

Neurons in the human amygdala selective for perceived emotion

Shuo Wang^a, Oana Tudusciuc^b, Adam N. Mamelak^c, Ian B. Ross^d, Ralph Adolphs^{a,b,e}, and Ueli Rutishauser^{a,c,e,f,1}

^aComputation and Neural Systems, California Institute of Technology, Pasadena, CA 91125; ^bDivision of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125; Departments of ^cNeurosurgery and ^dNeurology, Cedars-Sinai Medical Center, Los Angeles, CA 90048; ^eEpilepsy and Brain Mapping Program, Huntington Memorial Hospital, Pasadena, CA 91105; and ^fDivision of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125

Edited by Riitta Hari, Aalto University, Espoo, Finland, and approved May 28, 2014 (received for review January 7, 2014)

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 7 neurosurgical patients with implanted depth electrodes. We presented degraded fear and happy faces and asked subjects to discriminate their emotion by button press. During trials where subjects responded correctly, we found neurons that distinguished fear vs. happy emotions as expressed by the displayed faces. During incorrect trials, these neurons indicated the patients' subjective judgment. 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.

human single unit | medial temporal lobe | limbic system | hippocampus | intracranial

The human amygdala plays a crucial role in processing socially and emotionally salient stimuli (1, 2). A large literature, primarily from studies in animals, shows that the amygdala is critical for conditioned fear responses (3). However, a number of other studies show that it is involved also in broader aspects of social perception, notably aspects of face processing (4). These two themes converge in several human studies: there is an impairment in recognizing fear faces in subjects that lack a functional amygdala (5) in addition to the impairment of fear conditioning (6, 7). Neuroimaging studies have also reported significant activation of the amygdala to fear faces (8).

In humans, it has been reported that amygdala neurons are selective for a variety of visual stimuli (9, 10). 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 fear faces (5), although there is now a consensus that the amygdala is involved in processing many emotions from faces, not just fear (11). Electrophysiological recordings in monkeys have found single neurons that respond not only to faces as such (12, 13), but also to face identities, facial expressions, and gaze directions (14, 15). Single neurons in the human amygdala discriminate faces from inanimate objects (10). Furthermore, single neurons in the human amygdala were found to encode whole faces selectively (16) and show abnormal facial feature selectivity in autism (17). Thus, there is substantial evidence from neurophysiological,

lesion, and functional MRI 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 (18), suggesting that the amygdala might act as a detector of perceptual saliency and biological relevance (19, 20). This was complemented by a neuroimaging study showing that amygdala activity was specifically enhanced for fear faces when saccading from the mouth to the eye region (21). Patients with schizophrenia (22), social phobia (23), and autism (24) also show abnormal facial scanning patterns, which have been hypothesized to result from amygdala dysfunction (25). The functional role of the amygdala is supported by its connection with visual cortices specialized for face processing (26–28) as well as reciprocal connections with multiple visually responsive areas in the temporal (29–31) and frontal lobes (32). 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 (*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 (18, 33), in which patients discriminated emotions from sparsely sampled fear or happy faces. We first characterize neurons that distinguished fear vs. happy emotions expressed by

Significance

Primates have a dedicated system to process faces. Neuroimaging, lesion, and electrophysiological studies find that the amygdala processes facial emotions. Here we recorded 210 neurons from 7 neurosurgical patients and asked whether amygdala responses are driven primarily by properties of the stimulus or by the perceptual judgments of the perceiver. Our finding shows, for the first time to our knowledge, that neurons in the human amygdala encode the subjective judgment of emotions shown in face stimuli, rather than simply their stimulus features.

Author contributions: R.A. and U.R. designed research; S.W., O.T., A.N.M., I.B.R., and U.R. performed research; S.W. and U.R. analyzed data; and S.W., R.A., and U.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. E-mail: Ueli.Rutishauser@cshs.org.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1323342111/-DCSupplemental.

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.

Results

Behavioral Performance. We recorded single neurons in the human amygdala while neurosurgical patients performed an emotion discrimination task (Table S1; see Fig. S1A for recording sites for each patient). All patients (nine sessions from seven patients in total; two patients did two 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 (six sessions) served as behavioral controls and participated in the same experiment. Subjects were asked to judge, for every trial, whether the stimulus was fear or happy by pushing corresponding buttons as quickly and accurately as possible (Fig. 1). Each trial was fear 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; Methods) were excluded from analysis.

We showed randomly selected parts of faces (bubbles; Fig. 1B) that allowed us to derive a behavioral classification image (BCI) (33) based on accuracy and reaction time (RT) of the responses (derived separately for happy-face trials and fear-face trials; Fig.

1C). 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 [regions of interest (ROIs) used are shown in Fig. 1B; two-tailed unpaired *t* test comparing average *z* scores within the ROIs: for fear-face trials, $P = 0.51$ for eyes and $P = 0.36$ for mouth; for happy-face 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 used information revealed by eyes to judge fear faces, whereas they used more mouth information to judge happy faces, consistent with previous studies (34, 35).

Patients were able to learn the task normally compared with controls (Fig. 1D and F; SI Results), even though they were slightly slower to respond (Fig. 1E). Importantly, there was no difference in accuracy or RT between "fear" or "happy" responses for both correct trials and incorrect trials (SI Results), showing that neither patients nor controls had a response bias. Overall, the behavioral performance-related metrics confirmed that patients were alert and attentive and had largely normal ability to discriminate emotion from faces.

Emotion-Selective Neurons. Two hundred and ten single units were isolated from nine recording sessions in seven 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 (Fig. S1A). 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

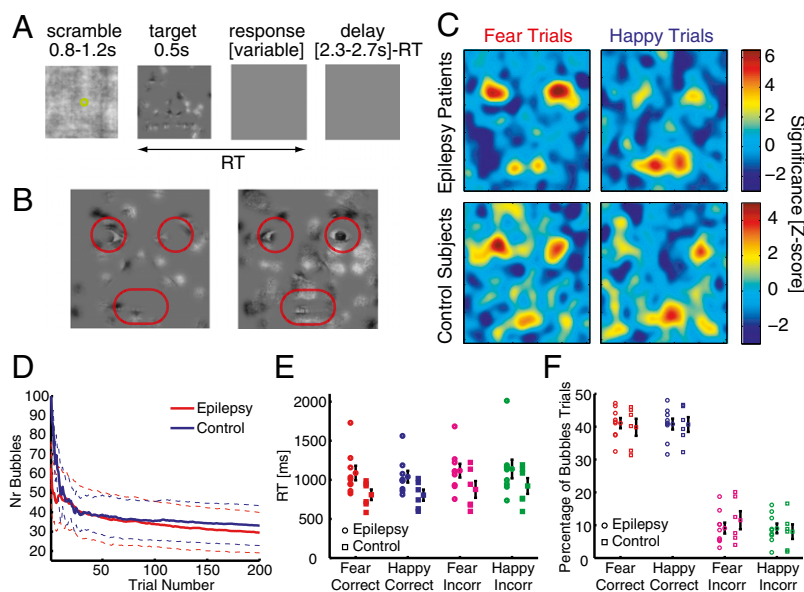


Fig. 1. Stimuli and behavioral performance. (A) Task structure. Immediately preceding the target image, a scrambled version of a face was presented for a variable time between 0.8 and 1.2 s. The target image was presented for 500 ms and showed either a fear (50%) or happy (50%) expression. Subjects indicated whether the presented face was happy or fear. (B) Example bubbles stimuli. The ROIs used for analysis are shown in red (not shown to subjects). (C) Behavioral classification images for fear- and happy-face 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, 1 session omitted here because the learning algorithm was disabled as a control; mean \pm 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. Fear correct: fear-face trials with a correct response; Happy correct: happy-face trials with a correct response; Fear incorrect: fear-face trials but incorrectly judged as happy; Happy incorrect: happy-face trials but incorrectly judged as fear. (F) Response choice for patients ($n = 9$ sessions, circles) and controls ($n = 6$ sessions, squares).

electrode and electrodes were thus likely sampling neurons in the midmedial part of the amygdala [basomedial nucleus or deepest part of the basolateral nucleus (36)]. The isolation criteria and other face-responsive characteristics of the same dataset were described previously (16, 17). 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 before 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 fear faces compared with happy faces. We selected emotion-selective units by comparing the total number of spikes in a time window 250- to 1,750-ms post-stimulus-onset between correct fear-face trials and correct happy-face 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 fear faces or happy faces separately, each with $\alpha = 0.05$. We found that 24 units showed significantly greater response to fear faces compared with 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 with fear faces. We refer to these units as neurons selective for fear expressions (“fear-selective” for short) (Fig. 2 *A* and *B*) and neurons selective for happy expressions (“happy-selective” for short) (Fig. 2 *C* and *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 (Table S2). However, we 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.

Fig. 2 shows four single-neuron examples (see Fig. S2 for more examples). The fear-selective neurons (Fig. 2 *A* and *B*) increased their activity for fear-face trials and decreased their activity in happy-face trials. In contrast, the happy-selective neurons (Fig. 2 *C* and *D*) increased their activity in happy-face trials. On average, significant differences in response between fear and happy faces appeared 625-ms post-stimulus-onset and lasted for up to 1.5 s (Fig. 3). For fear-selective neurons, the difference was mainly due to a suppression of activity in happy-face trials (Fig. 3*A*), whereas for happy-selective neurons, it was mainly due to an increase in activity for happy-face trials (Fig. 3*B*). A similar plot for all recorded neurons ($n = 185$, Fig. S3) showed no significant difference, indicating that overall mean activity was not different between the two conditions.

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 fear when shown the entire face for extended periods of time (37). 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 fear 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 Fig. 2 *A* and *B*). Similarly, when a happy face was incorrectly judged as fear, the neurons responded as if a fear face had been correctly judged as fear (compare green and red lines for the examples shown in Fig. 2 *C* and *D*). In Fig. 2*E*, lines connect the conditions with the

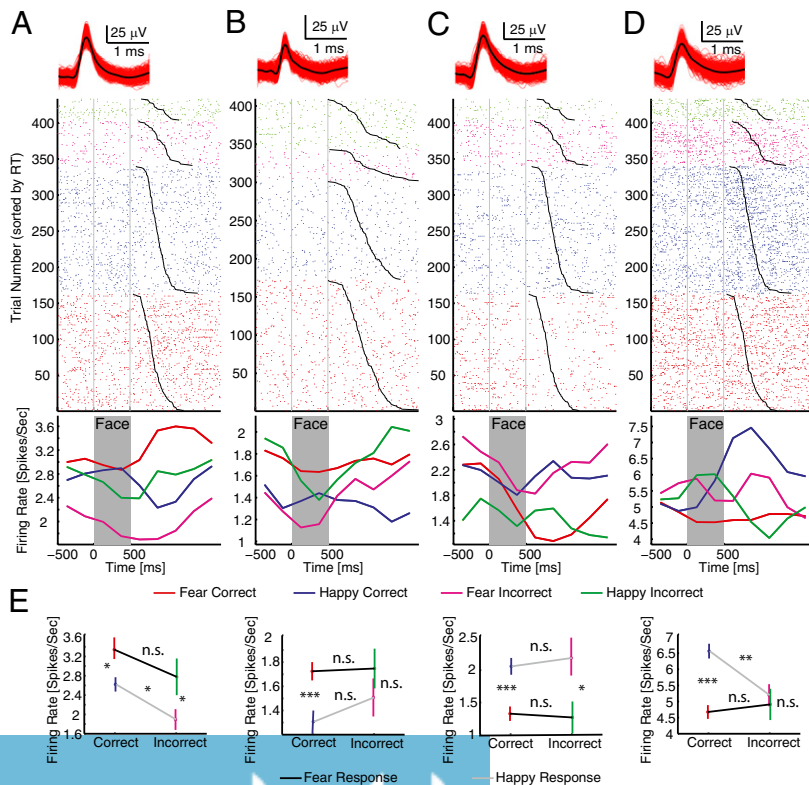


Fig. 2. Single-unit examples of emotion-selective neurons in the amygdala. (*A* and *B*) Example fear-selective neurons, which have a higher firing rate for correct fear-face trials compared with correct happy-face trials (selection *t* test: $P < 0.005$). (*C* and *D*) Example happy-selective neurons, which have a higher firing rate for correct happy-face trials compared with correct fear-face trials (selection *t* test: $P < 10^{-8}$). Each raster (Upper) and PSTH (Lower) 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 top of the raster plot. (*E*) Average firing rate 250- to 1,750-ms post-stimulus-onset for each unit. Red: fear-face trials with a correct response; blue: happy-face trials with a correct response; magenta: fear-face trials but incorrectly judged as happy; green: happy-face trials but incorrectly judged as fear. 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 \pm 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 ($P > 0.05$).

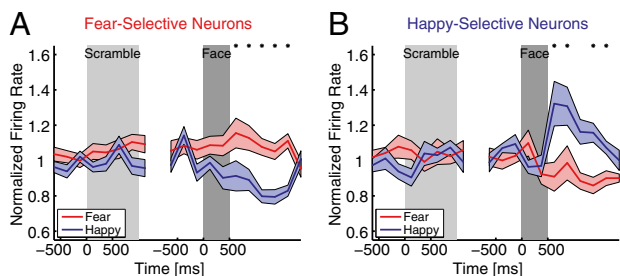


Fig. 3. Average PSTH of all emotion-selective neurons. (A) Mean response of all fear-selective units that increased their spike rate for correct fear-face trials compared with correct happy-face trials ($n = 24$ units; shaded area denotes \pm SEM; the firing rate was normalized to average baseline response 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 fear-face trials and happy-face trials in that bin ($P < 0.05$, two-tailed t test, Bonferroni-corrected).

same response (fear 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 fear-face correct trials compared with happy-face correct trials, it would also show a greater response in happy-face incorrect trials compared with fear-face incorrect trials. Thus, firing rate increased whenever a judgment of fear was made, regardless of whether it was correct or incorrect. The neuronal response of the examples shown in Fig. 2 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 (fear–happy) and accuracy of judgment (correct–incorrect) as tested by a 2×2 ANOVA with number of spikes fired in a 1.5-s window after stimulus onset (250- to 1,750-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 Fig. 2A, $F(1,405) = 7.09$ for Fig. 2B, $F(1,429) = 16.06$ for Fig. 2C, and $F(1,429) = 9.47$ for Fig. 2D]. We next quantified this phenomenon across the population.

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] \times [fear 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 fear-face trials and 13 of which responded with a higher firing rate in correct happy-face trials, hence denoted as fear interactive neurons and happy interactive neurons, respectively (Table S2).

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 (Fig. 4A and B). Each neuron contributed two data points: one for correct (red, blue) and one for incorrect trials (gray), respectively. By definition, fear interactive neurons increased their firing rate for correctly identified fear face trials (Fig. 4A, red). Similarly, happy interactive neurons increased their firing rate for correctly identified happy face trials (Fig. 4B, 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 (fear

interactive neurons: Fig. 4A; χ^2 -test on the number of neurons falling on each side of the diagonal line (gray bars), $P < 10^{-5}$; happy interactive neurons: Fig. 4B, $P < 0.01$). In each case, the mean of all incorrect trials from all neurons was on the opposite side of the diagonal shown in Fig. 4 from that for correct trials: the average normalized firing rate thus indicated the behavioral judgment of the subjects.

To summarize the population response, we next visualized the mean difference in response between fear and happy stimuli for both correct and incorrect trials (Fig. 4C). For fear interactive neurons, this response difference tended to be positive for correct and negative for incorrect trials (Fig. 4C, red) and vice versa for happy interactive neurons (Fig. 4C, blue). Thus, the response during incorrect trials tended to be similar to the correct trials

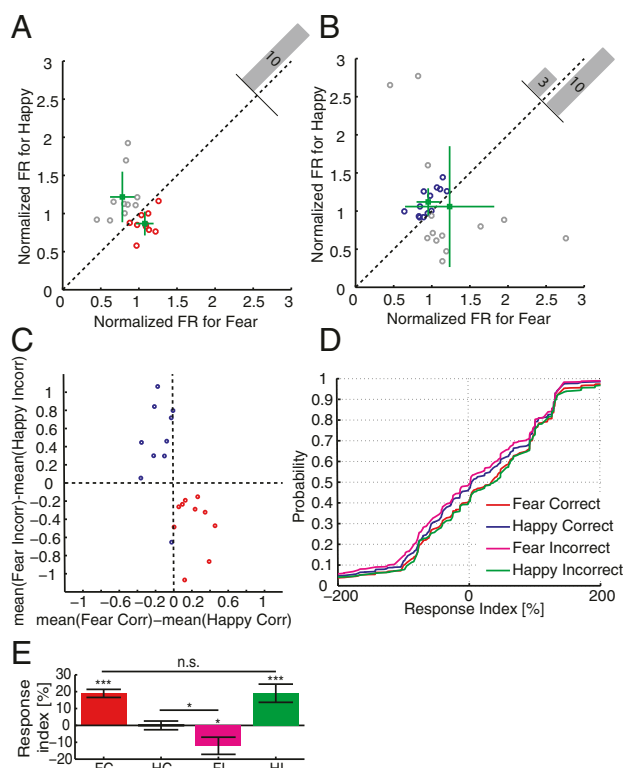


Fig. 4. Neuronal activity followed perceptual judgment rather than stimulus identity ($n = 23$ units, selected by a significant interaction term). (A and B) Scatter plot of mean normalized firing rate for fear- and happy-face trials, shown separately for fear interactive (A) and happy interactive (B) 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 with the correct trials. Error bars (green) are mean \pm 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 fear- and happy-face trials for correct (x-axis) and incorrect (y-axis) trials. Fear 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 (Methods). The response during incorrect trials was opposite 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 FC and HI trials, which shows that when a happy face was shown but judged as a fear face, the neurons responded as if a real fear face was shown. The same interpretation can be derived for HC trials and FI trials. FC: fear-face trials with a correct response; HC: happy-face trials with a correct response; FI: fear-face trials but incorrectly judged as happy; HI: happy-face trials but incorrectly judged as fear. Note that HC is equal to zero by definition (Methods). *: $P < 0.05$, **: $P < 0.01$, and ***: $P < 0.001$.

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 fear- or happy-face trials for each neuron. We used a simple response index R_i as a single-trial metric (Eq. 1), which takes into account the opposite signs of the two types of neurons—the fear 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-s 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-face trials divided by the average of the baseline (Eq. 1). For example, if a neuron doubles its firing rate for a fear 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. 2).

We next used 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 fear 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, Fig. 4D). Comparing the distributions using a cumulative distribution function (Fig. 4D; Methods) shows that the response during incorrect trials was similarly distributed compared with the correct trials of the opposite category. For example, happy-face incorrect trials (Fig. 4D, green curve) were similarly distributed to fear-face correct trials (Fig. 4D, red curve), and vice versa. Confirming this observation, there was no significant difference between happy-face incorrect and fear-face correct trials (KS test, $P = 0.62$) nor between fear-face incorrect and happy-face 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 fear-face correct and happy-face incorrect trials (in both cases the perceptual judgment was fear) had response indices significantly above 0 (Fig. 4E; two-tailed one-sample t test, $P < 10^{-13}$ for fear-face correct trials and $P < 0.0005$ for happy-face incorrect trials), and there was no significant difference between correct and incorrect trials for fear subjective judgments (two-tailed two-sample t test comparing 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 fear incorrect, $P = 0.027$), with fear-face incorrect trials significantly below 0 (t test against 0: $P < 0.05$). This was because the firing rate for fear-face incorrect trials was lower than it was for happy-face correct trials. Separate analyses for only fear- or happy-selective neurons led to similar results (Fig. S4), 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.

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 subgroup. To obtain a broader inventory, we next analyzed the previously described units ($n = 41$, among which 6 were fear interactive neurons and 3 were happy interactive neurons; Table S2) 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 in-

teractive 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 fear- and happy-face trials (Methods and Eq. S1). To assess statistical significance, we estimated the null distribution by first randomly shuffling the labels of trials (fear–happy) and then computing the population summary metric. We repeated this permutation test 1,000 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 (Fig. 5, gray). In contrast, the value of the population effect metric was 25.0% for correct trials [Fig. 5, 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); Fig. 5, blue]. Importantly, the metric from the incorrect trials was significantly negative and thus on the opposite side of the null distribution compared with the metric from correct trials (Fig. 5, blue). This shows that when the behavioral response was incorrect (opposite 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 (Fig. S5).

Were the results influenced by difficulty? The mean number of bubbles shown was 38.2 ± 34.2 (mean \pm 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 fear-correct vs. fear-incorrect and happy-correct vs. happy-incorrect). We found very similar results (Fig. S6 A–C), confirming that emotion-selective neurons signal the perceptual judgment independent of difficulty. We also re-

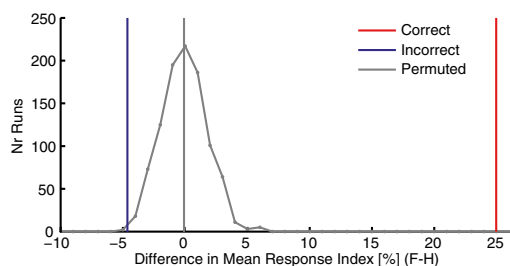


Fig. 5. Mean response across all emotion-selective neurons encoded subjective perceptual judgment. The gray distribution is the null distribution derived from a permutation test. The red and blue bars are the population summary metrics for correct and incorrect trials, respectively. Both were located 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). Note that the blue and red bars were located on opposite sides of the null distribution.

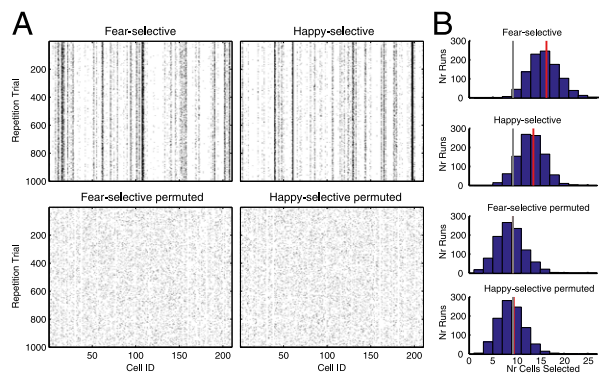


Fig. 6. Illustration of the split analysis method to compute the population response (see Fig. 7 for results). (A) Cells selected across runs. A black dot indicates that a particular cell was selected. There was substantial consistency of cells selected in the split analysis (Upper) but cell selection was evenly distributed across cells and runs in the permutation test (Lower). (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 whereas the number of cells selected in the permutation test was near chance.

peated the analysis by using equal numbers of trials for correct and incorrect to exclude any potential bias and we found very similar results (Fig. S6 D–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 (Fig. S6 G–I). Finally, two of the neurosurgical patients had a clinical diagnosis of autism (17). We repeated the analysis after excluding these two patients and again found very similar results (Fig. S6 J–L).

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 1,000 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 Fig. 5) 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 (1,000 runs) with randomly shuffled trial labels. However, 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 ensured our results against biases and false positives during selection because only out-of-sample errors were calculated.

Out of the 210 neurons recorded, we considered 185 cells with >0.2 -Hz firing rate for this analysis. Many cells were reliably selected over the 1,000 repetitions (Fig. 6A, Upper; 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 (Fig. 6A, Lower; 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. In contrast, for the

permutation test which randomly shuffled labels, each cell was equally likely to be selected with a probability of 0.05; the selected cells were evenly distributed across all 185 cells and across permutation runs (Fig. 6A, Lower) and did not show a bias toward those that could be selected with all trials. On average, 16.3 ± 3.1 (mean \pm 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; Fig. 6B). 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 (Fig. 6B, Middle), 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 (Fig. S7) showed that the permutation test was not biased.

We next quantified the responses of the groups of cells selected in each run using the population summary metric as described above (Fig. 7). The population summary metric is calculated as the difference between the average of response indices from all fear-face trials (either correct or incorrect) collapsed across all selected cells and the average of response indices from all happy-face trials (either correct or incorrect) collapsed across all selected cells (Eq. S1). 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 Fig. S7 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.

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 the predefined eye and mouth ROIs was revealed (shown in Fig. 1C). 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

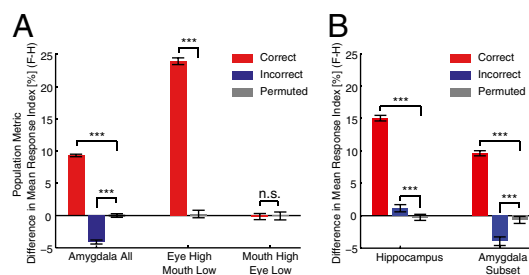


Fig. 7. Quantification of the population response using split analysis. (A) Emotion-selective neurons are primarily driven by information revealed by the eyes. (B) Hippocampal neurons also encode emotions but not subjective judgment. In contrast, a subset of amygdala neurons equal to the total number 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.

Mouth” (Fig. 7A; see Fig. S8 A and B for the distribution), and “Low Eye AND High Mouth” (Fig. 7A; see Fig. S8 C and 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. Because only correct trials were involved in the selection of ROI trials, the distributions in Fig. 7A 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.

Information conveyed by the eyes strongly modulated the neuronal response (Fig. 7A and Fig. S8 A and B). On average, 16.7 ± 3.2 (mean \pm SD) cells were selected as fear-selective neurons and 10.0 ± 2.4 cells were selected as happy-selective neurons, both significantly above chance ($P < 0.001$; Fig. S8E). 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 ($P < 0.0001$; Fig. 7A). The results held when analyzing fear-selective and happy-selective neurons separately (both $P < 0.0001$; Fig. S8 A and B).

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 (Fig. S8 C and D). On average, 10.8 ± 2.9 cells were selected as happy-selective neurons, which was significantly above chance ($P < 0.0001$; Fig. S8F). By contrast, only 8.3 ± 2.4 cells were selected as fear-selective neurons, which was significantly below the chance value of 9.25 ($P < 0.0001$; Fig. S8F), indicating that when eyes were absent and mouth was present, the neuronal response to fear faces was suppressed.

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

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 S1). Sixty-three 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 four fear-selective neurons (6.4%) and seven 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 1,000 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 fear-selective neurons were consistently selected across repetitions (Fig. S9A). Interestingly, only happy-selective neurons, but not fear-selective neurons were selected above chance (Fig. S9B). The selected neurons differentiated fear from happy faces in correct trials ($P < 0.001$; Fig. 7B; also see Fig. S9 D and E). Thus, a subset of hippocampal neurons distinguished happy from fear emotions in correct trials. Crucially, however, this was only the case for correct trials. In contrast with the amygdala neurons, the hippocampal neurons did not indicate the behavioral response made during incorrect tri-

als. Rather, the response indicated, albeit only weakly so (Fig. 7B), 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 with 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 (Fig. 7A). In conclusion, hippocampal neurons, unlike amygdala neurons, did not track the subjective judgment of facial emotion in incorrect trials.

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 with the entire population of amygdala neurons (Fig. 7B; 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.

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 found no significant correlation between firing rate and RT, and we found that the emotion-selective neurons were not lateralized or related to the output button response associated with the emotion (see SI Results for details). 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.

Discussion

In this study, we found that a subset of amygdala neurons encodes 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 fear faces and primarily used the mouth region to correctly judge happy faces, findings consistent with prior studies (34, 35). Forty-one out of 185 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 the 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 was 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 fear-incorrect; Fig. 4E). 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.

Possible Confounds. Our stimuli were based on the well-validated set of facial emotion images from Ekman and Friesen (37), from which we chose a subset depicting fear and happy emotions 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 sex 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 versus positive valence. Furthermore, because 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 (16)], 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 (38), 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 (38) 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; 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 (Fig. 6). The result (Fig. 7) 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 (Fig. 7), which is very unlikely to be driven by chance alone. Second, the sets of cells selected by the two different methods were comparable, showing that emotion-selective neurons were consistently selected even with a random subset of trials. Third, we rigorously established chance levels using permutation tests (Fig. 7) and found that the number of cells selected was well above chance (Fig. 6). 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 (Fig. 7B), showing that our results were not driven by a particular subset of neurons.

Selectivity of Amygdala Neurons. Faces can be readily characterized by independent attributes, such as identity, expression, and sex, which have segregated cortical representations (13, 39), and single-unit recordings in the primate amygdala have documented responses selective for faces, their identity, or emotional expression (10, 14). We previously showed that neurons in the human amygdala selectively respond to whole faces compared with facial parts, suggesting a predominant role of the amygdala in representing global information about faces (16). 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 with our previous neuronal classification images which were based on pixelwise analyses of face regions that drive neuronal response (17), 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 S2). 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 (18, 21, 34).

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 (40, 41), and there are reports of amygdala blood-oxygen-level dependent (BOLD) discrimination to the presentation of fear faces even if such faces are presented to patients in their blind hemifield in cases of hemianopia due to cortical lesions (42). 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 nonconscious 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 (43). 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 (44). In particular, because 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 (45) and electrical brain stimulation in fusiform areas in humans (46) have suggested a causal role of the temporal cortex in face categorization and perception. Future studies using 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.

Comparison with Neuroimaging Studies and Functional Role of the Amygdala. We further compare our study with neuroimaging studies and discuss the functional role of the amygdala in *SI Discussion*.

Conclusions

In conclusion, we suggest that the amygdala serves to integrate sensory information about faces, conveyed via temporal neocortex, with reward value (47), task, and social context (48), 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 (16) and others (43) 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 (49) and approachability (50). It will be critical to compare our findings to responses obtained from face-selective neurons in temporal cortex (51), which provide the primary visual input to the amygdala (52), and which in turn receive feedback from the amygdala (31). 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.

Methods

In this study we recorded single units from 10 neurosurgical patients who had chronically implanted depth electrodes in the amygdalae (Table S1). Three patients (total of three sessions) did not contribute well-isolated units and hence were excluded from analysis. Two patients completed two sessions, resulting in a total of nine recording sessions that we analyzed. All participants provided written informed consent according to protocols approved by the institutional review boards of the Huntington Memorial Hospital, Cedars-Sinai Medical Center, and the California Institute of Technology.

The subjects' electrophysiology as well as construction of bubbles stimuli, scrambled face stimuli, and classification images were described in our previous publications (16, 17).

Task. We used a facial emotion discrimination task in which patients were asked to judge fear or happy faces as quickly and accurately as possible from randomly selected parts of the face (bubbles; Fig. 1B). In each trial, a scrambled face with a central fixation circle was presented for 0.8–1.2 s (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 RT, the next trial started after an interval of 2.3–2.7 s after stimulus onset. If the patient did not respond by that time, a timeout was indicated by a beep (2.2% of trials were timeouts) (Fig. 1A). Each block contained 72 trials and patients completed 5–7 blocks. Timeout 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 eight face base images [chosen from the Ekman and Friesen stimulus set, four different individuals (two female and two male)] showing fear 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.

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-s 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 [poststimulus time histogram (PSTH); Figs. 2 and 3] were computed by counting spikes across all trials in consecutive 250-ms bins. 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 PSTHs 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).

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 (Figs. 2 and 3). 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 (fear or happy). Because only two emotions were shown, an incorrect trial always implies that the subject chose the opposite emotion.

The selection criterion for emotion-selective units (Figs. 2 and 3) 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 fear relative to happy faces after stimulus onset. By definition, fear-selective units responded significantly more in correct fear-face trials compared with correct happy-face 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] \times [fear stimuli vs. happy stimuli]) was used to probe for a significant interaction term with $P < 0.05$ (Fig. 2E).

Data Analysis: Response Index. We quantified for each neuron whether its response differed between fear-face and happy-face trials using a single-trial response index R_i (Eq. 1; Fig. 4). 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 (16, 53). The response index quantifies the response during trial i relative to the mean response to correct happy stimuli and baseline (a 1-s interval of blank screen right before scramble onset). The mean response and baseline was calculated individually for each unit.

$$R_i = \frac{FR_i - \text{mean}(FR_{\text{HappyCorrect}})}{\text{mean}(FR_{\text{Baseline}})} \cdot 100\%. \quad [1]$$

For each trial i , which can be either fear or happy, R_i is the baseline normalized firing rate (FR) during a 1.5-s 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 fear-face trials, the average value of R_i will be significantly different from 0. Because fear-selective neurons have more spikes in fear-face trials and happy-selective neurons have more spikes in happy-face 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. 2). This makes R_i 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.

$$R_i = - \frac{FR_i - \text{mean}(FR_{\text{HappyCorrect}})}{\text{mean}(FR_{\text{Baseline}})} \cdot 100\%. \quad [2]$$

After calculating R_i for every trial, we subsequently averaged all R_i s of trials that belong to the same category. We used four categories: fear correct (FC), fear 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-face correct trials (Eq. 2). The mean baseline firing rate was calculated across all trials. The same $FR_{\text{HappyCorrect}}$ was subtracted for both correct and incorrect trials.

The cumulative distribution function (CDF) (Fig. 4D and Fig. S4 A and 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 CDFs of fear and happy-face trials were compared using two-tailed two-sample KS tests. All error bars are \pm SE unless indicated otherwise.

Data Analysis: Split Analysis and Permutation Test. We used 1,000 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 of the incorrect trials for the selected cells. To summarize the population difference in response to fear compared with happy faces, we calculated a

summary population metric that provided a single number for a population of cells for every run of the permutation test (Figs. 5 and 7). The methods and equations are detailed in *SI Methods*.

We further quantified how sensitive neurons were to specific facial parts by repeating the permutation analysis with only a subset of trials that revealed the ROI of interest. The methods are detailed in *SI Methods*.

ACKNOWLEDGMENTS. We thank all patients for their participation; Drs. Linda Philpott and William Sutherland 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 the Conte Center from National Institute of Mental Health.

- Adolphs R (2010) What does the amygdala contribute to social cognition? *Ann N Y Acad Sci* 1191(1):42–61.
- Kling AS, Brothers LA (1992) The amygdala and social behavior. *The Amygdala: Neurobiological Aspects of Emotion, Memory and Mental Dysfunction*, ed Aggleton JP (Wiley-Liss, New York), pp 353–377.
- LeDoux JE (1993) Emotional memory systems in the brain. *Behav Brain Res* 58(1-2): 69–79.
- Rolls ET (1992) Neurophysiology and functions of the primate amygdala. *The Amygdala: Neurobiological Aspects of Emotion, Memory and Mental Dysfunction*, ed Aggleton JP (Wiley-Liss, New York), pp 143–165.
- 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(6507):669–672.
- Bechara A, et al. (1995) Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science* 269(5227): 1115–1118.
- LaBar KS, LeDoux JE, Spencer DD, Phelps EA (1995) Impaired fear conditioning following unilateral temporal lobectomy in humans. *J Neurosci* 15(10):6846–6855.
- Morris JS, et al. (1996) A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature* 383(6603):812–815.
- Kreiman G, Koch C, Fried I (2000) Category-specific visual responses of single neurons in the human medial temporal lobe. *Nat Neurosci* 3(9):946–953.
- Fried I, MacDonald KA, Wilson CL (1997) Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. *Neuron* 18(5):753–765.
- Fitzgerald DA, Angstadt M, Jelsone LM, Nathan PJ, Phan KL (2006) Beyond threat: Amygdala reactivity across multiple expressions of facial affect. *Neuroimage* 30(4): 1441–1448.
- 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(2):159–176.
- Rolls ET (1984) Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. *Hum Neurobiol* 3(4):209–222.
- Gothard KM, Battaglia FP, Erickson CA, Spitzer KM, Amaral DG (2007) Neural responses to facial expression and face identity in the monkey amygdala. *J Neurophysiol* 97(2):1671–1683.
- Hoffman KL, Gothard KM, Schmid MC, Logothetis NK (2007) Facial-expression and gaze-selective responses in the monkey amygdala. *Curr Biol* 17(9):766–772.
- Rutishauser U, et al. (2011) Single-unit responses selective for whole faces in the human amygdala. *Curr Biol* 21(19):1654–1660.
- Rutishauser U, et al. (2013) Single-neuron correlates of atypical face processing in autism. *Neuron* 80(4):887–899.
- Adolphs R, et al. (2005) A mechanism for impaired fear recognition after amygdala damage. *Nature* 433(7021):68–72.
- Sander D, et al. (2005) Emotion and attention interactions in social cognition: Brain regions involved in processing anger prosody. *Neuroimage* 28(4):848–858.
- Adolphs R (2008) Fear, faces, and the human amygdala. *Curr Opin Neurobiol* 18(2): 166–172.
- Gamer M, Büchel C (2009) Amygdala activation predicts gaze toward fearful eyes. *J Neurosci* 29(28):9123–9126.
- Sasson N, et al. (2007) Orienting to social stimuli differentiates social cognitive impairment in autism and schizophrenia. *Neuropsychologia* 45(11):2580–2588.
- Horley K, Williams LM, Gonsalves C, Gordon E (2004) Face to face: Visual scanpath evidence for abnormal processing of facial expressions in social phobia. *Psychiatry Res* 127(1-2):43–53.
- Pelphrey KA, et al. (2002) Visual scanning of faces in autism. *J Autism Dev Disord* 32(4):249–261.
- Baron-Cohen S, et al. (2000) The amygdala theory of autism. *Neurosci Biobehav Rev* 24(3):355–364.
- Hadj-Bouziane F, et al. (2012) Amygdala lesions disrupt modulation of functional MRI activity evoked by facial expression in the monkey inferior temporal cortex. *Proc Natl Acad Sci USA* 109(52):E3640–E3648.
- Moeller S, Freiwald WA, Tsao DY (2008) Patches with links: A unified system for processing faces in the macaque temporal lobe. *Science* 320(5881):1355–1359.
- 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(11):1271–1278.
- Desimone R, Gross CG (1979) Visual areas in the temporal cortex of the macaque. *Brain Res* 178(2-3):363–380.
- Amaral DG, Behnia H, Kelly JL (2003) Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey. *Neuroscience* 118(4): 1099–1120.
- Freese JL, Amaral DG (2006) Synaptic organization of projections from the amygdala to visual cortical areas TE and V1 in the macaque monkey. *J Comp Neurol* 496(5): 655–667.
- Ghashghaei HT, Barbas H (2002) Pathways for emotion: Interactions of prefrontal and anterior temporal pathways in the amygdala of the rhesus monkey. *Neuroscience* 115(4):1261–1279.
- Gosselin F, Schyns PG (2001) Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Res* 41(17):2261–2271.
- Scheller E, Büchel C, Gamer M (2012) Diagnostic features of emotional expressions are processed preferentially. *PLoS ONE* 7(7):e41792.
- Smith ML, Cottrell GW, Gosselin F, Schyns PG (2005) Transmitting and decoding facial expressions. *Psychol Sci* 16(3):184–189.
- Oya H, Kawasaki H, Dahdaleh NS, Wemmie JA, Howard MA, 3rd (2009) Stereotactic atlas-based depth electrode localization in the human amygdala. *Stereotact Funct Neurosurg* 87(4):219–228.
- Ekman P, Friesen WV (1975) *Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues* (Prentice Hall, Englewood Cliffs, NJ).
- Peck CJ, Lau B, Salzman CD (2013) The primate amygdala combines information about space and value. *Nat Neurosci* 16(3):340–348.
- Perrett DI, et al. (1984) Neurones responsive to faces in the temporal cortex: Studies of functional organization, sensitivity to identity and relation to perception. *Hum Neurobiol* 3(4):197–208.
- Whalen PJ, et al. (1998) Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *J Neurosci* 18(1):411–418.
- Morris JS, Ohman A, Dolan RJ (1998) Conscious and unconscious emotional learning in the human amygdala. *Nature* 393(6684):467–470.
- 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(Pt 6):1241–1252.
- Mormann F, et al. (2008) Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. *J Neurosci* 28(36):8865–8872.
- 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(11):773–783.
- Afraz S-R, Kiani R, Esteky H (2006) Microstimulation of inferotemporal cortex influences face categorization. *Nature* 442(7103):692–695.
- Parvizi J, et al. (2012) Electrical stimulation of human fusiform face-selective regions distorts face perception. *J Neurosci* 32(43):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(7078):865–870.
- Kim H, et al. (2004) Contextual modulation of amygdala responsivity to surprised faces. *J Cogn Neurosci* 16(10):1730–1745.
- Adolphs R, Tranel D, Damasio AR (1998) The human amygdala in social judgment. *Nature* 393(6684):470–474.
- Kennedy DP, Gläscher J, Tyszka JM, Adolphs R (2009) Personal space regulation by the human amygdala. *Nat Neurosci* 12(10):1226–1227.
- Tsao DY, Freiwald WA, Tootell RBH, Livingstone MS (2006) A cortical region consisting entirely of face-selective cells. *Science* 311(5761):670–674.
- Amaral DG, Price JL, Pitkanen A, Carmichael ST (1992) Anatomical organization of the primate amygdaloid complex. *The Amygdala: Neurobiological Aspects of Emotion, Memory and Mental Dysfunction*, ed Aggleton JP (Wiley-Liss, New York), pp 1–66.
- Rutishauser U, Schuman EM, Mamelak AN (2008) Activity of human hippocampal and amygdala neurons during retrieval of declarative memories. *Proc Natl Acad Sci USA* 105(11):329–334.