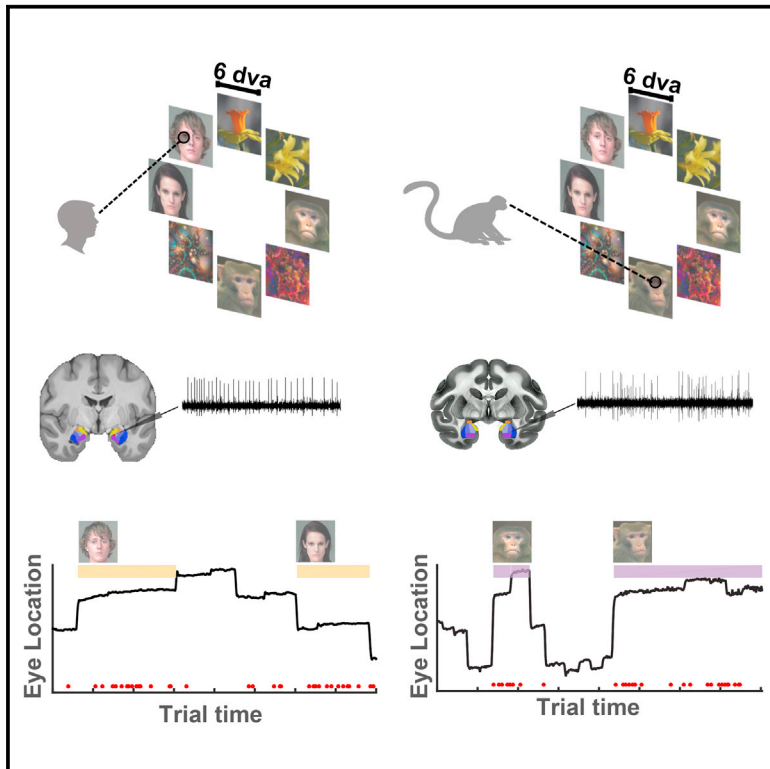# Fixations Gate Species-Specific Responses to Free Viewing of Faces in the Human and Macaque Amygdala

Graphical Abstract

Authors

Juri Minxha, Clayton Mosher,
Jeremiah K. Morrow, Adam N. Mamelak,
Ralph Adolphs, Katalin M. Gothard,
Ueli Rutishauser

Correspondence

ueli.rutishauser@cshs.org

In Brief

The role of the amygdala in social perception remains unknown. Minxha et al. reveal that the response of face-selective amygdala neurons in macaques and humans is gated by fixations and the species of the fixated face. This shows that primate amygdala responses to social stimuli are gated by attention.

Highlights

- Investigated the role of the human and macaque amygdala in face processing

- Response of face-selective neurons was gated by fixations during free viewing

- Response latencies were shortest when fixating on a face of the subject's own species

- Covert attention gated human amygdala responses to faces

CellPress

Fixations Gate Species-Specific Responses to Free Viewing of Faces in the Human and Macaque Amygdala

Juri Minxha,[1] Clayton Mosher,[2] Jeremiah K. Morrow,[2] Adam N. Mamelak,[3] Ralph Adolphs,[1,4,5] Katalin M. Gothard,[2] and Ueli Rutishauser[3,4,6,7,*]

[1]Computation and Neural Systems Program, California Institute of Technology, Pasadena, CA 90025, USA
[2]Department of Physiology, College of Medicine, University of Arizona, Tucson, AZ 85724, USA
[3]Department of Neurosurgery, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA
[4]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 90025, USA
[5]Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 90025, USA
[6]Department of Neurology, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA
[7]Lead Contact
*Correspondence: ueli.rutishauser@cshs.org
 http://dx.doi.org/10.1016/j.celrep.2016.12.083

SUMMARY

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

INTRODUCTION

Faces are important stimuli for primate social behavior. Humans and macaques share a homologous set of cortical regions specialized for processing faces (Tsao et al., 2008), and in macaques these "face patches" contain neurons almost entirely selective for faces (Tsao et al., 2006). Together, face patches constitute an interconnected system for constructing face representations from facial features (Freiwald et al., 2009; Moeller et al., 2008). A key unanswered question is how this cortical representation of faces guides social behavior. The amygdala is a

key structure in such subsequent processing: it is reciprocally connected with the cortical face patches (Grimaldi et al., 2016), contains a large proportion of neurons selective for faces (Gothard et al., 2007; Mosher et al., 2014; Rutishauser et al., 2011; Sanghera et al., 1979), and is critical for primate social behavior.

The amygdala processes stimuli with ecological significance, including not only social stimuli such as faces, but also conditioned and unconditioned rewards and punishments (Adolphs, 2010; Paton et al., 2006). Face-selective responses are prominent in the amygdala of both humans and monkeys (Fried et al., 1997; Gothard et al., 2007; Mosher et al., 2014; Rutishauser et al., 2011; Sanghera et al., 1979), as would be expected from the highly processed visual inputs the amygdala receives from the multiple areas where face-selective cells have been discovered (Bruce et al., 1981; Desimone, 1991; Gross et al., 1972; Perrett et al., 1982; Rolls, 1984). This picture suggests a limited contribution of the amygdala to face processing: all its face selectivity might be explained by the inputs from face-selective cortical regions. Also, it is commonly believed that the large receptive fields of the neurons that provide input to the amygdala would result in visual receptive fields of amygdala neurons that are not spatially restricted (Barraclough and Perrett, 2011; Boussaoud et al., 1991; Gross et al., 1969). It has been proposed that the amygdala responds to faces even when they are not attended (Vuilleumier et al., 2001) or consciously perceived (Tamietto and de Gelder, 2010). This view of the amygdala's function fits with a long-standing debate about whether the amygdala mediates rapid automatic and relatively coarse detection of significant stimuli through a route of subcortical inputs (Cauchoix and Crouzet, 2013). These views challenge observations that human amygdala neurons show exceedingly long visual response latencies (Mormann et al., 2008; Rutishauser et al., 2015a), and that fMRI activation of the amygdala appears to require visual attention (Pessoa et al., 2002). In the absence of comparative studies using the same stimuli and the same paradigm in both species, it is nearly impossible to determine whether these are differences between the two species or rather different experimental conditions. Resolving these disparate conclusions thus

CrossMark

requires a more comprehensive investigation, which no single study has yet accomplished: assessing amygdala responses to faces at the single-unit level across both monkeys and humans, and investigating selectivity and response latency in relation to fixation onset during free viewing with concurrent eye tracking.

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

In the context of natural vision, the amygdala is of particular interest because amygdala lesions are known to interfere with the efficient visual exploration of faces (Adolphs et al., 2005) and amygdala neurons respond to dynamic social signals such as eye contact (Mosher et al., 2014). Moreover, lesions of the amygdala produce a complex constellation of impairments in social behavior (Adolphs et al., 1994, 1998). The amygdala is thus a prime candidate for mediating between the perceptual representations of faces in the cortical face patch system, and the mediation of social behaviors based on such perception. Elucidating this role, however, requires both a more naturalistic presentation of stimuli, and a better quantification of how they are attended. Here, we achieved both these imperatives by allowing subjects to freely view a complex array of images that competed for attention while we monitored eye movements. Our focus was on the category selectivity of amygdala neurons during natural vision, with specific emphasis on the potential category selectivity for conspecific and heterospecific faces. We focused on faces not only because of their patent ecological relevance but also because this is the visual category of stimuli consistently explaining the largest proportion of variance of the responses of amygdala neurons (Gothard et al., 2007; Rutishauser et al., 2011).

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

RESULTS

Task and Behavior

We tested a total of 12 human epilepsy patients (28 sessions, 50 ± 1 trials per session, $\pm$SD) and 3 healthy macaques (16 sessions, 113 ± 13 trials per session, $\pm$SD). Subjects freely viewed complex visual stimuli (Figure 1A). Each stimulus consisted of a circular array of eight images chosen at random from two face categories (human and monkey faces) and two non-face categories (either flowers and fractals or fruits and cars, depending on the version of the task performed). The non-face categories were later pooled for analysis as "distractors" (Distractor Group #1 contained cars and fractals; and Distractor Group #2 contained fruits and flowers). Each image array was displayed for 3–4 s, and subjects were free to view any location. Although stimuli were identical for the two species, each necessitated slightly different task conditions (see below for behavioral controls). Macaques received a fixed amount of reward after conclusion of a trial if they maintained their gaze position anywhere on the screen during the entire stimulus period. Trials were aborted and no reward was given if a monkey moved its gaze off the screen within the first 3 s of stimulus onset. This achieved attention to the image array while encouraging free exploration. Human subjects were instructed to freely view the stimuli for a later memory test (Figure 1A). To verify that our two tasks produced largely comparable fixation behaviors in the two species, we compared the scan paths used by humans and monkeys to explore the image arrays (Figure 1B shows an example).

To further ensure that our memory task does not introduce fixation preference biases, we also asked two groups of healthy human control subjects to perform the identical task (Figure S5): one with the same instructions as the patients ("memory controls") and one that did not know that a memory test would follow ("free-viewing control"). Patients had good recognition memory (average across n = 14 sessions, 70%, p = 0.001 versus chance, one-sided binomial test), showing that they attended to the stimuli. The patients' performance was somewhat lower than that of the memory group but not the free-viewing group (Figure S5D,
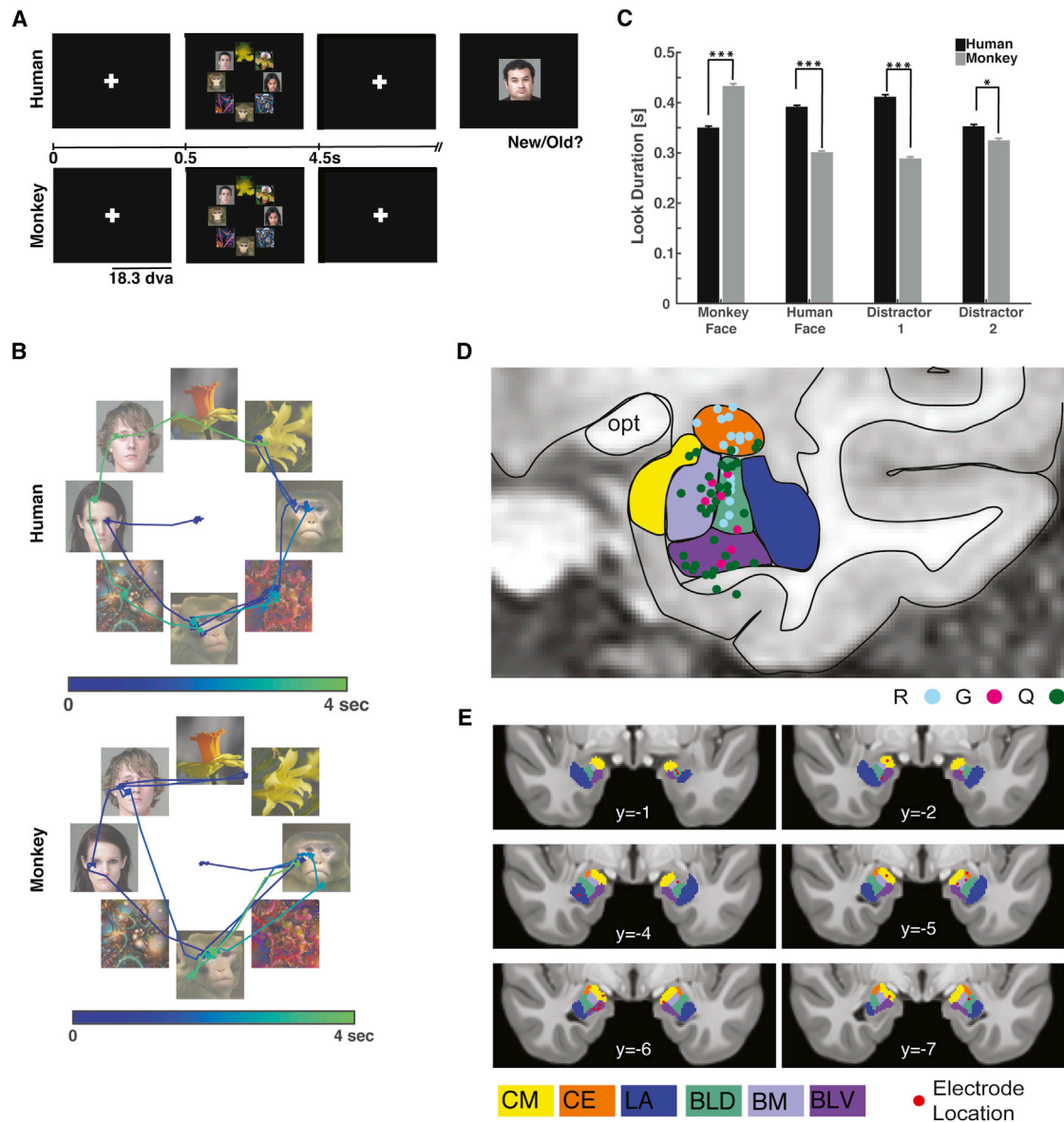
Figure 1. Task, Behavior, and Recording Locations

(A) Tasks performed by human and monkey subjects. Task version #1 is shown.

(B) Example scan paths from a human (top) and monkey (bottom) viewing the same stimuli. Dots are fixations, and lines are saccades. The first saccade (starting at the center) targets the face of a conspecific. Trial time is encoded by color.

(C) Look duration ("dwell time") on each stimulus category for humans (black) and monkeys (gray). Monkeys fixated longer on conspecific faces and on flowers, whereas humans fixated longer on conspecific faces and on fractals (three stars indicates $p < 0.001$ and one star indicates $p < 0.05$, two-tailed t test). Error bars indicate SEM.

(D and E) Recording locations. Amygdala nuclei are indicated in color. (D) Recording sites in the three monkey subjects (R, G, Q in different color dots) collapsed onto a representative coronal section. (E) Human recording sites (red dots) in MNI152 space. Abbreviations: LA, lateral nucleus; BLD, dorsal basolateral; BM, basomedial; CE, central; CM, cortical and medial nuclei; BLV, ventral basolateral.

$t(19) = 2.845$, $p = 0.01$, and $t(19) = 1.467$, $p = 0.16$ respectively). Crucially, the probability that the first fixation landed on a human face was not influenced by task instruction and did not differ between any of the groups ($t(19) = 0.8538$, $p = 0.40$, and $t(19) = 0.3013$, $p = 0.77$, respectively). Subjects in all three groups explored all stimuli extensively regardless of task instructions

(Figures S5A and S5B), confirming that informing subjects of a subsequent memory test without further specific information does not affect fixation preferences for faces.

The majority of all fixations landed on one of the eight images in both human patients and macaques (84% and 89%, respectively; Figure S1F) and subjects looked longer and earlier at faces

of conspecifics compared to faces of heterospecifics (Figure 1C, see legend for statistics). A reliable viewing pattern for both species was that the probability of looking first at a conspecific face was higher than that of first looking at a heterospecific face (32% versus 24%, t(54) = 3.63, p = 0.0006, in humans, and 32% versus 20%, t(30) = 2.77, p = 0.01, in monkeys, paired t tests; chance is 25%; Figures S1A and S1B). Also, in both species, some sequences of fixations (i.e., human followed by monkey face) were more likely than others (Figures S1C and S1D), which shows that fixation location was influenced by image content throughout. This suggests that even before launching a sequence of exploratory eye movements, conspecific faces competed successfully for fixation priority in both humans and monkeys. Together, this argues that the location of the faces on the screen was attended to by, and influenced the fixation patterns of, both humans and monkeys. We tested this mechanistic hypothesis further with a covert-attention task in the human subjects that is described later below.

Electrophysiology

We isolated 422 and 195 putative single units from the human and macaque amygdala, respectively (on average, 19 and 12 per session, respectively). We only analyzed cells with a mean firing rate of >0.5 Hz (290 and 148 units, respectively; mean firing rates, 2.28 and 5.6 Hz; range, 0.5–26.4 and 0.5–72.8 Hz). Throughout the manuscript, we use the terms neuron or cell to refer to a putative single unit, and we used only units satisfying multiple conservative criteria (see Experimental Procedures for details). To facilitate direct comparison between species, all spike detection, sorting, and data analysis were performed using the same methods for all recordings from both species (see Figure S4 for electrophysiological properties of cells in both species).

Fixation Target-Sensitive Neuronal Responses

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

We further characterized the category selectivity of fixation target-sensitive responses (Figures 2C and 2D). We first classified each fixation target-sensitive cell according to the category to which it responds most strongly (highest firing rate) at the point of time at which neurons provided most category information (see Experimental Procedures and Figure S6). The majority of fixation target-sensitive neurons preferred faces of conspecifics: 49% (n = 42/85) and 39% (n = 24/61) in humans and macaques, respectively (Figure 3A). A smaller proportion preferred faces of heterospecifics (that is, faces of the opposite species): 32% (n = 27/85) and 18% (n = 11/61) in humans and macaques,

respectively. Together, about 71% of all fixation-sensitive neurons preferred faces (Figure 3A, 81% in humans and 58% in macaques). In contrast, only 19% (n = 16/85, in humans) and 42% (n = 26/61, in macaques) were sensitive to the non-face categories we used (flowers, fractals, fruits, and cars). Because subjects were free to look at any of the images, we had no way of ensuring that they would sample uniformly from the different image categories. In order to ensure that the tuning of the cells was not confounded by the number of fixations on each category, we carried out a control analysis in which we selected cells after equalizing the number of fixations for each image category by subsampling. This revealed similar proportions: 50% ± 3% and 40% ± 2.2% of fixation-sensitive neurons preferred faces of conspecifics, respectively (±SD across 100 bootstraps). Thus, most primate amygdala fixation target-sensitive neurons responded to faces rather than nonsocial object categories, and there were two groups of such neurons: those that increase their firing rate whenever fixations are made onto faces of conspecifics, and those that increase their firing rate only when looking at faces of heterospecifics (Figures 3B–3E show the average response of all four types of face cells). For clarity, we label each type of face-sensitive cell with a capital letter to signify the species in which the cell was recorded (H or M) and with a lowercase letter to signify the tuning of the cell to human or monkey faces (h or m).

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

We found that neurons in the human amygdala differentiated between, on average, 3.47 ± 0.1 pairs of categories (out of 6), whereas neurons in the monkey amygdala differentiated between 4.15 ± 0.2 pairs (Figures 4A and 4B). Thus, neurons in humans differentiated between significantly fewer (p = 0.002, two-sample Kolmogorov-Smirnov [KS] test) pairs of categories compared to neurons in macaques. Similarly, the depth of selectivity (DOS), an index of the narrowness of tuning to a specific category, of all human neurons was larger than that of macaque neurons (0.54 ± 0.02 versus 0.43 ± 0.03, p = 0.0003, two-sample KS test, Figure 4C), but was at the same time significantly lower than 1 (p < 1e-37, two-sample KS test). Note that a DOS value of 1 means exclusive tuning to one stimulus, but no response to all other stimuli; in contrast, a DOS value of 0 implies no preferred tuning. We observed DOS values of 0.18–0.87 in humans and 0.11–0.90 in macaques (Figures 4C and 4D). DOS values for neurons recorded in humans were significantly larger than those for neurons recorded in monkeys (Figure 4E), a result compatible with the sparser response profile over categories as shown in Figure 4B. Although the DOS differed between species (see above), it did not differ significantly between cells tuned to conspecific and heterospecific faces (1 × 2 ANOVA,
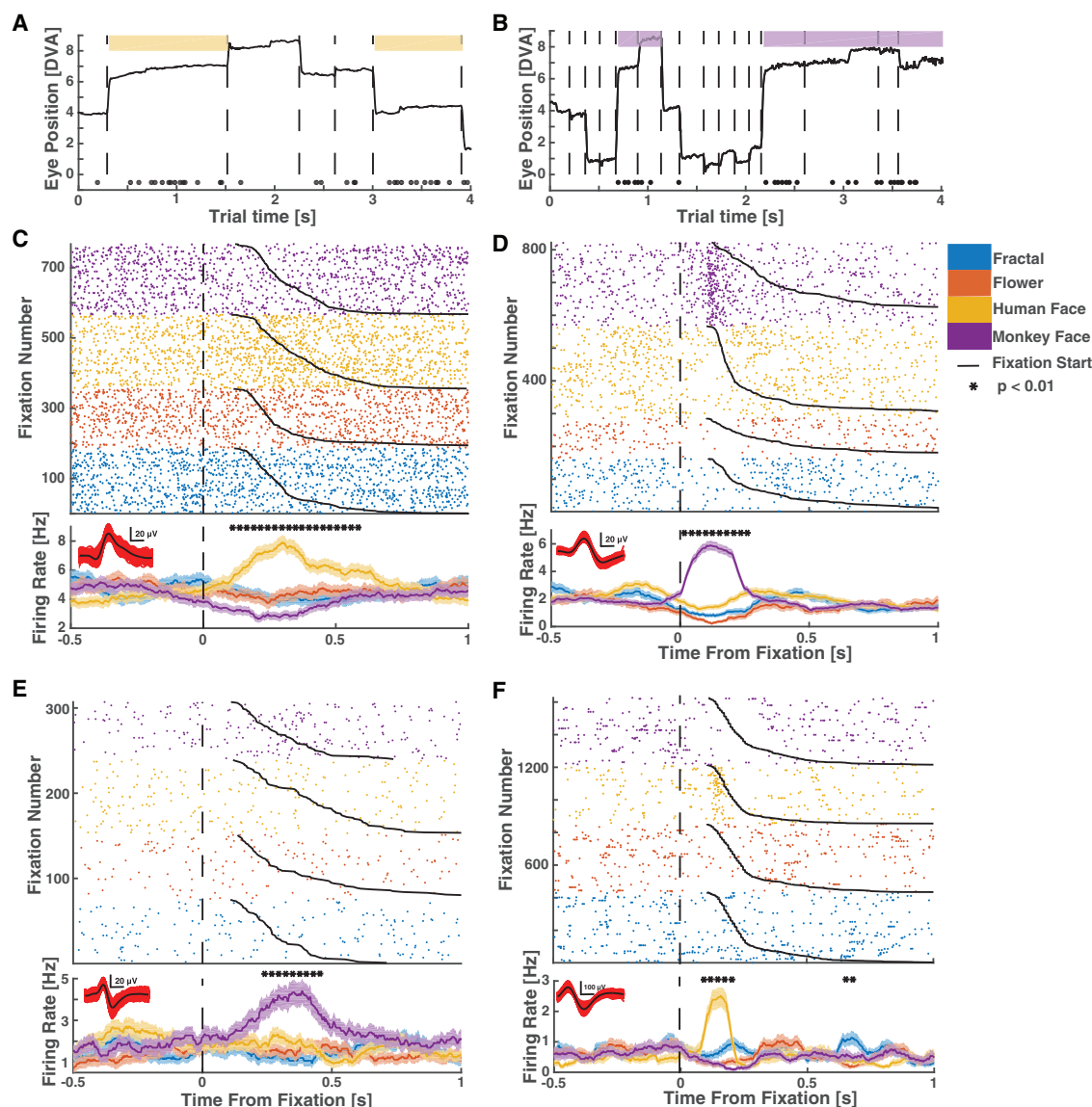
Figure 2. Example Single Neurons with Fixation-Related Activity

(A and B) Example trial from a human (A) and monkey (B) face-selective neuron. Spikes are indicated by black dots. Whenever gaze fell onto a conspecific face (colored patch), the neuron increased its activity.

(C–F) Rasters (top) and mean firing rate (PSTH, bottom) for neurons recorded in humans (C and E) and monkeys (D and F). Neuron are selective for conspecific (C and D) and heterospecific (E and F) faces. (C) and (D) show the activity of the neurons depicted in (A) and (B). t = 0 marks fixation onset. Trials were sorted by category of the fixated image (color code) and fixation duration (black line). Stars above the PSTHs indicate bins of neural activity (of 250-ms duration) with a significant (1 × 4 ANOVA, p < 0.01) difference in firing rate. Horizontal scale bar for waveforms is 0.2 mV. The four neurons are from different subjects. Shading indicates SEM.

$F_{(1,67)} = 1.75$, $p = 0.19$, and $F_{(1,33)} = 0.52$, $p = 0.47$ in humans and monkeys, respectively). We also estimated DOS values using fixations (50/50 split) not previously used to select neurons and found that DOS values are highly reliable and significantly larger than those of unselected cells (Figures 4C and 4D; see legends for statistics). Taken together, these observations show three important similarities between neurons in the human and monkey amygdalae: both contain fixation target-sensitive neurons; these neurons show category-specific responses; and

the largest subset of such neurons responds preferentially to conspecific faces. A difference between the species was that human neurons have a sparser response profile over categories.

Interspecies Comparison of Response Latencies of Face-Selective Neurons

We next compared the latency of the fixation target-sensitive neurons between species. We estimated the response latency for each cell to test whether the time at which the modulation of firing
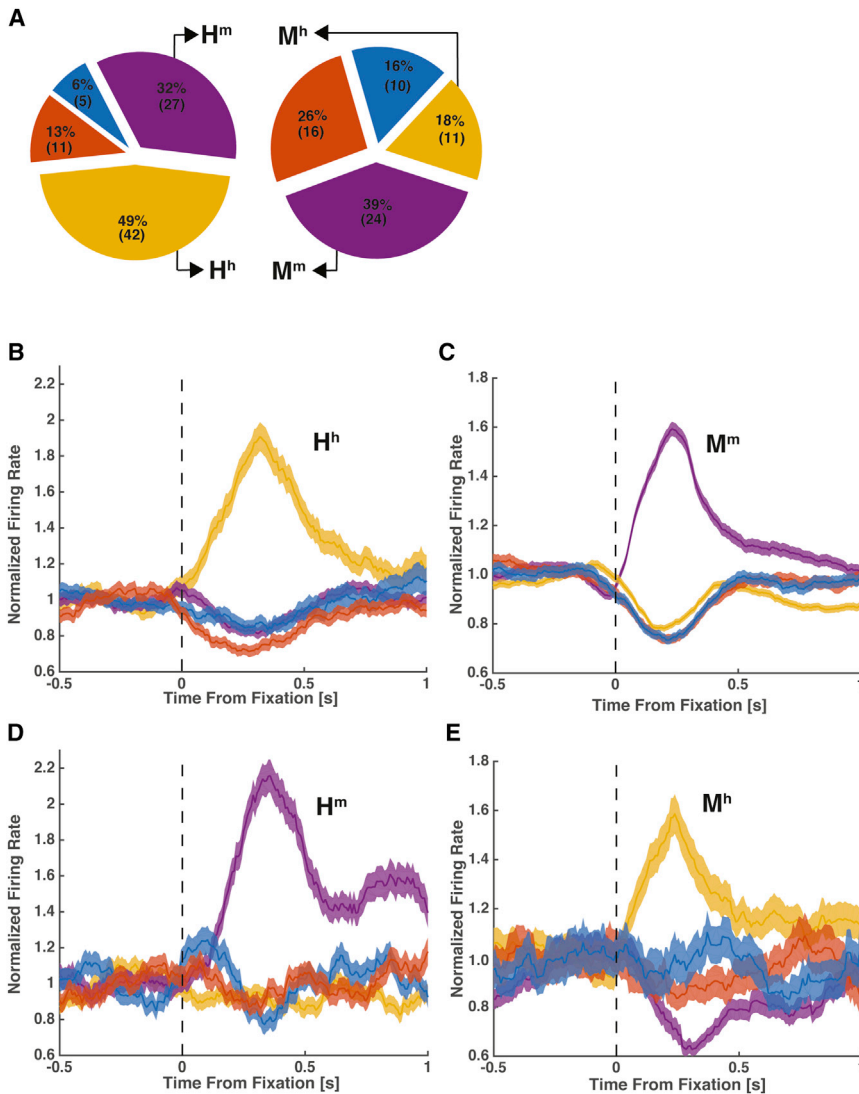
Figure 3. Population Analysis and Cross-Species Comparison of Fixation-Related Visual Category Selectivity

(A) Preferred stimulus of all recorded visually selective cells in human (left) and monkey (right) amygdala. The largest proportion of neurons responded maximally to faces of conspecifics: 49% (N = 42) and 39% (N = 24) of selective neurons in humans and monkeys, respectively.

(B–E) Average normalized PSTHs of the four groups of face cells we identified (H^h, M^m, H^m, M^h). The middle row (B and C) shows neurons selective for conspecifics in humans (left, H^h, N = 42) and monkeys (right, M^m, N = 24). The bottom row (D and E) shows cells selective for heterospecifics in humans (left, H^m, N = 27) and monkeys (right, M^m, N = 11). Error bars are ± SEM across cells. Notation: H^h, human cell selective for human faces; H^m, human cell selective for monkey faces; M^m, monkey cell selective for monkey faces; M^h, monkey cell selective for human faces.

test; Figure 5F). Together, this shows that, regardless of stimulus selectivity, cells in the human amygdala respond approximately 100 ms later relative to fixation onset compared to cells recorded in the macaque amygdala.

We next compared, within each species, whether there were latency differences between cells tuned for different stimuli. We used two methods (Figure 6) to measure the response latency difference between the two most prominent cell categories that we found: face cells for conspecific and heterospecific faces (H^h versus H^m in humans and M^m versus M^h in macaques). Using our selection criteria, we computed the number of cells that would be tuned for each category as we shifted the point of analysis from 500 ms before the onset of fixation, until 1,000 ms after the onset of fixation (step size of 8 ms). Our measure of latency was the point in time where the proportion of cells tuned exceeded that expected by chance for the first time (see Experimental Procedures). Using this approach, we found that cells that were selective for conspecific faces responded significantly earlier than cells that were selective for heterospecific faces in both species (Δ_{human} = 70 ms, Δ_{monkey} = 90 ms; Figures 6A and 6B). In addition, we also confirmed this result using a moving-window receiver operating characteristic (ROC) analysis and found a similar difference (Δ_{human} = 62 ms, Δ_{monkey} = 38 ms; Figures 6C and 6D). Together, this shows that information about conspecific faces is available at an earlier point of time relative to information about faces of other species in both humans and monkeys.

Category Preference of Fixation-Sensitive Neurons during Covert Attention

We next asked whether fixation target-sensitive cells retain their tuning for peripherally presented stimuli when these are
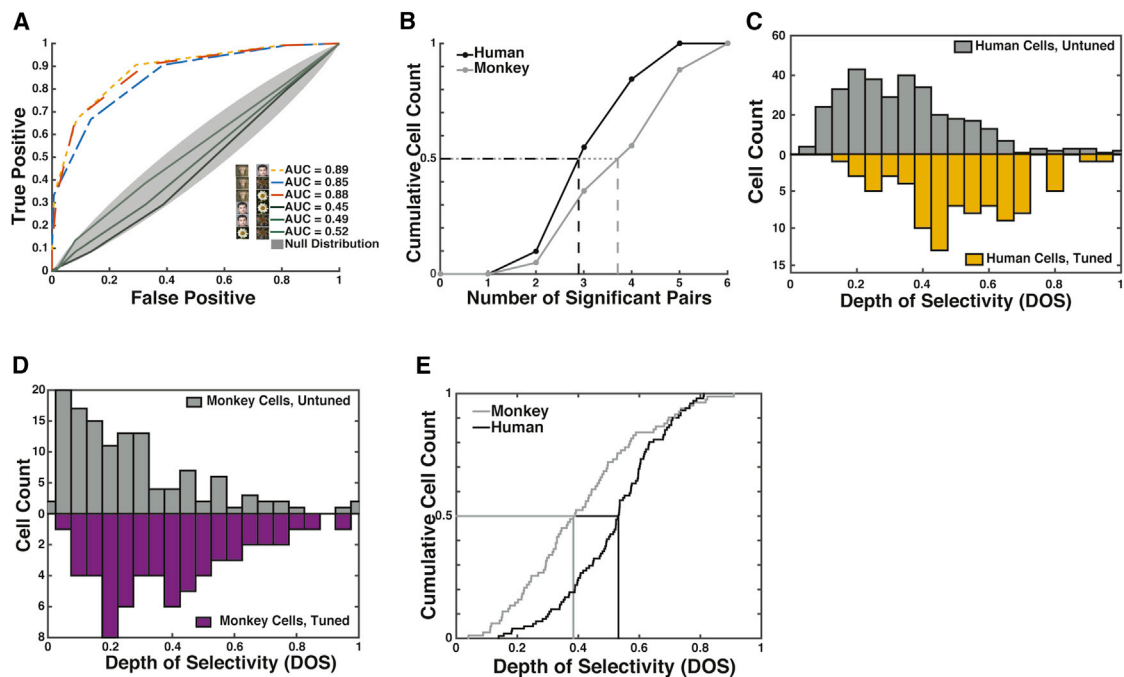
rate was first detectable systematically co-varied as a function of species and stimulus type. We first quantified latency differences using a single-neuron moving-window regression model to estimate the amount of variance in the firing rate that can be explained at any point of time by the visual category of the fixated stimulus (Figures 5A–5D; 1 × 4 ANOVA, moving window of 250 ms, step size of 8 ms). We then estimated the effect size ω^2 as a function of time separately for each neuron to determine the point of time, relative to fixation onset, at which ω^2 first became significant (Figures 5A–5D). We found that the onset of the fixation-sensitive neurons in macaques was, on average, 76 ms earlier than in humans (101 ± 7.5 ms versus 177 ± 8.7 ms, p = 5e-8, two-sample KS test; Figure 5D). Times reported are the center of the analysis window; i.e., spikes within ±125 ms are considered. Also, the proportion of neurons that became visually selective increased earlier in macaques compared to humans (Figure 5E), and the point of time at which neurons provided the most information (peak of ω^2) was 113 ms earlier in macaques compared to humans (209 ± 8.9 ms versus 322 ± 7.5 ms, p < 1e-14, two-sample KS

Figure 4. Monkey and Human Amygdala Cells Differ in Their Depth of Selectivity

(A) Single-cell ROC analysis example. The monkey cell shown (identical to that in Figure 2D) responded only to images of conspecifics, allowing it to discriminate three pairs of categories (dashed colored lines).

(B) Distribution of the number of significant contrasts (see A) for all visually tuned neurons in humans (black) and monkeys (gray). Cells recorded in monkeys differentiated significantly more contrasts (4.15 ± 0.2) than humans cells (3.47 ± 0.1, $p < 0.002$, two-sample KS test).

(C and D) Population summary. Comparison of depth of selectivity (DOS) values for tuned and untuned cells in human (C) and monkey (D). In both species, the DOS values are significantly greater in the tuned population ($p < 1e-16$ in humans and $p < 5e-5$ in monkeys, two-sample KS test). For tuned cells, DOS values were calculated for a subset of fixations that were not used in the selection of that cell (i.e., to determine its tuning).

(E) Depth of selectivity (DOS) for all visually tuned neurons in humans (black) and monkeys (gray). Human cells had significantly larger DOS values (0.51 ± 0.02 versus 0.43 ± 0.03, $p < 0.0003$, two-sample KS test).

attended but not fixated. In a separate experiment, we recorded 119 cells (six human subjects, eight sessions) during a covert-attention task with enforced central fixation. Images were identical to those used in the free-viewing task except that only a single image was shown at one (randomized) array location in isolation. Subjects maintained fixation at the center of the screen while a single stimulus was shown in the periphery (Figure 7A). We found that, of all the tuned cells in the covert condition (n = 31/119, 26%), 25/31 of neurons were tuned to either human or monkey faces (1 × 4 ANOVA, n = 16, Figures 7B and 7E; n = 9, Figures 7C and 7F, respectively). For a subset (n = 10) of these face-selective neurons, we also recorded responses during the free-viewing task. Of these 10 cells, all maintained their face selectivity across the two task conditions and a comparison of all cells recorded in both tasks (n = 31) revealed a high probability for cells to either be tuned in both tasks or neither tasks (p = 0.004, odds ratio: 22.8, Fisher's test of association). Also, the proportion of face-selective cells was not significantly different across the two tasks (25/31 and 69/85 in covert and free-viewing task, respectively; χ^2 = 0.0042, p = 0.94). Notably, cells responded significantly earlier in the free-viewing condition compared to the cells recorded in the covert-attention condition (Δ_{peak} = 191 ms; Figure 7D). This is expected because, during the

covert-attention condition, the location of the stimulus was unpredictable and thus deployment of covert attention could only be initiated following stimulus onset. In addition, the depth of selectivity was significantly larger in the covert compared to the free-viewing condition (p < 0.01 for all n = 422 cells in free viewing and n = 119 in covert condition, p < 0.05 for all n = 85 tuned cells in free viewing and n = 31 cells in covert condition, two-sample KS test). Together, these data support the hypothesis that amygdala neurons selective for faces and other complex visual objects are responsive to the currently attended visual stimulus both during free viewing as well as during covert attention.

DISCUSSION

Our results reveal that amygdala activity during active exploration of complex scenes is strongly modulated by the currently fixated stimulus. In contrast, previous studies in the amygdala of humans (Fried et al., 1997; Kreiman et al., 2000; Rutishauser et al., 2011) and macaques (Desimone, 1991; Gothard et al., 2007; Leonard et al., 1985; Nakamura et al., 1992; Rolls, 1984) relied on isolated single objects and were thus unable to investigate whether responses were modulated by gaze or not. Indeed,
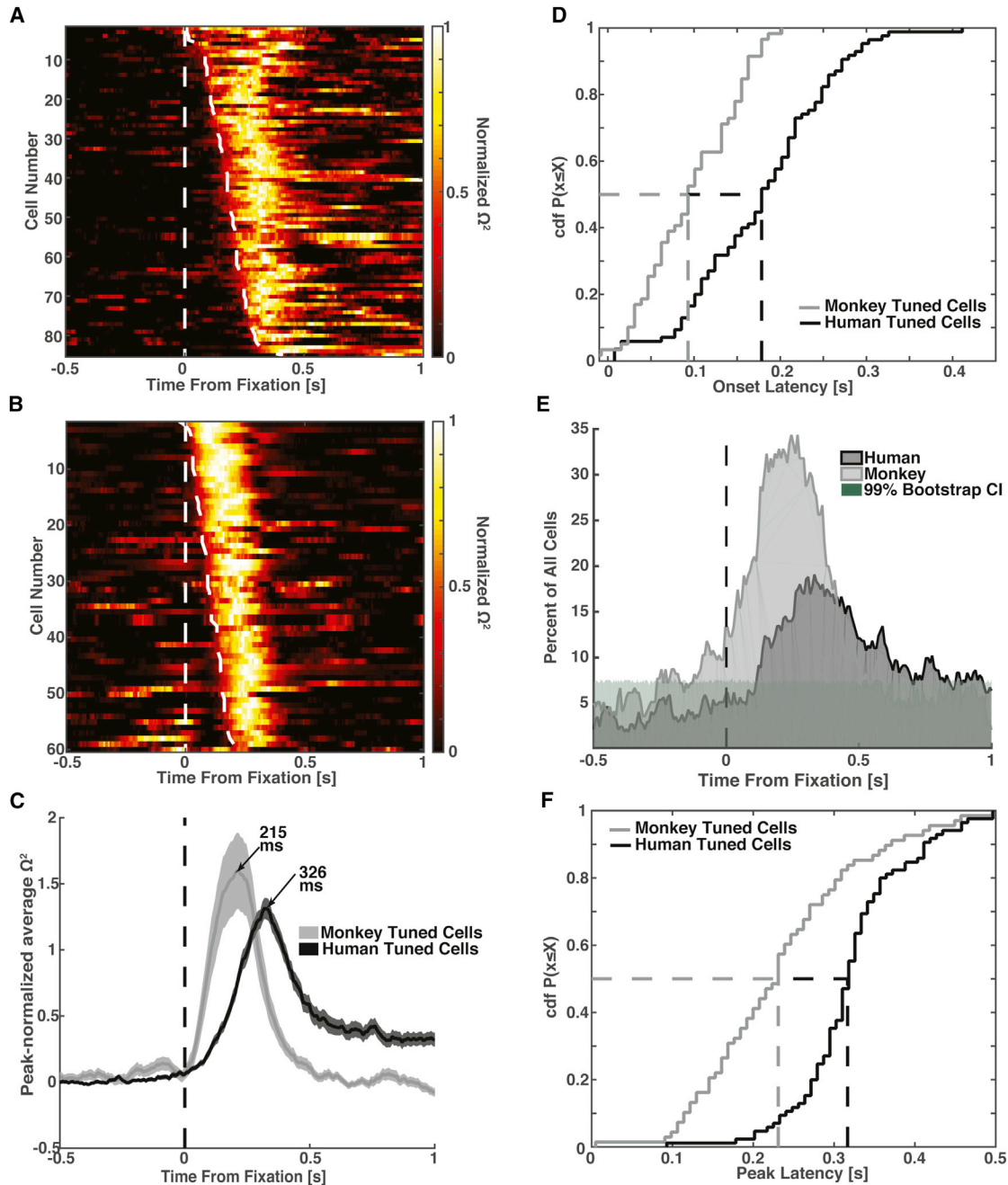
Figure 5. Interspecies Comparison of Response Latency Relative to Fixation Onset

(A and B) Effect size (Ω^2) for all visually selective human (A) and monkey cells (B) as a function of time and sorted by earliest point of significance (only cells that are significant at the $p < 0.01$ level are shown). Each cell's effect size is normalized to its peak.

(C) Mean normalized effect size for all visually selective cells recorded in humans (N = 85) and monkeys (N = 61). Shading indicates SEM.

(D) Cumulative distribution of the onset latency computed using the effect size (see A and B). The mean onset latency was significantly earlier in monkeys (101 ± 7.5 ms) compared to humans (177 ± 8.7 ms, $p < 5e\text{-}8$, two-sample KS test).

(E) Proportion of all recorded cells that were sensitive to the identity of the fixated stimulus as a function of time (bin size, 250 ms; step size, 8 ms). Shading shows the 99th percentile of the bootstrap distribution.

(F) Cumulative distribution of the time from fixation onset until peak effect size. Peak effect size was reached significantly earlier in monkeys compared to humans (209 ± 8.9 ms versus 322 ± 7.5 ms, $p < 1e\text{-}14$, two-sample KS test).
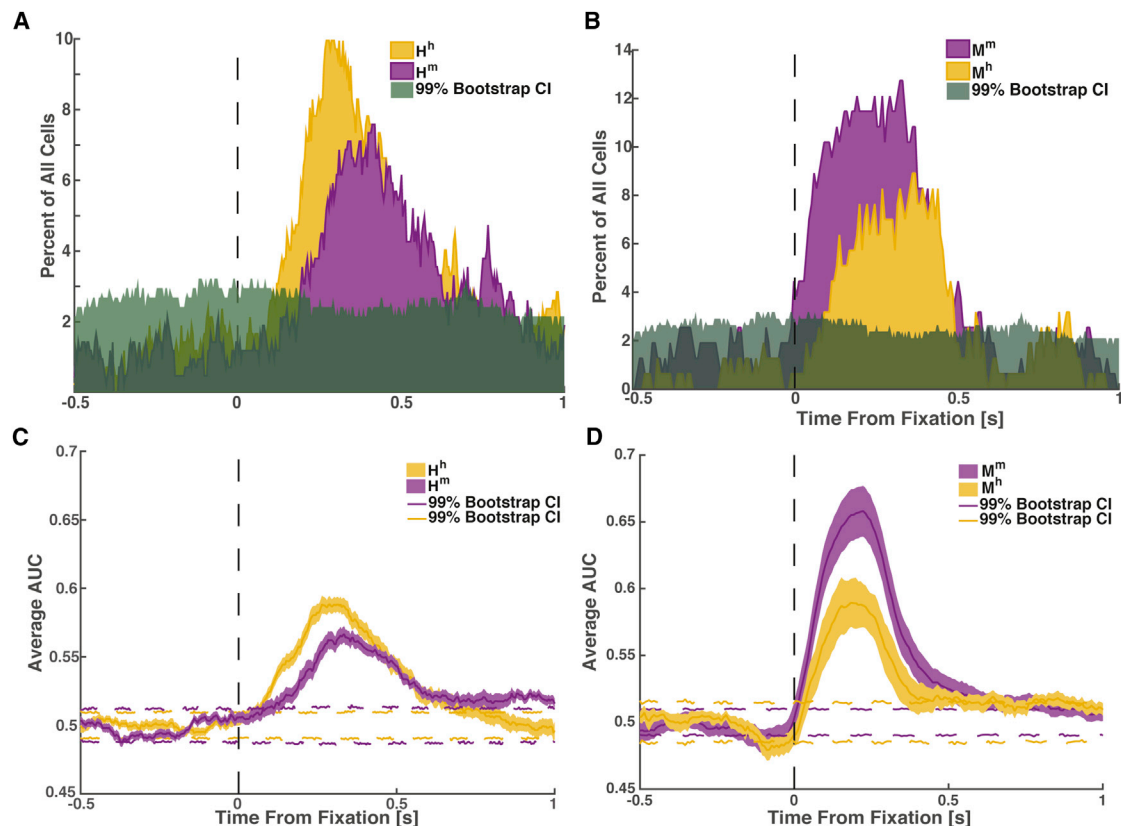
Figure 6. Face Cells Responded Earlier and More Strongly to Conspecific Compared to Heterospecific Faces

(A) Proportion of all recorded cells in humans (out of N = 422) selective for fixations on conspecific (H^h, yellow) and heterospecific faces (H^m, purple). Green shading indicates the 99% confidence interval.

(B) Proportion of all recorded cells in monkeys (out of N = 195) selective for fixations on conspecific (M^m, purple) and heterospecific (M^h, yellow) faces.

(C and D) The average AUC for all cells recorded in human that are selective for human faces (yellow) and monkey faces (purple) (C). The average AUC for all cells that are recorded in monkey that are selective for human faces (yellow) and monkey faces (purple) (D). Dotted colored lines indicate the 99% confidence interval.

the assumption so far has been that, because inferotemporal cortex neurons have large receptive fields for images shown in isolation (Tanaka, 1993; Tovee et al., 1994), the response of amygdala neurons should not depend on fixation location. However, here we find that the effective receptive field is relatively small in our task. This finding is similar to the response properties of "eye cells" in the macaque amygdala, which respond only when a monkey fixates on the eyes of another monkey (Mosher et al., 2014).

Little is known about the effective receptive field sizes and their dependence on stimulus density for human and macaque amygdala neurons. In higher visual cortical areas in macaques, receptive fields encompass the entire hemifield (Barraclough and Perrett, 2011; Boussaoud et al., 1991; Gross et al., 1969). At the same time, many such neurons have heightened sensitivity to information present at the fovea (Moran and Desimone, 1985; Rolls et al., 2003). Once animals are allowed to actively explore complex visual scenes, however, receptive fields of neurons in macaque TE can shrink considerably (Rolls et al., 2003; Sheinberg and Logothetis, 2001). Although it is possible that neurons in the amygdala inherit some of their properties from the same higher visual

cortical areas (Amaral et al., 1992; Barraclough and Perrett, 2011; Rolls et al., 2003), the significantly increased response latencies and complex selectivity changes we show make it unlikely that the responses we document are simply representing cortical input.

Our results show that the fixation-dependent responses were likely an effect of attention. This is because covert attention produced the same conclusions, even in some of the very same cells. The strength of this result is limited to humans, because we did not perform the same task in monkeys due to the difficulty of training monkeys accustomed to free viewing on a covert-attention task. An additional difference between covert and overt attention was that the sharpness of tuning (sparsity) was greater (more sparse) during covert compared to overt attention. A plausible explanation for this result is that the overt-attention task still permits some influence from the concurrently presented (unattended) other images. In contrast, this source of competition is removed in the covert task (because only a single stimulus was presented). Indeed, unattended task-irrelevant peripheral faces can impair performance in a variety of settings (Landman et al., 2014), and it is possible that the reduction of selectivity we observed here is a reason for this effect.
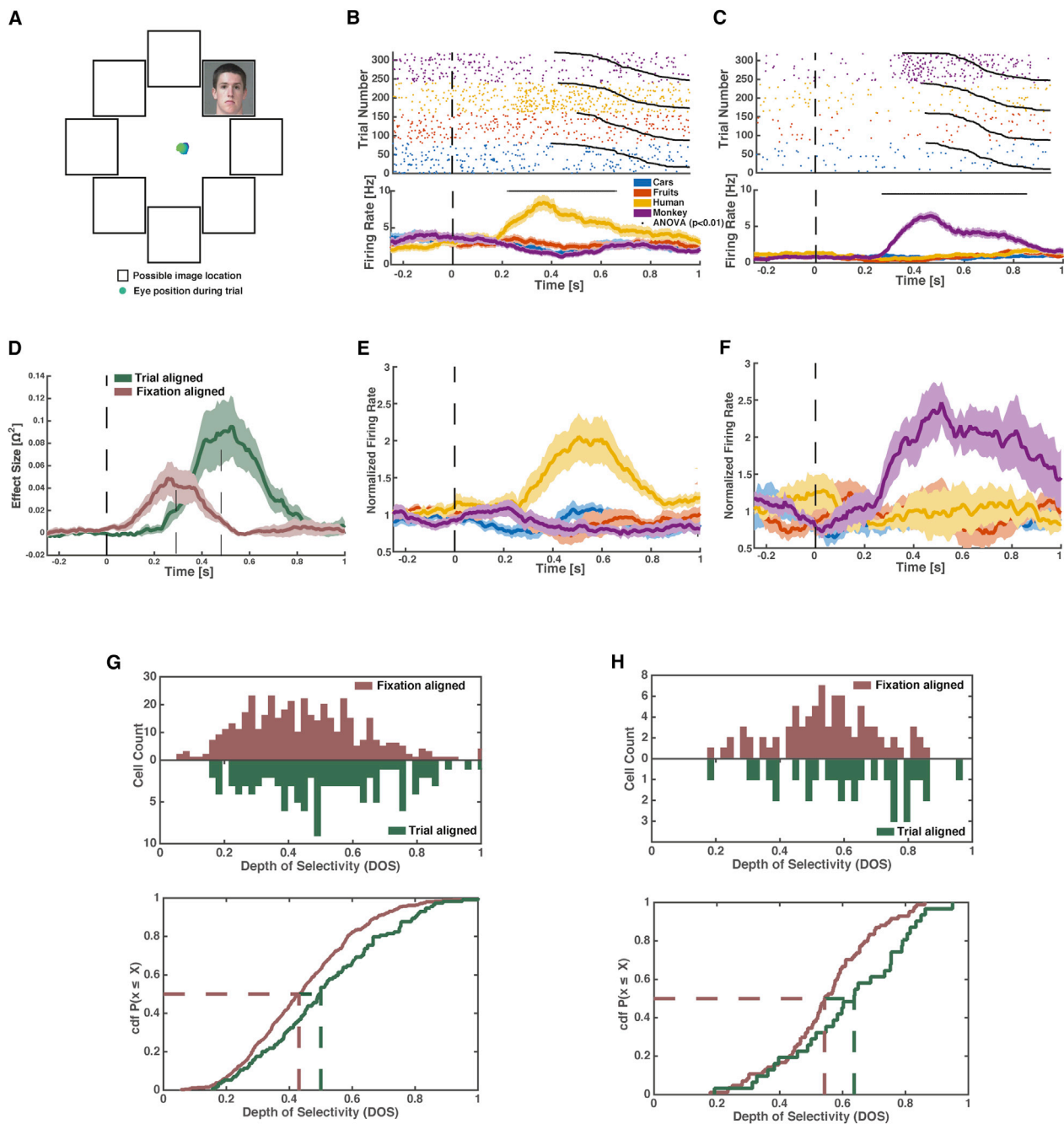
Figure 7. Face-Selective Amygdala Neurons Recorded in Humans Respond to Covertly Attended Faces

(A) Subjects fixated at the center of the screen and indicated by button press whether a peripheral image depicted a car. Shown is a single example trial, with eye-tracking data (blue) indicating that subjects maintained fixation.

(B and C) An example cell recorded in human that responds selectively to human faces in the covert attention task (B). An example cell recorded in human that responds selectively to monkey faces in the covert attention task (C). t = 0 is stimulus onset.

(D) Comparison of response of face cells in covert and free-viewing sessions for the subset of cells that were recorded in both tasks (4/7 sessions, n = 10). The average effect size is shown fixation and trial onset aligned.

(E) PSTH of all human face-selective neurons (H^h, n = 16) during the fixation-enforced covert-attention condition.

(F) PSTH of all monkey face-selective neurons (H^m, n = 9) during the fixation-enforced covert-attention condition.

Shading indicates SEM in (B)–(F).

(G and H) Population-level comparison between the covert and free-viewing tasks for all recorded (G) and only visually tuned (H) cells. (G) DOS values were significantly larger in the covert-attention task compared to the free-viewing task (p < 0.01, two-sample KS test). (H) DOS values were significantly higher in the covert-attention task (p < 0.05, two-sample KS test).

Role of Face Cells in Social Behavior

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

Information Represented by Face Cells

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

Latency Differences

We found that, in both species, face cells responded significantly earlier to faces of conspecifics relative to heterospecific faces. Behaviorally, both macaques and humans preferentially process faces of conspecifics more efficiently and are better at differentiating individuals of the same species (Dufour et al., 2006; Pascalis and Bachevalier, 1998). In macaques, face-sensitive cells in the inferotemporal cortex differentiate between human and macaque faces (Sigala et al., 2011) and respond earlier to human compared to non-primate animal faces (Kiani et al., 2005). The

same neurons, however, showed no latency difference when comparing humans versus macaques (Kiani et al., 2005). A new hypothesis motivated by our result is that the human versus macaque same-species latency advantage is first visible in the amygdala as the result of the higher social significance attributed by the amygdala to conspecific faces.

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

Conclusions

Faces are stimuli of high significance for primates, and the brains of several species contain multiple areas connected in a network specialized for face processing (Desimone, 1991; Emery, 2000; Tsao et al., 2006, 2008). Exploring the division of labor among the different nodes in this network has been a fruitful approach to capturing more general, circuit-level principles of neural computation. Indeed, a detailed analysis of face cells throughout the brain revealed a distributed but interconnected system of cortical face patches specialized for different components of face processing (Kanwisher and Yovel, 2006; Tsao et al., 2008). However, most of what is known about this network has been derived exclusively from work in macaques, even though it is often assumed that the properties of this system are the same in humans (Barraclough and Perrett, 2011). Here, we present critical direct evidence for significant differences and commonalities. It is likely that the face-responsive properties of amygdala neurons arise, at least in part, through convergent inputs from several cortical areas where face cells have been identified. However, the face cells in the amygdala do not merely recapitulate the response properties of face cells in cortical areas, but show pronounced effects of species-specific

relevance, and of attention. These findings revise our view of the amygdala's contribution to face processing from that of an automatic and broad detector, to that of a highly selective and attention-dependent filter. These effects likely constitute an essential ingredient for guiding processing in downstream regions, and ultimately for generating social behavior in real-world settings where many stimuli constantly compete for attention.

EXPERIMENTAL PROCEDURES

Detailed methods are provided in Supplemental Experimental Procedures.

Human Electrophysiology

Human subjects were patients being evaluated for surgical treatment of drug-resistant epilepsy that provided informed consent and volunteered for this study. Monocular gaze position was monitored at 500 Hz (EyeLink 1000; SR Research). The institutional review boards of Cedars-Sinai Medical Center and the California Institute of Technology approved all protocols. We recorded bilaterally from the amygdala using microwires embedded in macroelectrodes.

Monkey Electrophysiology

All surgical procedures were approved by the Institutional Animal Care and Use Committee at the University of Arizona. We performed single-neuron recordings as reported before (Gothard et al., 2007).

Behavioral Task: Free Viewing

Monkeys were trained to fixate on a white cross. If the monkeys maintained gaze on the fixation spot for at least 100 ms, a circular array of images subtending 23.4 × 23.4 degrees of visual angle (dva) was presented. Monkeys were allowed to freely scan the scene for 3–4 s but were required to keep their gaze within the boundaries of the array for at least 3 s. Monkeys received a 0.5- to 1-mL juice reward followed by a 3-s intertrial interval if this condition was met. If the monkey failed to fixate, or looked outside the boundary of the image, the trial was terminated, reward was withheld, and the array was repeated.

Humans were instructed to freely observe the arrays for a fixed amount of time (4 s). After each array, a blank screen with a fixation cross was displayed for 1 s.

Behavioral Task: Covert Attention

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

Spike Sorting and Single-Neuron Analysis

We used the same processing pipeline to process the monkey and human recordings (see Supplemental Experimental Procedures). All post-stimulus time histogram (PSTH) diagrams were computed using a 250-ms window with a step-size of 7.8 ms. No smoothing was applied.

Selection of Units

We determined whether a cell's response is sensitive to the identity of fixated stimuli using a 1 × 4 ANOVA of the spike counts during a 250-ms-long time window centered on the point of time at which mutual information (MI) between the spike rate and identity of fixated images was maximal for each species (t = 332 ms and t = 229 ms, respectively). We excluded successive fixations that fall on the same category (Figure S3C, conservative criteria). To achieve this, we included only fixations that were not preceded or succeeded by fixation(s) on an image of the same category for at least 100 ms. If the ANOVA was significant (p < 0.05), we determined the category with the largest mean response

in the same time window. This category was used as the preferred category of the cell.

Assessment of Selectivity

We used ROC analysis between all possible six pairs of stimulus categories to assess the number of pairwise comparisons that each neuron was able to differentiate. For each of the six possible comparisons, we computed the moving-window area under the curve (AUC) and compared this to the bootstrap distribution, which was generated by shuffling the fixation labels and computing the AUC 1,000 times. In addition, we quantified the DOS of each neuron i by $DOS_i = (n - (\sum_{j=1}^{n} R_j / R_{max}))/(n - 1)$, where n is the number of categories (n = 4), R_j is the mean response to category j, and R_{max} is the maximal mean response. D varies from 0 to 1, with 0 indicating an equal response to all categories and 1 exclusive response to one but none of the other categories. Thus, a DOS value of 1 is equal to maximal sparseness.

Regression Analysis

We used the regression model $S(t) = \alpha_0(t) + C$ to estimate whether the firing rate S was significantly related to the factor category (C, 1–4). Spike counts $S(t)$ were computed for a 200-ms window that was moved in steps of 50 ms.

REFERENCES

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

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

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

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

Amaral, D.G., Price, J.L., Pitkanen, A., and Carmichael, S.T. (1992). Anatomical organization of the primate amygdaloid complex. In The Amygdala:

Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction, J.P. Aggleton, ed. (Wiley-Liss), pp. 1–66.

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

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

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

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

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

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

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

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

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

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

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

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

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

Grimaldi, P., Saleem, K.S., and Tsao, D. (2016). Anatomical connections of the functionally defined "face patches" in the macaque monkey. Neuron 90, 1325–1342.

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

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

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

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

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

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

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

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

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

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

Mosher, C.P., Zimmerman, P.E., and Gothard, K.M. (2014). Neurons in the monkey amygdala detect eye contact during naturalistic social interactions. Curr. Biol. 24, 2459–2464.

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

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

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

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

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

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

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

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

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

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

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

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

Rutishauser, U., Tudusciuc, O., Wang, S., Mamelak, A.N., Ross, I.B., and Adolphs, R. (2013). Single-neuron correlates of atypical face processing in autism. Neuron 80, 887–899.

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

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

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

Self, M.W., Peters, J.C., Possel, J.K., Reithler, J., Goebel, R., Ris, P., Jeurissen, D., Reddy, L., Claus, S., Baayen, J.C., and Roelfsema, P.R. (2016). The effects of context and attention on spiking activity in human early visual cortex. PLoS Biol. 14, e1002420.

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

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

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

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

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

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

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

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

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