S., and Tsao, D. (2016). Anatomical Connections of the Functionally Defined "Face Patches" in the Macaque Monkey. Neuron 90, 1325-1342.

Gross, C.G., Bender, D.B., and Rocha-Miranda, C.E. (1969). Visual receptive fields of neurons in inferotemporal cortex of the monkey. Science 166, 1303-1306.

Gross, C.G., Rocha-Miranda, C.E., and Bender, D.B. (1972). Visual properties of neurons in inferotemporal cortex of the Macaque. J Neurophysiol 35, 96-111.

Kanwisher, N., and Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. Philos Trans R Soc Lond B Biol Sci 361, 2109-2128.

Kiani, R., Esteky, H., and Tanaka, K. (2005). Differences in onset latency of macaque inferotemporal neural responses to primate and non-primate faces. J Neurophysiol 94, 1587-1596.

Kreiman, G., Koch, C., and Fried, I. (2000). Category-specific visual responses of single neurons in the human medial temporal lobe. Nat Neurosci 3, 946-953.

Leonard, C.M., Rolls, E.T., Wilson, F.A., and Baylis, G.C. (1985). Neurons in the amygdala of the monkey with responses selective for faces. Behav Brain Res 15, 159-176.

Moeller, S., Freiwald, W.A., and Tsao, D.Y. (2008). Patches with links: a unified system for processing faces in the macaque temporal lobe. Science 320, 1355-1359.

Moran, J., and Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. Science 229, 782-784.

Mormann, F., Kornblith, S., Quiroga, R.Q., Kraskov, A., Cerf, M., Fried, I., and Koch, C. (2008). Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. Journal of Neuroscience 28, 8865-8872.

Mosher, C.P., Zimmerman, P.E., and Gothard, K.M. (2014). Neurons in the Monkey Amygdala Detect Eye Contact during Naturalistic Social Interactions. Current Biology.

Nakamura, K., Mikami, A., and Kubota, K. (1992). Activity of single neurons in the monkey amygdala during performance of a visual discrimination task. J Neurophysiol 67, 1447-1463.

Ohayon, S., Freiwald, W.A., and Tsao, D.Y. (2012). What makes a cell face selective? The importance of contrast. Neuron 74, 567-581.

Pascalis, O., and Bachevalier, J. (1998). Face recognition in primates: a cross-species study. Behav Processes 43, 87-96.

Paton, J.J., Belova, M.A., Morrison, S.E., and Salzman, C.D. (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. Nature 439, 865-870.

Perrett, D.I., Rolls, E.T., and Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. Exp Brain Res 47, 329-342.

Pessoa, L., McKenna, M., Gutierrez, E., and Ungerleider, L.G. (2002). Neural processing of emotional faces requires attention. Proc Natl Acad Sci U S A 99, 11458-11463.

Pessoa, L., and Ungerleider, L.G. (2004). Neuroimaging studies of attention and the processing of emotion-laden stimuli. Prog Brain Res 144, 171-182.

Quiroga, R.Q., Reddy, L., Kreiman, G., Koch, C., and Fried, I. (2005). Invariant visual representation by single neurons in the human brain. Nature 435, 1102-1107.

Rainer, G., Asaad, W.F., and Miller, E.K. (1998). Selective representation of relevant information by neurons in the primate prefrontal cortex. Nature 393, 577-579.

Rolls, E.T. (1984). Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. Hum Neurobiol 3, 209-222.

Rolls, E.T., Aggelopoulos, N.C., and Zheng, F. (2003). The receptive fields of inferior temporal cortex neurons in natural scenes. J Neurosci 23, 339-348.

Rutishauser, U., Mamelak, A.N., and Adolphs, R. (2015a). The primate amygdala in social perception - insights from electrophysiological recordings and stimulation. Trends Neurosci 38, 295-306.

Rutishauser, U., Tudusciuc, O., Neumann, D., Mamelak, A.N., Heller, A.C., Ross, I.B., Philpott, L., Sutherling, W.W., and Adolphs, R. (2011). Single-unit responses selective for whole faces in the human amygdala. Current Biology 21, 1654-1660.

Rutishauser, U., Tudusciuc, O., Wang, S., Mamelak, Adam N., Ross, Ian B., and Adolphs, R. (2013). Single-Neuron Correlates of Atypical Face Processing in Autism. Neuron 80, 887-899.

Rutishauser, U., Ye, S., Koroma, M., Tudusciuc, O., Ross, I.B., Chung, J.M., and Mamelak, A.N. (2015b). Representation of retrieval confidence by single neurons in the human medial temporal lobe. Nature Neuroscience 18, 1041-1050.

Sanghera, M.K., Rolls, E.T., and Roper-Hall, A. (1979). Visual responses of neurons in the dorsolateral amygdala of the alert monkey. Exp Neurol 63, 610-626.

Self, M.W., Peters, J.C., Possel, J.K., Reithler, J., Goebel, R., Ris, P., Jeurissen, D., Reddy, L., Claus, S., Baayen, J.C., and Roelfsema, P.R. (2016). The Effects of Context and Attention on Spiking Activity in Human Early Visual Cortex. PLoS Biol 14, e1002420.

Sheinberg, D.L., and Logothetis, N.K. (2001). Noticing familiar objects in real world scenes: the role of temporal cortical neurons in natural vision. J Neurosci 21, 1340-1350.

Sigala, R., Logothetis, N.K., and Rainer, G. (2011). Own-species bias in the representations of monkey and human face categories in the primate temporal lobe. J Neurophysiol 105, 2740-2752.

Tamietto, M., and de Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. Nat Rev Neurosci 11, 697-709.

Tanaka, K. (1993). Neuronal mechanisms of object recognition. Science 262, 685-688.

Tovee, M.J., Rolls, E.T., and Azzopardi, P. (1994). Translation invariance in the responses to faces of single neurons in the temporal visual cortical areas of the alert macaque. J Neurophysiol 72, 1049-1060.

Tsao, D.Y., Freiwald, W.A., Tootell, R.B., and Livingstone, M.S. (2006). A cortical region consisting entirely of face-selective cells. Science 311, 670-674.

Tsao, D.Y., Moeller, S., and Freiwald, W.A. (2008). Comparing face patch systems in macaques and humans. Proc Natl Acad Sci U S A 105, 19514-19519.

Vuilleumier, P., Armony, J.L., Driver, J., and Dolan, R.J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. Neuron 30, 829-841.

Zirnsak, M., and Moore, T. (2014). Saccades and shifting receptive fields: anticipating consequences or selecting targets? Trends Cogn Sci 18, 621-628.

Avants, B., Duda, J.T., Kim, J., Zhang, H., Pluta, J., Gee, J.C., and Whyte, J. (2008). Multivariate analysis of structural and diffusion imaging in traumatic brain injury. Acad Radiol 15, 1360-1375.

Bakker, R., Tiesinga, P., and Kotter, R. (2015). The Scalable Brain Atlas: Instant Web-Based Access to Public Brain Atlases and Related Content. Neuroinformatics 13, 353-366.

Hentschke, H., and Stuttgen, M.C. (2011). Computation of measures of effect size for neuroscience data sets. Eur J Neurosci 34, 1887-1894.

Konig, S.D., and Buffalo, E.A. (2014). A nonparametric method for detecting fixations and saccades using cluster analysis: Removing the need for arbitrary thresholds. Journal of Neuroscience Methods 227, 121-131.

Olejnik, S., and Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect size for some common research designs. Psychol Methods 8, 434-447.

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spatial Vision 10, 437-442.

Reuter, M., Rosas, H.D., and Fischl, B. (2010). Highly accurate inverse consistent registration: a robust approach. Neuroimage 53, 1181-1196.

Rohlfing, T., Kroenke, C.D., Sullivan, E.V., Dubach, M.F., Bowden, D.M., Grant, K.A., and Pfefferbaum, A. (2012). The INIA19 Template and NeuroMaps Atlas for Primate Brain Image Parcellation and Spatial Normalization. Front Neuroinform 6, 27.

Rutishauser, U., Cerf, M., and Kreiman, G. (2014). Data Analysis Techniques for Human Microwire Recordings: Spike detection and Sorting, Decoding, Relation between Neurons and Local Field Potential. In Single Neuron Studies of the Human Brain, I. Fried, U. Rutishauser, M. Cerf, and G. Kreiman, eds. (Boston: MIT Press), pp. 59-98.

Rutishauser, U., Schuman, E.M., and Mamelak, A.N. (2006). Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, in vivo. J Neurosci Methods 154, 204-224.

Segonne, F., Dale, A.M., Busa, E., Glessner, M., Salat, D., Hahn, H.K., and Fischl, B. (2004). A hybrid approach to the skull stripping problem in MRI. Neuroimage 22, 1060-1075.

Spitler, K.M., and Gothard, K.M. (2008). A removable silicone elastomer seal reduces granulation tissue growth and maintains the sterility of recording chambers for primate neurophysiology. J Neurosci Methods 169, 23-26.

Tyszka, M.J., and Pauli, W.M. (2016). A high resolution in vivo MRI atlas of the adult human amygdaloid complex. submitted.

Landman, R., Sharma, J., Sur, M., and Desimone, R. (2014). Effect of distracting faces on visual selective attention in the monkey. Proc Natl Acad Sci U S A 111, 18037-18042.

Chapter IV: Distinct Neuronal Populations Signal Choice for Recognition Memory and Categorization-based decisions in the Human Medial Frontal Cortex

4.1 Overview

Decision-making relies on the accumulation of evidence in favor of a particular choice (Gold and Shadlen 2007). While this process is relatively well-understood for perceptual decisions, little is known about decisions that rely on internal representations such as memories. Here, we investigated the neural mechanisms of decisions that depend on two types of internal representations: a recognition memory and a visual categorization decision. Across 23 sessions in 9 patients, we recorded simultaneously from single neurons in the human medial-frontal cortex (MFC, n = 399) and medial temporal lobe (MTL, n = 360). We show that (1) there are distinct populations of cells in the MFC encoding recognition memory (49/399) or categorization-based choices (38/399); (2) visually-selective MTL cells are insensitive to such task conditions; and (3) theta-band spike-field coherence between field potentials in the MTL and action potentials in the MFC are enhanced during the memory compared to visual categorization decisions. This suggests that inter-areal communication between MTL and MFC may be facilitated selectively in tasks that rely on recognition memory-based information. These results suggest that memory representations are conveyed from the MTL to the MFC, and that specific neurons within the MFC represent abstract action-independent choices.

4.2. Introduction

There are few things more pervasive in everyday life than having to make decisions. Some decisions are effortful and require a lot of deliberation ("Which college should I go to?") while others are trivial to make ("Is this portion of food bigger than that one?"). Most of what we know about the *cellular mechanisms* facilitating decision-making comes from the domain of perceptual decisions (see (Gold and Shadlen 2007) for a review). Despite being domain specific (i.e. perceptual), these studies have collectively sketched a conceptual framework for what constitutes a decision process. A decision process must include a

representation of the task-relevant information, a choice space ("yes" versus "no" or potentially more options), and in between, a way of accumulating evidence in favor of a particular decision (i.e. using a decision variable). There might also be additional facilitating processes; on the input end, task and sensory data would need to be processed into *task-relevant* information, and on the output end, there could be additional processes that map a choice onto a particular action (e.g. a button press or saccade). In this paper, we adopt a similar conceptual framework to study decisions that rely on internal representations of image category and recognition memory. We operationalized these two conditions by presenting an image and asking one of two different types of questions: "Is this an image of a car" and "Have you seen this image before?" The decision processes we study here are different from perceptual decisions in two important ways. The first has to do with the nature of the stimuli. In most perceptual decision tasks, the stimuli are dynamic whereas the stimuli we use here (i.e. a single static image) reveal all the available information at once. Secondly, in perceptual decisions, sensory processes give rise to momentary evidence, which can be accumulated in a decision variable. In the decision processes we study in this paper, the stimulus itself does not contain all the relevant information, and the moment-by-moment evidence is computed based on an association or memory retrieval process (Shadlen and Shohamy 2016). The key question that we address in this paper is whether we can observe evidence of this decision process (which relies on internal variables) at the level of single cells. With our recordings, we target several brain areas we believe are good candidates for observing single cell evidence for such a decision process. Given the memory component of our task, a prime candidate for representing the necessary evidence for such a decision is the medial temporal lobe. Specifically, we target two areas in the temporal lobe, the hippocampus (mid-body to anterior hippocampus) and the amygdala. It has been shown in human single-unit recordings, that both of these areas contain novelty-responsive cells (Rutishauser, Ye et al. 2015). Furthermore, given the time scales of the visual memories we are probing (on the order of minutes), we can expect them to be strongly MTL dependent (Squire 1992, Jeneson and Squire 2012). Given these findings, we believe that in our task, the MTL contains the signals that are precursors of choice. This speaks to the recognition memory component of the task. In a separate

condition, we also ask about the broad visual category of the image (e.g. "Is this a human face?"). Unlike the recognition memory questions, we do not believe this is MTL dependent. Even though there is a strong category signal in cells within the MTL, their response latency (Rutishauser, Ye et al. 2015) is not consistent with the decision times we observed here on the visual categorization trials.

In terms of the representation of task–demands and the necessary decision variables, a likely candidate is the medial frontal cortex, which has been implicated in both memory and decision-making (see (Euston, Gruber et al. 2012) for review[1]). Here we target two areas in the MFC, namely the dorsal anterior cingulate cortex (dACC) and pre-supplementary motor area (pre-SMA) in order to look for representations of task and decision variables. Evidence of representation of decisions variables from the medial frontal cortex of humans comes both from fMRI studies (Wunderlich, Rangel et al. 2009) as well as single-units recordings (Mukamel, Ekstrom et al. 2010).

One key component of this paper is to demonstrate functional interactions between these two areas (i.e. hippocampus/amygdala, pre-SMA/dACC). There is a plethora of anatomical and electrophysiological studies, primarily in rats, that demonstrate the existence of a monosynaptic pathway from the hippocampus to the medial pre-frontal cortex (Swanson 1981, Siapas, Lubenov et al. 2005). These functional interactions are also of particular interest since animal studies have shown that they are severely impaired in clinical conditions such as schizophrenia (Sigurdsson, Stark et al. 2010).

Given this past work, our hypotheses in this paper are four-fold: (1) representations of decision variables are present exclusively in the medial frontal cortical cells; (2) these representations are modulated by task demands; (3) MTL representations are not sensitive to task demands; and (4) the interaction between the two areas is mediated by theta-band phase-locking MFC cells and hippocampal local field potentials.

[1] The authors focus primarily on medial pre-frontal cortex whereas most of our electrodes are a little more posterior (dACC and pre-SMA).

4.3 Results

4.3.1 Task and Behavior

Subjects were shown single images of objects chosen from 4 visual categories (cars, fruits, human and monkey faces). For each image, subjects were asked to answer a single question. The question was either: "Have you seen this image before?" or, "Is this an image of a fruit?" The first question required access to recognition memory, whereas the second requires access to high-level visual representations to perform categorization. Subjects gave yes/no responses with a button press or a saccade (**Figure 4.1A**). Subjects were not given any feedback until the end of the experiment. Except when responding with saccades, they were asked to maintain fixation on the image. Subjects fixated well, as measured by the number of eye-tracking data points that fell within the boundary of the image during a trial (94 ± 1.7%, see **Figure 4.1C** for example session, hand trials are shown in the left panel and eye trials in the right panel). Accuracy was higher on the categorization trials (97.8 ± 0.6%) compared to the memory trials (72.7 ± 1.4%). This condition difference was also evident in the decision times, with response time on memory trials being significantly slower (~300ms) than categorization trials (**Figure 4.1F**, 1.24s ± 0.02, and 0.94s ± 0.03 respectively, $p<1e-22$, 2-sample t-test). Performance on the memory task was a function of image category, with the performance being lowest for images of monkeys and cars, and higher for human faces and images of fruits (**Figure 4.1H;** pairwise t-tests, monkey-car, 1.43e-7; car-human, 7.06e-8; human-fruits, 0.037). This category separation was also evident in the ROC analysis of the behavior. Subjects generated fewer false positives for images of fruits and human faces than they did for cars and monkeys (see **Figure 4.1 I, J**; $p = 1.45e-8$, paired t-test). When asked about strategies for remembering the items in the different image categories, subjects consistently mentioned the fact that images of fruits could be named (ex. "strawberry") as one of the reasons why they thought the fruits were easier to remember. As expected, the subjects' performance improved as a function of time, demonstrating incremental learning of the repeated stimuli (**Figure 4.1G**, $p =3.9e-49$, mixed effects logistic regression). Performance on the visual categorization trials, in contrast, did not change as a function of time, demonstrating that learning was not required to answer this question. We also observed

differences in reaction time within condition, between the two different responses (i.e. "yes" and "no"). In the memory condition, subjects were significantly slower at saying "no" than "yes" (**Figure 4.2 A**). This is a common pattern in recognition memory tasks, and thus shows that subjects utilize recognition memory (Stern and Hasselmo 2009) . In contrast, in the visual categorization condition, subjects were slower at saying "yes" than "no" (**Figure 4.2 B)**. These condition differences were evident at the level of individual sessions (**Figure 4.2C**, $p = 1.4e\text{-}7$ for memory condition, $p = 0.03$ for category condition, t-test). Making a particular image the target category on the visual categorization trials did not improve memory for that image category on the following memory blocks (**Figure 4.2D**). Together, this behavioral data shows that the subject's performance reflected the task demands, with longer decision times, error rates, and learning happening specifically during the memory condition.

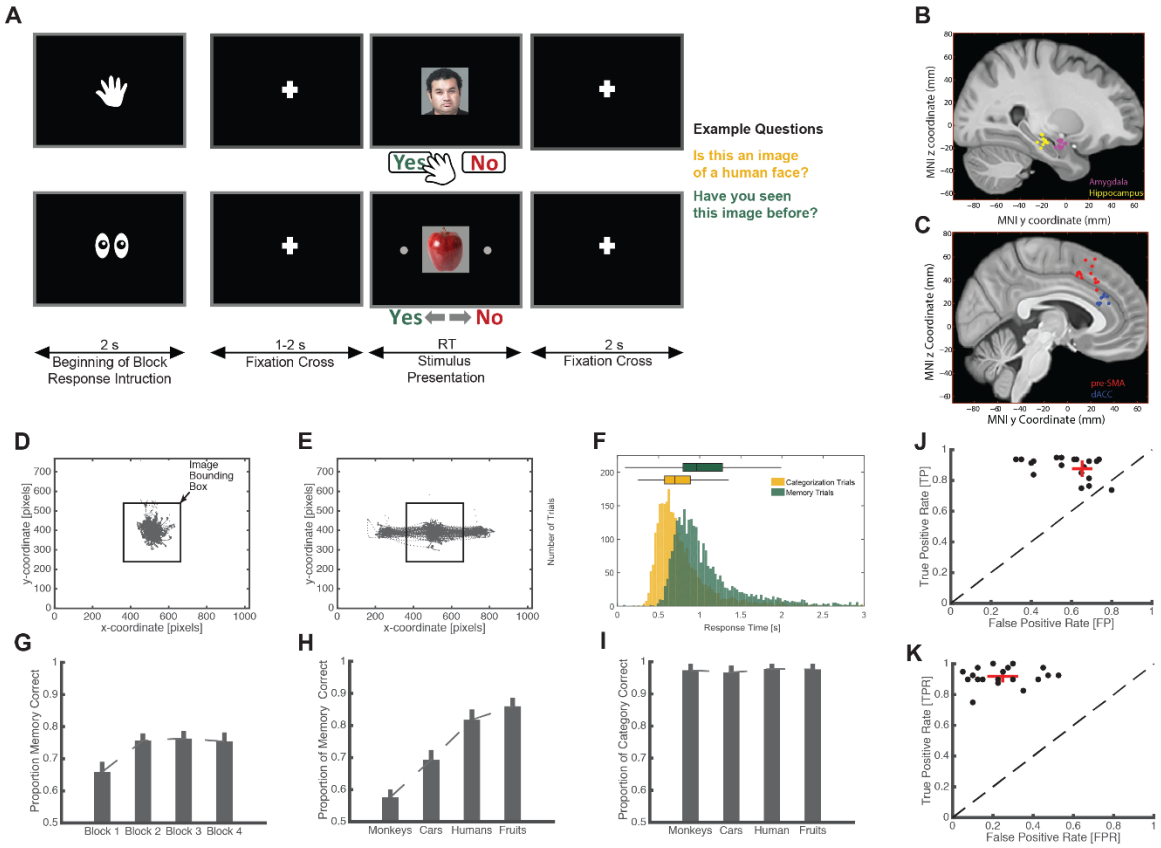


Figure 4.1: Task, behavior, and electrode locations (a) The subjects responded using either their hands (example trial on **top)** or by saccading to one of the target locations on the left and right of the image (example trial on the **bottom**). Yes and No responses are always on the left and right respectively. The subject is told *how* to respond (i.e. hand versus saccade) and to *what* question (i.e. recognition memory versus image categorization) to address at the beginning of each block (40 images/block). The question varied block by block, with category and recognition memory questions interleaved over the 8 blocks. The first block was always an image categorization block. For each of the categorization blocks, the target category was randomly assigned to one of the four available image categories. **(b)** Locations of the MTL electrodes, with hippocampal electrodes shown in yellow and amygdala electrodes shown in magenta. **(c)** MFC electrodes with pre-SMA electrodes shown in red and dACC electrodes shown in blue. **(d)** Eye-tracking data from an example session, showing only the trials with button press (160/320 trials). **(e)** Same as (c) but for eye response (160/320 trials). Note that the subject only breaks fixation on the center of the image in order to make the response. **(e)** Reaction time differences between the two conditions (memory condition, mean ± sem, 1.22 ±0.017s; category condition, 0.89 ± 0.02s; p = 2.5e-216, 2-sample KS test). **(g)** Performance on memory questions improves over the course of the experiment (p<1e-10, logistic regression, mixed effects model). **(h)** Memory performance by image category. **(i)** Performance on categorization questions across the 4 image categories. **(j)** Behavioral ROC for memory trials for the monkey and car categories. The red cross marks the mean and 95% confidence interval for the distribution of points **(k)** Same as (i) but for fruits and human faces.

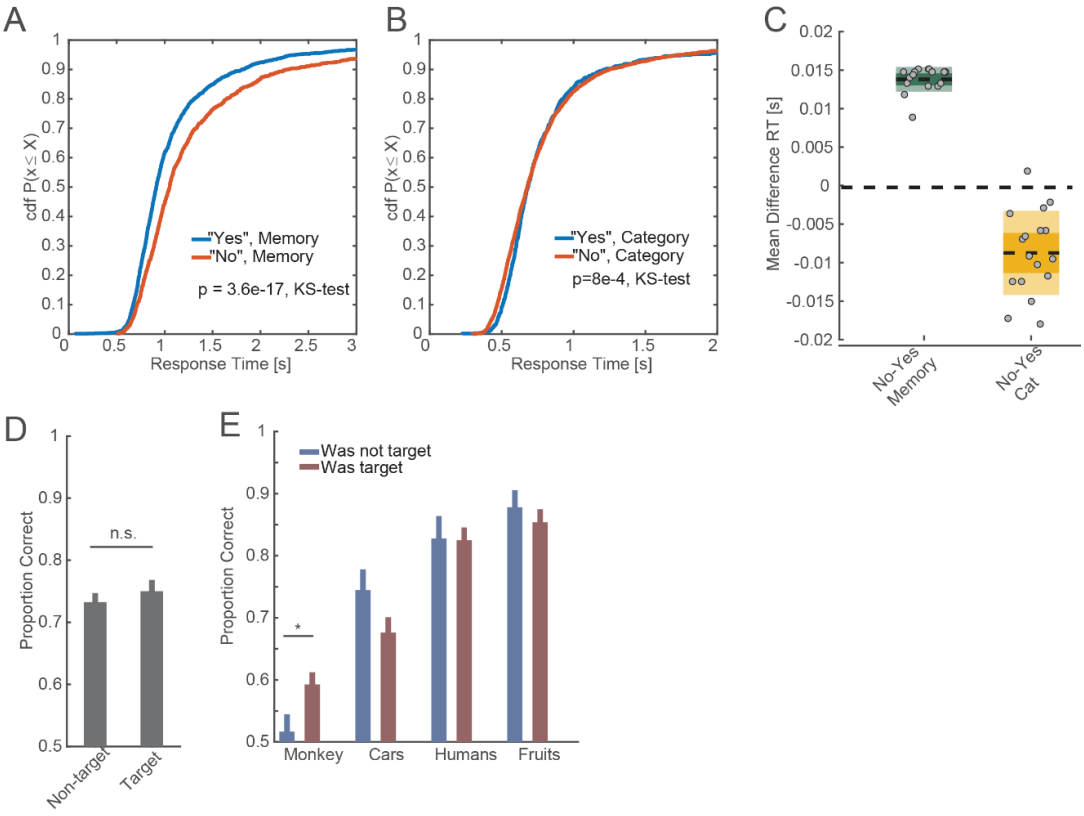


Figure 4.2: Yes vs. No differences in memory and categorization trials and effect of target
(a) Cumulative distribution of reaction times for "yes" and "no" responses during the memory trials shown for all included sessions. Subjects were significantly slower at saying "No, I have not seen that image before" than "Yes, I have seen that image before." **(b)** Same as (a) but for the categorization trials. **(c)** Plotted for each session, is the mean difference between "yes" and "no" responses shown separately for memory trials (green, p= 1.2773e-16, t-test) and categorization trials (yellow, p = 6.2e-6, t-test). **(d)** Making a category the target on a categorization block does not enhance memory for that category on follow-up memory blocks. **(e)** Same as (d) but shown separately for each category. The only category that receives a small boost is that of monkey images (p=0.03, paired t-test).

4.3.2 Electrophysiology

We recorded from 360 neurons in the amygdala and hippocampus (MTL, n=191 in amygdala and n=169 in hippocampus), and 399 neurons in the dorsal anterior cingulate cortex and pre-supplementary motor area (MFC, n=202 in dACC and n=197 in pre-SMA) in eight neurosurgical epilepsy patients, across 23 sessions. In MFC, we identified 78/399 (20%) cells that signaled the choice made by the subject. Selection for choice cells was done using a 1s window centered at 700ms after stimulus onset [200 – 1200ms]. The choice of window center was determined from the peak of the choice Ω^2-effect size, computed across the entire population, prior to selection. Spike counts in this bin were regressed against the response (binary, "yes" or "no") independently in the visual categorization and memory conditions using the linear model: $\mathbf{FR \sim 1 + \beta_1 \bullet Response + \beta_2 \bullet RT}$. We explicitly controlled for reaction time differences between the two different answers by including a reaction time term in the regression. Cells that showed a significant $\boldsymbol{\beta_1}$ coefficient were selected and included in the pool of choice cells. Response preference for "yes" or "no" was determined from the signal of the $\boldsymbol{\beta_1}$ coefficient (positive = yes, negative = no). Note that the selection was done using completely independent trials for memory choice cells and categorization choice cells. Of these selected cells, 49/78 signaled the choice made in the memory condition, 38/78 signaled the choice made in the visual categorization condition and 9/78 that signaled both (i.e. significant $\boldsymbol{\beta_1}$ coefficient in both selection regression models). The small overlap between these two populations highlights one of the key properties of the choice cells, namely that they specialize in 1 type of decision, either memory or visual categorization.

We isolated an additional set of 62/399 cells in the MFC that discriminated between the image categories. Selection was done using all available trials, with a 1x4 ANOVA and spike counts measured in a 1-second window [200 – 1200ms]. The set of cells selective for image category and choice in the MFC were largely disjointed, with only 15/399 showing tuning for both. In the MTL, we isolated 79/360 cells that showed visual selectivity for one of the four image categories (Ω^2 effect size = 0.08±0.01).

4.3.3 Distinct Populations of Cells in MFC Encode Choice in the Recognition Memory and Categorization Task

In the medial frontal cortex, 78/399 (20%) of neurons signaled the choice that the subject made. Here, by choice, we refer to the answer provided, which was either "yes" or "no". **Figure 4.3 A, B** shows two example choice cells recorded in the anterior cingulate cortex (left) and pre-supplementary motor area (right) in two different patients. The cell in Panel A shows a preference for a "yes" response in the memory condition, while the cell in Panel B shows a preference for a "no" response in the memory condition. **Figure 4.4 A, B** show three more examples of memory choice cells and categorization choice cells respectively. The t-statistic is shown in **Figure 4.3E** for both categorization choice cells (yellow), memory choice cells (green), for cells that single choice in both conditions (purple). The numbers for memory and categorization choice cells were both highly significant when compared to the null distribution (**Figure 4.4 E**, $p<0.001$). To demonstrate that the choice cells specialize in either condition, we show the average PSTH for preferred and non-preferred responses for both the categorization choice condition (**Figure 4.3C**) and memory choice condition (**Figure 4.3D**). The top panel of **Figure 4.3C** shows the PSTH for categorization choice cells, split by preferred versus non-preferred responses only during the categorization trials. Performing the same split for the memory trials (bottom panel) shows that these cells do not contain choice information in the memory condition. Visa versa, the memory choice cells in **Figure 4.3D** contain no choice information during the categorization trials.

We next tested whether the response of choice cells could be explained by activity related to the novelty/familiarity and/or the category of the stimulus. To test this possibility, we quantified the amount of information that cells carried about choices, new/old, and visual category separately for the two task conditions (**Figure 4.3G**, blue bar) (**Figure 4.3G**, red bar). For this, we used an ROC analysis approach, by which we quantified the amount of information carried by each cell using the area-under-the-curve (AUC) metric. The AUC was computed between preferred versus non-preferred responses during the memory condition, new versus old ground truth in the categorization condition, and new versus old ground truth in the memory condition. We first focused on memory-choice cells (n=49). We found that memory-choice signaling cells carried significantly more information about choices (yes versus no, regardless of whether this choice was correct or incorrect,

p = 0.59, paired t-test) compared to information about whether a stimulus was old or new (ground truth; 0.61 ± 0.01 vs , 0.55 ± 0.01; p=6.8321e-07, paired t-test). In addition, we examined whether memory-choice cells continued to provide information about the novelty/familiarity of stimuli during the categorization task, where this variable was not relevant. This revealed that the AUC value for new/old information during categorization trials (blue condition) was not significantly different from chance (0.51 ± 0.007, p = 0.34). Together, this shows that memory-choice cells predict choices better than ground truth, and that they do not signal new/old information during the categorization task.

We next performed a similar analysis for categorization-choice cells. The key difference between the two conditions is that while "yes'' and "no" responses can be mapped directly onto "new" and "old" respectively (allowing for direct comparison), no such clear mapping exists for category. The reason for this is that the target category was randomly selected for each categorization block.

In order to test whether category choice cells contain information about the ground truth (in this case, the image category) rather than choice, we collapsed the image categories onto a preferred and non-preferred binary variable. The preferred image category was dictated by whatever image category was selected as the target for the block. For example, asking the subject to answer the question, "Is this an image of a monkey?" would set the target to monkey faces, and therefore the preferred category. This allowed us to map "yes" responses to the preferred category, and "no" to the non-preferred category for a direct comparison. The amount of information for preferred versus non-preferred in categorization choice cells was not different from the null distribution (p = 0.54).

Next, we tested whether choice cells were sensitive to the response type, hand versus saccade. **Figure 4.3 F** shows the AUC values for preferred versus non-preferred responses in the subset of choice cells that were recorded in sessions with both hand and saccade response types (see **Figure 4.4 C** for a single cell example). While choice cells show some preference for one versus the other response type, there is no significant effect at the population level (p = 0.18, paired t-test).

Lastly, we compared the choice information at the population level across all MTL and MFC cells. To do so, we created a pseudo-population (Meyers, Freedman et al. 2008) from all available cells, separately in MTL and MFC (see Methods). We compared choice information for all trials and

separately only for memory trials in MFC and MTL (**Figure 4.4 D,** left panel included all trials, right panel is only for memory choice trials). In both tests, choice information was significantly stronger in the MFC population than in MTL (p = 1.87e-24 for all trials; p = 4.58e-21 mem trials; paired t-test, measured at t = 750ms after stimulus onset). This shows a clear functional separation between these two areas, with MTL cells representing features that pertain to the stimulus itself and MFC cells representing choice.

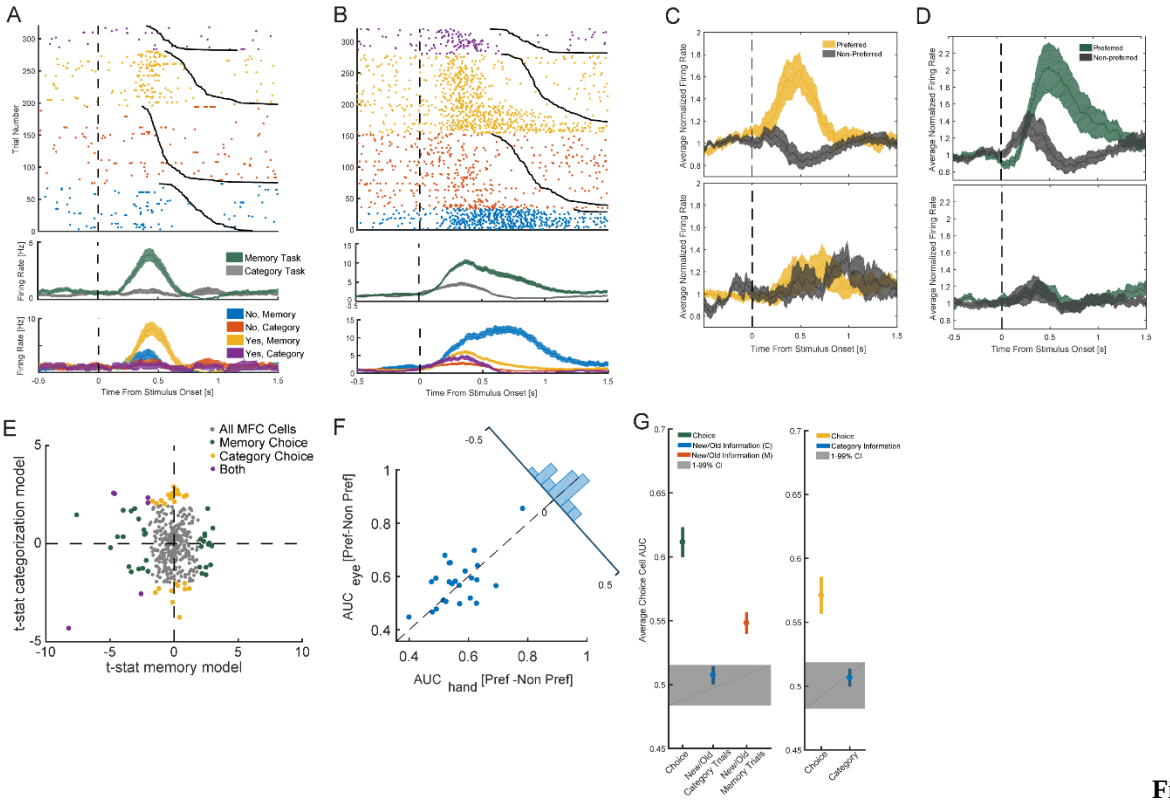


Figure 4.3: Separate neuronal populations in medial frontal cortex signal choice in a recognition memory and categorization task. (a) Example choice cells in MFC. This example ACC cell fires preferentially for "yes" (i.e. "I have seen this before") responses in the memory condition. Raster condition PSTH (i.e. memory versus visual categorization), and choice PSTH (each condition is split into "yes" versus "no" traces) correspond to the top, middle, and bottom panels respectively. **(b)** An example cell recorded from pre-SMA that preferentially fires for "no" (i.e. "I have not seen this image before") responses in the memory condition. **(c)** Average PSTH for cells that separate "yes" versus "no" in the visual categorization condition only (**top panel).** Also shown, is the average PSTH for the "yes" and "no" responses made in the memory condition (**bottom panel**). The PSTH collapses, suggesting that these cells do not have carry any choice information in the memory condition. **(d)** Same as (c) but for memory choice cells. **(e)** t-statistic shown for memory (green) and categorization (yellow) choice cells **(f)** Choice cells are insensitive to response type (eye vs. hand AUC, p = 0.54, sign test). **(g) On the left,** *green bar* = "yes" versus "no" information in memory choice cells (n=49) as measured by AUC (mean ± sem) on recognition memory trials, *blue bar* = new (corresponds to "no") versus old (corresponds to "yes") information in memory choice cells as measured by AUC on categorization trials, *red bar* = new versus old information in memory choice cells as measured by AUC on memory trials. Note that the red and green conditions share all *correctly* answered recognition memory questions. Memory choice cells do not contain memory

information on categorization trials (blue condition, p=0.32, permutation test of the mean). Memory choice cells represent choice and not a recognition memory signal during the recognition trials (green vs. red trace, p<0.005, paired t-test). **On the right,** *yellow bar* = "yes" versus "no" information in category choice cells (n=32) as measured by AUC (mean ± sem) on categorization trials, *blue bar* = "preferred category" versus "non-preferred category" on recognition memory trials, where preferred is defined by the target in the preceding categorization block. Category choice cells do not contain category information (blue condition, p=0.41 permutation test).

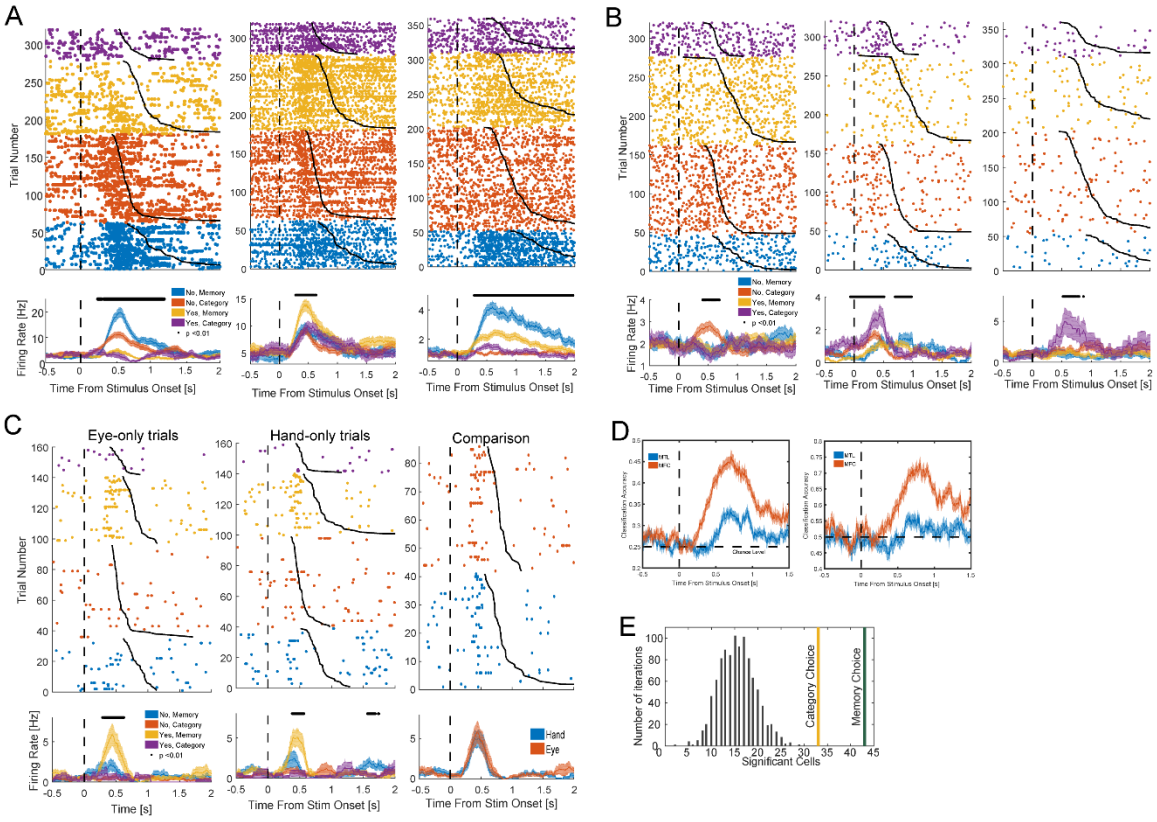


Figure 4.4: MFC and not MTL neurons represent choice. (a) Three additional example cells that were selected as memory choice cells. Example 1 satisfies criteria for both a categorization choice cell as well as a memory choice cell. It does however, show clear preference for one condition over the other. **(b)** Example categorization choice cells. The first example responds to "yes" while the second and third examples respond preferentially to the "no" response during the categorization condition. **(c)** An example cell that shows tuning for a "yes" memory choice is indifferent to response type. The first panel shows the raster and PSTH only for hand trials, while the second shows only eye trials. The third panel shows the preferred response, for both hand and eye trials (no significant difference). **(d)** Pseudo-population decoding of choice (4 possible outcomes, 25% chance level) using MFC cells (red trace) and MTL cells (blue trace). On the left, decoding results using all available trials. On the right, memory trials only (see Methods for detailed description of approach). (e) Null distribution of choice cell selection, 1000 permutations. Observed numbers are shown in yellow (categorization choice cells) and green (memory choice cells).

4.3.4 Cells in MTL encode image category but not choice

In the medial temporal lobe, we recorded from in total 360 cells from the amygdala (n=191) and hippocampus (n=169) combined. Of these, the response of 79 (22%) cells was visually selective. For selection, we used a 500ms window, centered at 350ms [100ms – 600ms] after stimulus onset. We performed a 1x4 ANOVA between spike counts and image category through all trials in the selection window and included all cells for which $p<0.05$. The preferred category was determined based on firing rate; each cell was assigned the image category that elicited the largest response as the preferred category (non-preferred included the other three image types, **Figure 4.5C**). Most cells in MTL preferred images of human faces (Minxha, Mosher et al. 2017) and fruits (**Figure 4.5 B). Figure 4.5A** shows an example of a human-face selective cell. Decoding accuracy was higher than chance level for all categories (**Figure 4.5D).** A cell that can discriminate across all image categories (i.e. has perfect information about all four categories) and a cell that can only discriminate 1 out of the 6 possible pairs (i.e. $n_{combos} = \binom{4}{2} = 6$) can equally qualify as visually selective using the selection criteria that we outline here. In order to determine the amount of information across all visually selective cells, we computed AUC between all six possible image pairs (**Figure 4.5E**). The peak of this distribution is at 3, which corresponds to a cell that can discriminate a single category from all the other three. This suggests that tuning in MTL is not mixed across categories, but individual cells tend to only respond to a single image category (see **Figure 4.6A** for 2 additional examples). Next we compared how visual representations in MTL change with task condition. Using a combination of metrics that include Ω^2- effect size (**Figure 4.5F**), AUC (**Figure 4.5G**), and population decoding (**Figure 4.6 B, C**), we compared the amount of image category information in the two different task conditions. Peak Ω^2- effect size was not significantly different between the two task conditions ($p = 0.81$, paired t-test, tested at t = 363ms). The AUC comparison between the two trial types also showed no difference between the two conditions across the whole population ($p = 0.33$, paired t-test, shown in black) and the visually selective cells ($p = 0.38$, paired t-test, shown in cyan).

We next asked if MTL cells represent choice-related variables. Using the same selection criteria that we used to select choice cells for the MFC population, we find n = 17/360 cells ($p=0.53$, permutation test) tuned for memory choice, and n = 23/360 ($p=0.1$, permutation test) cells for category choice

cells (**Figure 4.5H**). At the population level, representation of image category is much stronger in MTL than in MFC, both when we use all available trials (**Figure 4.6B**, p = 3.1e-66, paired t-test, evaluated at peak decoding accuracy for each area independently) and when we only use memory trials (**Figure 4.6C**, p = 6.7410e-38, paired t-test, evaluated at peak decoding accuracy). Together, this result shows that representations in the MTL population are stable and not sensitive to task demands. Furthermore, this data again highlight the difference between the stable, stimulus-specific representations in the MTL, and the representations of task-specific variables that we find in MFC cells.

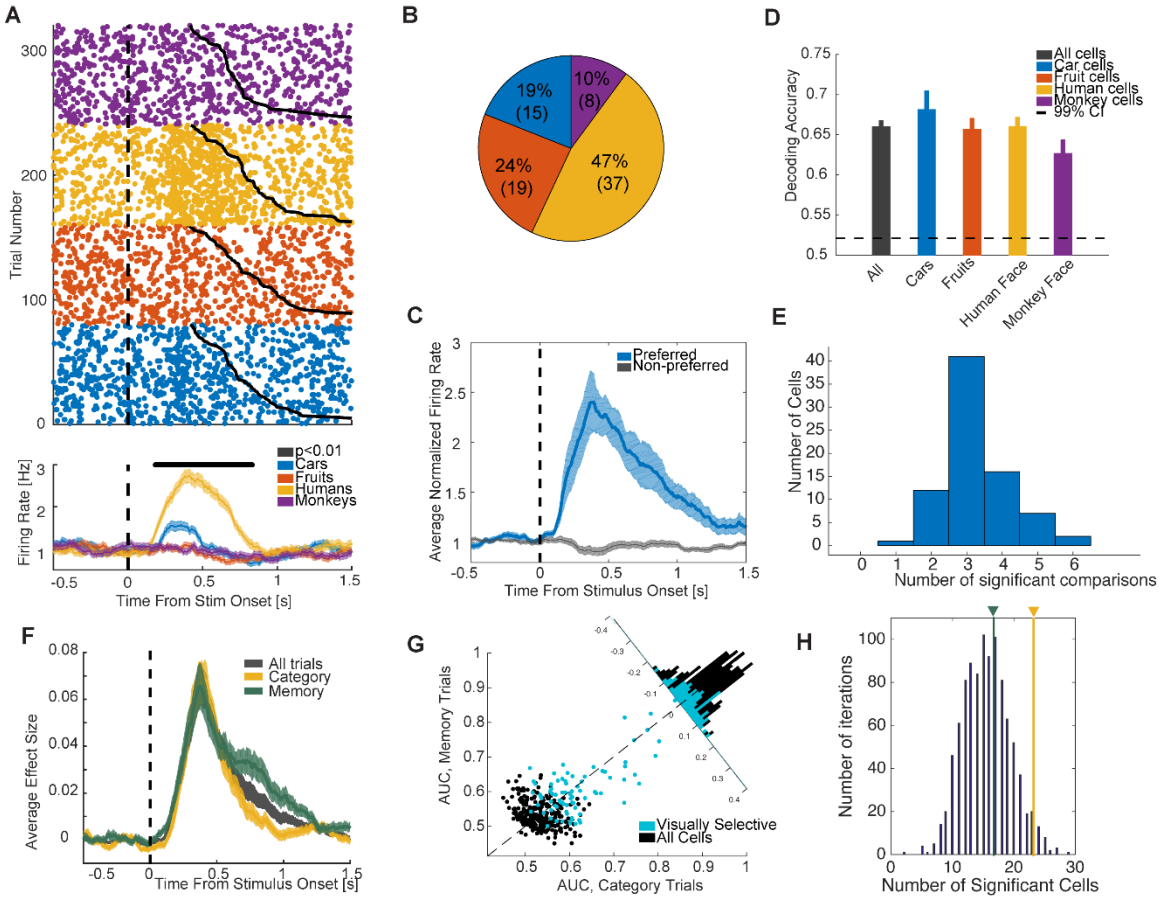


Figure 4.5: Visually selective cells in MTL are not sensitive to task demands
(a) Example cell, selective for images of human faces (left amygdala). **(b)** Proportion of tuned cells that are selective for each category with the largest one being for human faces. **(c)** Average normalized PSTH of the preferred category versus non-preferred categories across all visually selective cells (n = 79/360). Selection is done with a 1-way ANOVA (selection window is 100ms – 600ms after stimulus onset), and preferred category is based on maximum firing rate. **(d)**

Decoding results across each of the subpopulations that are tuned to the different image categories. Average decoding performance is above chance (99 % CI of the null distribution shown with the black dotted line) and qualitatively comparable across all categories. **(e)** Number of significant pairwise AUC comparisons (4 categories = 6 possible pairwise comparisons). **(f)** Average effect size for all visually selective cells, split by condition, shows no preference for task type. **(g)** Comparison of AUC values for category and memory trials shows no difference (p = 0.38, t-test). **(h)** Using the same selection criteria as that outlined for the choice cells, we find n = 17/360 cells (p=0.53) tuned for memory choice, and n = 23/360 (p=0.1) cells for category choice, much weaker than MFC cells.

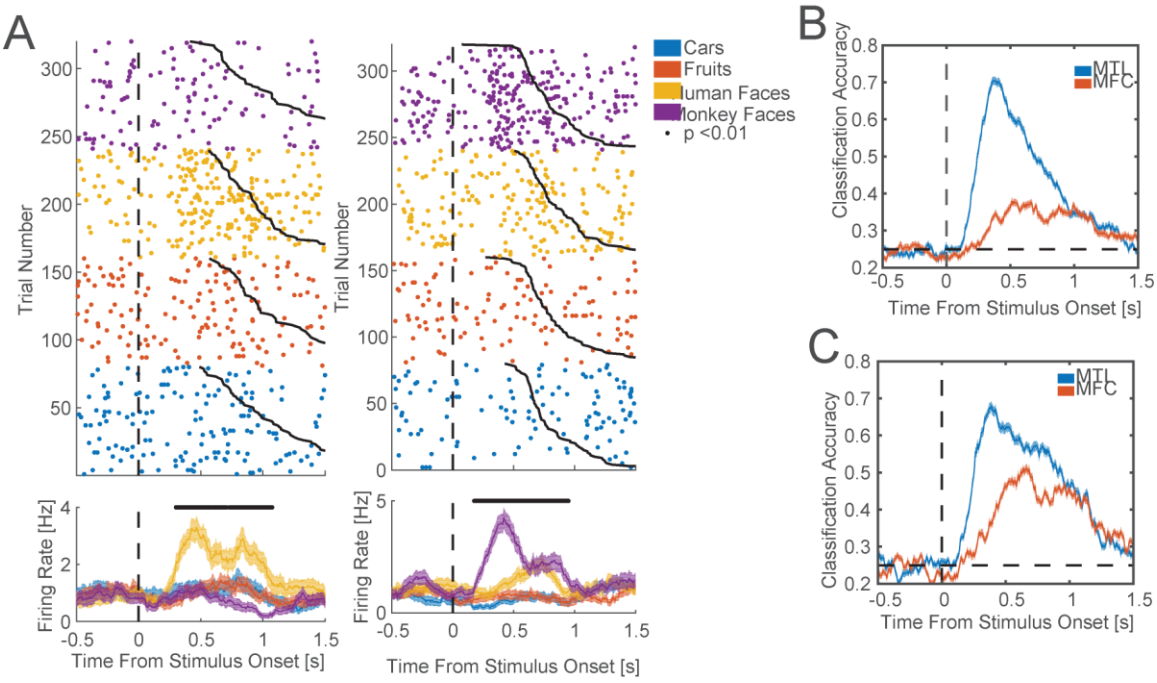


Figure 4.6: Visual information is stronger in MTL than it is in MFC
(a) Two additional example cells. Left is a cell tuned to faces of humans, right is a cell that responds selectively to monkey faces. **(b)** Pseudo-population decoding of image category (4 groups, 25% chance level). Category information is stronger and available earlier in MTL. **(c)** Same as (b) but computed only for memory trials. While there is more category information in MFC cells during the memory trials, MTL is still much stronger.

4.3.5 Category information in the MFC cells is modulated by task demands

We next tested to see if category information was also present in MFC cells. Selection for visually selective MFC cells was done using a 1s window centered at 700 after stimulus onset [200 – 1200ms]. Note that this is the same window that was used for the choice-selective cells in MFC. We isolated 62/399 visually selective cells using a 1x4 ANOVA with a threshold of $p<0.05$. Unlike the MTL cells, category tuning in the MFC was strongly modulated by task condition (see **Figure 4.7 A, B** for examples). When we computed Ω^2- effectize (**Figure 4.7 C**) for image category in the memory and categorization conditions separately, we found that it was signitifantly higher in the memory condition, both at the population level ($p = 1.5e-6$, paired t-test) and for the selected cells ($p = 1.7e-4$, paired t-test). When selecting for visually tuned cells indepednetly in the memory and categorization condition, we found a significantly larger proportion of cells in the memory condition (**Figure 4.7 D**, 60 vs. 29 cells; χ^2 – 12.15, $p = 4.9e-4$). This difference was also apparent at the population level, as revealed by the decoding results (**Figure 4.7 E**). For the decoding, we used the entire population of cells, not just the visually selective ones. The decoding method is the same as that outlined in the previous sections (see Methods).

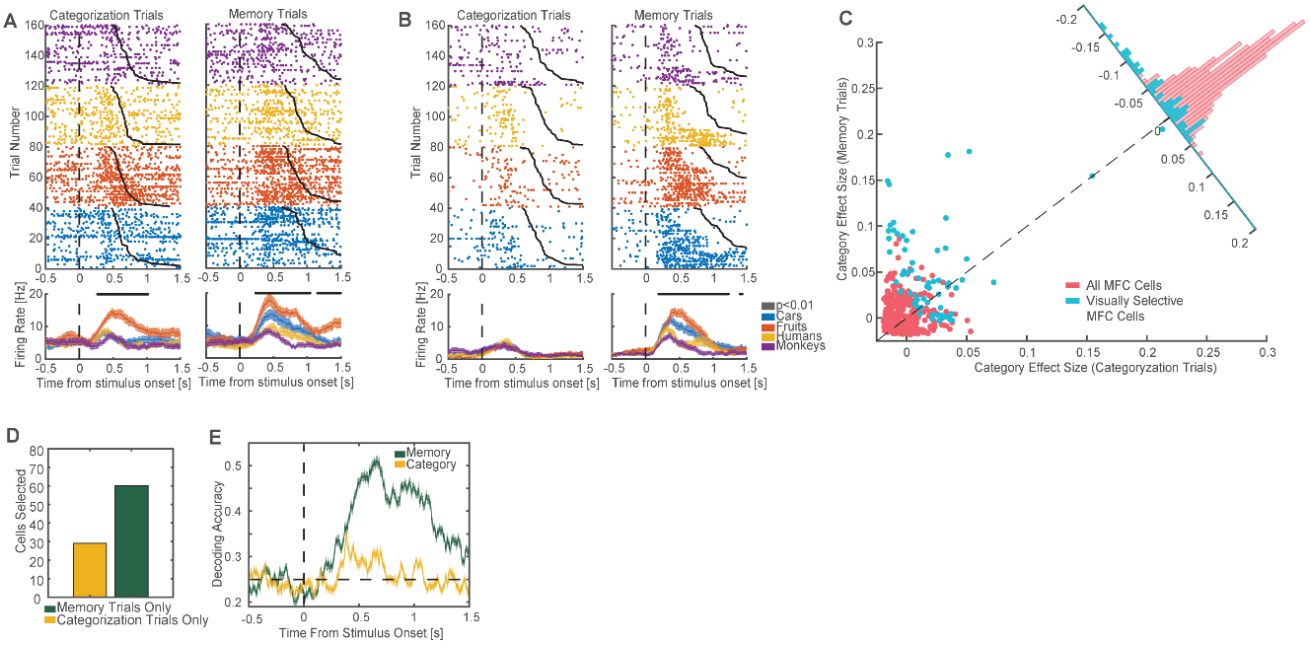


Figure 4.7: Category information in the MFC cells is modulated by task
(a) Example raster of a cell in the medial-frontal cortex. The raster plots are split by image category. The left panel shows all the trials from the visual categorization condition and the right panel shows all trials from the memory condition. **(b)** Same as (a). This particular cell is selected from the pool of cells that are both choice and category cells (example cell 1 from Figure 2). **(c)** Scatter of image category effect size of all MFC cells (pink) and MFC cells that are visually selective

(cyan). Cell selection was done using all trials. Effect size was then measured independently on memory trials and categorization trials. The effect size is greater during the memory condition. **d)** Selecting independently, using only categorization trials or memory trials yields a much larger number of cells in the memory condition versus the categorization condition. (e) Decoding accuracy across the entire pseudo-population of MFC cells is much higher during the memory condition than the categorization condition.

4.3.6 Choice cells are distinct from visually selective cells in the MFC

We have identified choice-selective and visually-selective cells in the medial frontal cortex. The next step is to determine if these two pieces of information are carried by the same set of cells or if they are distinct populations. To do so, we computed the Ω^2- effect size for choice (4 possible responses, yes/no $\otimes$ categorization\memory) as well as the Ω^2- effect size for image category (4 image categories). **Figure 4.8A** shows the effect sizes for choice (x-axis) and image category (y-axis) across the entire population (in pink), choice cells (in green), and image category cells (in blue). The populations were largely disjointed, with only 15 cells showing tuning for both image category and choice.

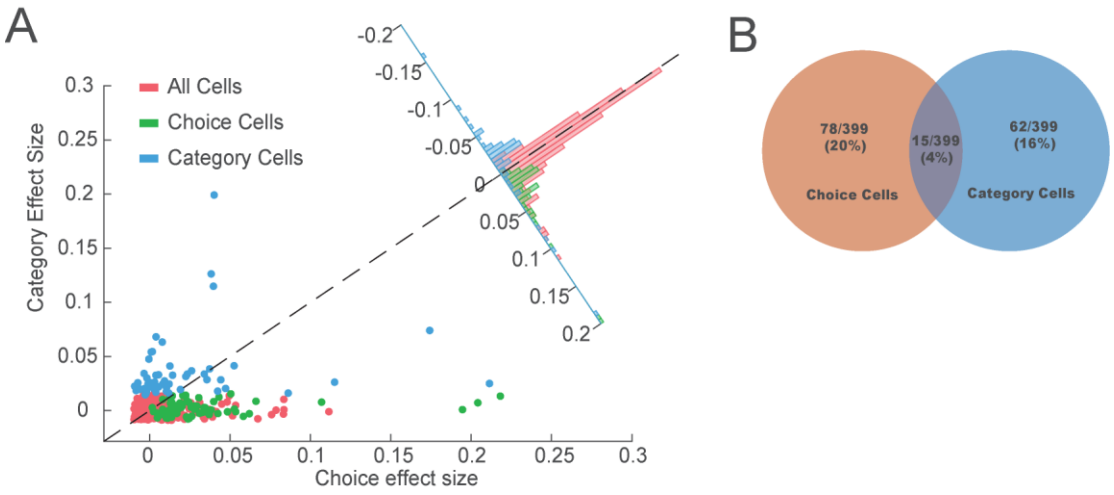


Figure 4.8: Choice cells are distinct from visually selective cells in the MFC. (a) Ω^2- effect size for image category and choice computed across all trial. **(b)** Proportion of cells selected as choice and as visually selective (here termed category cells) cells in the MFC.

4.3.7 Phase locking of hippocampal cells to local theta

Phase locking to theta is a prominent feature of both hippocampal and cortical neurons. In our recordings, as expected, hippocampal cells phase-locked to theta-band LFP activity (80/169

[47%] of hippocampal cells had a significant preference to theta-frequency LFP; See **Figure 4.9A-E**). This data is important for two reasons: (1) it shows that hippocampal theta is prominent in our task and that we are able to record it, and (2) it allows us to observe whether cells recorded in the MFC functionally interacted with the hippocampus and whether this interaction was modulated by task demands. Task modulation is not observed locally in the hippocampus, as shown in **Figure 4.9G-F**. **Figure 4.9F** shows that there is no difference in the spike triggered power of the theta oscillations while **Figure 4.9G** demonstrates that there is no difference in how well hippocampal cells cohere to theta as a function of task.

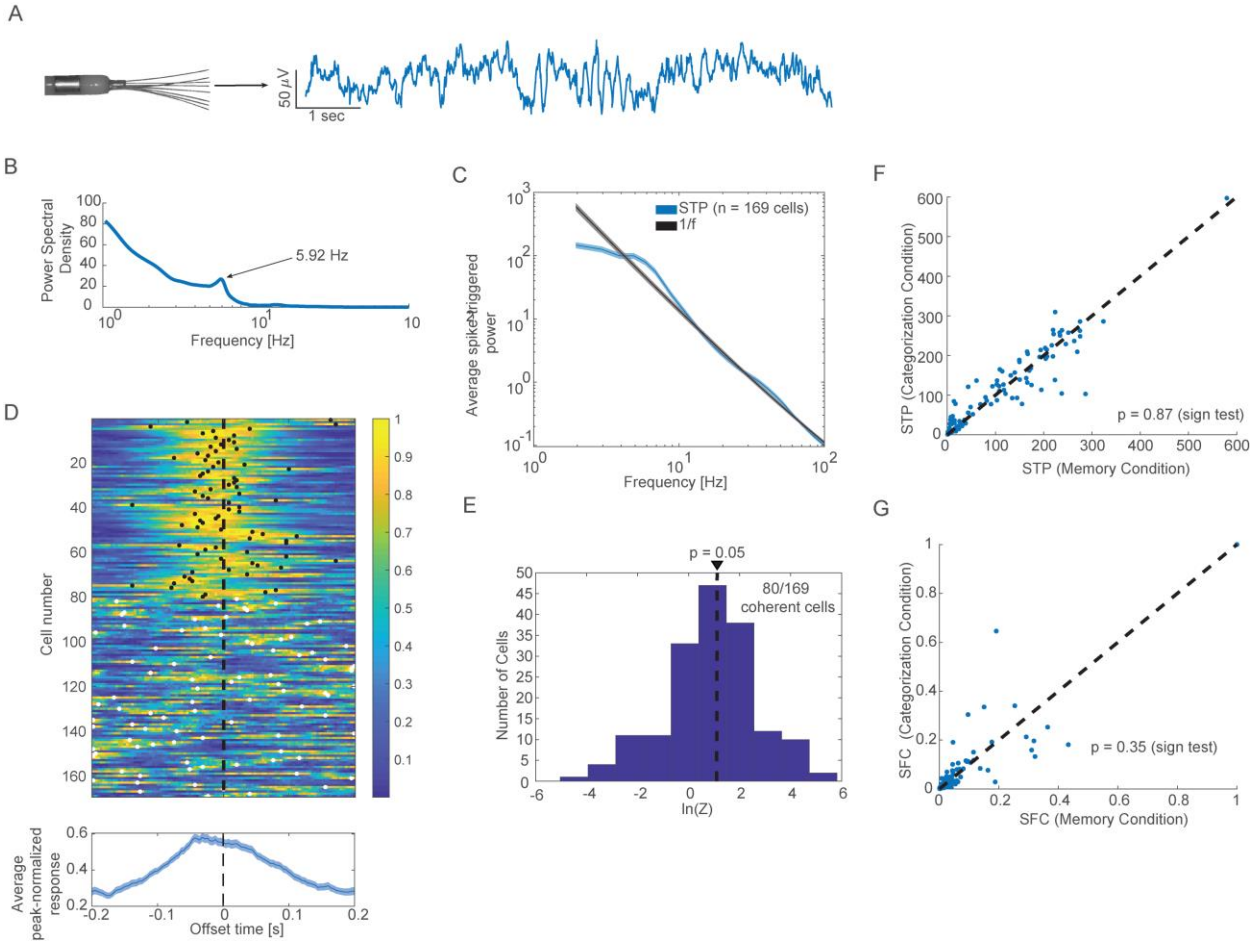


Figure 4.9: Hippocampal theta and phase locking of hippocampal cells. (a) Example LFP from a high-impedance microelectrodes placed in the hippocampus. (b) Spectra of the example trace shown in (a). (c) Average spike-triggered power (STP) across all n=169 hippocampal cells recorded. All spikes recorded were used for each cell. (d-e) Activity of hippocampal cells is phase-locked to hippocampal theta. (d) Rayleigh Z value as a function of time delay (-200 to 200 ms) for each cell for the 3-8Hz band. The peak coherence as a function of delay is indicated by the black dots for all cells that are significantly coherent, and with white dots for cells that are not significantly coherent. The bottom shows the average Z value across all cells. Z-values are peak normalized. (e) Histogram of Rayleigh's z-value for all cells (at

offset=0). Threshold for significance is set at p = 0 .05. Significance was determined using a null distribution computed by scrambling the timing of the spikes with respect to the phase of the theta oscillations (phase is estimated by computing the Hilbert transform of the bandpass-filtered LFP in the 3-8 Hz) in the local field potential recorded on the same electrode. (f) Spike-triggered power (averaged in the 3-8 Hz band) for all 169 cells, split by task condition. The two conditions were not significantly different (p=0.87, sign test). (g) Spike-field coherence for all cells, split by condition. The two conditions were not significantly different (p = 0.35, sign test).

4.3.8 Inter-area spike field coherence between MFC cells and MTL local-field potential is modulated by task demands

We next tested to see if the task effects were apparent in the inter-area coherence between MTL theta oscillations and MFC cells (see **Figure 4.10A** for an outline of the measurement we performed). For numerous methodological reasons, we decided to measure spike-field coherence with respect to the high impedance contacts and not the clinical contacts (see *LFP-processing pipeline* in Methods). We looked for coherence specifically in the 3-8Hz band of the hippocampal LFP (**Figure 4.10b**). Overall, MFC cells showed strong phase locking to hippocampal theta (54/399 cells selected, FDR corrected across number of channels, see **Figure 4.10B** bottom panel for an example). This proportion changed drastically however as we shifted the alignment of the spikes with respect to the hippocampal theta (**Figure 4.10C**, left panel) with the peak selection occurring -150ms. A negative shift in this case means that MFC cells cohered best to the *past* of the local field potential in the hippocampus. This result is particularly striking when we compare the selection numbers as a function of delay time for the local field potential measured in the MFC itself (**Figure 4.10C**, right panel). In the latter case, the selection numbers are highest at an offset of 0, which is what we would expect. The numbers of cells that are significantly coherence is also larger for the inter-area case than for the local case (103 cells at -150ms vs. 77 at 0ms). This finding has been documented in the rat as well (Siapas, Lubenov et al. 2005). Lastly, if we look at selection over time as a function of task condition, clear differences arise. Comparing the strength of coherence (Rayleigh z-value) across the two conditions reveals a significant difference at t = -150ms (**Figure 4.10D,** p = 0.003, t-test**).**

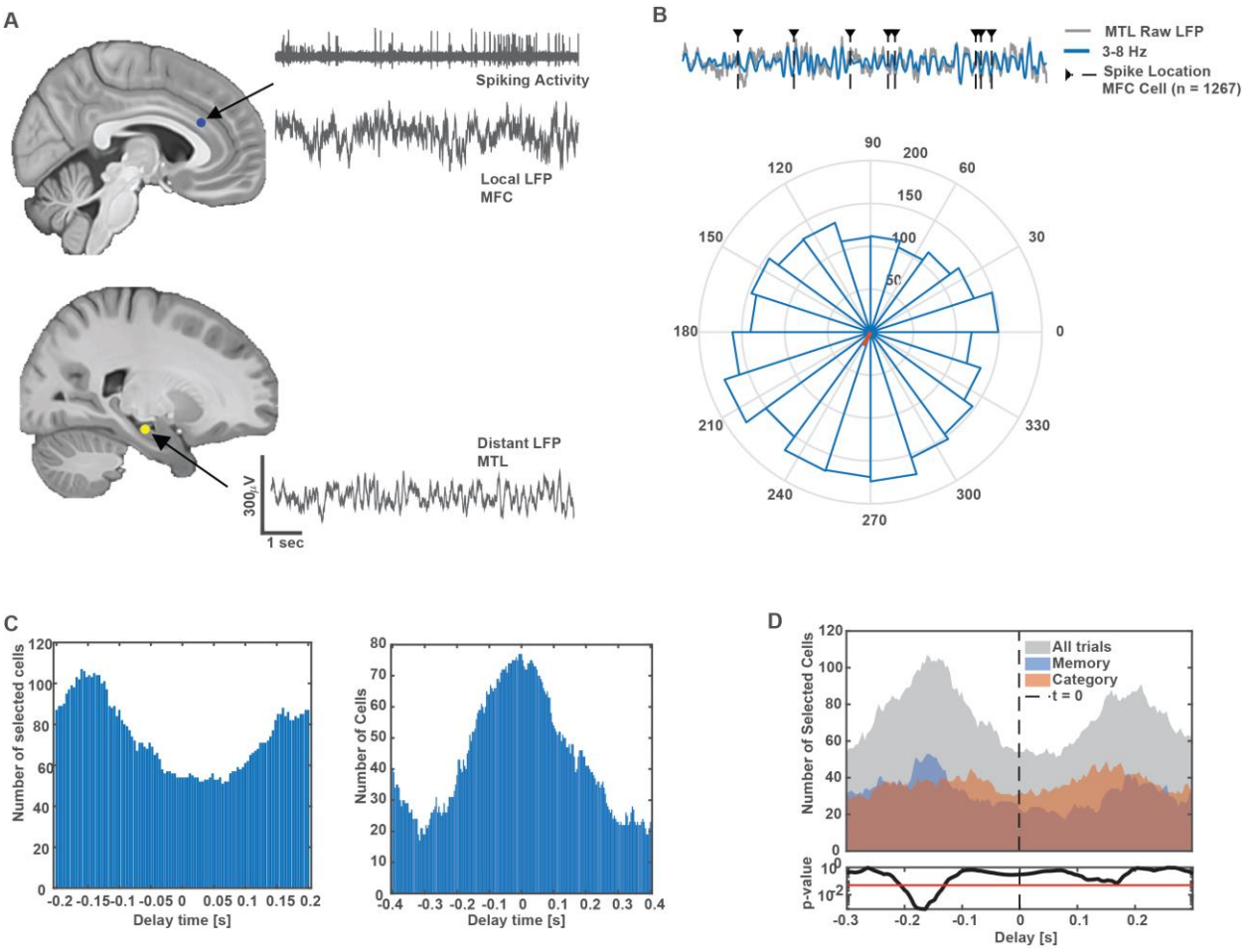


Figure 4.10: MFC cell coherence with MTL theta is modulated by task
(a) Outline of the inter-area spike-field coherence measurement between spikes recorded in MFC and local-field potential recorded in the hippocampus. **(b) Top:** Example aligned trace and spikes. Shown in light gray is the raw, unfiltered trace (apart from antialiasing filter prior to down sampling to 250 Hz) and the 3-8 Hz filtered blue trace. Superimposed on top of the raw and filtered trace, are the spike locations. **Bottom:** A polar histogram of the extracted phase from the filtered trace (blue) for every spike recorded from this cell. The mean resultant vector is shown in red. The p-value from the Rayleigh test for this cell is p = 1e-12, shown very strong phase preference. **(c)** Selecting MFC cells as a function of delay in both inter-area coherence with hippocampal LFP (left) and within area coherence (right). **(d)** Number of tuned cells as a function of delay time, shown for all spikes (gray area), memory condition (blue area), and categorization condition (red area). The trace at the bottom shows the p-value from a t-test comparing Rayleigh's z-value for selected cells, during the two task conditions.

4.4 Discussion

We show that human MFC cells signal abstract choices (decisions), with most doing so selectively for either categorization or recognition-memory based choices. Like in rodents, spike timing of human MFC cells was coordinated with hippocampal theta-frequency LFPs. Here, we demonstrate that such inter-areal coordination is increased selectively for decisions relying on

declarative memory. Importantly, such modulation occurred following verbal instruction alone in the presence of otherwise identical sensory inputs and motor outputs.

This shows that, in humans, MFC- hippocampal coordination can be engaged flexibly and selectively when needed, making this mechanism suitable for inter-areal information routing of memories. In contrast, the activity and phase-locking of MTL neurons was not modulated by task demands and was not related to choices, demonstrating specificity. Together, our work suggests that the human MFC plays a critical role in integrating memory-based internal variables to make decisions.

In primates, the strongest efferent connections from the hippocampus to the MFC are formed by the anterior hippocampus, which is the part of the hippocampus we recorded from in humans. In contrast, the large majority of work on MFC-hippocampal interactions in rodents has examined the dorsal hippocampus, which is thought to correspond to the posterior hippocampus. Future work will be needed to determine whether posterior hippocampal neurons in humans also engage with MFC in a similar manner than the anterior neurons examined here.

Our findings are of relevance also to the pathophysiology of psychiatric diseases. While it has long been appreciated that impaired frontal-hippocampal synchronization in schizophrenia has devastating consequences, the underlying reasons remain unclear. Mouse models (Sigurdsson, Stark et al. 2010) have provided much insight, but limitations in modeling human behaviors and significant evolutionary differences with unclear homologies in frontal cortex make it necessary to investigate this question directly in humans. Here, we were able to do so by observing in behaving humans action potentials of isolated neurons, a necessity to test the hypothesis advanced. We make two contributions towards this goal: coordination of spike timing relative to hippocampal theta-frequency oscillations is a mechanism of inter-areal communication in humans and this mechanism is engaged selectively whenever a decision requires access to declarative memories. This mechanistic insight indicates that selective on-demand restoration of inter-areal coordination by electrical stimulation should be investigated as potential treatment approach.

4.5 Conclusions

In our data, we have observed several components of a decision process that relies on internal-representations (recognition memory and image category). First, we observe two distinct populations of cells that communicated the subject's choice: one population specializes in

memory-based decisions (“Have you seen this image before?”) and another in visual categorization (“Is this an image of a human face?”). Secondly, we observed that while representation of the stimulus category is not dependent on task in the MTL, it is strongly modulated by task in the MFC. Thirdly, we observe that spike field coherence between MFC cells and MTL cells shows task-specific modulation of communication between the two areas. The task modulation is specific to the inter-area coherence, and cannot be observed locally when we measure spike-field coherence of MFC cells with local-field potentials recorded on the same electrode or when we compute SFC with hippocampal LFP using spikes from another brain area. This data shows for the first time in primates (to the best of our knowledge) that MFC-hippocampal phase locking is selectively engaged based on task demands.

Methods

Human electrophysiology

Human subjects were patients being evaluated for surgical treatment of drug-resistant epilepsy that provided informed consent and volunteered for this study. Monocular gaze position was monitored at 500Hz (EyeLink 1000, SR Research). The institutional review boards of Cedars-Sinai Medical Center and the California Institute of Technology approved all protocols. We recorded bilaterally from the amygdala and hippocampus in MTL and ACC and pre-SMA in MFC using microwires embedded in macroelectrodes. From each microwire, we recorded the broadband 0.1-9000Hz continuous extracellular signal with a sampling rate of 32kHz (Neuralynx Inc). We tested two different referencing techniques: (1) one microwire on each macroelectrode served as a local reference (bi-polar recording), (2) global referencing for all ipsilateral electrodes to a contact in pre-SMA. Given that all of our local-field potential analysis was carried out on virtual contacts (i.e. pairwise differences between electrodes), it did not matter much whether we used local or global references during recording. All included patients had clearly distinguishable spiking activity on at least one electrode in any of the MFC or MTL microwires.

Behavioral task

Subjects were given instructions on the task prior to the beginning of the experiment. They were told to follow the cues concerning effector (i.e. "eyes" or "hands") as well as what question they had to address (memory versus visual categorization). If they were cued to respond by hand, they registered their response using a low-latency response pad manufactured by Cedrus Inc. In the eye condition, they registered their response by saccading to the target on the left or right of the image. During the task, "yes" and "no" responses were always associated with the same side (i.e. no response remapping). The experiment was organized in 8 blocks, with each block containing 40 images (always 20 new and 20 old, expect for the very first block). The subjects always started with a categorization block. The response type (i.e. eyes or hands), was selected randomly for each block, and was changed only once halfway through the block. In between image presentations, subjects were instructed to look at the fixation cross in the center of the screen. We eliminated all trials where the subject was looking outside of the boundaries of the image or where we temporarily lost tracking of the eyes (<0.5% of trials).

Analysis of behavior

For the group analysis of behavior, we used mixed-effects models of the form $y = X\beta + Zb + \varepsilon$, where y is the response, X is the fixed-effects design matrix, β is the fixed-effects coefficients, Z is the random-effects design matrix, b is the random-effects coefficients, and ε is the error vector. In all analysis, we used a random intercept model with a fixed slope. The grouping variable for the random-effects was the session ID. The reported p-values in the main text correspond to the fixed-intercept for the relevant regressor. In the case of measuring the effect of number of expositions (i.e. number of times an image was seen) on the subject's accuracy during the memory trials, we used a mixed-effects logistic regression with the independent variable as an ordinal-valued whole number ranging from 1-7. The response was a logical value indicating success or failure on each memory question. Prior to running any analysis of reaction time data, we excluded outliers from the

distribution using the following procedure: a sample was considered an outlier if it was outside the 99^{th} percentile of the empirical distribution.

Spike sorting and single-neuron analysis

The raw signal was filtered with a zero-phase lag filter in the 300-3000Hz band and spikes were detected and sorted using a semi-automated template-matching algorithm (Rutishauser, Schuman et al. 2006, Rutishauser, Cerf et al. 2014). All PSTH diagrams were computed using a 250ms window with a step-size of 7.8ms. No smoothing was applied.

MRI processing and localization

Electrode localization was performed based on post-operative MRI scans that were performed immediately after implantation of the electrodes. These scans were registered to pre-operative MRI scans using Freesurfer's mri_robust_register (Reuter, Rosas et al. 2010) to allow accurate and subject-specific localization. To summarize electrode positions and to provide across-study comparability, we also aligned the pre-operative scan to the MNI152-aligned CIT168 template brain (Tyszka and Pauli 2016) using a concatenation of an affine transformation followed by a symmetric image normalization (SyN) diffeomorphic transform (Avants, Duda et al. 2008). This procedure provided the MNI coordinates that are reported here for every recording location.

Processing of human eye-tracking data

Calibration was performed using the built-in 9-point calibration grid and was only used if validation resulted in a measurement error of <1 dva (average validation error was 0.7 dva). We used the Eyelink systems which automatically annotate fixations and saccades from the continuous stream of data using a motion, velocity, and acceleration threshold (default thresholds).

Selection of visually selective units

Visually selective units in the MTL (amygdala and hippocampus combined) were selected using a fixed analysis window, 500ms in size, centered at 350ms after the stimulus onset (i.e. 100-600ms post stimulus onset). Cells were initially screened using a 1x4 ANOVA with a $p<0.05$ threshold. The preferred category of the cells that passed the ANOVA screening was defined as the category of images that had the highest mean firing rate across all trials (80 trials per category). All trials were used for the selection process.

Selection of choice cells

Choice cells were selected using a regression model applied to the firing rate in a 1s size window starting 200ms after stimulus onset. We fit the following regression model: spike count ~ $1 + \beta_1 \bullet \text{Response} + \beta_2 \bullet \text{RT}$, where the response is binary (yes or no) and RT is is the reaction time. We

fit this model separately to trials in the memory-and categorization condition. RT was included as a regressor to control for RT differences between the two possible responses. A cell qualified as a choice cell if the t-statistic of the β_1 term was significant at $p<0.05$ for at least one of the two task conditions. The response preference of significant cells for either yes or no was determined based on the sign of β_1 (positive = yes, negative = no). All trials regardless of whether the answer was correct or incorrect were used for the selection process. To estimate the significance of the number of selected cells, we generate a null distribution by repeating above selection process 1000 times after randomly re-shuffling the response label (keeping RT intact). We estimated this null distribution separately for choice cells in for the memory-and categorization condition and used each to estimate the significance of the number of selected cells of each type.

Choice cell selection bootstraps

To generate a null distribution for the number of selected cells, we repeated our selection process (see *Selection of choice cells*) 1000 times, each time shuffling the response label in the regression, while preserving the response time structure that the cells had. For both types of choice cells, we estimated likelihood of occurrence by comparing the true value with the null distribution.

Single-cell decoding

Single cell decoding (see Figure 4.3D) was done over the same window as that used for selection. In the case of the visually-selective cells, we counted spikes between 100-600ms after the stimulus onset. Category decoding was done using the naïve-bayes poisson decoder from the neural-decoding toolbox (Meyers 2013). The image category labels were converted to a binary label, indicating preferred versus non-preferred (chance level, 0.5). Since the number of examples from the preferred group is ¼ the size of the number of examples from non-preferred (1 out of 4 categories), the priors for the naïve-bayes decoder were not determined empirically but instead forced to be uniform.

Population-level decoding

Population decoding was based on a pseudo-population using an approach first described by Meyers et al. (Meyers, Freedman et al. 2008) Briefly, we started the procedure by randomly selecting 75% available cells in MTL and MFC (corresponds to 274 and 299 cells respectively). For each cell, we selected a subset of the available trials to be used for decoding. For category decoding, we used 50/80 instances (320 trials total) from each of the four categories. For choice decoding, the number of instances in each of the four conditions, response (yes/no) $\otimes$ condition (memory\categorization), were varied from subject to subject. We used 30 trials for each of the 4 groups because it was a number across all subjects. Prior to concatenating the responses across the pseudo-population, we normalized the firing rates for each cell individually, using only the trials contained within the training set. The spikes were counted using bins of 250ms in width. Using the randomly selected subset of cells and trials, we performed 10-fold cross validation decoding using a linear support vector machine (SVM) algorithm. Finally, the entire procedure was repeated 50 times with different subsets of cells and trials in order to produce a smooth estimate of classification accuracy.

Selection of coherent cells

In order to determine whether a cell was phase-locked to the local or distal field potential in the theta band (3-8 Hz), we used the Rayleigh statistic, corrected for multiple comparisons (Bonferroni, 8 electrodes). Phase estimation for each of the 8 ipsilateral electrodes was done by first bandpass-filtering in the 3-8 Hz band and then computing the instantaneous phase of the signal with the Hilbert transform. To compute coherence, we used all of the spikes recorded from a given cell throughout the task. We repeated this same procedure for several different comparisons: (1) spiking activity in MFC cells and oscillations recorded locally (ex. measure coherence between spiking activity of an SMA cell with respect to the 8 ipsilateral electrodes in SMA, one of which also recorded the spiking activity), (2) spiking activity in MFC and distal oscillations recorded in hippocampus, (3) spiking activity recorded in amygdala and oscillations recorded in hippocampus, and finally (4) spiking activity recorded in hippocampus and oscillations recorded locally (same as (1) but with LFP and spiking activity recorded in hippocampus).

LFP-processing pipeline

On our electrodes, we have the option of measuring local-field potentials on the low impedance (~10 kΩ), clinical contacts of the Ad-Tech electrode, or on the finer, high impedance (~500 kΩ) contacts. Given the lower impedance of the electrode, the first option has a much larger "listening sphere" and would reflect more of an intracranial-EEG type signal. The second option is to use the microwires, the same electrodes that record the spiking activity. Depending on the application, it is sometimes better to use one versus the other, but for the purposes of our study, we determined that using microwires is better or the following reasons: (1) activity measured on these electrodes is much closer to the canonical definition of LFP because the activity as a function of distance from electrode is attenuated much more aggressively; (2) given that we have 8 electrodes, we have 8 samples, and therefore a much better estimate of any underlying oscillations, whereas with the macroelectrode we only have one; and (3) pairwise subtractions can remove large unwanted artifacts (related to movement for example). In order to get a much more robust measurement of the local-field potential, we computed all 28 pairwise differences between the 8 microelectrodes. This was done after first high-pass filtering and normalizing (z-score) each trace, in order to remove low frequency artifacts and to match the amplitude across pairs prior to subtraction. We performed significance testing using two methods: (1) Rayleigh z-test on the phase estimate from each extracted snippet (the test assumes that the data is sampled from a von Mises distribution and that it is unimodal) and (2) we constructed a null distribution by extracting the phase at *n* instances (where n is the number of spikes emitted by a cell, in a given condition), selected randomly along the entire recording session. The second option controls for any non-uniformity that may be present in distribution of phase during the entire recording session. This procedure was repeated 1000 times in order to generate a null distribution individually for 2 metrics, (1) pairwise phase consistency (PPC), and (2) Rayleigh's z-value. Our sessions were composed of a mix of locally referenced and globally referenced electrodes; therefore we only performed this procedure on the globally referenced sessions.

References

Avants, B., et al. (2008). "Multivariate analysis of structural and diffusion imaging in traumatic brain injury." Academic radiology **15**(11): 1360-1375.

Euston, D. R., et al. (2012). "The role of medial prefrontal cortex in memory and decision making." Neuron **76**(6): 1057-1070.

Gold, J. I. and M. N. Shadlen (2007). "The neural basis of decision making." Annu. Rev. Neurosci. **30**: 535-574.

Jeneson, A. and L. R. Squire (2012). "Working memory, long-term memory, and medial temporal lobe function." Learning & Memory **19**(1): 15-25.

Meyers, E. M. (2013). "The neural decoding toolbox." Frontiers in neuroinformatics **7**.

Meyers, E. M., et al. (2008). "Dynamic population coding of category information in inferior temporal and prefrontal cortex." Journal of neurophysiology **100**(3): 1407-1419.

Minxha, J., et al. (2017). "Fixations gate species-specific responses to free viewing of faces in the human and macaque amygdala." Cell reports **18**(4): 878-891.

Mukamel, R., et al. (2010). "Single-neuron responses in humans during execution and observation of actions." Current biology **20**(8): 750-756.

Oostenveld, R., et al. (2011). "FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data." Computational intelligence and neuroscience **2011**: 1.

Reuter, M., et al. (2010). "Highly accurate inverse consistent registration: a robust approach." Neuroimage **53**(4): 1181-1196.

Rutishauser, U., et al. (2014). Data Analysis Techniques for Human Microwire Recordings: Spike detection and Sorting, Decoding, Relation between Neurons and Local Field Potential. Single Neuron Studies of the Human Brain. I. Fried, U. Rutishauser, M. Cerf and G. Kreiman. Boston, MIT Press**:** 59-98.

Rutishauser, U., et al. (2006). "Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, in vivo." J Neurosci Methods **154**(1-2): 204-224.

Rutishauser, U., et al. (2015). "Representation of retrieval confidence by single neurons in the human medial temporal lobe." Nature neuroscience **18**(7): 1041-1050.

Shadlen, M. N. and D. Shohamy (2016). "Decision making and sequential sampling from memory." Neuron **90**(5): 927-939.

Siapas, A. G., et al. (2005). "Prefrontal phase locking to hippocampal theta oscillations." Neuron **46**(1): 141-151.

Sigurdsson, T., et al. (2010). "Impaired hippocampal–prefrontal synchrony in a genetic mouse model of schizophrenia." Nature **464**(7289): 763-767.

Squire, L. R. (1992). "Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans." Psychological review **99**(2): 195.

Stern, C. E. and M. E. Hasselmo (2009). Recognition Memory A2 - Squire, Larry R. Encyclopedia of Neuroscience. Oxford, Academic Press**:** 49-54.

The capacity to accurately recognize an item as having been previously encountered is known as recognition memory. In this article, the concept of recognition memory is discussed, including the idea that recognition memory is composed of two main components: recollection and familiarity. Neuroscience research examining recognition memory in humans and in animal models is described, besides the computational models of recognition.

Swanson, L. (1981). "A direct projection from Ammon's horn to prefrontal cortex in the rat." Brain research **217**(1): 150-154.

Tyszka, J. M. and W. M. Pauli (2016). "In vivo delineation of subdivisions of the human amygdaloid complex in a high-resolution group template." Human brain mapping **37**(11): 3979-3998.

Wunderlich, K., et al. (2009). "Neural computations underlying action-based decision making in the human brain." Proceedings of the National Academy of Sciences **106**(40): 17199-17204.

Chapter V: General Discussion and Future Directions

In this chapter, I will briefly summarize the key results that we have shown in the thesis. I will also try to expand a little on the main topics that have been introduced so far in this thesis in the remaining sections of the discussion. These topics range from the task-dependence of eye movements and aligning neural data with fixations, to interpretations of the face response in the amygdala. Throughout, I will present some data which did not necessarily fit in elsewhere in the thesis but is relevant to what is being discussed.

5.1 Summary of results

One of the key findings in this thesis is related to the nature of the visual response in the amygdala and how this response is gated by fixations in a crowded visual scene. The key differences between our task and that used previously to study visual responses of this sort are: (1) the subjects can select the target they want to fixate on freely and (2) there is competition for attention. Consistent with our hypothesis, we found that amygdala cells responded to individual fixations and did not simply produce just a coarse response to the preference of faces in the scene. We also found that behavioral preferences (as measured by the way the subjects sampled the objects in the array) corresponded nicely with the neural representations that we found. Specifically, measures of eye movements, such as first fixation, showed that subjects preferred to look first at images of conspecifics. In turn, most of the visually selective cells we recorded from were tuned to images of conspecifics. In the section below, we talk about a different kind of preference, novelty, for which there is also strong representation in amygdala and more generally the medial-temporal lobe. In short, the fact that neurons in the amygdala responded to faces only when they were fixated, suggests that their activity was gated by visual attention. Our experiments in humans utilizing covert attention confirm this hypothesis.

Our second set of results looks jointly at representations in the medial temporal lobe and the medial frontal cortex. We find that representations of image category are stronger in the medial temporal lobe, and as expected, they are independent of what the task is. Medial frontal cells on the other hand also represented category information but in a task-dependent way, namely there was more information related to image category during trials where the subject had to make a memory-guided decision. Our key finding is the representation of choice in MFC cells, with distinct cells representing choice in either image categorization trials or recognition trials. We qualified this signal as a “choice”

signal, because it was not dependent on the action that the subject took (i.e. saccade or button press) but simply reflected the abstract notion of "no" or "yes". Our hypothesis was that to make such memory-guided choices, cells in MFC would need to access representations of memory that are available in the medial temporal lobe. To test our hypothesis, we measured the coherence between the MFC and MTL using metrics that measure spike-field interactions. We showed that MFC cells cohere to hippocampal theta more strongly during the memory trials than the visual categorization trials. In the discussion below, I will explore some possible extensions of this work that we are now taking into consideration.

5.2 The importance of tracking eye-movements

In Chapter III, we presented fixation-gated responses in amygdala neurons from monkeys and humans. This data underscores the importance of tracking eye-movements, but it also reveals how incredibly rich eye-movement data can be, even in a simple task. In a single session of the array task for example, a human subject may generate almost 1000 fixations and just as many saccades. A monkey can generate even more (inter-fixation time was shorter in monkeys). Understanding the sources of variance in the eye-movement data is of paramount importance, because it can reveal interesting dynamics that are not apparent on longer time-scales (over the time course of a single trial for example) just as easily as it can lead to artifacts and wrong conclusions if not properly controlled. These sources of variance can range from the trivial, such as level of crowdedness in the visual scene and task instruction, all the way to the individual preferences of the participant. **Figure 5.1** shows a trivial source of variance, where the level of crowdedness can change the nature of the eye-movement features. Here we focus on fixation duration and saccade duration, but there are many others features one could extract, such as blinks and pupil area. What becomes immediately apparent is that the fixation duration (**Figure 5.1C, D**) in the non-crowded case is longer than that during the crowded case (0.31s vs. 0.29s respectively, $p = 1.6e\text{-}7$, 2-sample KS-test). Furthermore, the difference between the saccade duration histograms for the two conditions (**Figure 5.1E**) reflects the different task structure, even though the actual images and their possible locations were the same.

In addition to task structure, eye-movement features are also modulated by the preferences of the subject. In **Figure 5.2A**, we reproduce a result from Chapter III, namely that human subjects tend to fixate first on images of conspecifics (i.e. human faces). They do this so much more often than the next competing category of images, monkey faces (0.32 vs. 0.25, $p < 8e\text{-}04$, paired t-test). Although

not shown here, we can also quantify a subject's look preferences by their sampling sequence (see Figure 3.3), number of within-stimulus fixations, and look-time for each category (see Figure 3.1).

Subjects also have preferences for novelty **(Figure 5.2 B, C, D**). When measured across the population, number of visitations is significantly correlated with the amount of time spent looking at the image (mixed-effects model across subjects, with random-effects on the intercept, p<2e-21). Furthermore, this effect can be seen on the individual categories, cars, humans, and fruits, but not on monkeys, which, across all subjects, are the hardest image category to remember (**Figure, 5.2E**, as measured on the recognition portion the subjects performed at the end of the array task). Looking time is often used in animal experiments as a proxy for whether something is perceived as familiar or novel (visual preference looking task, (Murray and Mishkin 1998, Zola, Squire et al. 2000).

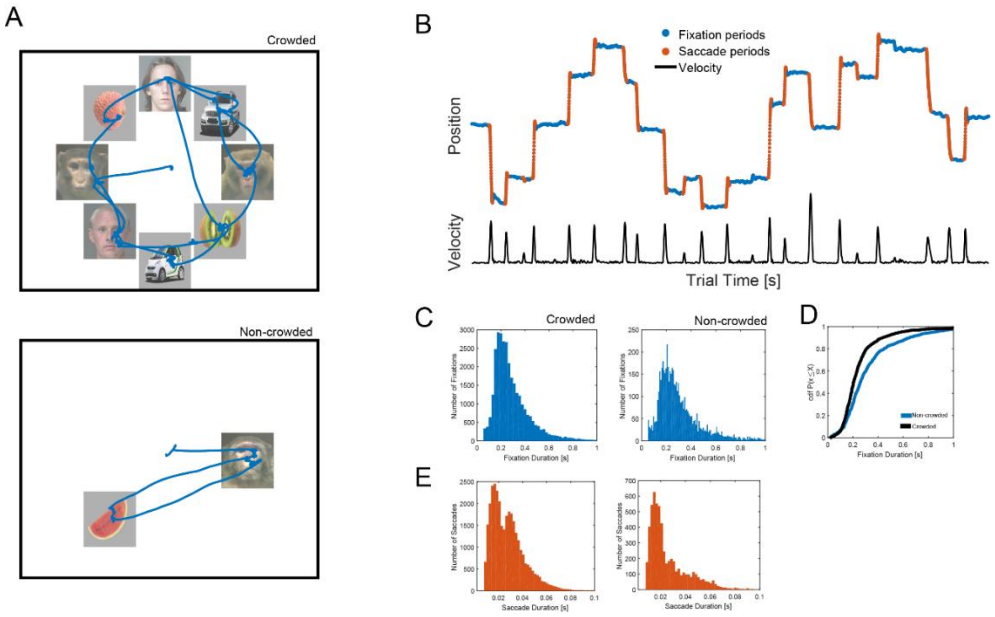


Figure 5.1: Labeling fixations and saccades from continuous eye-tracking data. (A) Example trial from the study described in Chapter III (top) and a less crowded version of the same task (2-3 images, bottom). Subjects are free to sample from any of the possible image locations for the duration of the trial. (B) The x-coordinate of the position of the subject's eye. Labeled in blue are fixation periods. Shown in red are periods when the subject was making a saccade. The black trace at the bottom shows the velocity of the eye movement. (C) Histogram of all fixation durations. (C) Histogram of all fixation durations in the crowded (left) and non-crowded (right) conditions. (D) For the same subjects, fixation duration was significantly longer in the non-crowded condition than the crowded one. (E) Saccade durations for the crowded (left) and non-crowded (right) conditions.

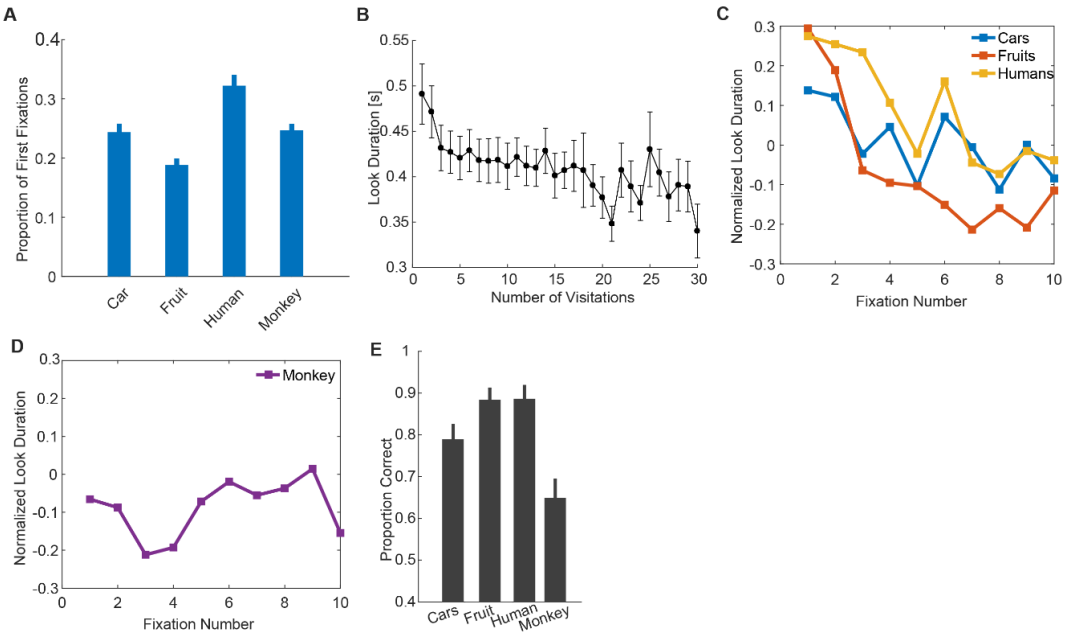


Figure 5.2: Preferences for certain image categories and for novel images are apparent in the subject's eye movements. (A) Proportion of first fixations on each of the four categories, computed over all recorded sessions. The preferred category is human faces, which is significantly more likely to be a target than the next best category, which is monkey faces ($p < 8e\text{-}04$, paired t-test). (B) Look duration as a function of the number of fixations that landed on a particular image (bars are s.e.m). (C) Same as (B) but shown separately for three of the categories. (D) Same as (C) and (B) but for monkey fixations. For (B), (C), (D), the look durations were first standardized within the subject prior to averaging across the whole population. (E) Proportion of category instances remembered in the recognition memory task that followed the free-viewing task.

Furthermore, just as the looking preference for faces in the human subjects was reflected in the population of visually selective cells (with most cells preferring human faces), the preference for novelty is also evident in cells recorded in the amygdala and hippocampus. **Figure 5.3** (panels A, B, C) shows 3 example cells that are modulated by the familiarity/novelty of the fixated stimulus, shortly after the onset of the fixation. Panel C shows an example of a cell that is both visually and novelty selective, firing preferentially for images of novel faces.

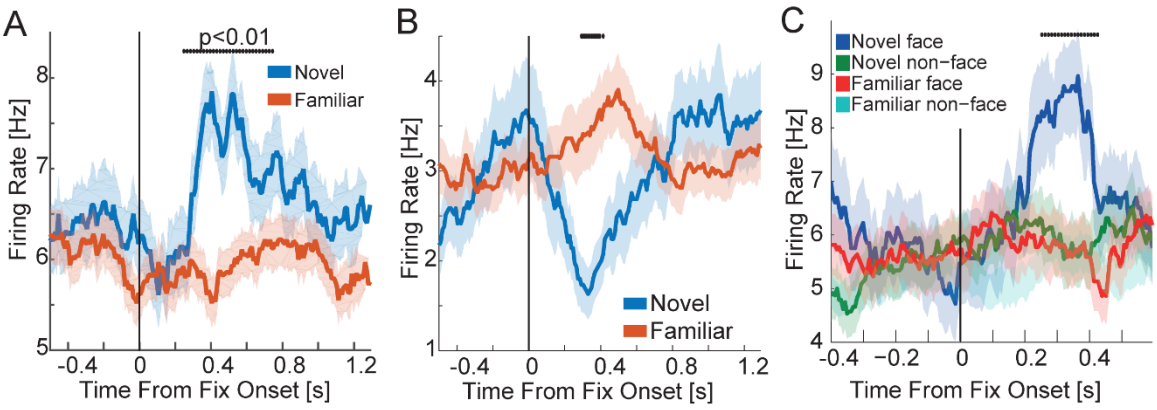


Figure 5.3: Fixation-aligned responses of cells in the human MTL encode the familiarity of the fixated item. (A,B) Example cells recorded in the human hippocampus (left) and amygdala (right) during the array task described in Chapter III. (C) A cell that is preferentially modulated by a novel face (recorded in amygdala).

5.3 Relating neural patterns with eye-movement data

The post-stimulus time histogram is one of the most commonly used tools in neuroscience in summarizing neural responses aligned with some event. It is an incredibly useful tool, but one that can easily lead to artifacts with fixation-aligned neural responses. In a free-viewing paradigm such as the one that we presented here, features of the eye-movements, such as fixation duration, are a random process that the experimenter cannot control. Since there are currently no better ways of summarizing the neural response however, selection of fixations used for analysis becomes very important. To illustrate this point, let us consider the response of a population of 129 visually selective amygdala cells recorded on the array task. **Figure 5.4** illustrates what the average response looks like for all selected cells when the subject fixated on the preferred image category (blue trace) and non-preferred image categories (black trace). In panel (A) of **Figure 5.4**, we can see what would happen if we were to include all fixations on the preferred category, independent of what came before and after. We notice that the baseline for fixations that land on preferred targets is actually lower than that landing on non-preferred targets. Furthermore, we notice a suppression of response for non-preferred fixations. Both of these features of the PSTH however, are artifacts that can easily be removed if we only consider instances that were preceded by fixations on non-preferred images. This problem can be exacerbated by highly stereotyped viewing preferences (imagine a cell that responds selectively to faces and a subject that tends to look at several faces on a row). The selection process might vary depending on the analysis that you want to do. **Figure 5.5** shows an example selection

process that was used for determining visually-selective cells in the array task. The selection criteria ranges from the most lax, which corresponds to a simple threshold on duration of fixation (or "look"), all the way to more strict criteria that only use the very first fixation on an image category.

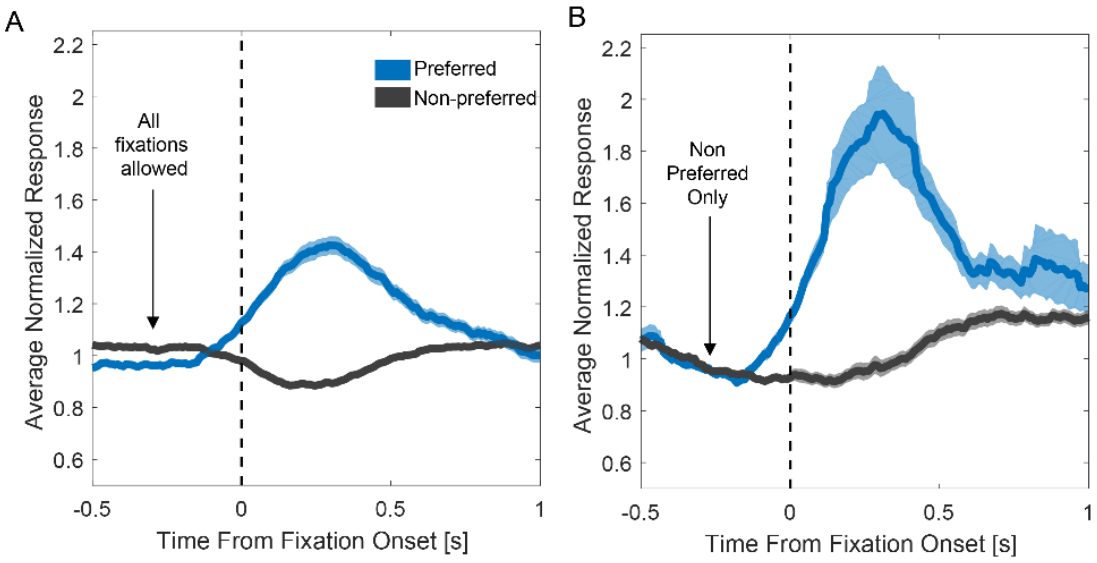


Figure 5.4: Averaging fixation-aligned neural responses can lead to artifacts. Fixation-aligned responses of 129 amygdala cells split by preferred (blue) and non-preferred (black) targets. (A) The average normalized response including all fixations on the preferred, independent of what preceded it. (B) Filtering the fixations and excluding all fixations that were preceded by another fixation on a preferred image category.

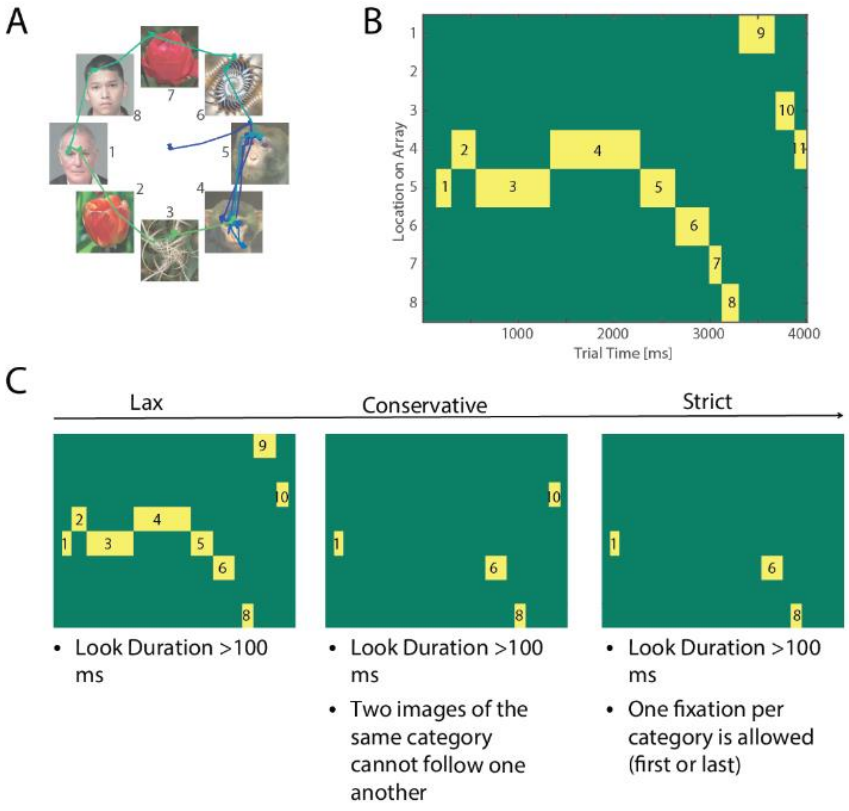


Figure 5.5: Subsampling sequences of fixations to prevent artifacts from averaging. (A) Example scan-path recorded from a monkey. Numbers indicate "location on array". (B) For trial shown in (A), summary of where the monkey looked. All successive fixations that fall on the same image are pooled into one "look" (Yellow patch). The looks are numbered 1-10 and the y-axis indicates the location on the array. (C) Different selection criteria for "looks" that can be included in the analysis. In the most lenient case, we can use all fixations that are longer than 100ms, and in the most stringent case, we can use only the first fixation for each category in addition to the duration requirement.

5.4 Covert spatial attention task with distractors

As an extension to the covert attention task that we present in Chapter III, we also collected data on a version of the task that included a distractor (**Figure 5.6A**) in addition to the target image. I have included this piece of data here because I think it highlights the role of attention in modulating these visual responses that we measured in the amygdala. The goal of this task was to highlight the effects of spatial attention on the responses in the amygdala by making the screen more crowded. Briefly, subjects were instructed to never break fixation from the cross in the center of the screen. They were first cued to the spatial location of the object for which they had to answer a yes or no question. The question was always about the visual category of the image, for example, "Is this an image of a human face?" We presented them with two images on opposite ends of the fixation cross (distance ~ 20^{o} of visual angle). Of the 120-amygdala cells that were recorded on this task, 22 of them showed

preferential tuning for one of the image categories (selected with 1x4 ANOVA and assigned to the image category that elicited highest firing rate). **Figure 5.6 B** shows the responses from two example cells split across the four possible conditions. The conditions are: (1) both the cued and the distractor were preferred (blue); (2) only the cued image was from the preferred category (red); (3) only the distractor was from the preferred category; and (4) neither the preferred image nor the distractor were from the preferred category.

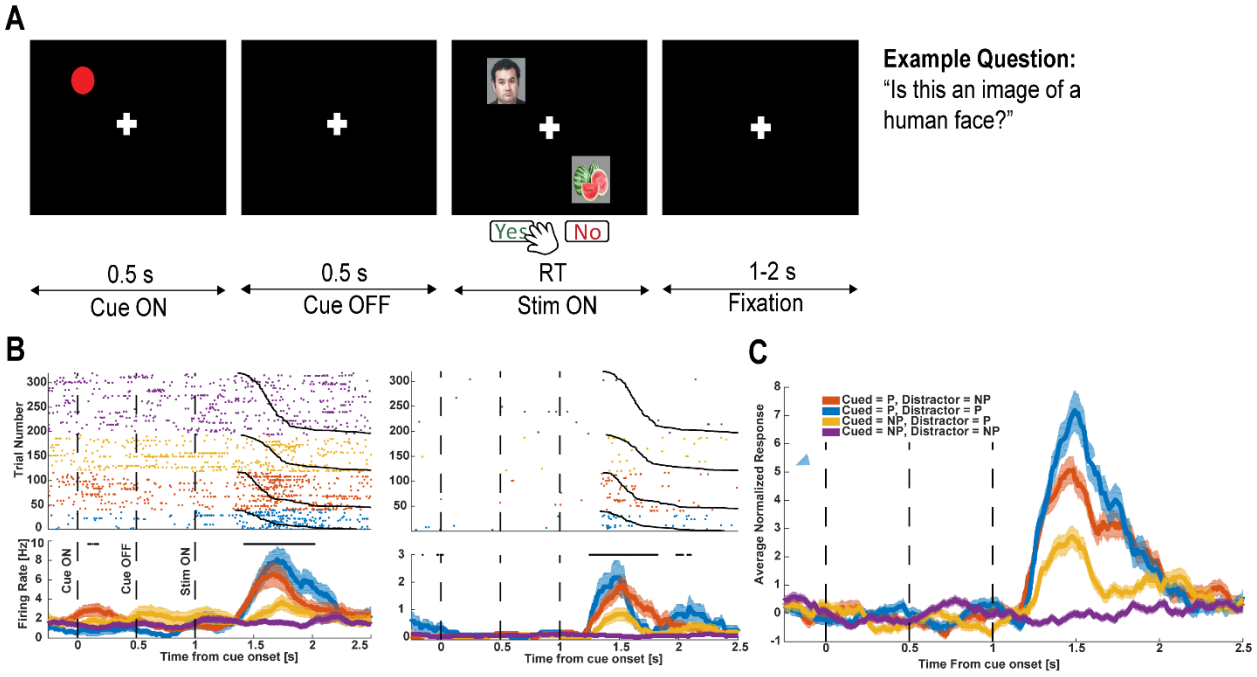


Figure 5.6: Covert attention task with distractors. (a) Task design for the covert attention task. Subjects were cued to the spatial location of the stimulus for which they had to answer a question (e.g. "Is this a human face?"). After a short delay, two objects were presented on opposite ends of the array. Subjects responded with a button press, while always maintaining fixation. **(b)** Raster for two example cells showing the response in the four task conditions. **(c)** Population average across 22/120 amygdala cells that showed preferential tuning for one of the four possible image categories.

5.5 On the role of face responses in the amygdala

It was clear from the data presented in Chapter III how prevalent the face response is in the amygdala. This does however bring up two key questions: (1) why are the latencies so long (~322ms, peak effect size, fixation aligned), and (2) what is this response being used for in the brain, and in what ways is it different from the cortical response to faces. Given the long latencies, it is unlikely that the face response in the amygdala is guiding moment-to-moment attention, or that it is necessary for a

perceptual process. One possibility is that face responses in the amygdala are actually facilitating learning for socially relevant stimuli (i.e. faces of conspecifics). Given the extensive connectivity between the amygdala and other brain areas, such as the hippocampus and the neocortex (Pessoa and Adolphs 2010), it has been suggested that the amygdala could provide a modulatory signal during memory encoding that is specific to a particular category of images (i.e. human faces, and perhaps specific identities) (Phelps 2004). We have made two observations in the population of amygdala cells that make this hypothesis very credible. The first is the novelty-dependent responses (see **Figure 5.3** for example responses) that we can directly measure during the array task. Of the 629 amygdala cells that we recorded, 59 of them (9%) show novelty-dependent responses. This response is present despite the absence of a clear emotional expression, since all of the face stimuli we used are neutral. Furthermore, 29% of all amygdala cells show significant (through Rayleigh z-test) phase locking to theta oscillations (3-8 Hz) recorded on the same electrode (see **Figure 5.5** for an example). Both the novelty response and the theta-phase locking are features of the amygdala cells that could be used to modulate learning about socially relevant stimuli. Given the population of cells that we have, the array task is a good candidate for testing this theory. Each frame of the array task contains two completely new stimuli, allowing us to see how these representations of familiarity are constructed fixation by fixation.

Another interesting hypothesis on the computational value of the face response in the amygdala is motivated by a recent observation on the nature of the face response in cells within the face-patch network. In their recent work (Chang and Tsao 2017) , Chang et al. show that the face response in the macaque face patches (specifically, middle lateral/middle fundus, and anterior medial) is actually a *measurement* that each cell performs along its preferred axis within the high dimensional face space. It is possible that the face responses in the amygdala could be doing a similar kind of measurement, but along axes that are more relevant for social behaviors. Recent evidence from single-unit recordings, for example, has shown that individual cells in the human amygdala parametrically encode the intensity of specific facial emotions (Wang, Yu et al. 2017). Not all cells might perform such a clearly interpretable measurement, but the key is that these measurements are made specifically for computing something in the social domain.

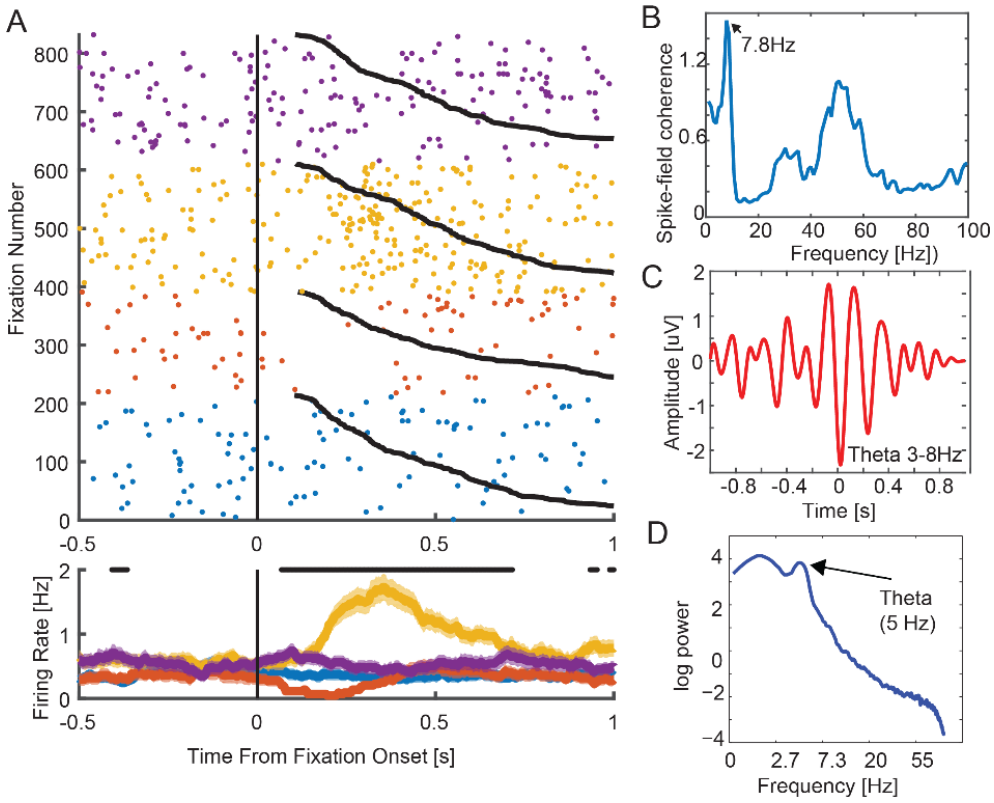


Figure 5.7: Theta-frequency spike field coherence of an amygdala cell. (a) Example cell recorded during the array task with strong response for human faces. **(b)** The peak of the spike-field coherence (measured on same electrode as the cell) is at 7.8 Hz. This particular cell also shows strong phase locking in the gamma band (40-60Hz) **(c)** The spike triggered average computed over all available spikes for the cell (n = 2913, across three sessions). **(d)** The spike-triggered power, peaking at approximately 5Hz, provides a measurement of the frequency of the theta oscillations recorded on this electrode.

References

Chang, L. and D. Y. Tsao (2017). "The Code for Facial Identity in the Primate Brain." Cell **169**(6): 1013-1028. e1014.

Murray, E. A. and M. Mishkin (1998). "Object recognition and location memory in monkeys with excitotoxic lesions of the amygdala and hippocampus." Journal of Neuroscience **18**(16): 6568-6582.

Pessoa, L. and R. Adolphs (2010). "Emotion processing and the amygdala: from a'low road'to'many roads' of evaluating biological significance." Nature reviews neuroscience **11**(11): 773-783.

Phelps, E. A. (2004). "Human emotion and memory: interactions of the amygdala and hippocampal complex." Current opinion in neurobiology **14**(2): 198-202.

Wang, S., et al. (2017). "The human amygdala parametrically encodes the intensity of specific facial emotions and their categorical ambiguity." Nature Communications **8**.

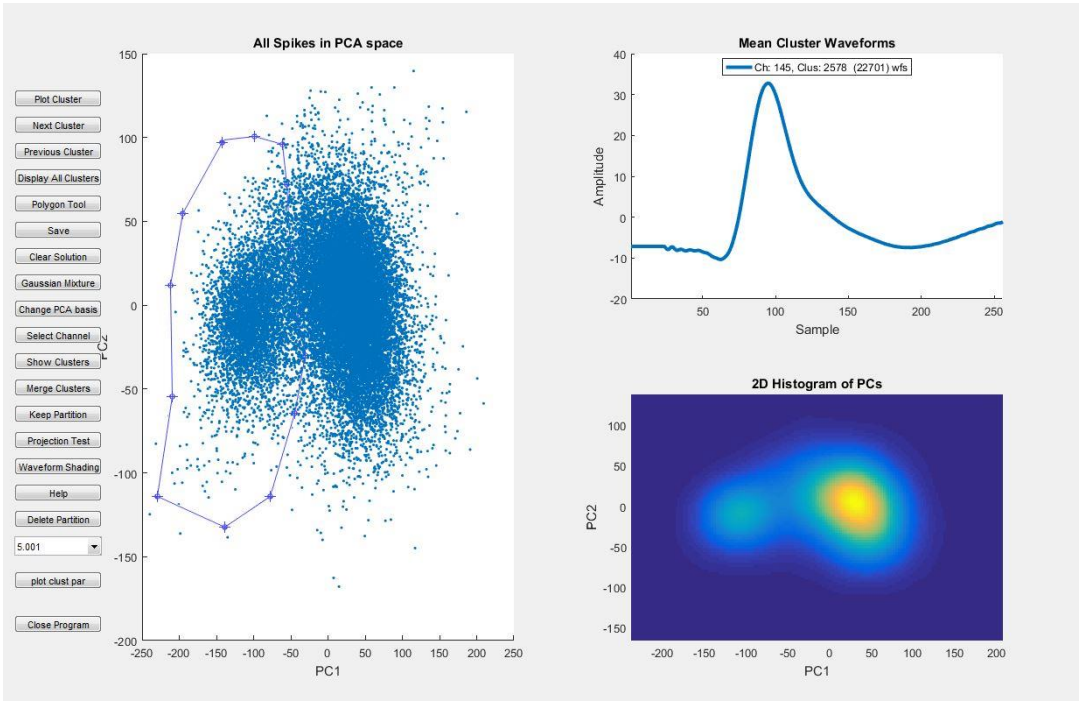
Zola, S. M., et al. (2000). "Impaired recognition memory in monkeys after damage limited to the hippocampal region." Journal of Neuroscience **20**(1): 451-463.

Appendix

There is a variety of methods and tools that were developed and either not used or used too sporadically to be mentioned in the main text of this thesis. I am including them here because I believe that they still represent a unique way of looking at behavioral and neural data and could potentially be useful for future applications.

7.1 A MATLAB interface for processing spikes

It seems that building a spike-sorting tool is a rite-of-passage for any graduate student working in electrophysiology. I was no exception, and building a tool that allowed me to combine both manual and automatic spike sorting was essential. Most of the tools we use for analyzing neural data had been developed in-house over the years. This particular tool was intended as a complimentary tool to Osort (Rutishauser, Schuman et al. 2006), which is the default method for the filtering, detecting, and sorting of spikes used in our lab. Osort provides automatic solutions, which then need to be curated manually. Often times, I needed to reject the automatic solutions because of mistakes in the sorting algorithm. Some of the most common manual operations include: (1) mergers of clusters, (2) cluster splits, (3) mixing across different sorting solutions. Often times a simpler method, such as drawing boundaries by hand (in PCA space), can be the best sorting solution. Other times, a parametric method, such as Gaussian Mixture Models, can yield the best results. The tool outlined below combines all of those possibilities into one easy-to-use platform.



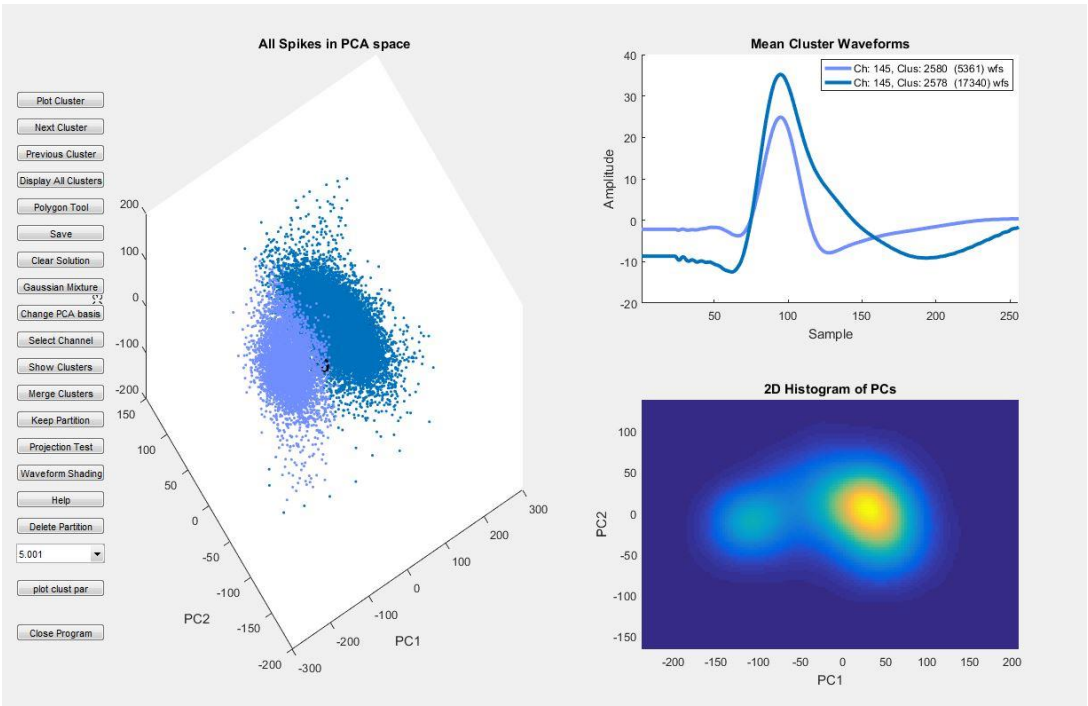


Figure 6.1: MATLAB-based spike sorting tool. The tool gives the user the ability to manually separate the clusters in PCA space **(top panel**, 2D or 3D drawing). It allows the users to scan through individual channels, merge across clusters, perform projection tests, change visualization schemes, or delete the entire solution on a channel in favor of a parametric method using a Gaussian Mixture Model (**bottom panel**).

7.2 A bin-free method for measuring onset latency

We developed a bin-less method to compute the latency of response of each single cell (relative to fixation onset). The method involves the following three steps for each cell. First, for each fixation i, compute the time since the last spike as a function (TSLS, Fig 7.4B) of time $TSLS(t,i)$. Second, use the interspike interval distribution P(ISI) to convert the TSLS to a probability: $P_{TSLS}(t,i) = P\big(ISI > TSLS(t,i)\big) = \int_{TSLS(t,i)}^{\infty} P(ISI > x)dx$ (Fig 7.1C-D). This probability is equivalent to the proportion of ISIs that is larger than the currently observed time since the last spike. The larger the deviation of $P_{TSLS}(t)$ from 0.5, the larger the deviation is of the firing of the cell from overall mean firing rate. Values >0.5 imply an increase in firing rate relative to baseline. Third, collapse the probabilities $P_{TSLS}(t,i)$ across all fixations n by multiplication of the individual probabilities: $P_{TSLS}(t) = \prod_{i=1}^{n} P_{TSLS}(t,i) = \sum_{i=1}^{n} log P_{TSLS}(t,i)$. This probabilistic interpretation of latency provides a robust latency estimate across all fixations regardless of tuning (Fig 7.4F).

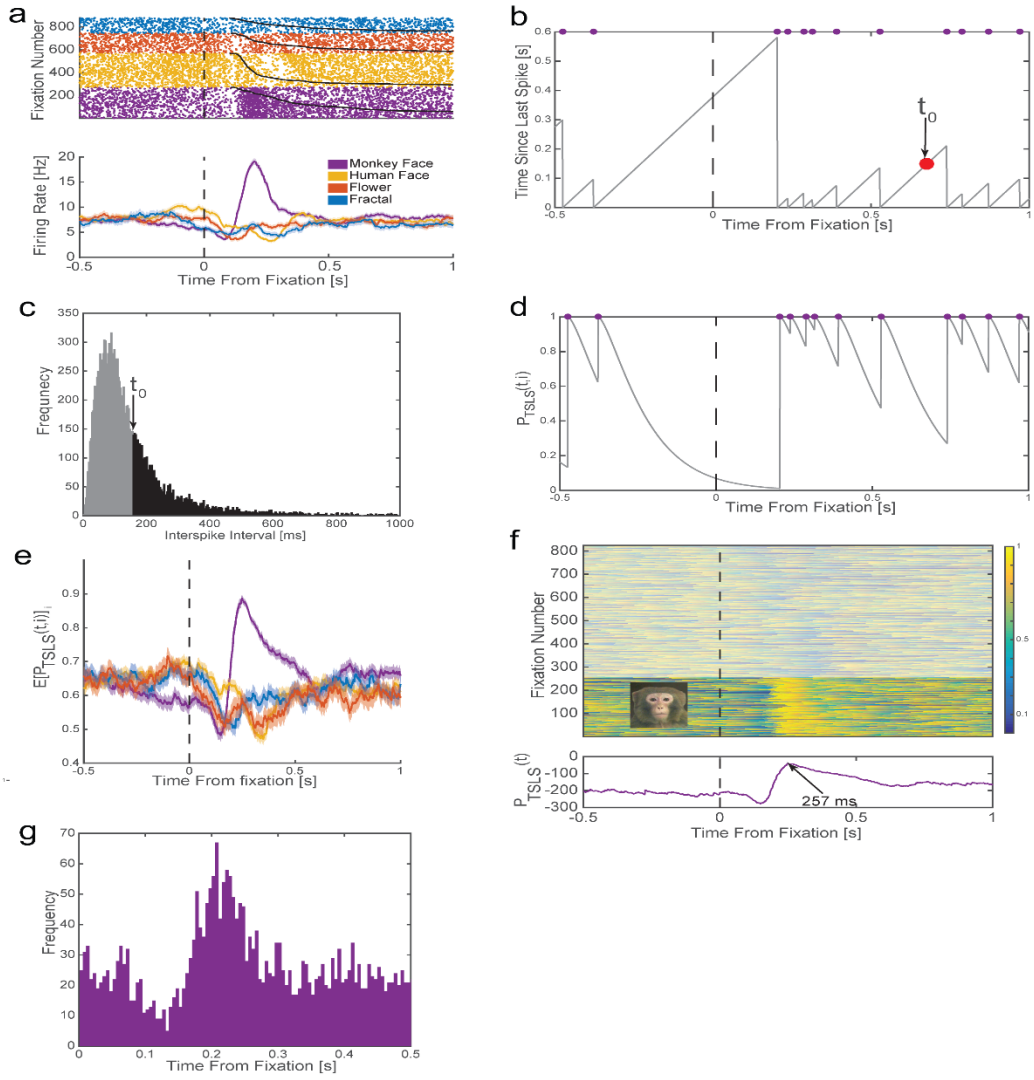


Figure 6.2: Bin-free time-since-last-spike (TSLS) latency estimation method. (A) Raster and PSTH of an example cell recorded from a monkey. (B) For each line (i.e. "look") in the raster in panel (A), we computed the TSLS(t) (gray line). Every time a spike occurs (magenta dots), the counter is reset to 0. (C) Histogram P(ISI) of all interspike-intervals (ISIs) observed for the cell shown in (A). (D) Same trial as (B), but converted to a probability. We converted the TSLS(t) values into probabilities based on how likely it is to observe a given ISI (see Methods for equation). For each observed TSLS value, this provides an estimate of how many ISIs are larger than the currently elapsed time. (E) Mean probability as a function of time and category. The TSLS method preserves information. The average of all traces within each category changes transiently shortly after fixation onset for the tuned category (i.e. monkey face). The firing rate for the tuned category increases, and consequently the average ISI decreases. This in turn results in a large probability (0.9 in this case), the ISIs of the cell are larger than the current TSLS. (F) The product of probabilities (log) across all fixations of a cell is calculated using an element-wise sum of the observed log likelihoods for each fixation. The peak of the log-likelihood function (arrow) provides a robust, bin-less, estimate of the latency of the cell. (G) Alternative method to determine single-cell latency using TSLS. For each fixation, the location of the peak (maximum) of the TSLS is computed and the peak of this histogram is used as the estimate of latency of this cell.

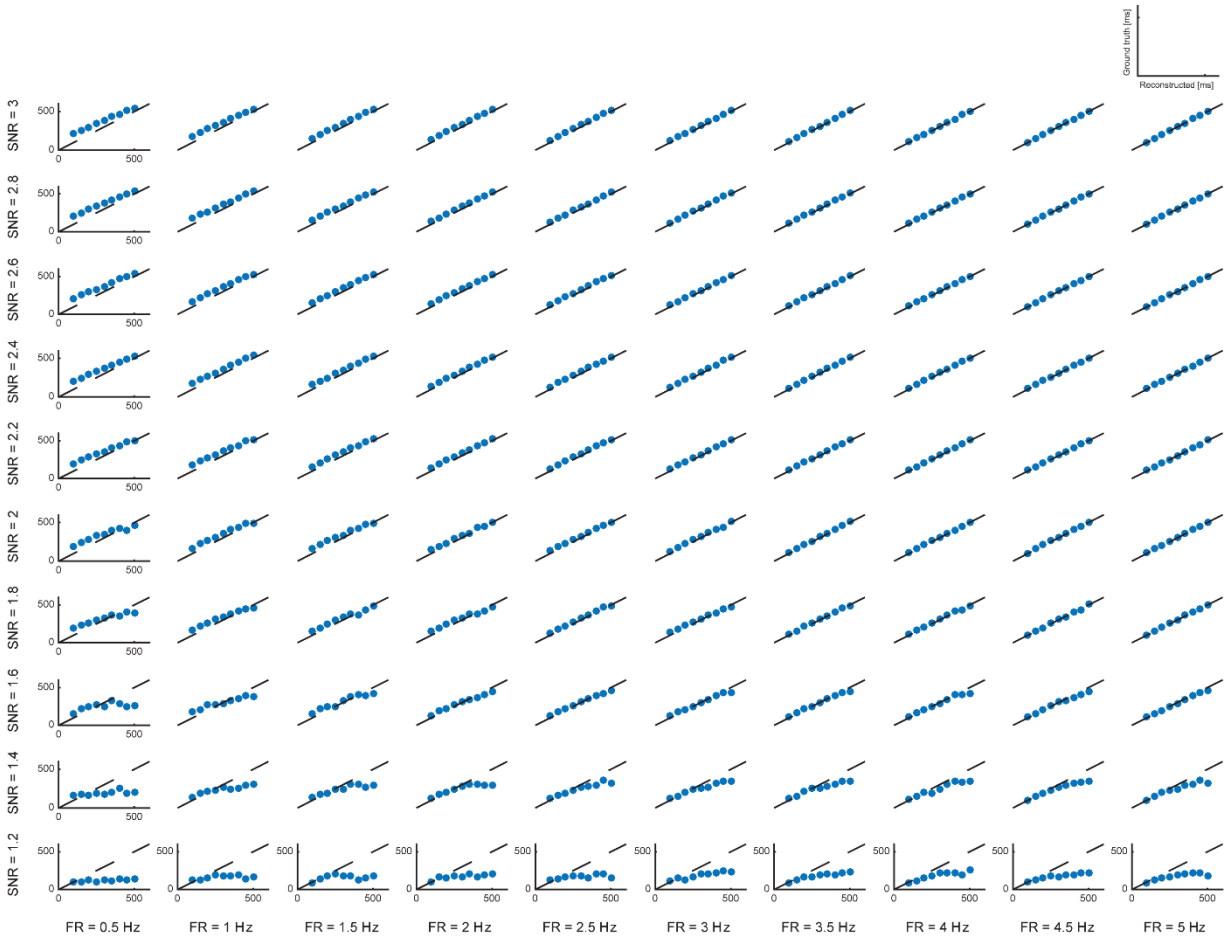


Figure 6.3: Reconstruction error of latency using the TSLS method. We simulated Poisson spiking with variable baseline firing rates (increasing from left to right) and variable signal-to-noise ratios (increasing from bottom to top). Plotted on each of the axes, is the true versus reconstructed latency of the cell for different combinations of baseline firing rate and signal to noise ratio.