Supplementary Figures and legends

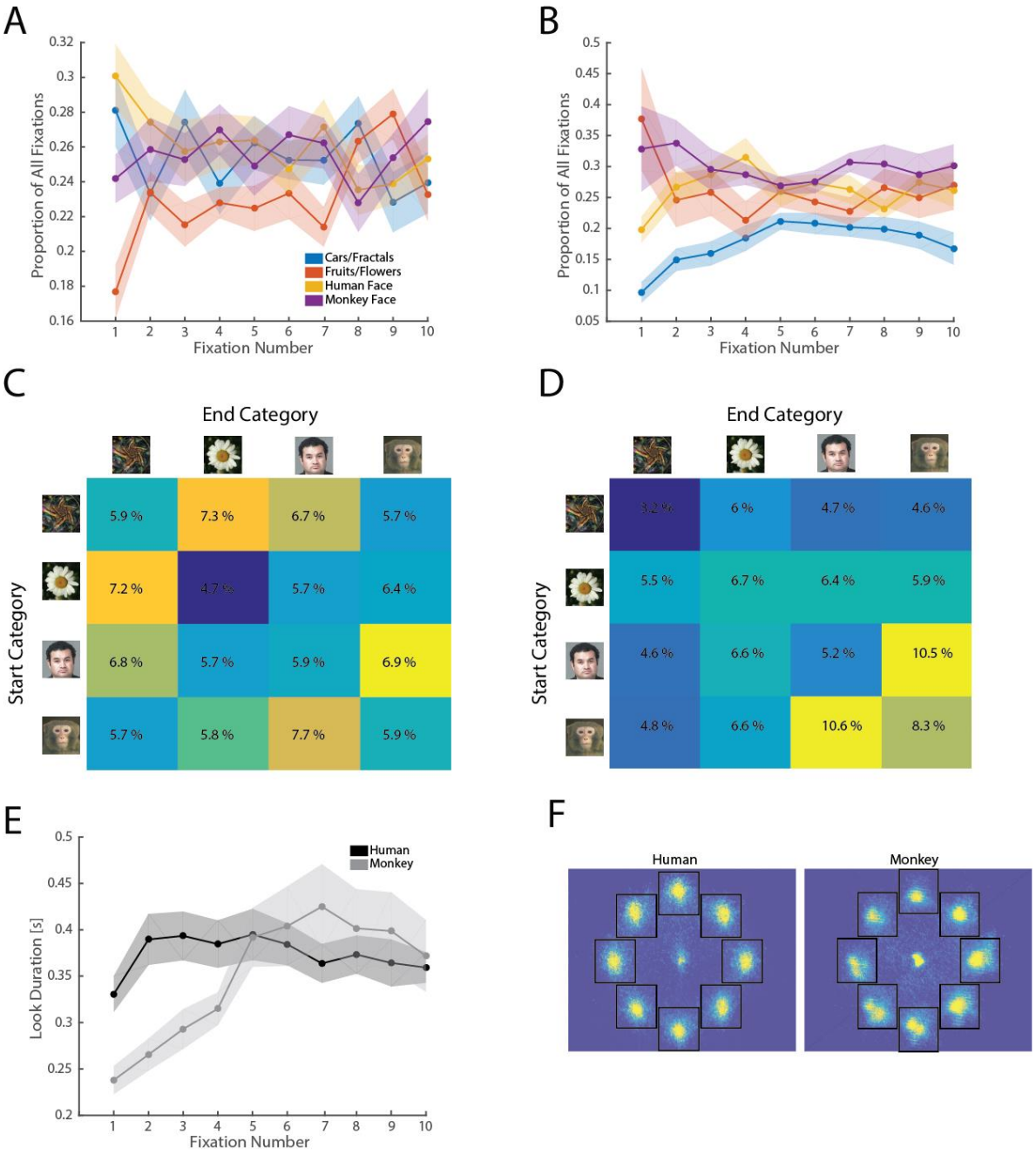

Figure S1: Comparison of human and monkey behavior, related to Figure 1. (A) Proportion of fixations that landed on a given image category as a function of fixation number and category for human subjects. (B) The same plot as in (A) but for monkeys. (A,B) Note how, for the first fixation, both species were more likely to look at faces of their own species compared to the other species. (C,D) Frequency with which each of the 16 possible category transitions occurred for human (C) and monkey (D) subjects. While all possible transitions occurred, some were more likely (i.e. human face – monkey face, and monkey face – human face). (E) Look duration as a function of fixation number. Both humans and macaques modulated look duration as a function of time in the trial (p=0.01 and P=10-7, respectively, 1x10 ANOVA). However, this difference was entirely accounted for by the difference between the first and second fixation in humans, but not macaques (p=0.09 and P=10-5, respectively,

1x9 ANOVA). (F) Heat maps showing the eye tracking data for all sessions in humans (left) and monkeys (right). Both humans and monkeys sampled all 8 images.

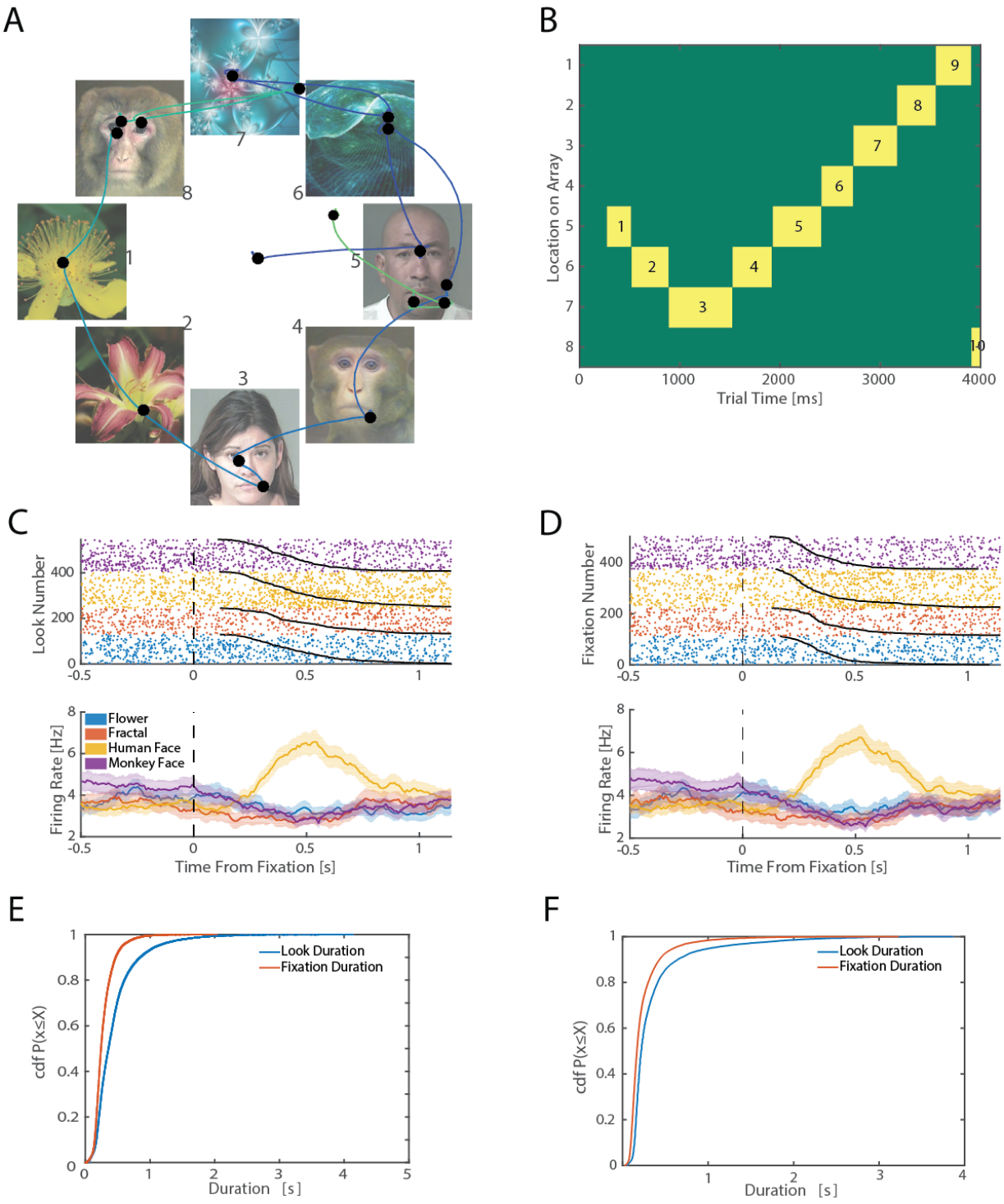

Figure S2: Comparison of Fixation onset and look onset methods, related to Figure 1. (A) Example scan path from a single trial from a human subject. See Fig 1 for notation. (B) Summary of eye tracking data into discrete periods of "looks" (yellow squares). Successive fixations that fall on the same image are pooled together into a single "look". The y-axis denotes the location of the look in the

array as indicated in (A). (C,D) Comparison of a single-cell response, aligned to fixation onset (C) and look onset (D). Note the virtually identical response of the cell using the two criteria. For each, Raster and PSTH are shown. (E) Cumulative distribution for fixation (red) and look (blue) duration. Look duration was longer because of the pooling of several fixations into one look. (F) Same as (E), but for monkeys.

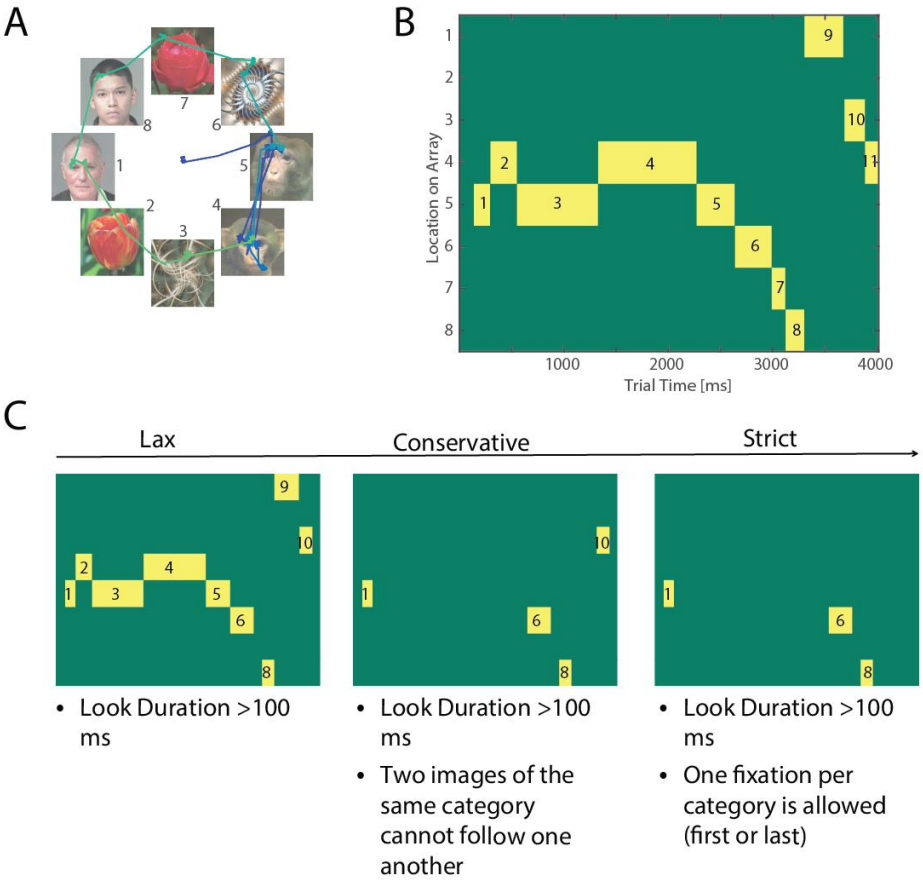

Figure S3: Criteria for looks/fixations that are included in analysis, related to Figure 2. (A) Example scan-path from a monkey. See Fig 1 for notation. Numbers indicate "location on array". (B) For trial shown in (A), summary of where the monkey looked. All successive fixations that fall on the same image are pooled into one "look" (Yellow patch). The looks are numbered 1-10 and the y-axis indicates the location on the array. (C) Different selection criteria for looks that are included in the analysis. In the most lenient case, we used all fixations that we longer than 100ms, and in the most stringent case, we used only used the first fixation for each category in addition to the duration requirement. All analysis, unless indicated, used the "conservative" (middle side) criteria.

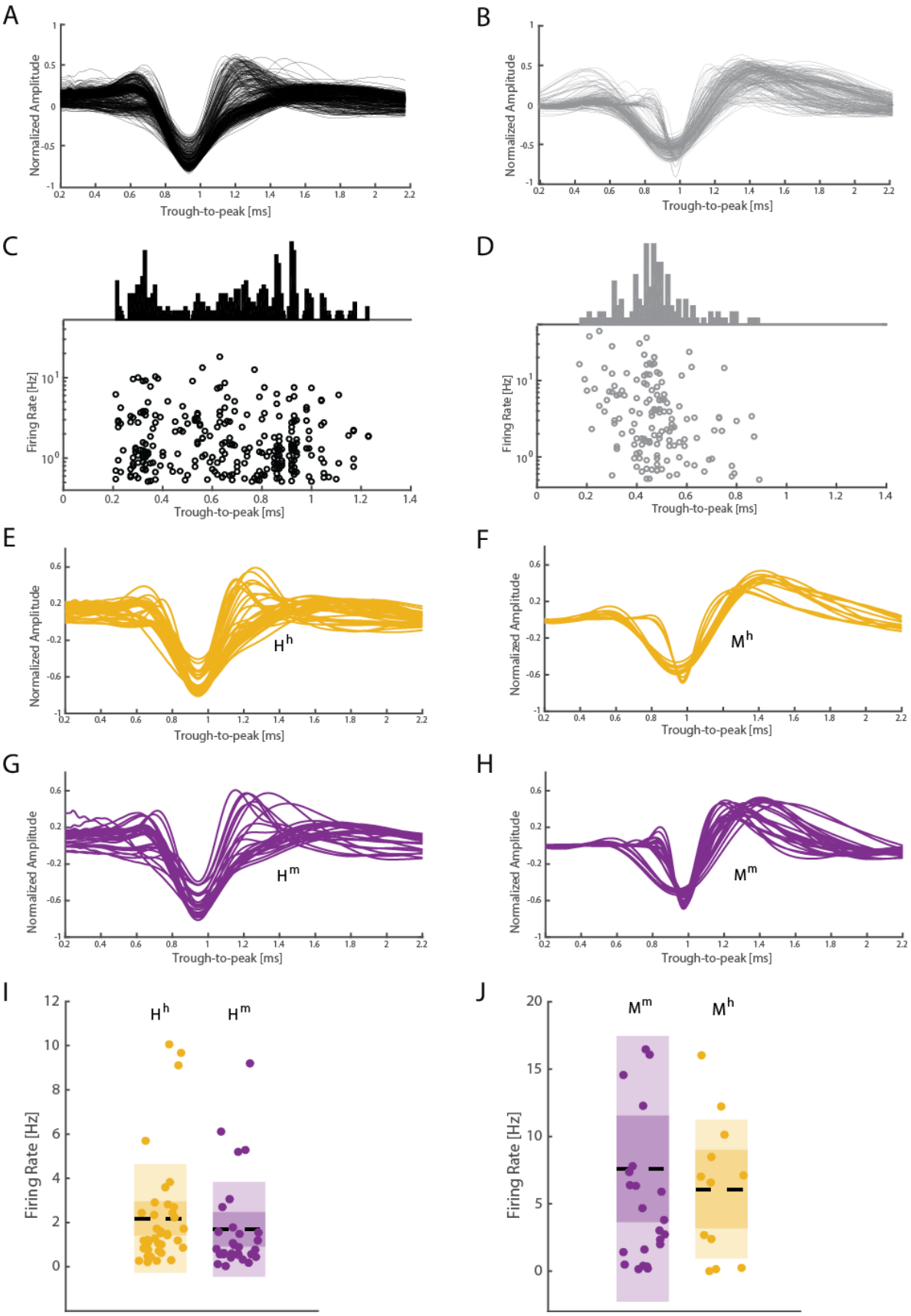

Figure S4: Electrophysiological properties of neurons in both species, related to Figure 2. (A)

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

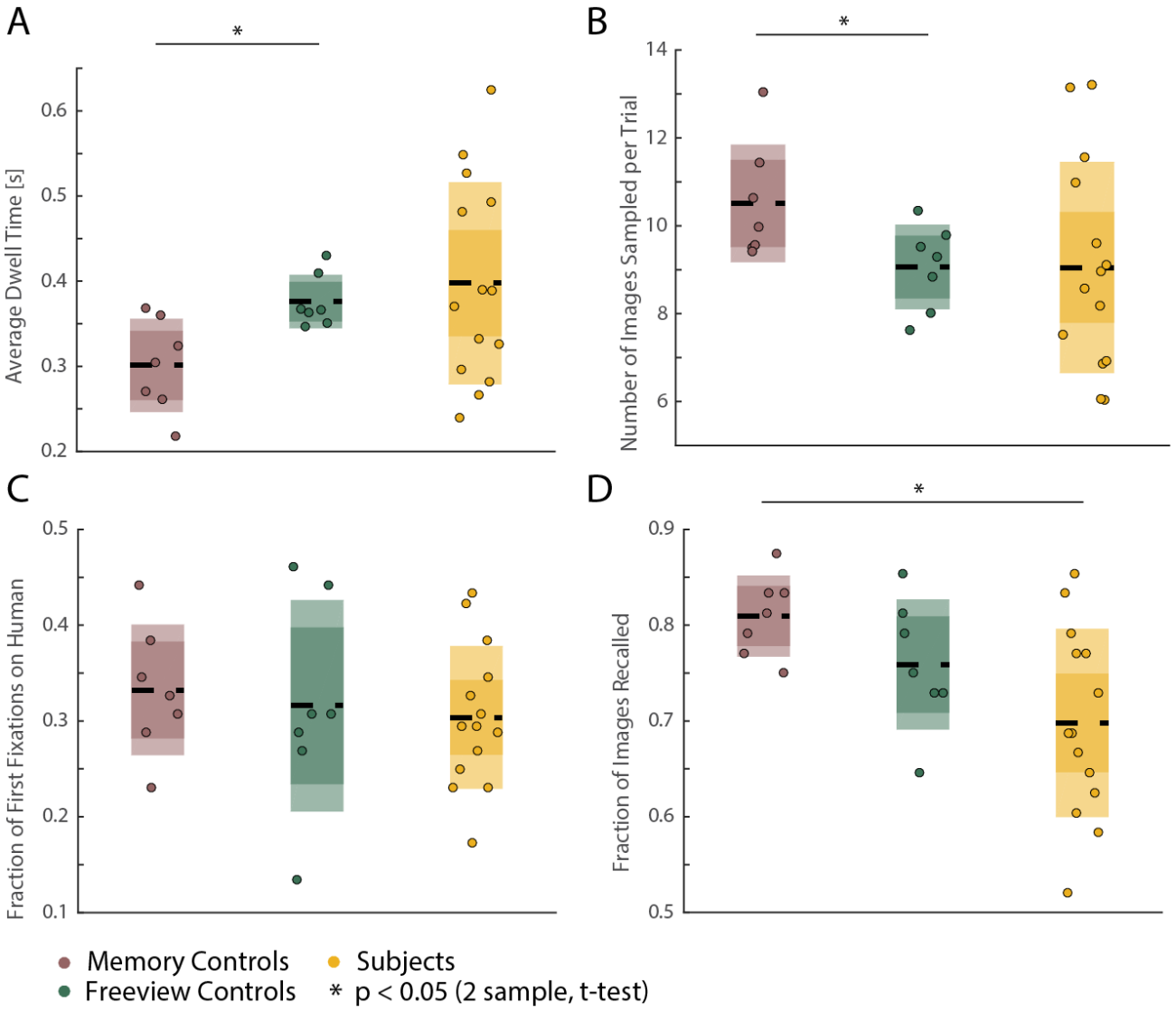

Figure S5: Comparison of behavior of human neurosurgical subjects with normal control subjects, related to Figure 1. (A) The average look time for each image for three groups of subjects:

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

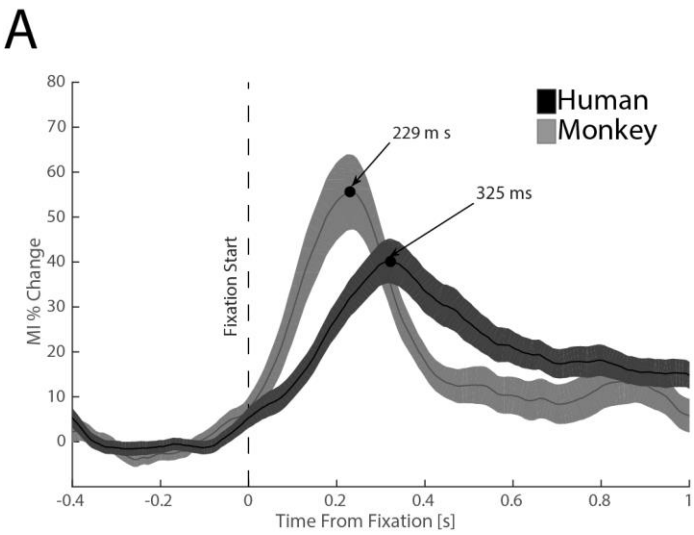

Figure S6: Using mutual information to determine the position of the analysis window for selectivity analysis, related to Figure 3. (A) Time-course of information, quantified as mutual information (MI; peak normalized) between the firing rate and visual category for all neurons recorded in monkeys (N=195, light gray) and humans (N=422, black). The point of time at which MI was maximal (t=325ms and 229ms, respectively) was used to place the analysis window for all further analysis.

Supplementary experimental procedures

Surgical procedures (macaques)

The stereotaxic coordinates of the amygdala of the three adult male monkeys (*Macaca mulatta*), Q, G, and R were calculated based on pre-surgical MRI scans. Each monkey was then implanted with a recording chamber above the right amygdala and 3 small titanium posts (Thomas Recording, Germany) for immobilizing the head during acute neurophysiological recordings. A craniotomy (~13 mm in diameter) was opened in the center of each chamber and sealed with a silicone elastomer to maintain sterility (Spitler and Gothard, 2008). A post-surgical MRI with contrast verified the orientation of the chamber relative to the amygdala. Visualization of electrode locations was further confirmed utilizing the INIA19 atlas (Bakker et al., 2015; Rohlfing et al., 2012), which was also used for visualization in the graphical abstract. All surgical procedures were carried out in compliance with NIH guidelines and were approved by the Institutional Animal Care and Use Committee at the University of Arizona.

Monkey electrophysiology

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

Human electrophysiology

From each micro-wire , we recorded the broadband 0.1-9000Hz continuous extracellular signal with a sampling rate of 32kHz (Neuralynx Inc). One microwire on each macroelectrode served as a local reference (bi-polar recording). All included patients had clearly distinguishable spiking activity on at least one electrode in at least one amygdala.

Spike sorting and single-neuron analysis

The raw signal was filtered with a zero-phase lag filter in the 300-3000Hz band and spikes were detected and sorted using a semi-automated template-matching algorithm (Rutishauser et al., 2014; Rutishauser et al., 2006). In humans, channels with interictal epileptic spikes in the LFP were excluded.

Localization of electrodes (humans)

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

in a post-operative scan overlayed on the MNI152-registered version of the CIT168 template brain (Tyszka and Pauli, 2016). We then used Freesurfer's freeview program to mark the electrodes as point sets to determine in which amygdala nucleus the tips of the microwires were located.

Behavioral tasks

The human tasks were implemented in MATLAB using the Psychophysics Toolbox (Pelli, 1997). The human subjects were instructed to look at the images on the array for a later memory test. They viewed 48-52 array trials followed by 48 single image trials for the memory test.

Processing of human eye tracking data

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

Using look-onset instead of fixations

In the monkey, fixations and saccades were annotated using Cluster Fix (Konig and Buffalo, 2014). In the human subjects, we used the annotation provided by the Eyelink system. However, to bypass potential differences in the two annotation approaches, we used the "look onset" instead of fixation onset for all analysis in this paper. The "look onset" is the point of time at which the first data point fell onto a particular image (Fig S2). This way, we pooled all successive fixations that fell on the same stimulus into a single "look" (Fig S2B). While the look duration is thus typically longer than fixation duration (Fig S2E-F), our analysis depends only on the onset and is thus insensitive to this difference. Aligning with look-onset instead of fixation onset resulted in qualitatively similar neural responses (Fig S2C-D). At the same time, using look-onsets instead of fixation onset has several advantages, including that it 1) is insensitive to idiosyncrasies that might arise from the two separate annotation methods in the two species, 2) provides a more conservative estimate of the neural response latency, and 3) ignores fine structure in the neural response that may be driven by successive fixations on the same stimulus. Throughout the manuscript, fixation onset refers to look onset unless mentioned otherwise (i.e. Fig S2).

Single-neuron ROC analysis

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

Mutual information (MI)

In order to determine the post-fixation window of analysis in monkeys and humans, we computed the mutual information between the spike counts (S) and the image category (C) using: $I(S,C) = \sum_{S,C} P(S,C) \cdot log_2 \frac{P(S|C)}{P(S)}$, where C is a discrete variable that can take 1 of 4 possible values, and S is also a discrete variable that can take 1 of N possible values, depending on the maximum firing rate of the cell. The mutual information was computed for each cell and at each point along the PSTH (from -0.5s to 1s around the fixation). The mutual information for each cell was then averaged to produce the mean trace. The location of the center of the fixed window for all follow-up analysis was set to the point of time at which MI was maximal (see Fig S6).

Estimation of latency

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

Behavioral Controls

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

Regression analysis

We quantified the effect size of the factor category using ω^2, which is less biased than percentage variance explained (Olejnik and Algina, 2003). Models were fit and effect sizes calculated using the effect size toolbox functions mes1way and mes2way (Hentschke and Stuttgen, 2011). The null distribution was estimated by randomly scrambling the fixation labels and fitting the same model 1000 times. Estimates of latency were based on the first time the actual value was located outside of the 99% confidence interval.

Supplemental References

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

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

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

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

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

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

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

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

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

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

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

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

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