# Social Saliency:
# Visual Psychophysics and Single-Neuron Recordings in Humans

**Thesis by**

**Shuo Wang**

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy


CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2014

(Defended May 27, 2014)

# ACKNOWLEDGEMENTS

My life at Caltech started in 2008. I was fortunate to get a SURF fellowship and worked with Naotsugu Tsuchiya and Ralph Adolphs over the summer. I really enjoyed the research environment at Caltech and learnt a lot from Nao. It was a productive and determinative summer for me, after which I decided to return to Caltech. Again, I was fortunate to get a chance to work with Nao and Ralph as a research assistant until I started officially as a graduate student in 2010.

I came to Ralph's lab with a passion for science. Ralph gave me the anomalous freedom to decide what I would like to work on and to do the projects in my own way. Ralph let me explore all the fields of my interest while always offering me inspiring guidance when I was wandering around with puzzled mind. Freedom came from his trust in the student, while guidance resulted in the trust of the student in him. I have never been intimidated by any hassle or difficulty during my PhD, because he is always supportive as an advisor. Besides the academic advice, Ralph is also broad-minded and outgoing, and always ensures that our PhD study is not boring, with numerous lab parties, camping trips and movies. I cannot imagine such a great advisor anywhere else. This thesis would not have been written without the advice and support of Ralph.

Nao was my mentor for the first few projects. I remember all the exciting moments when our experiments worked out. I learnt a lot of visual psychophysics and consciousness from Nao and was strongly influenced by his way of doing research. Without Nao's help, I could not have published my first paper.

I returned to Singapore and worked with Annett Schirmer in the summer of 2009 and have been collaborating with Annett since then. I learnt how to do EEG experiments and analysis, which is a very valuable experience.

Christof Koch was my co-advisor. Nao introduced me to Christof to work on a psychophysics project of optic flow. I am very happy that this project eventually turned

out to be a paper and became a chapter of my thesis. I would like to thank Christof for all his support and advice.

During my rotation, I am very grateful to have had an experience to work on monkeys with Doris Tsao, even though our idea of using fMRI to probe unconscious pressing of faces in face patches failed eventually after struggling for half a year. I thank Shay Ohayon for helping me to set up the experiment and Ty Basinger for helping me train the monkey very hard. I am also grateful to work with John O'Doherty and Mimi Liljeholm on human fMRI, which is a very valuable experience for me. I thank Mimi for teaching me the basics of analyzing fMRI data.

I have been working with Ueli Rutishauser on single-unit recordings. I am very happy that our paper has finally been accepted to PNAS after a long struggle. I thank Ueli for teaching me all the single unit analysis and discussing with me all the interesting ideas. I am very impressed by his passion in science and his meticulousness.

I thank Qi Zhao for teaching me the computational modeling of saliency and I have been collaborating with her on various saliency projects. I hope our current work will be published in the near future.

I have been fortunate to be surrounded my colleagues and talented peers from the CNS program, especially my classmates Bo Chen, Nikki Sullivan, Soyoung Park, Yong-Jun Lin and Dan McNamee. We had many mentally stimulating discussions and I have very cheerful memories of hanging out with my classmates.

I thank all of my colleagues in the Adolphs lab and Koch lab, past and present, for their generous help.

I am grateful to my thesis committee: Ralph Adolphs, Christof Koch, John O'Doherty, Shinsuke Shimojo, Doris Tsao and Ueli Rutishauser.

Finally, I would like to thank my parents for their generous support and encouragement.

# ABSTRACT

My thesis studies how people pay attention to other people and the environment. How does the brain figure out what is important and what are the neural mechanisms underlying attention? What is special about salient social cues compared to salient non-social cues? In Chapter I, I review social cues that attract attention, with an emphasis on the neurobiology of these social cues. I also review neurological and psychiatric links: the relationship between saliency, the amygdala and autism. The first empirical chapter then begins by noting that people constantly move in the environment. In Chapter II, I study the spatial cues that attract attention during locomotion using a cued speeded discrimination task. I found that when the motion was expansive, attention was attracted towards the singular point of the optic flow (the focus of expansion, FOE) in a sustained fashion. The more ecologically valid the motion features became (e.g., temporal expansion of each object, spatial depth structure implied by distribution of the size of the objects), the stronger the attentional effects. However, compared to inanimate objects and cues, people preferentially attend to animals and faces, a process in which the amygdala is thought to play an important role. To directly compare social cues and non-social cues in the same experiment and investigate the neural structures processing social cues, in Chapter III, I employ a change detection task and test four rare patients with bilateral amygdala lesions. All four amygdala patients showed a normal pattern of reliably faster and more accurate detection of animate stimuli, suggesting that advantageous processing of social cues can be preserved even without the amygdala, a key structure of the "social brain". People not only attend to faces, but also pay attention to others' facial emotions and analyze faces in great detail. Humans have a dedicated system for processing faces and the amygdala has long been associated with a key role in recognizing facial emotions. In Chapter IV, I study the neural mechanisms of emotion perception and find that single neurons in the human amygdala are selective for subjective judgment of others' emotions. Lastly, people typically pay special attention to faces and people, but

people with autism spectrum disorders (ASD) might not. To further study social attention and explore possible deficits of social attention in autism, in Chapter V, I employ a visual search task and show that people with ASD have reduced attention, especially social attention, to target-congruent objects in the search array. This deficit cannot be explained by low-level visual properties of the stimuli and is independent of the amygdala, but it is dependent on task demands. Overall, through visual psychophysics with concurrent eye-tracking, my thesis found and analyzed socially salient cues and compared social vs. non-social cues and healthy vs. clinical populations. Neural mechanisms underlying social saliency were elucidated through electrophysiology and lesion studies. I finally propose further research questions based on the findings in my thesis and introduce my follow-up studies and preliminary results beyond the scope of this thesis in the very last section, Future Directions.

TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# Chapter I: General Introduction

## 1.1 Overview

Saliency historically refers to the bottom-up visual properties of an object that automatically drive attention. It is an ordinal property that depends on the relative saliency of one object with respect to others in the scene. Simple examples are a red spot on a green background, a horizontal bar among vertical bars, or a sudden onset of motion. Researchers have introduced the idea of a saliency map, an abstract and featureless map of the 'winners' of attention competition, to model the dynamics of visual attention. The standard saliency map involves channels like color, orientation, size, shape, movement or unique onset. But how do complex stimuli, especially stimuli with social meaning such as faces, pop out and attract attention? Suppose you are attending a big party: your attention might be captured by someone in a fancy dress, someone looking at you, someone who is attractive, familiar, or distinctive in some way. This happens essentially automatically, and encompasses a huge number of different stimuli that are all competing for your attention. What determines which is the most salient, and how can we best measure this?

Humans are social animals. We constantly interact with other people and the environment and we are unceasingly bombarded with various socially salient stimuli: faces, gestures, emoticons, and socially relevant pieces of text. These capture our attention, are encoded preferentially into memory, and influences our thoughts and actions. What is it about such stimuli that accomplishes these multiple effects? In particular, are there mechanisms analogous to those known to operate for low-level (non-social) saliency (such as visual motion and contrast)? Is there a finite vocabulary of social attention-grabbing cues, or a small set of dimensions that render social stimuli especially salient? Finally, how does the brain figure out what is important and what are the underlying neural mechanisms?

To understand these questions, I conducted several studies in my thesis. First, I analyzed a strong non-social cue, optic flow, that is well known to attract attention. Second, I

compared animate social cues to inanimate non-social cues in an attentional change detection task and investigated their dependence on the amygdala, a neural structure that is recognized for processing socially relevant stimuli. Third, I analyzed in detail one particularly salient social stimuli—faces, and its neural representation in the amygdala. Lastly, I directly compared social stimuli of people and faces to non-social stimuli of, e.g., food, gadgets, and electronics, in an attentional visual search task. Deficiencies in processing social cues lead to complex disabilities such as autism. In the last study, I also included people with autism spectrum disorders (ASD) to explore the deficits in processing socially salient stimuli, and compared results from this clinical population directly with the results from patients with amygdala lesions.

The General Introduction is organized as follows. First, I begin by discussing the socially salient cues, including eye gazes, faces, head directions, finger gestures, postures, actions, biological motion, personal distance, social touch, and social rewards. I underscore the neural substrates underlying processing of such social stimuli. Second, I discuss the role of the amygdala in encoding saliency. Third, I discuss the deficits in processing socially salient stimuli in autism and the possible involvement of the amygdala in these deficits. Lastly, I outline my thesis and propose future directions.

### 1.2 Socially salient cues

In everyday life, we are constantly bombarded by social cues. Rich information can be derived through social cues. But how does the brain figure out the message conveyed by the social cues? How are social cues represented and integrated in the brain? In particular, is social interaction mediated through a mechanism that relies on saliency?

In this section, I discuss the socially salient cues, which include eye gazes, faces, head directions, finger gestures, postures, actions, biological motion, personal distance, social touch, and social rewards. I also discuss what is known about the functions and the neural

underpinnings of each social cue, as well as the developmental ontogeny and comparative neurobiology of the social cues.

### *1.2.1 Eye gaze*

Eye gaze plays important roles in social communication. It functions for information seeking, signaling, controlling the synchronizing of speech, cueing for intimacy, avoiding undue intimacy, and avoiding excess input of information (Argyle et al., 1973). Especially, eye gaze directs attention and provides important sources of social information. Behavioral studies show that chimpanzees spontaneously follow human gaze direction and share joint visual attention (Povinelli and Eddy, 1996). Rhesus monkeys can also follow gazes and use the attentional cues of other monkeys to orient their own attention to objects (Emery et al., 1997). Human studies show that newborns prefer faces with eyes open vs. eyes closed (Batki et al., 2000), and that infants as early as 10 weeks of age follow the gaze of others (Hood et al., 1998), arguing for innate aspects to gaze cognition. Deficiencies in processing eye gaze is associated with complex disabilities such as autism spectrum disorder (Pelphrey et al., 2002).

Friesen and Kingstone reported that humans attend reflexively to locations and objects that are being looked at by other people (Friesen and Kingstone, 1998). Humans infer other people's movement trajectories from their gaze direction and use this information to guide their own visual scanning of the environment (Nummenmaa et al., 2009). However, the specificity of eye gaze in orienting attention has been questioned and it has been reported that other non-social stimuli, such as arrows, trigger reflexive shifts in attention in a manner behaviorally identical to those triggered by eyes (Ristic et al., 2002). Further investigation has differentiated these two types of attentional cues neuronally, showing that the neural systems subserving eye gazes and arrows are not equivalent—with the superior temporal sulcus (STS) being engaged disproportionately when the fixation stimulus is perceived as eyes (Kingstone et al., 2004). This is further supported by a study

using a similar spatial cueing paradigm on a patient with circumscribed superior temporal gyrus damage, who showed a detection advantage only when cued by non-biological arrows but not gaze (Akiyama et al., 2006). The specificity of the STS in cueing by gaze is also illustrated through patients with parietal damage—perceived gaze in faces can still trigger automatic shifts of attention in the contra-lesional direction, even though parietal damage causes spatial neglect and impairs the representation of location on the contra-lesional side (Vuilleumier, 2002). This suggests a specific and anatomically distinct attentional mechanism through the STS. Interestingly, attentional shift by gaze can be triggered without awareness (Sato et al., 2007).

STS activation has consistently been demonstrated in the normal brain when viewing eyes and this brain region has been implicated in gaze processing. Recordings from single cells in awake, behaving monkeys have shown that this region of the temporal lobe is sensitive to faces and further modulated by head direction and gaze direction (Perrett et al., 1985). Similarly, STS lesions in the rhesus monkey impair gaze direction discrimination (Campbell et al., 1990). Human neuroimaging studies have demonstrated that a superior temporal region centered in the STS is activated when a subject views a face in which the eyes shift their gaze (Puce et al., 1998). In a spatial cueing paradigm, the STS has been demonstrated to be sensitive to the social context in which a gaze shift occurs (Pelphrey et al., 2003). Moreover, in a virtual reality environment, mutual gaze evokes greater activity in the STS than averted gaze, suggesting that the STS is involved in processing social information conveyed by gaze shifts within an overtly social context (Pelphrey et al., 2004). Multivariate pattern analysis of human functional imaging data has shown that anterior STS encodes the direction of another's attention regardless of how this information is conveyed (Carlin et al., 2011).

Perception of eye gaze also recruits the spatial cognition system in the intraparietal sulcus to encode and pay attention to the direction of another's gaze (Hoffman and Haxby, 2000). Research with split-brain patients suggests that lateralized cortical connections between temporal lobe subsystems specialized for processing gaze, and parietal lobe

subsystems specialized for orienting spatial attention, underlie the reflexive joint attention elicited by gaze (Kingstone et al., 2000). Besides the STS, contrasting between directed vs. averted gaze produces a tight cluster of activation corresponding to and restricted to the central nucleus and the bed nucleus of the stria terminalis (termed the lateral extended amygdala) in the monkey (Hoffman et al., 2007).

### *1.2.2 Faces*

People often form judgments of others based purely on facial features. People are able to pick up subtle changes in facial structures from faces varying along one dimension to another (Oosterhof and Todorov, 2008). Trait evaluations from faces can predict important social outcomes. Inferences of competence based solely on facial appearance predict the outcomes of elections (Todorov et al., 2005). Facial features can also influence sentencing decisions – inmates with more Afrocentric features received harsher sentences than those with less Afrocentric features (Blair et al., 2004). Remarkably, impressions and judgments of unfamiliar people can be formed by a very brief exposure to faces as short as 100 ms (Willis and Todorov, 2006).

Primates have a dedicated visual system to process faces (Tsao et al., 2006). Electrophysiological recordings in monkeys (Rolls, 1984, Leonard et al., 1985) and humans (Kreiman et al., 2000, Rutishauser et al., 2011) have found single neurons that respond not only to faces, but also to face identities, facial expressions and gaze directions (Gothard et al., 2007, Hoffman et al., 2007). Neuroimaging studies have revealed neural substrates for emotional attention, which might supplement but also compete with other sources of top-down control on perception (Vuilleumier, 2005). In particular, the amygdala plays a crucial role in emotional attention and is required for accurate social judgments of other individuals on the basis of their facial appearance. The amygdala also shares parallel roles in humans and other animals in emotional influences on attention and social behavior (see Phelps and LeDoux, 2005 for a review). Lesion

studies showed that patients with complete bilateral amygdala damage judge unfamiliar individuals to be more approachable and more trustworthy than do control subjects (Adolphs et al., 1998). The impairment is most striking for faces which normal subjects judge most unapproachable and untrustworthy. Besides the amygdala, recent findings show that orbitofrontal cortex lesions result in abnormal social judgments to emotional faces (Willis et al., 2010). The relationship between the amygdala and saliency is reviewed in more detail in the next section.

As reviewed in more detail below, people with autism have altered saliency representations towards faces compared to non-face objects, as shown by reduced attention to faces compared to inanimate objects (Dawson et al., 2005, Sasson, 2006), as well as circumscribed interests to a narrow range of inanimate subjects (e.g., gadgets, devices, electronics and Japanese animation, etc.) (Kanner, 1943, Lewis and Bodfish, 1998, South et al., 2005). It has even been shown in children and adolescents (Sasson et al., 2008), as well as in 2–5 year-olds (Sasson et al., 2011), that people with autism fixated faces or people less than controls when freely viewing arrays containing both faces and non-face objects. When looking within faces, the relative saliency of facial features is also altered in autism, as evidenced from both behavioral and neuronal findings: people with autism have an increased tendency to saccade away from (Spezio et al., 2007b) and actively avoid the eyes (Kliemann et al., 2010), but have an increased preference to fixate (Neumann et al., 2006) and rely on information from the mouth (Spezio et al., 2007a). The behavioral abnormality is supported by neuronal evidence of abnormal processing of information from the eye region of faces in single cells recorded from the amygdala in neurosurgical patients with ASD (Rutishauser et al., 2013).

Interestingly, gaze cues interact with facial emotion cues. It has been shown that direct gaze facilitates the processing of facially communicated approach-oriented emotions (e.g., anger and joy), whereas averted gaze would facilitate the processing of facially communicated avoidance-oriented emotions (e.g., fear and sadness), suggesting that gaze

cues combine with facial emotion cues in the processing of emotionally relevant facial information (Adams and Kleck, 2003).

### *1.2.3 Head direction*

Comparative research with non-human primates suggests that the orientation of the head might provide a stronger cue to another individual's attentional direction than eye-gaze alone (Langton et al., 2000). In humans, manipulating the face directions of emotional expressions in the unilateral visual fields allows us to alter the emotional significance of the facial expression for the observer without affecting the physical features of the expression. It has been shown that the left amygdala increases activity for angry expressions looking towards the subjects than angry expressions looking away from them, suggesting that the amygdala is involved in emotional processing for facial expressions (Sato et al., 2004). In infants, the emergence of the tendency to look where another person looks is a fundamental landmark in the development of referential communication. It has been found that normal infants 10 to 12 months old reliably look in the direction towards which adults turn their heads and eyes (Scaife and Bruner, 1975).

There is often interplay between gaze direction and head direction. Human neuroimaging studies suggest that right anterior STS is invariant to head view and physical image features (Carlin et al., 2011). Furthermore, head direction cues also interact with body direction cues—only when the head is rotated in the cuing person's reference frame but not the observer's frame, can head direction cues shift the observer's attention to the same direction (Hietanen, 2002).

### *1.2.4 Finger gestures*

Finger pointing provides social information and captures attention. Many animals can use experimenter-given cues in an experimental setup and are sensitive to human gestural

communication. In object-choice tasks, lowland gorillas complete the task better when the experimenter taps on or points at an object that contains a reward. Performance remains good when the experimenter gazes with eyes and head orients towards the correct object without manual gestures. In contrast, when only the experimenter's eye orientation serves as the cue, the gorillas do not appropriately complete the task (Peignot and Anderson, 1999) (but also see (Byrnit, 2009)). In a similar task, capuchin monkeys do not use the experimenter's gazing as a cue to find the correct baited object. In contrast, they do use gazing plus pointing, and it has been shown that pointing is necessary and sufficient under the conditions of the study (Anderson et al., 1995). Dogs are able to utilize pointing, bowing, nodding, head turning and glancing gestures of humans as cues for finding hidden food. Interestingly, this ability can be generalized from one person (owner) to another familiar person (experimenter) in using the same gestures as cues (Miklösi et al., 1998). Even fur seals were found to be able to follow human gestures—they are able to use cues involving a fully exposed arm or a head direction, but fail to use glance only, suggesting that a domestication process is not necessary to develop receptive skills to cues given by humans (Scheumann and Call, 2004). In humans, finger pointing gestures can interfere with speech in a Stroop-type paradigm, suggesting that verbal and non-verbal dimensions are integrated prior to the response selection stage of processing (Langton et al., 1996).

### *1.2.5 Postures, actions and biological motion*

People not only pay attention to their own motion in the environment (Wang et al., 2012a), but also pay attention to and automatically infer other people’s mental states such as intention from their motion and actions. Psychophysical and functional neuroimaging evidence shows that biological motion is processed as a special category, and the mechanism underlying the attribution of intentions to actions might rely on simulating the observed action and mapping it onto representations of our own intentions (Blakemore and Decety, 2001). In a visual search paradigm with point-light animations,

differentiating between actions requires attention in general and there are search asymmetries between actions (van Boxtel and Lu, 2011). Particularly, animated threatening boxer targets pop out from emotionally neutral walker distractors in a crowd, whereas walkers do not, showing that body cues signal important social information related to threat and survival (van Boxtel and Lu, 2012). It has been suggested that body cues rather than facial expressions discriminate between intense positive and negative emotions (Aviezer et al., 2012). Even laboratory rodents, the rat, the mouse, the guinea pig, and the golden hamster, use postures and acts to signal social information (Grant and Mackintosh, 1963).

### *1.2.6 Personal distance*

When interpersonal space is invaded, people feel uncomfortable and aroused. People automatically and reliably regulate the distance between one another during social interaction (Hall, 1966). The amygdala may be required to trigger strong emotional and arousal reactions when personal space is invaded, as evidenced by neuroimaging data showing amygdala activation in healthy individuals upon close personal proximity, and a lack of personal space in an individual with complete amygdala lesions (Kennedy et al., 2009). This is consistent with monkey studies showing that amygdalectomized monkeys demonstrated increased social affiliation, decreased anxiety, and increased confidence compared with control monkeys (Emery et al., 2001). These effects might arise from lack of a saliency signal mediated by the amygdala to personal space violation. Furthermore, the monkey amygdala mediates the approach and avoidance to ambiguous or threatening novel situations and people (Mason et al., 2006).

### *1.2.7 Social touch*

Social touch is a salient cue in our everyday social interactions since it plays a particularly important role in social bonding which, in turn, has a major impact on an individual's lifetime reproductive fitness (Dunbar, 2010). Interpersonal touch provides an effective means of influencing people's various social behaviors (see Gallace and Spence, 2010 for a review) and plays an important role in governing our emotional well-being (Field, 2003). Human orbitofrontal cortex represents affectively positive and negative touch in different areas (Rolls et al., 2003) (also see Rolls, 2010 for a review).

In addition to the fast-conducting myelinated afferent fibers responsible for tactile sensation, a system of slow-conducting unmyelinated tactile (CT) afferents is responsible for affective sensation, as supported by neuroimaging studies in a unique patient lacking large myelinated afferents. Those studies showed that touch activates brain regions implicated in emotional and social processing such as the insula, but not primary somatosensory areas (Olausson et al., 2002). Further electrophysiological studies in healthy individuals have shown that soft brush stroking activates CT afferents but not myelinated afferents, suggesting that CT afferents constitute a privileged peripheral pathway for pleasant tactile stimulation that is likely to signal affiliative social body contact (Loken et al., 2009). However, recent neuroimaging studies in humans have shown that the response in primary somatosensory cortices to a sensual caress is modified by the perceived sex of the caresser, arguing for a more important role that somatosensory areas might play in affective processing than previously thought (Gazzola et al., 2012). Indeed, the different components of social touch, such as somatosensory experience, the proximity to the person, and an attribution of the somatosensory experience to the person, have been teased apart (Schirmer et al., 2011). All these findings illustrate that touch is a special and salient social cue that enhances visual attention and sensitizes ongoing cognitive and emotional processes.

### *1.2.8 Social rewards*

People not only pay attention to concrete, physical social cues delivered by direct body-body interaction, but also pay attention to more abstract social cues such as rewards. Attentional control is at the center of the function of dopamine in reinforcement learning and animal approach behavior (Ikemoto and Panksepp, 1999, Wise, 2004). Dopamine systems mediate the incentive saliency of rewards by specifically changing the perceptual representation of reward-conditioned stimuli such that they become salient and draw attention (Berridge and Robinson, 1998). It has been shown that sensory and perceptual processing of reward-associated visual features is facilitated such that attention is deployed to objects characterized by these features, even when a strategic decision to attend to reward-associated features will be counterproductive and result in suboptimal performance (Hickey et al., 2010). Visual search for a salient target is slowed by the presence of an inconspicuous, task irrelevant distractor previously associated with monetary reward through learning, showing that the value of stimuli can modulate voluntary attention allocation (Anderson et al., 2011). Reward can even create oculomotor saliency and modulate saccade trajectories, suggesting low-level and non-strategic mechanisms that operate automatically (Hickey and van Zoest, 2012).

Complex social behavior and decision making may share the same neural basis as simple monetary evaluation and learning. The acquisition of one's good reputation activated striatum as robustly as did monetary rewards, suggesting a 'common neural currency' for rewards (Izuma et al., 2008). In healthy individuals, social and monetary reward learning share overlapping neural substrates (Lin et al., 2012b). In monkeys, neurons from orbitofrontal cortex signal both social values and juice rewards, and far more neurons signal social category than fluid value, despite the stronger impact of fluid reward on monkeys' choices (Watson and Platt, 2012). Interestingly, people can make optimal reward choices without being fully aware of the basis of their decision (Wang et al., 2012b). Moreover, people with autism show various impairments in social decision making and reward learning (see below) (Izuma et al., 2011, Lin et al., 2012a, Lin et al., 2012c).

### 1.3 Neural representation of saliency

A distributed network of visuomotor areas is proposed to encode a representation of saliency that combines bottom-up and top-down influences to identify locations for further processing. Neurons in the primate frontal eye field (FEF) exhibit the characteristics of a visual saliency map—they are not sensitive to specific features of visual stimuli, but their activity evolves over time to select the target of the search array (Thompson et al., 1996, Thompson and Bichot, 2005). Visual activity in the FEF not only signals location of targets for orienting, but also signals movement-related saccade preparation (Murthy et al., 2009). However, in an adjacent area, the supplementary eye field (SEF), only very few neurons selected the location of the search target (Purcell et al., 2012), showing a very limited role of the SEF in encoding target saliency. Moreover, it has also been shown that neurons in the lateral intraparietal (LIP) area reflect selection to salient stimuli defined by a target when animals have to make a saccade towards the salient stimulus (Thomas and Paré, 2007). Activity in the LIP correlates with the monkey's planning of memory-guided saccades to goal-directed salient locations in visual search tasks (Ipata et al., 2006) and these neurons only respond to stimuli that are behaviorally significant (Gottlieb et al., 1998). Studies have even found pure bottom-up saliency signals in LIP in a passive fixation task without any top-down instructions (Arcizet et al., 2011). Furthermore, individual neurons in monkey area 7a of the posterior parietal cortex encode the location of the salient stimulus and can thus provide spatial information required for orienting to a salient spot in a complex scene (Constantinidis and Steinmetz, 2001).

Besides cortical areas, subcortical superior colliculus (SC) encodes both stimulus identity and saccade goals during visual conjunction search (Shen and Paré, 2007). Neuronal activity in the SC signals selection or increased saliency of subsequent saccade goals even before the initial saccade has ended (McPeek and Keller, 2002), and a recent report has shown that the process can encompass at least two future saccade targets (Shen and

Paré, 2014), suggesting parallel processing of visual saliency in the SC. The causal functional role of the SC in target selection has been revealed by focal reversible inactivation in monkeys (McPeek and Keller, 2004). Interestingly, even substantia nigra pars reticulata (SNr) has been shown to change activity with target selection and saccade initiation, which in turn may make substantial and direct contributions to the SC (Basso and Wurtz, 2002).

On the other hand, neurons in the inferior temporal cortex have been suggested to play a critical role in representing and processing visual objects (Gross, 1994, Logothetis and Sheinberg, 1996, Tanaka, 1997), which holds for both isolated objects and objects in complex natural scenes (Sheinberg and Logothetis, 2001). Compared to visual areas in the temporal lobe, early visual areas such as V1 and V2 are generally accepted to represent low-level visual features, and V4 has been reported to show convergence of bottom-up and top-down processing streams that facilitate oculomotor planning for visual search (Mazer and Gallant, 2003). V4 neurons not only enhance responses to a preferred stimulus in their receptive field when the stimulus matched a feature of the target, but also enhance responses to candidate targets selected for saccades (Bichot et al., 2005). Comparing pop-out and conjunctive stimuli, V4 neurons encode pop-out saliency in a top-down attention-dependent manner (Burrows and Moore, 2009). In both single saccade tasks (Chelazzi et al., 1993, Tolias et al., 2001, Ogawa and Komatsu, 2004) and tasks with naturalistic free-viewing (Sheinberg and Logothetis, 2001, Mazer and Gallant, 2003, Bichot et al., 2005), several studies have reported that temporal cortical neurons enhanced responses to visual stimuli presaccadically when the stimulus in the receptive field becomes the target, suggesting that task-relevant target saliency is encoded by these neurons.

### 1.4 Saliency and the amygdala

The human amygdala is critical to process emotionally salient and socially relevant stimuli (Kling and Brothers, 1992, Adolphs, 2010). Earlier views of the amygdala in representing saliency emphasized a fear-related function, and the amygdala has generally been conceptualized as a fear-processing module. This view was supported by animal models of fear conditioning (LeDoux, 1993) and impairment of fear conditioning after amygdala damage in humans (Bechara et al., 1995, LaBar et al., 1995). Human studies demonstrated a selective impairment in recognizing fearful faces in subjects that lack a functional amygdala (Adolphs et al., 1994), mirrored by neuroimaging studies showing significant activation differences within the amygdala to fearful faces compared to happy faces (Morris et al., 1996). Interestingly, increased amygdala BOLD-fMRI to fearful stimuli was linked to serotonin transporter genes, which have been associated with anxiety-related behaviors (Hariri et al., 2002).

Recently, however, (Adolphs, 2010) argued that the amygdala plays a broader role in social cognition and processes a stimulus dimension related to saliency or relevance in general. The amygdala has been proposed to respond to a broader spectrum of social attributes such as facial emotions in general and regulating a person's personal space (Kennedy et al., 2009), rather than being specific for fearful faces (Fitzgerald et al., 2006). Electrophysiological recordings in monkeys have found single neurons that respond not only to faces (Rolls, 1984, Leonard et al., 1985), but also to face identities, facial expressions and gaze directions (Gothard et al., 2007, Hoffman et al., 2007). In humans, it has been reported that amygdala neurons are selective for a variety of visual stimuli (Kreiman et al., 2000). Single neurons in the human amygdala have been found to encode whole faces selectively (Rutishauser et al., 2011) and account for the abnormal face processing in autism (Rutishauser et al., 2013). A recent study has shown that neurons in the human amygdala encode subjective judgment of facial emotions, rather than simply their stimulus features (Wang et al., 2014c).

Salient social cues signal value, and the amygdala responds to values and rewards that are important to the organism (Baxter and Murray, 2002). The primate amygdala represents

the positive and negative value of visual stimuli during learning (Paton et al., 2006) and is sensitive to temporal reward structure (Bermudez et al., 2012). Monkeys with amygdala lesions showed impaired devaluation in selectively satiated food, indicating a failure to respond to the changing value of food rewards (Malkova et al., 1997, Baxter et al., 2000). These and other findings support a clear role for the amygdala in goal-directed instrumental learning, a function that is in addition to its even better established role in Pavlovian fear conditioning. Furthermore, amygdala neurons can predict the monkey's save-spend choices while monkeys choose between saving liquid reward with interest and spending the accumulated reward (Grabenhorst et al., 2012). In rats, rapid strengthening of thalamo-amygdala synapses mediates cue–reward learning (Tye et al., 2008). These various roles that the amygdala plays in aspects of reward learning no doubt support its function in social behavior.

Lastly, the amygdala processes more abstract attributes such as stimulus unpredictability (Herry et al., 2007). Amygdala lesions result in an absence or reduction of fixations on novel objects observed in monkeys (Bagshaw et al., 1972). It has also been shown that the amygdala mediates emotion-enhanced vividness (Todd et al., 2012), responds more to animate entities compared to inanimate ones (Yang et al., 2012b), and is even selective to animals (Mormann et al., 2011). The amygdala has also been reported to modulate consolidation of aspects of declarative memory, especially for highly emotionally arousing tasks (McGaugh, 2000, 2004, Roozendaal et al., 2008).

However, there are also many examples showing no obvious corresponding behavioral impairment when the amygdala is lesioned. Recent findings have shown that preferential attention to animals and people is independent of the amygdala (Wang et al., 2014b), and amygdala lesions do not lead to deficits in social attention as observed in people with autism (Wang et al., 2014d). These findings are consistent with preserved attentional capture by emotional stimuli and intact emotion-guided visual search in patients with acute amygdala lesions due to neurosurgical resection (Piech et al., 2010, Piech et al., 2011). Besides compensatory circuits that might account for the intact social attention in

amygdala lesion patients (Becker et al., 2012), these findings leave open the question of what are the essential structures mediating social saliency and to what extent the amygdala contributes to social saliency. These questions remain important topics for future studies.

Overall, the amygdala might act as a detector of perceptual saliency and biological relevance (Sander et al., 2005, Adolphs, 2008). The functional role of the amygdala is supported by its connection with the visual cortices specialized for face processing (Vuilleumier et al., 2004, Moeller et al., 2008, Hadj-Bouziane et al., 2012) as well as reciprocal connections with multiple visually responsive areas in the temporal (Desimone and Gross, 1979, Amaral et al., 2003, Freese and Amaral, 2006) and frontal lobes (Ghashghaei and Barbas, 2002).

### 1.5 Saliency and autism

Autism is a disorder characterized by impairments in social and communicative behavior and a restricted range of interests and behaviors (DSM-5, 2013). Individuals with autism show reduced attention to faces as well as to all other social stimuli such as the human voice and hand gestures, but pay more attention to inanimate objects (Dawson et al., 2005, Sasson, 2006). Some characteristics, such as preference for inanimate objects and a lack of interest in social objects, are often evident very early in infancy (Kanner, 1943, Osterling and Dawson, 1994). Children with autism displayed significantly fewer social and joint attention behaviors, including pointing, showing objects, looking at others, and orienting to name (Osterling and Dawson, 1994). People with autism also show circumscribed interests to a narrow range of inanimate subjects, a type of repetitive behavior occurring commonly in autism, and are fascinated with gadgets, devices, vehicles, electronics, Japanese animation and dinosaurs, etc. (Kanner, 1943, Lewis and Bodfish, 1998, South et al., 2005). The circumscribed interests are evident in children and adolescents (Sasson et al., 2008), as well as in 2–5 year-olds (Sasson et al., 2011), as

shown by fewer fixations onto faces or people compared to controls when they freely view arrays containing both faces and non-face objects. Moreover, two-year-olds with autism orient to non-social contingencies rather than biological motion (Klin et al., 2009). Taken together, people with autism show a different saliency representation of social stimuli vs. non-social stimuli compared to normals.

When the stimuli are restricted to faces, people with autism show impaired face discrimination and recognition and use atypical strategies for processing faces characterized by reduced attention to the eyes and piecemeal rather than configural strategies (Dawson et al., 2005). In particular, when viewing naturalistic social situations, people with autism demonstrate abnormal patterns of social visual pursuit (Klin et al., 2002). They viewed core feature areas of the faces (i.e., eyes, nose, and mouth) significantly less compared to neurotypical controls (Pelphrey et al., 2002). They showed a greater tendency to saccade away from the eyes when information was present in those regions (Spezio et al., 2007b), but showed increased preference to the location of the mouth (Neumann et al., 2006) and relied primarily on information from the mouth (Spezio et al., 2007a). Eye-tracking data revealed a pronounced influence of active avoidance of direct eye contact on atypical gaze in people with autism (Kliemann et al., 2010). These results again show a different saliency representation of faces in autism compared to normals.

In tasks with top-down instructions such as visual search, attention is guided towards likely targets by a limited set of stimulus attributes such as color and size (Wolfe and Horowitz, 2004, Wolfe, 2012). Several studies have shown superior visual search skills by individuals with autism (Plaisted et al., 1998, O'Riordan and Plaisted, 2001, O'Riordan et al., 2001, O'Riordan, 2004, Kemner et al., 2008), particularly in relatively difficult tasks. This superiority has been attributed to enhanced memory for distractor locations already inspected, and enhanced ability to discriminate between target and distractor stimulus features (O'Riordan and Plaisted, 2001), while it is also arguable that the superiority is due to the anomalously enhanced perception of stimulus features (Joseph et

al., 2009). However, studies investigating the role of top-down excitation and inhibition of stimulus representations in children with autism showed that children with autism did not differ from controls in excitatory or inhibitory top-down control of stimulus representations (O'Riordan, 2000), leaving open the possibility that the autism advantage in visual search (O'Riordan et al., 2001) derived from enhanced bottom-up perception of stimulus attributes (Joseph et al., 2009). It is important to note that the stimuli in the above-mentioned studies are low-level features and inanimate stimuli (e.g., letters and shapes) but not complex images or social stimuli. Using both social stimuli of faces and people and non-social autism special-interest stimuli as search objects, we have demonstrated that people with autism have reduced attention to target-congruent objects in the search array, especially social attention (Wang et al., 2014d). Furthermore, some studies employed visual search to investigate recognition abilities of facial expressions in children with autism and found that faces with certain emotions are detected faster than others (Farran et al., 2011, Rosset et al., 2011). However, when compared with age-matched controls, no significant differences were found anymore.

Social rewards are salient cues (see above) and they share a common neural basis with monetary rewards (Izuma et al., 2008, Lin et al., 2012b). However, people with autism show a disproportionate impairment in learning to choose social rewards, compared to monetary rewards (Lin et al., 2012a). Furthermore, people with autism are not influenced by the presence of an observer in a charity donation task as compared with healthy controls who donate significantly more in the observer's presence than absence, showing insensitivity in people with autism to social reputation (Izuma et al., 2011). People with autism also have reduced preference and sensitivity to donations to people charities compared with donations to the other charities (Lin et al., 2012c). In conclusion, people with autism also show altered saliency representation of more abstract social cues like social rewards.

### 1.6 Amygdala theory of autism

The amygdala is proposed to be part of a neural network comprising the "social brain" (Brothers, 1990), while autism is a neuropsychiatric condition that disrupts the development of social intelligence. It is thus plausible that autism may be caused, in part, by an amygdala abnormality (Baron-Cohen et al., 2000). This hypothesis is supported by the following evidence. The abnormal facial scanning patterns in people with autism (Adolphs et al., 2001, Klin et al., 2002, Pelphrey et al., 2002, Neumann et al., 2006, Spezio et al., 2007a, Spezio et al., 2007b, Kliemann et al., 2010) are rather similar as seen in patients with amygdala damage, who fail to fixate on the eyes in faces (Adolphs et al., 2005), while neuroimaging studies in healthy individuals have shown that amygdala activation is specifically enhanced for fearful faces when saccading from the mouth to the eye regions (Gamer and Büchel, 2009). Besides abnormal eye fixations onto faces, several studies have found reliable, but weak, deficits in the ability to recognize emotions from facial expressions in autism (Law Smith et al., 2010, Philip et al., 2010, Wallace et al., 2011, Kennedy and Adolphs, 2012) (for review, see (Harms et al., 2010)), while on the other hands, patients with amygdala lesions also show abnormal recognition of emotion from facial expressions (Adolphs et al., 1999), and abnormal recognition of mental states from the eye region of faces (Adolphs et al., 2002), providing further support for the amygdala's involvement in autism.

When directly testing the amygdala function in people with autism, the amygdala-mediated orientation towards eyes seen in BOLD-fMRI is reported to be dysfunctional in autism (Kliemann et al., 2012). Activation in the amygdala is strongly correlated with the time spent fixating the eyes in the autistic group (Dalton et al., 2005), but compared to neurotypically developed controls, the amygdala activation was significantly weaker in the people with autism (Kleinhans et al., 2011), consistent with behavioral findings of reduced fixations onto the eyes. Recent studies with single neuron recordings in the human amygdala have even found weaker response to the eyes but stronger response to the mouth in patients with autism compared to control patients (Rutishauser et al., 2013). Despite considerable variability in reports of abnormal face processing in autism, this

evidence largely supports a link between abnormal processing of faces in autism and amygdala function.

Although there is evidence for global dysfunction at the level of the whole brain in autism (Piven et al., 1995, Geschwind and Levitt, 2007, Amaral et al., 2008, Anderson et al., 2010), several studies emphasize abnormalities in the amygdala both morphometrically (Ecker et al., 2012) and in terms of functional connectivity (Gotts et al., 2012). The aberrant gaze patterns in individuals with autism has also been associated with an anatomical link supported by findings of similar gaze fixations, brain activation patterns and amygdala volume in their genetic but unaffected siblings, who demonstrate robust differences compared with typically developing control individuals (Dalton et al., 2007). Amygdala volume can predict gaze patterns in humans (Nacewicz et al., 2006), and even in monkeys (Zhang et al., 2012), consistent with a substantial literature showing structural abnormalities (Bauman and Kemper, 1985, Schumann et al., 2004, Schumann and Amaral, 2006, Amaral et al., 2008, Ecker et al., 2012) and atypical activation (Gotts et al., 2012, Philip et al., 2012) in the amygdala in autism.

Finally, it is important to note that autism is well known to be highly heterogeneous at the biological and behavioral levels and it is arguable that there will be no single genetic or cognitive cause for the diverse symptoms defining autism (Happe et al., 2006). No unanimously endorsed hypothesis for a primary deficit has emerged that can plausibly account for the full triad of social, communicative and rigid/repetitive difficulties (Happe, 2003). It is also worth noting that a bona fide lesion of the amygdala shows no autistic symptoms by clinical examination and autism diagnosis (Paul et al., 2010).

### 1.7 Thesis overview

My thesis investigates how people pay attention when interacting with other people and the environment. I am particularly interested in how socially relevant cues attract and compete for attention and how these cues stand out from non-social stimuli. It is

important to create a vocabulary of social cues including eye gaze, face, head direction, finger gesture, and body posture, in order to investigate whether there is also a saliency map for these social cues. It is also useful to extend the investigation to lexical cues: what sequence of words captures people's attention? Again, some examples are not hard to think of: your name, taboo words, exclamation marks all work fairly well. To break down the question, I am interested in quantifying the set of cues, or set of dimensions, that determine social saliency; in inquiring whether and how these are related to and interact with standard (non-social) "bottom-up" saliency; in exploring to what extent there are individual differences (e.g., in people from different cultures, in people with autism, or in males vs. females); and in understanding the neural mechanisms underlying these processes.

To approach these questions, I used four primary experimental techniques. One was high-resolution eye-tracking, which measures where people look, and in turn indicates where the attention goes. Using eye-tracking with high spatial and temporal resolution, we could understand the dynamic deployment of attention. A second was single-neuron recording in neurosurgical patients, which directly probes the neural correlates of perception and judgments. A third was a lesion approach, in particular in amygdala lesion patients, which tests the causal functional role of the amygdala, a key neural structure of the "social brain". A fourth was testing neuropsychiatric populations such as people with autism, which is able to reveal possible behavioral deficits, especially social deficits, and trace these ultimately to their neural source. Importantly, combining these techniques can answer the same question from different angles and thus have a more holistic view of the question under investigation. For example, testing patients with amygdala lesions and people with autism on the identical task can inform whether amygdala dysfunction will lead to social deficits in autism. Testing a neurological population with concurrent eye-tracking can explore possible deficits of visual attention, while recording single-neurons from neurosurgical patients co-morbid with autism can reveal neuronal mechanisms that lead to behavioral impairment. Therefore, combining different approaches can yield more insights and often leads to more exciting findings. Along with this idea, we have been

conducting single-neuron recordings with concurrent eye-tracking in neurosurgical patients, in order to elucidate the neural mechanisms of one of the most important questions—how do we direct our gaze rapidly to salient objects in our visual environment.

On the other hand, I also employed a variety of experiments to investigate visual attention. In Chapter II, I employed a cued speeded task, in which different locations with respect to the center of optic flow were probed by reaction times in order to study the deployment of spatial attention in optic flow. In Chapter III, I employed a "change detection" protocol, in which subjects were exposed to alternations between two complex scenes that switched back and forth and were entirely identical except for a single change. It is well known that people are remarkably bad at detecting the single item that is changing between the alternating scenes (hence the name, "change blindness"), and this method has been widely used to study which stimuli automatically capture attention and become objects of our conscious awareness. In Chapter IV, we showed degraded emotional faces (a 'bubbles task') and asked subjects to judge emotions shown in the faces. With an adaptive learning algorithm implemented to keep a roughly constant performance, we were able to induce enough errors to investigate how neurons responded in the error trials. In Chapter V, I adopted a standard visual search task and directly compared fixations onto social vs. non-social objects in the cluttered search array. Taken together, I combined diverse experimental strategies of cognitive psychology with multiple neuroscience techniques, together with specific neurological and psychiatric populations, to elucidate the neural mechanisms that come into play. The neurobiological approaches provide information on the brain-end of my question: just like I am interested in determining what it is about social stimuli that captures attention, I am interested in whether there are specialized systems in the human brain for processing social stimuli.

My thesis is organized as follows. In Chapter I, I reviewed the social cues and the neural structures, particularly the amygdala, involved in processing saliency. I argued that people with autism may have altered saliency representation of the visual environment

and that amygdala dysfunction may partially account for this deficit. Then I start to investigate saliency from non-social cues. People constantly move in the environment and pay attention to their locomotion. In Chapter II, I studied the spatial cues that attract attention during locomotion using a cued speeded discrimination task and found that motion cues indicating a forward motion are the strongest to attract attention. However, compared to inanimate objects and cues, people preferentially attend to animals and faces, processing in which the amygdala is thought to play an important role. In Chapter III, I employed a change detection protocol and tested four rare patients with bilateral amygdala lesions. All four amygdala patients showed a normal pattern of reliably faster and more accurate detection of animate stimuli. People not only attend to faces, but also pay attention to others' facial emotions. Humans have a dedicated system to process faces and the amygdala has long been associated with a key role in recognizing facial emotions. In Chapter IV, I studied the neural mechanism of emotion perception and found that single neurons in the human amygdala are selective for subjective judgment of others' emotions. Lastly, normal people pay more attention to faces and conspecifics than to inanimate objects, but people with autism spectrum disorders (ASD) might not. To further study social attention, in Chapter V, I employed a visual search task and revealed a deficit of social attention in people with ASD. This deficit is independent of the amygdala but dependent on task demands. Overall, through visual psychophysics with concurrent eye-tracking, my thesis found and analyzed socially salient cues and compared social vs. non-social cues and healthy vs. clinical populations. Neural mechanisms underlying social saliency were elucidated through electrophysiology and lesion studies.

# Chapter II: Spatial Attention is Attracted in a Sustained Fashion towards Singular Points in the Optic Flow

## 2.1 Overview

Understanding how the attentional system selects goal-directed information and allocates limited resources is an important question in cognitive neuroscience. Since the invention of Posner's cueing paradigm in 1970's, a large amount of knowledge has accumulated. However, the majority of these experiments have been carried out on static images or scenes.

People constantly move in the environment. In this chapter, we focused on the problem of how optic flow influences visual attention. Optic flow, one of the most fundamental properties of any natural visual scene, is associated with self-motion of an organism. It has been studied extensively for more than half a century using psychophysics, electrophysiology, functional neuroimaging, and computational modeling. The quantitative relationship between optic flow and visual attention, however, remains little explored.

With a series of experiments, we showed that expanding optic flow fields, consistent with forward self-motion, attract visual attention strongly in a sustained fashion. We concluded that motion itself, rather than depth structure or temporal evolution of the size of objects, is critical for this effect. We also provided evidence that this attentional effect has a sizable influence even in real life using a change detection paradigm (natural photo stimuli with free eye movements). Continued research on the potency of the FOE and other qualitative aspects of the optical flow to attract attention and gaze will further educate cognitive neuroscientists, engineers and film directors on the forces that shape and control where we attend and look

To understand what is special about salient social cues compared to salient non-social cues, I start to investigate saliency from non-social cues. This first empirical chapter find

a salient motion cue that can be further compared to salient social cues. Importantly, we discuss the neural computations underlying this salient non-social cue—the optic flow, which can be compared to neural computations underlying social saliency. We will directly compare social and non-social cues in Chapter III, where we also test the dependency of social saliency on the amygdala—a key neural structure of the "social brain".

This work has been published as (Wang et al., 2012a).

## 2.2 Summary

While a single approaching object is known to attract spatial attention, it is unknown how attention is directed when the background looms towards the observer as s/he moves forward in a quasi-stationary environment. In Experiment 1, we used a cued speeded discrimination task to quantify where and how spatial attention is directed towards the target superimposed onto a cloud of moving dots. We found that when the motion was expansive, attention was attracted towards the singular point of the optic flow (the focus of expansion, FOE) in a sustained fashion. The effects were less pronounced when the motion was contractive. The more ecologically valid the motion features became (e.g., temporal expansion of each dot, spatial depth structure implied by distribution of the size of the dots), the stronger the attentional effects. Further, the attentional effects were sustained over 1000 ms. Experiment 2 quantified these attentional effects using a change detection paradigm by zooming into or out of photographs of natural scenes. Spatial attention was attracted in a sustained manner such that change detection was facilitated or delayed depending on the location of the FOE only when the motion was expansive. Our results suggest that focal attention is strongly attracted towards singular points that signal the direction of forward ego-motion.

## 2.3 Introduction

The psychophysics of overt and covert attention is a well explored subject with deep roots (Yarbus, 1967). The physiological correlates of visual attention are beginning to be understood at both the single neuron (Colby and Goldberg, 1999, Maunsell and Cook, 2002) and at the brain regional level (Corbetta and Shulman, 2002). This has given rise to detailed computational models of the factors that control the allocation of bottom-up, saliency-driven attention in both artificial and natural static scenes (Itti et al., 1998, Itti and Koch, 2001, Foulsham and Underwood, 2008).

In our daily life, however, the visual inputs to the retina are rarely stationary due to eye, head, and body movements. Furthermore, any object in the scene is embedded in a 3D environment. Looming stimuli on a 2D display are often utilized in laboratory experiments to mimic approaching objects in 3D. Looming stimuli signify biological urgencies or dangers, especially when they approach closer to the body, implying a potential interaction between motion, the projected size of an object on the retina, and attention. Therefore, to fully understand how attention works in a realistic situation, it is necessary to study how the retinal optic flow that accompanies looming stimuli, ego motion and 3D scene structures affect and guide attentional mechanisms.

Looming stimuli typically attract attention and elicit avoidance responses. Many species, including *Drosophila*, locusts, fiddler crabs, fishes, frogs, turtles, chicks, monkeys and humans, persistently dodge looming stimuli (Schiff et al., 1962, Schiff, 1965, Hayes and Saiff, 1967, Tronick, 1967, Bower et al., 1970, Ball and Tronick, 1971, Dill, 1974, Ingle and Shook, 1985, Yamamoto et al., 2003, Nakagawa and Hongjian, 2010, Fotowat and Gabbiani, 2011, de Vries and Clandinin, 2012, Hemmi and Tomsic, 2012). Infant and adult rhesus monkeys manifest persistent avoidance responses to a rapidly expanding but not to rapidly contracting circular shadows (Schiff et al., 1962). This response appears in human infants as well (Ball and Tronick, 1971).

Indeed, the time-to-contact of an approaching object can be precisely estimated (Wang and Frost, 1992, Gray and Regan, 1998, Regan and Gray, 2000), using specialized visual mechanisms (Regan and Beverley, 1978, 1979). Lin *et al* showed that a looming stimulus captures visual attention of an observer only when it would collide with him or her (Lin et al., 2008). This effect was observed even when observers could not consciously discriminate whether or not the object was on a collision path with them (Lin et al., 2009).

While it is well known that a single looming stimulus attracts visual attention among static ones (Franconeri and Simons, 2003), little is known about whether and how visual attention is guided in the presence of an expanding optic flow where many objects loom together. Psychophysical (Schrater et al., 2001), imaging (Field and Wann, 2005) and physiological studies (Laurent and Gabbiani, 1998, Sun and Frost, 1998) provided evidences that expanding optic flow can be decomposed into separate optic features and each optical feature may be individually computed and represented in the brain. Although many conventional psychophysical and electrophysiological studies of ego-motion utilized random dots for expanding optic flow (Saito et al., 1986, Komatsu and Wurtz, 1988, Newsome et al., 1988, Tanaka and Saito, 1989, Duffy and Wurtz, 1991, Graziano et al., 1994, Duffy and Wurtz, 1995, Britten and van Wezel, 1998, Duffy, 1998, Lappe et al., 1999, Morrone et al., 1999, von Muhlenen and Lleras, 2007), such a visual stimulus is less ecological, in the sense that each individual dot does not expand in size and the distribution of the dot size is not consistent with the depth structure in the real world.

Here, we studied how attention is affected by the background visual stimuli that are composed of multiple elements. In a first experiment, we independently manipulated three features of the background dot stimuli: (1) movement of the dots away from or towards a singular point in the visual field (FOE or FOC); (2) expansion or contraction of the dots over time; (3) distribution of the size of the dots in each frame, to make it consistent or inconsistent with the depth structure of the scene in a 3D environment. We created stimuli that lacked or possessed each of the above features (see **Figure 2.1**). We

found the largest attentional effects when all three features were conjoint, emulating a situation where an observer moves towards a fronto-parallel surface in a 3D environment with depth structure.

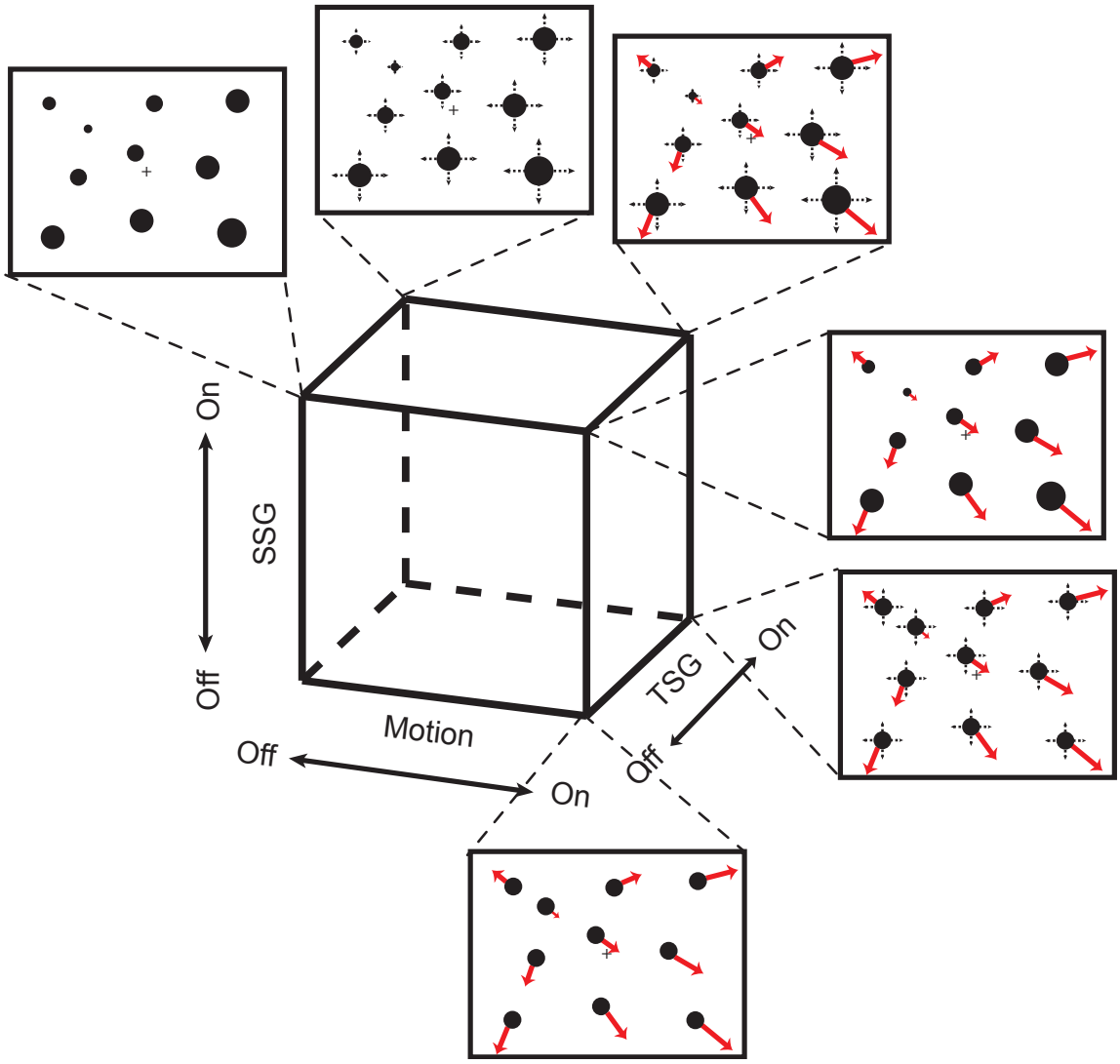


**Figure 2.1. Cube representation of expanding optic flow features.**

Motion, the change in object size over time (or temporal size gradient, TSG), and the spatial depth structure implied by object size distribution (or spatial size gradient, SSG) correspond to one of the three axes of the cube. They can be either on or off. Each corner

of the cube represents a certain combination of features. The specifications of the six conditions that went for testing are illustrated. Red arrows represent the motion. Black horizontal and vertical arrows represent the TSG. Different dot sizes represent the SSG. Note that the sizes of the dots are not to scale.

In a second experiment, we utilized a change detection paradigm using natural scenes (Rensink et al., 1997, Simons and Rensink, 2005, New et al., 2007). We zoomed into or out from a part of a natural scene and manipulated the location of the change in order to test if zooming motion affects spontaneous monitoring of object change. We found strong attentional effects only when the optic flow of the scene expanded (i.e., zooming towards a singular point in the scene) but not when it contracted (i.e., zooming away from the point).

## 2.4 Methods

### *2.4.1 Experiment 1: Speeded discrimination under background dot motion*

#### *2.4.1.1 Subjects and apparatus*

Subjects from the Caltech Community gave written informed consent. The experiments were approved by the Caltech Institutional Review Board. Fifteen subjects (6 females) and one of the authors (SW) participated in the experiments (7 subjects and SW took part in Exp 1a and the other 8 took part in Exp 1b). All subjects had good natural or corrected visual acuity.

Subjects sat 70 cm from a CRT display. The refresh rate of the display was 120Hz and the stimuli occupied the entire display (32° × 24°, visual angle). The stimuli were presented using MATLAB with the Psychtoolbox 3 (Brainard, 1997, Pelli, 1997, Cornelissen et al., 2002) (http://psychtoolbox.org).

We monitored the subjects' eye movements with a non-invasive infra-red eye-tracker (Eyelink-II system, SR Research, Canada) tracking both eyes at 250 Hz. We calibrated the eye tracker with the built-in 13-point grid method. During the main experiment, we repeated the calibration procedure when subjects had several fixation failures in a row.

*2.4.1.2 Task*

We employed a cued speeded discrimination task to quantify how attention is guided by the singular point defined by the flow field of dot motion (i.e., the focus of expansion (FOE) or contraction (FOC)) or by depth structures due to the size distribution of the dots. These features emulate some aspects of the ego-motion related optic flow and the depth structure of the 3D scene. In each trial, a singular point is randomly selected in one of the four quadrants (i.e., top-left (TL), top-right (TR), bottom-left (BL) and bottom-right (BR) corner of the screen). We define congruent, resp. incongruent, trials as those where the target was located in the same, resp. diagonally opposite, quadrant as the singular point. We define the attentional effect as the increase of the mean reaction time (RT) in the incongruent trials compared to the congruent trials.

$$Attentional\ Effect = Mean\ RT_{incongruent} - Mean\ RT_{congruent}$$

Overt eye movements are known to be attracted towards the singular point corresponding to the focus of expansion (Lappe et al., 1998, Niemann et al., 1999). To exclude a possibility that such an effect contaminates our measure of attentional effects, we monitored the gaze location and removed trials with poor fixation. We asked subjects to fixate within 1.6° from the central fixation cross and discarded trials when central fixation was broken.

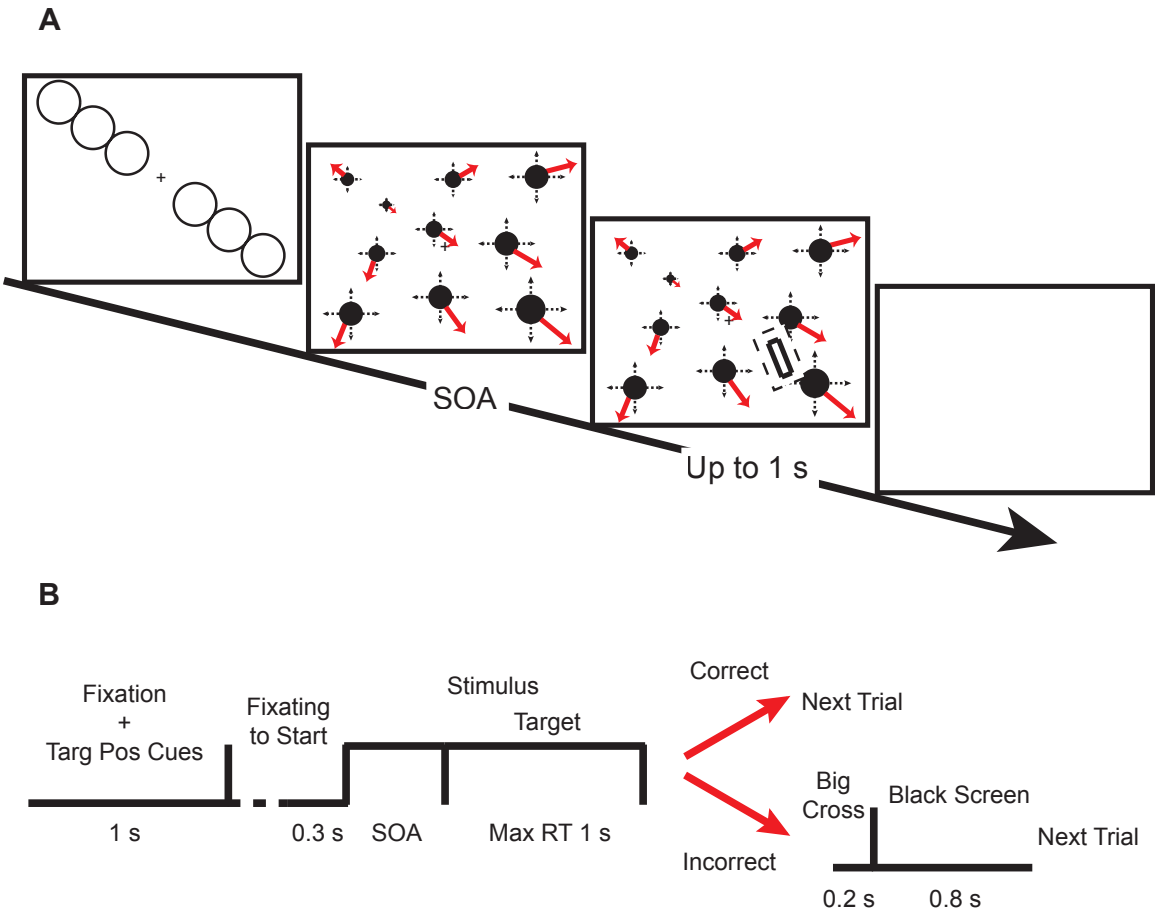


**Figure 2.2. Paradigm for Experiment 1.**

**(A)** Structure and time course of a trial. The dashed line surrounding the tilted target rectangle demonstrates the protection zone. **(B)** A central fixation cross, together with six possible locations of the target, was shown for 1 sec. To initiate a trial, subjects had to fixate 0.3 sec stably within a 1.6 deg radius from the center of the cross. Moving dots appeared subsequently. After various SOAs, a target rectangle appeared. Subjects were asked to discriminate the tilt orientation of the target rectangle by pressing the left or right arrow key as quickly as possible. Subjects had a maximum of 1 sec to respond. If responded correctly, the next trial started. Otherwise, a big cross, indicating the error, appeared, followed by a black screen.

**Figure 2.2** illustrates the task structure. Before each trial, a white central fixation and six thin white peripheral cueing circles (radius 2.6º) were presented for 1 sec. To test if attention is attracted exactly "to" the singular point or "towards" the side of the singular point, we measured the attentional effects at three eccentricities. The circles were positioned along the diagonal of the screen to remind the subjects of the potential locations of a target rectangle. In alternating trials, the potential locations were swapped between top-left vs. bottom-right and top-right vs. bottom-left. There were three potential locations in the top half of the screen and three in the bottom. According to their eccentricity, we refer them as 'far', 'middle' and 'near' cues. The singular point was always located at the 'middle' eccentricity, either in the same or diagonally opposite quadrant (e.g., at the location of the top left middle or bottom right middle circle in **Figure 2.2**). The attentional effect refers to the increase in RT between the inconsistent trials where the singular point was located in the opposite quadrant (e.g., in the top left) with respect to the target (e.g., in the bottom right, possibly near, middle or far locations) compared to the consistent trials where the singular point was located in the same quadrant as the target.

1 sec after the onset of fixation and cues, the subjects' eye positions were monitored. After 0.3 sec of stable fixation, a trial was initiated. The start of the trial was defined as a sudden replacement of the cues with white background dots. After a variable (0, 0.25, 0.5, 0.75, or 1 sec) and randomized stimulus onset asynchrony (SOA) with respect to the onset of the background dots, a target rectangle was presented. The target was a thick white rectangle (0.96º × 3.2º), tilted either 22º left or right. The surrounding area of the target was protected from background dots by a black rectangular zone (3.2º × 5.7º) tilted in the same orientation as the target to ensure its visibility. To facilitate stable fixation, the central fixation cross was also protected from the background dots with a black circular exclusion zone (radius 1.6º). Subjects had to discriminate the orientation of tilt of the target (by pressing the left or right arrow key) within 1 sec from the target onset as fast as possible. When they made a mistake, the data was discarded and the trial was repeated (see below). They were told that any attribute of the background dots was task-irrelevant

and independent of the location or the tilt of the target. They were asked to reduce blinks as much as possible and to keep fixation throughout the trial. **Figure 2.3A** illustrates the distribution of raw RTs and **Figure 2.3B** shows how the attentional effect is defined.

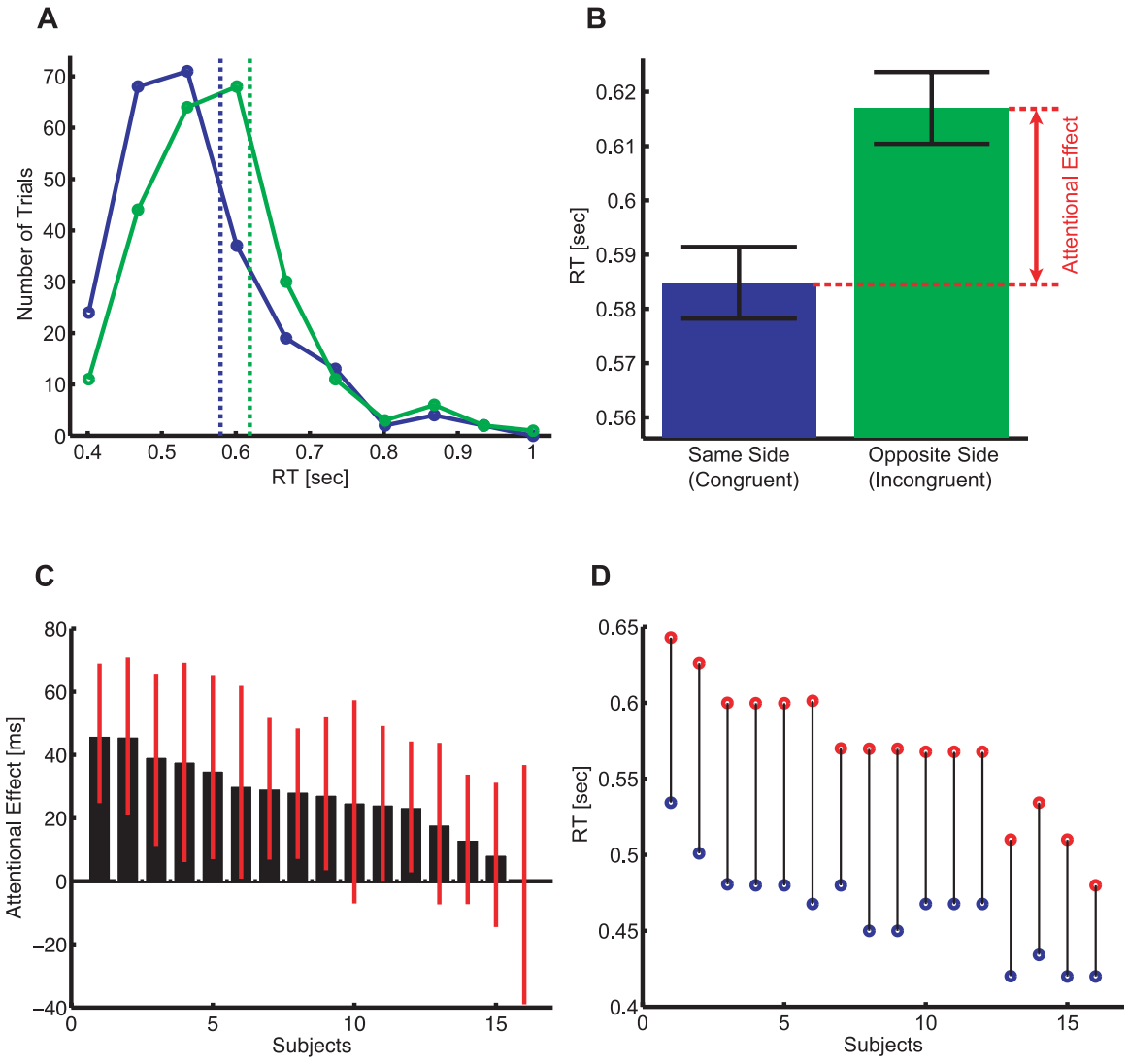


**Figure 2.3. Reaction times (RTs) and the attentional effect.**

**(A)** RT distribution of a single subject. Blue and green colors represent congruent and incongruent trials, respectively. The dashed vertical bars represent the mean RT. **(B)** The mean RT of congruent (blue) and incongruent (green) trials of the single subject shown in **(A)**. The attentional effect for each subject is defined as the increase of the mean RT in

the incongruent trials compared to that in the congruent trials. Error bars denote one s.e.m. across trials. **(C)** Individual results for the attentional effect in Condition 8 (motion = on, TSG = on, SSG = on). The black bars represent the mean attentional effect and the red error bars denote the 5 to 95 percentile intervals. The means and the errors were estimated by the bootstrap method (1000 repetition per subject) (Efron and Tibshirani, 1994). **(D)** Individual results of the 25th and 75th percentile of RT, shown in blue and red, respectively.

*2.4.1.3 Stimuli*

The background visual stimuli, which were irrelevant and non-informative for the discrimination task, consisted of a collection of dots. Across different conditions, we systematically manipulated three features of these dots. (1) The *motion* feature controlled the optic flow of the dots. In the 'motion on' condition, dots moved away from or towards the singular point in the display, which was located in one of the four quadrants. In the 'motion off' condition, the position of each dot remained the same, and did not define the location of a singular point. 2) The *temporal size gradient* (TSG) mimicked looming or receding of each dot. In the 'TSG on' condition, the radius of each dot increased or decreased over time. In the 'TSG off' condition, the radius of each dot remained the same. The TSG did not signify the location of the singular point. 3) The *spatial size gradient* (SSG) implied depth structure in the 3D environment. In the 'SSG on' condition, the size of the dots gradually increased proportionally to the distance from the singular point in the first frame of the stimulus movie. In the 'SSG off' condition, the size of the dots was uniform across the display, thus it did not signify the location of the singular point.

The SSG decides whether a scene structure is present in subsequent frames. To elucidate the relationships between the TSG and SSG, we want readers to note that the SSG decides the expanding rate of the TSG (see **Eqs. 2.2** and **2.3** and their conditions) and

hence ensures whether a scene structure is present or absent across frames: If the SSG is on, the TSG makes the dots grow *proportionally* according to the distance from the singular point (the perspective of the 3D space), preserving the presence of a scene structure; otherwise, the TSG makes the dots grow *uniformly*, preserving the absence of a scene structure.

For all possible 8 combinations of the features, we made sample demo movies. **Table 2.1** and its legend summarize each stimulus. To define the attentional effects, we need a singular point that is defined by the background dots. Therefore, either the 'motion' or the 'SSG' feature has to be on. Accordingly, we used 6 of the 8 conditions in our experiment. When the 'TSG' is turned on, it can enhance the ecological validity of the background dots. However, the TSG did not signify the location of the singular point.

**Table 2.1. The stimulus parameters for each condition.**

TSG and SSG stand for temporal and spatial size gradient, respectively. In Condition 1, uniformly distributed stationary dots are presented. As they do not cue the location of the singular point, this condition was not used in our experiment. In Condition 2, stationary dots with the size gradient imply a 3D scene structure, signifying the location of the singular point. In Condition 3, all the stationary dots expand their diameter at the same rate. As they do not cue the location of the singular point, this condition was not used in our experiment. In Condition 4, static dots are initially arranged with the size gradient, implying a 3D depth structure. Each dot changes its size as if it looms or recedes without changing its position. Condition 5 corresponds to a conventional random dot movie with uniform dot size, which does not change over time. In Condition 6, the initial frame has the size gradient to imply the 3D depth structure. However, each dot does not change its size as it moves, which is unlikely to happen in the real situation. In Condition 7, all the dots have the same size in the initial frame. As they start to move, they change the size together at the same rate, regardless of the distance to the singular point, which is

unlikely to happen in the real situation. In Condition 8, the dots are arranged to have the size gradient to imply the 3D depth structure. Each dot changes its size as it moves so that its diameter is proportional to the distance from the singular point. This is closest to the real situation where an observer moves in a 3D environment, which has the 3D depth structure.

| **Condition (Movie)** | **Motion** | **TSG** | **SSG** | **Motion Speed** | **Rate of Expansion** | **Dot Diameter** | **# Trials** |
|---|---|---|---|---|---|---|---|
| **1** (S1) | Off | Off | Off | 0 | 0 | 0.32º | not tested |
| **2** (S2) | Off | Off | On | 0 | 0 | 0.032º – 0.64º | 60 |
| **3** (S3) | Off | On | Off | 0 | 0.32 º/s | 0.032º – 0.64º | not tested |
| **4** (S4) | Off | On | On | 0 | 0 – 1.68 º/s | 0.032º – 1.60º | 120 |
| **5** (S5) | On | Off | Off | 0 – 29 º/s | 0 | 0.32º | 120 |
| **6** (S6) | On | Off | On | 0 – 29 º/s | 0 | 0.032º – 0.64º | 120 |
| **7** (S7) | On | On | Off | 0 – 29 º/s | 0.32 º/s | 0.032º – 0.64º | 120 |
| **8** (S8) | On | On | On | 0 – 29 º/s | 0 – 0.64 º/s | 0.032º – 0.64º | 120 |

Note that at any given time, the speed of motion of a dot was proportional to the distance to the singular point of the flow field

$$\frac{dx}{dt} = 1.05 \cdot x \qquad \textbf{(Eq. 2.1)}$$

in which $x$ (in the unit of pixels) is the distance from the center of the dot to the singular point and $t$ is in units of seconds. In the case of the contractive motion, the negative sign was added in **Eq. 2.1**. When the SSG was on (a scene structure was present), the dot size was proportional to the distance to the focus of the flow field

$$\theta = 0.02 \cdot x \qquad \textbf{(Eq. 2.2)}$$

in which θ is the diameter of the dot (in the unit of pixels). Thus, the rate of expansion (or contraction) of a dot was proportional to the distance to the singular point

$$\frac{d\theta}{dt} = \frac{d\theta}{dx} \cdot \frac{dx}{dt} = 0.02 \cdot 1.05x = 0.021x \qquad \textbf{(Eq. 2.3)}$$

When the SSG was off (no scene structure), all dots were of the same size across all frames and the rate of expansion (or contraction) of a dot was uniform regardless the distance to the focus of the flow field. The homogenous expansion (or contraction) ensures no size gradient at any frame.

In Condition 4, dots were stationary but kept on expanding in size, as if they were moving. The instantaneous rate of expansion and the dot size were proportional to the virtual distance to the singular point. This was the distance as if the dots kept on moving from their starting position. The virtual moving speed was proportional to the virtual distance to the singular point. Though no dots left the display frame, their rate of expansion increased exponentially over time. Thus, we had to terminate expansion in the middle of the trial at the frame when the largest dot reached 1.6º, in order to keep individual dots distinguishable. The maximum speed (here 1.6 º/s) refers to the rate of expansion of the dot that is farthest from the singular point in the last expanding frame, which had the largest expanding rate among all dots.

Conditions with the contracting motion were the reverse play of the corresponding expanding conditions. Note that, in Condition 4, since expansion stopped in the middle, its corresponding contraction started from the last expanding frame of the expansion, reverse-played all the expanding frames, and stopped and remained stationary with the first expanding frame for the rest of the time in the trial.

Five of the six tested conditions consisted of 120 trials (2 motion directions [expansion vs. contraction] × 6 target locations × 5 SOAs × 2 sides for the singular point [top left (or top right) vs. bottom right (or bottom left)]). Condition 2 (motion off, TSG off and SSG on) was tested only for 60 trials as it did not differ between the expansion and contraction

conditions (**Table 2.1**). The order of trials was fully randomized. Subjects continued until a correct trial was registered for each condition and took a break every 60 trials. In total, there were 480 or 660 correct trials (see below).

We performed two sub-experiments separately. Experiment 1a (480 trials) grouped all four conditions with motion on. We did not replace any moving dots when they moved off the screen. Over time, the background dot density decreased for expansion but increased for contraction. Experiment 1b (660 trials) grouped all six conditions, also without replacing any dots. Note that the density of dots was constant for the two conditions without motion, thus always higher than for those four conditions with motion. Everything else was the same as in Exp 1a.

*2.4.1.4 Data analysis*

We labeled trials with poor fixation (more than 1.6° deviation from the fixation cross or a blink) or incorrect responses (incorrect target discrimination, missing response, response before 0.1 sec or after 1 sec from the target onset) as error trials, and we removed these trials from the RT analysis.

We used MATLAB for t-tests and R (R Foundation for Statistical Computing, Vienna, Austria) for repeated ANOVAs.

*2.4.2 Experiment 2: Change detection with zooming in and out*

In Experiment 1, we tightly controlled stimuli and eye movements. Experiment 2 seeks to relax these constraints by using movies of natural scenes as stimuli and allowed eye movements in a change detection task.

*2.4.2.1 Subjects and apparatus*

Fifteen naive male subjects, none of whom took part in Experiment 1, participated. Subjects sat 80 cm from the display with a chin rest to minimize head movements. The refresh rate of the display was 50Hz and the images occupied the entire display (29° × 22°). Eye movements were not recorded.

*2.4.2.2 Zooming algorithm*

In Experiment 2, we chose to study the effects of attention based on natural scene images. As a consequence, we focused on Condition 8 in Experiment 1, where motion, TSG and SSG were all on. As a control, we also used Condition 1, where all three features were off. For Condition 8, singular points coincide with the FOE or FOC.

We used a zooming algorithm, based on the OpenGL function in the Psychtoolbox-3. During expansion, the camera speed of zooming was kept constant over time. The speed of expansion at each pixel was proportional to its distance (in the unit of visual angles) to the FOE and ranged from 0 to 5.4 °/s. Denoting the location of a pixel $p$ at time $t$ during the expanding period as $p(t)$, our zooming algorithm computes $p(t) = f + z^t(p(0) - f)$ where $f$ denotes the location of the FOE and $z$ denotes the zoom speed, thus the $p$ increases exponentially as $t$ increases. The zoom speed, $z$, was fixed at 2 [°/s]. The same algorithm was used for the contraction but with negative $t$.

*2.4.2.3 Procedure*

Subjects pressed a button to initiate a trial. Each trial started with a 0.6 sec movie sequence consisting of 15 frames, which was replaced by a uniform gray field for 0.28 sec. The last frame of the sequence that contained a single noticeable change was then presented for 0.6 sec. After this static image, another 0.28 sec blank period followed. A

complete cycle of this movie-blank-image-blank sequence was repeated until subjects pressed a space bar, indicating that they were sure that they have seen the change explicitly. When the space bar was pressed during the movie presentation or during the blank period immediately after the movie, the last frame of the movie was presented again on the screen and subjects had to indicate the change location via the mouse. When it was pressed during the stationary image or the blank period immediately following it, the stationary image was presented, on which subjects localized the change. This procedure prevented any visual transients that could be used to localize the change. If subjects could not detect the change after 52.8 sec, the trial was stopped.

*2.4.2.4 Stimuli*

For a given change detection image pair, we created 4 movie sequences for 4 different conditions corresponding to the FOE and FOC being close or far away from the location of the change. For example, when an image pair contained a change within a top-left quadrant, we created 4 movies as follows: 1. FOE-on by zooming into the top-left corner, 2. FOE-off by zooming into the bottom-right corner, 3. FOC-on by zooming out from the top-left corner, and 4. FOC-off by zooming out from the bottom-right corner. These sequences were carefully constructed such that the last frames of the 4 movies were identical. The stationary image that contained the change was also identical across all conditions. Thus, the size of the objects in the last frame of the movie and the critical change frame was identical across conditions, rendering the difficulty of the search comparable. The 5th condition, a stationary control, was created by presenting the last frame of the movie for 0.6 sec.

We prepared 55 image pairs (5 of them were used for practice). We presented each image pair to a particular subject in one of five conditions. In other words, each subject was tested ten times in each condition, but each subject only saw a given image pair once in one condition. To achieve balance across subjects, we created 3 groups of 5 subjects and

assigned image pairs to each group such that each image pair was seen under one experimental condition by only one member of the group. For the data analysis, the results from one group were considered as a single data point. To reflect this grouping process, the error bars are the standard deviation divided by the square root of the number of groups, which is 3.

#### *2.4.2.5 Data analysis*

Prior to data collection, we defined a region of acceptable click location for each image pair by delineating a rectangular area that encompassed the change. Out of 750 trials, 701 clicks (93.5%) were within the pre-defined areas and only 10 clicks (1.3%) were outside of the rectangle. In 39 trials (5.2%), subjects did not click any location within 52.8 sec.

A one-way ANOVA was performed on log-transformed RTs because RTs were heavily long-tailed as can be seen from the cumulative histogram, whose x-axis is the logarithm of RT. For display purpose, the means of log-transformed RT as well as the error bars were transformed back into a linear scale by exponentiation. We used non-parametric Kolmogorov-Smirnov test for post-hoc comparisons.

## 2.5 Results

### *2.5.1 Experiment 1: Speeded discrimination under background dots*

#### *2.5.1.1 Motion is a strong cue while TSG and SSG act as auxiliary cues*

Our main interest in this paper is how visual attention is attracted and guided by motion, looming stimuli, and depth structure. These cues represent some aspects of the visual input during navigation within the 3D environment. Our expansive or contractive motion as well as the size distribution of dots (the spatial size gradient, SSG) defined a singular point in the display, which may or may not attract attention. When the size of the dots

changed over time (the temporal size gradient, TSG), they did not signal the location of the singular point but they assisted the ecological interpretation of the motion and depth structure of the dots. We measured whether the singular point defined by the motion and/ or SSG attracted covert attention by measuring RTs in the discrimination task and by defining the attentional effect as the RT increase in the trials where the target was located in the opposite (or incongruent) side of the display from the singular point compared to where they were located in the same (or congruent) side (**Figure 2.3B**). Significant attentional effects were highly robust and measurable in almost all subjects as shown in **Figure 2.3C** (in Condition 8), with a confidence interval estimated by the bootstrap method (Efron and Tibshirani, 1994). We demonstrated raw RT range for each subject in **Figure 2.3D**.

Comparing the overlapping conditions between Exp 1a and 1b (four conditions with motion on), we did not find any difference in the attentional effect (four-way ANOVA; Experiment [1a vs. 1b] (between-subjects factor) X motion direction [expansive vs. contractive] X TSG X SSG: the p-value for the main effect of the Experiment was > 0.32). Post-hoc two-tailed t-tests confirmed no difference between each pair of overlapping conditions (all p-values were above 0.05). This analysis confirmed that our experiment was replicated by two independent samples.

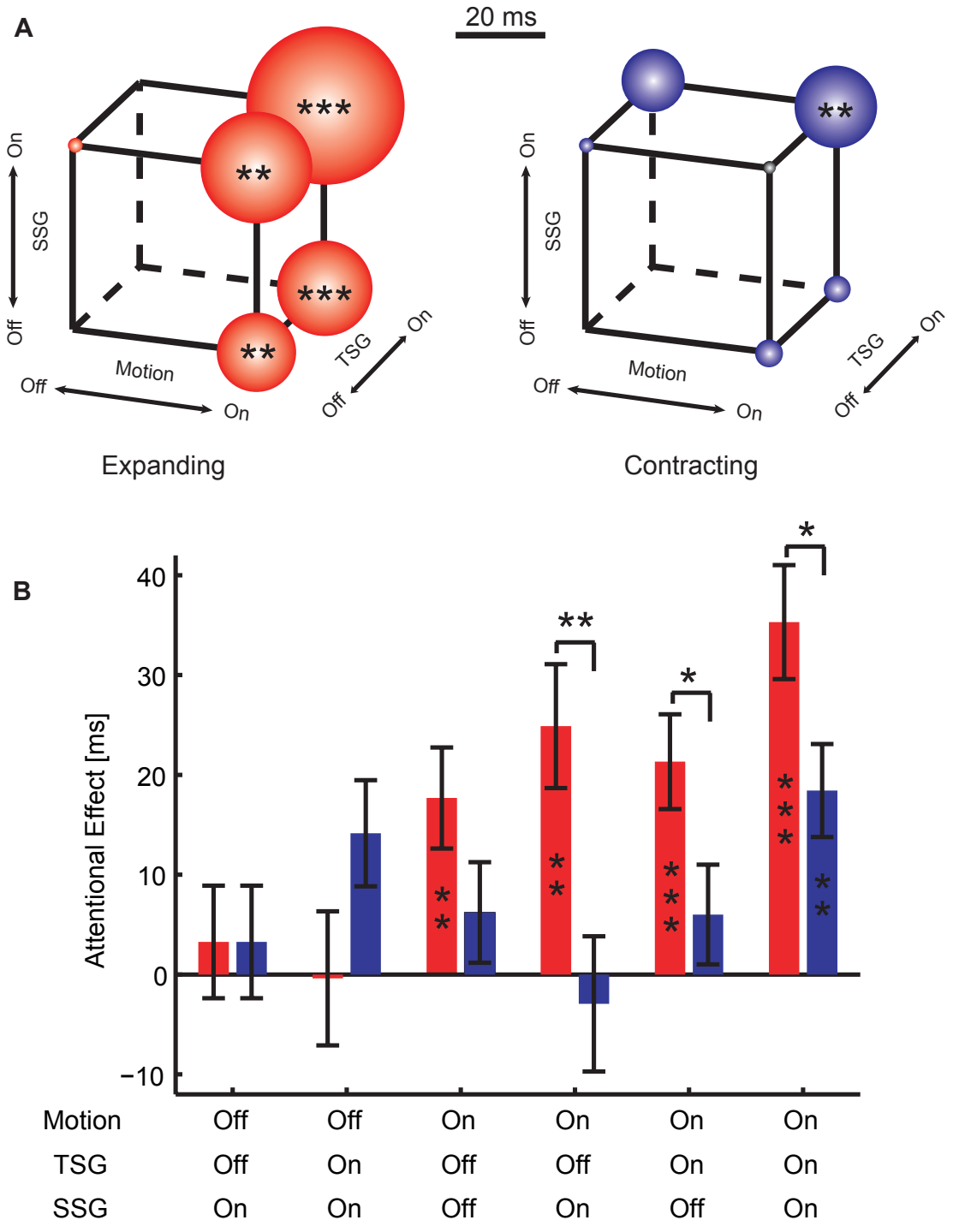


**Figure 2.4. The attentional effects of the expanding and contracting optic flow.**

**(A)** Cube representation of the attentional effect. Red and blue colors represent the positive attentional effect for the expanding and contracting conditions, respectively.

Black color represents the negative attentional effect. The area of the balls corresponds to the absolute magnitude of the attentional effect (the scale indicates 5 ms). p-values from two-tailed t-test against zero are represented by *, ** and *** indicating $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively. **(B)** Bar representation of the attentional effects. Red and blue bars are for the expanding and contracting conditions, respectively. The stars indicating the level of the p-values (*, ** and ***) from two-tailed t-test against zero are shown within the bars. Significant differences between the expanding and contracting conditions are denoted by stars above the bars. Error bars denote one s.e.m. across subjects.

In **Figure 2.4A**, we represent the attentional effects as the area of balls in a cube configuration, using the motion, TSG and SSG as the three axes. Note that the condition with motion off, TSG off and SSG on (a static perspective image) was identical for expanding and contracting motion.

As the first analysis, we tested if each condition produced reliable attentional effects (two-tailed t-tests against 0). For all the conditions with the expanding motion, we observed significant attentional effects (above 0, all ps < 0.01, the 4 red bars on the right in **Figure 2.4B**). Their magnitudes ranged from 18 to 35 ms for motion on, but less than 5 ms for motion off. With the contracting motion, we found significant attentional effect only when combined with the TSG and SSG [motion = on, TSG = on, SSG = on] ($p < 0.01$, 18 ms, the rightmost blue bar in **Figure 2.4B**). Separately for expanding and contracting motion, we compared the attentional effects between motion on and off, collapsing TSG and SSG, and found a highly significant difference for the expanding (paired t-test; $p < 0.0005$) but not for the contracting conditions ($p = 0.78$). We conclude that focal attention is critically captured by the focus of expansion signaled by the expanding motion, but not by the contracting motion.

Second, we investigated the effects of the TSG with repeated ANOVAs. We used a subset of balanced data from Exp 1b, with [motion = on/off, TSG = on/off] (i.e., the data points in the upper plane of the cube in **Figure 2.4A**). For the expanding condition, we found the main effect of motion to be significant (two-way ANOVA, $p = 0.016$), but neither for the main effect of TSG ($p = 0.86$) nor the interaction between motion and TSG ($p = 0.80$). For the contracting condition, we observed the main effect of TSG ($p = 0.0047$), but neither the main effect of motion ($p = 0.95$), nor the interaction between motion and TSG ($p = 0.55$). We conclude that the TSG was a critical feature to capture attention for the contracting but not for the expanding optic flow.

Third, we investigated the effects of the SSG on attention by confining the analysis to motion on, *i.e.*, the right side of the cube in **Figure 2.4A**. Three-way, within-subjects repeated ANOVA (motion direction [expansion vs. contraction] X TSG X SSG) revealed significant main effects of motion direction ($p < 0.01$) and TSG ($p = 0.038$) but not SSG ($p = 0.09$). There was a significant interaction between TSG and SSG ($p = 0.044$), but neither between motion direction and TSG ($p > 0.65$), motion direction and SSG ($p > 0.19$), nor a 3-way interaction ($p > 0.18$). To understand the nature of the interaction between TSG and SSG, we performed post-hoc two-way ANOVAs (TSG X SSG) separately for the expanding and contracting conditions. For the expanding condition, we found a significant main effect of SSG ($p = 0.021$) but not TSG ($p = 0.27$) or interaction ($p > 0.41$). For the contracting condition, we found a significant main effect of TSG ($p = 0.043$) and interaction ($p = 0.026$) but not the main effect of SSG ($p = 0.76$). Although significant, these effects tend to be small in magnitude (~10 ms) compared to the effects caused by the presence of motion itself (~28 ms; **Figure 2.4A**). To conclude, the SSG played a significant role only in the expanding motion condition.

Fourth, to further characterize the importance of the TSG and SSG, we compared the attentional effects in Condition 8 (motion = on, TSG = on, SSG = on) with the conditions that lacked only the TSG and/or SSG. The effect increased with both features [motion = on, TSG = off, SSG = off] (~18 ms, pairwise two-tailed t-test, $p = 0.026$) and [motion =

on, TSG = on, SSG = off] (~14 ms, p = 0.01), but not with [motion = on, TSG = off, SSG = on] (~10 ms, p = 0.23). This pattern seems consistent with an additive attentional effect of the TSG and SSG. To summarize, we found that the effects of the TSG and SSG were about 1/3 of the attentional effect due to the motion cue alone.

Fifth, we compared the magnitude of the attentional effects between the expanding and contracting conditions (**Figure 2.4B**, the red and blue bars are for the expanding and contracting motion, respectively). We found that the expanding motion attracted more attention only when the TSG and/or SSG were on (paired t-test, $p < 0.05$) but not when both the TSG and SSG were off ($p = 0.12$). This indicates an interdependence of motion direction, TSG and SSG. The TSG and SSG helped the expanding motion to attract attention. This further buttressed our claim that the motion plays the dominant role in the attentional effect while the TSG and SSG played an auxiliary role for the attentional attraction due to the expanding motion.

*2.5.1.2 Target eccentricity and SOA on the attentional effects*

So far, we showed that Condition 8 (motion = on, TSG = on, SSG = on), which is closest to the ecological condition, strongly attracts attention towards the singular point. In this section, we characterize the spatiotemporal characteristics of the attentional cueing (**Figure 2.5**). We analyzed the influence of the target eccentricity and SOA on the attentional effects separately for the expanding (**Figure 2.5 A,C,E** and **G**) and contracting (**Figure 2.5 B,D,F** and **H**) motion, by averaging across all conditions (**Figure 2.5 A–D**) or by focusing on Condition 8 (**Figure 2.5 E–H**).

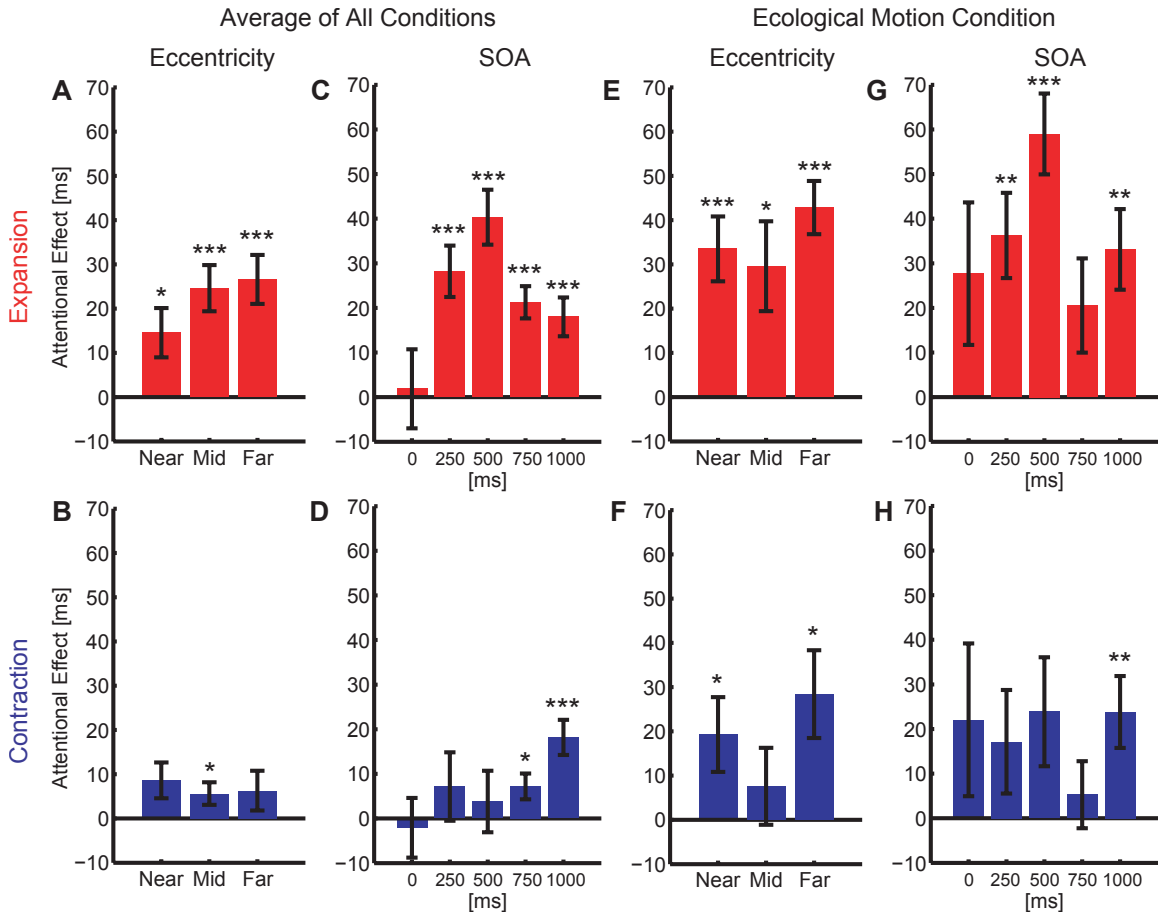


**Figure 2.5. Dependency of the attentional effects on target eccentricity and SOA.**

Dependency of the attentional effects on target eccentricity **(A,B,E,F)** and SOA **(C,D,G,H)** for all conditions averaged **(A–D)** and for Condition 8 (motion = on, SSG = on, TSG = on) **(E–H)**. Red bars are for the expanding conditions **(A,C,E,G)** and blue bars are for the contracting conditions **(B,D,F,H)**. The level of p-values from two-tailed t-test against zero are shown above the bars by *, ** and *** for $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively.

With the data averaged across all conditions **(Figure 2.5 A–D)**, the results of the three-way, within-subjects ANOVA (motion direction [expansion vs. contraction] X eccentricity [near, mid and far] X SOA [0, 250, 500, 750 and 1000 ms]) revealed

significant main effects of motion direction ($p = 0.013$) and SOA ($p = 0.0022$), but not eccentricity ($p = 0.16$). A significant interaction was observed between motion directions and SOA ($p = 0.012$), but no other significant interactions were observed. The lack of the main effect of the eccentricity implies that attention was attracted towards the side of the singular point and that attention was not attracted to the exact location of the singular point.

To characterize the nature of the interaction between motion direction and SOA, we performed post-hoc, one-way, within-subjects ANOVA on SOA separately for the expanding and contracting conditions (collapsing across the eccentricities): SOA dependence came from the expanding ($p = 7.4\times10^{-5}$, **Figure 2.5C**) but not contracting conditions ($p = 0.19$, **Figure 2.5D)**.

The expansive motion captured attention as soon as 250 ms after stimulus onset ($p < 0.001$ for t-tests testing that the attentional effects were above 0 at SOA = 0.25, 0.5, 0.75 and 1s, $p = 0.84$ at SOA = 0s; **Figure 2.5C**). Unlike other exogenous attentional cues, such as a flash of a bright square, the expansive motion attracted attention rapidly and in a sustained manner (see **Discussion)**. In contrast, the contracting motion field took a long time to capture attention ($p < 0.05$ at SOA = 0.75 and 1s, $p > 0.36$ at SOA = 0, 0.25 and 0.5s; **Figure 2.5D**). This slow orienting process is unlikely to be caused by bottom-up stimulus factor, suggesting a possible difference in the neuronal mechanisms of attentional capture for the expansive and contractive motions.

We repeated the above analysis, focusing on Condition 8 (motion = on, TSG = on, SSG = on), which is closest to the ecological condition, as these experiments produced the largest attentional effects (**Figure 2.5E–H**). The results were similar to those collapsing over all conditions: a marginally significant main effect of motion direction ($p = 0.061$) but not eccentricity ($p = 0.11$). Here we did not observe a significant dependency on SOA ($p = 0.16$). We found that the expanding motion started to attract attention for an SOA as short as 250 ms and lasting until 1s (all ps $< 0.05$ except $p = 0.071$ for SOA = 0.75s;

**Figure 2.5G**) while the contracting motion started to attract attention with the long SOA ($p < 0.01$ at SOA = 1s; **Figure 2.5H**).

*2.5.1.3 Laterality of the attentional effects*

We found an unexpected and sizable effect of laterality of the singular point. Averaging across all conditions, the attentional effects were stronger when the singular point appeared in the right visual field than the left, but they were similar between the upper and lower visual field: with a three-way within-subjects ANOVA (motion direction [expansion vs. contraction] X the horizontal [left vs. right] X the vertical [upper vs. lower] position of the singular point), we found significant main effects of motion direction ($p = 0.013$) and the horizontal ($p = 0.021$) but not the vertical position of the singular point ($p = 0.41$; **Figure 2.6A,B**).

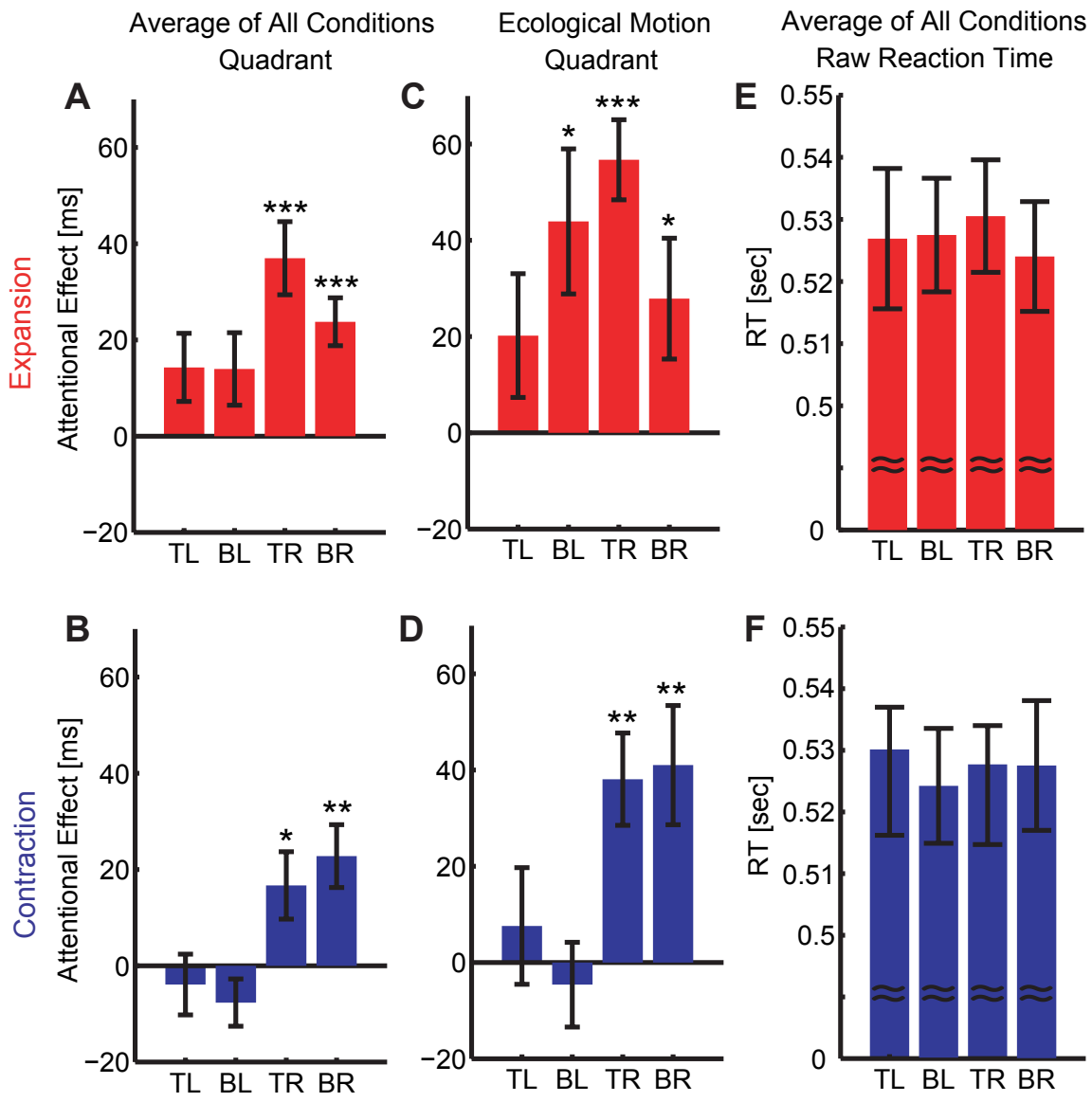


**Figure 2.6. The size of the attentional effects depends on the side of the singular point.**

**(A,B)** The attentional effects averaged for all conditions and **(C,D)** for Condition 8 (motion on, TSG on, and SSG on). **(E,F)** The effects are not explained by the difference in raw RTs (all conditions averaged). Red bars are for the expanding conditions **(A,C,E)** and blue bars are for the contracting conditions **(B,D,F)**. p-values from two-tailed t-test against zero are shown above the bars by *, ** and *** for $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively. TL: top-left. TR: top-right. BL: bottom-left. BR: bottom-right.

The most ecological motion (Condition 8, motion, TSG and SSG all on) also revealed this left-right asymmetry of the attentional effect: with a three-way within-subjects ANOVA, we found a significant main effect of the horizontal position of the singular point ($p = 0.030$) but not of the other factors (motion directions, $p = 0.080$ and the vertical position $p = 0.63$; **Figure 2.6C,D**).

This effect is not an artifact of using the right hand for response; the reaction time for target detection was comparable when the singular point appeared in any of the quadrants (three-way within-subjects ANOVA (motion direction [expansion vs. contraction] X the horizontal X the vertical position of the singular point, the main effect of motion directions: $p = 0.35$; horizontal: $p = 0.85$; vertical $p = 0.62$, no significant interactions (all ps > 0.12; **Figure 2.6E,F**)).

When we grouped the trials according to the horizontal position of the target and repeated the same analysis, we did not find any significant effects.

*2.5.1.4 Analysis of error trials*

There were five types of errors (1. fixation-break, 2. wrong discrimination, 3. missing response, 4. too early response, and 5. too late response). **Table 2.2** summarizes the error rates for each condition. All errors except fixation-break were well controlled below 5%, showing that subjects well understood and concentrated on the task. The mean rate of fixation-break was 19%. In this task, constant fixation was not easy and subjects were frequently reminded to keep a good fixation and reduce blinks. Condition 4 (motion = off, TSG = on, SSG = on) had a slightly higher error rate than the rest of conditions (two-way ANOVA (error types X conditions), the main effect of error types: $p < 2\times10^{-16}$, the main effect of conditions: $p = 0.017$), indicating that this condition was slightly more difficult to maintain constant fixation than others. No interaction was found between the

error types and conditions (p = 0.36). Separate analysis within Exp 1a and Exp 1b (two-way within-subjects ANOVA) revealed the same effect.

**Table 2.2. The error rates (in percentage) for each condition.**

| Condition | Motion | TSG | SSG | Fixation Break | Wrong Discrimination | Missing Response | Too Early Response | Too Late Response |
|---|---|---|---|---|---|---|---|---|
| **1** | Off | Off | Off | - | - | - | - | - |
| **2** | Off | Off | On | 20.6 | 2.17 | 0.116 | 0 | 0.614 |
| **3** | Off | On | Off | - | - | - | - | - |
| **4** | Off | On | On | 26.3 | 2.4 | 0.236 | 2.49 | 2.87 |
| **5** | On | Off | Off | 18.1 | 1.9 | 0.211 | 0.0324 | 0.541 |
| **6** | On | Off | On | 17.3 | 1.86 | 0.0403 | 0 | 0.792 |
| **7** | On | On | Off | 15.7 | 2.11 | 0.125 | 0 | 0.722 |
| **8** | On | On | On | 19.2 | 1.95 | 0.171 | 0 | 0.743 |
| **Average** | - | - | - | 19 | 2.02 | 0.134 | 0.265 | 0.898 |

*2.5.2 Experiment 2: Change detection with zooming in and out*

In Experiment 2, we investigated if the attentional effects revealed in Experiment 1 can be replicated in a more naturalistic setting. For this purpose, we used natural scene images and allowed subjects to move their eyes in a change detection paradigm **(Figure 2.7A)**.

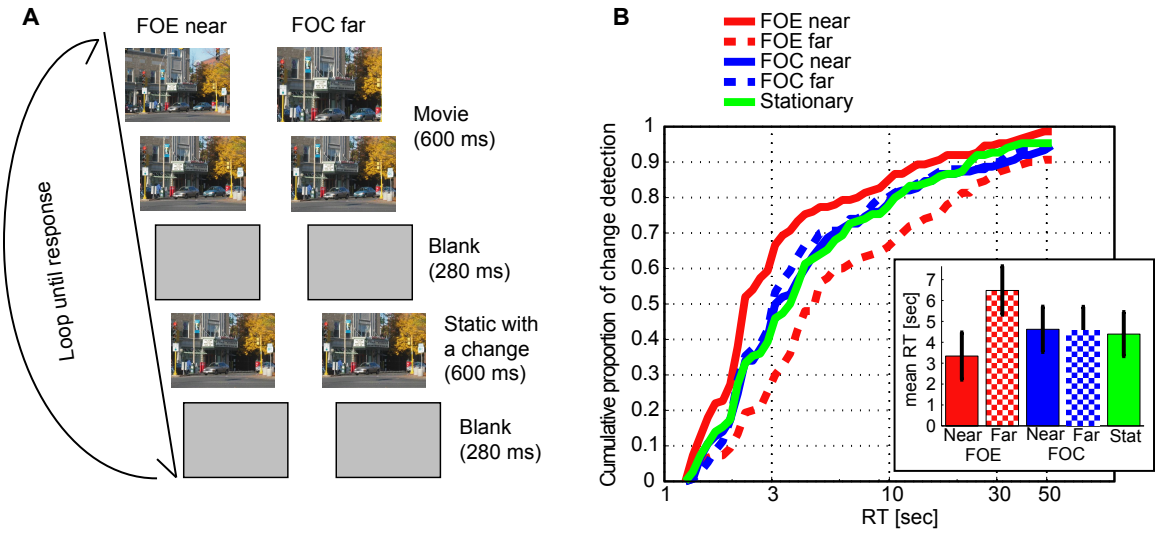


**Figure 2.7. Expansion but not contraction influences the speed of change detection.**

**(A)** For Experiment 2, a 0.6 sec movie expanded or contracted with the associated FOE or FOC located either at the corner of the same quadrant as the change or at the opposite corner (in this example, the car on the bottom right disappeared). After a 0.28 sec blank period, a stationary image with a single noticeable change from the last frame of the movie was presented for 0.6 sec, followed by another 0.28 sec blank. This loop was repeated until subjects responded. **(B)** The cumulative detection probability as a function of RT (log scale). When the FOE was close to the location of the change, detection was facilitated, while when the FOE was far away, it interfered with change detection. Contraction did not affect change detection, compared to the control stationary condition. (Inset) Mean RT (error bars are for s.e.m.).

As was expected from Experiment 1, subjects detected the change more quickly when it was close to the FOE (**Figure 2.7B**). Mean RTs across conditions (FOE-on: $3.34 \pm 1.13$ sec; FOE-off: $6.48 \pm 1.16$ sec; FOC-on: $4.62 \pm 1.08$ sec; FOC-off: $4.56 \pm 1.14$ sec; stationary: $4.39 \pm 1.06$ sec; mean ± standard error) differed significantly (one-way ANOVA, $p < 10^{-5}$). A post-hoc Kolmogorov-Smirnov test confirmed (i) that the RT was strongly influenced by the location of the FOE ($p < 10^{-9}$) but not by the FOC ($p > 0.9$),

(ii) that the RT in the FOE-on condition was faster than any other conditions ($p < 0.01$ for all comparisons) and (iii) that the RT in the FOE-off condition was slower than any other conditions ($p < 0.02$ for all comparisons). FOC-on, FOC-off, and stationary conditions did not differ among each other ($p > 0.27$ for all). We conclude that zooming into the change (FOE), but not zooming away from the change (FOC), guides covert and overt attention.

## 2.6 Discussion

In two separate experiments, visual attention was rapidly attracted in a sustained manner towards the focus of the expanding motion. The effect was largely specific to the expanding motion and was weak or absent for the contracting motion. The motion cue played a key role in capturing attention while the temporal evolution of object size (TSG) and depth structure (SSG) played an auxiliary role (Experiment 1). Change detection was substantially slowed or facilitated depending on the location of the FOE (focus of expansion), but not FOC (focus of contraction), relative to the changed object (Experiment 2).

### *2.6.1 Attention is attracted towards the singular point defined by the expansive, but not contractive, motion*

Throughout our experiments, we found a profound asymmetry between the strong attentional effects of expansive motion and the weak or inconsistent effects for contractive motion. This ruled out a possibility that the slower speed vector fields around the singular point attracted attention since both the contractive and expansive motion had slower motion field near the singular point, yet much larger attentional effects were found in the expanding motion. Our result is consistent with the asymmetric ease in visual search (e.g., it is easy to find an expanding object among receding ones and it is difficult

to find a receding object among expanding ones (Takeuchi, 1997)). Likewise, cortical neurons that prefer expanding radial motion outnumber neurons that prefer contracting motion (Saito et al., 1986, Graziano et al., 1994). The attentional and neuronal bias towards expansive motion might have been shaped through evolution reflecting ecological conditions, as contractive motion occurs only when moving backward, which happens much less often in the natural environment. This conjecture is supported by developmental studies of babies that prefer to look at expansive rather than contractive motion; even more, the developmental onset of expansive motion preference starts even before babies start moving by themselves and experiencing expansive optic flow (Brosseau-Lachaine et al., 2008), suggesting an innate bias towards expansive motion. Furthermore, in the real world, animals manifested a fine-tuned neural system to perceive expanding optic flow and control motion, for example during pigeon perching (Lee et al., 1993), fly landing (Wagner, 1982), gannet plunge-diving (Lee and Reddish, 1981) and during human landing from a fall (Sidaway et al., 1989), steering (Land and Lee, 1994) and braking a car (Lee, 1976, Yilmaz and Warren, 1995). Abundant psychophysical (Morrone et al., 1999) and physiological (Duffy and Wurtz, 1991, Orban, 1992) studies have shown that these expansionary motions are processed by specialized mechanisms in mammalian visual systems.

### *2.6.2 Sustained attentional effects*

Consistent with von Muhlenen & Lleras (von Muhlenen and Lleras, 2007) who used random dot motion, we found that the expanding optic flow field rapidly attracted attention towards the FOE in a sustained manner. While many exogenous cues attract attention, these cues tend to attract attention only during the initial several hundred milliseconds, usually acting in a repelling fashion after ~500 ms, a phenomenon called ‘inhibition-of-return (IOR)’ (Klein, 2000), which is believed to facilitate orientation towards novel locations, facilitating foraging and other search behaviors.

In Experiment 1, the attentional effects were sustained up to 1 sec, which suggests that IOR is not operating for the attentional mechanisms with the expansive motion. In Experiment 2, the attentional effects even amounted to 3 sec, implying that IOR was not operating over long period of time in this paradigm. On this point, we invite readers to look at our demo movies. We expect them to feel like they tend to look at the location around the FOE repeatedly although they know that there is no change to be detected around that location. The lack of IOR in our expansive motion implies that attention towards the FOE may be important in coordinating behavior by aligning the direction of gaze, head and body.

### *2.6.3 Mechanisms of computation of the FOE*

Optic flow is processed in a network of visual motion areas, V1, V3, MT, medial superior temporal area (MST) (Saito et al., 1986, Duffy and Wurtz, 1991, Graziano et al., 1994, Britten and van Wezel, 1998), the ventral intraparietal sulcus (VIP) (Schaafsma and Duysens, 1996, Bremmer et al., 2002, Zhang et al., 2004), area 7a and STP (for a review, see (Britten, 2008)). Recordings from neurons in the ventral intraparietal sulcus (VIP), which receives strong input from MSTd, also revealed strong tuning to the optic flow (Schaafsma and Duysens, 1996, Bremmer et al., 2002, Zhang et al., 2004). A recent fMRI study compared the response characteristics of these two regions and found that VIP is more consistent with the computation of FOEs than MSTd (Wall and Smith, 2008).

Given the known strong effects of attention in VIP (Colby and Goldberg, 1999, Maunsell and Cook, 2002) and other parietal areas, it is possible that the attentional effects of the FOE are mediated by neurons in this region. These overlapping regions for computing the FOE and attention raise the question of to which extent the FOE attracts focal attention and, if so, whether this depends on the task at hand.

*2.6.4 Advantage of our stimulus design*

Conventional studies often used homogeneous random-dot patterns without any size change over time (TSG off) and/or uniform size distribution over space (SSG off). We found that the size change over time (TSG on) and the size distribution over space (SSG on) maximize the attentional effect of the expansive motion. Future studies might be better able to simulate ego-motion in the real world by including temporal evolvement (TSG) and depth information (SSG).

Our decomposition paradigm begs a question: how is each optical feature represented in the brain? Human psychophysical studies showed perception of visual expansion without optic flow (Schrater et al., 2001), indicating that judgment of size (or scale) change is independent of local translational motion. Human fMRI studies have also tried to separate and control optical variables, such as time-to-contact, image expansion, motion in depth and rate of gap closure, in the case of looming (Field and Wann, 2005). In future research, it will be important to examine the neural mechanisms of each feature.

*2.6.5 Laterality effects of attention*

Unexpectedly, we found the attentional effects strongly depend on the laterality of the singular point (**Figure 2.6A-D**): when the singular point appears in the right visual field, the attentional effects became roughly twice as large (30ms vs. 15ms, for the expansion). Behaviorally, lateralized effects have been reported for the sensory and cognitive processing of language, face, and emotion (MacNeilage et al., 2009). Recent studies also report laterality effects in frogs, chickens, birds and monkeys, implying the evolutionary origin of the laterality (Vallortigara and Rogers, 2005). Laterality has been also well documented for the attentional mechanisms (Fox et al., 2006). In normal subjects, a strong asymmetry in the attentional resolution has been reported between the upper and lower visual field (He et al., 1996). While it is unclear why spatial attention is more strongly captured when the singular point locates in the right visual field, our findings

might be related to the ancestral origin of hemispheric lateralization for detecting unexpected predators vs. performing routine jobs (MacNeilage et al., 2009).

## 2.7 Conclusion

In this paper, we explored the attentional effects of the singular point defined by motion, object expansion and 3D depth structure. We found the strongest attentional effects in the condition that incorporates expansive motion with the 3D depth structure, which is most compatible with the visual input during forward ego motion in the 3D environment. While extensive studies have been performed on the mechanisms of attention, relatively less is explored on how attention is guided in the real 3D natural environment with the observer motion. Accordingly typical computational models of attention do not incorporate the factors we investigated here (Itti et al., 1998, Itti and Koch, 2001, Foulsham and Underwood, 2008). Our experiments revealed that expanding motion that accompanies forward ego motion is likely to guide attention strongly in everyday life. Further studies will be necessary to uncover how attention is guided and how we perceive the world in the natural environment.

## 2.8 Acknowledgements

We thank Claudia Wilimzig for offering us the stimuli for Experiment 2, Shima Hajimirza for collecting the data for Experiment 2 and Elnaz Nouri for performing the preliminary version of Experiment 1.

# Chapter III: Preferential attention to animals and people is independent of the amygdala

## 3.1 Overview

In Chapter II, we studied non-social spatial cues that attract attention during locomotion using a cued speeded discrimination task and found that motion cues indicating a forward motion are the strongest to attract attention. However, compared to inanimate objects and cues, people preferentially attend to animals and faces, a process in which the amygdala is thought to play an important role. In this chapter, we tested four rare patients with selective bilateral amygdala lesions, in order to address an important open question: is the amygdala critical for the preferential processing of animate stimuli? There is substantial evidence to suggest the amygdala's involvement in attention to, and detection of, animate stimuli (pictures of animals and people), but nobody has yet tested whether this involvement is necessary.

We used a well-validated experimental protocol (New et al., 2007), and supplemented this with additional control tasks as well as detailed eye-tracking measures. Compared to matched controls, our four patients all performed completely normally, showing a robust advantage for detecting animals and people. These results are particularly surprising because both fMRI studies as well as single-unit recordings have found strong evidence for amygdala responses that are tuned to these stimulus categories. This result is important, because it redirects the search for the neural substrates of animacy detection elsewhere, and we suggest some plausible candidate structures in this chapter.

Linking between Chapter II and Chapter IV, this chapter directly compares social (faces, people and head directions) vs. non-social cues (plants and artifacts). We show attentional advantage of social over non-social cues, but this attentional advantage is independent of the amygdala—a key structure of the "social brain". Then what is role of the amygdala in

processing animate stimuli, in particular faces? We will investigate this question in the next chapter.

This work has been published as (Wang et al., 2014b).

## 3.2 Summary

The amygdala is thought to play a critical role in detecting salient stimuli. Several studies have taken ecological approaches to investigating such saliency, and argue for domain-specific effects for processing certain natural stimulus categories, in particular faces and animals. Linking this to the amygdala, neurons in the human amygdala have been found to respond strongly to faces, and also to animals. Yet the amygdala's necessary role for such category-specific effects at the behavioral level remains untested.

Here we tested four rare patients with bilateral amygdala lesions on an established change-detection protocol. Consistent with prior published studies, healthy controls showed reliably faster and more accurate detection of people and animals, as compared to artifacts and plants. But so did all four amygdala patients: there were no differences in phenomenal change blindness, in behavioral reaction time to detect changes, or in eye-tracking measures. The findings provide decisive evidence against a critical participation of the amygdala in rapid, initial processing of attention to animate stimuli, suggesting that the necessary neural substrates for this phenomenon arise either in other subcortical structures (such as the pulvinar) or within cortex itself.

## 3.3 Introduction

The human amygdala clearly contributes to processing emotionally salient and socially relevant stimuli (Kling and Brothers, 1992, LeDoux, 1996, Adolphs, 2010). While most studies have investigated stimuli that are salient because they are emotionally arousing

(McGaugh, 2004), or involve reward-related valuation (Baxter and Murray, 2002, Paton et al., 2006), recent findings show that the amygdala processes salient stimuli even when there is no emotional component involved at all (Herry et al., 2007). Earlier notions that the amygdala specifically mediates fear processing have been replaced by recent accounts that it is involved in processing a broader spectrum of salient stimuli, such as biological values and rewards (Baxter and Murray, 2002), novel objects (Bagshaw et al., 1972), emotion-enhanced vividness (Todd et al., 2012), animate entities (Yang et al., 2012b), temporal unpredictability (Herry et al., 2007) and personal space (Kennedy et al., 2009). While some of these may involve fear processing, it has been argued that a more parsimonious explanation is that the amygdala instead acts as a detector of perceptual saliency and biological relevance (Sander et al., 2005, Adolphs, 2008).

One category of salient stimuli that have been recently investigated is animate (living) stimuli (New et al., 2007, Mormann et al., 2011). Subjects can rapidly detect animals in briefly presented novel natural scenes even when attentional resources are extremely limited (Li et al., 2002), suggesting that such detection may in fact be pre-attentive. Furthermore, images of animals and people are detected preferentially during change blindness tasks (New et al., 2007), an approach on which we capitalized here. The amygdala's role in such preferential detection is also related to a large literature of neuroimaging studies suggesting that amygdala activation to faces might be seen even under conditions of reduced attention or subliminal presentation (Morris et al., 1998, Whalen et al., 1998, Morris et al., 2001, Vuilleumier et al., 2001, Anderson et al., 2003, Jiang and He, 2006) (but see (Pessoa et al., 2006)). Importantly, direct recordings of single neurons in the human amygdala were recently shown to respond preferentially to images of animals (Mormann et al., 2011), and robust responses to images of faces have been shown as well (Rutishauser et al., 2011). This begs the question whether the strong neuronal responses tuned to animals in the amygdala (Mormann et al., 2011) have a behavioral consequence such as enhanced attention to animals (New et al., 2007). If so,

we would expect a reduced preferential detection of animals in patients with amygdala lesions.

Here we tested four rare patients with bilateral amygdala lesions on a flicker change-detection protocol (Grimes, 1996, Rensink et al., 1997) with concurrent eye-tracking to test the amygdala's role in rapid detection of animate stimuli. We found both healthy controls and all four amygdala patients showed reliably faster and more accurate detection of animals and people. Detailed eye-tracking analyses further corroborated the superior attentional processing of animals, people and faces, and again were equivalent in controls and amygdala patients.

## 3.4 Methods

### *3.4.1 Subjects*

We tested four rare patients, SM, AP, AM and BG, who all have bilateral amygdala lesions due to Urbach-Wiethe disease (Hofer, 1973), a condition that caused complete bilateral destruction of the basolateral amygdala and variable lesions of the remaining amygdala while sparing hippocampus and all neocortical structures (see **Figure 3.1** for MRI anatomical scans and **Table 3.1** for neuropsychological data). AM and BG are monozygotic twins whose lesions and neuropsychology have been described in detail previously (Becker et al., 2012): both AM and BG have symmetrical complete damage of the basolateral amygdala with some sparing of the centromedial amygdala. SM and AP are two women who have also been described previously (Hampton et al., 2007, Buchanan et al., 2009): SM has complete bilateral amygdala lesions, whereas AP has symmetrical bilateral lesions encompassing about 75% of the amygdala. Ten neurologically and psychiatrically healthy subjects were recruited as controls, matched in gender, age, IQ and education (see **Table 3.1**). Subjects gave written informed consent and the experiments were approved by the Caltech Institutional Review Board. All subjects had normal or corrected-to-normal visual acuity.

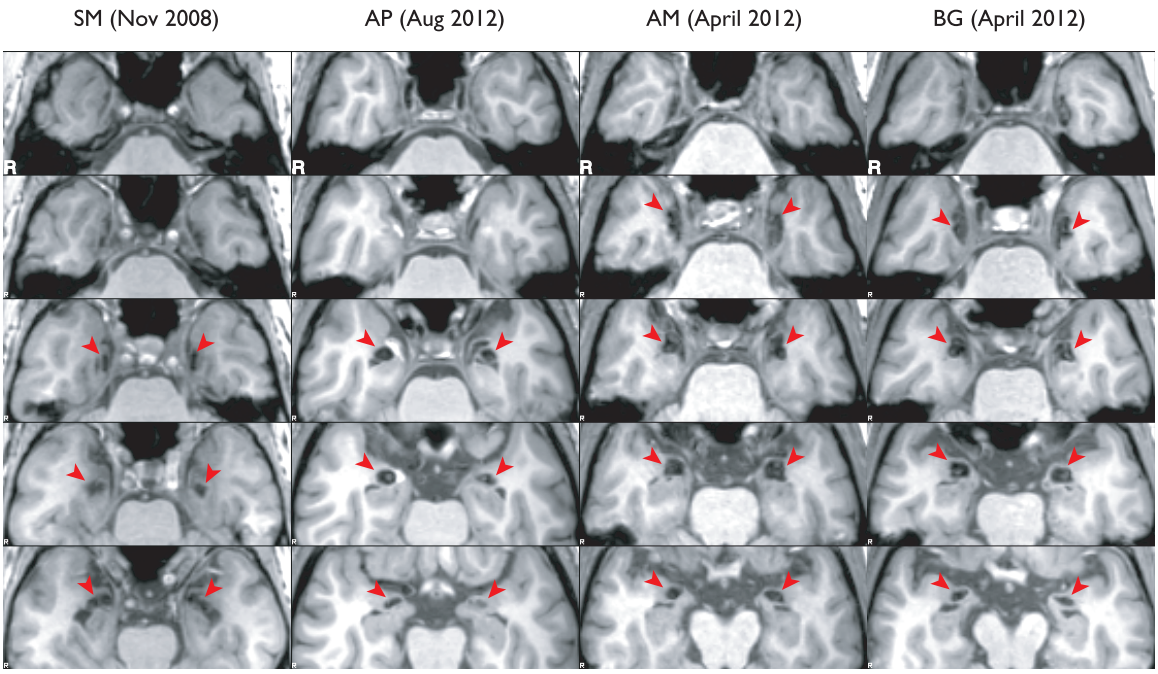


**Figure 3.1. MRI anatomical scans of the amygdala lesions.**

Displayed are high-resolution (0.5–1 mm isotropic) horizontal T1-weighted magnetic resonance imaging sections of the anterior medial temporal lobes. Red arrows index the focal bilateral amygdala calcification damage. R: right.

**Table 3.1. List of subject demographics and psychological evaluation.**

Intelligence was measured by the Wechsler Abbreviated Scale of Intelligence (WASI). SM's IQ was measured by the Wechsler Adult Intelligence Scale (WAIS-III). AM and BG's IQ was measured by the HAWIE-R ('Hamburg-Wechsler Intelligenztest für Erwachsene in revidierter Fassung'), a German-language adaptation of the WAIS-R (Wechsler Intelligence Test for Adults-Revised), which provides a measure of verbal, performance, and full-scale IQ.

Abbreviations: Age: age at testing. Hand: Dominant handedness (A: ambidextrous, L: left, R: right); Benton: Benton Facial Recognition Test, long form score. Benton scores 41–54 are in the normal range. WASI: IQ scores from the Wechsler Abbreviated Scale of Intelligence: full scale IQ (FSIQ), performance IQ (PIQ), verbal IQ (VIQ). n.a.: not available.

Subjects were tested individually. The four amygdala patients do not differ in age from the controls (amygdala patient mean age = 36.5 years, range 27–43; control mean age = 35.7 years, range 23–57; t-test p = 0.89). SM's IQ was measured by WAIS-III. The rest of the subjects were measured by WASI. The mean IQ of the other three amygdala patients was 98.3 (range 96–101), not different from that of their controls (mean IQ = 104.7, range 100–116; p = 0.13).

| ID | Age | Sex | Hand | Race | Education | Benton | WASI | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | FSIQ | PIQ | VIQ |
| SM | 43 | F | R | Caucasian | High School | 45 | 88 | 95 | 86 |
| AP | 27 | F | R | Asian/Pacific Islander | Bachelor's Degree | 50 | 98 | 106 | 92 |
| AM | 38 | F | A | Caucasian | 13 years of education in Germany | 36 | 101 | 103 | 99 |
| BG | 38 | F | R | Caucasian | 13 years of education in Germany | 41 | 96 | 97 | 94 |
| RA0067 | 57 | F | R | African American | Some College | 54 | 104 | n.a. | n.a. |
| RA0071 | 45 | F | R | Caucasian | Some College | 49 | 111 | 103 | 117 |
| RA0629 | 32 | F | A | Caucasian | Some College | n.a. | n.a. | n.a. | n.a. |
| RA0633 | 27 | F | R | Asian/Pacific Islander | Bachelor's Degree | n.a. | n.a. | n.a. | n.a. |
| RA0762 | 23 | F | A | Hispanic/ Latino | Some College | 50 | 100 | 105 | 95 |

| ID | Age | Sex | Hand | Race | Education | Benton | WASI | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | FSIQ | PIQ | VIQ |
| RA0764 | 31 | F | R | Caucasian | Master's Degree | n.a. | 102 | 103 | 101 |
| RA0829 | 29 | F | R | Caucasian | Bachelor's Degree | n.a. | 116 | 111 | 116 |
| RA0835 | 38 | F | R | Hispanic/ Latino | Bachelor's Degree | 49 | 102 | 99 | 104 |
| RA0848 | 40 | F | R | Caucasian | High School | n.a. | 101 | 104 | 98 |
| RA0851 | 35 | F | R | Caucasian | Bachelor's Degree | n.a. | 107 | 103 | 108 |

*3.4.2 Stimuli and apparatus*

We used a flicker change-detection task using natural scenes (**Figure 3.2**). Change targets were drawn from the following five categories: animals (32 images), artifacts (32 images), people (31 images), plants (29 images) and head directions (26 images). A subset of the images had been used in previous studies that showed reliably faster detection of animals and people (New et al., 2007, New et al., 2010). Targets were embedded in complex and natural scenes that contained items from non-target categories as well. The changes to the targets between alternating presentations of an image included both flips and disappearances. Construction and validity of the stimuli, stimulus properties and further control experiments using inverted stimuli have been discussed in previous studies (New et al., 2007, New et al., 2010).

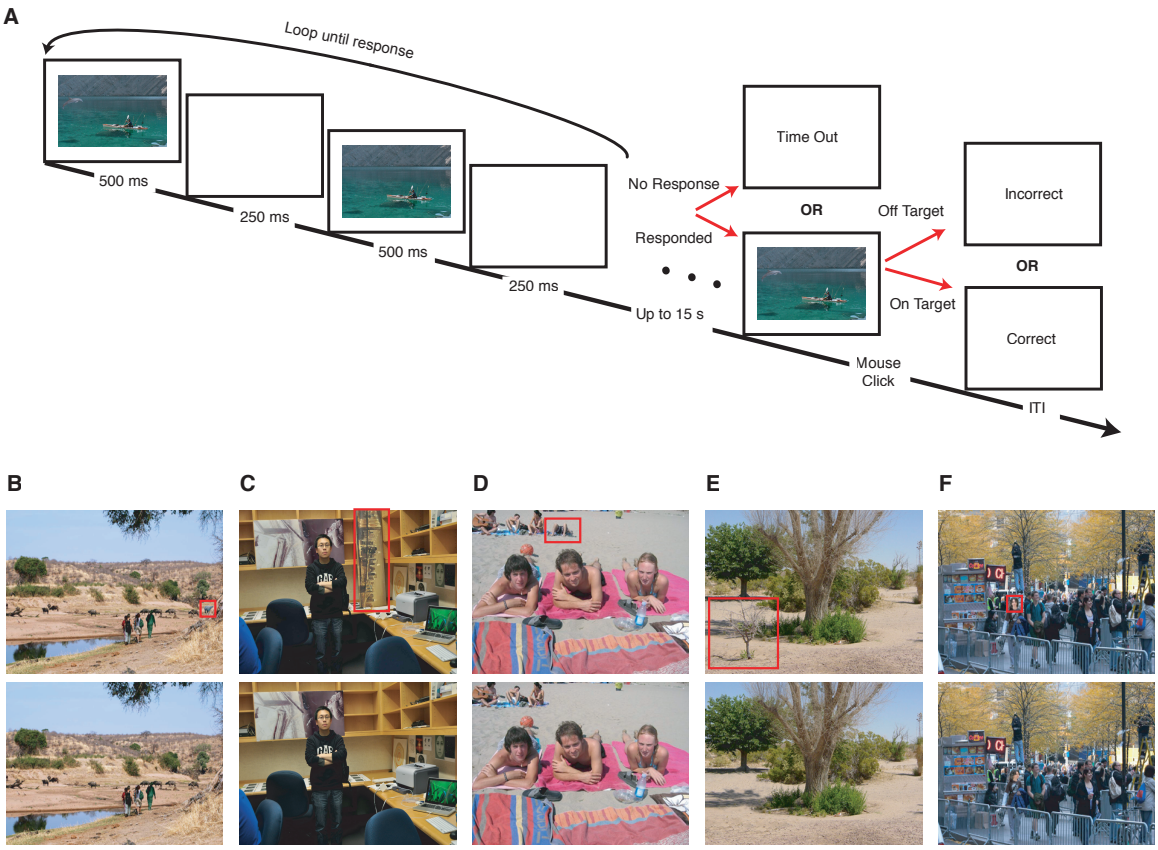


**Figure 3.2. Task and sample stimuli.**

**(A)** Task structure and timecourse. One target object either disappeared or changed its orientation between two alternating frames. These frames were separated by a blank frame. Note that the sizes of the stimuli are not to scale. Sample stimuli showing changes of **(B)** an animal, **(C)** artifact, **(D)** person, **(E)** plant and **(F)** head direction. The changes are labeled by a red box. Low-level saliency and eccentricity of the changes did not differ between categories, while plants were significantly larger in area, favoring easier detection.

We quantified low-level properties of all stimuli. Target categories did not differ in terms of bottom-up local saliency around the target region as quantified by the Itti-Koch bottom-up model of attention (Itti et al., 1998, Itti and Koch, 2001) (one-way ANOVA, p

= 0.44; mean saliency was normalized to 1 within each image), nor by mean distance from the center of the image ($p = 0.28$). Plants subtended a larger area on the screen than the other categories ($p < 0.05$). SM and SM controls were tested on a subset of the stimuli that had larger area for inanimate stimuli (artifacts and plants vs. animals and people; $p < 0.005$), but did not differ in Itti-Koch saliency (artifacts and plants vs. animals and people; $p = 0.77$) or distance to the center ($p = 0.13$). Overall, any low-level differences in area favored a faster detection of inanimate stimuli instead of the faster detection of animate stimuli we observed. We also note that our key comparison is between amygdala patients and their matched controls, and these two groups always saw identical stimuli in any case.

Subjects sat 65 cm from an LCD display (refresh rate 60 Hz, centrally presented stimuli subtending 14.9° × 11.2°). Stimuli were presented using MATLAB with Psychtoolbox 3 (Brainard, 1997) (http://psychtoolbox.org).

### *3.4.3 Task*

In each trial, we presented a sequence of the original scene image (500 ms), a blank screen (250 ms), the altered scene with a changed target (500 ms), and a blank (250 ms). This sequence was repeated until subjects detected the changed target (**Figure 3.2**). Subjects were asked to press the space bar as quickly as possible upon detecting the change. Subsequent to detection, subjects were asked to use a mouse to click on the location of the change on the original scene image, which was followed by a feedback screen for 1 second (the words, ‘accurate’, or ‘inaccurate’). If subjects did not respond within 15 seconds (20 seconds for SM and SM controls), a message ‘Time Out’ was displayed. An inter-trial-interval (ITI) was jittered between 1 and 2 seconds. Scene and category order were completely randomized for each subject. Subjects practiced 5 trials (one trial per stimulus category) for initial familiarization.

Patients AP, AM and BG and 8 matched controls performed the task as described above. Patient SM and two matched controls performed the task with a subset of the stimuli (identical setup and stimuli to (New et al., 2010), which did not contain the head direction change category).

### *3.4.4 Eye-tracking*

We tracked binocular eye positions using a Tobii TX300 system operating at 300 Hz with a 23-inch screen (screen resolution: 1920 × 1080). Fixations were detected using the Tobii Fixation Filter implemented in Tobii Studio (Olsson, 2007) which detects quick changes in the gaze point using a sliding window averaging method (velocity threshold was set to 35 pixels/sample and distance threshold was set to 35 pixels in our study).

### *3.4.5 Data analysis*

ROIs were defined for each image pair by delineating a rectangular area that encompassed the target change region. Out of 1818 trials, 1571 mouse clicks (86.4%) fell within these pre-defined ROIs (correct trials) and 111 clicks (6.11%) fell outside (incorrect trials); 136 trials (7.48%), were time-out trials. For all subsequent analyses, we only analyzed correct trials with RTs that fell within ± 2.5 SD; 61 correct trials (3.36% of all trials) were excluded due to this RT criterion. There was no difference between amygdala patients and matched control subjects in the proportion of any of the above trial types (all t-tests ps > 0.05). We used MATLAB for t-tests and one-way ANOVAs, and R (R Foundation for Statistical Computing, Vienna, Austria) for repeated-measures ANOVAs.

## 3.5 Results

*3.5.1 Phenomenological change blindness and conscious detectability*

To obtain a systematic characterization of awareness of, and attention to, the change target, we first quantified phenomenological change blindness—the most severe case of change blindness in which the target change is missed entirely. The full time course of change detection for each stimulus category is depicted in **Figure 3.3A,F**, which plots the cumulative proportion of changes detected as a function of time elapsed. Steeper slopes indicate faster change detection and higher asymptotes mean more changes eventually detected. For both amygdala patients and control subjects, the curves for animate targets rose more rapidly and reached higher asymptotes, compared to inanimate targets. At any given time, a greater proportion of changes was detected for animate targets than inanimate ones. Both amygdala patients and control subjects were entirely change-blind more often for inanimate targets than for animate ones (time-out rates, **Figure 3.3B,G**; amygdala: 5.4 ± 4.8% for animate vs. 11.0 ± 7.8% for inanimate; see **Table 3.2** for statistics) and there was no significant difference between amygdala patients and controls.

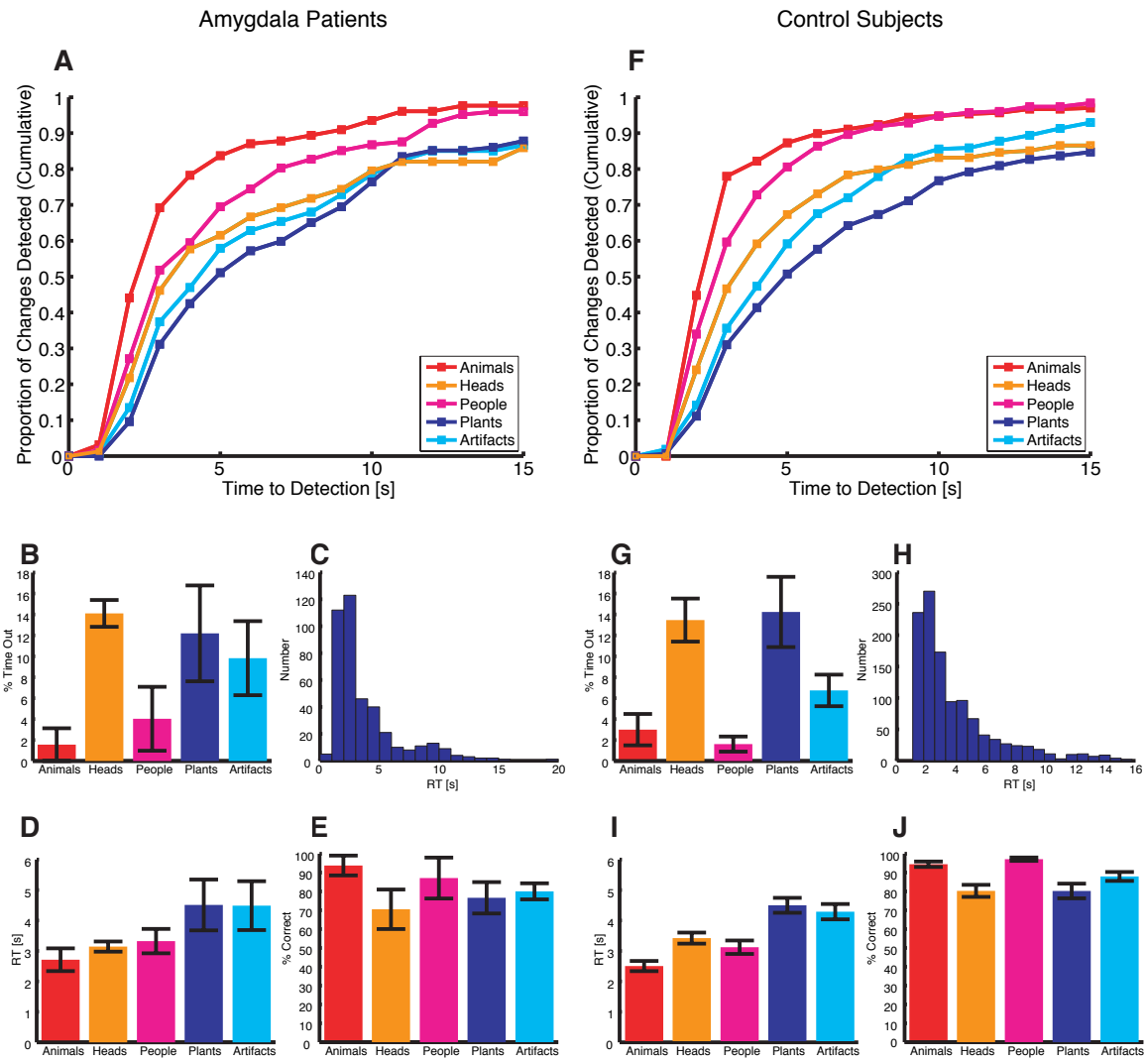


**Figure 3.3. Change detection is category-specific.**

Both amygdala lesion patients **(A–E)** (N = 4) and control subjects **(F–J)** (N = 10) showed advantageous change detection of animals, people and head directions over changes to plants and artifacts. **(A,F)** Graphs show proportion of changes detected as a function of time and semantic category. **(B,G)** Percentage of time-out for each category. **(C,H)** RT histogram across all trials. **(D,I)** Mean RT for each category. **(E,J)** Percentage of correct detection for each category. Error bars denote one s.e.m. across subjects.

**Table 3.2. ANOVA table.**

p-values in bold indicate a statistical significance at $p < 0.05$. d.f.: degree of freedom.

| Measure | Statistical Test | Effect | F-statistic (d.f.) | p-value |
|---|---|---|---|---|
| Change blindness | 5x2 mixed-model ANOVA of target category X group (amygdala lesion vs. control) | Main effect of target category | F(4,45) = 13.1 | **p = 3.76×10** |
| | | Main effect of subject group | F(1,12) = 0.053 | p = 0.82 |
| | | Interaction | F(4,45) = 0.46 | p = 0.76 |
| | One-way repeated-measures ANOVA in amygdala lesion group | Main effect of category | F(4,11) = 2.68 | p = 0.088 |
| | One-way repeated-measures ANOVA in control group | Main effect of category | F(4,34) = 11.4 | **p = 5.82×10** |
| Conscious detection | Mixed-model two-way ANOVA of target category X subject group | Main effect of category | F(4,36) = 21.1 | **p = 5.11×10** |
| | | Main effect of group | F(1,9) = 0.045 | p = 0.84 |
| | | Interaction | F(4,36) = 0.079 | p = 0.99 |
| | One-way repeated-measures ANOVA in amygdala lesion group | Main effect of category | F(4,8) = 6.73 | **p = 0.011** |
| | One-way repeated-measures ANOVA in control group | Main effect of category | F(4,28) = 14.8 | **p = 1.29×10** |
| RT | Mixed-model two-way ANOVA of target category X subject group | Main effect of category | F(4,45) = 44.4 | **p = 4.44×10** |
| | | Main effect of group | F(1,12) = 0.22 | p = 0.65 |
| | | Interaction | F(4,45) = 0.12 | p = 0.97 |
| | One-way repeated-measures ANOVA in amygdala lesion group | Main effect of category | F(4,11) = 7.57 | **p = 0.0035** |
| | One-way repeated-measures ANOVA in control group | Main effect of category | F(4,34) = 39.7 | **p = 2.26×10** |
| | Mixed-model two-way ANOVA of target category X subject group | Main effect of category | F(4,36) = 32.2 | **p = 1.95×10** |
| | | Main effect of group | F(1,9) = 0.15 | p = 0.71 |

| Measure | Statistical Test | Effect | F-statistic (d.f.) | p-value |
|---|---|---|---|---|
| Number of fixations | | Interaction | F(4,36) = 1.45 | p = 0.24 |
| | One-way repeated-measures ANOVA in amygdala lesion group | Main effect of category | F(4,8) = 4.19 | **p = 0.040** |
| | One-way repeated-measures ANOVA in control group | Main effect of category | F(4,28) = 31.6 | **p = 5.22×10** |
| Hit rates | Mixed-model two-way ANOVA (subject group X category) | Main effect of target category | F(4,45) = 17.2 | **p = 1.22×10** |
| | | Main effect of subject group | F(1,12) = 1.37 | p = 0.26 |
| | | Interaction | F(4,45) = 0.88 | p = 0.48 |
| | One-way repeated-measures ANOVA in amygdala lesion group | Main effect of category | F(4,11) = 5.64 | **p = 0.010** |
| | One-way repeated-measures ANOVA in control group | Main effect of category | F(4,34) = 12.5 | **p = 2.35×10** |
| Fixation order | Mixed-model two-way ANOVA of target category X subject group | Main effect of category | F(4,36) = 24.6 | **p = 7.14×10** |
| | | Main effect of group | F(1,9) = 0.049 | p = 0.83 |
| | | Interaction | F(4,36) = 2.65 | **p = 0.049** |
| | One-way repeated-measures ANOVA in amygdala lesion group | Main effect of category | F(4,8) = 2.27 | p = 0.15 |
| | One-way repeated-measures ANOVA in control group | Main effect of category | F(4,28) = 26.7 | **p = 3.32×10** |
| Latency | Mixed-model two-way ANOVA of target category X subject group | Main effect of category | F(4,36) = 11.2 | **p = 5.43×10** |
| | | Main effect of group | F(1,9) = 0.45 | p = 0.52 |
| | | Interaction | F(4,36) = 0.70 | p = 0.59 |
| | Mixed-model three-way ANOVA of category X subject group X horizontal position [left vs. right]; main effect of category | Main effect of category | F(4,102) = 38.4 | **p < 10** |
| | | Main effect of horizontal position | F(1,102) = 0.52 | p = 0.47 |
| | | Main effect of subject group | F(1,12) = 0.38 | p = 0.55 |

| Measure | Statistical Test | Effect | F-statistic (d.f.) | p-value |
|---|---|---|---|---|
| Horizontal position effect | | Interactions | | **all ps > 0.05** |
| | Two-way ANOVA of category X horizontal position in amygdala lesion group | Main effect of category | F(4,25) = 6.98 | **p = 0.0006** |
| | | Main effect of horizontal position | F(1,25) = 0.071 | p = 0.79 |
| | | Interaction | F(4,25) = 1.06 | p = 0.40 |
| | Two-way ANOVA of category X horizontal position in control group | Main effect of category | F(4,77) = 36.6 | **p < 10** |
| | | Main effect of horizontal position | F(1,77) = 1.70 | p = 0.20 |
| | | Interaction | F(4,77) = 2.07 | p = 0.093 |
| Vertical position effect | Mixed-model three-way ANOVA of category X subject group X vertical position [upper vs. lower] | Main effect of category | F(4,100) = 22.3 | **p = 3.48×10** |
| | | Main effect of vertical position | F(1,100) = 11.9 | **p = 0.00084** |
| | | Main effect of subject group | F(1,12) = 0.22 | **p = 0.64** |
| | | Interaction between category and vertical position | F(4,100) = 3.90 | **p = 0.0055** |
| | | Other interactions | | **all ps > 0.05** |
| | Two-way ANOVA of category X vertical position in amygdala lesion group | Main effect of category | F(4,25) = 7.92 | **p = 2.89×10** |
| | | Main effect of vertical position | F(1,25) = 1.48 | p = 0.23 |
| | | Interaction | F(4,25) = 1.13 | p = 0.37 |
| | Two-way ANOVA of category X vertical position in control group | Main effect of category | F(4,75) = 14.5 | **p = 8.56×10** |
| | | Main effect of vertical position | F(1,75) = 10.8 | **p = 0.0015** |

| Measure | Statistical Test | Effect | F-statistic (d.f.) | p-value |
|---|---|---|---|---|
| | | Interaction | F(4,75) = 3.16 | **p = 0.019** |

We further analyzed gaze patterns to elucidate a possible mechanism for faster conscious detectability of animate stimuli: having fixated a target, its change should be detected more efficiently for animate than inanimate stimuli. We quantified this by computing the percentage of trials having 'misses', which were defined as fixations onto the target area ROI (a rectangular ROI tightly surrounding the target) yet without the change detected. We excluded the last 3 fixations entering the ROI for misses since they may have been associated with subsequent detection of changes (subjects tended to fixate on the target for 1 to 3 fixations in order to confirm their selection. Thus, the last 1–3 fixations corresponded to the detection instead of misses of targets). For homogeneity of the data, we here only analyzed the data from AP, AM, BG and their matched controls, who all had identical stimuli and experimental setup.

**Figure 3.4A–B** shows that animate stimuli had a lower percentage of trials with misses and thus preferentially emerged into consciousness (see **Table 3.2** conscious detection analysis; animate vs. inanimate: 8.1 ± 9.2% vs. 28.8 ± 9.3%, t(2) = -4.26, p = 0.051 for amygdala patients, and 9.8 ± 6.5% vs. 29.3 ± 12.9%, t(7) = -6.63, p = $2.96\times10^{-4}$ for controls) and there was no difference between amygdala patients and control subjects. No target category showed any significant differences in the percentage of misses between amygdala patients and their matched controls (two-tailed t-tests, all ps > 0.67; bootstrap (Efron and Tibshirani, 1994) with 1000 runs, all ps > 0.30). The same pattern of results held when we repeated the analysis by computing the average number of misses instead of percentage of trials with misses as used above. Similarly, the same pattern held when we inflated the size of the ROI to a more lenient region of the image (a 50-pixel circular ROI (1.2 deg. visual angle) centered on the target). These results confirm that the

amygdala is not required for preferential conscious detection of biologically relevant stimuli.

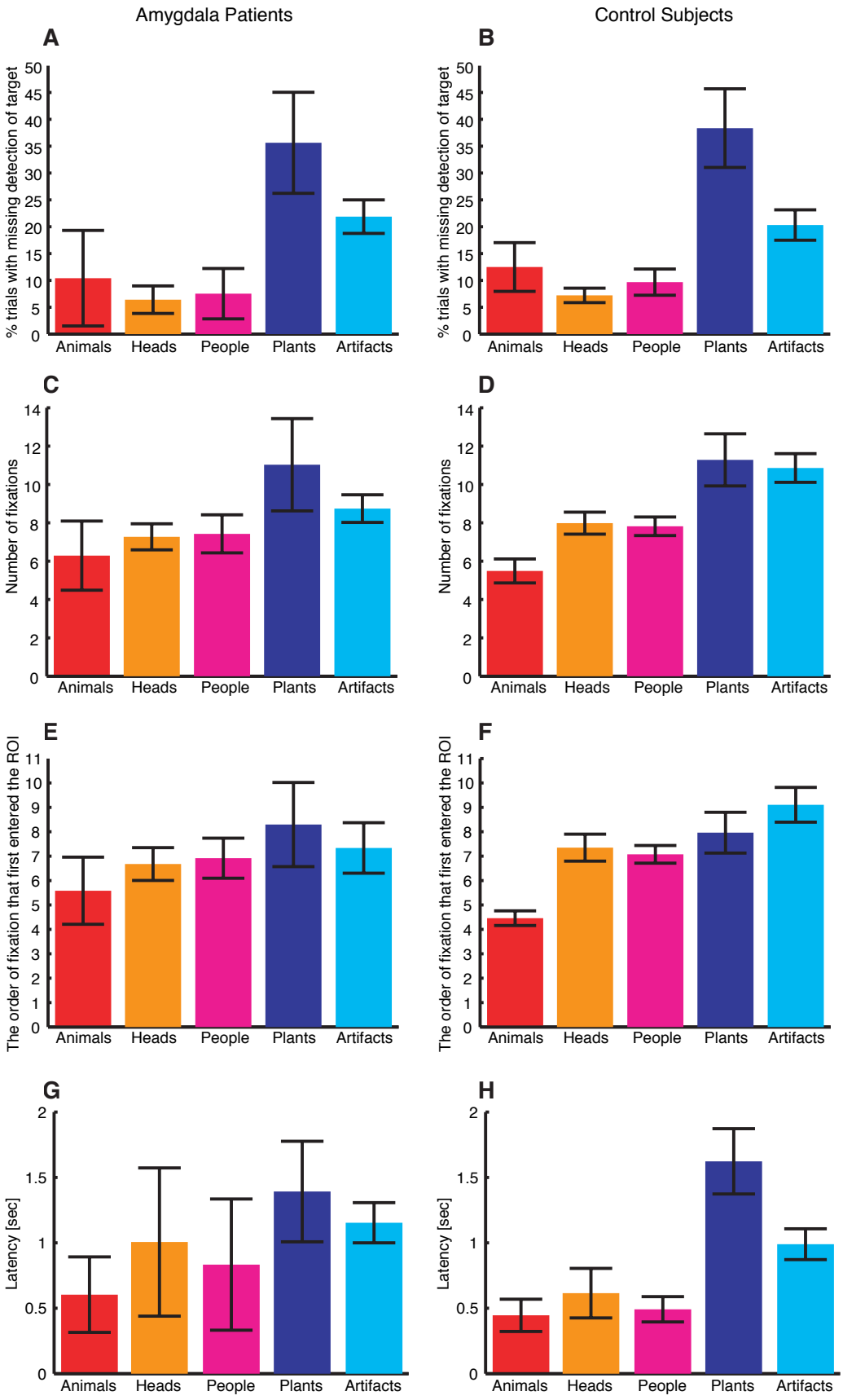
Amygdala Patients
Control Subjects
A
B
C
D
E
F
G
H
% trials with missing detection of target
Number of fixations
The order of fixation that first entered the ROI
Latency [sec]
Animals
Heads
People
Plants
Artifacts

**Figure 3.4. Quantification of fixation properties.**

**(A–B)** Percentage of trials with change blindness despite direct fixation on the change target. **(C–D)** Number of fixations before detecting changes. **(E–F)** The serial order of fixation that first entered the target ROI. **(G–H)** Latency from first fixation onto target to detection of target. **(A,C,E,G)** Amygdala lesion patients (N = 3). **(B,D,F,H)** Control subjects (N = 8). Error bars denote one s.e.m. across subjects.

*3.5.2 Rapid detection of animate stimuli by explicit behavioral reports of change detection*

We next quantified reaction times for the explicit behavioral reports of change detection. We found category-specific effects in reaction times (RT) in both subject groups (see **Table 3.2** RT analysis for statistics). There was a main effect of category but none of group nor any interaction. Category effects were significant when tested separately in the amygdala lesion group (**Figure 3.3D**) as well as in the control group (**Figure 3.3I**), with animate targets (animals, people and head directions) reliably showing faster detection than inanimate targets (artifacts and plants). Both amygdala-lesioned subjects and controls detected animate targets faster (amygdala: 3.13 ± 0.66 sec for animate and 4.50 ± 1.63 sec for inanimate; controls: 2.91 ± 0.52 sec for animate and 4.36 ± 0.70 sec for inanimate, mean ± SD). We confirmed this animacy effect for both groups using a summary statistic approach: the difference of the mean RT for animate and inanimate targets was significant both for the amygdala patients ($t(3) = -2.57$, $p = 0.041$, paired t-test) and control subjects ($t(9) = -12.94$, $p = 2.02\times10^{-7}$). All individual control subjects and amygdala patients except AM showed detection advantages of animate stimuli (two-tailed t-tests comparing animate vs. inanimate stimuli within each subject, all ps < 0.05). No target category showed any significant differences between amygdala patients and their matched controls (two-tailed t-tests, all ps > 0.47; bootstrap with 1000 runs, all ps >

0.24). All above effects also held when we used log-transformed RT as our dependent measure.

We quantified the number of fixations made before the explicit report of change detection (**Figure 3.4C,D**), and found a pattern which mirrored the RT results. There was a category effect as expected (see **Table 3.2**, number of fixations analysis) but no difference between amygdala patients and controls. No target category showed any significant differences between amygdala patients and their matched controls (two-tailed t-tests, all ps > 0.14; bootstrap with 1000 runs, all ps > 0.32). Category effects were prominent separately within amygdala patients (see **Figure 3.4C**) and within control subjects (see **Figure 3.4D**), with changes in animate stimuli requiring fewer numbers of fixation to be detected than those in inanimate stimuli. Direct comparisons collapsing all animate stimuli vs. inanimate stimuli revealed a significantly faster detection of animate stimuli for both amygdala patients (7.0 ± 2.0 vs. 9.9 ± 2.5 fixations, paired-sample two-tailed t-test, $t(2) = -9.20$, $p = 0.012$) and control subjects (7.1 ± 1.5 vs. 11.1 ± 2.9 fixations, $t(7) = -6.85$, $p = 2.42\times10^{-4}$).

Consistent with prior reports (New et al., 2007), more rapid detection of changes to animals and people was not accompanied by any loss of accuracy. On the contrary, both amygdala patients and control subjects were both faster (see **Figure 3.3D,I**) and more accurate for animate targets (hit rates, **Figure 3.3E,J**; amygdala: 86.2 ± 17.3% for animate vs. 78.3 ± 12.6% for inanimate; control: 91.6 ± 4.3% for animate vs. 84.1 ± 8.7% for inanimate; see **Table 3.2,** hit rates analysis, for statistics), and there was no difference between amygdala patients and control subjects. Thus, speed-accuracy trade-offs could not explain the faster detection of animate stimuli and the strong orienting towards animate stimuli resulted in both more rapid and accurate detection of changes.

Within animate targets, animals showed the greatest detection advantages. For both amygdala patients and control subjects, animals had the steepest cumulative detection rate curve (**Figure 3.3A,F**) and the shortest detection RT (**Figure 3.3D,I**, two-tailed

pairwise t-tests to compare animals vs. every other category; amygdala: $p = 0.041$ ($t(3) = -3.44$) for people and $ps < 0.081$ for all other comparisons; controls: $ps < 0.05$ for all comparisons). Further, animals featured a higher detection rate over artifacts, plants and head direction changes (**Figure 3.3E,J**, two-tailed paired-sample t-test; $ps < 0.05$ for all comparisons of both amygdala patients and controls) and a lower time-out rate over head direction changes (**Figure 3.3B,G**, $ps < 0.05$ for both amygdala patients and controls).

Finally, a series of direct and uncorrected t-tests showed no significant differences between amygdala patients and control subjects on change blindness (i.e., time-out), hit rates and RT for any categories (two-tailed unpaired t-tests, $ps > 0.11$ for all comparisons; confirmed by bootstrap with 1000 runs (all $ps > 0.19$)).

### *3.5.3 Implicit measures of change detection from eye-tracking*

While we did not find any impairment of change blindness in amygdala patients at the level of phenomenology or explicit detection response, it remained possible that they might be impaired on more implicit measures. To address this possibility, we analyzed the eye-tracking data in more detail: subjects might look at targets more rapidly for animate stimuli (an attentional mechanism of faster orienting that could in principle be distinct from the conscious detectability mechanism (Koch and Tsuchiya, 2007)). We quantified this by computing the serial order of fixation that first entered the target area.

Control subjects had earlier fixations onto animate than inanimate targets (**Figure 3.4F**; see **Table 3.2**, fixation order analysis; $6.3 \pm 1.3$ vs. $8.5 \pm 2.2$ for animate vs. inanimate, paired t-test: $t(7) = -4.31$, $p = 0.0035$) and animals attracted the earliest fixations (paired t-tests against every other category, $ps < 0.005$). We observed a similar pattern of earlier fixations onto animals and animate targets in the amygdala lesion patients (**Figure 3.4E**; $6.4 \pm 1.6$ vs. $7.8 \pm 2.1$ for animate vs. inanimate; paired t-test: $t(2) = -5.15$, $p = 0.036$), and we observed no difference between amygdala lesion patients and control subjects. No target category showed any significant differences between amygdala patients and their

matched controls (two-tailed t-tests, all ps > 0.22; bootstrap with 1000 runs, all ps > 0.19).

In the above analysis, we counted as a datapoint the last fixation of the trial even when the subject never fixated onto the target (i.e., time-out trials). When we repeated the above analysis by excluding all time-out trials, we obtained qualitatively the same pattern of results. Furthermore, when we repeated the above analysis with the absolute latency (in seconds) of the first fixation onto the target (instead of the serial order of the first fixation), we obtained qualitatively the same pattern of results.

So far, we have shown that detection advantages of animate stimuli could be attributed to either attention or conscious detection, but neither requires the amygdala. But how might initial attention and conscious detectability interact? We observed that faster detection of animate stimuli (by pushing a button) was typically preceded by more rapid initial fixation towards them (**Figure 3.4E,F**). Supporting a role for fast initial orientation in facilitating subsequent detection, there was a significant trial-by-trial correlation (on all correct trials) between the serial order of the first fixation onto the target ROI and the total number of fixations taken to detect the change (Pearson correlation; amygdala: $r = 0.89$, $p < 10^{-20}$; control: $r = 0.76$, $p < 10^{-20}$); similarly, there was a correlation between latency (absolute time elapsed in seconds) of the first fixation onto the target ROI and button press RT (amygdala: $r = 0.81$, $p < 10^{-20}$; control: $r = 0.78$, $p < 10^{-20}$). To further establish the role of initial orienting in conscious detectability, we next measured the latency from having first fixated onto the target ROI to detecting the target change on all correct trials (**Figure 3.4G,H**). Once the target ROI had been fixated, this latency should reflect the efficacy of conscious detectability. We found a category-specific effect on latency (see **Table 3.2** latency analysis), with animate stimuli featuring shorter latencies than inanimate stimuli. Again, there was no difference between amygdala patients and controls nor any interaction. No target category showed any significant differences between amygdala patients and their matched controls (two-tailed t-tests, all ps > 0.32; bootstrap with 1000 runs, all ps > 0.17). These results isolate a category-specific effect of

animate stimuli on the efficacy of conscious detectability, and furthermore demonstrate that this mechanism is independent of the amygdala.

*3.5.4 Detection advantages to animals were not lateralized*

Given that animal-selective neurons were discovered primarily in the right amygdala (Mormann et al., 2011), we expected that detection advantages might be lateralized to some extent. We thus divided target locations according to their horizontal positions. The category effects described above replicated for targets in either the left or right half of the image (see **Table 3.2**, horizontal position effect analysis), and there was no main effect of laterality (3.7 ± 1.2 vs. 3.6 ± 1.3 seconds (mean ± SD) for left vs. right) or subject group, nor any interactions. Similarly, no laterality effect was found separately within amygdala patients nor within control subjects. Further post-hoc paired-sample t-tests showed no difference in detecting the targets between left and right (ps > 0.05 for all categories and for both amygdala patients and control subjects, except one uncorrected $p = 0.022$ ($t(18) = 2.50$) for people detection from control subjects).

We repeated this analysis in relation to upper vs. lower parts of the image. The category effects were observed for both upper and lower parts (see **Table 3.2**, vertical position effect analysis). We found a main effect of category, and to our surprise, a main effect of vertical position (4.0 ± 1.4 vs. 3.6 ± 1.1 seconds (mean ± SD) for upper vs. lower) as well as an interaction between category and vertical position. Separate analyses within amygdala patients and control subjects confirmed both the category effect and the vertical position effect (amygdala: 4.1 ± 1.5 vs. 3.7 ± 1.3 seconds for upper vs. lower; controls: 4.0 ± 1.4 vs. 3.5 ± 0.9 seconds for upper vs. lower). This vertical position effect was primarily driven by faster detection of people and plants in the lower visual field. All above patterns held also with log-transformed RT as the dependent measure.

## 3.6 Discussion

On a flicker change-blindness protocol, all our control subjects showed an advantage in detecting animate stimuli (animals, people and head directions) over inanimate stimuli (artifacts and plants), consistent with the prior finding of category-specific attention towards animals (New et al., 2007). Interestingly, the amygdala lesion patients also showed the same detection advantages. Category effects were not lateralized. Eye-tracking data further dissociated two mechanisms contributing to these detection advantages: animate stimuli attracted initial gaze faster and were preferentially detected by button press. Amygdala lesions spared both of these components. Our findings argue against a critical participation of the amygdala in rapid, initial processing of attention to ecologically salient stimuli, and extend this conclusion to both initial orienting as well as to detectability.

### *3.6.1 Advantages of our change detection task and comparison with other tasks*

Compared to previous studies of change detection (New et al., 2007, New et al., 2010), our addition of eye-tracking to the design strongly expanded the scope of our analyses and allowed us to elucidate the mechanisms underlying change detection and provide interesting insights into the visual search performance in change detection. One advantage of using change detection in this study is to better link with previous studies—for instance, it permits comparisons with a large college population (New et al., 2007), a developmental population (i.e., 7–8 year olds) (New et al., 2010), and with individuals diagnosed with autism spectrum disorder (New et al., 2010). Most importantly, the change detection task allows us to quantify the percentage of misses to dissociate attention to animals from conscious detectability of them (eye-tracking vs. detection), which is difficult to probe with a free viewing task.

Ultra-rapid categorization of animals has been shown in a forced choice saccadic task in which human participants can reliably make saccades to the sides containing animals in

as little as 120 ms (Kirchner and Thorpe, 2006). Our response latency was considerably longer compared to this markedly different task, which explicitly tasks the participants with detecting the specific target category, and typically presents one large, central object in each image. It is very likely that the participants in this study would have performed that explicit task far more quickly, even with the natural and complex scenes used here. Conversely, had the change detection task been conducted with far simpler stimuli, such as two side-by-side objects, the animate bias could easily have been revealed through first fixation locations. Interestingly, in the first studies of this bias in healthy participants (New et al., 2007), the fastest responses (< 1 second) were for detecting animate than inanimate objects. Change detection within the first second likely required the target object to be the first attended item in the scene (New et al., 2007).

### *3.6.2 Possible caveats*

In this study, we have shown that the amygdala is not involved in rapid, initial processing of ecologically salient animate stimuli. Top-down contextual knowledge might have played a more important role (cf. (Kanan et al., 2009)) and the reliance on top-down control and contextual information in the task could have diminished the potential effect of amygdala lesions on detection performance. It has been shown that contextual knowledge can drive change detection performance (e.g., (Rensink et al., 1997)) and, interestingly, as a function of semantic inconsistency (Hollingworth and Henderson, 2000). However, in our stimuli, all of the targets were comparably semantically consistent with their scenes.

Top-down control and contextual knowledge are mostly effective when applied towards explicit tasks or targets. However, in our stimuli, the target from one category was often embedded in other distractor categories and the subject had no prior expectation of the target category to apply a specific contextual knowledge regarding that target category. In other words, since our natural scene stimuli mostly contain multiple categories of objects,

subjects could only apply a uniform strategy across all stimuli. For example, in a scene containing both faces and plants, subjects might look at faces first regardless of whether the target was a face or a plant. Therefore, top-down control involved in our study would be unlikely to affect within-subject comparisons between categories. It will be interesting to explore this issue further in future studies with quantitative analyses of the spatial layout of fixations with respect to the distribution of different target categories.

Our findings were not explained by category differences in low-level saliency. Our stimulus set was biased, if anything, towards low-level features favoring better detection of inanimate stimuli, the opposite of the effect we found, and detection advantages towards animate stimuli are known to be abolished with inverted stimuli, which preserve low-level stimulus properties (New et al., 2007), an effect we replicated in SM and SM's controls.

### *3.6.3 Lateralized effects of category attention*

We did not observe lateralized effects of category attention in this study, even though there is a lateralized distribution of animal-selective neurons in the right human amygdala (Mormann et al., 2011). Behaviorally, lateralized effects have been reported for the sensory and cognitive processing of language, face, and emotion (MacNeilage et al., 2009). Neurologically, laterality has been also well documented for attentional systems (Fox et al., 2006) as well as cortical components of face processing (De Renzi et al., 1994). Recent studies also report laterality effects in frogs, chickens, birds and monkeys, implying an evolutionarily preserved mechanism for detecting salient stimuli that shows an asymmetry for the right hemisphere (Vallortigara and Rogers, 2005). The absence of laterality effects in our data may be due to the limited visual angle subtended by our stimuli (none of the stimuli were far in the left or right periphery), the nature of the stimuli (e.g., none included threatening or strongly valenced stimuli), or the nature of the task. In healthy subjects, a strong asymmetry in attentional resolution has been reported

between the upper and lower visual field (He et al., 1996), a finding that may be related to the intriguing effect of vertical position of change targets in our study.

### *3.6.4 Amygdala lesions and plasticity*

All four amygdala patients have symmetrical complete damage of the basolateral amygdala and in general the damage is extensive, as documented in detail in prior publications (see **Methods**). Although, in the three patients other than SM, there is some sparing of the centromedial amygdala, it would seem unlikely that this remaining intact portion of the amygdala would be able to play the role required for attention or detectability in our task: since the basolateral amygdala is the primary source of visual input to the amygdala (Amaral et al., 1992) and all patients have complete lesions of the basolateral amygdala, this would effectively disconnect any remaining spared parts of the amygdala from temporal neocortex. Furthermore, patient SM has complete bilateral amygdala lesions yet her individual data still showed normal detection advantages for animate stimuli, demonstrating that the amygdala is indeed not necessary for the rapid detection of animate stimuli.

A final consideration concerns the issue of reorganization and plasticity. While we found entirely intact orientation to, and detection of, animate stimuli in all four amygdala patients, all of them had developmental-onset lesions arising from Urbach-Wiethe disease. On the one hand, this made for a homogenous population to study; on the other it introduces the possibility that, over time, compensatory function was provided by other brain regions in the absence of the amygdala. Indeed, evidence for compensatory function (on an unrelated task) has been reported in one of the patients we studied (Becker et al., 2012). Furthermore, normal recognition of prototypical emotional faces has been reported in some (Siebert et al., 2003) but not other (Adolphs et al., 1999) patients with amygdala lesions, and one study even reported a hyper-vigilance for fearful faces in three patients with Urbach-Wiethe disease (Terburg et al., 2012). A critical direction for future studies

will be to replicate our findings in patients with adult, and with acute-onset, amygdala lesions to investigate the added complexities introduced by developmental-onset amygdala lesions.

### *3.6.5 The role of the amygdala in attention and saliency*

Since the early 1990s, an influential view of the role of the amygdala in sensory processing was that it plays a rather automatic, non-conscious role (Dolan, 2002, Ohman, 2002) with long-standing debates about the amygdala's response to fearful faces being either independent of attention (Vuilleumier et al., 2001, Anderson et al., 2003) or requiring attention (Pessoa et al., 2002). A subcortical pathway through the superior colliculus and pulvinar to the amygdala is commonly assumed to mediate rapid, automatic and non-conscious processing of affective and social stimuli, and to form a specific subcortical 'low route' of information processing (LeDoux, 1996, Tamietto and de Gelder, 2010). However, the same patient SM we tested here, who has complete bilateral amygdala lesions, nonetheless showed normal rapid detection and non-conscious processing of fearful faces, suggesting that the amygdala does not process fear-related stimuli rapidly and non-consciously ((Tsuchiya et al., 2009), replicated in (Yang et al., 2012a)). A variety of evidence, including the long latencies that are observed from amygdala recordings in humans (Mormann et al., 2008, Rutishauser et al., 2011), further challenges the 'low route' account of amygdala function (Cauchoix and Crouzet, 2013). Instead, it has been proposed that the amygdala participates in an elaborative cortical network to evaluate the biological significance of visual stimuli (Pessoa and Adolphs, 2010)—a role that appears to necessarily require the amygdala when detailed social judgments need to be made about faces (Adolphs et al., 1994, Adolphs et al., 1998), but not when rapid detection or conscious visibility are assessed.

The human amygdala responds to both emotionally and socially significant information, and arguably social stimuli are often also emotionally salient. Yet there seem to be effects

of social saliency even independent of emotion: the human amygdala is more strongly activated for neutral social vs. non-social information but activated at a similar level when viewing socially positive or negative images (Vrticka et al., 2013). Socially relevant information in faces is expressed in large part in the eye region, including gaze directions (Argyle et al., 1973, Whalen et al., 2004), and viewers predominantly fixate the eyes, a tendency normally correlated with amygdala activation (Gamer and Büchel, 2009). A range of psychiatric disorders feature abnormal fixations onto faces, including abnormal fixations onto the eye region of faces, and several of these are hypothesized to involve the amygdala (Baron-Cohen et al., 2000, Baron-Cohen, 2004, Dalton et al., 2005). Patients with schizophrenia (Sasson et al., 2007), social phobia (Horley et al., 2004), and autism (Adolphs et al., 2001) all show abnormal facial scanning patterns. While by no means eliminating the amygdala as one structure contributing to social dysfunction in these diseases, the data from the present study do argue that it may not play a key online role in those components involving orienting and attentional mechanisms.

### 3.7 Conclusion

Our results show unambiguously that an intact amygdala is not required for rapid orientation towards, and conscious detection of, animate stimuli that normally show preferential processing for these measures. This conclusion leaves open the question of what are the essential structures mediating this effect. Three plausible candidates worth further study would be the pulvinar nucleus of the thalamus, prefrontal cortex, or visual cortices. Both the pulvinar (Tamietto and de Gelder, 2010) and prefrontal cortex (Bar, 2007) have been hypothesized to subserve rapid initial evaluation of stimuli, which can then influence subsequent processing; it is also possible that circuitry within visual cortices itself could suffice to detect salient stimulus categories. How such mechanisms

are initially set up during development, and whether any of them might be innate, remain important topics for future studies.

### 3.8 Acknowledgements

This research was supported by grants from NIMH and the Simons Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank Claudia Wilimzig for providing some of the stimuli, Ty Basinger for creating some of the stimuli, Peter Foley for help with the statistical analysis, and Mike Tyszka for providing the anatomical scans of the lesion patients. The authors declare no competing financial interests.

# Chapter IV: Neurons in the human amygdala selective for perceived emotion

## 4.1 Overview

In Chapter III, we analyzed how animate stimuli attract attention and we showed that amygdala patients have a normal pattern of reliably faster and more accurate detection of animate stimuli. However, people not only attend to people and faces, but also pay attention to others' facial emotions. Humans have a dedicated system to process faces and the amygdala has long been associated with a key role in recognizing facial emotions. In this chapter, we analyzed in detail how the neurons in the amygdala respond to facial emotions. This work is a continuation of our previous single-unit studies where we showed differential neuronal response to whole faces compared to facial parts (Rutishauser et al., 2011), and abnormal neuronal response in autism (Rutishauser et al., 2013). Here, we for the first time showed that neurons in the human amygdala encode the subjective judgment of emotions shown in face stimuli, rather than simply their stimulus features.

Our study makes three broad novel contributions. First, it tests a key hypothesized function of the amygdala, providing an important complement to studies of amygdala responses to faces in monkeys (where it is very difficult to ask the question that we asked). Second, it offers a clear result that suggests a specific transformation between visually responsive cortex in the temporal lobe, the main visual input, and the amygdala. Third, the amygdala is strongly implicated in mood and anxiety disorders, for which the present findings provide important detail.

Connecting with Chapter III, in which we found that the amygdala does not play a role in advantageous attention to social stimuli, here we illustrate what the amygdala does in

processing faces, a particular category of salient social stimuli. The role of the amygdala in social attention will be further investigated in the next chapter.

This work has been published as (Wang et al., 2014c).

### 4.2 Summary

The human amygdala plays a key role in recognizing facial emotions and neurons in the monkey and human amygdala respond to the emotional expression of faces. However, it remains unknown whether these responses are driven primarily by properties of the stimulus or by the perceptual judgments of the perceiver. We investigated these questions by recording from over 200 single neurons in the amygdalae of seven neurosurgical patients with implanted depth electrodes. We presented degraded fearful and happy faces and asked subjects to discriminate the emotion by button press. During trials where subjects responded correctly, we found neurons that distinguished fearful vs. happy emotions as expressed by the displayed faces. During incorrect trials, these neurons indicated the patients' subjective judgment, regardless of whether it was correct or incorrect. Additional analysis revealed that, on average, all neuronal responses were modulated most by increases or decreases in response to happy faces, and driven predominantly by judgments about the eye region of the face stimuli. Following the same analyses, we showed that hippocampal neurons, unlike amygdala neurons, only encoded emotions but not subjective judgment. Our results suggest that the amygdala specifically encodes the subjective judgment of emotional faces, but that it plays less of a role in simply encoding aspects of the image array. The conscious percept of the emotion shown in a face may thus arise from interactions between the amygdala and its connections within a distributed cortical network, a scheme also consistent with the long response latencies observed in human amygdala recordings.

## 4.3 Introduction

The human amygdala plays a crucial role in processing socially and emotionally salient stimuli (Kling and Brothers, 1992, Adolphs, 2010). A large literature, primarily from studies in animals, shows that the amygdala is critical for conditioned fear responses (LeDoux, 1993). However, a number of other studies show that it is involved also in broader aspects of social perception, notably aspects of face processing (Rolls, 1992). These two themes converge in several human studies: there is an impairment in recognizing fearful faces in subjects that lack a functional amygdala (Adolphs et al., 1994) in addition to the impairment of fear conditioning (Bechara et al., 1995, LaBar et al., 1995). Neuroimaging studies have also reported significant activation of the amygdala to fearful faces (Morris et al., 1996).

In humans, it has been reported that amygdala neurons are selective for a variety of visual stimuli (Fried et al., 1997, Kreiman et al., 2000). One category of stimuli that the amygdala plays a key role in analyzing is faces and facial emotions. Subjects with amygdala damage fail to recognize fearful faces (Adolphs et al., 1994), although there is now a consensus that the amygdala is involved in processing many emotions from faces, not just fear (Fitzgerald et al., 2006). Electrophysiological recordings in monkeys have found single neurons that respond not only to faces as such (Rolls, 1984, Leonard et al., 1985), but also to face identities, facial expressions and gaze directions (Gothard et al., 2007, Hoffman et al., 2007). Single neurons in the human amygdala discriminate faces from inanimate objects (Fried et al., 1997). Furthermore, single neurons in the human amygdala were found to encode whole faces selectively (Rutishauser et al., 2011) and show abnormal facial feature selectivity in autism (Rutishauser et al., 2013). Thus, there is substantial evidence from neurophysiological, lesion and fMRI studies for the involvement of the primate amygdala in face processing.

More detailed investigations suggest that impaired fear recognition after amygdala damage can be attributed to a failure to fixate on the eyes (Adolphs et al., 2005), suggesting that the amygdala might act as a detector of perceptual saliency and biological

relevance (Sander et al., 2005, Adolphs, 2008). This was complemented by a neuroimaging study showing that amygdala activity was specifically enhanced for fearful faces when saccading from the mouth to the eye region (Gamer and Büchel, 2009). Patients with schizophrenia (Sasson et al., 2007), social phobia (Horley et al., 2004), and autism (Pelphrey et al., 2002) also show abnormal facial scanning patterns, which have been hypothesized to result from amygdala dysfunction (Baron-Cohen et al., 2000). The functional role of the amygdala is supported by its connection with visual cortices specialized for face processing (Vuilleumier et al., 2004, Moeller et al., 2008, Hadj-Bouziane et al., 2012) as well as reciprocal connections with multiple visually responsive areas in the temporal (Desimone and Gross, 1979, Amaral et al., 2003, Freese and Amaral, 2006) and frontal lobes (Ghashghaei and Barbas, 2002). All of these findings, while supporting a clear role for the amygdala in face processing, also suggest that this role may be relatively specific for certain properties or features of faces, raising the question of what function distinguishes the amygdala's role in face processing from the better known role of temporal cortex in face processing (see **Discussion**). We focused on one particular question in the present study.

Neurons in the monkey and human amygdala respond to the emotional expression of faces, but it remains unknown whether these responses are driven primarily by image properties of the stimuli, by the perceptual judgments of the perceiver, or by behavioral categorization in terms of motor output. To investigate this question, we recorded 210 neurons from 7 neurosurgical patients with implanted depth electrodes on an established 'bubbles' task (Gosselin and Schyns, 2001, Adolphs et al., 2005), in which patients discriminated emotions from sparsely sampled fearful or happy faces. We first characterize neurons that distinguished fearful vs. happy emotions expressed by the displayed faces, on those trials where subjects responded correctly. Next we show that these neurons tracked the patients' subjective judgment regardless of whether it was correct or incorrect. Population permutation analysis confirmed the robustness of this result, on average, across the entire population of neurons. Our data suggest that neuronal responses within the human amygdala are selective for perceived emotion shown in faces

and track subjective judgment expressed by behavior rather than visual properties of the stimuli.

## 4.4 Methods

### *4.4.1 Subjects*

In this study we recorded single-units from 10 neurosurgical patients who had chronically implanted depth electrodes in the amygdalae (**Table 4.1**). 3 patients (total of 3 sessions) did not contribute well-isolated units and hence were excluded from analysis. 2 patients completed 2 sessions, resulting a total of 9 recording sessions that we analyzed.

The subjects' electrophysiology as well as construction of bubbles stimuli, scrambled face stimuli and classification images were described in our previous publications (Rutishauser et al., 2011, Rutishauser et al., 2013).

**Table 4.1. List of patient demographics, pathology, and neuropsychological evaluation.**

Abbreviations: Hand: Dominant handedness; Lang Dom: language dominance as determined by Sodium Amybarbital (Wada) test; Benton: Benton Facial Recognition Test, long form score; WAIS-III: IQ scores from the Wechsler Adult Intelligence Scale: performance IQ (PIQ), verbal IQ (VIQ), full scale IQ (FSIQ), perceptual organization index (POI), verbal comprehension index (VCI). Benton scores 41–54 are in the normal range. Tests indicated with n/a were not performed for clinical reasons. Rey-Osterrieth Complex Figures test are raw scores, subtests are copy, immediate recall reproduction (IR), and 30-minute delayed recall reproduction (DR). 36 possible points for each, 18+ is normal depending on age. Patients 20, 21 and 27 did not contribute neurons and were

thus excluded for analysis. Patients 17 and 28 were diagnosed with ASD. Patients 28 and 29 performed two sessions (each row of neurons represents a separate recording session).

| ID | Age | Sex | Hand | Lang Dom | Benton | Epilepsy diagnosis | WAIS-III | | | | | Rey-O | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | PIQ | VIQ | VCI | POI | FSIQ | Copy | IR | DR |
| P17 | 19 | M | R | L | 43 | Left inferior frontal | 128 | 131 | 122 | 133 | 134 | 34 | 23 | 21 |
| P18 | 40 | M | R | L | 52 | Right mesial temporal hippocampus | 69 | n/a | n/a | n/a | n/a | n/a | n/a | n/a |
| P19 | 34 | M | R | n/a | 39 | Left supplementary motor neocortex | 81 | 74 | 76 | 80 | 86 | 31 | 23 | 20.5 |
| P20 | 27 | M | R | L | 49 | Right mesial temporal hippocampus | 88 | 98 | 89 | 101 | 81 | 33 | 21 | 23.5 |
| P21 | 20 | M | R | n/a | 45 | Right dorsolateral neocortex | n/a | n/a | 93 | 89 | n/a | 34 | 27.5 | 27 |
| P23 | 35 | M | R | L | 41 | Left mesial temporal amygdala | n/a | n/a | 74 | 86 | n/a | 34 | n/a | 9.5 |
| P25 | 31 | M | R | L | 47 | Right dorsolateral neocortex | 81 | 91 | 98 | 82 | 87 | 36 | 9 | 5 |
| P27 | 41 | M | R | n/a | 49 | Bilateral independent temporal lobe | 86 | 91 | 86 | 88 | 89 | 36 | 5 | 5 |
| P28 | 23 | M | R | L | 47 | Right mesial temporal hippocampus | 79 | 77 | 78 | 80 | 76 | 34 | 9.5 | 13 |
| P29 | 18 | F | L | L | 49 | Left deep insula | 104 | 110 | 107 | 101 | 107 | 36 | 19.5 | 19.5 |

(continued)

| ID | Nr Amygdala Neurons | | | Nr Hippocampus Neurons | | |
|---|---|---|---|---|---|---|
| | Total | Fear | Happy | Total | Fear | Happy |
| P17 | 10 | 0 | 1 | 0 | n/a | n/a |
| P18 | 26 | 8 | 0 | 5 | 0 | 0 |
| P19 | 12 | 0 | 3 | 2 | 0 | 0 |
| P20 | 0 | n/a | n/a | 0 | n/a | n/a |
| P21 | 0 | n/a | n/a | 0 | n/a | n/a |
| P23 | 19 | 2 | 1 | 0 | n/a | n/a |
| P25 | 4 | 0 | 0 | 0 | n/a | n/a |

| | | | | | | |
|---|---|---|---|---|---|---|
| P27 | 0 | n/a | n/a | 0 | n/a | n/a |
| P28 | 46<br>27 | 8<br>0 | 1<br>4 | 9<br>23 | 0<br>0 | 1<br>2 |
| P29 | 31<br>34 | 2<br>4 | 3<br>4 | 17<br>11 | 3<br>1 | 2<br>2 |

*4.4.2 Task*

We employed a facial emotion discrimination task in which patients were asked to judge fearful or happy faces as quickly and accurately as possible from randomly selected parts of the face ('bubbles'; **Figure 4.2B**). In each trial, a scrambled face with a central fixation circle was presented for 0.8–1.2 second (randomized). Then the target face stimulus was presented for 500 ms and a blank gray screen followed. Patients started to respond after the target face stimulus onset and regardless of reaction time (RT), the next trial started after an interval of 2.3–2.7 second after stimulus onset. If the patient did not respond by that time, a time-out was indicated by a beep (2.2% of trials were timeouts) (**Figure 4.2A**). Each block contained 72 trials and patients completed 5–7 blocks. Time-out trials were excluded from analysis so all trials included had a behavioral response. We displayed the performance score to the patients at the end of each block as an incentive.

We used 8 face base images (chosen from the Ekman and Friesen stimulus set, 4 different individuals (2 female and 2 male)) showing fearful and happy expressions each. We normalized all faces for mean luminance, contrast, and position of eyes and mouth. We randomly flipped 50% of the stimuli along the vertical axis to prevent any influence of left-right asymmetries present in the faces. This resulted in 16 different face images in total and these face stimuli were then sparsely sampled and presented to participants.

*4.4.3 Data analysis: spikes*

Only single units with an average firing rate of at least 0.2 Hz (entire task) were considered. Trials were aligned to stimulus onset, except when comparing the baseline (a

1 second interval of blank screen right before scramble onset) to the scramble-response for which trials were aligned to scramble onset (which precedes the stimulus onset). Average firing rates (PSTH, see **Figure 4.4** and **Figure 4.6**) were computed by counting spikes across all trials in consecutive 250 ms bins. In order to investigate the temporal dynamics of the significant difference, pairwise comparison was made at each bin using a two-tailed t-test at $p < 0.05$ and Bonferroni-corrected for multiple comparisons across bins in the group PSTH (this is not the unit selection). The PSTH of individual neuron examples were smoothed by a Gaussian kernel with sigma 200 ms (for plotting purposes only, all statistics are based on the raw counts).

### *4.4.4 Data analysis: selection of emotion-selective and interactive units*

Statistical comparisons between the firing rates in response to different stimuli were based on the total number of spikes produced by each unit in a 1.5 s interval starting at 250 ms after stimulus onset (see **Figure 4.4** and **Figure 4.6**). Based on behavior, we categorized each trial as either correct or incorrect. In the following, correct/incorrect thus always refers to whether or not the subject successfully identified the correct emotion of the stimulus shown (fearful or happy). Since only two emotions were shown, an incorrect trial always implies that the subject chose the opposite emotion.

The selection criterion for emotion-selective units (see **Figure 4.4** and **Figure 4.6**) was based on the correct trials only, leaving the incorrect trials statistically independent. Units were defined as emotion-selective if they responded with a different firing rate to fearful relative to happy faces after stimulus onset. By definition, fear-selective units responded significantly more in correct fearful trials compared to correct happy trials, and vice versa for happy-selective units. One-tailed t-tests with $p < 0.05$ were used.

We also quantified whether units responded to emotions conditionally on behavior. For this, a two-way ANOVA ([correct vs. incorrect trials] X [fearful stimuli vs. happy

stimuli]) was used to probe for a significant interaction term with $p < 0.05$ (see **Figure 4.4E**).

*4.4.5 Data analysis: response index*

We quantified for each neuron whether its response differed between fearful and happy trials using a single-trial response index $R_i$ (**Eq. 4.3**; see **Figure 4.8**). The response index can facilitate group analysis and comparisons between different types of cells (i.e., fear and happy selective cells in this study), as motivated by previous studies (Rutishauser et al., 2008, Rutishauser et al., 2011). The response index quantifies the response during trial *i* relative to the mean response to correct happy stimuli and baseline (a 1 second interval of blank screen right before scramble onset). The mean response and baseline was calculated individually for each unit.

$$R_i = \frac{FR_i - mean(FR_{HappyCorrect})}{mean(FR_{Baseline})} \cdot 100\% \qquad \textbf{(Eq. 4.3)}$$

For each trial *i*, which can be either fearful or happy, $R_i$ is the baseline normalized firing rate (FR) during a 1.5-second interval 250 ms post stimulus-onset (the same time interval as cell selection). Different time intervals were tested as well, to ensure that results were qualitatively the same and not biased by particular spike bins.

If a neuron distinguishes happy from fearful trials, the average value of $R_i$ will be significantly different from 0. Since fear-selective neurons have more spikes in fearful trials and happy-selective neurons have more spikes in happy trials (the selection process is described above), on average $R_i$ is positive for fear-selective neurons and negative for happy-selective neurons. To get an aggregate measure of activity that pools across neurons, $R_i$ was multiplied by -1 if the *neuron* is classified as a happy-selective neuron (**Eq. 4.4**). This makes *Ri* on average positive for both types of emotion-selective neurons. Notice that the factor -1 depends only on the *neuron* type, which is determined by t-tests

on correct trials as described above, but not trial type. Thus, negative $R_i$ values are still possible.

$$R_i = -\frac{FR_i - mean(FR_{HappyCorrect})}{mean(FR_{Baseline})} \cdot 100\% \qquad \textbf{(Eq. 4.4)}$$

After calculating $R_i$ for every trial, we subsequently averaged all $R_i$ of trials that belong to the same category. We used four categories: fearful correct (FC), fearful incorrect (FI), happy correct (HC) and happy incorrect (HI). By definition, the average value of $R_i$ for HC trial will be equal to zero because the definition of $R_i$ is relative to the response to happy correct trials (see **Eq. 4.4**). The mean baseline firing rate was calculated across all trials. The same $FR_{HappyCorrect}$ was subtracted for both correct and incorrect trials.

The cumulative distribution function (CDF) (see **Figure 4.8D** and **Figure 4.9A,C**) was constructed by calculating for each possible value $x$ of the response index how many examples are smaller than $x$. That is, $F(x) = P(X \leq x)$, where $X$ is a vector of all response index values. The CDF of fearful and happy trials were compared using two-tailed two-sample Kolmogorov–Smirnov tests. All error bars are ± SE unless indicated otherwise.

### *4.4.6 Data analysis: split analysis and permutation test*

We used 1000 runs for the permutation analysis. In each run, we randomly selected half of the correct trials to identify emotion-selective units and to determine the neuron type (as described above). We then used the remaining half of correct trials to calculate the response indices. This makes the response index values statistically independent of the cell selection. We also calculated the responses indices for all the incorrect trials for the selected cells.

To summarize the population difference in response to fearful compared to happy faces, we calculated a summary metric that provided a single number for a population of cells (see **Figure 4.10** and **Figure 4.16**). This provided a single quantity for every run of the

permutation test. The population summary metric is equal to the difference between the average of response indices from all fearful trials (either correct or incorrect) collapsed across all selected cells and the average of response indices from all happy trials (either correct or incorrect) collapsed across all selected cells (**Eq. 4.5**).

$$Metric = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N_i^F} R_{i,j}^F}{N_{i,j}^F} - \frac{\sum_{i=1}^{N}\sum_{j=1}^{N_i^H} R_{i,j}^H}{N_{i,j}^H} \qquad \textbf{(Eq. 4.5)}$$

in which $R_{i,j}^F$ is the response index for the *j*-th fearful trial of the *i*-th selected cell, $R_{i,j}^H$ is the response index for the *j*-th happy trial of the *i*-th selected neuron, $N$ is the total number of selected cells, $N_i^F$ is the total number of fearful trials of the *i*-th selected cell, $N_i^H$ is the total number of happy trials of the *i*-th selected cell, $N(R_{i,j}^F)$ is the total number of viable fearful trials from all selected cells, and $N(R_{i,j}^H)$ is the total number of viable happy trials from all selected cells.

Note that the summary metric can be calculated either for correct or incorrect trials. Thus, for the summary metric for correct or incorrect trials, the fearful/happy trials refer only to correct or incorrect trials, respectively. We analyzed correct trials and incorrect trials separately to derive population metric distributions shown in **Figure 4.15** and **Figure 4.17**.

The selected cells can be fear-selective neurons, happy-selective neurons and pooled populations of fear and happy selective neurons. Note that for all happy-selective neurons, we flipped the sign of all the response indices for that neuron (refer to **Eq. 4.4**), so that we can combine them with fear-selective neurons to get a pooled population.

To quantify how sensitive neurons were to specific facial parts, we repeated the permutation analysis with only a subset of trials that revealed the ROI of interest. First, we selected trials according to the overlap of the bubbles with the specified eye and

mouth ROIs (as shown in **Figure 4.2B**). The more overlap between bubbles and ROIs, the more is revealed within the ROI. We chose two categories of ROI trials: those where predominantly only the eye or the mouth was shown. This was achieved by enforcing either ‘High Eye AND Low Mouth’ overlap or ‘Low Eye AND High Mouth’ overlap. ‘High’ or ‘Low’ here was above or below the median of the overlapping values across all correct trials. Selection of trials based on ROIs revealed was only based on the stimulus shown to the patient and did not involve the neuronal response. We subsequently repeated the permutation analysis as described above on the subset of trials that revealed the desired ROI. There were too few incorrect trials that satisfied the strict ROI criteria and these are thus not considered for the ROI analysis.

To estimate the expected difference in the population summary metric as well as the number of significant units, we performed the same analysis but randomly scrambled the trial labels. This resulted in an empirical estimate of the null distribution. We used 1000 runs.

### *4.4.7 Electrode localization from structural MRIs*

To identify electrode recording sites in the amygdala, T2 relaxation times were measured using spin-echo dual-echo sequences on a 1.5-T Toshiba MR scanner. 25 contiguous axial slices were acquired (0.575 × 0.575 mm in-plane, 5 mm thick, TR = 5777.5 ms, TE = 105 ms, flip angle = 90°). The imaging slices covered the entire brain, including the amygdala. The electrodes were clearly visible as dark lines in the T2 scans. Images were subsequently processed using SPM8 (Friston, 2007). Scans were first segmented and normalized to the standard MNI space. The electrode tip coordinates were visualized and manually labeled in FSL (Jenkinson et al., 2012). A mask was created in MATLAB for each patient with each recording site as a 3 × 3 × 3 mm cube centered on the identified electrode tip. All masks were then overlaid on the standard MNI152 template with 1 mm isotropic resolution.

*4.4.8 Eye tracking*

Eye tracking was carried out using a non-invasive infra-red remote Tobii X300 system. It was recorded at 300 Hz with a 23-inch screen (screen resolution: 1920x1080). Both eyes were recorded. A professional visualization software (Tobii Studio™ 2.2) manufactured by Tobii was used together to record the eye movements and perform the gaze analysis. Fixations were detected by Tobii Fixation Filter implemented in Tobii Studio. The Tobii Fixation Filter is a classification algorithm proposed by Olsson (Olsson, 2007) and detects quick changes in the gaze point using a sliding window averaging method. Velocity threshold was set to 35 [pixels/samples] and distance threshold was set to 35 [pixels] in our study.

Fixations were smoothed by a 40-pixel Gaussian kernel with a standard deviation of 10 pixels (the same as is used for display of the stimuli to the subjects and in the analysis of the spikes). Each heat map indicates the probability of fixating a given location (in arbitrary units). The fixation probability is calculated based on the number and duration of fixations. The heat maps show an average over all face trials. Maps M(x,y) are calculated as following: i) the average map is initialized M(x,y) = 0 at all locations. ii) Each fixated location is marked by setting $M(x,y) = M(x,y)+n_i$, where $n_i$ is the number of samples of fixation i (corresponding to the duration of fixation i). iii) the number of fixations from each subject was normalized by the total number of fixations N of this subject by setting M(x,y) = M(x,y)/N, resulting in the probability of fixating each pixel; the fixation maps M(x,y) from all subjects were then averaged and lastly smoothly by the Gaussian kernel. This procedure ensured an equal contribution from each subject and thus the statistical independence between subjects.

We performed a post-hoc drift-correction procedure for each trial. Before the presentation of faces, a fixation circle superimposed on a scrambled face image was presented for a random duration between 800 to 1200 ms. We assumed subjects fixated on the fixation

circle during this period (which was confirmed by visual inspection) and hence we subtracted the mean fixation position of the last 500 ms during this fixation period from all subsequent fixations during the face presentation period. We excluded trials in which more than 70% of the fixations were not within the face image.

To quantitatively compare the fixation densities within certain parts of the face, we defined three regions of interest (ROIs): eyes (left and right), mouth and center. Each ROI is round and has a radius of 30 pixels. Paired t-tests were performed to compare the fixation densities within the ROIs between fearful and happy trials during the 500 ms stimulus presentation period.

## 4.5 Results

### *4.5.1 Behavioral performance*

We recorded single neurons in the human amygdala while neurosurgical patients performed an emotion discrimination task (**Table 4.1**; see **Figure 4.1** for recording sites for each patient). All patients (9 sessions from 7 patients in total; 2 patients did 2 sessions; neurons from each individual recording session are considered independent even if they are from the same patient) were undergoing epilepsy monitoring and had normal basic ability to discriminate faces. Six healthy subjects (6 sessions) served as behavioral controls and participated in the same experiment. Subjects were asked to judge, for every trial, whether the stimulus was fearful or happy by pushing corresponding buttons as quickly and accurately as possible (**Figure 4.2**). Each trial was fearful or happy with 50% probability. No other attribute of the stimuli (such as identity) predicted the emotion. Each stimulus was preceded by a phase-randomized baseline image of equal luminance and complexity (‘scramble’). Trials with no response (timeouts, see **Methods**) were excluded from analysis.

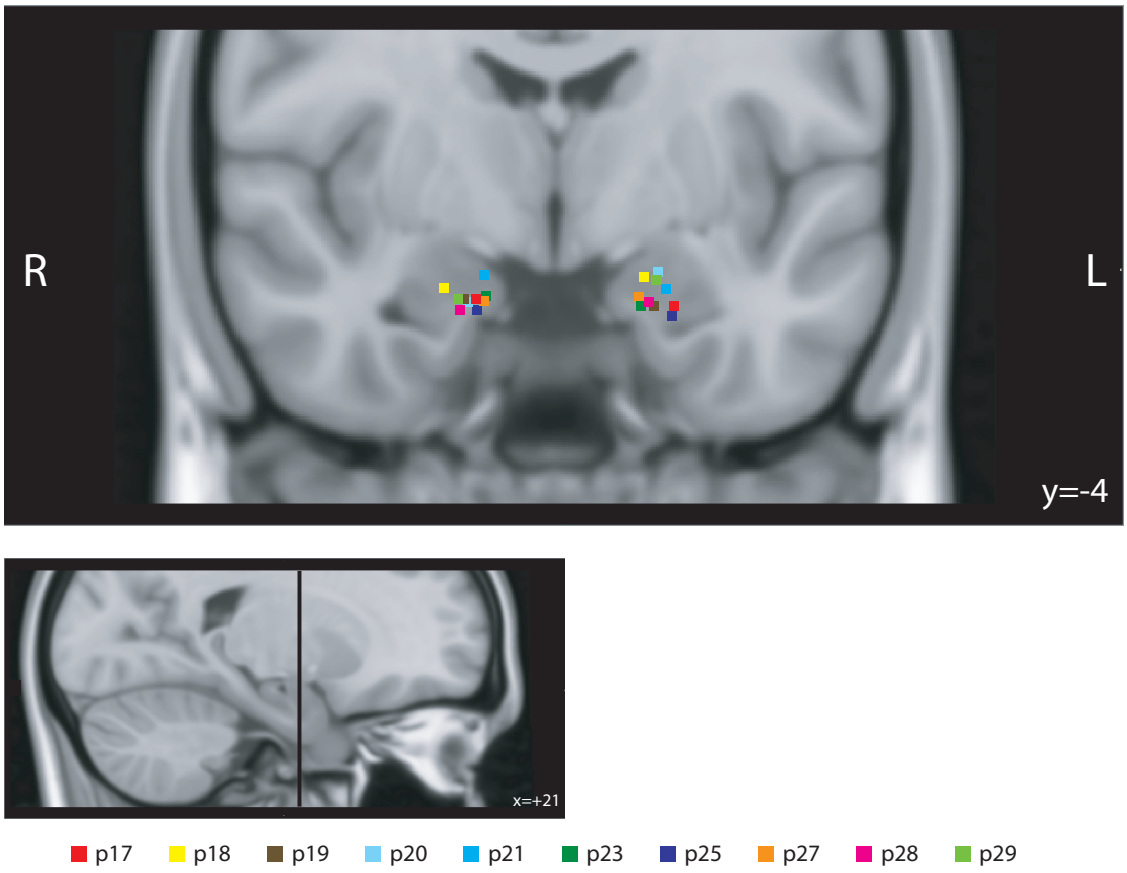


**Figure 4.1. Recording sites mapped from post-implantation MRIs.**

Each square represents the recording site of the electrode; different colors correspond to individual patients (codes indicated at bottom legend).

We showed randomly selected parts of faces ('bubbles'; **Figure 4.2B**) that allowed us to derive a behavioral classification image (BCI) (Gosselin and Schyns, 2001) based on accuracy and reaction time (RT) of the responses (derived separately for happy trials and fearful trials; **Figure 4.2C**). The BCI shows, for every pixel, whether revealing this pixel is likely to increase accuracy and decrease RT. The higher a pixel's value, the more it contributed to behavioral judgment in the task. BCIs from patients and controls did not differ within key facial features (ROIs used are shown in **Figure 4.2B**; two-tailed unpaired t-test comparing average z-scores within the ROIs: for fearful trials, $p = 0.51$ for

eyes and p = 0.36 for mouth; for happy trials, p = 0.68 for eyes and p = 0.14 for mouth), confirming that patients performed the task with a normal strategy. Both patients and controls primarily utilized information revealed by eyes to judge fearful faces, while they utilized more mouth information to judge happy faces, consistent with previous studies (Smith et al., 2005, Scheller et al., 2012).

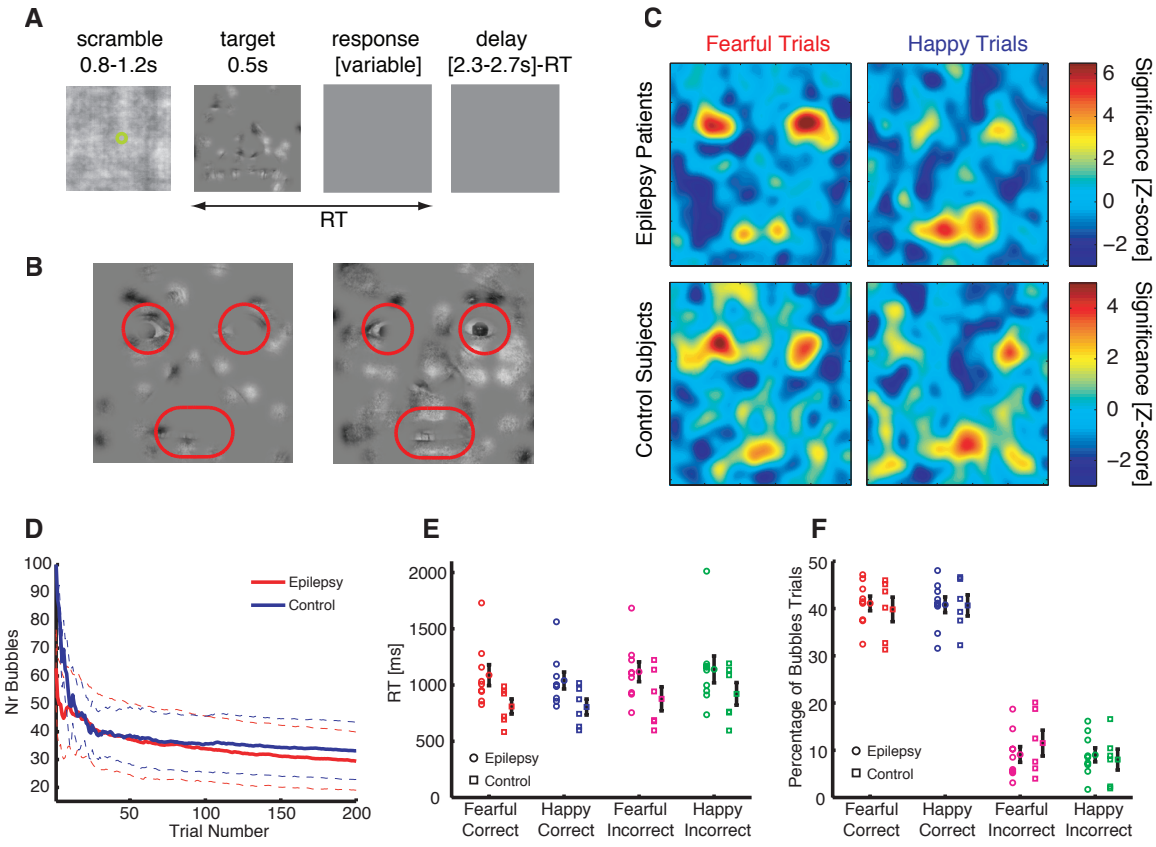


**Figure 4.2. Stimuli and behavioral performance.**

**(A)** Task structure. Immediately preceding the target image, a scrambled version of a face was presented for a variable time between 0.8 and 1.2 s. The target image was presented for 500 ms and showed either a fearful (50%) or happy (50%) expression. Subjects indicated whether the presented face was happy or fearful. **(B)** Example bubbles stimuli. The regions of interest (ROIs) used for analysis are shown in red (not shown to subjects).

**(C)** Behavioral classification images for fearful and happy trials for the neurosurgical patients and control subjects. Color code is the *z* scored correlation between the presence or absence of a particular region of the face and behavioral performance. **(D)** Learning curve for both patients (n = 8 sessions, one session omitted here because the learning algorithm was disabled as a control; mean ± standard error of the mean [SEM]) and controls (n = 6 sessions). Only first 200 trials are shown. **(E)** Reaction time for patients (n = 9 sessions, circles) and controls (n = 6 sessions, squares). Each data point represents a single recording session and the error bars denote SEM of the mean. Fearful Correct: fearful trials with a correct response; Happy Correct: happy trials with a correct response; Fearful Incorrect: fearful trials but incorrectly judged as happy; Happy Incorrect: happy trials but incorrectly judged as fearful. **(F)** Response choice for patients (n = 9 sessions, circles) and controls (n = 6 sessions, squares).

The proportion of the face revealed in the bubbles stimuli was adaptively modified to achieve an asymptotic target performance of 80% correct; the number of bubbles required to achieve this criterion decreased, on average, over trials (**Figure 4.2D**). Patients completed on average a total of 401 trials, and the average number of bubbles required ranged from 100 at the beginning to 29.4 ± 29.5 on the 200$^{th}$ trial (n = 8 sessions, mean ± standard deviation [SD]; one session omitted here because the learning algorithm was disabled as a control). Control subjects completed on average a total of 216 trials, and the average number of bubbles was 33.2 ± 25.2 on the 200$^{th}$ trial, showing no statistical difference compared to controls (t-test, $p = 0.81$).

RT was in general longer for patients (n = 9 sessions, 1047 ± 197 ms, mean ± SD, relative to stimulus onset) than control subjects (n = 6 sessions, 793 ± 152 ms; $p < 0.05$; **Figure 4.2E**). This is likely due to the uncontrolled hospital environment compared to well-controlled lab environment. However, there is no significant RT difference between fearful and happy trials or correct and incorrect trials for both patients and control

subjects (two-way ANOVA; Emotion [fearful vs. happy] X Correctness [correct vs. incorrect]: for Patients, $F(1,32) = 0.0193$, $p = 0.89$ for Emotion, $F(1,32) = 0.455$, $p = 0.51$ for Correctness, and $F(1,32) = 0.128$, $p = 0.72$ for interaction; for Controls, $F(1,20) = 0.0491$, $p = 0.83$ for Emotion, $F(1,20) = 1.09$, $p = 0.31$ for Correctness, and $F(1,20) = 0.0759$, $p = 0.79$ for interaction).

Average accuracy across all trials was (by design) 81.9 ± 8.3% for patients (n = 9 sessions, mean ± standard deviation [SD]) and 80.5 ± 4.1% for control subjects. There was no significant difference in accuracy between patients and control subjects (two-tailed t-test: $p = 0.64$). Importantly, there was no difference in the proportion of "fearful" or "happy" responses for both correct trials ($p = 0.87$ for patients and $p = 0.86$ for controls) and incorrect trials ($p = 0.95$ for patients and $p = 0.49$ for controls), showing that neither patients nor controls had any response bias (**Figure 4.2F**). Overall, the behavioral performance-related metrics confirmed that patients were alert and attentive and had largely normal ability to discriminate emotion from faces.

### *4.5.2 Eye tracking*

We asked three patients to conduct an eye-tracking experiment in the laboratory after completion of their surgery. Patients were shown the identical stimuli they saw in the hospital ("replay"), while we recorded eye movements. We also recruited six neurotypically developed control subjects for this task. We computed the fixation density maps during the 500 ms stimulus presentation period and subsequently performed ROI analysis.

Epileptic patients had similar fixation patterns as control subjects, showing that patients used the same strategy as neurotypical subjects to discriminate emotional faces and the eye movement patterns were not affected by the epileptic morbidity or the experimental conditions in the hospital. Further, the density maps are similar between fearful and happy trials for both patients and control subjects (see **Figure 4.3**).

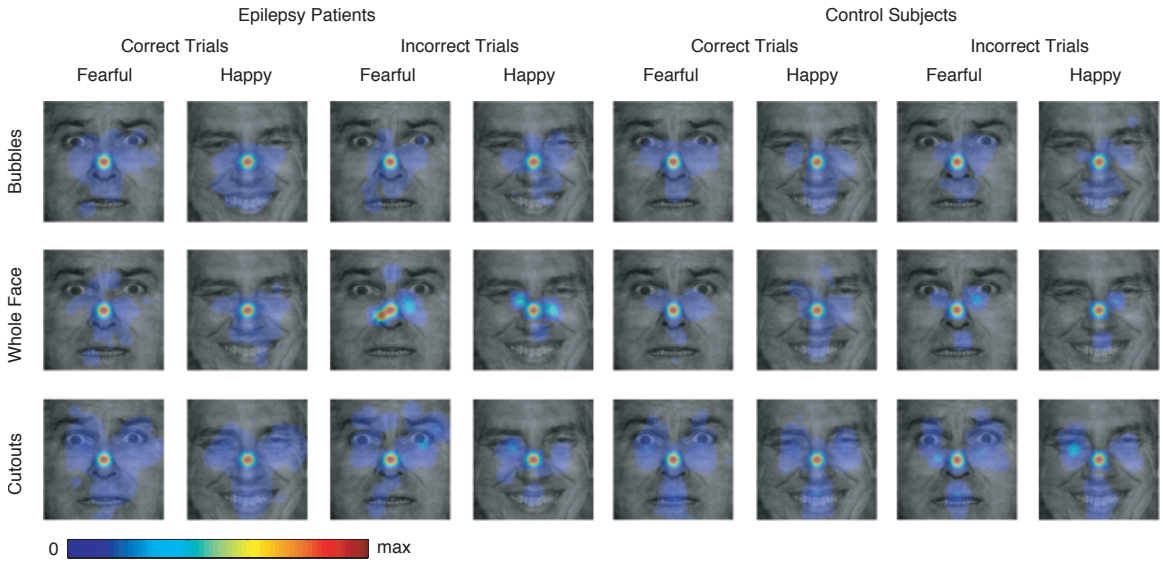


**Figure 4.3. Fixation density maps.**

Each heat map indicates the probability of fixating a given location in arbitrary units (blue: zero probability, and red: maximal probability). The sum of density is 1 in each plot. Fixations were smoothed by a 40-pixel Gaussian kernel with a standard deviation of 10 pixels (the same as is used for display of the stimuli to the subjects and in the analysis of the spikes).

For epileptic patients, quantitative ROI analysis comparing fearful and happy bubbles trials showed that the large majority of fixations within the 500 ms stimulus presentation time remained on the center of the face (69.5% for fearful correct trials and 71.6% for happy correct trials, $p = 0.21$). While few saccades away from the center were made, these nevertheless showed the expected differences between fear vs. happy: mainly, eyes were more likely to be fixated in fearful compared to happy trials (5.26% vs. 1.75%, respectively, $p = 0.02$); whereas the fixation probability for the mouth was 1.82% vs. 3.5% for fearful and happy trials, respectively, showing a tendency of higher density on the mouth ($p = 0.075$). These results are consistent with the behavioral classification

images and consistent with previous studies (Scheller et al., 2012). Interestingly, for incorrect trials this difference was no longer significant (fear vs. happy: 2.13% vs. 5.76% for eyes ($p = 0.33$) and 1.33% vs. 1.23% for mouth ($p = 0.51$)).

Similar results were observed for control subjects: the large majority of fixations also remained on the center of the face (77.6% for fearful correct trials and 78.4% for happy correct trials, $p = 0.81$). Higher fixation density was observed on eyes in fearful compared to happy trials (4.9% vs. 1.47% for fearful vs. happy, respectively, $p = 0.0068$), whereas the fixation probability for the mouth was 1.86% vs. 5.92% for fearful vs. happy, respectively ($p = 0.2$). Still, for incorrect trials, the significant difference for eyes no longer exists (fear vs. happy: 5.1% vs. 1.23% for eyes ($p = 0.12$) and 1.46% vs. 4.46% for mouth ($p = 0.3$)).

### *4.5.3 Emotion-selective neurons*

210 single units were isolated from 9 recording sessions in 7 patients. Of these, 185 units (102 in the right amygdala, 83 in the left) that had an average firing rate of at least 0.2 Hz were chosen for further analysis. Structural MRI analyses of the amygdala with the electrodes in situ showed that recordings were mostly from the basomedial and basolateral nucleus of the amygdala (**Figure 4.1**). Electrodes were positioned such that their tips were located in the upper third to center of the deep amygdala, ~7 mm from the uncus. Microwires projected medially out at the end of the depth electrode and electrodes were thus likely sampling neurons in the midmedial part of the amygdala (basomedial nucleus or deepest part of the basolateral nucleus; (Oya et al., 2009)). The isolation criteria and other face-responsive characteristics of the same dataset were described previously (Rutishauser et al., 2011, Rutishauser et al., 2013). To analyze neuronal responses, we aligned all trials to the onset of the face. The firing rate was normalized by dividing by average baseline (the firing rate 500 ms prior to scramble onset) across all trials, separately for each unit.

We here investigate the response characteristics of the amygdala neurons to emotions. We define emotion-selective units as those that responded differentially to fearful faces compared to happy faces. We selected emotion-selective units by comparing the total number of spikes in a time window 250 ms to 1750 ms post stimulus-onset between correct fearful trials and correct happy trials. A trial was classified as correct if the subject indicated the emotion associated with the stimulus displayed (ground truth). We used a one-tailed t-test to identify units with a greater response to fearful faces or happy faces separately, each with $\alpha = 0.05$. We found that 24 units showed significantly greater response to fearful faces compared to happy faces (13.0%, binomial test on the number of significant cells: $p < 0.00001$) and 17 units (9.2%, $p < 0.01$) that showed a greater response to happy faces compared to fearful faces. We refer to these units as neurons selective for fearful expressions ("fear-selective" for short) (**Figure 4.4A–B**) and neurons selective for happy expressions ("happy-selective" for short) (**Figure 4.4C–D**), respectively. The probability of observing 41 emotion-selective neurons in a population of 185 neurons by chance is very low ($p < 10^{-6}$, estimated by a binomial distribution with false positive rate of 0.1 for each neuron due to performing two one-tailed tests at $p < 0.05$), indicating that amygdala neurons signal information about emotions (see **Table 4.2**). However, it is important to emphasize that we do not know the response selectivity of the same neurons to other stimuli. In particular, it is possible that the same neurons would also respond to other emotions that we did not test in this study. Our labels of units as fear- or happy-selective are not meant to imply that these units would not respond to other, not tested, emotions or stimuli.

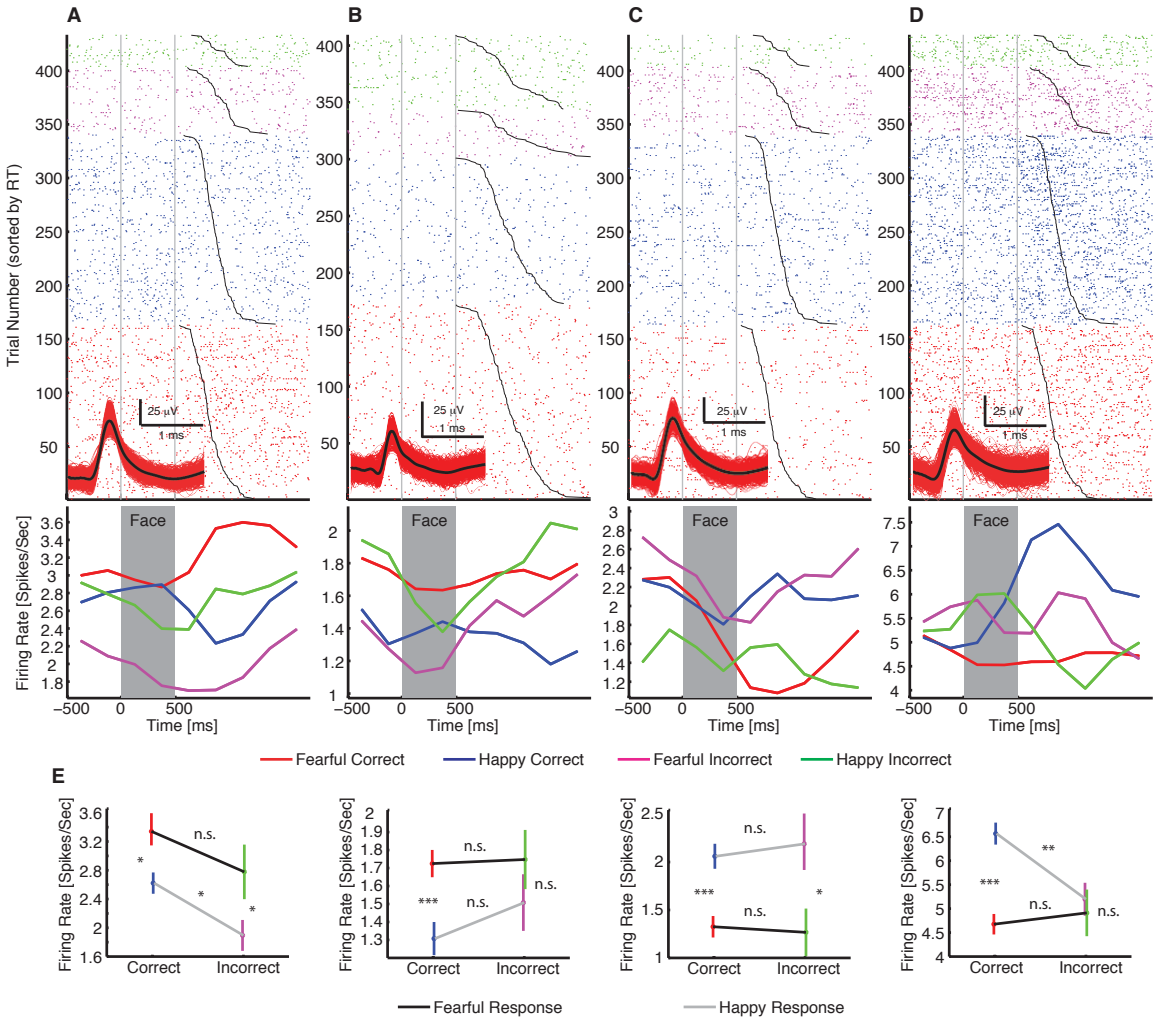


**Figure 4.4. Single-unit examples of emotion-selective neurons in the amygdala.**

**(A–B)** Example fear-selective neurons, which have a higher firing rate for correct fearful trials compared to correct happy trials (selection t-test: $p < 0.005$). **(C–D)** Example happy-selective neurons, which have a higher firing rate for correct happy trials compared to correct fearful trials (selection t-test: $p < 10^{-8}$). Each raster (top) and post-stimulus time histogram (PSTH) (bottom) is shown with color coding as indicated. Trials are aligned to face stimulus onset (dark gray shade, fixed 500 ms duration). Trials within each stimulus category are sorted according to reaction time (black line). Waveforms for each unit are shown at the bottom of the raster plot. **(E)** Average firing rate 250–1750 ms post stimulus-onset for each unit. Red: fearful trials with a correct response; blue: happy

trials with a correct response; magenta: fearful trials but incorrectly judged as happy; green: happy trials but incorrectly judged as fearful. Black lines connect conditions with the same response: fear (black) and happy (gray). Note that the lines do not cross, implying that whatever response tuning the neuron had was maintained regardless of whether the response was correct or not. Error bars denote ± SEM across trials. Two-tailed t-tests were applied to compare between conditions. *: $p < 0.05$, **: $p < 0.01$ and ***: $p < 0.001$. n.s.: not significant ($p > 0.05$).

**Table 4.2. Summary of neuronal response characteristics.**

All percentages are derived from a total of 185 units. Comparisons between scramble and face stimuli was performed using a two-tailed paired t-test at $p < 0.05$. Selection of interactive neurons was performed using a two-tailed t-test at $p < 0.05$. Selection of emotion-selective cells was performed using a one-tailed t-test at $p < 0.05$.

Patients with autism show abnormal processing of faces and we previously showed that neurons in the patients with autism responded significantly more to the mouth but less to the eyes (Rutishauser *et al.* 2013). In total, there are 83 cells (40.0% of 210 cells) recorded from two neurosurgical patients with autism (3 sessions), and among these cells, 62 (74.7%) are from the left amygdala and 21 (25.3%) are from the right amygdala. Regarding emotion coding, 8 cells (9.64%) are fear-selective and 6 cells (7.23%) are happy-selective. Compared to neurosurgical patients without autism, 127 (60%) cells were recorded from 6 sessions, among which 30 (23.6%) are from the left amygdala and 97 (76.4%) are from the right amygdala. 16 cells (12.6%) are fear-selective and 11 cells (8.67%) are happy-selective (Note for all 210 cells, 24 (11.4%) cells are fear-selective and 17 cells (8.1%) cells are happy-selective). The percentage of emotion-selective cell from each subject is about equal between patients with vs. without autism (a three-way mixed model ANOVA: autism X emotion X laterality; main effect of autism: $F(1,17) = 7.67\times10^{-6}$, $p = 1.00$), as well as between left vs. right amygdala (main effect of laterality:

F(1,17) = 2.86, p = 0.11). There is no main effect of emotion (F(1,17) = 0.027, p = 0.80) nor interactions (all ps > 0.52). Furthermore, there is no difference across subjects (nested factor: F(7,17) = 0.31, p = 0.94).

| **Response Characteristics** | **Nr cells** | **% cells** |
|---|---|---|
| Face responsive (face stimulus vs. scramble, all trial categories pooled) | 95 | 51.4% |
| Interactive neurons | 23 total<br>10 fearful<br>13 happy | 12.4%<br>5.4%<br>7.0% |
| Emotion-selective neurons | 41 total<br>24 fearful-selective<br>17 happy-selective | 22.2%<br>13.0%<br>9.2% |
| Emotion-selective AND Interactive neurons | 9 | 4.9% |
| Emotion-selective AND Whole-face neurons<br>(see **Discussion** and (Rutishauser *et al.* 2011)) | 8 | 4.3% |
| Emotion-selective AND autism neurons (see below) | 14 | 7.6% |
| Laterality | 83 left amygdala<br>102 right amygdala | 44.9%<br>55.1% |

**Figure 4.4** shows four single-neuron examples (see **Figure 4.5** for more examples). The fear-selective neurons (**Figure 4.4A–B**) increased their activity for fearful trials and decreased their activity in happy trials. In contrast, the happy-selective neurons (**Figure 4.4C–D**) increased their activity in happy trials. On average, significant differences in response between fearful and happy faces appeared 625 ms post stimulus-onset and lasted for up to 1.5 seconds (**Figure 4.6**). For fear-selective neurons, the difference was mainly due to a suppression of activity in happy trials (**Figure 4.6A**), whereas for happy-selective neurons, it was mainly due to an increase in activity for happy trials (**Figure 4.6B**). A similar plot for all recorded neurons (n = 185, **Figure 4.7**) showed no significant difference, indicating that overall mean activity was not different between the two conditions.

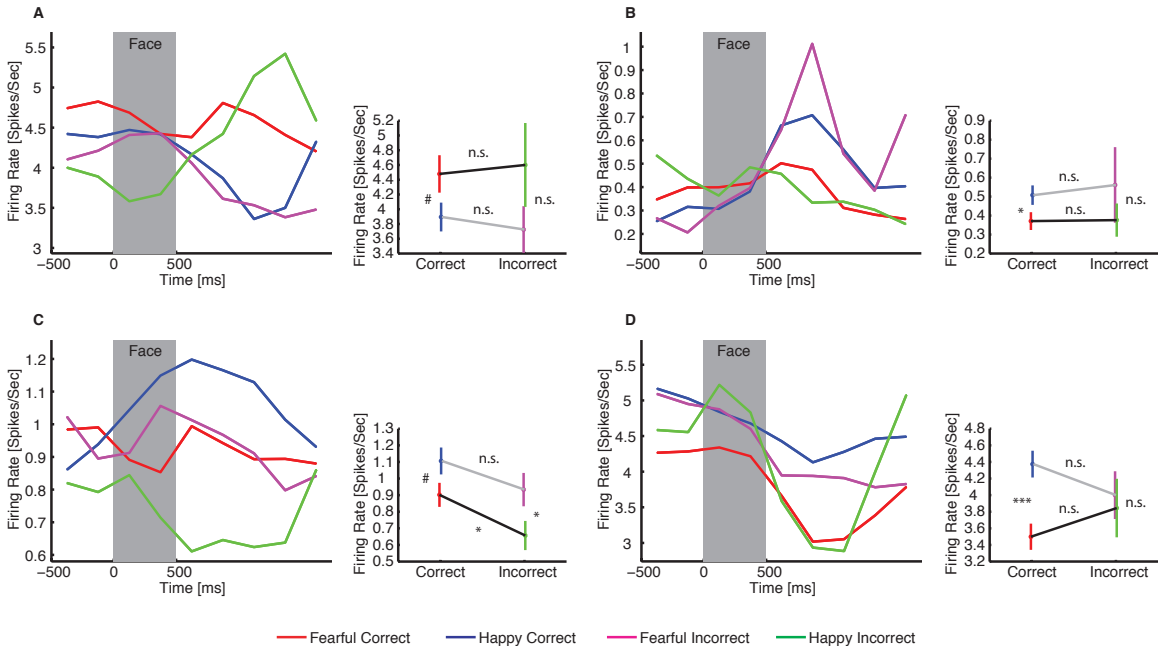


**Figure 4.5. More single-unit examples of emotion-selective neurons in the amygdala.**

**(A–D)** Each panel represents a single-unit example (all from different patients). Post-stimulus time histograms (PSTH) (left) are shown with color coding as indicated. Trials are aligned to face stimulus onset (dark gray shade, fixed 500 ms duration). On the right are plots of average firing rate in a time window 250 ms to 1750 ms post stimulus-onset of each unit. Red: fearful trials with a correct response; blue: happy trials with a correct response; magenta: fearful trials but incorrectly judged as happy; green: happy trials but incorrectly judged as fearful. Lines connect conditions with the same response: black lines for a fearful response and gray lines for a happy response. Note that the lines do not cross, implying that whatever response tuning the neuron had was maintained regardless of whether the response was correct or not. Error bars denote standard error of the mean. Two-tailed t-tests were applied to compare between conditions. #: $p < 0.1$, *: $p < 0.05$, **: $p < 0.01$ and ***: $p < 0.001$. n.s.: not significant ($p > 0.05$).

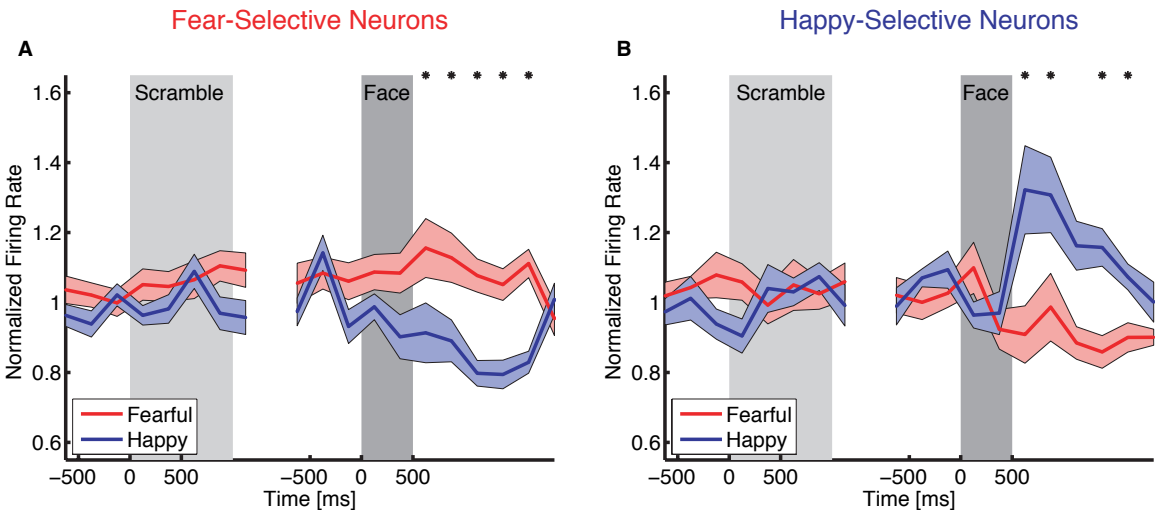


**Figure 4.6. Average PSTH of all emotion-selective neurons.**

**(A)** Mean response of all fear-selective units that increased their spike rate for correct fearful trials compared to correct happy trials (n = 24 units; shaded area denotes ± SEM; the firing rate was normalized to average baseline response for each unit separately). **(B)** Mean response of all happy-selective units for correct trials (n = 17 units). Asterisk indicates a significant difference between the response to fearful trials and happy trials in that bin ($p < 0.05$, two-tailed t-test, Bonferroni-corrected).

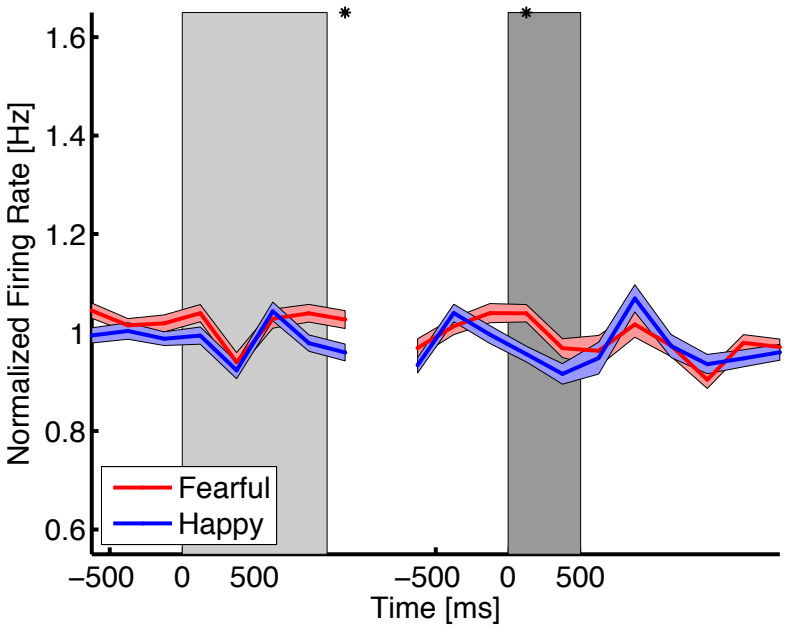


**Figure 4.7. Average PSTH of all 185 neurons (correct trials only).**

The only significant difference appears 125 ms after stimulus onset, showing no bias of the population as a whole towards one of the emotions (Bonferroni corrected).

So far we only considered trials where patients judged the emotion expressed correctly. Here, correctness was assessed by the ground truth of the stimuli, which control subjects have classified unequivocally as either happy or fearful when shown the entire face for extended periods of time (Ekman and Friesen, 1975). How did the same neurons respond during errors in emotional judgment? We next compared the neuronal response during incorrect trials to the response during correct trials (for which the neurons were selected in the first place, see above). We found that the neuronal response during incorrect trials was similar to the one for the same behavioral response during correct trials rather than the actual emotion shown. For example, when a fearful face was incorrectly judged as happy, the neurons responded as if a happy face was correctly judged as happy (and vice-versa; compare magenta and blue lines for the examples shown in **Figure 4.4A–B**). Similarly, when a happy face was incorrectly judged as fearful, the neurons responded as

if a fearful face had been correctly judged as fearful (compare green and red lines for the examples shown in **Figure 4.4C–D**). In **Figure 4.4E**, lines connect the conditions with the same response (fearful or happy). Note that the lines do not intersect, indicating that the relationship between the responses for the two emotions was similar in correct and incorrect trials, regardless of overall mean firing rate. For example, if a neuron showed a greater response in fearful correct trials compared to happy correct trials, it would also show a greater response in happy incorrect trials compared to fearful incorrect trials. Thus, firing rate increased whenever a fearful judgment was made, regardless of whether it was correct or incorrect. The neuronal response of the examples shown in **Figure 4.4** thus indicated the subjective perceptual judgment that subjects made, rather than the ground truth of the emotion shown in the stimulus. A significant interaction between stimulus emotion (fearful/happy) and accuracy of judgment (correct/incorrect) as tested by a 2x2 ANOVA with number of spikes fired in a 1.5 second window after stimulus onset (250–1750 ms post stimulus onset) confirmed this impression: the interaction term was significant for all example neurons at $p < 0.01$ ($F(1,429) = 9.04$ for **Figure 4.4A**, $F(1,405) = 7.09$ for **Figure 4.4B**, $F(1,429) = 16.06$ for **Figure 4.4C**, and $F(1,429) = 9.47$ for **Figure 4.4D**). We next quantified this phenomenon across the population.

### *4.5.4 Interactive neurons encode perceptual judgment of emotions other than ground truth shown in stimulus*

We next assessed for all neurons whether there was a significant interaction between the emotion shown and the correctness of the subject's judgment using a two-way ANOVA ([correct vs. incorrect trials] X [fearful stimuli vs. happy stimuli]). Units with a significant interaction term are referred to as "interactive units" henceforth, and reflect the subjective judgment regardless of the emotion shown in the image. There were 23 units with a significant interaction term (12.4%, binomial test $p < 0.00005$), 10 of which responded with a higher firing rate in correct fearful trials and 13 of which responded

with a higher firing rate in correct happy trials, hence denoted as fearful interactive neurons and happy interactive neurons, respectively (see **Table 4.2**).

To further quantify the response of the interactive neurons, we next plotted the average baseline-normalized firing rate for correct and incorrect trials for each interactive neuron (**Figure 4.8A,B**). Each neuron contributed two data points: one for correct (red, blue) and one for incorrect trials (gray), respectively. By definition, fearful interactive neurons increased their firing rate for correctly identified fearful face trials (**Figure 4.8A**, red). Similarly, happy interactive neurons increased their firing rate for correctly identified happy face trials (**Figure 4.8B**, blue). Incorrect trials (gray dots), in contrast, tended to have greater firing rates for the emotion opposite to the one actually shown in the stimulus (Fearful interactive neurons: **Figure 4.8A**; $\chi^2$-test on the number of neurons falling on each side of the diagonal line (gray bars), $p < 10^{-5}$; Happy interactive neurons: **Figure 4.8B,** $p < 0.01$). In each case, the mean of all incorrect trials from all neurons was on the opposite side of the diagonal shown in **Figure 4.8** from that for correct trials: the average normalized firing rate thus indicated the behavioral judgment of the subjects.

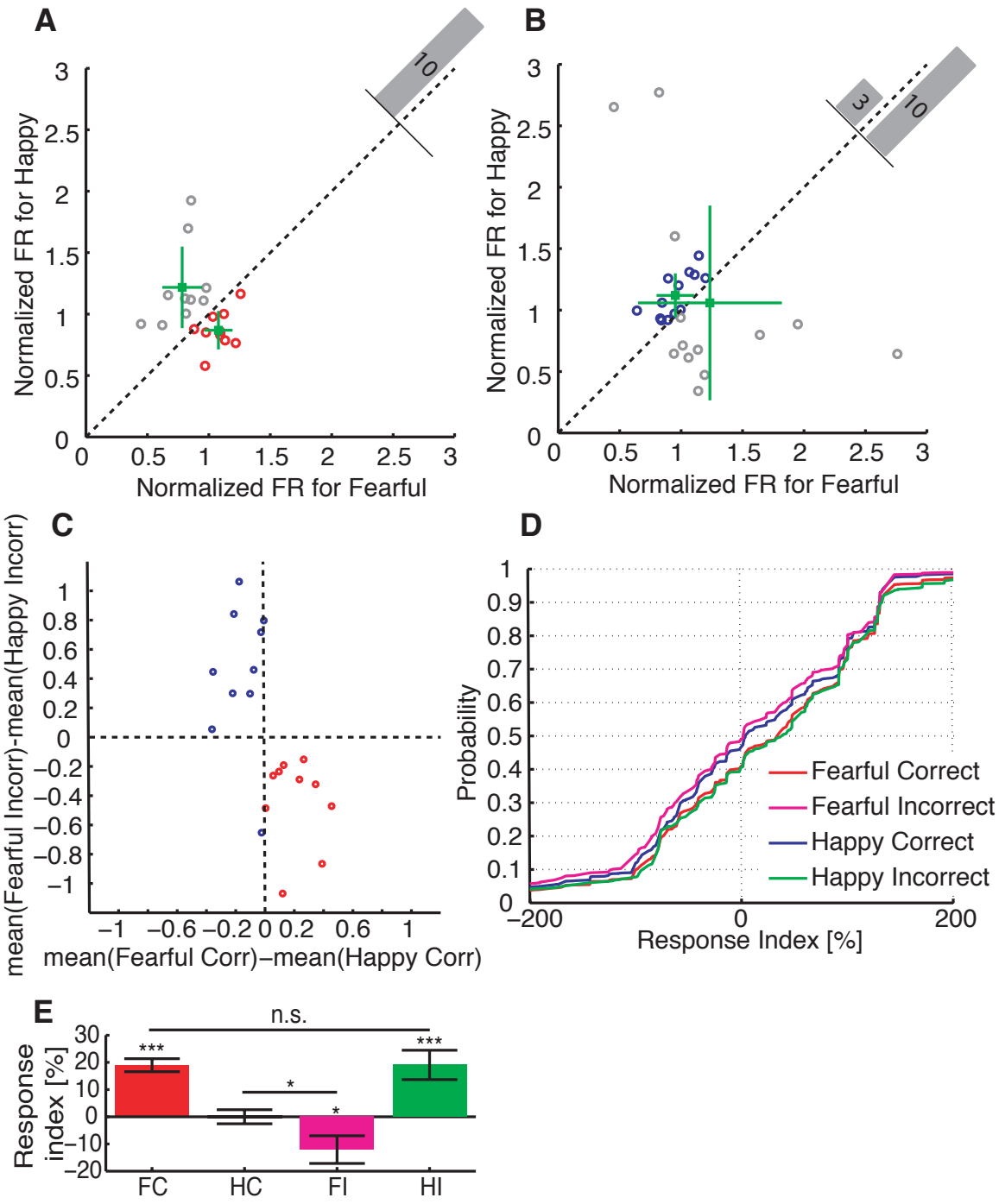


**Figure 4.8. Interactive neurons followed perceptual judgment rather than stimulus identity.**

N = 23 units, selected by a significant interaction term. **(A,B)** Scatter plot of mean normalized firing rate for fearful and happy trials, shown separately for fearful interactive

**(A)** and happy interactive **(B)** neurons. Red and blue dots denote correct trials, which were distributed below and above the diagonal, respectively (by definition). Gray dots denote incorrect trials, which tended to be distributed on the opposite side of the diagonal line compared to the correct trials. Error bars (Green) are mean ± SD. Gray bars (upper right) show the number of neurons falling on each side of the diagonal. **(C)** Scatter plot of the normalized firing rate difference comparing the response to fearful and happy trials for correct (x-axis) and incorrect (y-axis) trials. Fearful interactive neurons (red) and happy interactive neurons were largely located in the lower right and upper left, respectively. **(D)** Cumulative distribution of the response index (see **Methods**). The response during incorrect trials was opposite to the one during correct trials, implying that the neuronal response followed the behavioral judgment. **(E)** Mean response index across all trials. Note similar response magnitude for fearful correct (FC) and happy incorrect trials (HI), which shows that when a happy face was shown but judged as a fearful face, the neurons responded as if a real fearful face was shown. The same interpretation can be derived for happy correct trials (HC) and fearful incorrect trials (FI). FC: fearful trials with a correct response; HC: happy trials with a correct response; FI: fearful trials but incorrectly judged as happy; HI: happy trials but incorrectly judged as fearful. Note that HC is equal to zero by definition (see **Methods**). *: $p < 0.05$, **: $p < 0.01$ and ***: $p < 0.001$.

To summarize the population response, we next visualized the mean difference in response between fearful and happy stimuli for both correct and incorrect trials (**Figure 4.8C**). For fearful interactive neurons, this response difference tended to be positive for correct and negative for incorrect trials (**Figure 4.8C**, red) and vice-versa for happy interactive neurons (**Figure 4.8C**, blue). Thus, the response during incorrect trials tended to be similar to the correct trials of the opposite emotional category. This result shows that interactive neurons code the subjective judgment of emotion.

To directly relate neuronal responses to the emotion judgments made on the task, we next performed a single-trial analysis that permits analysis of response variability. In contrast, the analysis discussed so far was based on an average across all fearful or happy trials for each neuron. We used a simple response index $R_i$ as a single-trial metric (see **Eq. 4.3**), which takes into account the opposite signs of the two types of neurons—the fearful type and the happy type—and normalizes for different baseline firing rates. The response index is a function of a neuron's response in a 1.5 second interval starting 250 ms after stimulus-onset (the same interval used above for selecting emotion-selective and interactive cells). $R_i$ is equal to the firing rate during a particular trial i, minus the mean firing rate of all correct happy trials divided by the average of the baseline (**Eq. 4.3**). For example, if a neuron doubles its firing rate for a fearful stimulus and remains at baseline for a happy stimulus, the response index would equal 100%. By definition, $R_i$ is negative for happy units, and thus we multiplied $R_i$ by -1 if the unit was previously classified as a happy unit (**Eq. 4.4**).

We next utilized the response index as defined above to quantify trial-by-trial variability by comparing the distribution of $R_i$ between different conditions. For the interactive neurons (n = 23), the distribution for fearful and happy stimuli was significantly different for both correct and incorrect trials (two-tailed two-sample Kolmogorov–Smirnov (KS) test, $p < 0.0005$ for both correct and incorrect trials, **Figure 4.8D**). Comparing the distributions using a cumulative distribution function (**Figure 4.8D**, see **Methods**) shows that the response during incorrect trials was similarly distributed compared to the correct trials of the opposite category. For example, happy incorrect trials (**Figure 4.8D**, green curve) were similarly distributed to fearful correct trials (**Figure 4.8D**, red curve), and vice-versa. Confirming this observation, there was no significant difference between happy incorrect and fearful correct trials (KS-test, $p = 0.62$) nor between fearful incorrect and happy correct trials ($p = 0.087$, uncorrected). Thus, single-trial neuronal responses confirmed the previous cell-by-cell findings. The mean of the distribution of response indices for both fearful correct and happy incorrect trials (in both cases the perceptual judgment was *fearful*) had response indices significantly above 0 (**Figure 4.8E**; two-

tailed one-sample t-test, $p < 10^{-13}$ for fearful correct trials and $p < 0.0005$ for happy incorrect trials), and there was no significant difference between correct and incorrect trials for fearful subjective judgments (two-tailed two-sample t-test comparing fearful correct and happy incorrect, $p = 0.99$). Interestingly, there was a significant difference between the two types of happy subjective judgments (comparing happy correct and fearful incorrect, $p = 0.027$), with fearful incorrect trials significantly below 0 (t-test against 0: $p < 0.05$). This was because the firing rate for fearful incorrect trials was lower than it was for happy correct trials. Separate analyses for only fearful or happy-selective neurons led to similar results (see **Figure 4.9**), with both classes of neurons showing the same pattern of response independently. In conclusion, we found that the neurons with a significant interaction term encoded the perceptual judgment made by the patient rather than the stimulus identity, at both the single-neuron and single-trial level.

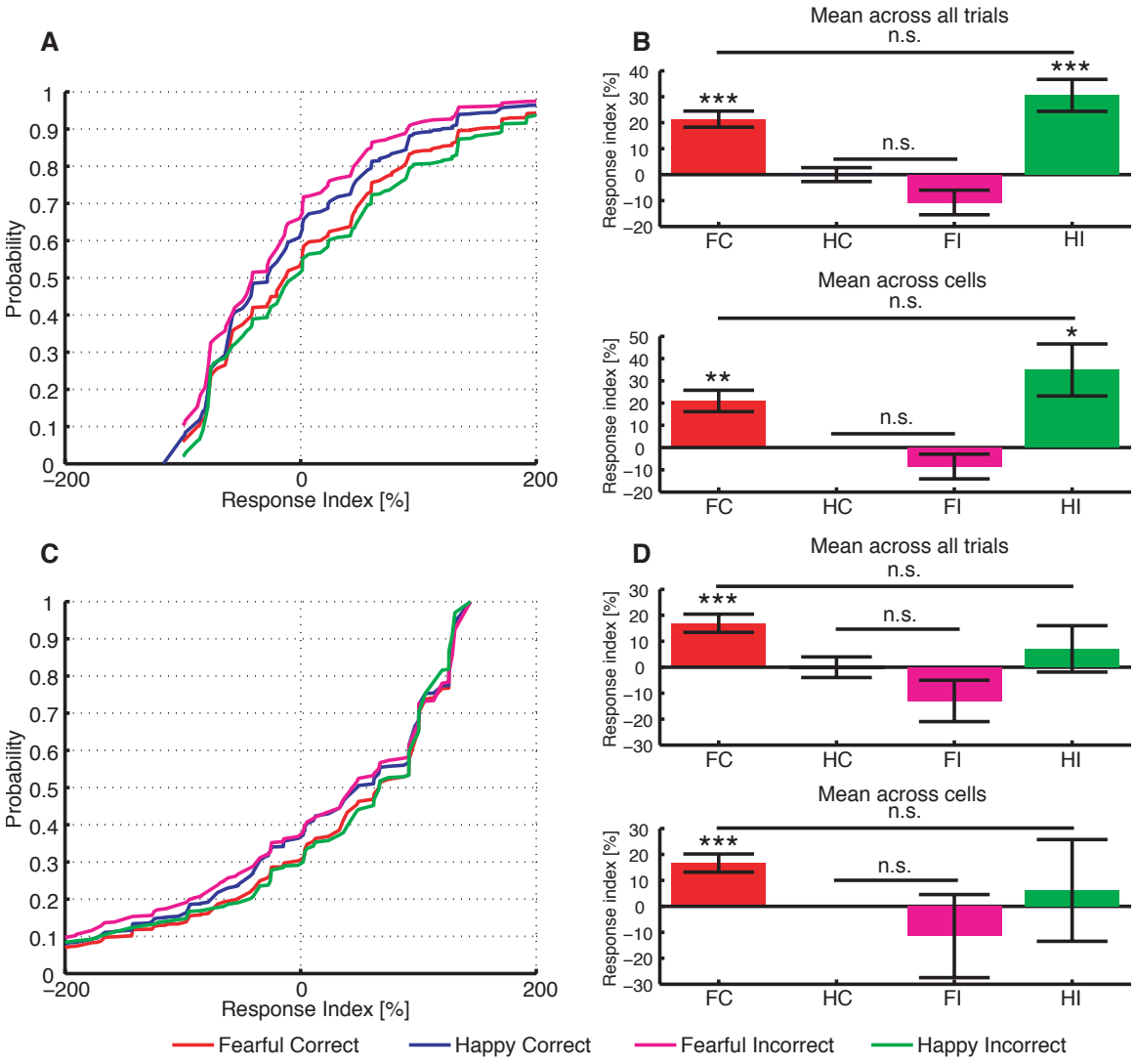


**Figure 4.9. Analysis of neurons with significant ANOVA interaction, broken down by fear-selective and happy-selective.**

Analyses for **(A–B)** fearful interactive neurons (n = 10 units), and **(C–D)** happy interactive neurons (n = 13 units). **(A,C)** Cumulative distribution function (CDF) of the response index, which was calculated as the baseline-normalized difference in response to fearful trials compared to happy trials. The response was significantly different between correct and incorrect for both happy and fearful trials for either type of neuron (KS-test, $p < 0.05$ for all comparisons). In contrast, comparing trials according to their subjective judgment (i.e., fearful correct with happy incorrect, and happy correct with fearful incorrect) resulted in no significant difference (KS-test, $p > 0.05$ for all

comparisons). **(B,D)** Bar plots of the response index across all trials (top) and across all cells (bottom). Pairwise t-tests are indicated by *: $p < 0.05$, **: $p < 0.01$ and ***: $p < 0.001$. FC: fearful trials with a correct response; HC: happy trials with a correct response; FI: fearful trials but incorrectly judged as happy; HI: happy trials but incorrectly judged as fearful.

*4.5.5 Emotion-selective neurons encode perceptual judgment*

Are all emotion-selective neurons sensitive to subjective judgment, or is this a property only of a subset of neurons? Above, we explicitly selected for a significant interaction to begin with, and subsequently analyzed this sub-group. To obtain a broader inventory, we next analyzed the previously described units (n = 41, among which 6 were fearful interactive neurons and 3 were happy interactive neurons; see **Table 4.2**) that were only selected for emotion selectivity on correct trials (incorrect trials were not used for this selection). We computed the response indices for every trial and pooled across all trials as described above in our analysis of interactive neurons. We then computed a population summary metric that summarized the response difference across a group of cells as the mean difference between the response index for fearful and happy trials (see **Methods**, **Eq. 4.5**). To assess statistical significance, we estimated the null distribution by first randomly shuffling the labels of trials (fearful/happy) and then computing the population summary metric. We repeated this permutation test 1000 times. We then compared the observed value of the metric with this null distribution of metrics. The chance values of the null distribution were clustered around 0 as expected (**Figure 4.10**, gray). In contrast, the value of the population effect metric was 25.0% for correct trials (**Figure 4.10**, red; $p < 0.001$ (estimated by counting the number of permutation runs from the null distribution that had a population metric *greater* than the observed value)), which is expected as the cells were selected for this effect in the first place. However, as cells were selected considering only correct trials, incorrect trials remain statistically independent. We found that the population response metric of incorrect trials was significantly negative (-4.63%,

p = 0.002 (estimated by counting the number of permutation runs from null distribution that had a population metric *smaller* than observed value); **Figure 4.10**, blue). Importantly, the metric from the incorrect trials was significantly negative and thus on the opposite side of the null distribution compared to the metric from correct trials (**Figure 4.10**, blue). This shows that when the behavioral response was incorrect (opposite as what was shown on the screen), the neuronal response was consistent with the behavioral response rather than the ground truth (if the blue bar were on the same side as the red bar, by contrast, it would indicate that neuronal responses instead tracked the emotion shown in the stimulus). We thus conclude that the 41 emotion-selective neurons signaled the subjective emotional judgment. We found similar results when we considered fear and happy selective neurons separately (see **Figure 4.11**).

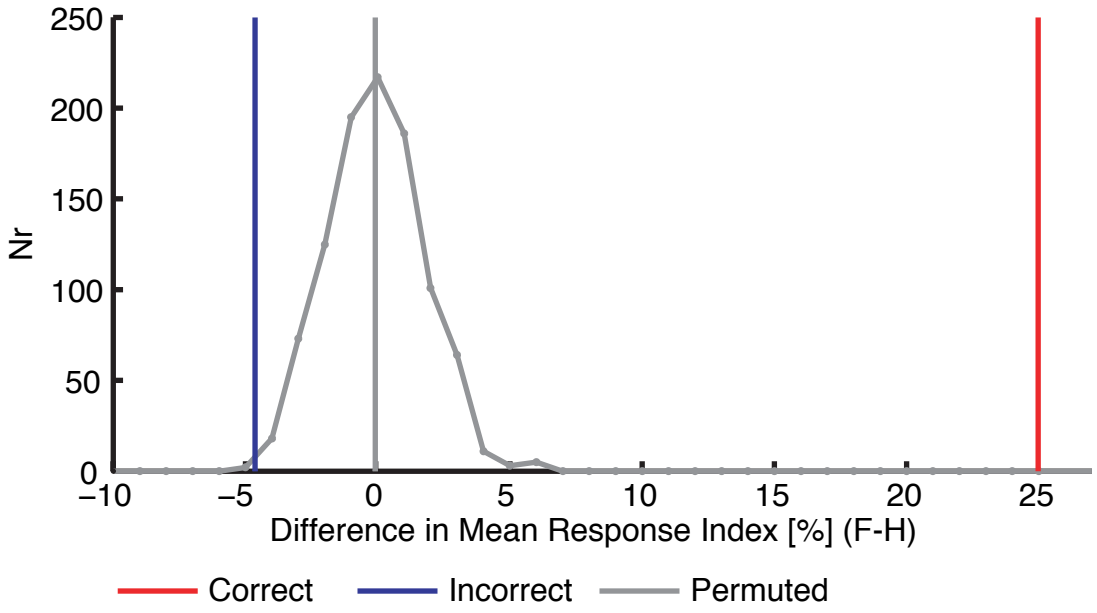


**Figure 4.10. The mean response across all emotion-selective neurons encoded subjective perceptual judgment.**

The gray distribution is the null distribution from permutation tests. The red bar is the metric from observed correct trials and the blue bar is the metric from observed incorrect

trials. Both the red bar and blue bar were located outside the null distribution ($p < 0.005$ for all, estimated by counting the number of permutation runs from the null distribution that had a population metric greater/smaller than the observed value). The blue bar was located on the opposite side of the red bar.

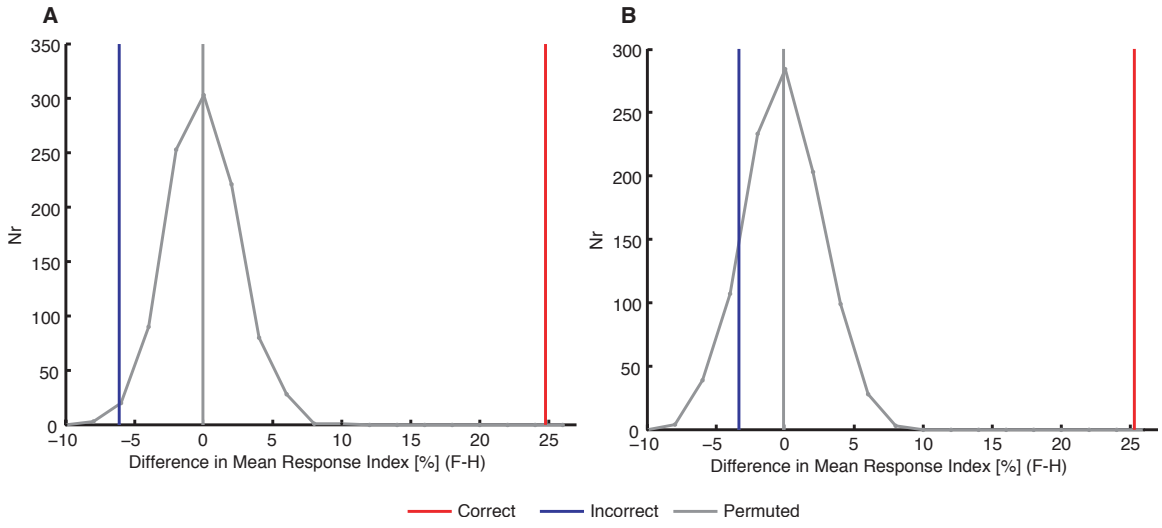


**Figure 4.11. Population summary metrics, compared to permutation distribution, shown separately for fear-selective and happy-selective cells.**

Both fear-selective cells **(A)** and happy-selective cells **(B)** showed coding of subjective perceptual judgment. The gray distribution is the null distribution from permutation tests. The red bar is the metric from observed correct trials and the blue bar is the metric from observed incorrect trials. Both the red bar ($p < 0.001$ in both cases) and blue bar were located outside the null distribution ($p = 0.007$ and $p = 0.11$, respectively). Importantly, the blue bar was located on the opposite side of the red bar.

Were the results influenced by difficulty? The mean number of bubbles shown was 38.2 ± 34.2 (mean ± SD) for correct and 21.9 ± 21.2 for incorrect trials ($p < 10^{-10}$, unpaired t-test). Thus, as expected, incorrect trials tended to occur when less visual information was

revealed. As a control, we repeated our analysis by using only a subset of trials such that, on average, equal amounts of the eye and mouth ROIs were revealed (on average, 28.92 ± 26.90 for correct trials and 28.86 ± 26.91 for incorrect trials, two-tailed paired t-test: $p > 0.05$; and for each individual session, $p > 0.05$ for both fearful correct vs. fearful incorrect and happy correct vs. happy incorrect). We found very similar results (**Figure 4.12A–C**), confirming that emotion-selective neurons signal the perceptual judgment independent of difficulty. We also repeated the analysis by using equal numbers of trials for correct and incorrect to exclude any potential bias and we found very similar results (**Figure 4.12D–F**). We further repeated the analysis by excluding any recordings obtained from epileptic tissue (31 out of a total of 210 units were from tissue subsequently resected as part of the epileptic focus, among which 10 units were fear-selective and 1 unit was happy-selective). The results were qualitatively the same (see **Figure 4.12G–I**). Finally, two of the neurosurgical patients had a clinical diagnosis of autism (Rutishauser et al 2013). We repeated the analysis after excluding these two patients and again found very similar results (**Figure 4.12J–L**).

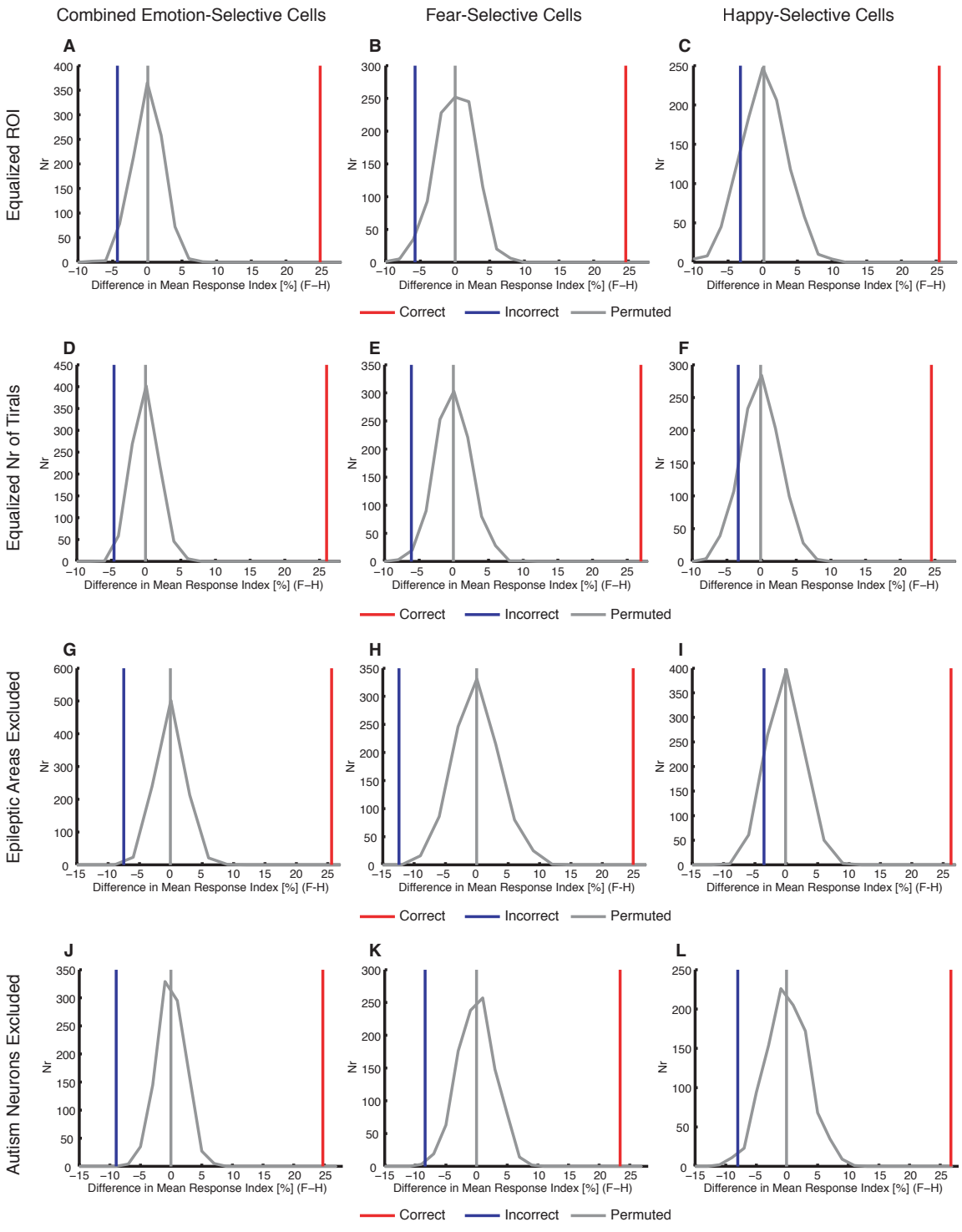


**Figure 4.12. Control analyses, using population summary metrics.**

Emotion-selective cells still showed coding of subjective perceptual judgment when equalizing the proportion of the face shown in eye and mouth ROIs **(A–C)**, when

equalizing the number of correct and incorrect trials **(D–F)**, when excluding epileptic areas **(G–I)**, and when excluding neurons from patients with autism **(J–L)**. **(A,D,G,J)** Combined emotion-selective cells. **(B,E,H,K)** Fear-selective cells. **(C,F,I,L)** Happy-selective cells. The gray distribution is the null distribution from permutation tests. The red bar is the metric from observed correct trials and the blue bar is the metric from observed incorrect trials. Both the red bar and blue bar stood outside the null distribution. Importantly, the blue bar stood on the opposite side of the red bar. **(A–C)**: for correct trials, $p < 0.001$ in all cases; for incorrect trials, $p < 0.05$ for combined emotion-selective cells ($n = 41$) and fear-selective cells ($n = 24$), and $p = 0.158$ for happy-selective cells ($n = 17$). **(D–F)**: for correct trials, $p < 0.001$ in all cases; for incorrect trials, $p < 0.005$ for combined emotion-selective cells ($n = 41$), $p < 0.01$ for fear-selective cells ($n = 24$), and $p = 0.111$ for happy-selective cells ($n = 17$). **(G–I)**: for correct trials, $p < 0.001$ in all cases; for incorrect trials, $p < 0.001$ for combined emotion-selective cells ($n = 30$) and fear-selective cells ($n = 14$), and $p = 0.106$ for happy-selective cells ($n = 16$). **(J–L)**: for correct trials, $p < 0.001$ in all cases; for incorrect trials, $p < 0.001$ for combined emotion-selective cells ($n = 27$), $p < 0.005$ for fear-selective cells ($n = 16$), and $p < 0.05$ for happy-selective cells ($n = 11$).

### *4.5.6 A full inventory of neurons in the amygdala that encode perceptual judgment*

How representative were the subsets of cells described so far of the entire population of amygdala neurons recorded? We next conducted a permutation analysis on the entire population of cells to assess the likely effect size across the population. This analysis used independent subsets of trials for cell selection and response quantification during each repetition of the permutation. We ran 1000 iterations in total. In each, we randomly selected half of the correct trials to select emotion-selective units and to classify them as either fear or happy selective. Subsequently, we calculated the response indices for the remaining half of the correct trials and all incorrect trials. Again, we calculated the population summary metric (as shown in **Figure 4.10**) but only using this independent

subset of trials not previously used for selecting the cells. For the null distribution, we did the same permutation test (1000 runs) with randomly shuffled trial labels. But here, we still use half of the trials to select cells and the other half to predict response indices. The complete independence between selection and prediction insured our results against biases and false positives during selection since only out-of-sample errors were calculated.

Out of our total 210 neurons recorded, we considered 185 cells with >0.2Hz firing rate for this analysis. Many cells were reliably selected over the 1000 repetitions (**Figure 4.13A**, upper panels; 40 and 34 cells were selected in at least 10% of runs for fear and happy conditions, respectively). In contrast, selection was random in the control condition with permuted labels (**Figure 4.13A**, lower panels; no cells were selected in at least 10% of the runs). Not surprisingly, there was considerable overlap between the cells consistently selected by the present split analysis and the cells selected with all trials (n = 41) as analyzed previously (**Figure 4.14**, upper panels). In contrast, for the permutation test which randomly shuffled labels, each cell was equally likely to be selected with a probability of 0.05 (lower panels of **Figure 4.14**); the selected cells were evenly distributed across all 185 cells and across permutation runs (lower panels of **Figure 4.13A**) and did not show a bias towards those that could be selected with all trials (**Figure 4.14**). On average, 16.3 ± 3.1 (mean ± SD) units (8.8% of 185) were categorized as fear-selective and 13.5 ± 2.8 (7.3% of 185) as happy-selective, above the chance estimate of 9.25 cells for each category ($p < 0.01$ for fear-selective and $p = 0.077$ for happy-selective; **Figure 4.13B**). In contrast, the control permutation test resulted in 9.2 ± 3.0 units that were fear-selective and 9.4 ± 2.8 units that were happy-selective (**Figure 4.13B**, middle panels), with no difference between the two categories ($p = 0.14$) and the chance value 9.25 ($p > 0.05$ for both). Furthermore, the symmetric shape of the null distribution (see **Figure 4.15**) showed that the permutation test was not biased.

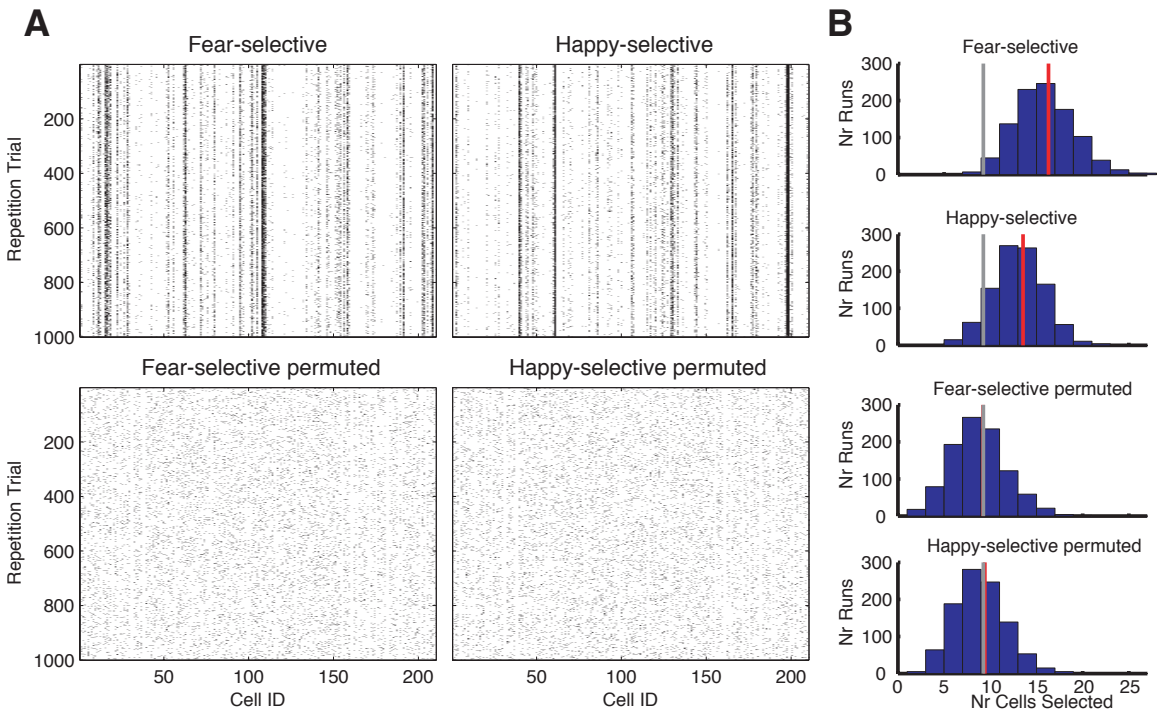


**Figure 4.13. Illustration of the split analysis method to compute the population response.**

**(A)** Cells selected across runs. Each black dot means a particular cell was selected. There was substantial consistency of cells selected in the split analysis (upper panels) but cell selection was evenly distributed across cells and runs in the permutation test (lower panels). **(B)** Summary of the number of cells selected across all runs. Gray and red vertical line indicates the mean of the chance and actual distribution, respectively. The number of cells selected in the split analysis was well above chance while the number of cells selected in the permutation test was near chance. See **Figure 4.16** for results.

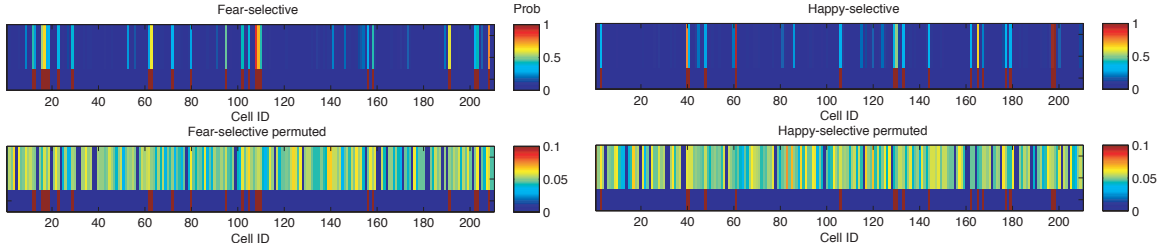

**Figure 4.14. Summary of the likelihood of each cell being selected.**

In the split analysis (upper panels), cells were consistently selected and there was substantial overlap with cells selected by all trials (shown as red bars at the bottom of each color-map with probability equal to 1). But in the permutation test (lower panels), each cell was equally likely to be selected with the predetermined false discovery rate of 0.05. Also, the selection was not biased towards the cells selected by all trials.

We next quantified the responses of the groups of cells selected in each run using the population summary metric as described above (**Figure 4.16**). The population summary metric is calculated as the difference between the average of response indices from all fearful trials (either correct or incorrect) collapsed across all selected cells and the average of response indices from all happy trials (either correct or incorrect) collapsed across all selected cells (see **Eq. 4.5**). The population metric here combined both fear and happy selective cells. The population response was significantly different from the null distribution, for both correct trials and incorrect trials (unpaired two-tailed t-test, $p < 0.0001$). The distribution of the incorrect trials was shifted in the opposite direction relative to the distribution of the correct trials. This also held separately for fear and happy selective neurons (see **Figure 4.15** for population metric distributions separately for fear and happy selective neurons). Thus, the neural signals always followed the behavioral response instead of stimulus ground truth, regardless of whether the behavioral response was correct or incorrect. We thus conclude that emotion-selective neurons in the amygdala encode perceptual judgment robustly.

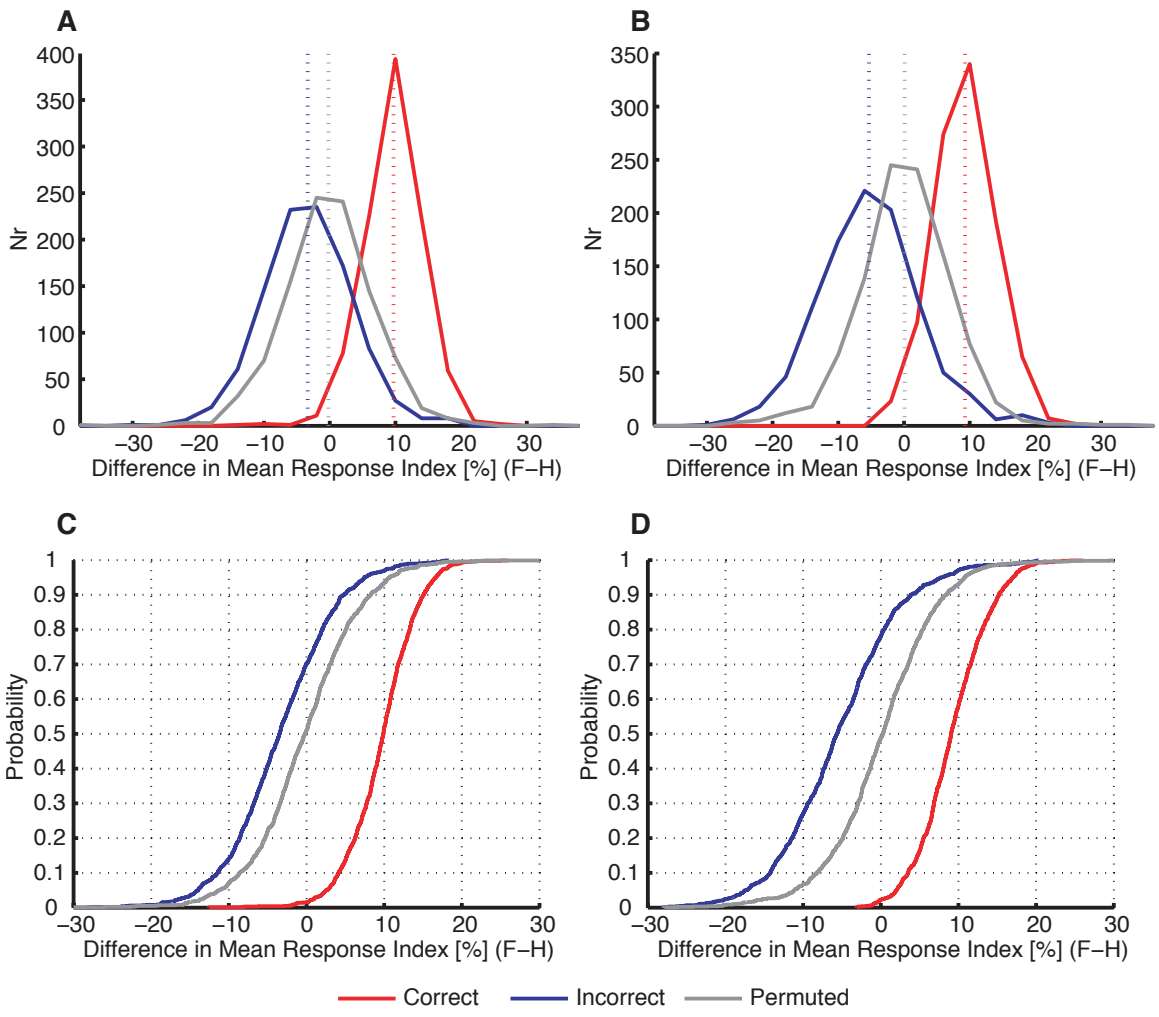


**Figure 4.15. Population summary metric for amygdala neurons separately for fear- and happy- selective neurons.**

**(A,C)** Fear-selective neurons. **(B,D)** Happy-selective neurons. **(A,B)** Histogram representation. **(C,D)** Cumulative distribution function (CDF) representation. The gray distribution is the null distribution from permutation tests. The red distribution is from correct trials and the blue distribution is from incorrect trials. Both the red distribution and blue distribution shifted away from the null distribution. Importantly, the blue distribution was on the opposite side of the red distribution.

*4.5.7 Neuronal response characteristics dependent on facial information revealed*

Were the emotion-selective units predominantly driven by information conveyed by specific parts of the face? We randomly revealed parts of the face, allowing us to select subsets of trials that reveal only specific parts of the face. We selected trials according to how much of pre-defined eye and mouth ROIs was revealed (shown in **Figure 4.2C**). The more overlap between bubbles and ROIs, the more is revealed within the ROIs specified. We picked two types of ROI trials: 'High Eye AND Low Mouth' (**Figure 4.16A** and see **Figure 4.17A–B** for the distribution), and 'Low Eye AND High Mouth' (**Figure 4.16A** and see **Figure 4.17C–D** for the distribution). 'High' or 'Low' here refer to above or below the median across all correct trials of each subject. The conjunction between one high facial feature and one low facial feature ensured that the neuronal response was primarily driven by one facial feature only. We subsequently repeated the split analysis as described above on these ROI trials. Since only correct trials were involved in the selection of ROI trials, the distributions in **Figure 4.16A** did not involve incorrect trials (the incorrect trials may not obey the above division according to eye and mouth ROIs). The population metric here combined both fear and happy selective cells.

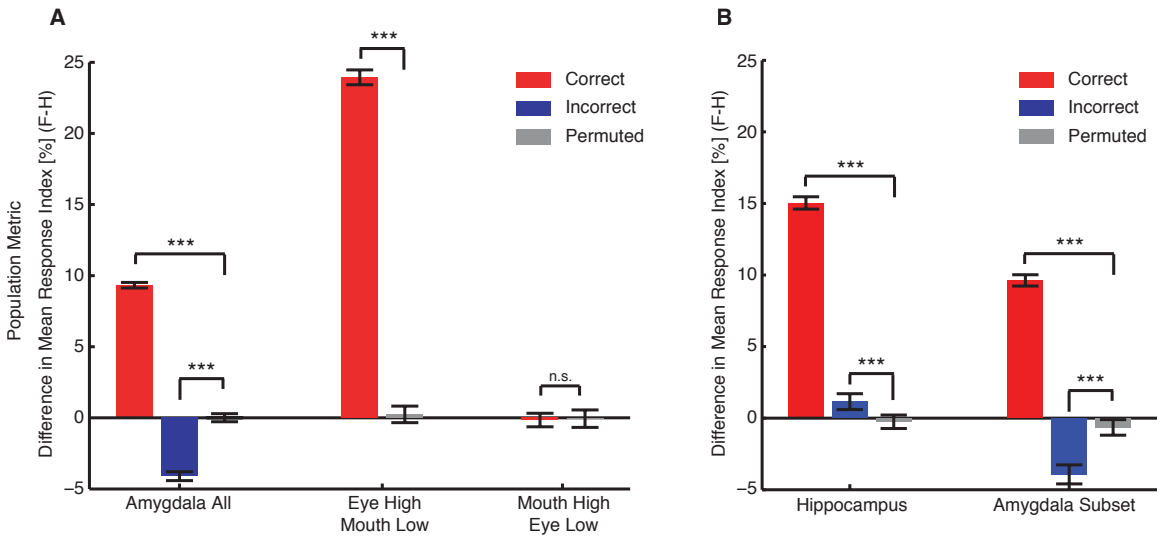


**Figure 4.16. Quantification of the population response using split analysis.**

**(A)** Emotion-selective neurons are primarily driven by information revealed by the eyes. **(B)** Hippocampal neurons also encode emotions but not subjective judgment. In contrast, a subset of amygdala neurons equal to the total number of hippocampal neurons (n = 67) could encode both emotion and subjective judgment as computed from the entire amygdala neuron population. Red: population metric from correct trials. Blue: population metric from incorrect trials. Gray: population metric from trials with shuffled labels. Error bars denote 95% confidence interval. ***: $p < 0.001$. Only correct trials were analyzed for the ROI-restricted analysis.

Information conveyed by the eyes strongly modulated the neuronal response (**Figure 4.16A** and **Figure 4.17A–B**). On average, 16.7 ± 3.2 (mean ± SD) cells were selected as fear-selective neurons and 10.0 ± 2.4 cells were selected as happy-selective neurons, both significantly above chance ($p < 0.001$, **Figure 4.17E**). For this subset, the separation between the distribution of correct trials (red) and null distribution (gray) was prominent (unpaired two-sample t-test: $p < 0.0001$) and the difference was much larger than with all trials ($p < 0.0001$, **Figure 4.16A**). The results held when analyzing fear-selective and happy-selective neurons separately (both $p < 0.0001$, **Figure 4.17A–B**).

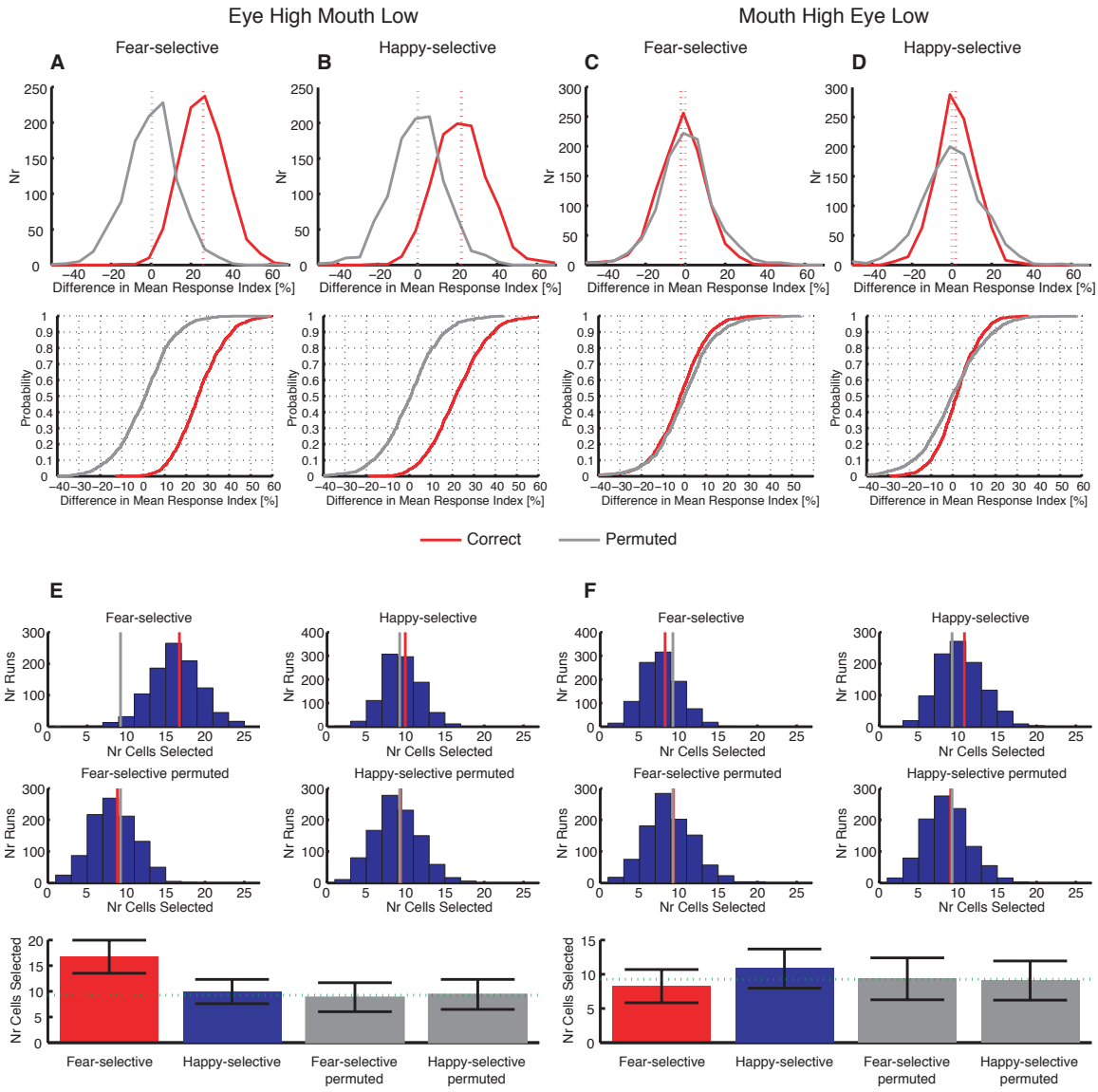


**Figure 4.17. Comparison of neuronal responses during trials where predominantly the eyes or the mouth was shown.**

Information revealed by the eyes but not the mouth area most strongly modulated the neuronal response that differentiated fearful and happy faces. **(A–D)** Population summary metric. The gray distribution is the null distribution from permutation tests and the red distribution is from correct trials. The shift of the red distribution from the gray indicates the modulation of neuronal response. **(E–F)** Summary of the number of cells selected.

In contrast, information provided by the mouth did not modulate neuronal response strongly, as shown by the overlap between the distribution of correct trials and the null distribution ($p = 0.59$). Although we observed a statistically significant difference when analyzing separately for fear-selective neurons and happy-selective neurons, the difference was very small (**Figure 4.17C–D**). On average, $10.8 \pm 2.9$ cells were selected as happy-selective neurons, which was significantly above chance ($p < 0.0001$, **Figure 4.17F**). By contrast, only $8.3 \pm 2.4$ cells were selected as fear-selective neurons, which was significantly *below* the chance value of 9.25 ($p < 0.0001$, **Figure 4.17F**), indicating that when eyes were absent and mouth was present, the neuronal response to fearful faces was suppressed.

In conclusion, we found that information conveyed by eyes but not mouth modulated emotion-selective neuronal responses in the amygdala.

### *4.5.8 Specificity of the amygdala neurons in coding subjective judgment*

How specific were amygdala neurons in encoding subjective judgment? We next analyzed neurons from an adjacent brain region—the hippocampus—to test the specificity of amygdala neurons in coding subjective judgment. We recorded in total 67 single neurons in 6 sessions from 4 patients (2 patients had 2 sessions; see **Table 4.1**). 63 cells had a firing rate greater than 0.2 Hz and were used for the subsequent analyses. Using identical criteria as for the analysis of amygdala neurons, we found 4 fear-selective neurons (6.4%) and 7 happy-selective neurons (11.1%).

We repeated the split analysis for the entire population of hippocampal cells using a random subset of 50% of the trials to select the neurons and the remaining 50% of the trials to quantify the response. We ran 1000 iterations in total. For the null distribution, we conducted the same permutation test with randomly shuffled trial labels. The result shows that both happy-selective neurons and fearful-selective neurons were consistently selected across repetitions (**Figure 4.18A**). Interestingly, only happy-selective neurons,

but not fear-selective neurons were selected above chance (**Figure 4.18B**). The selected neurons differentiated fearful from happy faces in correct trials ($p < 0.001$; **Figure 4.16B;** also see **Figure 4.19**). Thus, a subset of hippocampal neurons distinguished happy from fearful emotions in correct trials. Crucially, however, this was only the case for correct trials. In contrast to the amygdala neurons, the hippocampal neurons did not indicate the behavioral response made during incorrect trials. Rather, the response indicated, albeit only weakly so (**Figure 4.16B**), what the correct response would have been (ground truth, as shown on the screen). The crucial difference, however, is that the distribution of the response during incorrect trials was shifted in the *same* direction ($p < 0.001$) relative to the distribution of the correct trials. This is in contrast to the amygdala neurons, for which the distribution of the response during the incorrect trials was shifted in the *opposite* direction relative to the distribution of the correct trials (**Figure 4.16A**). In conclusion, hippocampal neurons, unlike amygdala neurons, did not track the subjective judgment of facial emotion in incorrect trials.

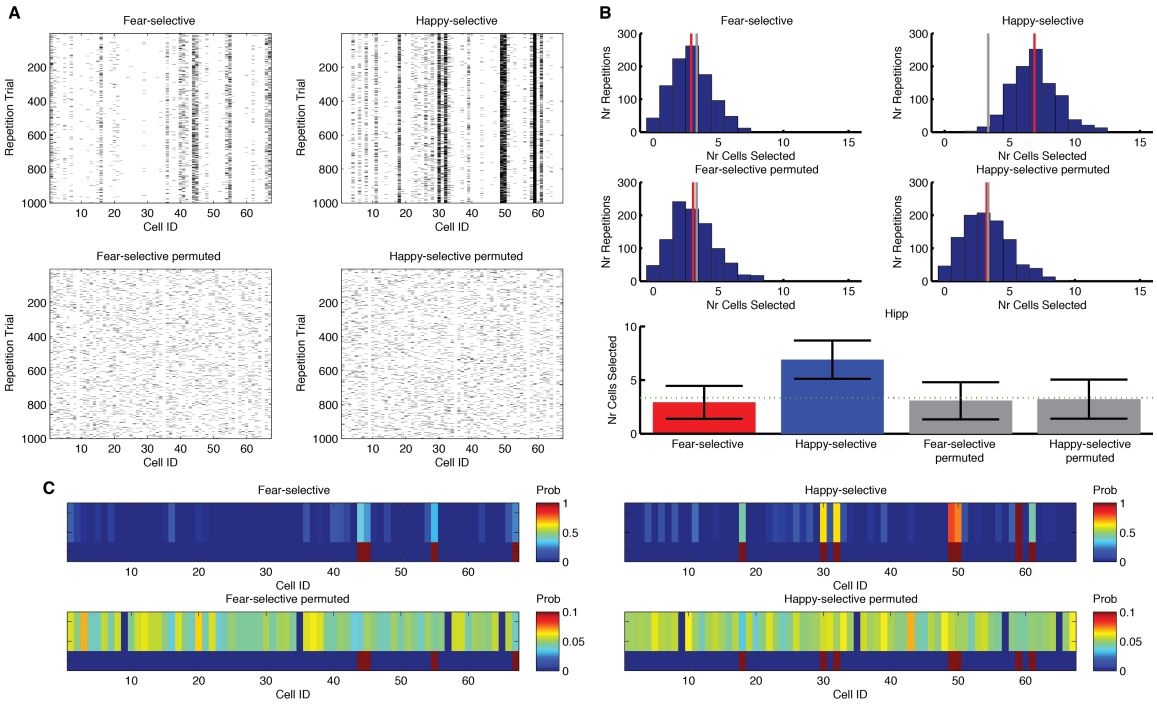

**Figure 4.18. Quantification of the split analysis and permutation test for hippocampal neurons.**

**(A)** Cells selected across runs. Each black dot means a particular cell was selected. There was substantial consistency of cells selected in the split analysis (upper panels) but cell selection was evenly distributed across cells and runs in the permutation test (lower panels). **(B)** Summary of the number of cells selected. The number of cells selected in the split analysis was well above chance while the number of cells selected in the permutation test was near chance. **(C)** Summary of the likelihood of each cell being selected. In the split analysis (upper panels), cells were consistently selected and there was substantial overlap with cells selected by all trials (shown as red bars at the bottom of each color-map with probability equal to 1). But in the permutation test (lower panels), each cell was equally likely to be selected with the predetermined false discovery rate of 0.05. Also, the selection was not biased towards the cells selected by all trials.

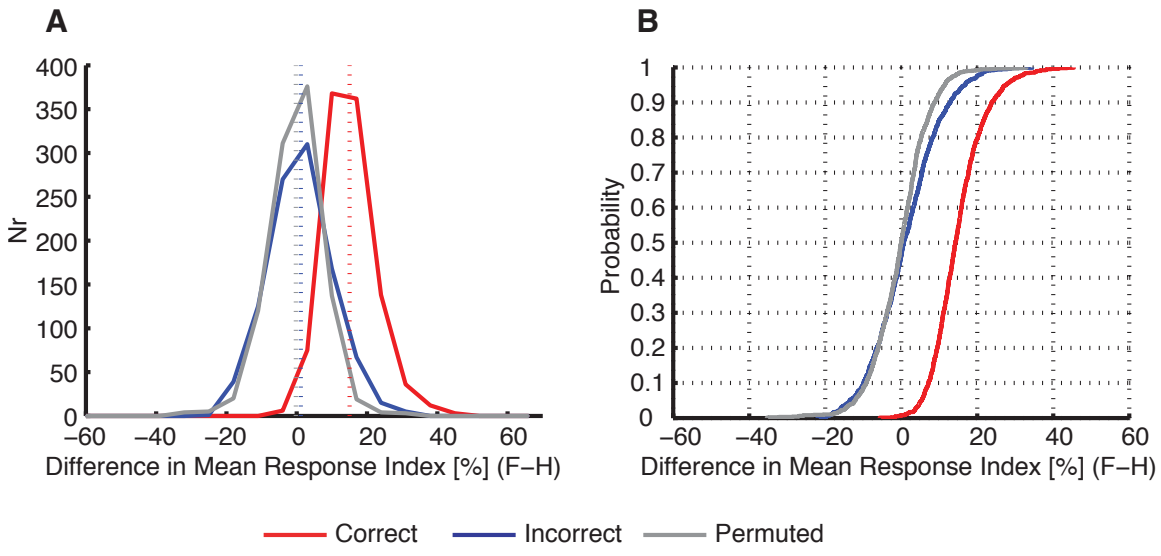


**Figure 4.19. Emotion coding in hippocampal neurons (combined emotion-selective neurons), quantified using population summary metric.**

**(A)** Histogram representation. **(B)** Cumulative distribution function (CDF) representation.

The gray distribution is the null distribution from permutation tests. The red distribution is from correct trials and the blue distribution is from incorrect trials. Only the correct trials (red) were different from the null distribution.

There were fewer hippocampal neurons than amygdala neurons (67 vs. 210), which could have biased the effect size. We thus next repeated the analysis of the amygdala neurons by randomly selecting a subset of 67 amygdala neurons in each run of the split analysis. We found very similar results compared to the entire population of amygdala neurons (**Figure 4.16B**; but note the larger variance due to fewer number of neurons), and again found a different pattern of results than what was seen in the hippocampus (with an identical number of selected neurons).

In conclusion, we found that only amygdala neurons, but not hippocampal neurons, indicated the subjective judgment of emotions.

### *4.5.9 Reaction time (RT) and laterality analysis*

In an attempt to distinguish perceptual judgments from motor outputs, we lastly analyzed whether the response of emotion-selective units was correlated with behavioral output. We analyzed the correlation between RT and spike timing or spike counts for the 41 emotion-selective neurons (Pearson correlation, false positive rate = 0.05, uncorrected for multiple comparisons). When correlating RT with peak firing time (center of the 250 ms bin which had the highest firing rate), we observed only two fear-selective neurons with a significant positive correlation (binomial test on the number of significant cells, $p = 0.12$; two neurons with a significant negative correlation, $p = 0.12$), while we observed no happy-selective neurons having a significant positive correlation ($p = 0.58$; one neuron had a significant negative correlation, $p = 0.21$). When correlating RT with the total number of spikes in a time window 500 ms prior to button press, we found only one

significant positive correlation for fear-selective neurons ($p = 0.34$) but no significant positive correlation for happy-selective neurons ($p = 0.58$; five fear-selective neurons ($p < 0.001$) and two happy-selective neurons ($p = 0.050$) had a significant negative correlation). Separate analyses for fearful trials and happy trials showed very similar results. We conclude that there was no significant correlation between firing rate and RT.

One confounding factor of our experimental setup is that the same motor action is always associated with the same emotion (the left button denoted fearful and the right button happy in our setup). To our knowledge there is no evidence that the amygdala encodes such specific motor actions. We further analyzed the distribution of emotion-selective neurons to investigate whether the button presses were associated with emotion coding in the amygdala. If the emotion neurons are associated with the button press, they should appear contralaterally to the pressed buttons. However, we did not observe such laterality. Of the total 210 cells, 92 cells were recorded from the left amygdala and 118 cells were recorded from the right amygdala (see **Figure 4.20**). Among the emotion-selective cells, 6 fear-selective cells were from the left amygdala, 18 fear-selective cells were from the right amygdala, 6 happy-selective cells were from the left amygdala and 11 happy-selective cells were from the right amygdala. The proportion of emotion-selective cell did not differ between left vs. right amygdala in any of these categories (Fisher's exact test, $p = 0.51$; paired two-tailed t-test across patients on the percentages, all $ps > 0.1$), showing that the emotion neurons are not lateralized nor related to the output button response associated with the emotion. Further, the same results held when excluding neurons from the epileptic areas (all $ps > 0.25$).

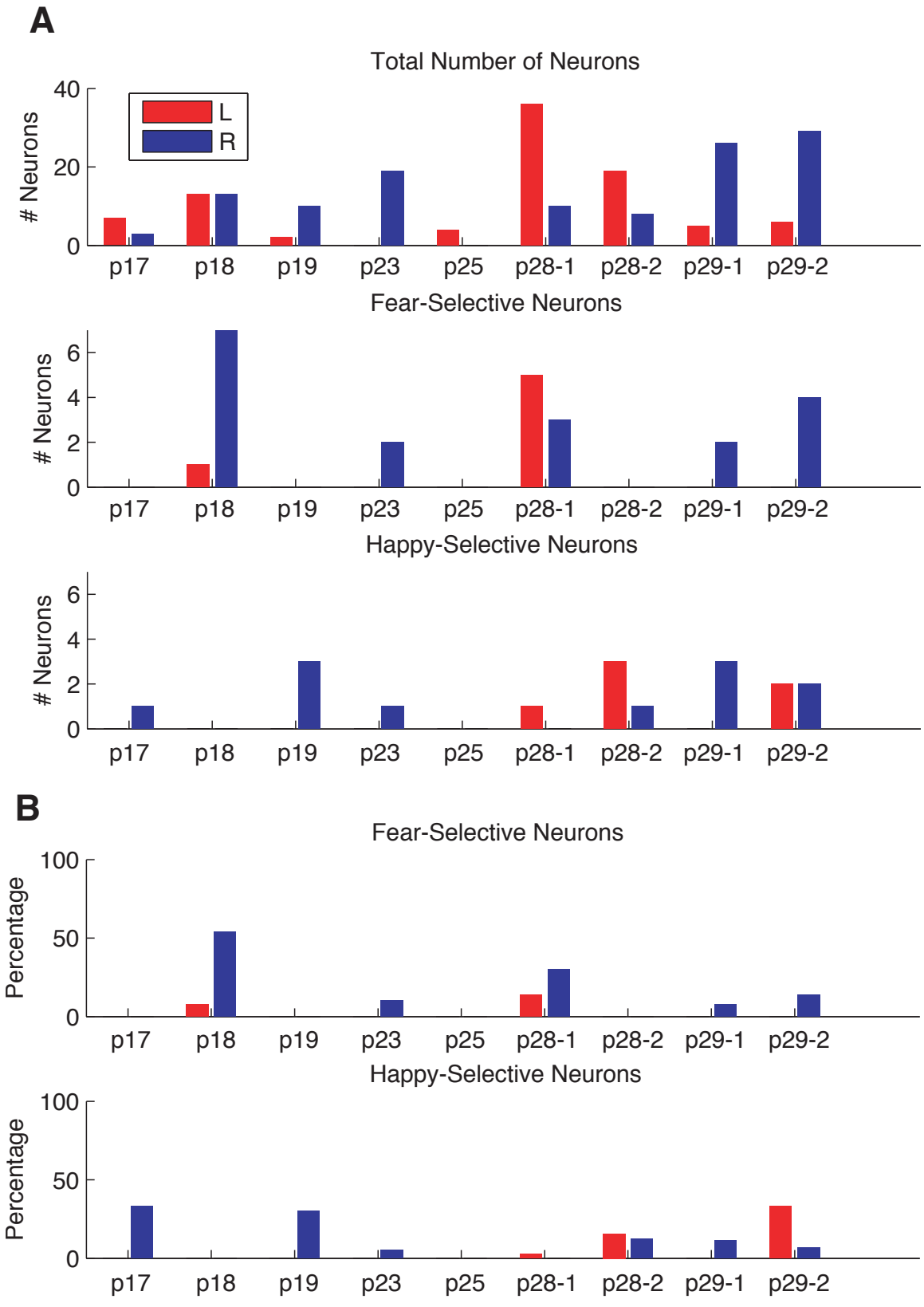
A
Total Number of Neurons
# Neurons
L
R
p17
p18
p19
p23
p25
p28-1
p28-2
p29-1
p29-2
Fear-Selective Neurons
Happy-Selective Neurons
B
Percentage

**Figure 4.20. Distribution of emotion-selective neurons.**

**(A)** Number of emotion-selective neurons from each patient. **(B)** Percentage of emotion-selective neurons. Red: neurons recorded from the left amygdala. Blue: neurons recorded from the right amygdala. Patients 17 and 28 were diagnosed with ASD.

Our results suggest that the amygdala encodes the subjective judgment of emotional faces, but that it plays less of a role in helping to program behavioral responses.

## 4.6 Discussion

In this study, we found that a subset of amygdala neurons encode the subjective judgment of the emotion shown in faces. Behaviorally, our epilepsy patients did not differ from healthy controls in terms of learning performance on the task, and both epilepsy patients and control subjects primarily used the eye region of the stimuli to correctly judge fearful faces and primarily used the mouth region to correctly judge happy faces, findings consistent with prior studies (Smith et al., 2005, Scheller et al., 2012). 41 cells significantly differentiated the two emotions and subsequent analyses indicated that these cells encoded the patients' subjective judgment regardless of whether it was correct or incorrect. Population permutation analysis with full independence between selection and prediction confirmed the robustness of this result when tested across the entire population. ROI analysis revealed that eyes but not mouth strongly modulated population neuronal responses to emotions. Lastly, when we carried out identical recordings, in the same patients, from neurons within the hippocampus, we found responses driven only by the objective emotion shown in the face stimulus, and no evidence for responses driven by subjective judgment.

It is notable that the population response metric for the correct trials was further away from the null distribution relative to the incorrect trials (25.0% vs. -4.63%). It is not

surprising that the strength of emotion coding in incorrect trials was weaker given fewer incorrect trials and thus potentially increased variability and decreased reliability. In addition, incorrect trials were likely a mixture of different types of error trials, such as true misidentifications of emotion, guesses, or accidental motor errors. Regardless, on average, the neural response during incorrect trials reliably indicated the subjectively perceived emotion. This suggests that a proportion of error trials were likely true misidentifications of the emotion rather than pure guesses.

Interestingly, there was a significant difference between the two types of happy subjective judgments (comparing happy correct and fearful incorrect, **Figure 4.8E**). This might reflect a different strategy used by subjects to compare the two emotions in our specific task. Future studies with a range of different tasks will be needed to understand how relative coding of emotion identity and task demands may interact in shaping neuronal responses.

### *4.6.1 Possible confounds*

Our stimuli were based on the well validated set of facial emotion images from Ekman and Friesen (Ekman and Friesen, 1975), from which we chose a subset depicting fear and happiness with the highest reliability. We normalized these base faces for luminance, orientation, color and spatial frequency, eliminating these low-level visual properties as possible confounds. Likewise, we showed a balanced number of male and female faces, and multiple identities, ensuring that neither gender nor individual identity of the face was driving the responses we report (each of these was completely uncorrelated with the emotion shown in the face). Nonetheless, it remains possible that our findings reflect higher-level properties that are correlated with the emotions fear and happiness—such as negative vs. positive valence. Furthermore, since we only tested two facial emotions, our conclusions can only speak to the emotions that we tested and are relative to the task that we used. Different facial regions would have likely been informative for other facial

emotions (had the task been a discrimination task that required a choice between, say, surprise and happiness) and we do not know whether the cells studied here might contribute to perceptual decisions for other emotions. A larger set of emotions, as well as of facial expressions without emotional meaning, would be important to study in future studies.

Our results suggest that emotion-selective neurons were not merely encoding the motor output associated with the perceived emotions (button press), as corroborated by the lack of correlation between the neuronal and behavioral response (consistent with similar prior findings (Rutishauser et al., 2011)), and the lack of lateralization of emotion neurons given the lateralized and fixed motor output actions. Although there has been a recent report of an interaction between spatial laterality and reward coding in the primate amygdala probed with lateralized reward cues (Peck et al., 2013), that effect appeared primarily as a difference in latency but not as the lateralization of reward coding neurons to the reward-predicting cues. It will be interesting to investigate in future studies whether these findings with basic rewards (Peck et al., 2013) can be generalized to emotions or other salient stimuli.

We initially selected emotion-selective neurons using a one-tailed t-test of fear vs. happy for correct trials only. Clearly, some cells surviving this test will be false positives and to quantify the robustness of the effect we thus conducted several additional analyses. First, we conducted a 50/50 split analysis procedure, which keeps the trials used for selection and prediction independent (**Figure 4.13**). The result (**Figure 4.16**) is an out-of-sample estimate of the true effect size and would thus not be expected to be different from chance if all selected cells were false positives. In contrast, we observed a highly reliable effect (**Figure 4.16**), which is very unlikely to be driven by chance alone. Second, the set of cells selected by the two different methods were comparable (see **Figure 4.14**), showing that emotion-selective neurons were consistently selected even with a random subset of trials. Third, we rigorously established chance levels using permutation tests (**Figure 4.16**) and found that the number of cells selected was well above chance (**Figure**

**4.13**). Fourth, we conducted additional control analyses using a time window -250 ms to 750 ms relative to scramble onset (no information about the upcoming face was available during this time window). The number of selected cells was as expected by chance and we did not find the significant patterns we report in the case of responses to faces. Similarly, we also did not replicate the pattern of amygdala responses to faces when we analyzed responses from hippocampal neurons. Taken together, the last two findings provide both stimulus specificity and neuroanatomical specificity to our conclusions. Lastly, we conducted analyses using a random subset of the amygdala neurons (n = 67, the number of hippocampal neurons recorded) at each permutation run and we derived qualitatively the same results (**Figure 4.16B**), showing that our results were not driven by a particular subset of neurons.

### *4.6.2 Comparison with neuroimaging studies*

While the relationship between BOLD responses and single-neuron activity in the amygdala is complex and largely unknown, our findings are nevertheless consistent with an fMRI study which found that the overall BOLD response in the amygdala to fearful faces followed subjective judgment (Pessoa et al., 2006). That study divided trials into hits, misses, correct rejects and false alarms, and showed that not only were amygdala BOLD responses during hit trials greater than those during physically identical miss trials, but also that responses during false alarm trials were greater than those during misses, even though in the former no actual fearful face was present while in the latter it was. While our single-neuron level findings are compatible with this interpretation, our study in addition showed that there are at least two sub-populations of emotion-sensitive neurons (responding to judged happiness and judged fearfulness, respectively) and that these neurons both code subjective emotion. Our findings are also consistent with previous work showing that presenting eyes but not the mouth result in a significant BOLD response in the amygdala (Morris et al., 2002, Whalen et al., 2004).

Interestingly, in contrast to increased neuroimaging responses of fearful faces compared to happy faces (Morris et al., 1996), we found neuronal selectivity for fearful faces in the amygdala comes mainly from a suppression of activity in happy trials (**Figure 4.6A**), whereas selectivity for happy faces is mainly due to an increase in activity for happy trials (**Figure 4.6B**). Such complex relationships between single-neuron dynamics and stimulus dimensions are not visible with an aggregate signal such as the BOLD response. Pooling across all recorded neurons (a condition more directly comparable to BOLD-fMRI, **Figure 4.7**) does not differentiate the two emotions either, a finding consistent with some more recent neuroimaging studies (Fitzgerald et al., 2006). This might be due to the population of neurons that we sampled, our stimuli with only sparsely sampled parts, or the non-linear response characteristics of amygdala neurons responding to facial parts (Rutishauser et al., 2011).

### *4.6.3 Selectivity of amygdala neurons*

Faces can be readily characterized by independent attributes, such as identity, expression, and gender, which have segregated cortical representation (Fried et al., 1982, Perrett et al., 1984, Rolls, 1984, Baylis et al., 1985, Hasselmo et al., 1989, Young and Bruce, 1991), and single-unit recordings in the primate amygdala have documented responses selective for faces, their identity or emotional expression (Fried et al., 1997, Gothard et al., 2007). We previously showed that neurons in the human amygdala selectively respond to whole faces as compared to facial parts, suggesting a predominant role of the amygdala in representing global information about faces (Rutishauser et al., 2011). How do these whole-face-selective cells overlap with the emotion-selective cells we report in the present work? We found 3 out of 24 (12.5%) fear-selective cells and 5 out of 17 (29.4%) happy-selective cells are also whole-face-selective, a ratio of whole-face cells similar to that found in the entire population (36 out of 185, 19.5%). This suggests that amygdala neurons encode whole-face information and emotion independently.

We found that face information conveyed by the eyes, but not the mouth region, modulated emotion-selective neuronal responses. Compared to our previous neuronal classification images which were based on pixel-wise analyses of face regions that drive neuronal response (Rutishauser et al., 2013), we here used a fully independent permutation test to further illustrate that when eyes are more visible, the population of neurons can discriminate the emotions better (also see **Table 4.2**). Together with a substantial prior literature, this finding supports the idea that amygdala neurons synthesize their responses based substantially on information from the eye region of faces (Morris et al., 2002, Whalen et al., 2004, Adolphs et al., 2005, Gamer and Büchel, 2009, Scheller et al., 2012).

### *4.6.4 Functional role of the amygdala*

The amygdala is in a pivotal position to modulate perceptual processing of faces. It sends output to many areas along the visual cortical pathways (Amaral and Price, 1984, Amaral et al., 2003, Catani et al., 2003, Pessoa and Adolphs, 2010). Monkey electrophysiological data show that stimulation of face-selective regions in temporal cortex (the anterior medial (AM) face patch) can induce activation in the lateral nucleus of the amygdala (Moeller et al., 2008). The amygdala could enhance sensory processing of emotional stimuli by prioritizing emotional stimuli over neutral stimuli in perception (Fox, 2000, Anderson and Phelps, 2001, Vuilleumier and Schwartz, 2001). Taken together, these findings suggest that feedback signals from the amygdala that encode emotional information can modulate face representations in temporal cortex, a prediction supported by findings in humans (Vuilleumier et al., 2004) and monkeys (Hadj-Bouziane et al., 2012).

Conversely, our findings also argue for responses emerging at the level of the amygdala that are unlikely to be present at earlier stages of processing. The long latency of the amygdala responses we observed already argues for considerable synthesis, consistent

with the integration of face input from temporal cortex with signals from other brain regions, as well as substantial processing internal to the amygdala. In this regard, it is worth noting a recent neuroimaging study, which found that whereas temporal cortical signals track the physical dimension of morphed facial expressions of emotion, signal in the amygdala showed a distinctly categorical response that sharply separated the stimuli according to their judged emotion (Harris et al., 2012).

The amygdala's proposed modulation of face representations by the subjectively judged emotion is in line with a large literature documenting the pervasive modulation by this structure of a host of cognitive processes. Amygdala activity is correlated with long-term, free recall of emotional information at the encoding phase (Cahill et al., 1996) and it influences memory storage by regulating parahippocampal and frontal regions (Kilpatrick and Cahill, 2003). The primate amygdala represents both positive and negative values of visual stimuli during learning (Paton et al., 2006), and is modulated by expectations to pleasant or aversive stimuli (Belova et al., 2007). Neuroimaging studies in patients with bilateral destruction of the amygdala showed that the human amygdala contributes to reward expectancy and choice signals in prefrontal cortex, which in turn may influence behavioral decision making (Hampton et al., 2007). Both the synthesis of long-latency amygdala responses sensitive to subjectively perceived emotion, and the subsequent effects these responses have on cognitive processing, will require future analyses of signals obtained concurrently from the amygdala and other brain regions.

### *4.6.5 The amygdala, consciousness and perception*

Does the amygdala's response to emotional faces require, or contribute to, conscious awareness? Some studies have suggested that emotional faces can modulate amygdala activity without explicit awareness of the stimuli (Morris et al., 1998, Whalen et al., 1998), and there are reports of amygdala BOLD discrimination to the presentation of fearful faces even if such faces are presented to patients in their blind hemifield in cases

of hemianopia due to cortical lesions (Morris et al., 2001). Our finding that amygdala neurons track subjective perceptual judgment argues for a key role in conscious perception, although it does not rule out a role in non-conscious processing as well. Further support for a role in contributing to our conscious awareness of the stimuli comes from the long response latencies we observed, consistent with previous findings on long latencies in the medial temporal lobe (Mormann et al., 2008). Our findings suggest that the amygdala might interact with visual cortices in the temporal lobe to construct our conscious percept of the emotion shown in a face, an interaction that likely requires additional components such as frontal cortex, whose identity remains to be fully investigated (Pessoa and Adolphs, 2010). In particular, since we failed to find any coding of subjectively perceived emotion in the hippocampus, it will be an important future direction to record from additional brain regions to fully understand how the amygdala responses we report might be synthesized.

Microstimulation of inferotemporal cortex in monkeys (Afraz et al., 2006) and electrical brain stimulation in fusiform areas in humans (Parvizi et al., 2012) have suggested a causal role of the temporal cortex in face categorization and perception. Future studies utilizing direct stimulation of the amygdala will be important to further determine the nature of its contribution to the subjective perception of facial emotion. Given the long average response latency observed in the amygdala neurons we analyzed, it may well be that the responses we report here reflect perceptual decisions that were already computed at an earlier time epoch. We would favor a distributed view, in which the subjective perceptual decision of the facial emotion emerges over some window of time, and drawing on a spatially distributed set of regions. The neuronal responses we report in the amygdala may be integral part of such computations, or they may instead reflect the readout of processes that have already occurred elsewhere in the brain. Only concurrent recordings from multiple brain regions will be able to fully resolve this issue in future studies.

## 4.7 Conclusion

In conclusion, we suggest that the amygdala serves to integrate sensory information about faces, conveyed via temporal neocortex, with reward value (Paton et al., 2006), task, and social context (Ochsner et al., 2002, Schaefer et al., 2002, Kim et al., 2004), through its dense web of connectivity with structures such as basal ganglia and prefrontal cortex. Such processing would underlie the synthesis of subjective judgments about the emotion shown in faces, as our present findings demonstrate, and would also account for the remarkably long neuronal response latencies that we (Rutishauser et al., 2011) and others (Mormann et al., 2008) have described previously. Responses tracking subjective judgments of emotion, in turn, could form the basis for other social judgments that have been linked to the amygdala, such as trustworthiness (Adolphs et al., 1998, Winston et al., 2002) and approachability (Kennedy et al., 2009), as well as to our conscious percept of the face (Pessoa et al., 2006) and the conscious experience of the emotion induced by seeing the face (Feinstein et al., 2011). It will be critical to compare our findings to responses obtained from face-selective neurons in temporal cortex (Tsao et al., 2006), which provide the primary visual input to the amygdala (Amaral et al., 1992), and which in turn receive feedback from the amygdala (Freese and Amaral, 2006). It may be that subjective percepts of facial emotion are represented through iterative cycles of processing between the amygdala, temporal cortex, and other brain structures involved in valuation and social inference.

## 4.8 Acknowledgements

We thank all patients for their participation; Drs. Linda Philpott and William Sutherling for neurological referral and evaluation of the patients; the staff of the Huntington Memorial Hospital Epilepsy and Brain Mapping Program for excellent support with participant testing; Erin Schuman for providing some of the recording equipment; and Frederic Gosselin for advice on the bubbles method. This research was supported by

grants from the Pfeiffer Family Foundation, the Simons Foundation, the Department of Neurosurgery at the Cedars-Sinai Medical Center and a Conte Center from NIMH. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

# Chapter V: Autism spectrum disorder, but not amygdala lesions, impairs social attention in visual search

## 5.1 Overview

In the Chapter III, we have shown that neural-typically developed people pay special attention to faces and conspecifics than inanimate stimuli. This process is independent of the amygdala, a neural structure that has long been associated with social cognition. But how about people with autism spectrum disorders (ASD), who have been diagnosed with impairments in social interactions? In this chapter, we tested two independent groups of people with ASD and three rare patients with selective bilateral amygdala lesions, in order to address an important open question: is there an amygdala-dependent attentional deficit in autism? There is substantial evidence both for a reduced orientation towards social stimuli in autism, and also for the amygdala's involvement in attention to social stimuli. Here, for the first time, we investigated both of these issues, by testing groups of high-functioning adults with ASD, as well as three rare patients with amygdala lesions, on the same stimuli and task to enable direct comparisons. We found reliable deficits of attention, especially social attention, in people with ASD, but not amygdala lesion patients. Our findings further contributed to the understanding of social attention and its dependent neural structures.

In the previous chapters, we have compared social cues vs. non-social cues and its dependence on the amygdala. We find that the amygdala does not play a role in attention but encodes facial emotions. As the last chapter, here we investigate again the role of the amygdala in social attention in a different task (visual search), and further include a clinical population—people with ASD who are known to have social deficits—to directly compare with patient with amygdala lesions. Once again, we find intact social attention in patients with amygdala lesions, corroborating the findings in Chapter III that the

amygdala is not involved in social attention, but instead involved in encoding perceptual judgments as shown in Chapter IV.

### 5.2 Summary

People with autism spectrum disorders (ASD) have pervasive impairments in social interactions, a diagnostic component that may have its roots in atypical social motivation and attention. One of the brain structures implicated in the social abnormalities seen in ASD is the amygdala. To further characterize the impairment of people with autism in social attention, and to explore the possible role of the amygdala, we employed a visual search task with both social (faces and people with different postures, emotions, ages, and genders) and non-social targets (e.g., electronics, food, utensils), which participants were asked to find in an array of 24 objects. We defined target-relevant effects as the difference in the percentage of fixations that fell on target-congruent vs. target-incongruent objects in the array. In Experiment 1, we tested 8 high-functioning adults with ASD, 3 adults with focal bilateral amygdala lesions, and 19 controls. Controls rapidly oriented to target-congruent objects and showed a strong and sustained preference for fixating them. Strikingly, people with autism oriented significantly more slowly to target-congruent objects, an effect driven primarily by reduced orientation towards social objects and not evident from global attentional effects independent of the target-directed search. By contrast, patients with amygdala lesions performed indistinguishably from controls. In Experiment 2, we recruited a different sample of 13 people with autism and 8 autism-matched controls, and tested them on the same search arrays but with all stimuli equated for low-level saliency to rule out possible effects of brightness, size or eccentricity. The results replicated those of Experiment 1. In Experiment 3, we recruited 13 people with autism, 8 autism-matched controls, 3 amygdala lesion patients and another group of 11 controls and tested them on a simpler array with only 12 array objects. Here our group effect for ASD strongly diminished and all four subject groups showed similar target-relevant effects. These findings argue for an attentional deficit in ASD that is

disproportionate for social stimuli, cannot be explained by low-level visual properties of the stimuli, and is more severe with high-load top-down task demands. Furthermore, this deficit appears to be independent of the amygdala.

### 5.3 Introduction

People with autism spectrum disorders (ASD) are characterized by pervasive impairments in social interaction and communication, together with restricted interests and repetitive behaviors (DSM-5, 2013). Laboratory-based measures reflecting the social impairments have documented abnormal eye-tracking to social videos (Klin et al., 2002) as well as static faces (Pelphrey et al., 2002). Work from our laboratory has argued for an increased tendency in adults with ASD to saccade away from the eye region of faces when information was present in those regions (Spezio et al., 2007b), and instead an increased preference to fixate the location of the mouth (Neumann et al., 2006), together with reliance of information from the mouth (Spezio et al., 2007a). Similarly, other eye-tracking studies have found active avoidance of fixating the eyes in faces in people with ASD (Kliemann et al., 2010).

These abnormalities in how eyes are fixated by people with ASD may be related to the more subtle and heterogeneous findings in the literature regarding face processing. In particular, several studies have found reliable, but weak, deficits in the ability to recognize emotions from facial expressions (Law Smith et al., 2010, Philip et al., 2010, Wallace et al., 2011, Kennedy and Adolphs, 2012) (for review, see (Harms et al., 2010)). The recognition of more complex mental states from faces may show a more reliable impairment in ASD, particularly if only the eye region of faces is shown (Baron-Cohen et al., 2001). Interestingly, abnormal fixations onto faces (Adolphs et al., 2005), abnormal recognition of emotion from facial expressions (Adolphs et al., 1999), and abnormal recognition of mental states from the eye region of faces (Adolphs et al., 2002), have also all been reported in rare patients with amygdala lesions, providing some support for a

long-standing hypothesis about the amygdala's involvement in ASD (Baron-Cohen et al., 2000).

Although there is evidence for global dysfunction at the level of the whole brain in ASD (Piven et al., 1995, Geschwind and Levitt, 2007, Amaral et al., 2008, Anderson et al., 2010), several studies emphasize abnormalities in the amygdala both morphometrically (Ecker et al., 2012) and in terms of functional connectivity (Gotts et al., 2012). Tying together the abnormal eye fixations onto faces in ASD mentioned above, and a correlation with amygdala processing, functional neuroimaging studies have found associations between abnormal fixation behavior and abnormal amygdala activation in people with ASD (Dalton et al., 2005, Kliemann et al., 2012). One recent study even found evidence for abnormal processing of information from the eye region of faces in single cells recorded from the amygdala in neurosurgical patients with ASD (Rutishauser et al., 2013). Despite considerable variability in reports of abnormal face processing in ASD, and despite the fact that there is brain dysfunction at a more global level in ASD, studies largely support (a) abnormal processing of faces in ASD, and (b) a link between this abnormality and amygdala function.

Much of the work cited above has focused on abnormal social processing in ASD in relation to the features of faces. Yet it is clear that the impairment is broader than this: two-year-olds with autism orient to non-social contingencies rather than biological motion (Klin et al., 2009), and attention to pictures of people is reduced in relation to pictures that are non-social when these compete for visual attention in arrays (Sasson et al., 2008, Sasson et al., 2011). We capitalized on these prior findings, and used the identical stimuli developed in these prior studies, with slight modification (see **Methods** for further details). Notably, these images provided stimuli that fell into three categories: social, non-social, and special interest. The prior findings had shown, both in children and adolescents (Sasson et al., 2008), as well as in 2–5 year-olds (Sasson et al., 2011), that participants with ASD fixated social images less than controls when freely viewing the arrays. Our approach here extends this prior work in four important respects:

(1) We assessed high-functioning adults with ASD, and also manipulated the difficulty of our task (number of items in the array) to test whether abnormal social attention would be revealed even in high-functioning adults;

(2) We provide a comparison to a small sample (three) of subjects with bilateral amygdala lesions, to enable comparisons between these two populations in light of the prior findings we reviewed above;

(3) We modified the experiment so that all subjects were performing a uniform search task for either social or non-social targets (rather than free viewing);

(4) We added a control experiment that equates the items in the search array for low-level visual properties (standard saliency, size, and distance to center).

Visual search tasks are not new to autism research. While a sizable literature in ASD has investigated search for simple, non-social objects (shapes and letters, etc.) and only manipulated low-level attributes of the stimuli (Plaisted et al., 1998, O'Riordan and Plaisted, 2001, O'Riordan et al., 2001, O'Riordan, 2004, Manjaly et al., 2007, Kemner et al., 2008), far fewer studies have examined visual search with social stimuli. In the present study, we used a more general framework that does not restrict the stimuli to specific facial emotions, or investigate internal features of faces, but tests competition for attention between natural social (faces and people with various emotions and poses) and non-social (e.g., furniture, toys and food) objects when presented simultaneously in a search array. We found that people with ASD had significantly fewer and slower fixations towards socially relevant objects, while fixations towards non-social objects were less impaired. This impairment was not evident in patients with amygdala lesions. The findings were replicated in a separate experiment with search arrays in which social objects and non-social objects had equal saliency, distance to center and sizes. Taken together, our findings argue that people with autism have attentional deficits that are disproportionate for social stimuli and that this deficit cannot be explained by low-level visual properties of the stimuli nor amygdala dysfunction.

## 5.4 Methods

### *5.4.1 Subjects*

In Experiment 1, eight high-functioning people with ASD were recruited (see **Table 5.1** and **Table 5.2**). All ASD participants met DSM-IV/ICD-10 diagnostic criteria for autism, and all met the cutoff scores for ASD on both the Autism Diagnostic Observation Schedule (ADOS) (Lord et al., 1989) and the Autism Diagnostic Interview-Revised (ADI-R) (LeCouteur et al., 1989) (**Table 5.1**). We assessed IQ for participants using the Wechsler Abbreviated Scale of Intelligence (WASI™). The ASD group had a full scale IQ of 106.9 ± 11.8 (mean ± SD).

**Table 5.1. List of ASD diagnosis and evaluation.**

Autism traits were evaluated by the Autism Diagnostic Observation Schedule (ADOS) and the Autism Diagnostic Interview-Revised (ADI-R). Cutoff scores for ASD on ADOS are 2 for A (communication) and 4 for B (social interaction). C is total (sum of A and B), and D is for stereotyped behavior. Cutoff scores for ASD on ADI-R are 10 for A (social interaction), 8 for B (communication) and 3 for C (stereotyped behavior). Higher score means more autistic.

Abbreviations: Exp: Experiments in which the subject participated. SCQ: Social Communication Questionnaire (cutoff score = 14). AQ: Autism Spectrum Quotient. SRS A-SR: Social Responsiveness Scale-2 Adult Form (Self Report). n.a: not available.

| | | ADOS | | | | ADI-R | | | | | SRS | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Exp** | **ID** | **A** | **B** | **C** | **D** | **A** | **B** | **C** | **D** | **SCQ** | **A-SR** | **AQ** |
| 1 | RA0780 | 5 | 11 | 16 | 1 | 29 | 18 | 10 | 4 | 31 | 63 | 17 |
| 1 | RA0796 | 4 | 9 | 13 | 0 | n.a. | n.a. | n.a. | n.a. | 7 | 71 | n.a. |
| 1 | RA0364 | 6 | 11 | 17 | 0 | 21 | 20 | 7 | 3 | 19 | 99 | 30 |
| 1 | RA0083 | 4 | 8 | 12 | 0 | 12 | 12 | 2 | 1 | n.a. | 78 | 27 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | RA0844 | 6 | 13 | 19 | 0 | n.a. | n.a. | n.a. | n.a. | 20 | 71 | 26 |
| 1 | RA0100 | 7 | 14 | 21 | 3 | 25 | 18 | 3 | 3 | 24 | 67 | 28 |
| 1 | RA0101 | 7 | 13 | 20 | 3 | 24 | 18 | 4 | 3 | 23 | 32 | 21 |
| 1,3 | RA0846 | 4 | 11 | 15 | 0 | n.a. | n.a. | n.a. | n.a. | n.a. | 60 | 33 |
| 2,3 | RA0582 | 3 | 5 | 8 | 3 | n.a. | n.a. | n.a. | n.a. | n.a. | 116 | n.a. |
| 2,3 | RA0784 | 2 | 5 | 7 | 0 | n.a. | n.a. | n.a. | n.a. | 26 | 94 | 26 |
| 2,3 | RA0085 | 4 | 9 | 13 | 1 | 21 | 11 | 6 | 3 | 12 | 114 | n.a. |
| 2,3 | RA0880 | 3 | 6 | 9 | 2 | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | 20 |
| 2,3 | RA0843 | 3 | 6 | 9 | 1 | n.a. | n.a. | n.a. | n.a. | n.a. | 66 | 29 |
| 2,3 | RA0584 | 3 | 4 | 7 | 3 | 14 | 12 | 5 | 0 | 21 | 92 | 28 |
| 2,3 | RA0080 | 6 | 14 | 20 | 2 | 16 | 14 | 5 | 1 | 15 | 110 | 39 |
| 2,3 | RA0869 | 3 | 8 | 11 | 0 | n.a. | n.a. | n.a. | n.a. | n.a. | 71 | 20 |
| 2,3 | RA0847 | 5 | 7 | 12 | 1 | n.a. | n.a. | n.a. | n.a. | n.a. | 90 | 31 |
| 2 | RA0871 | 2 | 6 | 8 | 2 | n.a. | n.a. | n.a. | n.a. | n.a. | 79 | 19 |
| 2,3 | RA0626 | 3 | 11 | 14 | 0 | 28 | 22 | 8 | 3 | 21 | 78 | 25 |
| 2,3 | RA0090 | 3 | 8 | 11 | 4 | 8 | 10 | 3 | 0 | n.a. | n.a. | 16 |
| 2,3 | RA0849 | 5 | 8 | 13 | 1 | n.a. | n.a. | n.a. | n.a. | 30 | 97 | 28 |

**Table 5.2. List of demographics and psychological evaluation for people with ASD and matched controls.**

Intelligence was measured by the Wechsler Abbreviated Scale of Intelligence (WASI).

Abbreviations: Exp: Experiments in which the subject participated. Age: age at testing. Hand: Dominant handedness (A: ambidextrous, L: left, R: right). WASI: IQ scores from the Wechsler Abbreviated Scale of Intelligence: full scale IQ (FSIQ), performance IQ (PIQ), verbal IQ (VIQ). n.a.: not available.

| Subject Category | Exp | ID | Age | Sex | Hand | Race | Education | WASI | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | FSIQ | PIQ | VIQ |
| | 1 | RA0780 | 25 | M | R | Asian/Pacific Islander | High School | 103 | 125 | 87 |
| | 1 | RA0796 | 26 | M | R | Caucasian | Bachelor's Degree | 133 | 127 | 131 |
| | 1 | RA0364 | 31 | M | A | Caucasian | Bachelor's Degree | 106 | 99 | 111 |
| | 1 | RA0083 | 26 | M | R | Caucasian | Some College | 106 | 118 | 94 |
| | 1 | RA0844 | 24 | M | R | Caucasian | n.a. | 107 | 103 | 109 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ASD | 1 | RA0100 | 23 | F | R | Caucasian | Some College | 107 | 110 | 102 |
| | 1 | RA0101 | 23 | F | R | Caucasian | Some College | 102 | 103 | 101 |
| | 1,3 | RA0846 | 33 | M | R | Caucasian | Bachelor's Degree | 91 | 111 | 50 |
| | 2,3 | RA0582 | 32 | M | R | Asian/Pacific Islander | Master's Degree | 124 | 115 | 127 |
| | 2,3 | RA0784 | 27 | M | R | Caucasian | Master's Degree | 128 | 121 | 129 |
| | 2,3 | RA0085 | 38 | F | A | Caucasian | Bachelor's Degree | 133 | 122 | 135 |
| | 2,3 | RA0880 | 28 | M | R | Caucasian | Bachelor's Degree | 108 | 99 | 114 |
| | 2,3 | RA0843 | 20 | F | A | Multiracial | Some College | 124 | 114 | 128 |
| | 2,3 | RA0584 | 26 | F | R | Caucasian | Bachelor's Degree | 125 | 119 | 123 |
| | 2,3 | RA0080 | 30 | M | L | Caucasian | Some College | 115 | 109 | 117 |
| | 2,3 | RA0869 | 32 | F | R | Hispanic/Latino | Bachelor's Degree | 88 | 85 | 95 |
| | 2,3 | RA0847 | 21 | M | R | Asian/Pacific Islander | Some College | 90 | 97 | 86 |
| | 2 | RA0871 | 44 | M | R | Caucasian | Associate's Degree | 89 | 80 | 101 |
| | 2,3 | RA0626 | 21 | M | A | Asian/Pacific Islander | Middle School | 125 | 119 | 123 |
| | 2,3 | RA0090 | 46 | M | R | Caucasian | Some College | 56 | 60 | 57 |
| | 2,3 | RA0849 | 21 | M | R | Hispanic/Latino | n.a. | 108 | 103 | 110 |
| ASD Controls | 2,3 | RA0782 | 32 | M | L | Hispanic/Latino | Bachelor's Degree | 104 | 114 | 95 |
| | 2,3 | RA0817 | 24 | M | R | Caucasian | Bachelor's Degree | 109 | 106 | 109 |
| | 2,3 | RA0829 | 30 | F | R | Caucasian | Bachelor's Degree | 116 | 111 | 116 |
| | 2,3 | RA0749 | 59 | M | R | Caucasian | Associate's Degree | 120 | 128 | 109 |
| | 2,3 | RA0548 | 46 | M | R | Caucasian | Some College | 97 | 109 | 85 |
| | 2,3 | RA0830 | 25 | M | R | Caucasian | Bachelor's Degree | 125 | 121 | 122 |
| | 2,3 | RA0842 | 32 | M | R | Caucasian | Bachelor's Degree | 117 | 110 | 119 |
| | 2,3 | RA0835 | 39 | F | R | Hispanic/Latino | Bachelor's Degree | 102 | 99 | 104 |

AP, AM and BG are three patients with selective bilateral amygdala lesions as a result of Urbach-Wiethe disease (Hofer, 1973). AM and BG are monozygotic twins. The details of these patients have been described previously (Buchanan et al., 2009, Becker et al., 2012). The anatomical scans of the lesions are shown in **Figure 5.1**. The amygdala group had a full scale IQ of 98.3 ± 2.5 (mean ± SD).

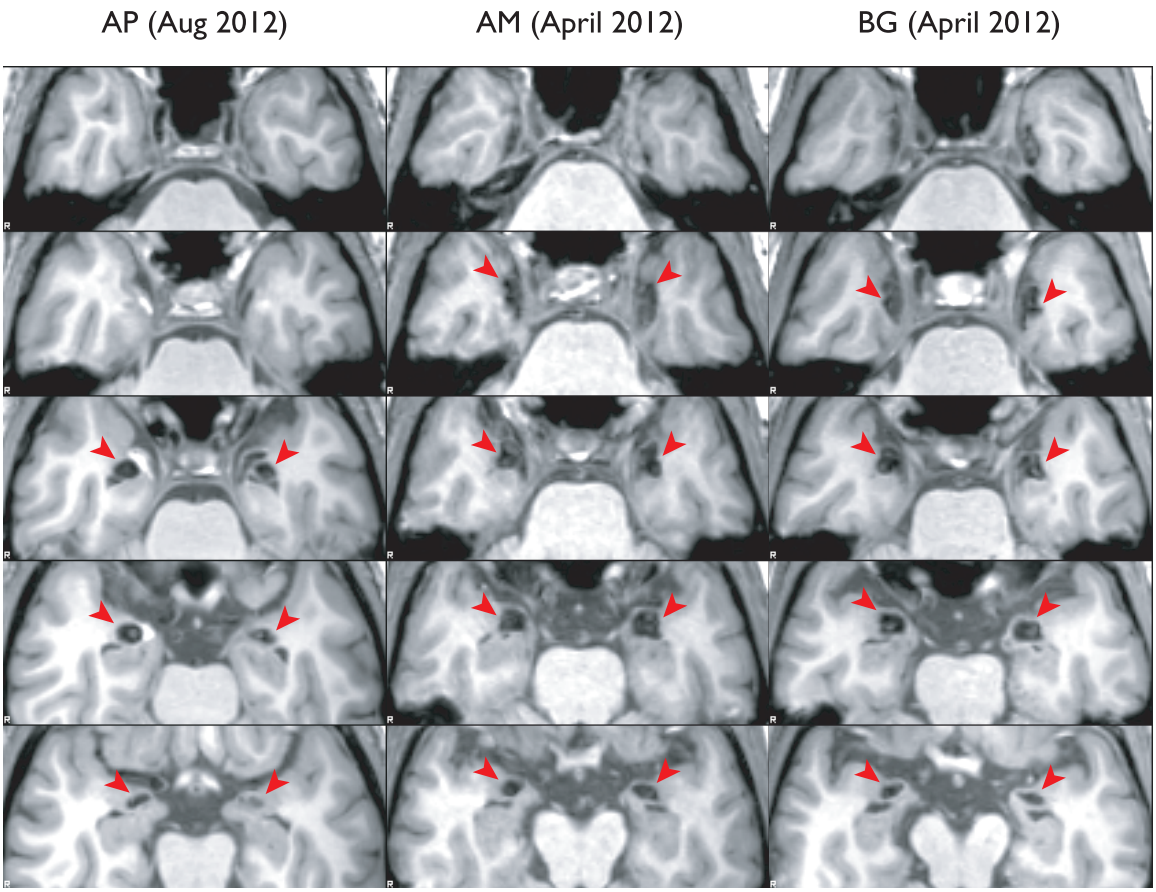


**Figure 5.1. MRI anatomical scans of the amygdala lesions.**

Displayed are high-resolution (0.5–1 mm isotropic) horizontal T1-weighted magnetic resonance imaging sections of the anterior medial temporal lobes with red arrows indexing the focal bilateral amygdala calcification damage. R: right.

Eight healthy subjects were recruited as general controls for both people with ASD and amygdala lesion patients, matched on IQ (full scale: 104.7 ± 6.1 (mean ± SD); t-test: $p = 0.68$ for people with ASD and $p = 0.13$ for amygdala patients) and education (**Table 5.3**).

**Table 5.3. List of demographics and psychological evaluation for amygdala lesion patients (AP, AM and BG) and general controls.**

Intelligence was measured by the Wechsler Abbreviated Scale of Intelligence (WASI). AM and BG's IQ was measured by the HAWIE-R ('Hamburg-Wechsler Intelligenztest für Erwachsene in revidierter Fassung'), a German-language adaptation of the WAIS-R (Wechsler Intelligence Test for Adults-Revised), which provides a measure of verbal, performance, and full-scale IQ.

Abbreviations: Age: age at testing. Hand: Dominant handedness (A: ambidextrous, L: left, R: right). Benton: Benton Facial Recognition Test, long form score. Benton scores 41–54 are in the normal range. WASI: IQ scores from the Wechsler Abbreviated Scale of Intelligence: full scale IQ (FSIQ), performance IQ (PIQ), verbal IQ (VIQ). n.a.: not available.

| ID | Age | Sex | Hand | Race | Education | Benton | WASI | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | FSIQ | PIQ | VIQ |
| AP | 27 | F | R | Asian/Pacific Islander | Bachelor's Degree | 50 | 98 | 106 | 92 |
| AM | 38 | F | A | Caucasian | 13 years of education in Germany | 36 | 101 | 103 | 99 |
| BG | 38 | F | R | Caucasian | 13 years of education in Germany | 41 | 96 | 97 | 94 |
| RA0629 | 32 | F | A | Caucasian | Some College | n.a. | n.a. | n.a. | n.a. |

| ID | Age | Sex | Hand | Race | Education | Benton | WASI | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | FSIQ | PIQ | VIQ |
| RA0633 | 27 | F | R | Asian/Pacific Islander | Bachelor's Degree | n.a. | n.a. | n.a. | n.a. |
| RA0762 | 23 | F | A | Hispanic/ Latino | Some College | 50 | 100 | 105 | 95 |
| RA0764 | 31 | F | R | Caucasian | Master's Degree | n.a. | 102 | 103 | 101 |
| RA0829 | 29 | F | R | Caucasian | Bachelor's Degree | n.a. | 116 | 111 | 116 |
| RA0835 | 38 | F | R | Hispanic/ Latino | Bachelor's Degree | 49 | 102 | 99 | 104 |
| RA0848 | 40 | F | R | Caucasian | High School | n.a. | 101 | 104 | 98 |
| RA0851 | 35 | F | R | Caucasian | Bachelor's Degree | n.a. | 107 | 103 | 108 |

Eleven students from the National University of Singapore (NUS) were tested for all three versions of the task (Experiment 1–3) to provide an independent reference group. Subjects gave written informed consent and the experiments were approved by the Caltech and NUS Institutional Review Boards. All subjects had normal or corrected-to-normal visual acuity.

In Experiment 2 we tested thirteen high-functioning people with ASD (different from those who participated in Experiment 1; see **Table 5.1** and **Table 5.2**), eight healthy ASD controls (**Table 5.2**) and eleven NUS control subjects (the same as Experiment 1; experiment order counterbalanced). The ASD group had a full scale IQ of 108.7 ± 22.3 (mean ± SD) and ASD controls had a comparable full scale IQ of 111.3 ± 9.8 (t-test, $p = 0.76$). The ASD group had a mean age of 29.7 ± 8.6 years and ASD controls had a mean

age of 35.9 ± 11.8 years (t-test, p = 0.18). ASD controls also matched on gender, race and education.

In Experiment 3, we tested the same three amygdala lesion patients from Experiment 1 (AP, AM and BG), thirteen high-functioning people with ASD (see **Table 5.1** and **Table 5.2**), eight healthy ASD-matched controls (the same as Experiment 2; **Table 5.2**), and eleven NUS control subjects (the same as Experiment 1 and 2; experiment order counterbalanced). The ASD group had a full scale IQ of 108.8 ± 22.1 (mean ± SD) and ASD controls had a comparable full scale IQ of 111.3 ± 9.8 (t-test, p = 0.78). The ASD group had a mean age of 28.8 ± 7.6 years and ASD controls had a mean age of 35.9 ± 11.8 years (t-test, p = 0.11). ASD controls also matched on gender, race and education.

*5.4.2 Stimuli and apparatus*

We used 20 distinct visual search arrays. In each array there were 24 objects whose spatial locations were randomized between the 20 arrays. 12 objects were social (faces and people with different postures, emotions, ages, and genders, etc.) and 12 objects were non-social (furniture, toys, food, etc.). These social and non-social objects composing the array stimuli have been characterized and described previously (Sasson et al., 2012) and were obtained from two prior studies that investigated visual attention in infants and children with ASD (Sasson et al., 2008, Sasson et al., 2011). From each array stimulus, we randomly assigned 4 social objects and 4 non-social objects as targets (on 8 distinct trials). For each array, we also had 2 catch trials, i.e., the target was not among the objects in the search array (one catch trial with a social target, and one with a non-social target). Therefore, in total we had 100 trials with social targets and 100 trials with non-social targets, and 20% of trials were catch trials.

The experimental setup of Experiment 2 was identical to Experiment 1 except that low-level properties of social and non-social objects were equalized within each search array. The social and non-social objects did not differ in standard low-level saliency as

quantified by the Itti-Koch model (Itti et al., 1998, Itti and Koch, 2001), distance to the center or size (all ps > 0.79; **Figure 5.2A–C**).

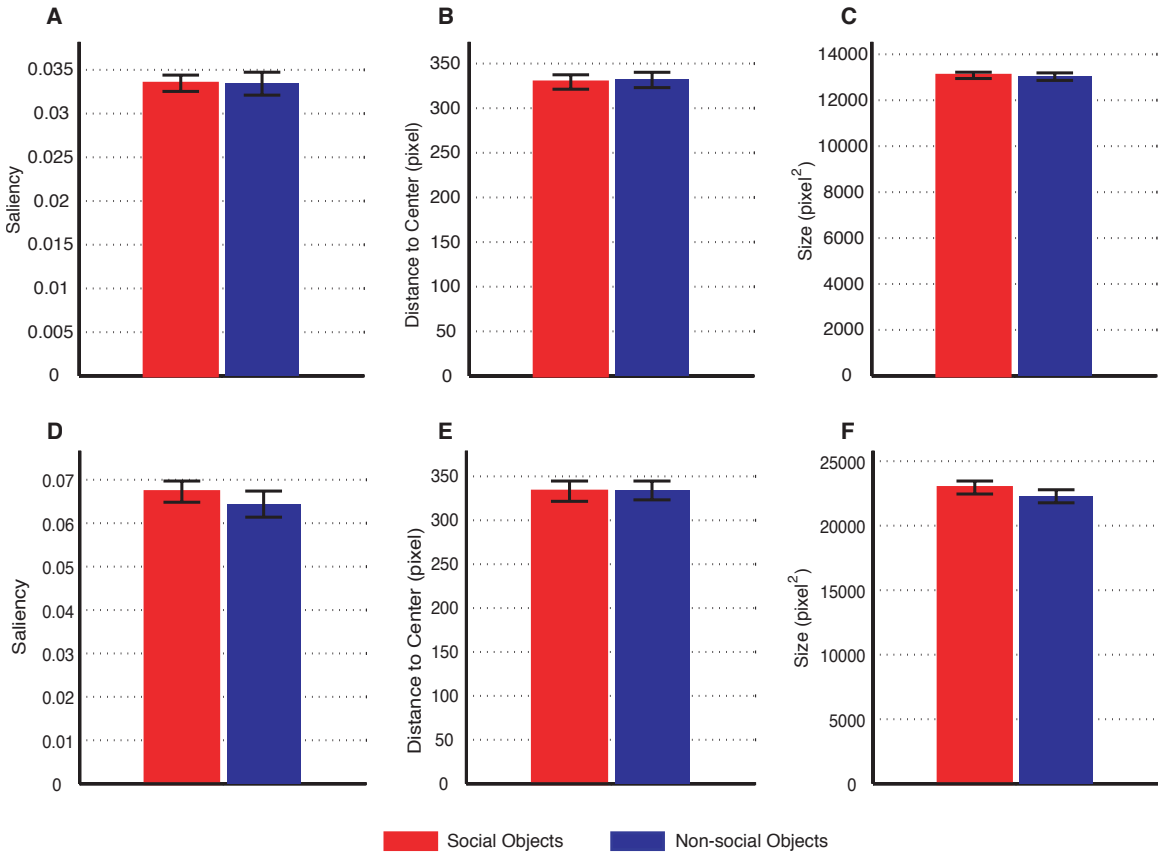


**Figure 5.2. Low-level properties of the stimuli.**

**(A–C)** Standard arrays used in Experiment 2. **(D–F)** Simpler arrays used in Experiment 3. **(A)** Standard low-level saliency measured with Itti-Koch model (Itti et al., 1998, Itti and Koch, 2001) did not differ between social and non-social objects in the search array (two-tailed t-test, p = 0.98 for standard arrays and p = 0.46 for simpler arrays). The sum of saliency of all objects was normalized to 1 for each search array. **(B)** Distance to center did not differ between social and non-social objects (measured in pixel, p = 0.85 for standard arrays and p = 0.96 for simpler arrays). **(C)** Object size did not differ between

social and non-social objects (measured in pixel$^2$, $p = 0.79$ for standard arrays and $p = 0.34$ for simpler arrays).

The experimental setup of Experiment 3 was identical to Experiment 2 except that there were only 12 objects in total in each search array (6 social and 6 non-social). Low-level properties of social and non-social objects were also equalized within each search array, as we had done for Experiment 2 (**Figure 5.2D–F**). The social and non-social objects did not differ in standard low-level saliency, distance to the center, or size (all $ps > 0.34$).

Subjects sat approximately 65 cm from an LCD display with a 23-inch screen (screen resolution: 1920x1080). The refresh rate of the display was 60 Hz and the stimuli occupied the center of the display (14.9° × 11.2° visual angle). Stimuli were presented using MATLAB with the Psychtoolbox 3 (Brainard, 1997) (http://psychtoolbox.org).

*5.4.3 Task*

We used a standard visual search task (**Figure 5.3**). A target was presented for 1 second followed by the search array. Subjects were instructed to find the object in the array that matched the target and explicitly told that the array might or might not contain the target. The search array stayed up for at most 14 seconds, or until the subject responded, either by pushing the space bar to indicate that the target was found in the array, or by pushing the button 'N' to indicate the target was absent in the array. If they pushed the space bar in target-present trials, subjects were asked to click on the target object in the array with a mouse. If subjects clicked on the correct target, a message 'Correct' was displayed to the subjects for 1 second. Otherwise, a message 'Incorrect' was displayed for 1 second. Subjects were instructed to respond as quickly and accurately as possible. If subjects did not respond within 14 seconds after array onset, a message 'Time Out' was displayed. An inter-trial-interval (ITI) was jittered between 1 to 2 seconds. The array and target orders

were completely randomized for each subject. Subjects practiced 5 trials before the experiment to familiarize themselves with the task.

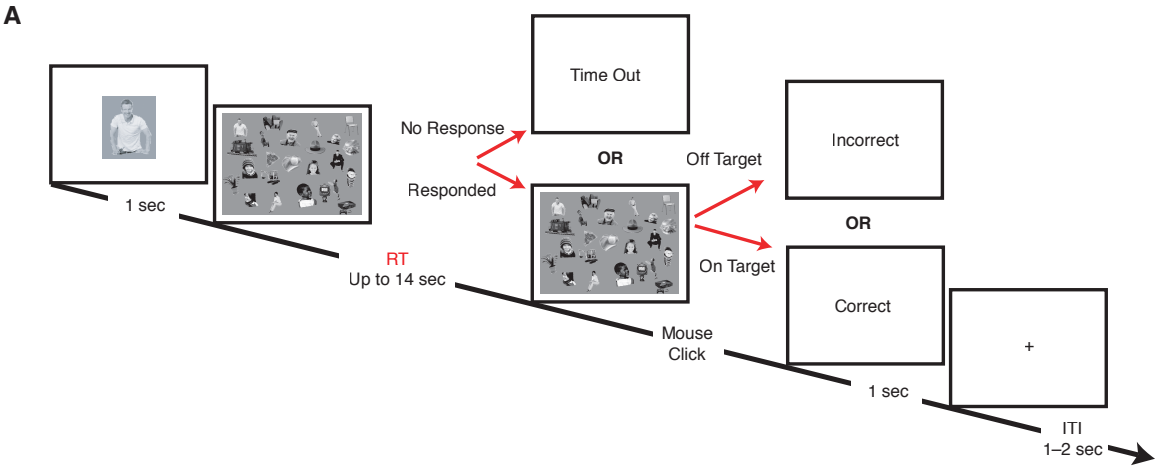


**Figure 5.3. Task and sample stimuli.**

**(A)** Task structure. A target is presented for 1 second followed by the search array. Subjects have a maximum of 14 seconds to respond by pressing the space bar to indicate that the target is present, or the letter 'N' to indicate that the target is absent. Following target detection, subjects provide a mouse click on targets. A feedback message of 'Correct', 'Incorrect' or 'Time Out' is displayed for 1 second before an ITI of 1 to 2 seconds. **(B)** Sample visual search arrays with fixations. Left: standard array used in Experiment 2. Right: simple array used in Experiment 3. Each circle represents a fixation.

Green circle: start fixation. magenta circle: end fixation. Yellow line: eye movement (saccade). Red box: target.

### *5.4.4 Eye tracking*

Eye tracking was carried out using a non-invasive infra-red remote Tobii X300 system which recorded binocular gaze at 300 Hz. The Tobii visualization software (Tobii Studio™ 2.2) was used to record eye movements and perform gaze analysis. Fixations were detected by Tobii Fixation Filter implemented in Tobii Studio. The Tobii Fixation Filter is a classification algorithm proposed by (Olsson, 2007) and detects quick changes in the gaze point using a sliding window averaging method. Velocity threshold was set to 35 [pixels/samples] and distance threshold was set to 35 [pixels] in our study.

NUS control subjects were recorded with a non-invasive infra-red Eyelink 1000 system (SR Research, Canada). One of the eyes was tracked at 2000 Hz. The eye tracker was calibrated with the built-in 9-point grid method.

### *5.4.5 Data analysis*

Prior to data collection, we defined a rectangular region that encompassed each target as the target region to define acceptable mouse click locations for each search. In Experiment 1, out of 4800 target-present trials, in 4547 trials (94.73%) subjects found the target and clicked within the pre-defined areas (correct trials). Subjects missed targets altogether (judged target-present trials as target-absent) in 183 trials (3.81%) and correctly reported target presence but clicked outside the target rectangle in 69 trials (1.44%) (both are incorrect trials). Subjects did not respond within 14 seconds after array onset (time-out trials) in only 1 trial (0.021%). Out of 1200 target-absent trials, subjects had 1129 (94.08%) correct trials, 70 (5.83%) false-alarm trials (reported target presence in target-absent trials), and 1 (0.08%) time-out trial. We found similar percentages of

correct, incorrect, time-out and false-alarm trials for Experiment 2 and Experiment 3. We only analyzed correct target-present trials (correct target-present response followed by correct identification of the target). Further, we only included trials with reaction times (RTs, with respect to search array onset) within ± 2.5 standard deviations for all analyses (in Experiment 1, 114 trials were excluded, 2.51%). There was no difference between participants with ASD, amygdala patients and control subjects in any of the above proportions (all ps > 0.10).

## 5.5 Results

### *5.5.1 Behavioral Performance: Accuracy and Reaction Time*

We first analyzed the behavioral performance of all subject groups. Across all three experiments, all subject groups (ASD, ASD controls, amygdala lesions, general controls and NUS students) had an average performance above 90% (**Figure 5.4**), indicating that they were able to perform the task without difficulty. In Experiment 1 (**Figure 5.4A**), only a marginal difference was found between social targets and non-social targets (two-way mixed ANOVA (target type X subject group); main effect of target type: $F(1,26) = 4.21$, $p = 0.051$, effect size $\eta^2 = 0.030$), and no difference was found between subject groups (main effect of subject group: $F(3,26) = 1.58$, $p = 0.22$, $\eta^2 = 0.12$) or interaction ($F(3,26) = 0.94$, $p = 0.44$, $\eta^2 = 0.020$). Similarly, in Experiment 2 (**Figure 5.4C**), no difference was found between social targets and non-social targets (main effect of target type: $F(1,29) = 0.17$, $p = 0.69$, $\eta^2 = 0.0022$), and no difference was found between subject groups (main effect of subject group: $F(2,29) = 0.86$, $p = 0.43$, $\eta^2 = 0.034$) or any interaction ($F(2,29) = 0.32$, $p = 0.73$, $\eta^2 = 0.0086$). Finally, also in Experiment 3 (**Figure 5.4E**), no difference was found between social targets and non-social targets (main effect of target type: $F(1,31) = 3.59$, $p = 0.068$, $\eta^2 = 0.036$), nor between subject groups (main effect of subject group: $F(3,31) = 0.15$, $p = 0.93$, $\eta^2 = 0.0089$) or interaction ($F(3,31) =$

0.58, p = 0.63, $\eta^2$ = 0.018), showing that overall people with ASD and amygdala lesion patients still had similar performance in terms of accuracy compared to controls.

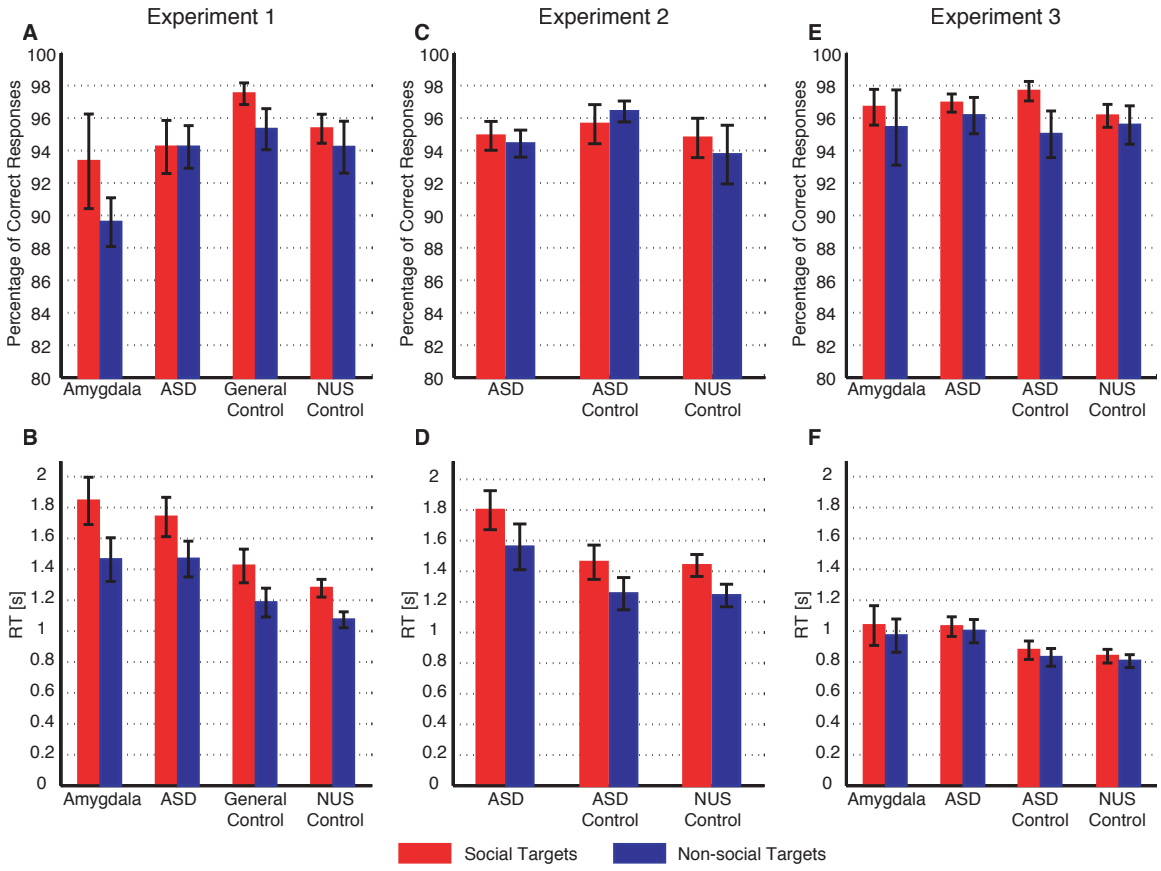


**Figure 5.4. Behavioral performance.**

**(A–B)** Experiment 1. **(C–D)** Experiment 2. **(E–F)** Experiment 3. **(A,C,E)** Percentage of correct response. **(B,D,F)** Reaction time (RT). Error bars denote one SEM of the mean.

In Experiment 1, non-social targets were detected more quickly by all subject groups (**Figure 5.4B**; two-way mixed ANOVA (target type X subject group); main effect of target type: F(1,26) = 199.4, p = $1.05\times10^{-13}$, $\eta^2$ = 0.13), an effect that showed only a weak interaction with subject group (F(3,26) = 2.83, p = 0.058, $\eta^2$ = 0.0057). General control

subjects and NUS control subjects showed overall faster detection of targets (main effect of subject group: $F(3,26) = 5.40$, $p = 0.0050$, $\eta^2 = 0.32$), but there was no difference between amygdala patients vs. general controls, amygdala patients vs. people with ASD, people with ASD vs. general controls, or general controls vs. NUS controls (two-tailed t-test, all ps > 0.05). In Experiment 2, non-social targets still featured faster detection due to their being more distinct from one another (**Figure 5.4D**; main effect of target type: $F(1,29) = 75.4$, $p = 1.47\times10^{-9}$, $\eta^2 = 0.068$), but the faster detection of non-social targets did not depend on subject groups (the interaction of target type and subject group: $F(2,29) = 0.31$, $p = 0.74$, $\eta^2 = 5.60\times10^{-4}$; main effect of subject group: $F(2,29) = 3.01$, $p = 0.065$, $\eta^2 = 0.16$). Notably, across independent samples of people with ASD, we found no difference between Experiment 1 and Experiment 2 in detection accuracy (unpaired t-test: $t(19) = -0.69$, $p = 0.50$, effect size in Hedges's g (standardized mean difference): $g = -0.30$) or RT ($t(19) = -0.39$, $p = 0.70$, $g = -0.17$). This argues against any influence of low-level visual properties on our task.

With simpler arrays in Experiment 3, non-social targets that were more distinct from one another retained their advantage to be detected faster (**Figure 5.4F**; two-way mixed ANOVA (target type X subject group); main effect of target type: $F(1,31) = 13.2$, $p = 9.82\times10^{-4}$, $\eta^2 = 0.0078$). ASD controls and NUS controls showed marginally faster detection of targets (main effect of subject group: $F(3,31) = 2.38$, $p = 0.088$, $\eta^2 = 0.18$), but there was no interaction ($F(3,31) = 0.38$, $p = 0.77$, $\eta^2 = 6.80\times10^{-4}$) or significant difference between amygdala patients vs. people with ASD, people with ASD vs. ASD controls, or ASD controls vs. NUS controls (two-tailed t-tests separately for social vs. non-social targets, all ps > 0.05). Lower task difficulty was confirmed with a shorter RT compared to Experiment 1 (paired t-test for NUS controls: $t(10) = 10.2$, $p = 1.38\times10^{-6}$, $g = 2.11$; paired t-test between amygdala patients: $t(2) = 17.7$, $p = 0.0032$, $g = 2.27$; unpaired t-test between people with ASD: $t(19) = 4.56$, $p = 2.13\times10^{-4}$, $g = 1.97$) and Experiment 2 (paired t-test for ASD controls: $t(7) = 6.13$, $p = 4.76\times10^{-4}$, $g = 1.95$; paired t-test for NUS controls: $t(10) = 10.8$, $p = 7.83\times10^{-7}$, $g = 2.53$).

*5.5.2 Eye tracking: general social preference does not differ between subject groups*

We first tested whether people with ASD had reduced global preference to look at social objects in the search array. For each fixation, we calculated a social bias in attention as the difference between the percentage of all fixations (within a specified number of serial order of fixations) on social objects as compared to non-social objects (**Figure 5.5**). In Experiment 1 (**Figure 5.5A**), we observed an overall reduced proportion of fixations onto social objects for people with ASD (one-way ANOVA across four subject groups on the average social bias for fixations number 2 to 10: ASD: 5.97 ± 2.31 (mean ± SEM), amygdala: 15.10 ± 4.91, general control: 17.24 ± 1.39, NUS control: 14.06 ± 1.96; $F(3,26) = 5.02$, $p = 0.007$, $\eta^2 = 0.37$; two-tailed t-test compared to general controls: $t(14) = -4.18$, $p = 9.17\times10^{-4}$, $g = -1.98$). This reduced social preference persisted over time as both early fixations (average of fixations 2 to 5) and late fixations (average of fixations 6 to 10) showed a difference compared to general controls (Early: ASD: 2.66 ± 1.90, general control: 10.66 ± 1.55; $t(14) = -3.26$, $p = 0.0057$, $g = -1.54$; Late: ASD: 8.63 ± 3.11, general control: 23.50 ± 3.19; $t(14) = -3.34$, $p = 0.0049$, $g = -1.58$), although fixation-by-fixation comparisons across subject groups (one-way ANOVA) and with general controls did not reveal reliable differences when corrected for multiple comparisons with false discovery rate (FDR) of 0.05 (Benjamini and Hochberg, 1995) (all statistical comparisons are listed in **Table 5.4**). Comparing people with ASD and amygdala lesion patients, we observed differences only for early fixations (ASD: 2.66 ± 1.90, amygdala: 10.55 ± 2.05; $t = 2.32$, $p = 0.045$, $g = 1.44$; also difference at the 2$^{nd}$ and 4$^{th}$ fixations). However, we observed no difference in social preference between amygdala lesion patients and general controls (two-tailed t-test; $p > 0.05$ for all fixations and averages; see statistics in **Table 5.4**). Furthermore, we observed no difference between general controls and NUS controls ($p > 0.05$ for all fixations and averages). Our results suggest a possibly mildly reduced bias for social preference in ASD.

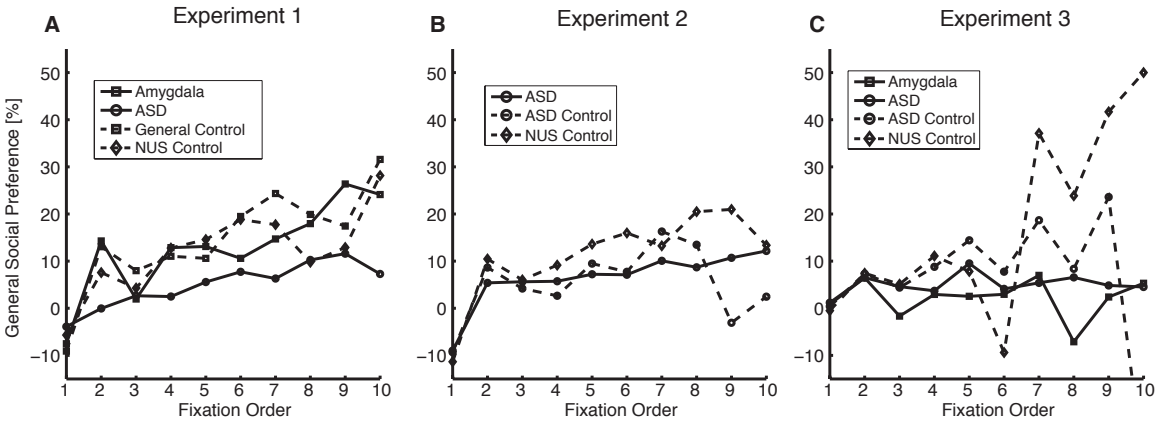


**Figure 5.5. General social preference.**

**(A)** Experiment 1. **(B)** Experiment 2. **(C)** Experiment 3. We calculated social preference as the average number of fixations (irrespective of task condition) across all trials that fell onto social stimuli, minus the average number of fixations that fell onto non-social stimuli, expressed as a percentage.

**Table 5.4. Statistical results for general social preference.**

*All* is the average of fixation 2 to 10. *Early* is the average of fixation 2 to 5, and *Late* is the average of fixation 6 to 10.

| **Experiment 1** | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *One-Way ANOVA* | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| F-statistic F(3,26) | 1.00 | 5.61 | 0.82 | 3.84 | 2.04 | 1.45 | 2.35 | 0.61 | 0.54 | 1.30 | 5.02 | 5.85 | 2.80 |
| p-value | 0.41 | 0.0042 | 0.50 | 0.021 | 0.13 | 0.25 | 0.10 | 0.61 | 0.66 | 0.30 | 0.0070 | 0.0034 | 0.060 |
| Effect Size | 0.10 | 0.39 | 0.086 | 0.31 | 0.19 | 0.14 | 0.21 | 0.066 | 0.060 | 0.13 | 0.37 | 0.40 | 0.24 |
| *Amygdala vs. ASD* | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| t-statistic t(9) | -0.71 | 2.72 | -0.13 | 2.34 | 1.43 | 0.31 | 1.27 | 0.67 | 1.67 | 1.66 | 1.92 | 2.32 | 1.53 |
| p-value | 0.50 | 0.024 | 0.90 | 0.044 | 0.19 | 0.76 | 0.23 | 0.52 | 0.13 | 0.13 | 0.086 | 0.045 | 0.16 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Effect Size Hedges's g | -0.44 | 1.68 | -0.08 | 1.45 | 0.88 | 0.19 | 0.79 | 0.42 | 1.03 | 1.03 | 1.19 | 1.44 | 0.95 |
| *Amygdala vs. General control* | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| t-statistic t(9) | 0.38 | 0.23 | -1.14 | 0.46 | 0.55 | -0.96 | -0.94 | -0.17 | 0.90 | -0.57 | -0.60 | -0.040 | -0.71 |
| p-value | 0.72 | 0.82 | 0.29 | 0.66 | 0.60 | 0.36 | 0.37 | 0.87 | 0.39 | 0.59 | 0.57 | 0.97 | 0.50 |
| Effect Size Hedges's g | 0.23 | 0.14 | -0.70 | 0.28 | 0.34 | -0.59 | -0.58 | -0.11 | 0.56 | -0.35 | -0.37 | -0.025 | -0.44 |
| *ASD vs. General control* | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| t-statistic t(14) | 1.66 | -3.24 | -1.23 | -2.63 | -1.77 | -1.79 | -2.68 | -1.41 | -0.83 | -2.59 | -4.18 | -3.26 | -3.34 |
| p-value | 0.12 | 0.0059 | 0.24 | 0.020 | 0.10 | 0.094 | 0.018 | 0.18 | 0.42 | 0.022 | $9.17 \times 10$ | 0.0057 | 0.0049 |
| Effect Size Hedges's g | 0.78 | -1.53 | -0.58 | -1.24 | -0.84 | -0.85 | -1.26 | -0.66 | -0.41 | -1.26 | -1.98 | -1.54 | -1.58 |
| *General control vs. NUS Control* | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| t-statistic t(17) | -1.35 | 1.78 | 1.07 | -0.52 | -1.06 | 0.10 | 0.91 | 1.14 | 0.44 | 0.23 | 1.23 | 0.46 | 1.21 |
| p-value | 0.19 | 0.094 | 0.30 | 0.61 | 0.30 | 0.92 | 0.37 | 0.27 | 0.67 | 0.82 | 0.24 | 0.65 | 0.24 |
| Effect Size Hedges's g | -0.60 | 0.79 | 0.48 | -0.23 | -0.47 | 0.046 | 0.41 | 0.50 | 0.20 | 0.10 | 0.54 | 0.21 | 0.54 |
| **Experiment 2** | | | | | | | | | | | | | |
| *One-Way ANOVA* | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| F-statistic F(2,29) | 0.54 | 0.98 | 0.15 | 1.88 | 0.80 | 2.25 | 0.70 | 1.92 | 4.06 | 0.35 | 2.28 | 1.14 | 2.14 |
| p-value | 0.59 | 0.39 | 0.86 | 0.17 | 0.46 | 0.12 | 0.50 | 0.16 | 0.03 | 0.71 | 0.12 | 0.33 | 0.14 |
| Effect Size | 0.036 | 0.064 | 0.010 | 0.11 | 0.052 | 0.13 | 0.05 | 0.12 | 0.22 | 0.02 | 0.14 | 0.073 | 0.13 |
| *ASD vs. ASD Control* | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| t-statistic t(19) | 0.12 | -0.97 | 0.39 | 0.84 | -0.52 | -0.13 | -1.09 | -0.80 | 1.66 | 1.20 | 0.35 | -0.09 | 0.51 |
| p-value | 0.90 | 0.34 | 0.70 | 0.41 | 0.61 | 0.90 | 0.29 | 0.43 | 0.11 | 0.25 | 0.73 | 0.93 | 0.62 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Effect Size Hedges's g | 0.053 | -0.42 | 0.17 | 0.36 | -0.22 | -0.055 | -0.47 | -0.35 | 0.72 | 0.52 | 0.15 | -0.037 | 0.22 |
| *ASD Control vs. NUS Control* | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| t-statistic t(17) | -0.74 | 0.41 | 0.48 | 2.21 | 0.67 | 2.00 | -0.62 | 1.03 | 2.90 | 0.65 | 2.30 | 1.24 | 2.20 |
| p-value | 0.47 | 0.69 | 0.64 | 0.041 | 0.51 | 0.062 | 0.54 | 0.32 | 0.010 | 0.52 | 0.034 | 0.23 | 0.042 |
| Effect Size Hedges's g | -0.33 | 0.18 | 0.21 | 0.98 | 0.30 | 0.89 | -0.28 | 0.46 | 1.29 | 0.29 | 1.02 | 0.55 | 0.98 |
| **Experiment 3** | | | | | | | | | | | | | |
| *One-Way ANOVA* | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| F-statistic F(3,31) | 0.36 | 0.050 | 0.79 | 2.10 | 0.69 | 1.13 | 1.92 | 0.51 | 0.63 | 1.21 | 1.20 | 1.13 | 0.62 |
| p-value | 0.79 | 0.98 | 0.51 | 0.12 | 0.56 | 0.35 | 0.15 | 0.68 | 0.60 | 0.34 | 0.32 | 0.35 | 0.61 |
| Effect Size | 0.032 | 0.0048 | 0.071 | 0.17 | 0.063 | 0.10 | 0.18 | 0.056 | 0.083 | 0.22 | 0.10 | 0.10 | 0.057 |
| *ASD vs. ASD Control* | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| t-statistic t(19) | 0.52 | -0.24 | 0.064 | -1.44 | -0.78 | -0.47 | -1.09 | -0.10 | -0.73 | 1.15 | -0.73 | -1.09 | -0.63 |
| p-value | 0.61 | 0.81 | 0.95 | 0.17 | 0.45 | 0.64 | 0.29 | 0.92 | 0.48 | 0.27 | 0.48 | 0.29 | 0.54 |
| Effect Size Hedges's g | 0.22 | -0.10 | 0.028 | -0.62 | -0.34 | -0.20 | -0.48 | -0.043 | -0.37 | 0.83 | -0.31 | -0.47 | -0.27 |
| *Amygdala vs. ASD* | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| t-statistic t(14) | -0.14 | -0.008 | -1.50 | -0.15 | -1.11 | -0.18 | 0.10 | -0.63 | -0.13 | 0.032 | -0.69 | -1.24 | -0.21 |
| p-value | 0.89 | 0.99 | 0.16 | 0.88 | 0.29 | 0.86 | 0.92 | 0.54 | 0.90 | 0.97 | 0.50 | 0.23 | 0.83 |
| Effect Size Hedges's g | -0.09 | -0.005 | -0.91 | -0.090 | -0.67 | -0.11 | 0.069 | -0.45 | -0.10 | 0.023 | -0.42 | -0.75 | -0.13 |
| *Amygdala vs. ASD Control* | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| t-statistic t(9) | 0.27 | -0.18 | -1.44 | -1.34 | -1.06 | -0.33 | -0.49 | -0.38 | -0.34 | 0.58 | -0.67 | -1.60 | -0.41 |
| p-value | 0.79 | 0.86 | 0.19 | 0.21 | 0.32 | 0.75 | 0.64 | 0.71 | 0.75 | 0.62 | 0.52 | 0.14 | 0.69 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Effect Size Hedges's g | 0.17 | -0.11 | -0.89 | -0.83 | -0.66 | -0.20 | -0.35 | -0.27 | -0.24 | 0.33 | -0.41 | -0.99 | -0.26 |

However, this was not borne out in Experiment 2 (**Figure 5.5B**). When low-level saliency between social and non-social objects was equalized, people with ASD showed entirely normal general social preference as compared to ASD controls (one-way ANOVA across three subject groups, $p > 0.05$ for all fixations and averages; two-tailed t-tests compared to ASD controls, $p > 0.05$ for all fixations and averages; see **Table 5.4**).

Similarly, in Experiment 3 (**Figure 5.5C**) in which low-level saliency between social and non-social objects was also equalized, people with ASD showed normal general social preference to our stimuli as compared to ASD controls (one-way ANOVA across four subject groups, $p > 0.05$ for all fixations and averages; two-tailed t-tests compared to ASD controls, $p > 0.05$ for all fixations and averages; see **Table 5.4**). Amygdala lesion patients also had normal social preference compared to ASD controls ($p > 0.05$ for all fixations and averages) and similar social preference compared to people with ASD ($p > 0.05$ for all fixations and averages), suggesting that neither people with ASD nor amygdala lesion patients have global deficits in social preference.

### *5.5.3 Reduced orientation towards target-relevant objects in visual search*

The above analysis showed that people with ASD do not have globally reduced social preferences, once low-level saliency was equalized. But how might social attention interact with task demands during visual search? We next analyzed target-relevant effects to answer this question.

All subjects oriented to social objects rapidly and kept on searching within social objects if the target was social (**Figure 5.6** upper row). Pronounced differences in the proportion

of fixations onto social and non-social objects were evident as early as the 2nd fixation and lasted until the 10th fixation. Symmetrically, when searching for a non-social target (**Figure 5.6** lower row), subjects oriented to non-social objects and kept on searching within non-social objects.

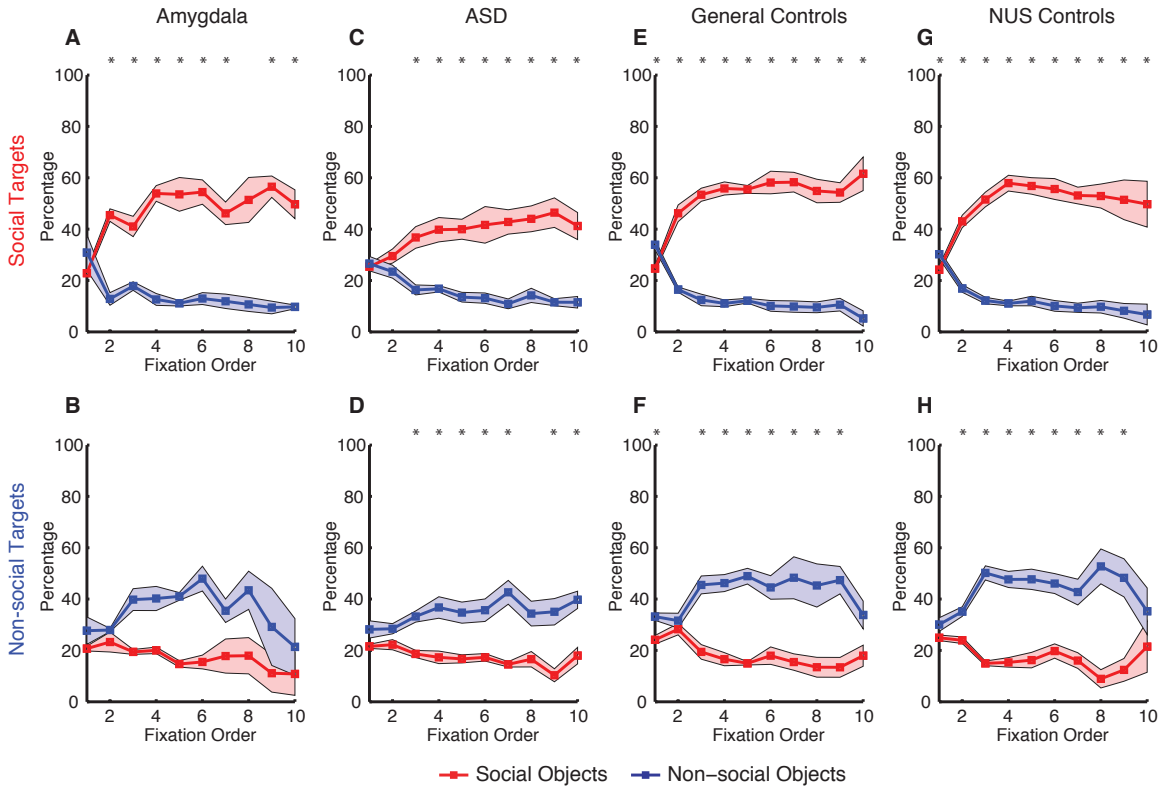


**Figure 5.6. Social and non-social target effects in Experiment 1.**

All subjects looked at target-congruent objects in a fast and sustained manner. **(A–B)** Amygdala patients. **(C–D)** People with ASD. **(E–F)** General controls. **(G–H)** NUS controls. Red: social objects. Blue: non-social objects. Upper row **(A,C,E,G)**: when searching for social targets. Lower row **(B,D,F,H)**: when searching for non-social targets. Asterisk indicates significant difference between target congruent objects and target incongruent objects (two-tailed paired t-test: $p < 0.05$, FDR corrected). Shaded area denotes ± SEM over the group of subjects.

We define a target-relevant effect as the difference in the percentage of fixations on target-congruent objects and the percentage of fixations on target-incongruent objects. All subjects showed rapid and sustained target-relevant effects, for both social targets and non-social targets (**Figure 5.7**). In Experiment 1, we found disproportionate target-relevant effects between social and non-social stimuli across fixations (two-way mixed ANOVA (target type X subject group); main effect of target type; average of fixations 2 to 10: social: 37.84 ± 2.31, non-social: 24.69 ± 1.72; $F(1,26) = 55.4$, $p = 6.63\times10^{-8}$, $\eta^2 = 0.26$; see **Table 5.5** for statistics), showing stronger attention towards social objects than non-social objects when searching for their respective targets. Both early (average of fixations 2 to 5: social: 33.54 ± 2.39, non-social: 21.97 ± 1.88; $F(1,26) = 43.9$, $p = 4.97\times10^{-7}$, $\eta^2 = 0.20$) and late fixations (average of fixations 6 to 10: social: 41.53 ± 2.58, non-social: 27.27 ± 2.52; $F(1,26) = 26.3$, $p = 2.38\times10^{-5}$, $\eta^2 = 0.21$) showed stronger social target-relevant effects, which persisted through fixation 7. However, here we also found pronounced target-relevant effects that differed between subject groups (main effect of subject group; average of fixations 2 to 10 collapsing social and non-social: ASD: 22.20 ± 3.30, amygdala: 28.81 ± 1.02, general control: 35.64 ± 3.05, NUS control: 35.34 ± 2.53; $F(3,26) = 4.76$, $p = 8.94\times10^{-3}$, $\eta^2 = 0.21$), especially during early fixations (average of fixations 2 to 5: ASD: 16.78 ± 3.54, amygdala: 26.38 ± 1.35, general control: 31.47 ± 2.62, NUS control: 33.41 ± 2.58; $F(3,26) = 6.79$, $p = 1.57\times10^{-3}$, $\eta^2 = 0.28$), with people with ASD showing reduced target-relevant effects. The reduced target-relevant effect in people with ASD persisted from the 2$^{nd}$ fixation to the 5$^{th}$ fixation, showing that they did not look at relevant targets as rapidly as controls during the initial fixations of their search. However, there was no difference between people with ASD and controls for later fixations (average of fixations 6 to 10: ASD: 26.54 ± 3.56, amygdala: 30.76 ± 2.14, general control: 40.11 ± 5.39, NUS control: 36.96 ± 2.81; $F(3,26) = 2.38$, $p = 0.093$, $\eta^2 = 0.12$; also see **Table 5.5** for fixation-by-fixation analysis), showing that people with ASD could catch up at later points in time. Although the impaired target-congruency effect in ASD was qualitatively more pronounced for social than for non-social search (cf. **Figure**

**5.7**; social–non-social for average of fixations 2 to 10: ASD: 6.88 ± 4.08, amygdala: 18.57 ± 7.42, general control: 17.85 ± 2.85, NUS control: 12.80 ± 2.55), there was no significant interaction between target type and subject group ($F(3,26) = 2.07$, $p = 0.13$, $\eta^2 = 0.030$; also see **Table 5.5** for fixation-by-fixation analysis).

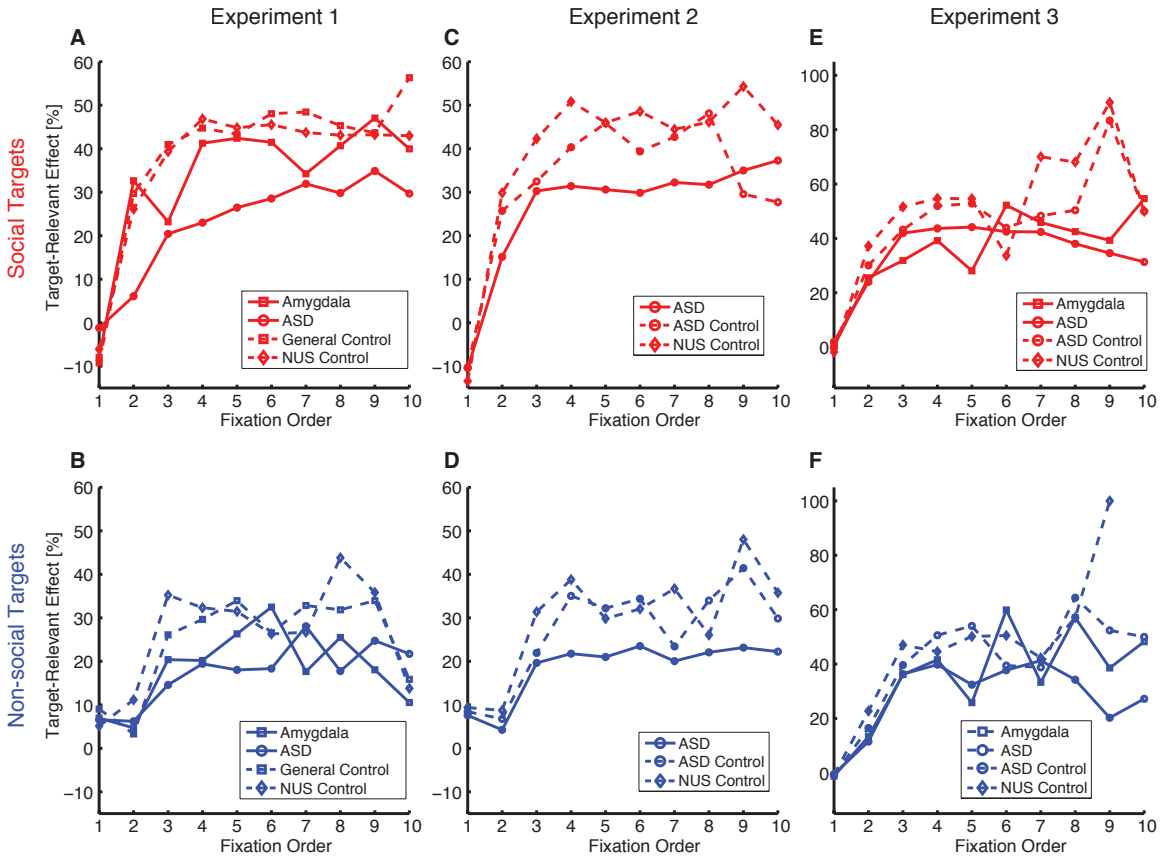


**Figure 5.7. Target-relevant effects.**

**(A–B)** Experiment 1. **(C–D)** Experiment 2. **(E–F)** Experiment 3. People with ASD had reduced attention towards social objects when searching for social targets **(A,C)**, an impairment that was less severe when searching for non-social targets **(B,D)** and with simpler search arrays **(E,F)**.

**Table 5.5. Statistical results for target-relevant effects.**

*All* is the average of fixation 2 to 10. *Early* is the average of fixation 2 to 5, and *Late* is the average of fixation 6 to 10. NaN: values not available (NUS controls did not have 10 fixations for non-social targets in Experiment 3).

| **Experiment 1** | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Two-way ANOVA (target type X subject group): All subject groups | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| Main effect of target type | | | | | | | | | | | | | |
| F-statistic F(1,26) | 29.6 | 34.1 | 7.38 | 20 | 13 | 16.4 | 11.4 | 1.92 | 3.27 | 6.97 | 55.4 | 43.9 | 26.3 |
| p-value | 1.0E-05 | 3.70E-06 | 0.0116 | 0.000135 | 0.00128 | 0.00041 | 0.00234 | 0.178 | 0.0824 | 0.015 | 6.63E-08 | 4.97E-07 | 2.38E-05 |
| Effect Size | 0.388 | 0.262 | 0.0613 | 0.150 | 0.124 | 0.183 | 0.112 | 0.0291 | 0.0517 | 0.174 | 0.264 | 0.199 | 0.213 |
| Main effect of subject group | | | | | | | | | | | | | |
| F-statistic F(3,26) | 0.731 | 3.73 | 7.57 | 4.95 | 3.86 | 1.91 | 0.94 | 2.20 | 0.525 | 0.595 | 4.76 | 6.79 | 2.38 |
| p-value | 0.543 | 0.0236 | 0.000851 | 0.00752 | 0.0207 | 0.152 | 0.435 | 0.112 | 0.669 | 0.625 | 0.00894 | 0.00157 | 0.0927 |
| Effect Size | 0.018 | 0.124 | 0.325 | 0.227 | 0.192 | 0.0924 | 0.0599 | 0.112 | 0.0318 | 0.0177 | 0.206 | 0.285 | 0.121 |
| Interaction | | | | | | | | | | | | | |
| F-statistic F(3,26) | 1.00 | 5.47 | 0.995 | 1.38 | 0.240 | 0.519 | 0.693 | 0.445 | 0.356 | 0.553 | 2.07 | 2.50 | 0.704 |
| p-value | 0.407 | 0.00475 | 0.411 | 0.270 | 0.867 | 0.673 | 0.565 | 0.723 | 0.785 | 0.651 | 0.129 | 0.0817 | 0.558 |
| Effect Size | 0.0394 | 0.126 | 0.0248 | 0.0311 | 0.00686 | 0.0173 | 0.0204 | 0.0203 | 0.0169 | 0.0414 | 0.0296 | 0.034 | 0.0171 |
| Two-way ANOVA (target type X subject group): ASD vs. General control | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| Main effect of target type | | | | | | | | | | | | | |
| F-statistic F(1,14) | 17.0 | 10.5 | 5.76 | 7.58 | 8.35 | 10.5 | 4.99 | 3.38 | 2.29 | 8.34 | 24.7 | 16.6 | 18.2 |
| p-value | 0.00103 | 0.00585 | 0.0309 | 0.0156 | 0.0119 | 0.00584 | 0.0423 | 0.0872 | 0.154 | 0.0127 | 0.000205 | 0.00113 | 0.000781 |
| Effect Size | 0.417 | 0.175 | 0.105 | 0.0798 | 0.0945 | 0.165 | 0.0576 | 0.0774 | 0.0728 | 0.226 | 0.209 | 0.150 | 0.171 |
| Main effect of subject group | | | | | | | | | | | | | |
| F-statistic F(1,14) | 1.76 | 5 | 9.55 | 6.44 | 10.5 | 3.77 | 1.4 | 2.92 | 1.7 | 2.23 | 8.97 | 11.1 | 4.56 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| p-value | 0.206 | 0.0422 | 0.00798 | 0.0237 | 0.0059 | 0.0725 | 0.256 | 0.109 | 0.214 | 0.159 | 0.00964 | 0.00489 | 0.0508 |
| Effect Size | 0.0191 | 0.109 | 0.251 | 0.234 | 0.320 | 0.126 | 0.0691 | 0.104 | 0.0595 | 0.0454 | 0.247 | 0.298 | 0.166 |
| Interaction | | | | | | | | | | | | | |
| F-statistic F(1,14) | 2.75 | 10.7 | 1.09 | 2.89 | 0.0232 | 1.34 | 1.81 | 0.0103 | 0.00109 | 4.1 | 4.86 | 5.5 | 2.32 |
| p-value | 0.119 | 0.00551 | 0.313 | 0.111 | 0.881 | 0.267 | 0.199 | 0.921 | 0.974 | 0.0638 | 0.0447 | 0.0342 | 0.150 |
| Effect Size | 0.0675 | 0.178 | 0.0199 | 0.0304 | 0.000263 | 0.0209 | 0.0210 | 0.000235 | 3.48E-05 | 0.111 | 0.0411 | 0.0498 | 0.0218 |
| Two-way ANOVA (target type X subject group): ASD vs. NUS Control | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| Main effect of target type | | | | | | | | | | | | | |
| F-statistic F(1,17) | 11.2 | 7.57 | 2.19 | 6.78 | 6.48 | 8.33 | 4.58 | 0.417 | 1.15 | 1.85 | 20.8 | 14.5 | 9.77 |
| p-value | 0.00388 | 0.0137 | 0.157 | 0.0185 | 0.0209 | 0.0103 | 0.0471 | 0.527 | 0.299 | 0.195 | 0.000275 | 0.00139 | 0.00616 |
| Effect Size | 0.269 | 0.105 | 0.0280 | 0.083 | 0.0946 | 0.139 | 0.103 | 0.00941 | 0.0272 | 0.0865 | 0.164 | 0.100 | 0.168 |
| Main effect of subject group | | | | | | | | | | | | | |
| F-statistic F(1,17) | 1.64 | 10.1 | 26.1 | 11.6 | 6.64 | 3.14 | 0.766 | 6.48 | 1.09 | 0.264 | 10.3 | 15.2 | 5.68 |
| p-value | 0.217 | 0.00545 | 8.75E-05 | 0.00335 | 0.0196 | 0.0942 | 0.394 | 0.0209 | 0.311 | 0.615 | 0.00509 | 0.00115 | 0.0290 |
| Effect Size | 0.0277 | 0.216 | 0.457 | 0.278 | 0.183 | 0.0884 | 0.0207 | 0.163 | 0.0343 | 0.00419 | 0.260 | 0.361 | 0.134 |
| Interaction | | | | | | | | | | | | | |
| F-statistic F(1,17) | 0.335 | 5.66 | 0.0687 | 2.02 | 0.300 | 0.695 | 1.47 | 0.762 | 0.0293 | 0.688 | 1.67 | 2.52 | 0.387 |
| p-value | 0.570 | 0.0294 | 0.796 | 0.173 | 0.591 | 0.416 | 0.242 | 0.395 | 0.866 | 0.421 | 0.213 | 0.131 | 0.542 |
| Effect Size | 0.00807 | 0.0787 | 0.000877 | 0.0247 | 0.00438 | 0.0116 | 0.033 | 0.0172 | 0.000694 | 0.0322 | 0.0132 | 0.0174 | 0.00666 |
| Two-way ANOVA (target type X subject group): Amygdala vs. General control | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| Main effect of target type | | | | | | | | | | | | | |
| F-statistic F(1,9) | 21.6 | 32.7 | 6.04 | 20.9 | 10.1 | 9.11 | 10.2 | 1.99 | 3.24 | 12.6 | 41.1 | 38.5 | 21.9 |
| p-value | 0.00121 | 0.000287 | 0.0363 | 0.00134 | 0.0112 | 0.0145 | 0.011 | 0.191 | 0.110 | 0.00745 | 0.000123 | 0.000158 | 0.00115 |
| Effect Size | 0.614 | 0.609 | 0.144 | 0.389 | 0.344 | 0.346 | 0.13 | 0.093 | 0.150 | 0.457 | 0.521 | 0.546 | 0.306 |
| Main effect of subject group | | | | | | | | | | | | | |
| F-statistic F(1,9) | 0.0201 | 0.128 | 2.16 | 1.07 | 1.26 | 0.00292 | 1.2 | 0.216 | 0.399 | 1.24 | 1.74 | 1.28 | 1.04 |
| p-value | 0.890 | 0.729 | 0.176 | 0.327 | 0.292 | 0.958 | 0.303 | 0.653 | 0.545 | 0.298 | 0.219 | 0.286 | 0.335 |
| Effect Size | 0.00028 | 0.00311 | 0.118 | 0.0462 | 0.0398 | 9.04E-05 | 0.0885 | 0.0114 | 0.0205 | 0.0328 | 0.0592 | 0.0407 | 0.0588 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Interaction | | | | | | | | | | | | | |
| F-statistic F(1,9) | 0.153 | 0.022 | 1.3 | 0.528 | 0.718 | 0.861 | 0.00812 | 0.0061 | 1.05 | 0.233 | 0.0129 | 0.00824 | 0.0169 |
| p-value | 0.705 | 0.885 | 0.284 | 0.486 | 0.419 | 0.378 | 0.930 | 0.939 | 0.335 | 0.642 | 0.912 | 0.930 | 0.900 |
| Effect Size | 0.00436 | 0.000409 | 0.0311 | 0.00983 | 0.0244 | 0.0327 | 0.000104 | 0.000285 | 0.0487 | 0.00845 | 0.000163 | 0.000117 | 0.000236 |
| One-Way ANOVA: Social Targets | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| F-statistic F(3,26) | 1.7259 | 6.9582 | 5.3581 | 6.5733 | 3.6054 | 1.8027 | 1.9806 | 1.1955 | 0.295 | 1.1938 | 5.0931 | 7.8157 | 2.8907 |
| p-value | 0.1863 | 0.0014 | 0.0052 | 0.0019 | 0.0266 | 0.1714 | 0.1416 | 0.3309 | 0.8286 | 0.3324 | 0.0066 | 7.03E-04 | 0.0545 |
| Effect Size | 0.1661 | 0.4453 | 0.382 | 0.4313 | 0.2938 | 0.1722 | 0.186 | 0.1212 | 0.0342 | 0.1253 | 0.3701 | 0.4742 | 0.2501 |
| One-Way ANOVA: Non-social Targets | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| F-statistic F(3,26) | 0.3116 | 1.0076 | 4.9305 | 1.6222 | 1.7633 | 0.7273 | 0.3568 | 1.4464 | 0.5923 | 0.1074 | 2.6043 | 3.2903 | 0.9206 |
| p-value | 0.8168 | 0.4052 | 0.0077 | 0.2084 | 0.1789 | 0.5449 | 0.7846 | 0.2521 | 0.6258 | 0.9549 | 0.0733 | 0.0364 | 0.4447 |
| Effect Size | 0.0347 | 0.1042 | 0.3626 | 0.1577 | 0.1691 | 0.0774 | 0.0395 | 0.143 | 0.0664 | 0.0144 | 0.2311 | 0.2752 | 0.096 |
| | | | | | | | | | | | | | |
| **Experiment 2** | | | | | | | | | | | | | |
| Two-way ANOVA (target type X subject group): All subject groups | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| Main effect of target type | | | | | | | | | | | | | |
| F-statistic F(1,29) | 93.9 | 28.2 | 19.9 | 14.2 | 12.1 | 7.64 | 12.4 | 9.04 | 0.646 | 2.33 | 31.3 | 39.2 | 15 |
| p-value | 1.34E-10 | 1.07E-05 | 0.000112 | 0.000748 | 0.00161 | 0.00982 | 0.00144 | 0.0054 | 0.428 | 0.139 | 4.91E-06 | 7.88E-07 | 0.000574 |
| Effect Size | 0.699 | 0.315 | 0.118 | 0.0847 | 0.134 | 0.0749 | 0.135 | 0.113 | 0.00570 | 0.0452 | 0.169 | 0.224 | 0.105 |
| Main effect of subject group | | | | | | | | | | | | | |
| F-statistic F(2,29) | 0.250 | 4.82 | 2.82 | 7.18 | 4.31 | 3.63 | 4.89 | 2.47 | 3.36 | 0.422 | 6.00 | 6.63 | 4.52 |
| p-value | 0.78 | 0.0156 | 0.0763 | 0.00293 | 0.0229 | 0.0392 | 0.0148 | 0.102 | 0.0487 | 0.660 | 0.00662 | 0.00426 | 0.0196 |
| Effect Size | 0.00131 | 0.084 | 0.116 | 0.244 | 0.124 | 0.124 | 0.135 | 0.0749 | 0.133 | 0.0134 | 0.195 | 0.189 | 0.162 |
| Interaction | | | | | | | | | | | | | |
| F-statistic F(2,29) | 0.56 | 1.13 | 0.00234 | 0.531 | 0.299 | 1.1 | 0.768 | 0.44 | 1.81 | 0.644 | 0.567 | 0.569 | 0.669 |
| p-value | 0.578 | 0.337 | 0.998 | 0.593 | 0.744 | 0.348 | 0.473 | 0.648 | 0.182 | 0.533 | 0.573 | 0.572 | 0.520 |
| Effect Size | 0.00832 | 0.0252 | 2.76E-05 | 0.00634 | 0.00661 | 0.0215 | 0.0167 | 0.0110 | 0.0319 | 0.025 | 0.00614 | 0.00651 | 0.00939 |
| Two-way ANOVA (target type X subject group): ASD vs. ASD Control | | | | | | | | | | | | | |

| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Main effect of target type | | | | | | | | | | | | | |
| F-statistic F(1,19) | 50 | 19.7 | 10.1 | 6.04 | 9.7 | 1.69 | 11.5 | 4.67 | 0.215 | 1.31 | 14.1 | 20.5 | 6.62 |
| p-value | 1.00E-06 | 0.000279 | 0.00487 | 0.0238 | 0.0057 | 0.209 | 0.00309 | 0.0437 | 0.648 | 0.266 | 0.00133 | 0.000231 | 0.0186 |
| Effect Size | 0.677 | 0.285 | 0.109 | 0.0729 | 0.102 | 0.0298 | 0.183 | 0.0792 | 0.00354 | 0.0291 | 0.138 | 0.188 | 0.0861 |
| Main effect of subject group | | | | | | | | | | | | | |
| F-statistic F(1,19) | 0.142 | 3.12 | 0.129 | 4.49 | 4.58 | 2.95 | 1.52 | 4.59 | 0.544 | 0.0133 | 2.7 | 3.61 | 1.85 |
| p-value | 0.710 | 0.0934 | 0.724 | 0.0475 | 0.0456 | 0.102 | 0.232 | 0.0453 | 0.47 | 0.909 | 0.117 | 0.0729 | 0.19 |
| Effect Size | 0.000485 | 0.0590 | 0.00461 | 0.132 | 0.135 | 0.0854 | 0.0374 | 0.116 | 0.0174 | 0.000364 | 0.0838 | 0.102 | 0.0581 |
| Interaction | | | | | | | | | | | | | |
| F-statistic F(1,19) | 0.0169 | 1.55 | 0.000241 | 0.419 | 0.298 | 0.0203 | 0.617 | 0.167 | 3.60 | 1.27 | 0.233 | 0.147 | 0.779 |
| p-value | 0.898 | 0.229 | 0.988 | 0.525 | 0.592 | 0.888 | 0.442 | 0.687 | 0.0731 | 0.274 | 0.635 | 0.706 | 0.388 |
| Effect Size | 0.00023 | 0.0223 | 2.59E-06 | 0.00506 | 0.00312 | 0.000357 | 0.00985 | 0.00284 | 0.0593 | 0.0281 | 0.00228 | 0.00134 | 0.0101 |
| Two-way ANOVA (target type X subject group): ASD vs. NUS Control | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| Main effect of target type | | | | | | | | | | | | | |
| F-statistic F(1,22) | 80.0 | 15.7 | 21.7 | 15.8 | 7.15 | 6.62 | 6.05 | 6.27 | 2.4 | 3.69 | 23.8 | 24.6 | 13.2 |
| p-value | 8.85E-09 | 0.000653 | 0.00012 | 0.000636 | 0.0138 | 0.0173 | 0.0222 | 0.0202 | 0.136 | 0.0698 | 7.16E-05 | 5.78E-05 | 0.00145 |
| Effect Size | 0.710 | 0.269 | 0.118 | 0.0966 | 0.120 | 0.0849 | 0.0853 | 0.102 | 0.0278 | 0.0990 | 0.161 | 0.193 | 0.115 |
| Main effect of subject group | | | | | | | | | | | | | |
| F-statistic F(1,22) | 0.167 | 9.88 | 5.04 | 12.4 | 6.11 | 5.84 | 8.73 | 1.85 | 6.12 | 0.711 | 9.43 | 10.9 | 7.27 |
| p-value | 0.687 | 0.00473 | 0.0352 | 0.00194 | 0.0217 | 0.0245 | 0.00732 | 0.187 | 0.0215 | 0.410 | 0.0056 | 0.00322 | 0.0132 |
| Effect Size | 0.000641 | 0.101 | 0.142 | 0.276 | 0.110 | 0.129 | 0.171 | 0.041 | 0.156 | 0.0141 | 0.206 | 0.208 | 0.172 |
| Interaction | | | | | | | | | | | | | |
| F-statistic F(1,22) | 1.11 | 1.69 | 0.004 | 0.194 | 0.479 | 1.4 | 0.28 | 0.804 | 0.213 | 0.00367 | 0.345 | 0.937 | 0.0427 |
| p-value | 0.303 | 0.207 | 0.950 | 0.664 | 0.496 | 0.249 | 0.602 | 0.379 | 0.649 | 0.952 | 0.563 | 0.344 | 0.838 |
| Effect Size | 0.00988 | 0.0288 | 2.17E-05 | 0.00118 | 0.00802 | 0.0180 | 0.00395 | 0.0131 | 0.00247 | 9.83E-05 | 0.00233 | 0.00735 | 0.000371 |
| One-Way ANOVA: Social Targets | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| F-statistic F(2,29) | 0.6669 | 3.731 | 2.1064 | 5.2707 | 2.8341 | 3.692 | 1.4271 | 2.5374 | 2.6888 | 1.392 | 3.5797 | 4.5291 | 2.7592 |
| p-value | 0.521 | 0.0361 | 0.1399 | 0.0112 | 0.0751 | 0.0373 | 0.2564 | 0.0965 | 0.0849 | 0.2653 | 0.0408 | 0.0194 | 0.0800 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Effect Size | 0.0440 | 0.2047 | 0.1268 | 0.2666 | 0.1635 | 0.2029 | 0.0896 | 0.1489 | 0.1564 | 0.0904 | 0.198 | 0.238 | 0.1599 |
| One-Way ANOVA: Non-social Targets | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| F-statistic F(2,29) | 0.2385 | 0.6854 | 2.2795 | 5.6748 | 2.1057 | 1.5259 | 9.5718 | 0.781 | 3.0718 | 0.5727 | 8.0094 | 5.8843 | 4.7523 |
| p-value | 0.7893 | 0.5119 | 0.1204 | 0.0083 | 0.140 | 0.2344 | 6.43E-04 | 0.4673 | 0.0617 | 0.5707 | 0.0017 | 0.0072 | 0.0164 |
| Effect Size | 0.0162 | 0.0451 | 0.1359 | 0.2813 | 0.1268 | 0.0952 | 0.3976 | 0.0511 | 0.1748 | 0.0407 | 0.3558 | 0.2887 | 0.2468 |
| | | | | | | | | | | | | | |
| **Experiment 3** | | | | | | | | | | | | | |
| Two-way ANOVA (target type X subject group): All subject groups | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| Main effect of target type | | | | | | | | | | | | | |
| F-statistic F(1,31) | 0.429 | 32.9 | 4.07 | 3.78 | 3.03 | 0.268 | 1.24 | 0.0262 | 1.18 | 0.305 | 6.50 | 20.9 | 0.297 |
| p-value | 0.517 | 2.59E-06 | 0.0524 | 0.0611 | 0.0916 | 0.609 | 0.277 | 0.873 | 0.301 | 0.596 | 0.0159 | 7.38E-05 | 0.590 |
| Effect Size | 0.0084 | 0.204 | 0.0168 | 0.0214 | 0.0196 | 0.00239 | 0.0179 | 0.000541 | 0.0554 | 0.00917 | 0.0432 | 0.065 | 0.00311 |
| Main effect of subject group | | | | | | | | | | | | | |
| F-statistic F(3,31) | 0.912 | 2.61 | 1.5 | 1.11 | 3.39 | 0.251 | 0.337 | 1.8 | 8.27 | 0.52 | 3.54 | 2.52 | 2.42 |
| p-value | 0.446 | 0.0693 | 0.234 | 0.360 | 0.0301 | 0.860 | 0.799 | 0.177 | 0.00367 | 0.680 | 0.0259 | 0.076 | 0.0855 |
| Effect Size | 0.0295 | 0.121 | 0.107 | 0.0762 | 0.189 | 0.0173 | 0.0250 | 0.115 | 0.279 | 0.112 | 0.189 | 0.163 | 0.131 |
| Interaction | | | | | | | | | | | | | |
| F-statistic F(3,31) | 0.371 | 0.0487 | 0.596 | 0.925 | 0.810 | 0.918 | 0.385 | 0.111 | 0.181 | 0.702 | 0.528 | 0.750 | 0.352 |
| p-value | 0.774 | 0.986 | 0.622 | 0.440 | 0.498 | 0.444 | 0.765 | 0.953 | 0.907 | 0.577 | 0.667 | 0.531 | 0.788 |
| Effect Size | 0.0218 | 0.000905 | 0.00741 | 0.0158 | 0.0157 | 0.0246 | 0.0167 | 0.00686 | 0.0255 | 0.0633 | 0.0105 | 0.00702 | 0.0111 |
| Two-way ANOVA (target type X subject group): ASD vs. ASD Control | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| Main effect of target type | | | | | | | | | | | | | |
| F-statistic F(1,19) | 1.45 | 14.7 | 3.39 | 0.992 | 3.19 | 0.319 | 0.101 | 0.00325 | 1.33 | 0.235 | 2.30 | 13.7 | 0.0445 |
| p-value | 0.243 | 0.00110 | 0.0811 | 0.332 | 0.0898 | 0.579 | 0.755 | 0.955 | 0.276 | 0.641 | 0.146 | 0.00152 | 0.835 |
| Effect Size | 0.0405 | 0.230 | 0.0286 | 0.00931 | 0.0297 | 0.00615 | 0.00224 | 0.000107 | 0.0945 | 0.00971 | 0.0371 | 0.0758 | 0.00122 |
| Main effect of subject group | | | | | | | | | | | | | |
| F-statistic F(1,19) | 0.812 | 1.74 | 0.143 | 2.5 | 4.15 | 0.0206 | 0.000245 | 2.06 | 3.12 | 0.00966 | 5.15 | 2.66 | 4.21 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| p-value | 0.379 | 0.203 | 0.709 | 0.131 | 0.0558 | 0.887 | 0.988 | 0.172 | 0.108 | 0.924 | 0.0352 | 0.120 | 0.0541 |
| Effect Size | 0.0173 | 0.0395 | 0.00607 | 0.0941 | 0.138 | 0.00068 | 8.91E-06 | 0.0606 | 0.0421 | 0.000689 | 0.139 | 0.0997 | 0.087 |
| Interaction | | | | | | | | | | | | | |
| F-statistic F(1,19) | 0.290 | 0.0256 | 0.144 | 0.177 | 2.70 | 0.000237 | 0.0624 | 0.163 | 0.242 | 2.13 | 0.134 | 1.15 | 0.0428 |
| p-value | 0.597 | 0.875 | 0.708 | 0.678 | 0.117 | 0.988 | 0.806 | 0.692 | 0.634 | 0.182 | 0.719 | 0.297 | 0.838 |
| Effect Size | 0.00808 | 0.000400 | 0.00121 | 0.00166 | 0.0251 | 4.57E-06 | 0.00139 | 0.00536 | 0.0172 | 0.0882 | 0.00216 | 0.00636 | 0.00117 |
| Two-way ANOVA (target type X subject group): ASD vs. NUS Control | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| Main effect of target type | | | | | | | | | | | | | |
| F-statistic F(1,22) | 0.209 | 23.4 | 4.70 | 4.54 | 3.74 | 0.938 | 1.02 | 0.00971 | 0.764 | NaN | 8.87 | 19.1 | 1.18 |
| p-value | 0.652 | 7.85E-05 | 0.0413 | 0.0445 | 0.0662 | 0.344 | 0.326 | 0.923 | 0.405 | NaN | 0.00694 | 0.000246 | 0.289 |
| Effect Size | 0.00622 | 0.178 | 0.0271 | 0.0398 | 0.0471 | 0.008 | 0.0226 | 0.000261 | 0.0424 | NaN | 0.0604 | 0.0877 | 0.0117 |
| Main effect of subject group | | | | | | | | | | | | | |
| F-statistic F(1,22) | 2.72 | 6.31 | 3.14 | 1.69 | 5.38 | 0.0101 | 1.1 | 4.56 | 24.7 | NaN | 8.25 | 5.02 | 5.67 |
| p-value | 0.113 | 0.0198 | 0.0903 | 0.207 | 0.0301 | 0.921 | 0.308 | 0.0485 | 0.000774 | NaN | 0.00886 | 0.0355 | 0.0268 |
| Effect Size | 0.0344 | 0.146 | 0.106 | 0.0541 | 0.131 | 0.000375 | 0.0349 | 0.126 | 0.332 | NaN | 0.215 | 0.151 | 0.165 |
| Interaction | | | | | | | | | | | | | |
| F-statistic F(1,22) | 0.857 | 0.125 | 0.049 | 0.969 | 0.702 | 4.34 | 1.28 | 0.0845 | 0.0764 | NaN | 0.359 | 7.54E-05 | 0.396 |
| p-value | 0.365 | 0.727 | 0.827 | 0.336 | 0.411 | 0.0497 | 0.273 | 0.775 | 0.788 | NaN | 0.555 | 0.993 | 0.536 |
| Effect Size | 0.0255 | 0.000952 | 0.000283 | 0.00849 | 0.00884 | 0.037 | 0.0283 | 0.00227 | 0.00424 | NaN | 0.00245 | 3.47E-07 | 0.00391 |
| Two-way ANOVA (target type X subject group): Amygdala vs. ASD Control | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| Main effect of target type | | | | | | | | | | | | | |
| F-statistic F(1,9) | 0.498 | 9.55 | 0.164 | 0.00920 | 0.00348 | 0.00523 | 0.269 | 0.312 | 0.387 | 10.6 | 0.228 | 2.51 | 0.0575 |
| p-value | 0.498 | 0.0129 | 0.695 | 0.926 | 0.954 | 0.944 | 0.620 | 0.600 | 0.597 | 0.190 | 0.645 | 0.148 | 0.816 |
| Effect Size | 0.0198 | 0.308 | 0.00256 | 0.000132 | 2.87E-05 | 0.000222 | 0.0121 | 0.0299 | 0.118 | 0.126 | 0.011 | 0.0232 | 0.00374 |
| Main effect of subject group | | | | | | | | | | | | | |
| F-statistic F(1,9) | 0.000244 | 0.502 | 0.615 | 1.41 | 3.97 | 0.417 | 0.00699 | 0.0527 | 0.343 | 0.316 | 1.31 | 2.60 | 0.27 |
| p-value | 0.988 | 0.496 | 0.453 | 0.265 | 0.0776 | 0.535 | 0.936 | 0.827 | 0.617 | 0.674 | 0.282 | 0.141 | 0.616 |
| Effect Size | 1.68E-05 | 0.0212 | 0.0538 | 0.118 | 0.283 | 0.0271 | 0.000670 | 0.00508 | 0.0237 | 0.178 | 0.0694 | 0.200 | 0.0118 |
| Interaction | | | | | | | | | | | | | |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| F-statistic F(1,9) | 0.0623 | 0.0214 | 0.962 | 0.243 | 0.145 | 0.115 | 0.041 | 0.0477 | 0.359 | 10.2 | 0.154 | 0.208 | 0.115 |
| p-value | 0.808 | 0.887 | 0.352 | 0.634 | 0.712 | 0.743 | 0.845 | 0.836 | 0.61 | 0.193 | 0.704 | 0.659 | 0.742 |
| Effect Size | 0.00248 | 0.00069 | 0.015 | 0.00348 | 0.00119 | 0.00487 | 0.00184 | 0.00457 | 0.110 | 0.121 | 0.00744 | 0.00192 | 0.0075 |
| One-Way ANOVA: Social Targets | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| F-statistic F(3,31) | 0.8863 | 2.0504 | 1.4601 | 1.5952 | 1.7811 | 0.3717 | 2.0223 | 1.2824 | 5.5773 | 0.2578 | 3.1216 | 2.184 | 2.4871 |
| p-value | 0.459 | 0.1272 | 0.2445 | 0.2105 | 0.1713 | 0.774 | 0.1345 | 0.303 | 0.0069 | 0.8543 | 0.0400 | 0.1098 | 0.0789 |
| Effect Size | 0.079 | 0.1656 | 0.1238 | 0.1337 | 0.147 | 0.0347 | 0.1835 | 0.1382 | 0.4817 | 0.0606 | 0.232 | 0.1745 | 0.194 |
| One-Way ANOVA: Non-social Targets | | | | | | | | | | | | | |
| Fixation Order | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | All | Early | Late |
| F-statistic F(3,31) | 0.0296 | 1.6797 | 1.2487 | 0.6742 | 3.7904 | 0.4411 | 0.0451 | 0.8797 | 3.8646 | 0.8791 | 2.2979 | 2.4435 | 1.166 |
| p-value | 0.993 | 0.1917 | 0.309 | 0.5744 | 0.0200 | 0.7253 | 0.987 | 0.4661 | 0.0332 | 0.4449 | 0.097 | 0.0827 | 0.339 |
| Effect Size | 0.0029 | 0.1398 | 0.1078 | 0.0613 | 0.2684 | 0.0422 | 0.0056 | 0.1029 | 0.453 | 0.1495 | 0.1819 | 0.1912 | 0.1044 |

We further compared people with ASD to general controls and found that social objects attracted more attention (two-way mixed ANOVA (target type X subject group); main effect of target type; average of fixations 2 to 10: social: 35.11 ± 3.52 (mean ± SEM), non-social: 22.74 ± 2.64; $F(1,14) = 24.7$, $p = 2.05\times10^{-4}$, $\eta^2 = 0.21$) and people with ASD had reduced target-relevant effects (ASD: 22.20 ± 3.30, general control: 35.64 ± 3.05; $F(1,14) = 8.97$, $p = 9.64\times10^{-3}$, $\eta^2 = 0.25$). Fixation-by-fixation analysis revealed that the impairment in people with ASD mainly came from initial fixations of their search (fixation 2 to 5: 16.78 ± 3.54; see **Table 5.5**). However, there was a weak interaction (average of fixations 2 to 10: $F(1,14) = 4.86$, $p = 0.045$, $\eta^2 = 0.041$), suggesting that compared to general controls, people with ASD were more impaired in orienting to socially relevant targets.

When analyzing target-relevant effects separately for social targets (**Figure 5.7A**) and non-social targets (**Figure 5.7B**), fixation-by-fixation analysis revealed that the target-relevant effect was reduced in people with ASD for social targets in the first few fixations (one-way ANOVA across subject groups, $p < 0.05$ for fixations 2 to 4; FDR corrected)

but not for non-social targets ($p > 0.05$ for all fixations, FDR corrected), further demonstrating a more severe impairment of people with ASD in social attention. Strikingly, there was no difference between people with ASD and controls for later fixations, showing that people with ASD could catch up gradually. Similar results were derived when comparing people with ASD to NUS controls, where we found a significant interaction between subject group and target type (social or non-social), with the impairment in the ASD group most pronounced again for the social case (**Table 5.5**).

We next compared amygdala lesion patients with general controls. Social targets still attracted more attention than non-social targets (two-way mixed ANOVA (target type X subject group); main effect of target type; average of fixations 2 to 10: social: 42.80 ± 1.89, non-social: 24.76 ± 3.37; $F(1,9) = 41.1$, $p = 1.23\times10^{-4}$, $\eta^2 = 0.52$), but there was no difference between amygdala patients and controls for the average of all fixations (main effect of subject group: amygdala: 28.81 ± 1.02, general control: 35.64 ± 3.05; $F(1,9) = 1.74$, $p = 0.22$, $\eta^2 = 0.059$), early fixations (amygdala: 26.38 ± 1.35, general control: 31.47 ± 2.62; $F(1,9) = 1.28$, $p = 0.29$, $\eta^2 = 0.041$), late fixations (amygdala: 30.76 ± 2.14, general control: 40.11 ± 5.39; $F(1,9) = 1.04$, $p = 0.33$, $\eta^2 = 0.059$), nor at any individual fixation ($p > 0.05$ for all fixations, FDR corrected). There was no interaction for all averages or at any fixations (see **Table 5.5**). Separate analysis within social targets (**Figure 5.7A**) and non-social targets (**Figure 5.7B**) found no reduced attention towards target-relevant objects, for neither social targets nor non-social targets (t-test with general controls: $p > 0.05$ for all fixations, FDR corrected). Further, we observed no difference between general controls and NUS controls ($p > 0.05$ for all fixations, FDR corrected).

The above results were robust to several factors. First, when controlling for the overall fewer numbers of fixations made by people with ASD (**Figure 5.6** and **Figure 5.8A,B**) with normalized fixation percentages, we obtained the same pattern of findings (**Figure 5.8C,D**). Likewise, our results were robust to the particular size of the ROI that defined each object (we tried several different sizes, all producing qualitatively the same results).

Finally, our analysis was based on target-present trials only; again, the target-relevant effects above all held when we analyzed target-absent catch trials only.

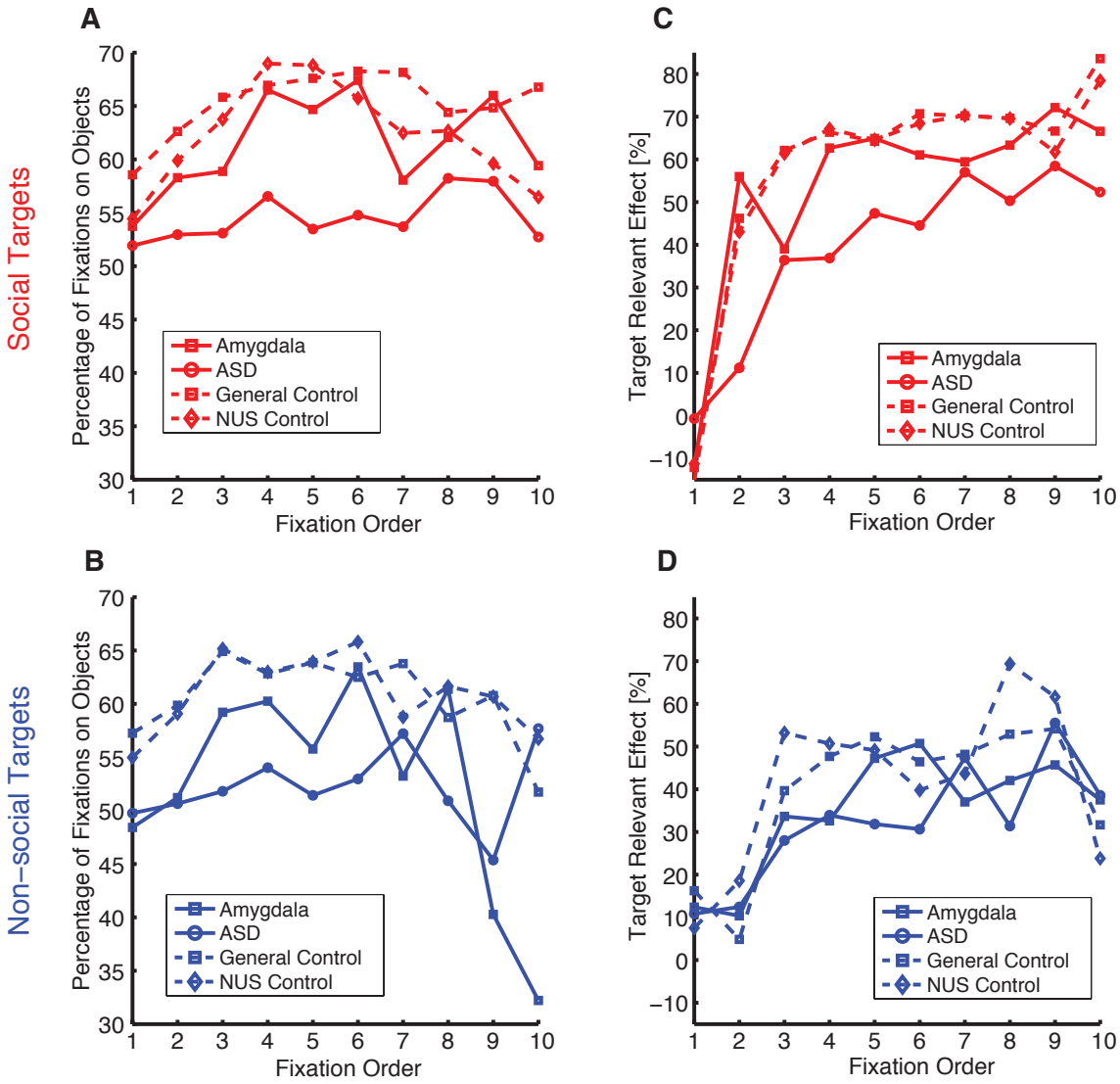


**Figure 5.8. Target relevant effect in Experiment 1 was preserved after normalization of fixation percentage.**

People with ASD had reduced percentage of fixations on objects, for both **(A)** social targets and **(B)** non-social targets. However, people with ASD still had reduced gazes

towards social objects when searching for social targets **(C)**, but they were not different from controls when searching for non-social targets **(D)**.

In conclusion, we found people with ASD did not orient towards target-relevant objects as rapidly as controls, an abnormality that was present for all stimuli but most pronounced for social stimuli, and furthermore, that this impairment was not evident in patients with amygdala lesions.

### *5.5.4 The attentional deficit in ASD could not be explained by low-level visual properties of the stimuli*

In Experiment 1, we observed reduced rapid orientation towards target-relevant objects in people with ASD, especially for social targets. To check whether this might be due to low-level saliency differences, we conducted Experiment 2 in which low-level properties of social and non-social objects were equalized within each search array (**Figure 5.2A-C**).

Target-relevant effects were replicated in Experiment 2. All subjects showed rapid and sustained target-relevant effects, for both social targets (**Figure 5.9** upper row) and non-social targets (**Figure 5.9** lower row). Even with equal low-level saliency, social targets still featured greater target-relevant effects (two-way mixed ANOVA (target type X subject group); main effect of target type; average of fixations 2 to 10: social: 37.19 ± 2.62, non-social: 26.18 ± 1.64; $F(1,29) = 31.3$, $p = 4.91\times10^{-6}$, $\eta^2 = 0.17$; see **Table 5.5** for statistics) and for both early fixations (social: 34.45 ± 2.47, non-social: 22.10 ± 1.56; $F(1,29) = 39.2$, $p = 7.88\times10^{-7}$, $\eta^2 = 0.22$) and late fixations (social: 39.43 ± 2.94, non-social: 29.49 ± 2.22; $F(1,29) = 15.0$, $p = 5.74\times10^{-4}$, $\eta^2 = 0.10$), showing persistent stronger attention towards social objects than non-social objects. Consistent with Experiment 1, the stronger social attention persisted through the 8$^{th}$ fixation. Still, people

with ASD had reduced overall target-relevant effects (main effect of subject group; average of fixations 2 to 10 collapsing social and non-social: ASD: 25.07 ± 3.55, NUS control: 38.68 ± 2.36, ASD control: 32.82 ± 1.51; $F(2,29) = 6.00$, $p = 6.62\times10^{-3}$, $\eta^2 = 0.20$) and for both early fixations (ASD: 21.76 ± 3.14, NUS control: 34.71 ± 2.07, ASD control: 30.03 ± 2.09; $F(2,29) = 6.63$, $p = 4.26\times10^{-3}$, $\eta^2 = 0.19$) and late fixations (ASD: 27.71 ± 4.01, NUS control: 42.00 ± 3.26, ASD control: 35.05 ± 2.06; $F(2,29) = 4.52$, $p = 0.020$, $\eta^2 = 0.16$). Comparing people with ASD to ASD controls alone revealed a marginally significant reduction of overall target-relevant effect during early fixations (ASD: 21.76 ± 3.14, ASD control: 30.03 ± 2.09; $F(1,19) = 3.61$, $p = 0.073$, $\eta^2 = 0.10$; see **Table 5.5**). Comparing people with ASD to NUS controls showed similar results and revealed a significant reduction of overall target-relevant effects for all fixations, early fixations and late fixations (**Table 5.5**). Separate analysis within social targets (**Figure 5.7C**) and non-social targets (**Figure 5.7D**) showed that the deficit mainly came from social targets (see **Table 5.5**), replicating Experiment 1.

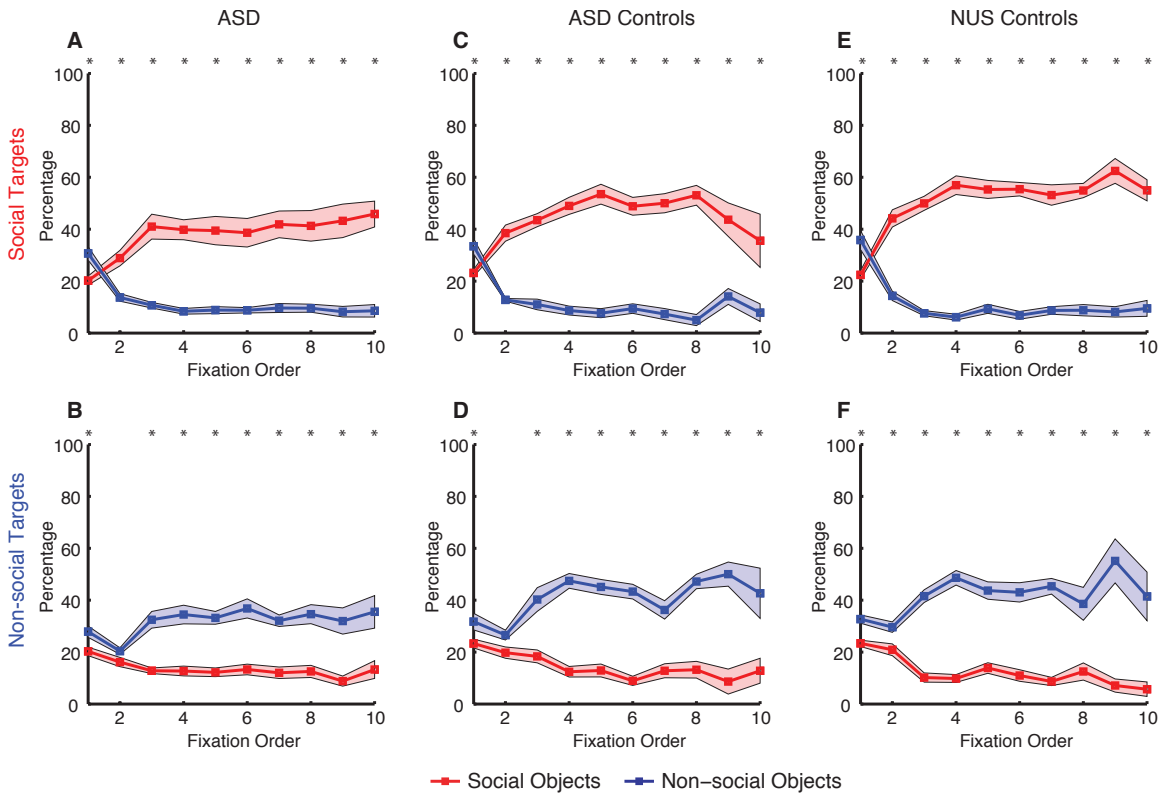


**Figure 5.9. In Experiment 2, all subjects looked at target congruent objects in a fast and sustained manner.**

**(A–B)** People with ASD. **(C–D)** ASD controls. **(E–F)** NUS controls. Red: social objects. Blue: non-social objects. Upper row **(A,C,E)**: when searching for social targets. Lower row **(B,D,F)**: when searching for non-social targets. Asterisk indicates significant difference between target congruent objects and target incongruent objects (two-tailed paired t-test: $p < 0.05$, FDR corrected). Shaded area denotes ± SEM.

Notably, no difference was found between Experiment 1 and Experiment 2 at all fixations (excluding the very first fixation) for people with ASD (unpaired two-tailed t-test: $p > 0.05$, FDR corrected) or NUS controls (paired two-tailed t-test: $p > 0.05$, FDR corrected), for both social targets and non-social targets.

In conclusion, Experiment 2 replicated the findings of Experiment 1 and thus corroborated our claim of reduced rapid orientation to target-relevant objects, especially when these were social, in people with ASD. Importantly, Experiment 2 demonstrated that the findings in Experiment 1 were not due to the low-level stimulus properties.

### *5.5.5 The attentional deficit in ASD is more severe with high task demands*

Experiment 1 and 2 showed that people with ASD, but not with amygdala lesions, have reduced attention to target-relevant objects. Do these effects depend on cognitive load? To test this hypothesis, we further designed simpler arrays with fewer objects to make the search easier. We still equalized low-level saliency, distance to center and size for these simpler search arrays.

As in Experiment 1 and 2, all subjects oriented to social objects rapidly and kept on searching within social objects if the target was social (**Figure 5.10** upper row) and oriented to non-social objects if the target was non-social (**Figure 5.10** lower row), showing rapid and sustained target-relevant effects for both social targets and non-social targets. In contrast to Experiment 1 and 2, with fewer objects in the search array, the difference between social target-relevant effects and non-social target-relevant effects became very small (social–non-social, Experiment 1: 13.15 ± 1.82, Experiment 2: 11.01 ± 1.94, Experiment 3: 6.30 ± 2.42; two-way mixed ANOVA (target type X subject group); main effect of target type; average of fixations 2 to 10: $F(1,31) = 6.50$, $p = 0.016$, $\eta^2 = 0.043$), and no difference was found at the single fixation level (see **Table 5.5** for statistics). The deficit of target-relevant orientation in people with ASD also became very small (main effect of subject group; average of fixations 2 to 10: ASD: 35.55 ± 3.32, amygdala: 39.38 ± 6.88, NUS control: 50.75 ± 4.21, ASD control: 46.53 ± 2.94; $F(3,31) = 3.54$, $p = 0.026$, $\eta^2 = 0.19$; only the 2$^{nd}$ fixation showed a difference) and there was no interaction.

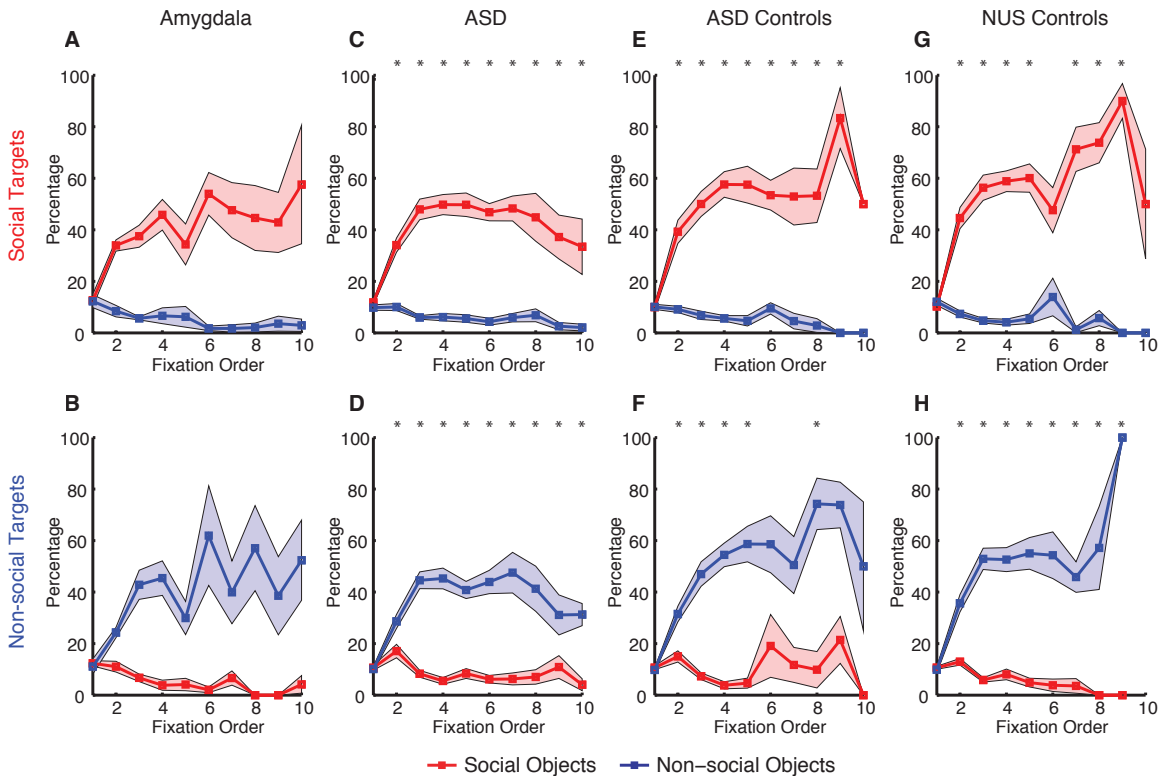


**Figure 5.10. In Experiment 3, all subjects looked at target congruent objects in a fast and sustained manner.**

**(A–B)** Amygdala patients. **(C–D)** People with ASD. **(E–F)** ASD controls. **(G–H)** NUS controls. Red: social objects. Blue: non-social objects. Upper row **(A,C,E,G)**: when searching for social targets. Lower row **(B,D,F,H)**: when searching for non-social targets. Asterisk indicates significant difference between target congruent objects and target incongruent objects (two-tailed paired t-test: $p < 0.05$, FDR corrected). Shaded area denotes ± SEM.

Comparing people with ASD and ASD controls also revealed a small but significant difference in target-relevant effects (average of fixations 2 to 10: ASD: 35.55 ± 3.32, ASD control: 46.53 ± 2.94; $F(1,19) = 5.15$, $p = 0.035$, $\eta^2 = 0.14$), and there was no fixation-by-fixation difference (**Table 5.5**). Further, consistent with Experiment 1 and 2, we found no difference in target-relevant effects between amygdala patients and ASD

controls, or between amygdala patients and NUS controls, for the average of all fixations, nor at each fixation ($p > 0.05$ for all fixations, FDR corrected; **Table 5.5**). Amygdala lesion patients had similar target-relevant effects as people with ASD at all fixations (**Table 5.5**).

Separate analysis within social targets (**Figure 5.7E**) showed that the target-relevant effect was not reduced in people with ASD for social targets – we observed no difference across all subject groups (ASD: 38.82 ± 3.96, amygdala: 38.94 ± 6.58, NUS control: 55.65 ± 4.57, ASD control: 48.43 ± 4.71; one-way ANOVA, $p > 0.05$ for all fixations, FDR corrected), and when comparing people with ASD and ASD controls, we observed no difference in target-relevant effects ($p > 0.05$ for all fixations, FDR corrected). Still, amygdala lesion patients did not show a different target-relevant effect compared to ASD controls, NUS controls, or people with ASD (**Table 5.5**). Similarly, separate analysis within non-social targets (**Figure 5.7F**) showed no difference across all subject groups (ASD: 32.28 ± 3.82, amygdala: 39.81 ± 7.68, NUS control: 45.86 ± 4.52, ASD control: 44.62 ± 4.39; $p > 0.05$ for all fixations, FDR corrected). There was no difference for people with ASD vs. ASD controls, amygdala patients vs. people with ASD, amygdala patients vs. ASD controls, nor ASD controls vs. NUS controls (for all comparisons: $p > 0.05$ for all fixations, FDR corrected; see **Table 5.5**).

In conclusion, we were able to find impaired attention to target-relevant stimuli in ASD only for the larger search array, but not for the smaller search array of Experiment 3. Likely explanation for the lack of an effect in Experiment 3 is reduced cognitive load.

### *5.5.6 Missing detection of targets was not prominent in amygdala lesion patients*

In some trials, targets failed to be detected even if the subject looked at the target object in the array. We further explored this mechanism by computing the percentage of trials having ‘misses’, which were defined as fixations that landed on the target even though

the target was not detected. We excluded the last 3 fixations landing on the target for misses since they corresponded to target detection.

**Figure 5.11** summarizes the percentage of trials with misses across subject groups. In Experiment 1 (**Figure 5.11A**), no difference was found between social and non-social targets (two-way repeated ANOVA (target type X subject group); main effect of target type: $F(1,26) = 0.28$, $p = 0.60$, effect size $\eta^2 = 8.90\times10^{-4}$) nor interaction ($F(3,26) = 0.50$, $p = 0.68$, effect size $\eta^2 = 0.0049$). However, NUS controls had significantly fewer misses (main effect of subject group: $F(3,26) = 5.45$, $p = 0.0048$, $\eta^2 = 0.35$; t-test against general controls: $t(17) = 3.44$, $p = 0.0032$, effect size in Hedges's g (standardized mean difference): $g = 1.53$ for social targets, and $t(17) = 2.21$, $p = 0.041$, $g = 0.98$ for non-social targets), which was likely due to the faster RT (**Figure 5.4B**; see **Discussion**). But compared to general controls, neither people with ASD ($t(14) = 0.59$, $p = 0.56$, $g = 0.28$ for social targets, and $t(14) = 0.53$, $p = 0.60$, $g = 0.25$ for non-social targets) nor amygdala lesion patients ($t(9) = 1.15$, $p = 0.27$, $g = 0.71$ for social targets, and $t(9) = 0.51$, $p = 0.62$, $g = 0.32$ for non-social targets) had more misses, suggesting that the amygdala is not essential for preferential coding of biologically relevant stimuli into conscious perception in this visual search.

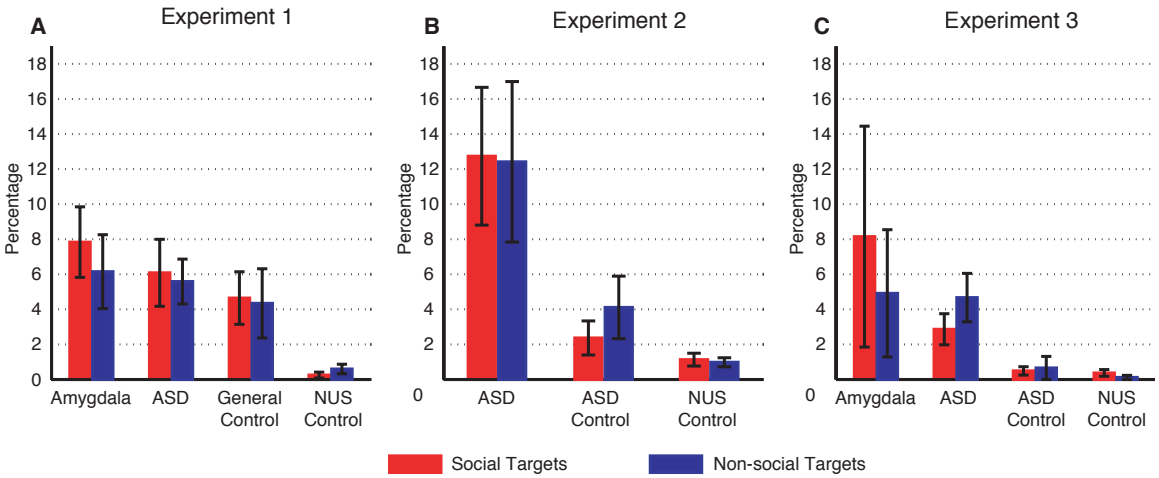


**Figure 5.11. Percentage of trials with misses.**

**(A)** Experiment 1. **(B)** Experiment 2. **(C)** Experiment 3. Red: social objects. Blue: non-social objects. Error bars denote one SEM of the mean.

We repeated the analysis by excluding the last 2 fixations landing on the target for misses and we derived qualitatively the same results.

Similarly, in Experiment 2 (**Figure 5.11B**), no difference was found between social and non-social targets (main effect of target type: $F(1,29) = 0.10$, $p = 0.75$, $\eta^2 = 1.34\times10^{-4}$) nor interaction ($F(2,29) = 0.60$, $p = 0.56$, $\eta^2 = 0.0015$), but NUS controls had significantly fewer misses than people with ASD and ASD controls (main effect of subject group: $F(2,29) = 4.57$, $p = 0.019$, $\eta^2 = 0.23$). People with ASD had comparable misses to those seen in ASD controls (unpaired t-test: $t(19) = 2.03$, $p = 0.057$, $g = 0.87$ for social targets and $t(19) = 1.37$, $p = 0.19$, $g = 0.59$ for non-social targets). Notably, with an independent sample of people with ASD, Experiment 2 had comparable percentages of trials with misses as Experiment 1 (unpaired t-test: $t(19) = -1.23$, $p = 0.23$, $g = -0.53$).

With an easier task in Experiment 3 (**Figure 5.11C**), we found the percentage of trials with misses decreased compared to Experiment 2 (both experiments had equal saliency between social and non-social objects; two-way ANOVA (experiment X subject group (ASD, ASD controls and NUS controls)); main effect of experiment: $F(1,58) = 4.99$, $p = 0.029$, $\eta^2 = 0.060$; paired t-test for people with ASD: $t(12) = 2.76$, $p = 0.017$, $g = 0.78$; paired t-test for ASD controls: $t(7) = 1.99$, $p = 0.087$, $g = 0.90$; paired t-test for NUS controls: $t(10) = 3.28$, $p = 0.0082$, $g = 1.15$), consistent with the idea that the percentage of misses is a function of the task difficulty (Rutishauser and Koch, 2007). We found no difference between social and non-social targets (two-way repeated ANOVA (target type X subject group); main effect of target type: $F(1,31) = 0.77$, $p = 0.39$, $\eta^2 = 0.0020$), suggesting that social and non-social targets had equal strength to be encoded into consciousness. However, ASD controls and NUS controls had significantly fewer misses (main effect of subject group: $F(3,31) = 4.57$, $p = 0.0092$, $\eta^2 = 0.27$), which was likely

due to the faster RT (**Figure 5.4F**; see **Discussion**). People with ASD had more misses than did ASD controls (t-test, $t(19) = 2.05$, $p = 0.054$, $g = 0.89$ for social and $t(19) = 2.17$, $p = 0.043$, $g = 0.94$ for non-social), but had similar numbers of misses as did amygdala lesion patients ($t(14) = 1.62$, $p = 0.13$, $g = 0.98$ for social and $t(14) = 0.073$, $p = 0.94$, $g = 0.044$ for non-social). ASD controls had similar number of misses as did NUS controls ($t(17) = 0.39$, $p = 0.70$, $g = 0.17$ for social and $t(17) = 0.94$, $p = 0.36$, $g = 0.42$ for non-social).

We lastly performed a subject-by-subject correlation analysis to confirm that the percentage of misses is a function of task difficulty. Task difficulty is typically measured by the time required to find the target (Treisman, 1988, 1998, Wolfe, 1998). In Experiment 1, there was strong subject-by-subject correlation between RT and the percentage of misses (Pearson correlation; all subjects: $r = 0.72$, $p = 8.31\times10^{-6}$; amygdala lesion patients: $r = 0.97$, $p = 0.15$; people with ASD: $r = 0.67$, $p = 0.068$; general controls: $r = 0.57$, $p = 0.14$; NUS controls: $r = 0.22$, $p = 0.51$). Strong correlations were observed in Experiment 2 (all subjects: $r = 0.84$, $p = 2.29\times10^{-9}$; people with ASD: $r = 0.89$, $p = 4.53\times10^{-5}$; ASD controls: $r = 0.69$, $p = 0.060$; NUS controls: $r = 0.87$, $p = 5.52\times10^{-4}$) and Experiment 3 (all subjects: $r = 0.76$, $p = 9.60\times10^{-8}$; amygdala lesion patients: $r = 0.90$, $p = 0.29$; people with ASD: $r = 0.89$, $p = 5.58\times10^{-5}$; ASD controls: $r = 0.89$, $p = 0.0034$; NUS controls: $r = 0.23$, $p = 0.50$) as well. These results showed that the percentage of misses is a function of task difficulty.

## 5.6 Discussion

In this study we found that people with ASD had reduced attention to target-relevant objects in visual search. Bilateral lesions of the amygdala did not result in a similar deficit. The impairment seemed most pronounced for social targets, although there was a deficit for non-social targets as well. The effect was not attributable to low-level properties of the stimuli. With arrays containing a reduced number of objects, we found a

weaker deficit. Overall, we revealed a search-dependent attentional deficit in people with ASD that was dependent on task demands.

Our findings suggest some clear future directions, as well as points of contact with prior findings and theories of ASD. With respect to future directions, there are in our view three core extensions of our study that would be important to undertake, aside from sheer replication. The first is replication together with generalization: that is, replicate our finding in a sample of people with ASD who are younger, and/or lower functioning, and/ or have more substantial comorbidity. This direction would be perhaps the most important from a clinical perspective. The second extension would be to broaden the difficulty of the search task. It is worth noting that (a) we only observed clear deficits in the ASD group for our larger search array (24 items; Experiments 1 and 2), but not for the smaller array (12 items; Experiment 3); and (b) all groups were close to ceiling in overall performance accuracy. Would one find a much larger deficit if more severe time constraints were imposed, or if arrays larger than 24 items were used? This might substantially increase the sensitivity of the task to detect abnormalities in ASD. The third extension of our study would be to probe in more detail the neural substrates of the effect, thus shedding light on the neurological basis of impaired social attention in ASD. The fact that we found no impairment in patients with amygdala lesions suggests that the amygdala is not essential here, but this of course does not rule out the possibility that the amygdala nonetheless plays a role in brains without amygdala lesions, including people with ASD. Translating our task into an fMRI study would thus be an informative future direction.

With respect to points of contact with the related literature in autism research, we take up the following issues in more detail below: relation to studies of visual search in ASD, and the connection with the amygdala. We also discuss missing detection of targets and task difficulty.

*5.6.1 Visual search in autism*

In a typical visual search task, an observer looks for a target item among an array of distractor items and responds by indicating whether a target is present or absent. In "classic guidance", attention is guided towards likely targets by a limited set of stimulus attributes such as color and size (Wolfe and Horowitz, 2004, Wolfe, 2012). Several studies have suggested superior visual search skills in individuals with ASD (Plaisted et al., 1998, O'Riordan and Plaisted, 2001, O'Riordan et al., 2001, O'Riordan, 2004, Kemner et al., 2008), particularly in relatively difficult tasks. Among various efforts to explain the differences, O'Riordan and Plaisted (O'Riordan and Plaisted, 2001) proposed two processing differences that could potentially explain the performance advantage: (1) enhanced memory for distractor locations already inspected, and (2) enhanced ability to discriminate between target and distractor stimulus features. Later, Joseph *et al.* (Joseph et al., 2009) argued that the superiority is due to the anomalously enhanced perception of stimulus features.

While most studies of visual search in autism focused on low-level features and inanimate stimuli (e.g., letters and shapes) (Plaisted et al., 1998, O'Riordan and Plaisted, 2001, O'Riordan et al., 2001, O'Riordan, 2004, Manjaly et al., 2007, Kemner et al., 2008), far fewer studies have examined complex images and social stimuli. Some studies employed visual search to investigate recognition abilities of facial expressions in children with ASD and found that faces with certain emotions are detected faster than others (Farran et al., 2011, Rosset et al., 2011). However, when compared with age-matched controls, no significant differences were found anymore.

Semantic-level features like faces can be considerably more potent than low-level cues to attract gaze to complex stimuli (Cerf et al., 2009, Judd et al., 2009, Zhao and Koch, 2011, 2012, Xu et al., 2014). In this study, we not only included social stimuli, but instead of isolated facial emotions, used natural social (faces and people with various emotions and poses) and non-social (e.g., furniture, toys and food) objects. In Experiment 2, we equalized the low-level saliency, object size and location of items, thus helping to isolate

effect to the semantic level. Our results suggest that reduced target-congruent attention in people with ASD is mostly restricted to the social domain and the semantic level.

Taken together, our findings and the prior literature then suggest that there may be two types of effects that distinguish visual search in people with ASD. One effect is that search is more efficient when it is based on low-level features and does not involve social content. A second effect is that search is less efficient when it is based on semantic-level features, and perhaps in particular when it involves social content. Respectively, these two putative effects bear some similarity to the two core aspects of the ASD diagnosis: augmented interests and focus on certain non-social patterns of stimuli and/or behavior; and diminished interest and focus on social communicative aspects.

### *5.6.2 Missing detection of targets and task difficulty*

Subjects could look at the target during search without detecting it, failures of attention despite fixation that occurred surprisingly frequently in our task, especially with increased task load (Experiments 1 and 2). Task difficulty is typically measured by the time required to find the target and the RT correlates with the size of the search array (Treisman, 1988, 1998, Wolfe, 1998). Thus, a search task is more difficult than another, or more difficult to one subject than another, if more time is required to find the target. Consistent with previous findings (Rutishauser and Koch, 2007), we found in our experiments that the percentage of misses is a function of the task difficulty—when RT is shorter, the percentage of missed detections is lower, as shown by a strong correlation within each subject group, as well as pronounced differences between subject groups such that NUS controls who had fastest RT also showed the smallest percentage of misses. Furthermore, in the simpler arrays (Experiment 3), not only were the targets easier to detect (shown by a significantly shorter RT), but also the percentage of missed detections was lower. The missed detections might be explained by a capacity limitation (Rutishauser and Koch, 2007). With greater task difficulty, the target object might not

effectively be reported as it failed to emerge into "access consciousness", a failure to transfer from iconic to working memory.

### *5.6.3 The amygdala and saliency*

A distinctive aspect of our studies was the direct comparison between subjects with ASD, and rare patients with bilateral amygdala lesions. The human amygdala has been quite broadly implicated in processing emotionally salient and socially relevant stimuli (Kling and Brothers, 1992, Adolphs, 2010). Studies of a patient with bilateral amygdala lesions demonstrated a selective impairment in recognizing fearful faces (Adolphs et al., 1994), congruent with early neuroimaging studies (Morris et al., 1996).

Recently, however, the amygdala has been proposed to respond to a broader spectrum of social attributes such as facial emotions in general (Fitzgerald et al., 2006) and regulating a person's personal space (Kennedy et al., 2009). Electrophysiological recordings in monkeys (Rolls, 1984, Leonard et al., 1985) and humans (Kreiman et al., 2000, Rutishauser et al., 2011) have found single neurons that respond not only to faces, but also to face identities, facial expressions and gaze directions (Gothard et al., 2007, Hoffman et al., 2007). Further, the amygdala processes more abstract attributes such as stimulus unpredictability (Herry et al., 2007). Amygdala lesions result in an absence or reduction of fixations on novel objects observed in monkeys (Bagshaw et al., 1972). It has also been shown that the amygdala mediates emotion-enhanced vividness (Todd et al., 2012) and responds more to animate entities compared to inanimate ones (Mormann et al., 2011, Yang et al., 2012b). Overall, the amygdala might act as a detector of perceptual saliency and biological relevance (Sander et al., 2005, Adolphs, 2008)—a reasonable substrate also for the altered preferences evident in people with ASD.

Our search arrays contained people and faces with various identities, expressions and gaze directions, but our data did not find any impairments in the three amygdala patients in deploying attention to target-relevant objects, either for social or non-social targets.

While our findings show that the amygdala cannot be essential in our task, we acknowledge that we are limited by statistical power given our small subject sample. It is also worth noting that compensatory circuits might account for the intact social attention in amygdala lesion patients (Becker et al., 2012) and a recent finding (Wang et al., 2014a) has also shown that amygdala lesion patients have intact preferred attention towards animals, even though these findings would not be expected on the basis of neuronal responses observed in the amygdala to animals (Mormann et al., 2011). Our finding is consistent also with preserved attentional capture by emotional stimuli and intact emotion-guided visual search in patients with acute amygdala lesions due to neurosurgical resection (Piech et al., 2010, Piech et al., 2011). Taken together, there are now numerous examples of a discrepancy between engagement of the amygdala (e.g., in functional neuroimaging studies) in tasks for which there is no obvious corresponding behavioral impairment when the amygdala is lesioned. This of course poses some challenges also for how to view the possible role of the amygdala in ASD, a final topic to which we turn next.

### *5.6.4 Amygdala theory of autism*

The abnormal facial scanning patterns generally reported in people with ASD (Adolphs et al., 2001, Klin et al., 2002, Pelphrey et al., 2002, Neumann et al., 2006, Spezio et al., 2007a, Spezio et al., 2007b, Kliemann et al., 2010) may plausibly be related to amygdala dysfunction (Baron-Cohen et al., 2000). This hypothesis is supported by rather similar patterns of deficits seen in patients with amygdala damage, who fail to fixate on the eyes in faces (Adolphs et al., 2005), single neuron recordings in the human amygdala showing weaker response to eyes in people with ASD (Rutishauser et al., 2013), as well as neuroimaging studies showing that amygdala activation is specifically enhanced for fearful faces when saccading from the mouth to the eye regions (Gamer and Büchel, 2009). This amygdala-mediated orientation towards eyes seen in BOLD-fMRI is dysfunctional in ASD (Kliemann et al., 2012). Activation in the amygdala has also been

reported to be correlated with the time spent fixating the eyes in ASD (Dalton et al., 2005). The idea of amygdala abnormalities in autism is supported by a substantial literature showing structural abnormalities (Bauman and Kemper, 1985, Schumann et al., 2004, Schumann and Amaral, 2006, Amaral et al., 2008, Ecker et al., 2012) and atypical activation (Gotts et al., 2012, Philip et al., 2012) in the amygdala in ASD.

While actual amygdala lesions did not result in search-related attentional deficits in our tasks, it is important to keep in mind that people with ASD of course do not have amygdala lesions. It is thus still conceivable that more subtle malfunction (including hyperactivation) of the amygdala contributes to ASD, even though a bona fide lesion of the amygdala has no effect that bears similarity to ASD (see also (Paul et al., 2010)). Finally, autism spectrum disorders are well known to be highly heterogeneous at the biological and behavioral levels, and it is likely that there will be no single genetic or cognitive cause for the diverse symptoms defining autism (Happe et al., 2006). No unanimously endorsed hypothesis for a primary deficit has emerged that can plausibly account for the full triad of social, communicative and rigid/repetitive difficulties (Happe, 2003). Nonetheless, our present findings argue for at least one further feature at the cognitive level that can be used to describe ASD: an inability to use semantic-level task demands, especially with high cognitive load and especially for social stimuli, in order to efficiently guide attention selection during visual search. As we noted at the beginning of our **Discussion**, it will be important to extend these studies to additional measures in the future, notably including neuroimaging studies of people with ASD during visual search.

### 5.7 Conclusion

While a sizable literature in ASD has investigated search for simple, non-social objects (shapes and letters, etc.) and only manipulated low-level attributes of the stimuli, far fewer studies have examined visual search with social stimuli. In this study, we used a visual search protocol with well-validated social stimuli. We observed reliable attentional

deficits in people with ASD, especially social attention. Our findings were further corroborated by the following evidence and manipulations: (1) We replicated the attentional deficit in ASD in an independent sample of ASD subjects. (2) We conducted a control experiment that equated the stimuli in the search array for low-level visual properties and ruled out the potential influence from low-level features. (3) Importantly, we performed the identical task on amygdala lesion patients, thus enabling a direct comparison. Taken together, our study has tested a key hypothesized function of the amygdala in autism, and offered both a clear test of this hypothesis, and a convincing result—a reliable attentional deficit in ASD, which does not depend on low-level properties, nor the amygdala, but on task demands. Our present findings have further advanced the understanding of how the brain processes socially salient stimuli and argue for at least one further well-characterized deficit of social attention in people with ASD.

## 5.8 Acknowledgements

We thank Jed Elison for providing the original stimuli, Ty Basinger for creating some of the stimuli, Lynn Paul for psychological assessments, and Mike Tyszka for providing the anatomical scans of the lesion patients. This research was supported by grants from NSF, the Pfeiffer Family Foundation, the Simons Foundation, an NIMH Conte Center, and Singapore MoE AcRF Tier 1. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

# Chapter VI: Future Directions

My thesis has investigated a few socially salient cues, including faces and people, using different psychological designs and neuroscience techniques. There are still remaining questions for future studies. Some of these questions have been put under investigation, and I will outline the questions and some preliminary findings in this section.

## 6.1 Computational modeling of saliency

In Chapter III, we used a change detection task to compare three animate cues (animals, people and head direction) to two inanimate cues (plants and objects), and in Chapter V, we used social (faces and people with different postures, emotions, ages, and genders) and a variety of non-social objects (e.g., electronics, food, utensils) for visual search. In order to further understand what it is that is special about social cues compared to non-social cues, it is important to test a larger vocabulary of social cues including eye gaze, emotional faces, finger gestures, and body postures. It is also useful to extend the investigation to lexical cues. With a larger vocabulary of social cues tested, we can then investigate whether there is also a saliency map for these social cues. Note that in both Chapter III and Chapter V, we computed and controlled for low-level saliency, in order to facilitate comparisons on the semantic level.

To computationally model saliency and create a comprehensive saliency map including social and semantic cues, I was also involved in modeling low-, object- and semantic-level saliency in natural scenes (Xu et al., 2014). To further understand the social aspect of the saliency map, we are extending the current framework to viewers with autism spectrum disorder (ASD): is there a difference in saliency representation in ASD? The computational model will be fit on the basis of behavior, and then queried in relation to the neuronal responses obtained from single-neuron recordings in humans. We will acquire two behavioral measures: (i) explicit ratings of interest and saliency from the

subjects when they see stimuli, and (ii) implicit saliency derived from where people look: their fixations onto stimuli when presented in arrays will index how salient the stimuli are (more fixated = more salient). We hypothesize that the behavior of people with ASD can best be fit by a model that has different weights for semantic categories (e.g., lower weights for faces or humans but higher weights for autism special-interest objects; the higher the weights, the more salient to the subject) as well as generally higher weights for low-level saliency (corresponding to increased fixations to visual features such as edges and flickering stimuli). We will also correlate the attribute values (a vector for each image) from this computational model of saliency trial-by-trial with the neuronal responses from the amygdala—the best correlated attributes would indicate that the amygdala represents them as the most salient. We will analyze this not only categorically (diagnosis of ASD), but also along dimensions in the non-autistic population (e.g., with questionnaires like the Autism Quotient (AQ) that show a distribution across the population). Preliminary analysis has shown a general similarity between people with ASD and healthy individuals, but revealed small differences in saliency weights of social attributes.

### 6.2 Investigating face perception

In Chapter III and Chapter V, we have shown that faces are particularly salient stimuli. In Chapter IV, we showed that single-neurons in the human amygdala encode subjectively perceived fearful and happy emotions. This opens many ensuing questions: does the amygdala track other emotions and to what extent does the amygdala track emotions? Is the subjectivity encoded in the amygdala restricted to emotions and faces? Does the amygdala track visual saliency or distinctiveness independently of, or disproportionally to, other stimulus attributes? Does the amygdala track visual saliency for all stimuli or specifically or disproportionally for social stimuli? Are bottom-up saliency and top-down saliency both coded in the amygdala? What parameters describing visual stimuli best correlate with amygdala activation? Even harder questions include what the neuronal

population coding of faces is in the amygdala and where the subjective judgment of faces commences. Since we have only tested one set of faces ("bubbled" faces), do amygdala neurons track subjective judgment on other face stimuli, especially ambiguous face stimuli in which different subjective judgments are possible on an identical face? In this case, how can perceptual decisions be formed based on eye movements? On the other hand, are there individual differences in subjective face judgment, in particular when comparing to neurological populations (e.g., people with autism and patients with amygdala lesions)? Are there in-group, out-group, or cultural effects? Neuroimaging studies have shown that some brain regions track stimulus strength and represent faces in a continuous manner, such as the posterior superior temporal sulcus for morphed emotions (Harris et al., 2012) and the fusiform gyrus for morphed genders (Freeman et al., 2010), and some other brain regions form categorical judgments, such as the amygdala (Harris et al., 2012) and the orbitofrontal cortex (Freeman et al., 2010). What are the response characteristics of the amygdala neurons to faces gradually changing along one dimension? Is it all the same for faces along many different dimensions? To answer these questions, we have conducted the following studies using eye-tracking and single-neuron recordings with concurrent eye-tracking. We are also testing the same paradigms on people with autism and amygdala lesion patients. These paradigms are extendable for future fMRI studies as well.

### *6.2.1 Investigating how neurons in the amygdala respond to morphed faces*

In Chapter IV, we used "bubbled faces" as stimuli. In each trial, only a certain parts of faces were revealed, and different parts of faces were revealed in different trials. These stimuli were not particularly natural, and indeed they just look artificial. Also, it is difficult to control the relationship between stimulus input and behavioral output given variable facial parts shown in each trial. Therefore, in this new study, we are using morphed faces, which have ambiguous facial emotions, and most importantly, different subjective judgments can be made on an identical face. We have created 5 levels of fear-

happy morphs (ranging from 30% fear–70% happy to 70% fear–30% happy) and 2 anchor faces of fear and happy without ambiguity. We present patients with these morphed faces and ask them to judge whether the emotion is ‘fear’ or ‘happy’. Using these faces, we can confirm whether amygdala neurons encode subjective judgment.

The gradient of morphs further allows us to test whether amygdala neurons track stimulus physical strength in a continuous manner or encode subjective categorization in a binary manner. Further research questions involve testing behavioral and electrophysiological abnormalities in people with ASD for emotion perception and judgment, with a focus on the amygdala. We hypothesize that people with ASD will judge the morphed faces differently (more biased towards one emotion and relatively insensitive to blends) and that the neuronal response to faces will differ in accordance with the behavior. We will specifically examine evoked responses in amygdala, in prefrontal cortex, and in temporal cortex, as well as functional connectivity between these regions. This will help to determine at what stage of processing the behavioral impairment arises.

We have recorded more than 250 neurons from 16 sessions from 10 patients. Our preliminary analysis has shown that about a third of amygdala neurons track the gradual change of emotion in the stimulus. Interestingly, testing the same task on amygdala lesion patients and people with autism has revealed increased percentage of fearful judgment on the same face, indicating a functional role of the amygdala in face judgment and possible deficits of face judgment in people with autism which might be due to amygdala dysfunction.

### *6.2.2 Investigating faces along many dimensions*

In everyday life, people constantly form judgments of others based on purely facial features. The objective of this project is to determine which dimension of faces the amygdala best tracks and investigate whether there is a difference in people with ASD. Established software FaceGen was used to create sets of faces that vary from one

dimension to another: “happy vs. fearful”, “anger vs. disgust”, “anger vs. fearful”, “male vs. female” and more complex social dimensions, like dominance and trustworthiness. This expands the dimensions under investigation and we will still record single neurons using these sets of faces. We present subjects with these parameterized faces and ask them to judge the face as well as rate the confidence they have in their judgment. Preliminary behavioral data that we have collected from healthy individuals have shown that subjects are able to pick up subtle changes in facial structures regarding emotion and gender in their judgments. We are collecting data from people with ASD and we hypothesize that people with ASD will judge some facial dimensions differently compared to controls (which would be revealed as shifted psychometric curves) and that neuronal responses in the amygdala will track these behavioral differences. Specifically, we will measure spike rate from single neurons within the amygdala, and in addition examine evoked responses from prefrontal cortex and temporal cortex, as well as connectivity between these regions. Furthermore, we will record eye movements when subjects view and judge these faces to answer the questions listed in the section below.

### *6.2.3 Eye-tracking*

One of the very novel aspects in our single-neuron recording setup is to simultaneously record eye movements. We can thus investigate how eye movements are related to face perception and judgments, and the underlying neural basis of such eye movements. We have shown in our previous work that amygdala neurons are tuned to facial features (Rutishauser et al., 2013). However, without concurrent eye-tracking, we were only able to infer the relationship between neuronal response and eye movement by recalling recorded neurosurgical patients to the laboratory for eye-tracking experiments, but not able to tell exactly how neurons respond to the fixated facial features.

In this study, we will investigate how the eye-movement patterns and the underlying neural responses determine their decisions. The questions we can analyze are (1) how do

fixation density maps / fixation density in specific regions of interest (ROIs) evolve with parametrically synthesized faces along different dimensions? (2) For a given face (identical visual input), do subjects judge the emotion differently if they look at faces differently—is the behavioral choice influenced by the gaze pattern? (3) How is the neural response related to the fixation? Specifically, do neurons fire in preparation of a fixation towards a facial part, or only once the facial part has in fact been fixated? (4) What is the relationship between how facial features drive neuronal response, and how they drive fixations by the viewer? We can compare people with ASD and without ASD for each question, and we can correlate the eye movement patterns (e.g., fixation density in eye ROI) with autism scores (measured by autism questionnaires like Autism Quotient (AQ) and Social Responsiveness Scale (SRS)). To investigate these questions, we have recorded from five neurosurgical patients with concurrent eye-tracking. Furthermore, in the laboratory (not at the hospital), we are administering the task behaviorally in a larger sample (N = 10–20) of people with autism and matched controls (neither surgical patients) in order to check that the surgical patients are representative in terms of their behavior.

### 6.3 Investigating saliency in visual search

In Chapter V, we have shown that healthy individuals can adopt an efficient search strategy with top-down task instructions. But what are the underlying neural mechanisms of this behavior? It is known that the amygdala plays an important role in detecting saliency and biological relevance. How do amygdala neurons respond to salient and target-relevant objects (compared to task-irrelevant objects) in the visual search? Furthermore, is there a difference in response between social and non-social objects? Is there a difference in neuronal response between people with ASD and controls? How do neurons respond to eye movements? Specifically, do amygdala neurons fire in preparation of a fixation towards a salient target, or only once the target has in fact been fixated (that is, are amygdala responses cause or consequence of visual attention). If

neurosurgical patients co-morbid with autism are available, we expect that the proportion of face-responsive neurons in the amygdala will be lower in ASD, whereas the proportion of neurons responsive to autism special-interest objects will be higher. This would indicate that the amygdala represents salient stimuli, and that what is salient differs in ASD. We would also expect reduced response to socially relevant objects, in accordance with the behavioral findings shown in Chapter V.

To answer these questions, we have been conducting further experiments with single-neuron recordings in neurosurgical patients with concurrent eye-tracking. It is well known that the human medial temporal lobe plays a key role in analyzing and recognizing visual objects, purported to convey saliency and object selection signals through its dense network with other cortical regions. Using the same task as in Chapter V, we have recorded from over 150 single neurons in the amygdalae and hippocampi of four neurosurgical patients (five sessions) with implanted depth electrodes. Behaviorally, we found patients rapidly oriented to, and persistently searched among, target-congruent objects, consistent with the findings of healthy individuals shown in Chapter V. Trial-by-trial analysis showed that 11.7% (8.3% for amygdala and 16.7% for hippocampus) of neurons responded only when a target was found. Fixation analysis revealed neurons that responded only when a fixation fell on a target but not a distractor. By comparing the average number of spikes in a time window of the entire fixation duration between fixations on targets and distractors, we selected 24.2% of these target neurons (9.7% for amygdala and 45.8% for hippocampus; two-tailed t-test at $p < 0.05$; among which 82.3% increased firing rate to targets and 17.2% decreased). Since the same objects can be either targets or distractors on different trials, this reveals the top-down driven nature of the response, which suggests that task-relevant target saliency was encoded by a subset of neurons. We further conducted two control experiments in the same patients: one ruled out motor confounds by applying a fixed duration of search in the same task without button press, and the second ruled out working memory or object matching confounds by employing a pop-out search task in which the target was defined as one face among vehicles or one vehicle among faces (thus no pre-defined target). We also found 21.7%

(27.8% for amygdala and 12.5% for hippocampus) of neurons that responded more to social objects than to non-social objects. Interestingly, there was a small population (4%) of target-responsive neurons that differentiated social vs. non-social objects both during single target presentation and subsequent target detection in the search array. Taken together, we found compelling evidence that neurons in the human medial temporal lobe encode object categories and saliency signals that contribute to attention.

### 6.4 More ecologically valid stimuli

A last note is that all the stimuli used in my thesis are in 2D. To be more ecologically valid, stimuli in 3D or face-to-face real person interactions will be necessary. Furthermore, we have been using eye-tracking with high temporal and spatial precision, but it is more desirable to have people wear small eye-trackers in glasses while they walk around the real world. These are challenging future extensions that I am not immediately planning to pursue, but they are important to think about nonetheless.

### 6.5 Conclusion

Findings from my research will have applications to how we interact with other people and the environment, in both the real world and in virtual environments such as through the internet. How can one grab and guide another's attention most efficiently? How can one influence another's attention and memory through structured presentation of items? How can we help and train people with autism given the identified attentional deficits? Can we build biologically plausible computational models of social saliency and build social robots? My thesis contributes to understanding these questions.

# REFERENCES

Adams RB, Kleck RE (2003) Perceived Gaze Direction and the Processing of Facial Displays of Emotion. Psychological Science 14:644-647.

Adolphs R (2008) Fear, faces, and the human amygdala. Current Opinion in Neurobiology 18:166-172.

Adolphs R (2010) What does the amygdala contribute to social cognition? Annals of the New York Academy of Sciences 1191:42-61.

Adolphs R, Baron-Cohen S, Tranel D (2002) Impaired Recognition of Social Emotions following Amygdala Damage. Journal of Cognitive Neuroscience 14:1264-1274.

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

Adolphs R, Sears L, Piven J (2001) Abnormal Processing of Social Information from Faces in Autism. Journal of Cognitive Neuroscience 13:232-240.

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

Adolphs R, Tranel D, Damasio H, Damasio A (1994) Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. Nature 372:669-672.

Adolphs R, Tranel D, Hamann S, Young AW, Calder AJ, Phelps EA, Anderson A, Lee GP, Damasio AR (1999) Recognition of facial emotion in nine individuals with bilateral amygdala damage. Neuropsychologia 37:1111-1117.

Afraz S-R, Kiani R, Esteky H (2006) Microstimulation of inferotemporal cortex influences face categorization. Nature 442:692-695.

Akiyama T, Kato M, Muramatsu T, Saito F, Umeda S, Kashima H (2006) Gaze but not arrows: A dissociative impairment after right superior temporal gyrus damage. Neuropsychologia 44:1804-1810.

Amaral DG, Behniea H, Kelly JL (2003) Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey. Neuroscience 118:1099-1120.

Amaral DG, Price JL (1984) Amygdalo-cortical projections in the monkey (Macaca fascicularis). J Comp Neurol 230:465-496.

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

Amaral DG, Schumann CM, Nordahl CW (2008) Neuroanatomy of autism. Trends in Neurosciences 31:137-145.

Anderson AK, Christoff K, Panitz D, De Rosa E, Gabrieli JDE (2003) Neural Correlates of the Automatic Processing of Threat Facial Signals. The Journal of Neuroscience 23:5627-5633.

Anderson AK, Phelps EA (2001) Lesions of the human amygdala impair enhanced perception of emotionally salient events. Nature 411:305-309.

Anderson BA, Laurent PA, Yantis S (2011) Value-driven attentional capture. Proceedings of the National Academy of Sciences.

Anderson JR, Sallaberry P, Barbier H (1995) Use of experimenter-given cues during object-choice tasks by capuchin monkeys. Animal Behaviour 49:201-208.

Anderson JS, Druzgal TJ, Froehlich A, DuBray MB, Lange N, Alexander AL, Abildskov T, Nielsen JA, Cariello AN, Cooperrider JR, Bigler ED, Lainhart JE (2010) Decreased Interhemispheric Functional Connectivity in Autism. Cerebral Cortex.

Arcizet F, Mirpour K, Bisley JW (2011) A Pure Salience Response in Posterior Parietal Cortex. Cerebral Cortex 21:2498-2506.

Argyle M, Ingham R, Alkema F, McCallin M (1973) The Different Functions of Gaze. In: Semiotica, vol. 7, p 19.

Aviezer H, Trope Y, Todorov A (2012) Body Cues, Not Facial Expressions, Discriminate Between Intense Positive and Negative Emotions. Science 338:1225-1229.

Bagshaw MH, Mackworth NH, Pribram KH (1972) The effect of resections of the inferotemporal cortex or the amygdala on visual orienting and habituation. Neuropsychologia 10:153-162.

Ball W, Tronick E (1971) Infant responses to impending collision. Science 171:818-820.

Bar M (2007) The proactive brain: using analogies and associations to generate predictions. Trends in Cognitive Sciences 11:280-289.

Baron-Cohen S (2004) Autism: research into causes and intervention. Developmental Neurorehabilitation 7:73-78.

Baron-Cohen S, Ring HA, Bullmore ET, Wheelwright S, Ashwin C, Williams SCR (2000) The amygdala theory of autism. Neuroscience & Biobehavioral Reviews 24:355-364.

Baron-Cohen S, Wheelwright S, Hill J, Raste Y, Plumb I (2001) The "Reading the Mind in the Eyes" Test Revised Version: A Study with Normal Adults, and Adults with Asperger Syndrome or High-functioning Autism. Journal of Child Psychology and Psychiatry 42:241-251.

Basso MA, Wurtz RH (2002) Neuronal Activity in Substantia Nigra Pars Reticulata during Target Selection. The Journal of Neuroscience 22:1883-1894.

Batki A, Baron-Cohen S, Wheelwright S, Connellan J, Ahluwalia J (2000) Is there an innate gaze module? Evidence from human neonates. Infant Behavior and Development 23:223-229.

Bauman M, Kemper TL (1985) Histoanatomic observations of the brain in early infantile autism. Neurology 35:866-874.

Baxter MG, Murray EA (2002) The amygdala and reward. Nat Rev Neurosci 3:563-573.

Baxter MG, Parker A, Lindner CCC, Izquierdo AD, Murray EA (2000) Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. J Neurosci 20:4311-4319.

Baylis GC, Rolls ET, Leonard CM (1985) Selectivity between faces in the responses of a population of neurons in the cortex in the superior temporal sulcus of the monkey. Brain Research 342:91-102.

Bechara A, Tranel D, Damasio H, Adolphs R, Rockland C, Damasio A (1995) Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. Science 269:1115-1118.

Becker B, Mihov Y, Scheele D, Kendrick KM, Feinstein JS, Matusch A, Aydin M, Reich H, Urbach H, Oros-Peusquens A-M, Shah NJ, Kunz WS, Schlaepfer TE, Zilles K, Maier W, Hurlemann R (2012) Fear Processing and Social Networking in the Absence of a Functional Amygdala. Biological Psychiatry 72:70-77.

Belova MA, Paton JJ, Morrison SE, Salzman CD (2007) Expectation Modulates Neural Responses to Pleasant and Aversive Stimuli in Primate Amygdala. Neuron 55:970-984.

Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological) 57:289-300.

Bermudez MA, Gobel C, Schultz W (2012) Sensitivity to Temporal Reward Structure in Amygdala Neurons. Current biology : CB 22:1839-1844.

Berridge KC, Robinson TE (1998) What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Research Reviews 28:309-369.

Bichot NP, Rossi AF, Desimone R (2005) Parallel and Serial Neural Mechanisms for Visual Search in Macaque Area V4. Science 308:529-534.

Blair IV, Judd CM, Chapleau KM (2004) The Influence of Afrocentric Facial Features in Criminal Sentencing. Psychological Science 15:674-679.

Blakemore S-J, Decety J (2001) From the perception of action to the understanding of intention. Nat Rev Neurosci 2:561-567.

Bower TGR, Broughton JM, Moore MK (1970) Infant responses to approaching objects: an indicator of response to distal variables. Perception & psychophysics 9:193-196.

Brainard DH (1997) The Psychophysics Toolbox. Spat Vis 10:433-436.

Bremmer F, Duhamel JR, Ben Hamed S, Graf W (2002) Heading encoding in the macaque ventral intraparietal area (VIP). The European journal of neuroscience 16:1554-1568.

Britten KH (2008) Mechanisms of self-motion perception. Annual review of neuroscience 31:389-410.

Britten KH, van Wezel RJ (1998) Electrical microstimulation of cortical area MST biases heading perception in monkeys. Nature neuroscience 1:59-63.

Brosseau-Lachaine O, Casanova C, Faubert J (2008) Infant sensitivity to radial optic flow fields during the first months of life. Journal of vision 8:5 1-14.

Brothers L (1990) The social brain: A project for integrating primate behavior and neurophysiology in a new domain. Concepts in Neuroscience 1:27-51.

Buchanan TW, Tranel D, Adolphs R (2009) The human amygdala in social function. In: The human amygdala (Whalen, P. W. and Phelps, L., eds), pp 289–320 New York: Oxford University Press.

Burrows BE, Moore T (2009) Influence and Limitations of Popout in the Selection of Salient Visual Stimuli by Area V4 Neurons. The Journal of Neuroscience 29:15169-15177.

Byrnit J (2009) Gorillas' (Gorilla gorilla) use of experimenter-given manual and facial cues in an object-choice task. Anim Cogn 12:401-404.

Cahill L, Haier RJ, Fallon J, Alkire MT, Tang C, Keator D, Wu J, McGaugh JL (1996) Amygdala activity at encoding correlated with long-term, free recall of emotional information. Proc Natl Acad Sci USA 93:8016-8021.

Campbell R, Heywood CA, Cowey A, Regard M, Landis T (1990) Sensitivity to eye gaze in prosopagnosic patients and monkeys with superior temporal sulcus ablation. Neuropsychologia 28:1123-1142.

Carlin JD, Calder AJ, Kriegeskorte N, Nili H, Rowe JB (2011) A Head View-Invariant Representation of Gaze Direction in Anterior Superior Temporal Sulcus. Current biology : CB 21:1817-1821.

Catani M, Jones DK, Donato R, Ffytche DH (2003) Occipito-temporal connections in the human brain. Brain 126:2093-2107.

Cauchoix M, Crouzet SM (2013) How plausible is a subcortical account of rapid visual recognition? Frontiers in Human Neuroscience 7.

Cerf M, Frady EP, Koch C (2009) Faces and text attract gaze independent of the task: Experimental data and computer model. Journal of Vision 9.

Chelazzi L, Miller EK, Duncan J, Desimone R (1993) A neural basis for visual search in inferior temporal cortex. Nature 363:345-347.

Colby CL, Goldberg ME (1999) Space and attention in parietal cortex. Annual review of neuroscience 22:319-349.

Constantinidis C, Steinmetz MA (2001) Neuronal Responses in Area 7a to Multiple-stimulus Displays: I. Neurons Encode the Location of the Salient Stimulus. Cerebral Cortex 11:581-591.

Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. Nature reviews 3:201-215.

Cornelissen FW, Peters EM, Palmer J (2002) The Eyelink Toolbox: eye tracking with MATLAB and the Psychophysics Toolbox. Behav Res Methods Instrum Comput 34:613-617.

Dalton KM, Nacewicz BM, Alexander AL, Davidson RJ (2007) Gaze-Fixation, Brain Activation, and Amygdala Volume in Unaffected Siblings of Individuals with Autism. Biological Psychiatry 61:512-520.

Dalton KM, Nacewicz BM, Johnstone T, Schaefer HS, Gernsbacher MA, Goldsmith HH, Alexander AL, Davidson RJ (2005) Gaze fixation and the neural circuitry of face processing in autism. Nat Neurosci 8:519-526.

Dawson G, Webb SJ, McPartland J (2005) Understanding the Nature of Face Processing Impairment in Autism: Insights From Behavioral and Electrophysiological Studies. Developmental Neuropsychology 27:403-424.

De Renzi E, Perani D, Carlesimo GA, Silveri MC, Fazio F (1994) Prosopagnosia can be associated with damage confined to the right hemisphere--An MRI and PET study and a review of the literature. Neuropsychologia 32:893-902.

de Vries SEJ, Clandinin TR (2012) Loom-Sensitive Neurons Link Computation to Action in the Drosophila Visual System. Current biology : CB 22:353-362.

Desimone R, Gross CG (1979) Visual areas in the temporal cortex of the macaque. Brain Research 178:363-380.

Dill LM (1974) The escape response of the zebra danio (Brachydanio rerio). I. The stimulus for escape. Anim Behav 22:771-722.

Dolan RJ (2002) Emotion, Cognition, and Behavior. Science 298:1191-1194.

DSM-5 (2013) Diagnostic and Statistical Manual of Mental Disorders: DSM-5: American Psychiatric Association.

Duffy CJ (1998) MST Neurons Respond to Optic Flow and Translational Movement. Journal of neurophysiology 80:1816-1827.

Duffy CJ, Wurtz RH (1991) Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. Journal of neurophysiology 65:1329-1345.

Duffy CJ, Wurtz RH (1995) Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. J Neurosci 15:5192-5208.

Dunbar RIM (2010) The social role of touch in humans and primates: Behavioural function and neurobiological mechanisms. Neuroscience & Biobehavioral Reviews 34:260-268.

Ecker C, Suckling J, Deoni SC, Lombardo MV, Bullmore ET, Baron-Cohen S, Catani M, Jezzard P, Barnes A, Bailey AJ, Williams SC, Murphy DGM (2012) Brain anatomy and its relationship to behavior in adults with autism spectrum disorder: A multicenter magnetic resonance imaging study. Archives of General Psychiatry 69:195-209.

Efron B, Tibshirani RJ (1994) An Introduction to the Bootstrap (Chapman & Hall/CRC Monographs on Statistics & Applied Probability): Chapman and Hall/CRC.

Ekman P, Friesen WV (1975) Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues. Englewood Cliffs, NJ: Prentice Hall.

Emery NJ, Capitanio JP, Mason WA, Machado CJ, Mendoza SP, Amaral DG (2001) The effects of bilateral lesions of the amygdala on dyadic social interactions in rhesus monkeys (Macaca mulatta). Behavioral Neuroscience 115:515-544.

Emery NJ, Lorincz EN, Perrett DI, Oram MW, Baker CI (1997) Gaze Following and Joint Attention in Rhesus Monkeys (Macaca mulatta). Journal of Comparative Psychology 111:286-293.

Farran EK, Branson A, King BJ (2011) Visual search for basic emotional expressions in autism; impaired processing of anger, fear and sadness, but a typical happy face advantage. Research in Autism Spectrum Disorders 5:455-462.

Feinstein JS, Adolphs R, Damasio A, Tranel D (2011) The Human Amygdala and the Induction and Experience of Fear. Current biology : CB 21:34-38.

Field DT, Wann JP (2005) Perceiving Time to Collision Activates the Sensorimotor Cortex. Curr Biol 15:453-458.

Field T (2003) Touch. Cambridge, MA: MIT Press.

Fitzgerald DA, Angstadt M, Jelsone LM, Nathan PJ, Phan KL (2006) Beyond threat: Amygdala reactivity across multiple expressions of facial affect. NeuroImage 30:1441-1448.

Fotowat H, Gabbiani F (2011) Collision Detection as a Model for Sensory-Motor Integration. Annual review of neuroscience 34:1-19.

Foulsham T, Underwood G (2008) What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. Journal of vision 8:6 1-17.

Fox E (2000) Facial expressions of emotion: are angry faces detected more efficiently? Cognit Emotion 14:61-92.

Fox MD, Corbetta M, Snyder AZ, Vincent JL, Raichle ME (2006) Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. Proceedings of the National Academy of Sciences 103:10046-10051.

Franconeri SL, Simons DJ (2003) Moving and looming stimuli capture attention. Perception & psychophysics 65:999-1010.

Freeman JB, Rule NO, Adams RB, Ambady N (2010) The Neural Basis of Categorical Face Perception: Graded Representations of Face Gender in Fusiform and Orbitofrontal Cortices. Cerebral Cortex 20:1314-1322.

Freese JL, Amaral DG (2006) Synaptic organization of projections from the amygdala to visual cortical areas TE and V1 in the macaque monkey. The Journal of Comparative Neurology 496:655-667.

Fried I, MacDonald KA, Wilson CL (1997) Single Neuron Activity in Human Hippocampus and Amygdala during Recognition of Faces and Objects. Neuron 18:753-765.

Fried I, Mateer C, Ojemann G, Wohns R, Fedio P (1982) Organization of visuospatial functions in human cortex. Evidence from electrical stimulation. Brain 105:349-371.

Friesen CK, Kingstone A (1998) The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. Psychonomic Bulletin & Review 5:490-495.

Friston KJ (2007) Statistical parametric mapping : the analysis of functional brain images. Amsterdam; Boston: Elsevier/Academic Press.

Gallace A, Spence C (2010) The science of interpersonal touch: An overview. Neuroscience & Biobehavioral Reviews 34:246-259.

Gamer M, Büchel C (2009) Amygdala Activation Predicts Gaze toward Fearful Eyes. The Journal of Neuroscience 29:9123-9126.

Gazzola V, Spezio ML, Etzel JA, Castelli F, Adolphs R, Keysers C (2012) Primary somatosensory cortex discriminates affective significance in social touch. Proceedings of the National Academy of Sciences.

Geschwind DH, Levitt P (2007) Autism spectrum disorders: developmental disconnection syndromes. Current Opinion in Neurobiology 17:103-111.

Ghashghaei HT, Barbas H (2002) Pathways for emotion: interactions of prefrontal and anterior temporal pathways in the amygdala of the rhesus monkey. Neuroscience 115:1261-1279.

Gosselin F, Schyns PG (2001) Bubbles: a technique to reveal the use of information in recognition tasks. Vision Research 41:2261-2271.

Gothard KM, Battaglia FP, Erickson CA, Spitler KM, Amaral DG (2007) Neural Responses to Facial Expression and Face Identity in the Monkey Amygdala. Journal of Neurophysiology 97:1671-1683.

Gottlieb JP, Kusunoki M, Goldberg ME (1998) The representation of visual salience in monkey parietal cortex. Nature 391:481-484.

Gotts SJ, Simmons WK, Milbury LA, Wallace GL, Cox RW, Martin A (2012) Fractionation of social brain circuits in autism spectrum disorders. Brain 135:2711-2725.

Grabenhorst F, Hernadi I, Schultz W (2012) Prediction of economic choice by primate amygdala neurons. Proceedings of the National Academy of Sciences.

Grant EC, Mackintosh JH (1963) A Comparison of the Social Postures of Some Common Laboratory Rodents. Behaviour 21:246-259.

Gray R, Regan D (1998) Accuracy of estimating time to collision using binocular and monocular information. Vision research 38:499-512.

Graziano MS, Andersen RA, Snowden RJ (1994) Tuning of MST neurons to spiral motions. J Neurosci 14:54-67.

Grimes J (1996) On the Failure to Detect Changes in Scenes across Saccades. In: Perception: Vancouver Studies in Cognitive Science, vol. 5 (Akins, K., ed), pp 89–110 New York: Oxford University Press.

Gross CG (1994) How Inferior Temporal Cortex Became a Visual Area. Cerebral Cortex 4:455-469.

Hadj-Bouziane F, Liu N, Bell AH, Gothard KM, Luh W-M, Tootell RBH, Murray EA, Ungerleider LG (2012) Amygdala lesions disrupt modulation of functional MRI activity evoked by facial expression in the monkey inferior temporal cortex. Proceedings of the National Academy of Sciences 109:E3640-E3648.

Hall ET (1966) The hidden dimension. Garden City, N.Y.: Doubleday.

Hampton AN, Adolphs R, Tyszka JM, O'Doherty JP (2007) Contributions of the Amygdala to Reward Expectancy and Choice Signals in Human Prefrontal Cortex. Neuron 55:545-555.

Happe F (2003) Cognition in autism: one deficit or many? Novartis Found Symp 251:198–207.

Happe F, Ronald A, Plomin R (2006) Time to give up on a single explanation for autism. Nat Neurosci 9:1218-1220.

Hariri AR, Mattay VS, Tessitore A, Kolachana B, Fera F, Goldman D, Egan MF, Weinberger DR (2002) Serotonin Transporter Genetic Variation and the Response of the Human Amygdala. Science 297:400-403.

Harms M, Martin A, Wallace G (2010) Facial Emotion Recognition in Autism Spectrum Disorders: A Review of Behavioral and Neuroimaging Studies. Neuropsychol Rev 20:290-322.

Harris RJ, Young AW, Andrews TJ (2012) Morphing between expressions dissociates continuous from categorical representations of facial expression in the human brain. Proceedings of the National Academy of Sciences 109:21164-21169.

Hasselmo ME, Rolls ET, Baylis GC (1989) The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. Behavioural Brain Research 32:203-218.

Hayes WN, Saiff EI (1967) Visual alarm reactions in turtles. Anim Behav 15:102-108.

He S, Cavanagh P, Intriligator J (1996) Attentional resolution and the locus of visual awareness. Nature 383:334-337.

Hemmi JM, Tomsic D (2012) The neuroethology of escape in crabs: from sensory ecology to neurons and back. Current Opinion in Neurobiology 22:194-200.

Herry C, Bach DR, Esposito F, Di Salle F, Perrig WJ, Scheffler K, Luthi A, Seifritz E (2007) Processing of Temporal Unpredictability in Human and Animal Amygdala. The Journal of Neuroscience 27:5958-5966.

Hickey C, Chelazzi L, Theeuwes J (2010) Reward Changes Salience in Human Vision via the Anterior Cingulate. The Journal of Neuroscience 30:11096-11103.

Hickey C, van Zoest W (2012) Reward creates oculomotor salience. Current Biology 22:R219-R220.

Hietanen J (2002) Social attention orienting integrates visual information from head and body orientation. Psychological Research 66:174-179.

Hofer PA (1973) Urbach-Wiethe disease (lipoglycoproteinosis; lipoid proteinosis; hyalinosis cutis et mucosae). A review. Acta Derm Venereol Suppl (Stockh) 53:1-52.

Hoffman EA, Haxby JV (2000) Distinct representations of eye gaze and identity in the distributed human neural system for face perception. Nat Neurosci 3:80-84.

Hoffman KL, Gothard KM, Schmid MC, Logothetis NK (2007) Facial-Expression and Gaze-Selective Responses in the Monkey Amygdala. Current biology : CB 17:766-772.

Hollingworth A, Henderson JM (2000) Semantic Informativeness Mediates the Detection of Changes in Natural Scenes. Visual Cognition 7:213-235.

Hood BM, Willen JD, Driver J (1998) Adult's Eyes Trigger Shifts of Visual Attention in Human Infants. Psychological Science 9:131-134.

Horley K, Williams LM, Gonsalvez C, Gordon E (2004) Face to face: visual scanpath evidence for abnormal processing of facial expressions in social phobia. Psychiatry Research 127:43-53.

Ikemoto S, Panksepp J (1999) The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. Brain Research Reviews 31:6-41.

Ingle DJ, Shook BL (1985) In: Brain Mechanisms of Spatial Vision (Ingle, D. J. et al., eds), pp 229-258 Dordrecht: Martinus Nijhoff.

Ipata AE, Gee AL, Goldberg ME, Bisley JW (2006) Activity in the Lateral Intraparietal Area Predicts the Goal and Latency of Saccades in a Free-Viewing Visual Search Task. The Journal of Neuroscience 26:3656-3661.

Itti L, Koch C (2001) Computational modelling of visual attention. Nature reviews 2:194-203.

Itti L, Koch C, Niebur E (1998) A model of saliency-based visual attention for rapid scene analysis. IEEE Trans Patt Anal Mach Intell 20:1254-1259.

Izuma K, Matsumoto K, Camerer CF, Adolphs R (2011) Insensitivity to social reputation in autism. Proceedings of the National Academy of Sciences 108:17302-17307.

Izuma K, Saito DN, Sadato N (2008) Processing of Social and Monetary Rewards in the Human Striatum. Neuron 58:284-294.

Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM (2012) FSL. NeuroImage 62:782-790.

Jiang Y, He S (2006) Cortical Responses to Invisible Faces: Dissociating Subsystems for Facial-Information Processing. Current Biology 16:2023-2029.

Joseph RM, Keehn B, Connolly C, Wolfe JM, Horowitz TS (2009) Why is visual search superior in autism spectrum disorder? Developmental Science 12:1083-1096.

Judd T, Ehinger K, Durand F, Torralba A (2009) Learning to predict where humans look. In: Computer Vision, 2009 IEEE 12th International Conference on, pp 2106-2113.

Kanan C, Tong MH, Zhang L, Cottrell GW (2009) SUN: Top-down saliency using natural statistics. Visual Cognition 17:979-1003.

Kanner L (1943) Autistic disturbances of affective contact. The Nervous Child 2:217–250.

Kemner C, van Ewijk L, van Engeland H, Hooge I (2008) Brief report: eye movements during visual search tasks indicate enhanced stimulus discriminability in subjects with PDD. J Autism Dev Disord 38:553-557.

Kennedy DP, Adolphs R (2012) Perception of emotions from facial expressions in high-functioning adults with autism. Neuropsychologia 50:3313-3319.

Kennedy DP, Glascher J, Tyszka JM, Adolphs R (2009) Personal space regulation by the human amygdala. Nat Neurosci 12:1226-1227.

Kilpatrick L, Cahill L (2003) Amygdala modulation of parahippocampal and frontal regions during emotionally influenced memory storage. NeuroImage 20:2091-2099.

Kim H, Somerville LH, Johnstone T, Polis S, Alexander AL, Shin LM, Whalen PJ (2004) Contextual Modulation of Amygdala Responsivity to Surprised Faces. Journal of Cognitive Neuroscience 16:1730-1745.

Kingstone A, Friesen CK, Gazzaniga MS (2000) Reflexive Joint Attention Depends on Lateralized Cortical Connections. Psychological Science 11:159-166.

Kingstone A, Tipper C, Ristic J, Ngan E (2004) The eyes have it!: An fMRI investigation. Brain and Cognition 55:269-271.

Kirchner H, Thorpe SJ (2006) Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. Vision Research 46:1762-1776.

Klein RM (2000) Inhibition of return. Trends in cognitive sciences 4:138-147.

Kleinhans NM, Richards T, Johnson LC, Weaver KE, Greenson J, Dawson G, Aylward E (2011) fMRI evidence of neural abnormalities in the subcortical face processing system in ASD. NeuroImage 54:697-704.

Kliemann D, Dziobek I, Hatri A, Baudewig Jr, Heekeren HR (2012) The Role of the Amygdala in Atypical Gaze on Emotional Faces in Autism Spectrum Disorders. The Journal of Neuroscience 32:9469-9476.

Kliemann D, Dziobek I, Hatri A, Steimke R, Heekeren HR (2010) Atypical Reflexive Gaze Patterns on Emotional Faces in Autism Spectrum Disorders. The Journal of Neuroscience 30:12281-12287.

Klin A, Jones W, Schultz R, Volkmar F, Cohen D (2002) Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. Arch Gen Psychiatry 59:809-816.

Klin A, Lin DJ, Gorrindo P, Ramsay G, Jones W (2009) Two-year-olds with autism orient to non-social contingencies rather than biological motion. Nature 459:257-261.

Kling AS, Brothers LA (1992) The Amygdala: Neurobiological Aspects of Emotion, Memory and Mental Dysfunction.

Koch C, Tsuchiya N (2007) Attention and consciousness: two distinct brain processes. Trends in Cognitive Sciences 11:16-22.

Komatsu H, Wurtz RH (1988) Relation of cortical areas MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. Journal of neurophysiology 60:580-603.

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

LaBar K, LeDoux J, Spencer D, Phelps E (1995) Impaired fear conditioning following unilateral temporal lobectomy in humans. The Journal of Neuroscience 15:6846-6855.

Land MF, Lee DN (1994) Where we look when we steer. Nature 369:742-744.

Langton SRH, O'Malley C, Bruce V (1996) Actions speak no louder than words: Symmetrical cross-modal interference effects in the processing of verbal and gestural information. Journal of Experimental Psychology: Human Perception and Performance 22:1357-1375.

Langton SRH, Watt RJ, Bruce V (2000) Do the eyes have it? Cues to the direction of social attention. Trends in cognitive sciences 4:50-59.

Lappe M, Bremmer F, van den Berg AV (1999) Perception of self-motion from visual flow. Trends in cognitive sciences 3:329-336.

Lappe M, Pekel M, Hoffmann K-P (1998) Optokinetic Eye Movements Elicited by Radial Optic Flow in the Macaque Monkey. Journal of neurophysiology 79:1461-1480.

Laurent G, Gabbiani F (1998) Collision-avoidance: nature's many solutions. Nature neuroscience 1:261-263.

Law Smith MJ, Montagne B, Perrett DI, Gill M, Gallagher L (2010) Detecting subtle facial emotion recognition deficits in high-functioning Autism using dynamic stimuli of varying intensities. Neuropsychologia 48:2777-2781.

LeCouteur A, Rutter M, Lord C (1989) Autism diagnostic interview: A standardized investigator-based instrument. J Autism Dev Disord 19:363–387.

LeDoux JE (1993) Emotional memory systems in the brain. Behav Brain Res 58:69-79.

LeDoux JE (1996) The Emotional Brain: The Mysterious Underpinnings of Emotional Life. Simon & Schuster.

Lee DN (1976) A theory of visual control of braking based on information about time-to-collision. Perception 5:437-459.

Lee DN, Davies M, Green PR, Van Der Weel FR (1993) Visual control of velocity of approach by pigeons when landing. J Exp Biol 180:85-104.

Lee DN, Reddish PE (1981) Plummeting gannets: a paradigm of ecological optics. Nature 293:293-294.

Leonard CM, Rolls ET, Wilson FA, Baylis GC (1985) Neurons in the amygdala of the monkey with responses selective for faces. Behav Brain Res 15:159-176.

Lewis MH, Bodfish JW (1998) Repetitive behavior disorders in autism. Mental Retardation and Developmental Disabilities Research Reviews 4:80-89.

Li FF, VanRullen R, Koch C, Perona P (2002) Rapid natural scene categorization in the near absence of attention. Proceedings of the National Academy of Sciences 99:9596-9601.

Lin A, Adolphs R, Rangel A (2012a) Impaired learning of social compared to monetary rewards in autism. Frontiers in Neuroscience 6.

Lin A, Adolphs R, Rangel A (2012b) Social and monetary reward learning engage overlapping neural substrates. Social Cognitive and Affective Neuroscience 7:274-281.

Lin A, Tsai K, Rangel A, Adolphs R (2012c) Reduced social preferences in autism: evidence from charitable donations. Journal of Neurodevelopmental Disorders 4:8.

Lin JY, Franconeri S, Enns JT (2008) Objects on a Collision Path With the Observer Demand Attention. Psychol Sci 19:686-692.

Lin JY, Murray SO, Boynton GM (2009) Capture of Attention to Threatening Stimuli without Perceptual Awareness. Curr Biol 19:1118-1122.

Logothetis NK, Sheinberg DL (1996) Visual object recognition. Annual Review of Neuroscience 19:577-621.

Loken LS, Wessberg J, Morrison I, McGlone F, Olausson H (2009) Coding of pleasant touch by unmyelinated afferents in humans. Nat Neurosci 12:547-548.

Lord C, Rutter M, Goode S, Heemsbergen J, Jordan H, Mawhood L (1989) Autism diagnostic observation schedule: A standardized observation of communicative and social behavior. J Autism Dev Disord 19:185–212.

MacNeilage PF, Rogers LJ, Vallortigara G (2009) Origins of the left & right brain. Scientific American 301:60-67.

Malkova L, Gaffan D, Murray EA (1997) Excitotoxic lesions of the amygdala fail to produce impairments in visual learning for auditory secondary reinforcement but interfere with reinforcer devaluation effects in rhesus monkeys. J Neurosci 17:6011-6020.

Manjaly ZM, Bruning N, Neufang S, Stephan KE, Brieber S, Marshall JC, Kamp-Becker I, Remschmidt H, Herpertz-Dahlmann B, Konrad K, Fink GR (2007) Neurophysiological correlates of relatively enhanced local visual search in autistic adolescents. NeuroImage 35:283-291.

Mason WA, Capitanio JP, Machado CJ, Mendoza SP, Amaral DG (2006) Amygdalectomy and responsiveness to novelty in rhesus monkeys (Macaca mulatta): Generality and individual consistency of effects. Emotion 6:73-81.

Maunsell JH, Cook EP (2002) The role of attention in visual processing. Philos Trans R Soc Lond B Biol Sci 357:1063-1072.

Mazer JA, Gallant JL (2003) Goal-Related Activity in V4 during Free Viewing Visual Search: Evidence for a Ventral Stream Visual Salience Map. Neuron 40:1241-1250.

McGaugh JL (2000) Memory--a Century of Consolidation. Science 287:248-251.

McGaugh JL (2004) The amygdala modulates the consolidation of memories of emotionally arousing experiences. Annual Review of Neuroscience 27:1-28.

McPeek RM, Keller EL (2002) Superior Colliculus Activity Related to Concurrent Processing of Saccade Goals in a Visual Search Task. Journal of Neurophysiology 87:1805-1815.

McPeek RM, Keller EL (2004) Deficits in saccade target selection after inactivation of superior colliculus. Nat Neurosci 7:757-763.

Miklösi A, Polgárdi R, Topál J, Csányi V (1998) Use of experimenter-given cues in dogs. Anim Cogn 1:113-121.

Moeller S, Freiwald WA, Tsao DY (2008) Patches with Links: A Unified System for Processing Faces in the Macaque Temporal Lobe. Science 320:1355-1359.

Mormann F, Dubois J, Kornblith S, Milosavljevic M, Cerf M, Ison M, Tsuchiya N, Kraskov A, Quiroga RQ, Adolphs R, Fried I, Koch C (2011) A category-specific response to animals in the right human amygdala. Nat Neurosci 14:1247-1249.

Mormann F, Kornblith S, Quiroga RQ, Kraskov A, Cerf M, Fried I, Koch C (2008) Latency and Selectivity of Single Neurons Indicate Hierarchical Processing in the Human Medial Temporal Lobe. The Journal of Neuroscience 28:8865-8872.

Morris JS, deBonis M, Dolan RJ (2002) Human Amygdala Responses to Fearful Eyes. NeuroImage 17:214-222.

Morris JS, DeGelder B, Weiskrantz L, Dolan RJ (2001) Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. Brain 124:1241-1252.

Morris JS, Frith CD, Perrett DI, Rowland D, Young AW, Calder AJ, Dolan RJ (1996) A differential neural response in the human amygdala to fearful and happy facial expressions. Nature 383:812-815.

Morris JS, Ohman A, Dolan RJ (1998) Conscious and unconscious emotional learning in the human amygdala. Nature 393:467-470.

Morrone MC, Burr DC, Di Pietro S, Stefanelli M (1999) Cardinal directions for visual optic flow. Curr Biol 9:763-766.

Murthy A, Ray S, Shorter SM, Schall JD, Thompson KG (2009) Neural Control of Visual Search by Frontal Eye Field: Effects of Unexpected Target Displacement on Visual Selection and Saccade Preparation. Journal of Neurophysiology 101:2485-2506.

Nacewicz BM, Dalton KM, Johnstone T, Long M, McAuliff E, Oakes T, Alexander AL, Davidson RJ (2006) Amygdala volume and nonverbal social impairment in adolescent and adult males with autism. Archives of General Psychiatry 63:1417-1428.

Nakagawa H, Hongjian K (2010) Collision-Sensitive Neurons in the Optic Tectum of the Bullfrog, Rana catesbeiana. Journal of neurophysiology 104:2487-2499.

Neumann D, Spezio ML, Piven J, Adolphs R (2006) Looking you in the mouth: abnormal gaze in autism resulting from impaired top-down modulation of visual attention. Social Cognitive and Affective Neuroscience 1:194-202.

New J, Cosmides L, Tooby J (2007) Category-specific attention for animals reflects ancestral priorities, not expertise. Proceedings of the National Academy of Sciences 104:16598-16603.

New JJ, Schultz RT, Wolf J, Niehaus JL, Klin A, German TC, Scholl BJ (2010) The scope of social attention deficits in autism: Prioritized orienting to people and animals in static natural scenes. Neuropsychologia 48:51-59.

Newsome WT, Wurtz RH, Komatsu H (1988) Relation of cortical areas MT and MST to pursuit eye movements. II. Differentiation of retinal from extraretinal inputs. Journal of neurophysiology 60:604-620.

Niemann T, Lappe M, Büscher A, Hoffmann KP (1999) Ocular responses to radial optic flow and single accelerated targets in humans. Vision research 39:1359-1371.

Nummenmaa L, Hyönä J, Hietanen JK (2009) I'll Walk This Way: Eyes Reveal the Direction of Locomotion and Make Passersby Look and Go the Other Way. Psychological Science 20:1454-1458.

O'Riordan M (2000) Superior modulation of activation levels of stimulus representations does not underlie superior discrimination in autism. Cognition 77:81-96.

O'Riordan M, Plaisted K (2001) Enhanced discrimination in autism. Q J Exp Psychol A 54:961-979.

O'Riordan M, Plaisted K, Driver J, Baron-Cohen S (2001) Superior visual search in autism. Journal of Experimental Psychology: Human Perception and Performance 27:719–730.

O'Riordan MA (2004) Superior Visual Search in Adults with Autism. Autism 8:229-248.

Ochsner KN, Bunge SA, Gross JJ, Gabrieli JDE (2002) Rethinking Feelings: An fMRI Study of the Cognitive Regulation of Emotion. Journal of Cognitive Neuroscience 14:1215-1229.

Ogawa T, Komatsu H (2004) Target Selection in Area V4 during a Multidimensional Visual Search Task. The Journal of Neuroscience 24:6371-6382.

Ohman A (2002) Automaticity and the Amygdala: Nonconscious Responses to Emotional Faces. Current Directions in Psychological Science 11:62-66.

Olausson H, Lamarre Y, Backlund H, Morin C, Wallin BG, Starck G, Ekholm S, Strigo I, Worsley K, Vallbo AB, Bushnell MC (2002) Unmyelinated tactile afferents signal touch and project to insular cortex. Nat Neurosci 5:900-904.

Olsson P (2007) Real-time and offline filters for eye tracking. vol. Msc: KTH Royal Institute of Technology.

Oosterhof NN, Todorov A (2008) The functional basis of face evaluation. Proceedings of the National Academy of Sciences 105:11087-11092.

Orban GA (1992) First-order analysis of optical flow in monkey brain. Proc Natl Acad Sci USA 89:2595-2599.

Osterling J, Dawson G (1994) Early recognition of children with autism: A study of first birthday home videotapes. J Autism Dev Disord 24:247-257.

Oya H, Kawasaki H, Dahdaleh NS, Wemmie JA, Howard Iii MA (2009) Stereotactic Atlas-Based Depth Electrode Localization in the Human Amygdala. Stereotactic and Functional Neurosurgery 87:219-228.

Parvizi J, Jacques C, Foster BL, Withoft N, Rangarajan V, Weiner KS, Grill-Spector K (2012) Electrical Stimulation of Human Fusiform Face-Selective Regions Distorts Face Perception. The Journal of Neuroscience 32:14915-14920.

Paton JJ, Belova MA, Morrison SE, Salzman CD (2006) The primate amygdala represents the positive and negative value of visual stimuli during learning. Nature 439:865-870.

Paul L, Corsello C, Tranel D, Adolphs R (2010) Does bilateral damage to the human amygdala produce autistic symptoms? Journal of Neurodevelopmental Disorders 2:165-173.

Peck CJ, Lau B, Salzman CD (2013) The primate amygdala combines information about space and value. Nat Neurosci 16:340-348.

Peignot P, Anderson JR (1999) Use of experimenter-given manual and facial cues by gorillas (Gorilla gorilla) in an object-choice task. Journal of Comparative Psychology 113:253-260.

Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10:437-442.

Pelphrey K, Sasson N, Reznick JS, Paul G, Goldman B, Piven J (2002) Visual Scanning of Faces in Autism. J Autism Dev Disord 32:249-261.

Pelphrey KA, Singerman JD, Allison T, McCarthy G (2003) Brain activation evoked by perception of gaze shifts: the influence of context. Neuropsychologia 41:156-170.

Pelphrey KA, Viola RJ, McCarthy G (2004) When Strangers Pass: Processing of Mutual and Averted Social Gaze in the Superior Temporal Sulcus. Psychological Science 15:598-603.

Perrett D, Smith P, Potter D, Mistlin A, Head A, Milner A, Jeeves M (1984) Neurones responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception. Hum Neurobiol 3:197-208.

Perrett DI, Smith PAJ, Potter DD, Mistlin AJ, Head AS, Milner AD, Jeeves MA (1985) Visual Cells in the Temporal Cortex Sensitive to Face View and Gaze Direction.

Proceedings of the Royal Society of London Series B, Biological Sciences 223:293-317.

Pessoa L, Adolphs R (2010) Emotion processing and the amygdala: from a 'low road' to 'many roads' of evaluating biological significance. Nat Rev Neurosci 11:773-783.

Pessoa L, Japee S, Sturman D, Ungerleider LG (2006) Target Visibility and Visual Awareness Modulate Amygdala Responses to Fearful Faces. Cerebral Cortex 16:366-375.

Pessoa L, McKenna M, Gutierrez E, Ungerleider LG (2002) Neural Processing of Emotional Faces Requires Attention. Proceedings of the National Academy of Sciences of the United States of America 99:11458-11463.

Phelps EA, LeDoux JE (2005) Contributions of the Amygdala to Emotion Processing: From Animal Models to Human Behavior. Neuron 48:175-187.

Philip RCM, Dauvermann MR, Whalley HC, Baynham K, Lawrie SM, Stanfield AC (2012) A systematic review and meta-analysis of the fMRI investigation of autism spectrum disorders. Neuroscience & Biobehavioral Reviews 36:901-942.

Philip RCM, Whalley HC, Stanfield AC, Sprengelmeyer R, Santos IM, Young AW, Atkinson AP, Calder AJ, Johnstone EC, Lawrie SM, Hall J (2010) Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. Psychological Medicine 40:1919-1929.

Piech RM, McHugo M, Smith SD, Dukic MS, Van Der Meer J, Abou-Khalil B, Most SB, Zald DH (2011) Attentional capture by emotional stimuli is preserved in patients with amygdala lesions. Neuropsychologia 49:3314-3319.

Piech RM, McHugo M, Smith SD, Dukic MS, Van Der Meer J, Abou-Khalil B, Zald DH (2010) Fear-enhanced visual search persists after amygdala lesions. Neuropsychologia 48:3430-3435.

Piven J, Arndt S, Bailey J, Havercam S, Andreasen N, Palmer P (1995) An MRI study of brain size in autism. American Journal of Psychiatry 152:1145-1149.

Plaisted K, O'Riordan M, Baron-Cohen S (1998) Enhanced visual search for a conjunctive target in autism: a research note. Journal of Child Psychology and Psychiatry 39:777–783.

Povinelli DJ, Eddy TJ (1996) Chimpanzees: Joint Visual Attention. Psychological Science 7:129-135.

Puce A, Allison T, Bentin S, Gore JC, McCarthy G (1998) Temporal Cortex Activation in Humans Viewing Eye and Mouth Movements. The Journal of Neuroscience 18:2188-2199.

Purcell BA, Weigand PK, Schall JD (2012) Supplementary Eye Field during Visual Search: Salience, Cognitive Control, and Performance Monitoring. The Journal of Neuroscience 32:10273-10285.

Regan D, Beverley KI (1978) Illusory motion in depth: aftereffect of adaptation to changing size. Vision research 18:209-212.

Regan D, Beverley KI (1979) Binocular and monocular stimuli for motion in depth: changing-disparity and changing-size feed the same motion-in-depth stage. Vision research 19:1331-1340.

Regan D, Gray R (2000) Visually guided collision avoidance and collision achievement. Trends in cognitive sciences 4:99-107.

Rensink RA, O'Regan JK, Clark JJ (1997) To see or not to see: The need for attention to perceive changes in scenes. Psychol Sci 8:1217-1219.

Ristic J, Friesen C, Kingstone A (2002) Are eyes special? It depends on how you look at it. Psychonomic Bulletin & Review 9:507-513.

Rolls E (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 ET (1992) Neurophysiology and functions of the primate amygdala. In: The Amygdala: Neurobiological Aspects of Emotion, Memory and Mental Dysfunction (Aggleton, J. P., ed), p 143−165 New York: Wiley-Liss.

Rolls ET (2010) The affective and cognitive processing of touch, oral texture, and temperature in the brain. Neuroscience & Biobehavioral Reviews 34:237-245.

Rolls ET, O'Doherty J, Kringelbach ML, Francis S, Bowtell R, McGlone F (2003) Representations of Pleasant and Painful Touch in the Human Orbitofrontal and Cingulate Cortices. Cerebral Cortex 13:308-317.

Roozendaal B, Castello NA, Vedana G, Barsegyan A, McGaugh JL (2008) Noradrenergic activation of the basolateral amygdala modulates consolidation of object recognition memory. Neurobiology of Learning and Memory 90:576-579.

Rosset D, Santos A, Da Fonseca D, Rondan C, Poinson F, Deruelle C (2011) More than just another face in the crowd: Evidence for an angry superiority effect in children with and without autism. Research In Autism Spectrum Disorders 5:949-956.

Rutishauser U, Koch C (2007) Probabilistic modeling of eye movement data during conjunction search via feature-based attention. Journal of Vision 7.

Rutishauser U, Schuman EM, Mamelak AN (2008) Activity of human hippocampal and amygdala neurons during retrieval of declarative memories. Proceedings of the National Academy of Sciences 105:329-334.

Rutishauser U, Tudusciuc O, Neumann D, Mamelak AN, Heller AC, Ross IB, Philpott L, Sutherling WW, Adolphs R (2011) Single-Unit Responses Selective for Whole Faces in the Human Amygdala. Current biology : CB 21:1654-1660.

Rutishauser U, Tudusciuc O, Wang S, Mamelak AN, Ross IB, Adolphs R (2013) Single-Neuron Correlates of Atypical Face Processing in Autism. Neuron 80:887-899.

Saito H, Yukie M, Tanaka K, Hikosaka K, Fukada Y, Iwai E (1986) Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. J Neurosci 6:145-157.

Sander D, Grandjean D, Pourtois G, Schwartz S, Seghier ML, Scherer KR, Vuilleumier P (2005) Emotion and attention interactions in social cognition: Brain regions involved in processing anger prosody. NeuroImage 28:848-858.

Sasson N (2006) The Development of Face Processing in Autism. J Autism Dev Disord 36:381-394.

Sasson N, Dichter G, Bodfish J (2012) Affective Responses by Adults with Autism Are Reduced to Social Images but Elevated to Images Related to Circumscribed Interests. PLoS ONE 7:e42457.

Sasson N, Tsuchiya N, Hurley R, Couture SM, Penn DL, Adolphs R, Piven J (2007) Orienting to social stimuli differentiates social cognitive impairment in autism and schizophrenia. Neuropsychologia 45:2580-2588.

Sasson NJ, Elison JT, Turner-Brown LM, Dichter GS, Bodfish JW (2011) Brief Report: Circumscribed Attention in Young Children with Autism. J Autism Dev Disord 41:242-247.

Sasson NJ, Turner-Brown LM, Holtzclaw TN, Lam KSL, Bodfish JW (2008) Children with autism demonstrate circumscribed attention during passive viewing of complex social and nonsocial picture arrays. Autism Research 1:31-42.

Sato W, Okada T, Toichi M (2007) Attentional shift by gaze is triggered without awareness. Experimental Brain Research 183:87-94.

Sato W, Yoshikawa S, Kochiyama T, Matsumura M (2004) The amygdala processes the emotional significance of facial expressions: an fMRI investigation using the interaction between expression and face direction. NeuroImage 22:1006-1013.

Scaife M, Bruner JS (1975) The capacity for joint visual attention in the infant. Nature 253:265-266.

Schaafsma SJ, Duysens J (1996) Neurons in the ventral intraparietal area of awake macaque monkey closely resemble neurons in the dorsal part of the medial superior temporal area in their responses to optic flow patterns. Journal of neurophysiology 76:4056-4068.

Schaefer SM, Jackson DC, Davidson RJ, Aguirre GK, Kimberg DY, Thompson-Schill SL (2002) Modulation of Amygdalar Activity by the Conscious Regulation of Negative Emotion. Journal of Cognitive Neuroscience 14:913-921.

Scheller E, Buchel C, Gamer M (2012) Diagnostic Features of Emotional Expressions Are Processed Preferentially. PLoS ONE 7:e41792.

Scheumann M, Call J (2004) The use of experimenter-given cues by South African fur seals (Arctocephalus pusillus). Anim Cogn 7:224-230.

Schiff W (1965) Perception of impending collision: A study of visually directed avoidant behavior. Psychological Monographs: General & Applied 79:1-26.

Schiff W, Caviness JA, Gibson JJ (1962) Persistent Fear Responses in Rhesus Monkeys to the Optical Stimulus of "Looming". Science 136:982-983.

Schirmer A, Teh KS, Wang S, Vijayakumar R, Ching A, Nithianantham D, Escoffier N, Cheok AD (2011) Squeeze me, but don't tease me: Human and mechanical touch enhance visual attention and emotion discrimination. Social Neuroscience 6:219-230.

Schrater PR, Knill DC, Simoncelli EP (2001) Perceiving visual expansion without optic flow. Nature 410:816-819.

Schumann CM, Amaral DG (2006) Stereological Analysis of Amygdala Neuron Number in Autism. The Journal of Neuroscience 26:7674-7679.

Schumann CM, Hamstra J, Goodlin-Jones BL, Lotspeich LJ, Kwon H, Buonocore MH, Lammers CR, Reiss AL, Amaral DG (2004) The Amygdala Is Enlarged in Children But Not Adolescents with Autism; the Hippocampus Is Enlarged at All Ages. The Journal of Neuroscience 24:6392-6401.

Sheinberg DL, Logothetis NK (2001) Noticing Familiar Objects in Real World Scenes: The Role of Temporal Cortical Neurons in Natural Vision. The Journal of Neuroscience 21:1340-1350.

Shen K, Paré M (2007) Neuronal activity in superior colliculus signals both stimulus identity and saccade goals during visual conjunction search. Journal of Vision 7.

Shen K, Paré M (2014) Predictive Saccade Target Selection in Superior Colliculus during Visual Search. The Journal of Neuroscience 34:5640-5648.

Sidaway B, McNitt-Gray J, Davis G (1989) Visual Timing of Muscle Preactivation in Preparation for Landing. Ecological Psychology 1:253 - 264.

Siebert M, Markowitsch HJ, Bartel P (2003) Amygdala, affect and cognition: evidence from 10 patients with Urbach-Wiethe disease. Brain 126:2627-2637.

Simons DJ, Rensink RA (2005) Change blindness: past, present, and future. Trends in cognitive sciences 9:16-20.

Smith ML, Cottrell GW, Gosselin F, Schyns PG (2005) Transmitting and Decoding Facial Expressions. Psychological Science 16:184-189.

South M, Ozonoff S, McMahon W (2005) Repetitive Behavior Profiles in Asperger Syndrome and High-Functioning Autism. J Autism Dev Disord 35:145-158.

Spezio ML, Adolphs R, Hurley RSE, Piven J (2007a) Abnormal Use of Facial Information in High-Functioning Autism. J Autism Dev Disord 37:929-939.

Spezio ML, Adolphs R, Hurley RSE, Piven J (2007b) Analysis of face gaze in autism using "Bubbles". Neuropsychologia 45:144-151.

Sun H, Frost BJ (1998) Computation of different optical variables of looming objects in pigeon nucleus rotundus neurons. Nature neuroscience 1:296-303.

Takeuchi T (1997) Visual search of expansion and contraction. Vision research 37:2083-2090.

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

Tanaka K (1997) Mechanisms of visual object recognition: monkey and human studies. Current Opinion in Neurobiology 7:523-529.

Tanaka K, Saito H (1989) Analysis of motion of the visual field by direction, expansion/ contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. Journal of neurophysiology 62:626-641.

Terburg D, Morgan BE, Montoya ER, Hooge IT, Thornton HB, Hariri AR, Panksepp J, Stein DJ, van Honk J (2012) Hypervigilance for fear after basolateral amygdala damage in humans. Transl Psychiatry 2:e115.

Thomas NWD, Paré M (2007) Temporal Processing of Saccade Targets in Parietal Cortex Area LIP During Visual Search. Journal of Neurophysiology 97:942-947.

Thompson KG, Bichot NP (2005) A visual salience map in the primate frontal eye field. In: Progress in Brain Research, vol. Volume 147 (J. van Pelt, M. K. C. N. L. A. v. O. G. J. A. R. and Roelfsema, P. R., eds), pp 249-262: Elsevier.

Thompson KG, Hanes DP, Bichot NP, Schall JD (1996) Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. Journal of Neurophysiology 76:4040-4055.

Todd RM, Talmi D, Schmitz TW, Susskind J, Anderson AK (2012) Psychophysical and Neural Evidence for Emotion-Enhanced Perceptual Vividness. The Journal of Neuroscience 32:11201-11212.

Todorov A, Mandisodza AN, Goren A, Hall CC (2005) Inferences of Competence from Faces Predict Election Outcomes. Science 308:1623-1626.

Tolias AS, Moore T, Smirnakis SM, Tehovnik EJ, Siapas AG, Schiller PH (2001) Eye Movements Modulate Visual Receptive Fields of V4 Neurons. Neuron 29:757-767.

Treisman A (1988) Features and Objects: The Fourteenth Bartlett Memorial Lecture. The Quarterly Journal of Experimental Psychology Section A 40:201-237.

Treisman A (1998) Feature binding, attention and object perception. Philosophical Transactions of the Royal Society of London Series B: Biological Sciences 353:1295-1306.

Tronick E (1967) Approach response of domestic chicks to an optical display. J Comp Physiol Psychol 64:529-531.

Tsao DY, Freiwald WA, Tootell RBH, Livingstone MS (2006) A Cortical Region Consisting Entirely of Face-Selective Cells. Science 311:670-674.

Tsuchiya N, Moradi F, Felsen C, Yamazaki M, Adolphs R (2009) Intact rapid detection of fearful faces in the absence of the amygdala. Nat Neurosci 12:1224-1225.

Tye KM, Stuber GD, de Ridder B, Bonci A, Janak PH (2008) Rapid strengthening of thalamo-amygdala synapses mediates cue-reward learning. Nature 453:1253-1257.

Vallortigara G, Rogers LJ (2005) Survival with an asymmetrical brain: Advantages and disadvantages of cerebral lateralization. Behav Brain Sci 28:575–589.

van Boxtel JJA, Lu H (2011) Visual search by action category. Journal of Vision 11.

van Boxtel JJA, Lu H (2012) Signature Movements Lead to Efficient Search for Threatening Actions. PLoS ONE 7:e37085.

von Muhlenen A, Lleras A (2007) No-onset looming motion guides spatial attention. Journal of experimental psychology 33:1297-1310.

Vrticka P, Sander D, Vuilleumier P (2013) Lateralized Interactive Social Content and Valence Processing within the Human Amygdala. Frontiers in Human Neuroscience 6.

Vuilleumier P (2002) Perceived gaze direction in faces and spatial attention: a study in patients with parietal damage and unilateral neglect. Neuropsychologia 40:1013-1026.

Vuilleumier P (2005) How brains beware: neural mechanisms of emotional attention. Trends in Cognitive Sciences 9:585-594.

Vuilleumier P, Armony JL, Driver J, Dolan RJ (2001) Effects of Attention and Emotion on Face Processing in the Human Brain: An Event-Related fMRI Study. Neuron 30:829-841.

Vuilleumier P, Richardson MP, Armony JL, Driver J, Dolan RJ (2004) Distant influences of amygdala lesion on visual cortical activation during emotional face processing. Nat Neurosci 7:1271-1278.

Vuilleumier P, Schwartz S (2001) Emotional facial expressions capture attention. Neurology 56:153-158.

Wagner H (1982) Flow-field variables trigger landing in flies. Nature 297:147-148.

Wall MB, Smith AT (2008) The representation of egomotion in the human brain. Curr Biol 18:191-194.

Wallace G, Case L, Harms M, Silvers J, Kenworthy L, Martin A (2011) Diminished Sensitivity to Sad Facial Expressions in High Functioning Autism Spectrum Disorders is Associated with Symptomatology and Adaptive Functioning. J Autism Dev Disord 41:1475-1486.

Wang S, Fukuchi M, Koch C, Tsuchiya N (2012a) Spatial Attention Is Attracted in a Sustained Fashion toward Singular Points in the Optic Flow. PLoS ONE 7:e41040.

Wang S, Krajbich I, Adolphs R, Tsuchiya N (2012b) The role of risk aversion in non-conscious decision-making. Frontiers in Psychology 3.

Wang S, Tsuchiya N, New J, Hurlemann R, Adolphs R (2014a) Preferential attention to animals and people is independent of the amygdala. Social Cognitive and Affective Neuroscience.

Wang S, Tsuchiya N, New J, Hurlemann R, Adolphs R (2014b) Preferential attention to animals and people is independent of the amygdala. Social Cognitive and Affective Neuroscience in press.

Wang S, Tudusciuc O, Mamelak A, Ross I, Adolphs R, Rutishauser U (2014c) Neurons in the human amygdala selective for perceived emotion. Proc Natl Acad Sci USA.

Wang S, Xu J, Jiang M, Zhao Q, Hurlemann R, Adolphs R (2014d) Autism spectrum disorder, but not amygdala lesions, impairs social attention in visual search. In: The Social & Affective Neuroscience Society (SANS) Conference Denver.

Wang Y, Frost BJ (1992) Time to collision is signalled by neurons in the nucleus rotundus of pigeons. Nature 356:236-238.

Watson KK, Platt ML (2012) Social Signals in Primate Orbitofrontal Cortex. Current biology : CB 22:2268-2273.

Whalen PJ, Kagan J, Cook RG, Davis FC, Kim H, Polis S, McLaren DG, Somerville LH, McLean AA, Maxwell JS, Johnstone T (2004) Human Amygdala Responsivity to Masked Fearful Eye Whites. Science 306:2061.

Whalen PJ, Rauch SL, Etcoff NL, McInerney SC, Lee MB, Jenike MA (1998) Masked Presentations of Emotional Facial Expressions Modulate Amygdala Activity without Explicit Knowledge. The Journal of Neuroscience 18:411-418.

Willis J, Todorov A (2006) First Impressions: Making Up Your Mind After a 100-Ms Exposure to a Face. Psychological Science 17:592-598.

Willis ML, Palermo R, Burke D, McGrillen K, Miller L (2010) Orbitofrontal cortex lesions result in abnormal social judgements to emotional faces. Neuropsychologia 48:2182-2187.

Winston JS, Strange BA, O'Doherty J, Dolan RJ (2002) Automatic and intentional brain responses during evaluation of trustworthiness of faces. Nat Neurosci 5:277-283.

Wise RA (2004) Dopamine, learning and motivation. Nat Rev Neurosci 5:483-494.

Wolfe JM (1998) What Can 1 Million Trials Tell Us About Visual Search? Psychological Science 9:33-39.

Wolfe JM (2012) The Rules of Guidance in Visual Search. In: Perception and Machine Intelligence, vol. 7143 (Kundu, M. K. et al., eds), pp 1-10: Springer Berlin Heidelberg.

Wolfe JM, Horowitz TS (2004) What attributes guide the deployment of visual attention and how do they do it? Nat Rev Neurosci 5:495-501.

Xu J, Jiang M, Wang S, Kankanhalli MS, Zhao Q (2014) Predicting human gaze beyond pixels. Journal of Vision 14.

Yamamoto K, Nakata M, Nakagawa H (2003) Input and Output Characteristics of Collision Avoidance Behavior in the Frog Rana catesbeiana. Brain, Behavior and Evolution 62:201-211.

Yang E, McHugo M, Dukic M, Blake R, Zald D (2012a) Advantage of fearful faces in breaking interocular suppression is preserved after amygdala lesions. Vision Sciences Society Annual Meeting

Yang J, Bellgowan PSF, Martin A (2012b) Threat, domain-specificity and the human amygdala. Neuropsychologia 50:2566-2572.

Yarbus AL (1967) Eye Movements and Vision. New York: Plenum Press.

Yilmaz EH, Warren WH (1995) Visual control of braking: A test of the !·π´ hypothesis. Journal of Experimental Psychology: Human Perception and Performance 21:996-1014.

Young AW, Bruce V (1991) Perceptual categories and the computation of "grandmother". European Journal of Cognitive Psychology 3:5-49.

Zhang B, Noble PL, Winslow JT, Pine DS, Nelson EE (2012) Amygdala volume predicts patterns of eye fixation in rhesus monkeys. Behavioural Brain Research 229:433-437.

Zhang T, Heuer HW, Britten KH (2004) Parietal area VIP neuronal responses to heading stimuli are encoded in head-centered coordinates. Neuron 42:993-1001.

Zhao Q, Koch C (2011) Learning a saliency map using fixated locations in natural scenes. Journal of Vision 11.

Zhao Q, Koch C (2012) Learning visual saliency by combining feature maps in a nonlinear manner using AdaBoost. Journal of Vision 12.