

# Shared neural coding for social hierarchy and reward value in primate amygdala

Jérôme Munuera<sup>1\*</sup>, Mattia Rigotti<sup>2</sup> and C. Daniel Salzman<sup>1,3,4,5,6\*</sup>

**The social brain hypothesis posits that dedicated neural systems process social information. In support of this, neurophysiological data have shown that some brain regions are specialized for representing faces. It remains unknown, however, whether distinct anatomical substrates also represent more complex social variables, such as the hierarchical rank of individuals within a social group. Here we show that the primate amygdala encodes the hierarchical rank of individuals in the same neuronal ensembles that encode the rewards associated with nonsocial stimuli. By contrast, orbitofrontal and anterior cingulate cortices lack strong representations of hierarchical rank while still representing reward values. These results challenge the conventional view that dedicated neural systems process social information. Instead, information about hierarchical rank—which contributes to the assessment of the social value of individuals within a group—is linked in the amygdala to representations of rewards associated with nonsocial stimuli.**

Human and nonhuman primates tend to live in large social groups that possess dominant and subordinate members<sup>1</sup>. In many settings, such as interviewing for a job or attending a family reunion, appropriate social interactions and emotional behavior depend upon an assessment of the hierarchical ranks of both oneself and the other people present. The hierarchical ranks of individuals and the accurate evaluation of the hierarchical ranks of others affect the ability to thrive—for example, by influencing fluid and food access, mating priority and defensive behavior<sup>2–5</sup>.

The assessment of hierarchical rank requires accurate identification of social group members. Primates are masters at recognizing identity based on faces<sup>6,7</sup>. Brain areas engaged during face processing include the occipital, fusiform and superior temporal sulcus face areas, as well as the amygdala and orbitofrontal cortex<sup>8–15</sup>. The superior temporal sulcus is anatomically connected to brain areas implicated in social processing, including the anterior cingulate and orbitofrontal cortices (ACC and OFC) and the amygdala<sup>16–18</sup>. These connections may link representations of facial identity to neural circuits processing motivational and emotional information about nonsocial stimuli. For example, the amygdala, OFC and ACC all provide neural representations of the value of conditioned stimuli acquired through reinforcement learning<sup>19–21</sup>.

How and where neural representations of social variables more complex than facial identity—such as hierarchical rank—emerge in the brain is an enduring mystery. The social brain hypothesis posits that the evolution of large brains in mammals reflects increasingly complex social demands<sup>22</sup>, a proposal consistent with the discovery in primates of brain areas specialized for representing faces. In further support of this hypothesis, gray matter brain volume in human and nonhuman primates varies as a function of social status and/or social experiences<sup>23–27</sup>. However, the extent to which more complex social variables such as hierarchical rank are processed in specialized neural circuits remains unclear. Human neuroimaging studies indicate that the amygdala helps establish artificial hierarchies acquired during performance of a task in the laboratory<sup>28–32</sup>,

but the amygdala's role in representing naturally established hierarchies has not been demonstrated. It is possible that some social variables, such as the propensity to share rewards with a partner, may engage distinct neuronal ensembles<sup>14,33–36</sup>, but other social variables may not engage neuronal circuits dedicated to processing social information alone. Indeed, rewarding stimuli and stimuli that possess hierarchical information have been observed to modulate the same neuronal ensembles in brain areas such as OFC, striatum and the lateral intraparietal area<sup>12,14,35,37</sup>. In these experiments, however, neural responses to stimuli spanning a full social hierarchy were not examined, leaving unclear whether the studied neural ensembles in fact represent hierarchical rank per se.

Monkeys (and other primates) almost certainly learn the social rank of other group members through experience. This learning process may entail assigning group members different 'values' through experience, a process bearing resemblance to learning about the value of nonsocial stimuli. Here we provide evidence that the primate amygdala encodes the hierarchical rank of individuals in the same neuronal ensembles that also encode the rewards associated with nonsocial stimuli. By contrast, despite representing the rewards associated with nonsocial stimuli, OFC and ACC lack strong representations of hierarchical rank. These results challenge strong versions of the social brain hypothesis<sup>22</sup>, the traditional view that the processing of social stimuli occurs exclusively in dedicated neural systems. Instead, although some aspects of social processing may occur in distinct neural circuits, some complex social variables, such as hierarchical rank, may be represented by neuronal ensembles that also process nonsocial stimuli.

## Results

Two rhesus monkeys who had lived in the same stable group of ten monkeys for many years performed a task in which they viewed, in different blocks of trials, images of the faces of other members of their group or fractal images predicting different reward amounts (Fig. 1a,b). The face images were neutral in expression

<sup>1</sup>Department of Neuroscience, Columbia University, New York, NY, USA. <sup>2</sup>IBM T.J. Watson Research Center, Yorktown Heights, NY, USA. <sup>3</sup>Kavli Institute for Brain Sciences, Columbia University, New York, NY, USA. <sup>4</sup>Department of Psychiatry, Columbia University, New York, NY, USA. <sup>5</sup>New York State Psychiatric Institute, New York, NY, USA. <sup>6</sup>Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA. \*e-mail: [jerome.munuera@gmail.com](mailto:jerome.munuera@gmail.com); [cgs2005@columbia.edu](mailto:cds2005@columbia.edu)

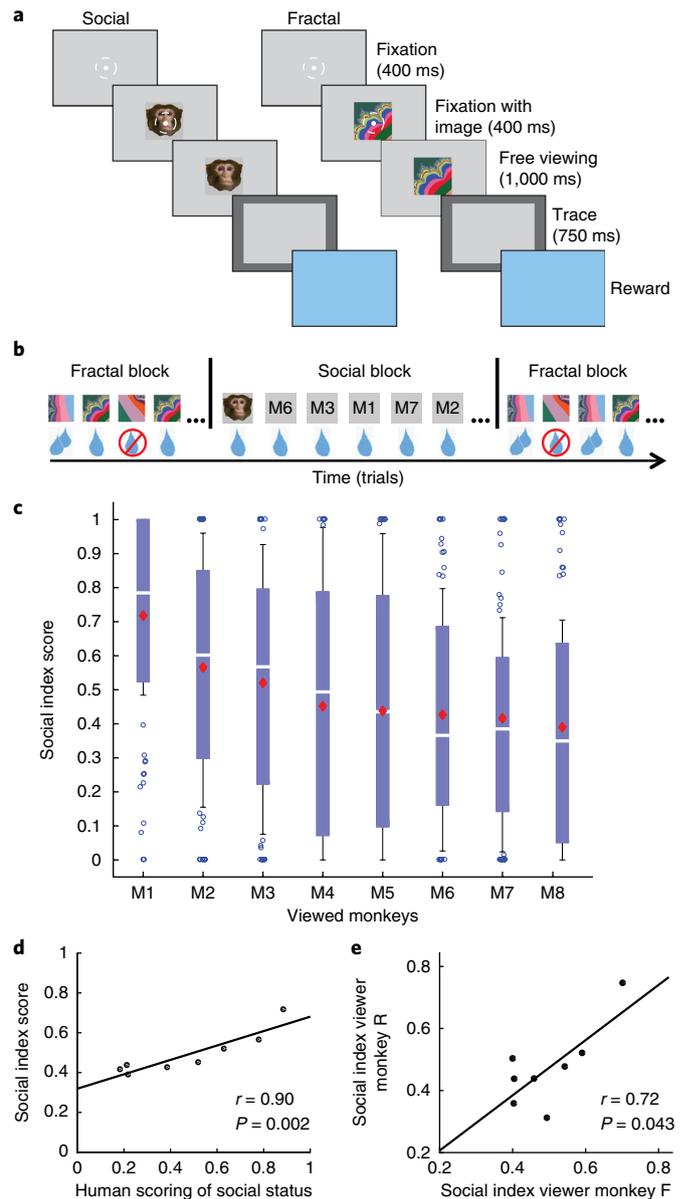
and did not depict emotionally relevant expressions (see Methods). Viewer monkeys fixated for the first 400 ms of image presentation for all trials and then were permitted free viewing of the images for an additional 1,000 ms. Successful fixation resulted in delivery of one drop of liquid reward on face image trials or zero, one or two drops of liquid reward on fractal trials (depending upon the image). Anticipatory licking demonstrated that monkeys learned the different reward associations of the fractal images (Kruskal-Wallis test,  $P < 1 \times 10^{-5}$ , Supplementary Fig. 1).

**The assessment of social hierarchy.** Human observers scored hierarchical rank using established methodology that determined the winners of agonistic interactions between monkeys (Table 1 and Supplementary Table 1)<sup>38</sup>. We also used the oculomotor behavior of viewer monkeys during electrophysiological recording sessions to construct an index of the monkeys' assessment of hierarchical rank of the face images. Several measures of viewer monkeys' oculomotor behavior varied depending upon the face image being displayed (Fig. 1c and Methods)<sup>39</sup>. These measures capitalized on the fact that monkeys tended to look at the face images of dominant monkeys, that viewing dominant monkeys can be more arousing, and that the propensity to look at the eyes of an individual is modulated by hierarchical rank (Methods and Supplementary Fig. 2)<sup>40</sup>. The social status indices for each viewer monkey were significantly correlated ( $r = 0.72$ ,  $P = 0.043$ ; Fig. 1e), perhaps reflecting that the two viewer monkeys both were in the middle of the hierarchy (Table 1). Furthermore, the index constructed from oculomotor measures was correlated with the human scoring of hierarchical rank ( $r = 0.90$ ,  $P = 0.002$ ; Fig. 1d) and was not significantly correlated with monkey age or weight ( $r = 0.41$ ,  $P = 0.305$  and  $r = 0.071$ ,  $P = 0.867$ , respectively). A behavioral experiment showed that the oculomotor measures obtained from viewer monkeys did not distinguish dominant from submissive monkeys if the faces displayed on the screen were unfamiliar to the viewer monkeys (Supplementary Fig. 3). The correlation of the social status index with a traditional scoring method of hierarchical organization by human observers provides strong evidence that during the electrophysiological recordings the viewer monkeys were actively assessing the hierarchical ranks of the monkeys depicted in the facial images.

### Neural representations of social hierarchy and nonsocial value.

While monkeys performed the fixation task, we recorded the responses of individual neurons in the amygdala (196 neurons), OFC (134 neurons) and ACC (187 neurons) to face images (Fig. 2a–c, left, and Supplementary Fig. 4). Across the populations of neurons, the proportion of neurons responsive to faces during fixation was higher in the amygdala than in OFC ( $z$ -test,  $z = 3.139$ ,  $P = 0.002$ ) or in ACC ( $z$ -test,  $z = 4.031$ ,  $P < 1 \times 10^{-4}$ ) (Fig. 2d, top row). OFC and ACC were statistically indistinguishable in this measure ( $z$ -test,  $z = 0.614$ ,  $P = 0.539$ ). However, responses to faces were usually not restricted to a single face image: 85.7% of face-responsive neurons in the amygdala responded to more than one face (80.2% of neurons in OFC and 59.4% of neurons in ACC exhibited the same property; Supplementary Fig. 5a). Only a minority of neurons responded to only a single face image (for example, Supplementary Fig. 5b). The proportion of neurons exhibiting a significant response to each face image did not differ across the face images for any brain area (chi-squared test,  $\chi^2 = 2.421$ ,  $P > 0.9$  in amygdala;  $\chi^2 = 3.930$ ,  $P > 0.5$  in OFC;  $\chi^2 = 2.964$ ,  $P > 0.5$  in ACC; Fig. 2e).

Although neurons often responded to more than one face image, an analysis using a linear decoder demonstrated that the population of amygdala neurons could be used to selectively discriminate which facial image was being viewed shortly after image onset (Fig. 2f and Methods)<sup>41</sup>. Decoding performance for face image identity was markedly superior in the amygdala relative to OFC or ACC (Fig. 2f, left), suggesting that on average the amygdala



**Fig. 1 | Task and behavioral measures.** **a**, Sequence of events in behavioral task. Monkeys view images of the faces of other monkeys who reside in their housing room (left) or fractal images (right). Successful fixation results in reward delivery. **b**, Behavioral task interleaving social and nonsocial (fractal) trial blocks. In fractal blocks, different images are associated with different reward amounts, unlike in the social block. **c**, Social index from behavioral measures of hierarchical assessment observed during the recording sessions plotted for each viewed monkey, M1 through M8, where M1 is the alpha monkey and M8 the most submissive one. Kruskal-Wallis test ( $\chi^2(7,535) = 47.11$ ,  $P < 1 \times 10^{-7}$ ), Dunn's post hoc: M1 different from M3 ( $P = 0.023$ ), M4 ( $P < 1 \times 10^{-3}$ ), M5 ( $P < 1 \times 10^{-3}$ ), M6 ( $P < 1 \times 10^{-4}$ ), M7 ( $P < 1 \times 10^{-4}$ ), M8 ( $P < 1 \times 10^{-4}$ ); M2 versus M8,  $P = 0.074$ ). Data represent the average of the four behavioral measures used to compute the social index across the sessions ( $n = 17$ ) (see Methods and Supplementary Fig. 2). Red diamonds represent mean, white lines represent median, blue bars and whiskers represent 75th and 85th percentiles respectively, blue circles represent data points beyond the whiskers limits. **d**, Social index score plotted as a function of the scoring of the colony hierarchy by a human observer (see Methods, Pearson's correlation coefficient (two-sided),  $r = 0.90$ ,  $P = 0.002$ ). **e**, Index score for one viewer monkey plotted against the other (Pearson's correlation coefficient (two-sided),  $r = 0.72$ ,  $P = 0.043$ ).

**Table 1 | Hierarchical rank determined by human observers**

Monkey	Number of won agonistic interactions observed	Number of agonistic interactions	Relative dominance status
Viewed: M1	61	69	88.4
Viewed: M2	53	68	77.9
Viewed: M3	22	35	62.9
Viewer: MF	45	83	54.2
Viewed: M4	27	51	51.9
Viewer: MR	42	94	44.7
Viewed: M5	10	47	21.3
Viewed: M6	17	44	38.6
Viewed: M7	9	49	18.4
Viewed: M8	12	55	21.8

Summary of the overall observed social interactions used in the computation of the social status scored by human observer. MF and MR are the two viewer monkeys, monkeys F and R.

communicated more facial identity information per neuron than the prefrontal areas.

We next explored whether a neural representation of hierarchical rank was present in each brain area. A regression analysis was used to quantify the relationship between neural responses during fixation of the face images (100–400 ms after image onset) and the social position (1–8) of the viewed face, where position 1 corresponds to the most dominant monkey (M1). The regression coefficient  $\beta_{\text{soc}}$  characterized neuronal tuning to the hierarchical rank of face images. For the amygdala neuron depicted in Fig. 2a (left),  $\beta_{\text{soc}}$  was significantly less than 0 ( $P < 10^{-4}$ ), indicating that the neuron tended to fire more strongly to face images from dominant monkeys. Overall, 30.6% of amygdala neurons, corresponding to 36% of face-responsive neurons, exhibited statistically significant selectivity for hierarchical rank (Supplementary Table 2). Out of those 30.6% neurons, 63.6% responded more strongly to dominant face images and 36.4% more strongly to submissive face images (Supplementary Table 2). Neurons in OFC and ACC exhibited selectivity for hierarchical rank much less frequently (16.7% of OFC neurons, representing 23% of face-responsive OFC neurons, and 6.3% of ACC neurons, representing 9% of face-responsive ACC neurons).

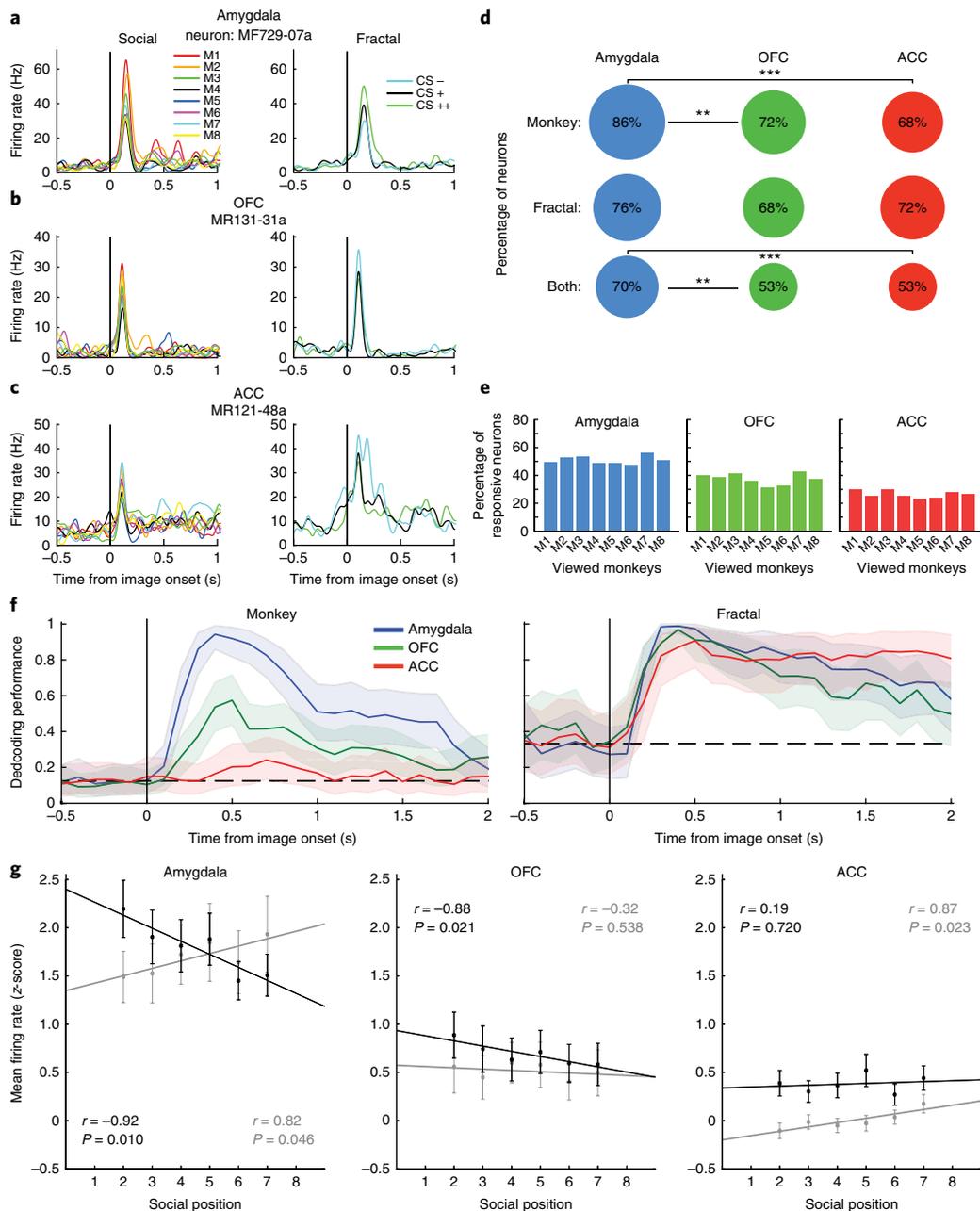
We confirmed the presence of a representation of hierarchical rank across the population of amygdala cells by performing an out-of-sample test: we analyzed separately the neurons that responded more strongly to M1 (most dominant monkey) than to M8 (most submissive monkey) and those that responded more strongly to M8 than M1 (see Methods). We then assessed whether neural responses to faces from the remaining out-of-sample viewed monkeys (M2–M7) were correlated with hierarchical rank. For the amygdala, responses to monkey face images correlated significantly with the social position of viewed monkeys ( $P = 0.010$  and  $P = 0.046$  for dominant- and submissive-monkey-preferring neurons, respectively; Fig. 2g, left). These results indicate that neurons classified according to preferential responding to M1 or M8 exhibit correlated neural activity with hierarchical ranking when only considering responses to M2–M7. The relationship between neural activity and hierarchical rank is therefore not merely driven by responses to the most dominant (M1) and most submissive (M8) monkeys.

Unlike the amygdala, OFC and ACC did not provide a compelling representation of social hierarchy (Fig. 2g, middle and right). Moreover, only the amygdala exhibited regressions with oppositely signed slopes for dominant- and submissive-preferring neurons. We used a Fisher's  $r$  to  $z$  transformation test to assess the statistical significance of the difference between the correlation coefficients in

each brain area. This test revealed a statistically significant difference between the amygdala and ACC for dominant-preferring neurons and between the amygdala and OFC for submissive-preferring neurons (difference between correlations for dominant-preferring neurons: amygdala ( $r = -0.92$ ) vs. OFC ( $r = -0.88$ ),  $z = -0.26$ ,  $P = 0.398$ ; amygdala ( $r = -0.92$ ) vs. ACC ( $r = 0.19$ ),  $z = -2.18$ ,  $P = 0.015$ ; OFC ( $r = -0.88$ ) vs. ACC ( $r = 0.19$ ),  $z = -1.92$ ,  $P = 0.028$ ; for submissive-preferring neurons: amygdala ( $r = 0.82$ ) vs. OFC ( $r = -0.32$ ),  $z = 1.82$ ,  $P = 0.034$ ; amygdala ( $r = 0.82$ ) vs. ACC ( $r = 0.87$ ),  $z = 0.22$ ,  $P = 0.413$ ; OFC ( $r = -0.32$ ) vs. ACC ( $r = 0.87$ ),  $z = -2.04$ ,  $P = 0.021$ ). We also combined the data from dominant-preferring and submissive-preferring neurons by sign-flipping the data from the submissive-preferring neurons and then averaging these responses with those of dominant-preferring neurons. A correlation coefficient of these combined data was computed for each area. The amygdala was the only area significantly correlated with the social position of the viewed monkeys ( $r = -0.91$ ,  $P = 0.013$  for amygdala;  $r = -0.61$ ,  $P = 0.20$  for OFC; and  $r = -0.60$ ,  $P = 0.21$  for ACC).

Our electrophysiological recordings were obtained from the same parts of the amygdala, OFC and ACC shown previously to contain individual neurons that respond differentially to a conditioned stimulus (CS) depending upon its reward association<sup>19,20,41</sup>, a finding we also observed in this study (Fig. 2a–c, right). We examined the encoding of reward magnitudes associated with fractal images by using a regression analysis to estimate a coefficient ( $\beta_{\text{frac}}$ ) quantifying the relationship between neural activity and the reward associated with fractals (see Methods).  $\beta_{\text{frac}}$  was significantly different from 0 in 38.9%, 25.8% and 39.2% of neurons in amygdala, OFC and ACC, respectively, with different neurons exhibiting positive or negative valence preference (Supplementary Table 2). The proportion of neurons responsive to one or more fractal images was not statistically distinguishable across the three brain structures (Fig. 2d, middle row, amygdala vs. OFC,  $z = 1.517$ ,  $P = 0.129$ ; amygdala vs. ACC,  $z = 0.855$ ,  $P = 0.393$ ; OFC vs. ACC,  $z = -0.723$ ,  $P = 0.470$ ;  $z$ -test for all comparisons), but the proportion of neurons responsive to one or more face and one or more fractal images was significantly greater in the amygdala than OFC and ACC, which were indistinguishable from each other (Fig. 2d, bottom row;  $z$ -test: amygdala vs. OFC,  $z = 3.126$ ,  $P = 0.002$ , amygdala vs. ACC,  $z = 3.411$ ,  $P < 1 \times 10^{-3}$ ; OFC vs. ACC,  $z = 0.008$ ,  $P = 0.994$ ). A linear decoder analysis revealed that neuronal populations in all three brain areas could be used to decode fractals associated with different rewards beginning shortly after image onset, consistent with prior observations (Fig. 2f, right)<sup>41</sup>.

**The same neuronal ensembles in the amygdala represent social hierarchy and reward value.** We next hypothesized that neural representations of hierarchical rank and of the rewards associated with nonsocial (fractal) images might be encoded by overlapping neuronal ensembles. To test this hypothesis, we trained a linear decoder to distinguish between trials when fractals were associated with large or no reward. The training procedure sets the weight of each neuron as well as a global threshold for distinguishing between the two conditions used for training (Fig. 3a). The decoder was then tested on held-out trials of the same types, as well as on trials in which monkeys viewed two different selected monkey face images (Fig. 3b). This decoding analysis was performed in a 300-ms sliding window (see Methods). As expected, decoding performance was high in all three brain areas when testing occurred on held-out fractal trials (Fig. 3c–e), consistent with previously described neural representations of expected reward in all three brain areas<sup>41</sup>. When the decoder was trained to distinguish between trials with fractals associated with large or no reward and was tested to distinguish between trials in which two monkey faces were presented that differed greatly in hierarchical rank (M1 vs. M8), decoding performance in the amygdala was significantly above chance shortly after



**Fig. 2 | The amygdala represents the hierarchical status of individuals. a–c.** Example neurons from the amygdala (**a**), OFC (**b**) and ACC (**c**) that respond to both face (left) and fractal images (right). The amygdala neuron responds most strongly to the dominant monkey face image and to the most highly rewarded fractal image, with responses decreasing monotonically with decreases in the hierarchical status of face images or the reward associated with fractals. **d**, The percentage of neurons in the three brain structures that respond to at least one monkey, one fractal or one of both types of image.

\*\* $P < 0.01$  and \*\*\* $P < 0.0001$  (two-tailed z-test, top row: amygdala versus OFC,  $z = 3.139$ ,  $P = 0.002$ ; amygdala vs. ACC,  $z = 4.031$ ,  $P < 1 \times 10^{-4}$ ; OFC vs. ACC,  $z = 0.614$ ,  $P = 0.539$ ; middle row: amygdala vs. OFC,  $z = 1.517$ ,  $P = 0.129$ ; amygdala vs. ACC,  $z = 0.855$ ,  $P = 0.393$ ; OFC vs. ACC,  $z = -0.723$ ,  $P = 0.470$ ; bottom row: amygdala vs. OFC,  $z = 3.126$ ,  $P = 0.002$ , amygdala vs. ACC,  $z = 3.411$ ,  $P < 1 \times 10^{-3}$ ; OFC vs. ACC,  $z = 0.008$ ,  $P = 0.994$ . Analyses have been performed on the entire neuronal population for each brain area: amygdala,  $n = 196$ ; OFC,  $n = 134$ ; ACC,  $n = 187$ ; see Methods). **e**, The proportion of neurons in each brain area that exhibit a significant response to each of the eight viewed faces. **f**, Average decoding performance for each brain area for classifying monkey image (left) or fractal (right) identity ( $n = 1,000$  iterations). Left, curves show the decoding performance for an eight-way classification of monkey face images (equalized number of neurons for the three brain areas,  $n = 110$  neurons). Right, average decoding performance for each brain structure of discriminating between the three fractal images (three-way classifier; equalized number of neurons for the three brain areas,  $n = 131$  neurons). Shading, 95% confidence intervals (bootstrap). Dotted lines, chance decoding level. **g**, Average z-scored firing rate plotted as a function of hierarchical status for neurons that respond more strongly (black line and dots) or weakly (gray line and dots) to a more highly ranked monkey (M1) than to a lower ranked monkey (M8) in amygdala (left), OFC (center) and ACC (right). Since neural responses to M1 and M8 are used to classify the cells, these data are excluded from the regression analyses. Pearson's correlation coefficient (two-sided): amygdala,  $r = -0.92$ ,  $P = 0.010$ ,  $n = 112$  neurons (black),  $r = 0.82$ ,  $P = 0.046$ ,  $n = 83$  (gray); OFC,  $r = -0.88$ ,  $P = 0.021$ ,  $n = 79$  (black),  $r = -0.32$ ,  $P = 0.538$ ,  $n = 49$  (gray); ACC,  $r = 0.19$ ,  $P = 0.720$ ,  $n = 97$  (black),  $r = 0.87$ ,  $P = 0.023$ ,  $n = 87$  (gray). Error bars, s.e.m. Neural responses were analyzed during the time epoch extending from 100–400 ms after image onset for all panels in this figure; fixation was required during this interval.

image onset and extending through the fixation interval (Fig. 3f). During the beginning of free viewing, however, decoding performance was significantly below chance, indicating that the opposite classification was made in this time interval, but that a relationship still existed between the encoding of fractals associated with different rewards and hierarchical rank. By contrast, when the same analysis was applied to neural responses to monkeys who had similar hierarchical rank (M4 and M5), decoding performance was barely different from chance (Fig. 3i). The same analysis performed on OFC and ACC neurons failed to classify hierarchical status when training the decoder on fractal image trials (Fig. 3g,h,j,k).

We used the analysis in Fig. 3f to define two time epochs: decoder above chance, which occupied 100–400 ms after image onset and occurred during fixation; and decoder below chance, which occupied 400–700 ms after image onset (0–300 ms after fixation point extinction). We then trained the decoder on large-reward and no-reward fractal image trials separately during each epoch and tested decoding for all pairs of face images during the corresponding time intervals, using the neuronal populations recorded from each brain area. In general, for the amygdala, decoding performance became increasingly different from chance with increasing distance between the social positions of two monkeys (Fig. 4a). Moreover, consistent with the analysis in which the decoder was tested on M1 vs. M8 (Fig. 3f), there was an opposite sign in decoding accuracy for the two time epochs. The same analysis applied to OFC and ACC demonstrated markedly less robust decoding of hierarchical status (Fig. 4b,c). These data indicate that the ability to decode the reward magnitude associated with an image does not necessarily result in the capacity to decode hierarchical rank, although it does so in the amygdala.

Decoder performance was quantitatively related to the social status index calculated from oculomotor behavior (see Fig. 1c). For each face image, we computed the average decoding performance for that image compared to each of the other faces when the decoder was trained to classify fractal images associated with large or no reward. In the amygdala, this average decoding performance was correlated with the computed social status index for that face image ( $P=0.006$  and  $P=0.003$ , Fig. 4d and Fig. 4g, respectively). The regression slopes were opposite in the two time epochs, consistent with a change in sign in relationship between reward value and hierarchical status encoding at the ensemble level. A significant relationship between average decoding performance and social status index was not observed for OFC and ACC ( $P=0.816$  for OFC, Fig. 4e;  $P=0.298$  for ACC, Fig. 4f;  $P=0.234$  for ACC, Fig. 4i), except for during the second time epoch in OFC ( $P=0.042$ , Fig. 4h). In addition, decoding performance between monkey face images was linearly correlated with the social distance between monkeys in the amygdala for the first and second time epochs, again with opposite signs for the two time epochs. (This analysis resampled the identities of the monkeys in the data points; Supplementary Fig. 6,  $r=0.84$ ,  $P=0.018$  and  $r=-0.92$ ,  $P=0.003$ .)

We considered two possible explanations for the intriguing change in sign of the relationship between the decoding of trials with different reward magnitudes and the decoding of hierarchical rank in the amygdala that occurred as a function of time (during and after required fixation). One possibility was that individual neurons may change their response preferences over time with respect to either hierarchical rank or expected reward magnitude. Alternatively, response preferences might be preserved, but instead different subensembles of neurons might contribute to the relationship between hierarchical rank and reward expectation during the two time intervals. The evidence supports this latter possibility.

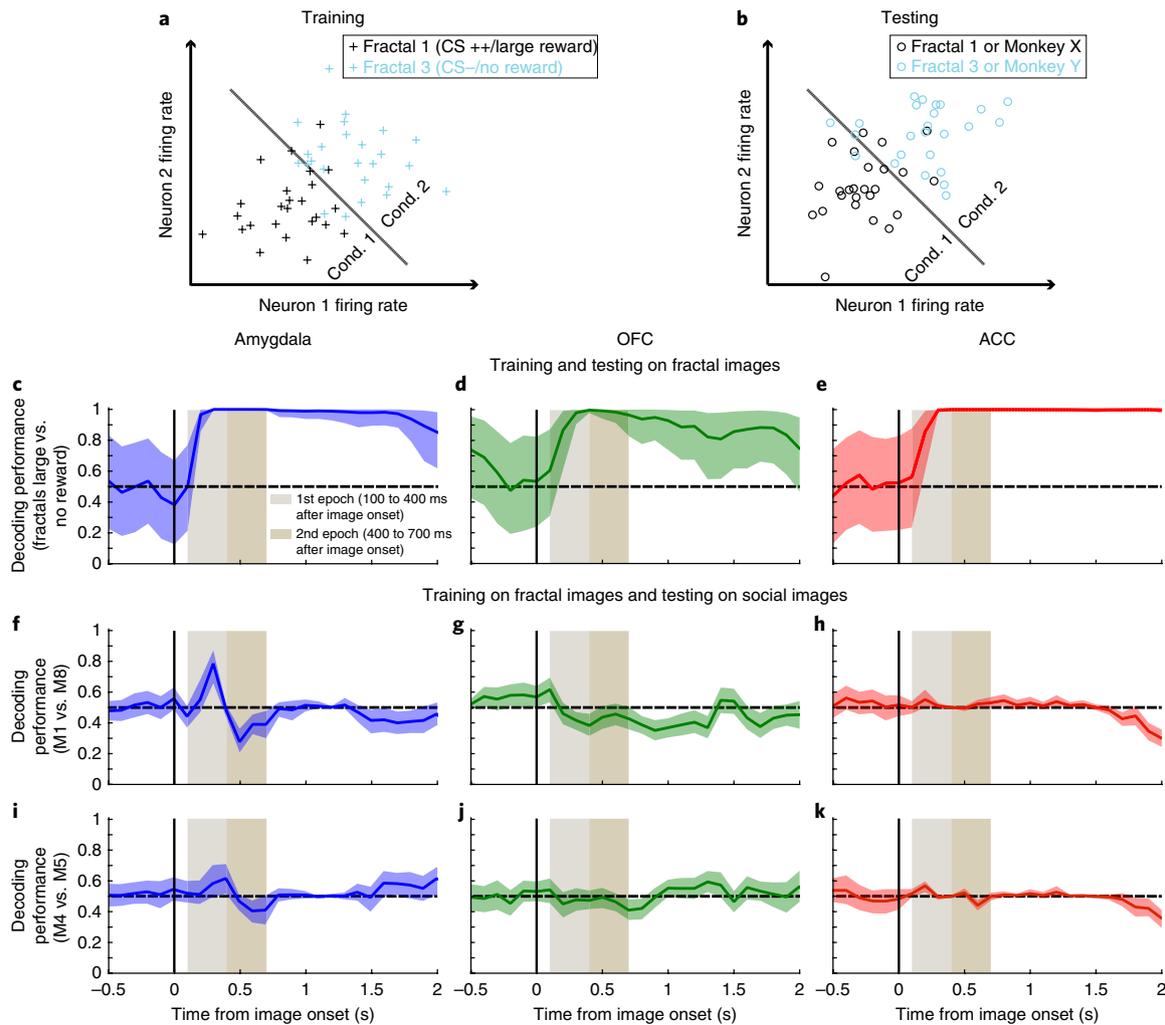
First, the preferred response selectivity of neurons tended to be preserved across time intervals for both reward value and hierarchical rank preference. Each of the regression coefficients  $\beta_{\text{frac}}$  and  $\beta_{\text{soc}}$ , which quantify the relationship between firing rate and either

the reward associated with fractals or the hierarchical status of face images, were correlated across the two time intervals on a cell-by-cell basis ( $r=0.49$ ,  $P<1 \times 10^{-6}$  for  $\beta_{\text{frac}}$ , Supplementary Fig. 7a, and  $r=0.79$ ,  $P<1 \times 10^{-15}$  for  $\beta_{\text{soc}}$ , Supplementary Fig. 7b; regression analyses). A change in sign in the relationship between decoding reward magnitude and hierarchical rank is therefore not likely to be accounted for by a change in the underlying response preferences of individual neurons.

Second, in the first time epoch (during fixation), for the dominant-preferring amygdala neurons ( $\beta_{\text{soc}}<0$ ), we observed a significant correlation between the two regression coefficients  $\beta_{\text{frac}}$  and  $\beta_{\text{soc}}$ , such that neurons that fired more for larger expected rewards fired more for dominant face images ( $P=0.032$ ,  $r=-0.21$ ). This was not observed in OFC and ACC for the subpopulations of neurons exhibiting a weak correlation with rank (Fig. 2g) ( $P=0.194$ ,  $r=0.17$  for dominant-preferring neurons in OFC and  $P=0.337$ ,  $r=0.11$  for submissive-preferring neurons in ACC, see Supplementary Fig. 8a–d). Consistent with this, dominant-preferring amygdala neurons responded significantly more to the large-reward fractal than to the fractal predicting no reward ( $P=0.045$ , paired *t*-test, Supplementary Fig. 9a), unlike submissive-preferring amygdala neurons ( $P=0.202$ , paired *t*-test in all cases, Supplementary Fig. 9b) or the subpopulations of OFC and ACC neurons that exhibited a weak correlation with hierarchical rank (Fig. 2g) ( $P=0.226$  for OFC and  $P=0.452$  for ACC, paired *t*-test, Supplementary Fig. 9c,d).

Finally, in the second time epoch (after the fixation requirement), a significant relationship was observed between  $\beta_{\text{frac}}$  and  $\beta_{\text{soc}}$  when considering all amygdala neurons ( $r=0.15$ ,  $P=0.039$ , Supplementary Fig. 10a). This relationship seemed to be driven predominantly by neurons responding most strongly to submissive monkeys ( $r=0.20$ ,  $P=0.066$ , Supplementary Fig. 10a) as opposed to neurons responding most strongly to dominant monkeys ( $r=0.04$ ,  $P=0.694$ , Supplementary Fig. 10a), which were the subensemble of neurons correlated with reward selectivity during the first time epoch (see above). Furthermore, neurons that responded more strongly to M8 than M1 (submissive-preferring neurons) during the second time epoch responded significantly more strongly to the large-reward fractal than to the no-reward fractal ( $P=0.030$ , paired *t*-test, Supplementary Fig. 10c), a relationship not observed for the dominant-preferring neurons ( $P=0.434$ , paired *t*-test, Supplementary Fig. 10b). Overall, therefore, the change in sign in decoding performance in the amygdala for hierarchical rank during the second compared to the first time epoch appears to be accounted for by a change in the relative contribution of dominant- and submissive-preferring subensembles of neurons to decoding performance in the two time epochs. These results can also be observed at a qualitative level by examining average neural responses to dominant and submissive face images separately for populations of neurons that prefer large or no reward (Supplementary Fig. 11).

We considered the possibility that the differential engagement of subpopulations of neurons during the two time epochs was related to differences in oculomotor behavior that could emerge as a result of the initiation of a free viewing interval. However, we did not observe any differences in oculomotor metrics that were related to social hierarchy (Supplementary Fig. 12) or reward amount of the fractal images. Even during the time interval extending from 700–1,000 ms after image onset, significant differences in oculomotor behavior that were related to hierarchical rank were not observed (Supplementary Fig. 13). The time window used to compute the social index, where oculomotor differences emerged, was at the very end of the free viewing epoch, more than 1,000 ms after image onset. Despite the fact that the observed patterns of neural activity could not be accounted for by oculomotor behavior, covert changes in cognitive processes, such as a desire to shift gaze or attention depending upon the hierarchical rank of a viewed face, may occur upon fixation point offset that could cause differential engagement of amygdala neuronal subensembles.



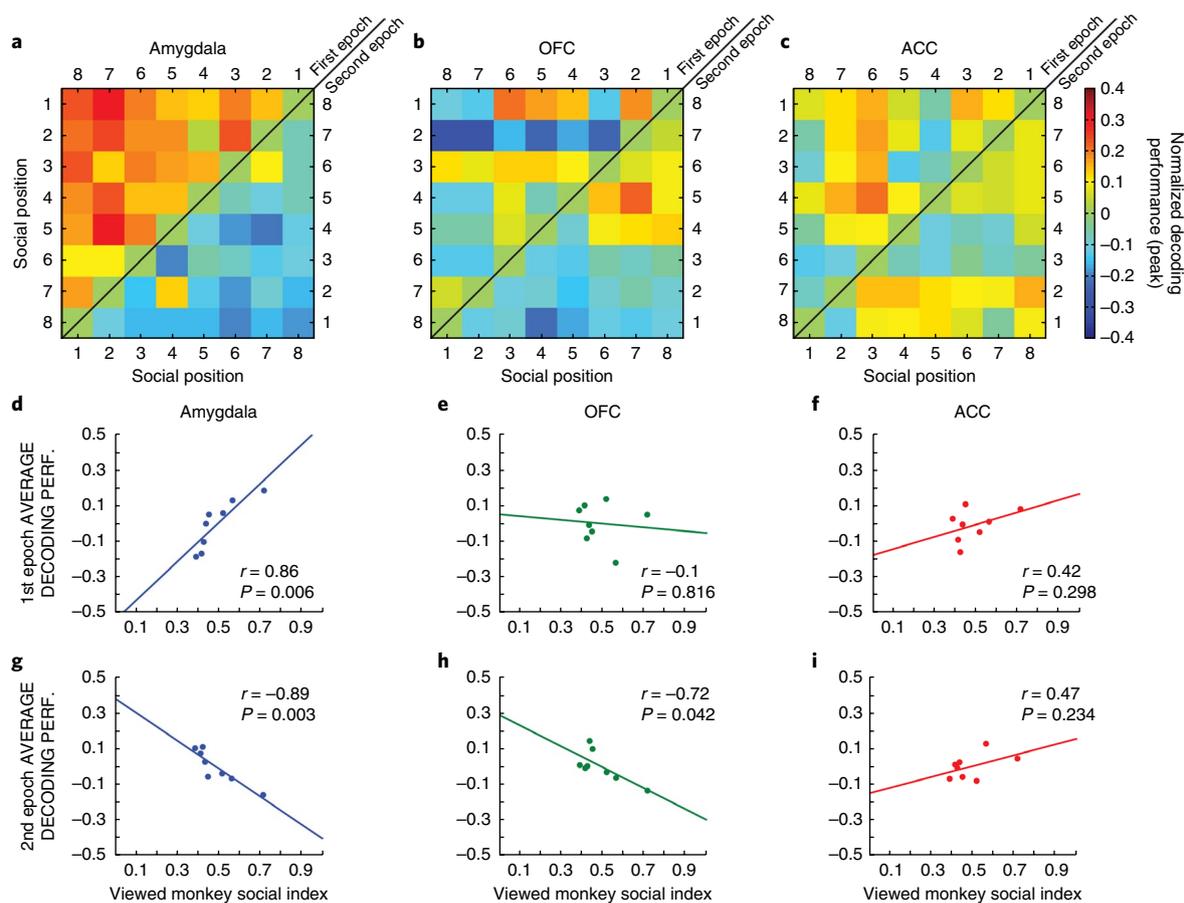
**Fig. 3 | The same neuronal ensemble in the amygdala can decode reward value and differences in hierarchical status.** **a, b.** Procedure for determining whether same neural ensemble represents the reward associated with fractals and the hierarchical status of viewed face images. A linear decoder is trained to discriminate between fractals associated with large and no reward on a subset of trials (**a**). Decoding performance (**b**) was tested on held-out fractal trials or on trials with any two monkey face images (for example, monkey X and monkey Y). **c–e.** Decoding performance plotted as a function of time for each brain area when testing the decoder on held-out large- and no-reward trials. **f–h.** Decoding performance for each brain area when testing the decoder on high- vs. low-ranked monkey image trials (M1 vs. M8; i.e., the most distant monkeys in terms of social status). **i–k.** Decoding performance when testing the decoder on two adjacent middle-ranked monkeys (M4 vs. M5). Shading, 95% confidence intervals (bootstrap). Dotted lines, chance decoding level. For all these analyses the decoder has been trained and tested on the entire neuronal population (i.e., without downsampling,  $n = 1,000$  iterations; see Methods).

## Discussion

This study took advantage of the fact that ten monkeys had lived together in the same social group for many years, enabling us to investigate the neurophysiological encoding of social hierarchy. Human observers and a social status index computed from measures of monkeys' oculomotor behavior as they viewed face images of their social cohort indexed the social hierarchy within this group in a consistent manner. Electrophysiological data revealed that the primate amygdala provided a neural representation of the hierarchical rank of individuals. As in prior studies, neurons in the amygdala also represented the reward magnitudes associated with different nonsocial (fractal) images. The representation of hierarchical rank was apparent in the same neuronal ensemble that represent rewards associated with nonsocial stimuli. Training a linear decoder to discriminate between trials in which different nonsocial stimuli predict large or no reward permitted decoding of the hierarchical relationship between individuals in the monkeys' social group. These data

provide strong evidence that the same neuronal ensembles that encode expected reward magnitudes may mediate aspects of social behavior by representing a complex social variable, hierarchical rank<sup>42</sup>. The data therefore argue against strong versions of the social brain hypothesis by demonstrating that, at least for the encoding of hierarchical rank, the same neuronal substrate in the amygdala appears to process social and nonsocial information.

Our results indicate that neuronal ensembles that represent reward magnitudes associated with nonsocial stimuli may also be used to represent the hierarchical ranks of individuals. Stimuli that depict individuals with known hierarchical rank convey motivational meaning to a subject, just as reward-predictive stimuli convey motivational meaning. Both social and nonsocial stimuli can influence subsequent decisions, actions, and emotional responses more broadly. However, two lines of evidence indicate that representations of the reward magnitudes associated with nonsocial stimuli do not inherently represent hierarchical rank in all brain



**Fig. 4 | The relationship between the encoding of reward value and hierarchical status. a–c,** Peak decoding performance in the amygdala, OFC and ACC tested after training the decoder to classify large- vs. no-reward trials and testing the decoder for all pairs of monkey face images (for example, M1 vs. M8, M1 vs. M7, etc.; see Methods). Color scale indicates decoding performance minus 0.5, so a value of 0 corresponds to chance decoding. Warm and cool colors above and below the identity diagonal reflect decoding above and below chance in the two time intervals, respectively. Decoding performance during fixation (100–400 ms after image onset), upper left of each plot; decoding performance during free viewing (400–700 ms after image onset), lower right of each plot. **d–i,** Average decoding performance for each monkey face image (across all comparisons) plotted as a function of that monkey’s social status index for the first (**d–f**) and second (**g–i**) time epochs in the three brain areas. Note the inversion of sign in the amygdala between the first (**d**) and second (**g**) time epochs. For each panel  $n = 8$ , Pearson’s correlation coefficient (two-sided): amygdala,  $r = 0.86$ ,  $P = 0.006$  (first epoch),  $r = -0.89$ ,  $P = 0.003$  (second epoch); OFC,  $r = -0.1$ ,  $P = 0.816$  (first epoch),  $r = -0.72$ ,  $P = 0.042$  (second epoch); ACC,  $r = 0.42$ ,  $P = 0.298$  (first epoch),  $r = 0.47$ ,  $P = 0.234$  (second epoch).

structures. First, the OFC and ACC both represent expected reward magnitudes strongly, but neither structure provided a compelling representation of hierarchical rank; training a linear decoder to discriminate between large and no reward fractal image trials did not allow us to decode hierarchical rank in those structures (Figs. 2–4). Second, we observed that the relationship between representations of reward magnitude and hierarchical rank in the amygdala changed as a function of time during a trial (Figs. 3 and 4). This observation argues strongly that the results cannot be easily accounted for by neural signals that merely encode variables such as arousal, attention, valuation or other related processes. It is very unlikely, for example, that an arousing and attention-grabbing alpha monkey becomes not arousing or attention-grabbing within a few hundred milliseconds, which would be required for signals encoding arousal or attention to account for our data. Instead, cell-by-cell regression analyses suggested that separate populations of amygdala neurons may be engaged at different times during a trial to represent hierarchical rank, and these subpopulations have differential relationships with the encoding of expected reward magnitudes (Supplementary Figs. 7–10).

Our data are also unlikely to be accounted for by a low-level visual feature of face images that somehow conveys hierarchical rank. Training a decoder on neural responses to fractal images that predict different reward amounts should not be expected to allow decoding of hierarchical rank based on a low-level visual feature. In short, it is difficult to imagine that a particular sensory feature of face images that conveys hierarchical rank information has a consistent relationship with reward magnitudes associated with fractal images. However, the shared encoding in the amygdala of hierarchical rank and the rewards associated with nonsocial stimuli could reflect the provision of a common currency for stimuli of many different types that is present transiently in the amygdala (as the representation of social rank was not sustained across the entire trial, unlike reward value; Fig. 3c,f)<sup>43</sup>, but further experiments are needed. To wit, the neuronal ensembles we have characterized may also represent other social variables, such as body gestures or complex facial expressions that convey emotional and motivational meaning, and it is unknown whether the representations of these social variables are related to encoding of the reward values of nonsocial stimuli. It remains possible that some aspects of primate social behavior

share the same neuronal encoding as nonsocial processes, with other aspects of social behavior being mediated by different neural representations<sup>42,44</sup>.

Monkeys almost certainly learn the hierarchical rank of other group members through experience, so the acquisition of a neural representation of hierarchical rank may occur in a similar manner as when a subject learns the relationship between nonsocial stimuli and rewards through reinforcement learning<sup>45</sup>. Of course, learning about hierarchical rank may also be facilitated through observational learning<sup>46</sup>. The sharing of a neuronal substrate for learning about nonsocial rewards and about social hierarchy information has mechanistic implications for understanding the pathophysiology and potential treatments for some patients with psychiatric disorders. For example, patients with autism spectrum disorder or schizophrenia often exhibit deficits in social interactions, sometimes struggling to appropriately register social situations or cues or to respond to such situations adaptively<sup>47,48</sup>. Some of these patients may also possess deficits in learning about the significance (for example, reward value) of nonsocial stimuli<sup>49</sup>. Indeed, applied behavior analysis treatments—based on reinforcing correct behavior with different types of reward—have been shown to benefit many patients with autism spectrum disorder<sup>50</sup>, suggesting that focused behavioral manipulations may compensate for deficits in learning mechanisms in these patients. However, the neural underpinnings of deficits in social processing that also requires learning has remained unclear, in large part because we have understood relatively little about how the brain represents and utilizes social information. The discovery of neural representations of hierarchical rank in the amygdala observed in the same neural ensembles that represent the rewards associated with nonsocial stimuli may thereby provide a significant step toward describing the neural networks critical for understanding psychiatric pathology and treatment.

## Methods

Methods, including statements of data availability and any associated accession codes and references, are available at <https://doi.org/10.1038/s41593-018-0082-8>.

Received: 3 July 2017; Accepted: 21 December 2017;  
Published online: 19 February 2018

## References

- Thierry, B. Covariation of conflict management patterns across macaque species. in *Natural Conflict Resolution* (ed. Aureli, F. & de Waal, F.B.M.) 106–128 (University of California Press, Berkeley, CA, USA, 2000).
- Bercovitch, F. B. Social-stratification, social strategies, and reproductive success in primates. *Ethol. Sociobiol.* **12**, 315–333 (1991).
- Cowlshaw, G. & Dunbar, R. Dominance rank and mating success in male primates. *Anim. Behav.* **41**, 1045–1056 (1991).
- Sapolsky, R. M. The influence of social hierarchy on primate health. *Science* **308**, 648–652 (2005).
- Byrne, R. & Whiten, A. *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans*. (Oxford Science Publications, Oxford, 1989).
- Schell, A., Rieck, K., Schell, K., Hammerschmidt, K. & Fischer, J. Adult but not juvenile Barbary macaques spontaneously recognize group members from pictures. *Anim. Cogn.* **14**, 503–509 (2011).
- Pfefferle, D., Kazem, A. J., Brockhausen, R. R., Ruiz-Lambides, A. V. & Widdig, A. Monkeys spontaneously discriminate their unfamiliar paternal kin under natural conditions using facial cues. *Curr. Biol.* **24**, 1806–1810 (2014).
- Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H. & Livingstone, M. S. A cortical region consisting entirely of face-selective cells. *Science* **311**, 670–674 (2006).
- Gothard, K. M., Battaglia, F. P., Erickson, C. A., Spitzer, K. M. & Amaral, D. G. Neural responses to facial expression and face identity in the monkey amygdala. *J. Neurophysiol.* **97**, 1671–1683 (2007).
- Hoffman, K. L., Gothard, K. M., Schmid, M. C. & Logothetis, N. K. Facial-expression and gaze-selective responses in the monkey amygdala. *Curr. Biol.* **17**, 766–772 (2007).
- Rutishauser, U. et al. Single-unit responses selective for whole faces in the human amygdala. *Curr. Biol.* **21**, 1654–1660 (2011).
- Azzi, J. C., Sirigu, A. & Duhamel, J. R. Modulation of value representation by social context in the primate orbitofrontal cortex. *Proc. Natl. Acad. Sci. USA* **109**, 2126–2131 (2012).
- Tsao, D. Y., Schweers, N., Moeller, S. & Freiwald, W. A. Patches of face-selective cortex in the macaque frontal lobe. *Nat. Neurosci.* **11**, 877–879 (2008).
- Watson, K. K. & Platt, M. L. Social signals in primate orbitofrontal cortex. *Curr. Biol.* **22**, 2268–2273 (2012).
- Chang, L. & Tsao, D. Y. The code for facial identity in the primate brain. *Cell* **169**, 1013–1028.e1014 (2017).
- Amaral, D., Price, J., Pitkanen, A. & Carmichael, S. Anatomical organization of the primate amygdaloid complex. in *The Amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction* (ed. Aggleton, J.) 1–66 (Wiley-Liss, New York, 1992).
- Ghashghaie, H. T., Hilgetag, C. C. & Barbas, H. Sequence of information processing for emotions based on the anatomic dialogue between prefrontal cortex and amygdala. *Neuroimage* **34**, 905–923 (2007).
- Stefanacci, L. & Amaral, D. G. Some observations on cortical inputs to the macaque monkey amygdala: an anterograde tracing study. *J. Comp. Neurol.* **451**, 301–323 (2002).
- Paton, J. J., Belova, M. A., Morrison, S. E. & Salzman, C. D. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* **439**, 865–870 (2006).
- Morrison, S. E., Saez, A., Lau, B. & Salzman, C. D. Different time courses for learning-related changes in amygdala and orbitofrontal cortex. *Neuron* **71**, 1127–1140 (2011).
- Cai, X. & Padoa-Schioppa, C. Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex. *J. Neurosci.* **32**, 3791–3808 (2012).
- Dunbar, R. I. M. The social brain hypothesis. *Evol. Anthropol.* **6**, 178–190 (1998).
- Sallet, J. et al. Social network size affects neural circuits in macaques. *Science* **334**, 697–700 (2011).
- Noonan, M. P. et al. A neural circuit covarying with social hierarchy in macaques. *PLoS Biol.* **12**, e1001940 (2014).
- Bickart, K. C., Wright, C. I., Dautoff, R. J., Dickerson, B. C. & Barrett, L. F. Amygdala volume and social network size in humans. *Nat. Neurosci.* **14**, 163–164 (2011).
- Powell, J. L., Lewis, P. A., Dunbar, R. I. M., García-Fiñana, M. & Roberts, N. Orbital prefrontal cortex volume correlates with social cognitive competence. *Neuropsychologia* **48**, 3554–3562 (2010).
- Kanai, R., Bahrami, B., Roylance, R. & Rees, G. Online social network size is reflected in human brain structure. *Proc. Biol. Sci.* **279**, 1327–1334 (2012).
- Watanabe, N. & Yamamoto, M. Neural mechanisms of social dominance. *Front. Neurosci.* **9**, 154 (2015).
- Kumaran, D., Melo, H. L. & Duzel, E. The emergence and representation of knowledge about social and nonsocial hierarchies. *Neuron* **76**, 653–666 (2012).
- Zink, C. F. et al. Know your place: neural processing of social hierarchy in humans. *Neuron* **58**, 273–283 (2008).
- Ligneul, R., Obeso, I., Ruff, C. C. & Dreher, J. C. Dynamical representation of dominance relationships in the human rostromedial prefrontal cortex. *Curr. Biol.* **26**, 3107–3115 (2016).
- Kumaran, D., Banino, A., Blundell, C., Hassabis, D. & Dayan, P. Computations underlying social hierarchy learning: distinct neural mechanisms for updating and representing self-relevant information. *Neuron* **92**, 1135–1147 (2016).
- Chang, S. W. et al. Neural mechanisms of social decision-making in the primate amygdala. *Proc. Natl. Acad. Sci. USA* **112**, 16012–16017 (2015).
- Chang, S. W., Gariépy, J. F. & Platt, M. L. Neuronal reference frames for social decisions in primate frontal cortex. *Nat. Neurosci.* **16**, 243–250 (2013).
- Klein, J. T. & Platt, M. L. Social information signaling by neurons in primate striatum. *Curr. Biol.* **23**, 691–696 (2013).
- Yoshida, K., Saito, N., Iriki, A. & Isoda, M. Social error monitoring in macaque frontal cortex. *Nat. Neurosci.* **15**, 1307–1312 (2012).
- Klein, J. T., Deaner, R. O. & Platt, M. L. Neural correlates of social target value in macaque parietal cortex. *Curr. Biol.* **18**, 419–424 (2008).
- Zumpe, D. & Michael, R. P. Dominance index – a simple measure of relative dominance status in primates. *Am. J. Primatol.* **10**, 291–300 (1986).
- Shepherd, S. V. & Platt, M. L. Spontaneous social orienting and gaze following in ringtailed lemurs (*Lemur catta*). *Anim. Cogn.* **11**, 13–20 (2008).
- Deaner, R. O., Khera, A. V. & Platt, M. L. Monkeys pay per view: adaptive valuation of social images by rhesus macaques. *Curr. Biol.* **15**, 543–548 (2005).
- Saez, A., Rigotti, M., Ostojic, S., Fusi, S. & Salzman, C. D. Abstract context representations in primate amygdala and prefrontal cortex. *Neuron* **87**, 869–881 (2015).

42. Dichter, G. S., Felder, J. N., Bodfish, J. W., Sikich, L. & Belger, A. Mapping social target detection with functional magnetic resonance imaging. *Soc. Cogn. Affect. Neurosci.* **4**, 59–69 (2009).
43. Rangel, A. & Camerer, C. & Montague, P.R. A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* **9**, 545–556 (2008).
44. Behrens, T. E., Hunt, L. T. & Rushworth, M. F. The computation of social behavior. *Science* **324**, 1160–1164 (2009).
45. Sutton, R. & Barto, A. *Reinforcement Learning*. (MIT Press, Cambridge, MA, USA, 1998).
46. Gariépy, J. F. et al. Social learning in humans and other animals. *Front. Neurosci.* **8**, 58 (2014).
47. Gaigg, S. B. The interplay between emotion and cognition in autism spectrum disorder: implications for developmental theory. *Front. Integr. Neurosci.* **6**, 113 (2012).
48. Dodell-Feder, D., Tully, L. M. & Hooker, C. I. Social impairment in schizophrenia: new approaches for treating a persistent problem. *Curr. Opin. Psychiatry* **28**, 236–242 (2015).
49. Solomon, M., Smith, A. C., Frank, M. J., Ly, S. & Carter, C. S. Probabilistic reinforcement learning in adults with autism spectrum disorders. *Autism Res.* **4**, 109–120 (2011).
50. Foxx, R. M. Applied behavior analysis treatment of autism: the state of the art. *Child Adolesc. Psychiatr. Clin. N. Am.* **17**, 821–834, ix (2008).

## Acknowledgements

We thank R. Adolphs, E. Issa and members of the Salzman lab for critical comments on this work and the manuscript. This work was supported by the Fyssen Foundation, Philippe Foundation (J.M.), the Simons Foundation and National Institutes of Health grant R01 MH082017 (C.D.S.).

## Author contributions

J.M. and C.D.S. designed the study; J.M. performed the experiments; J.M., M.R. and C.D.S. designed analyses; J.M. and M.R. performed analyses; J.M., M.R. and C.D.S. wrote the manuscript.

## Competing interests

The authors declare no competing financial interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41593-018-0082-8>.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Correspondence and requests for materials** should be addressed to J.M. or C.D.S.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Methods

**Animals and surgical procedures.** All experiment procedures were in accordance with the guidelines of the National Institutes of Health guide for the care and use of laboratory animals and the Animal Care and Use Committees at New York State Psychiatric Institute and Columbia University. Two adult male rhesus monkeys (*Macaca mulatta*) weighing 9.6 kg and 10.2 kg (monkey R and monkey F) underwent surgical sessions under isoflurane general anesthesia and using aseptic techniques to implant a plastic head post and an oval recording chamber positioned over a craniotomy that provided access to the right amygdala and prefrontal cortex. Analgesics and antibiotics were administered during recovery from surgery. TheBrainsight 2 system (Rogue Research) was used in combination with presurgery magnetic resonance imaging (MRI) to localize the brain structures and determine coordinates of chamber placement. A second MRI was performed after surgery to verify correct placement of the recording chamber and to determine electrode coordinate locations using the Brainsight 2 software for recording from the amygdala, orbitofrontal cortex and anterior cingulate cortex. Another MRI was performed on the monkeys at the end of the experimental sessions to validate localization of our recording sites.

These experiments involved electrophysiological recordings in two monkeys each of whom lived in the same social group of ten male monkeys for more than 5 years. In addition, no monkeys entered or exited this stable social group of ten monkeys. Monkeys were housed either individually or, when possible, in pairs. In the present study, monkey F was paired with M5 and Monkey R was occasionally paired with M3. Human observers have noted a consistent hierarchy within the group of ten monkeys, with the same monkeys being viewed as most dominant or submissive. Their room was equipped with mirrors on the wall as well as Plexiglas panels dividing the different cages to allow frequent visual contact between all members of the colony. Play cages were adjacent to all apartment cages in the room, and every monkey had regular access to a play cage.

**Experimental procedures.** The experimental sessions were conducted in a darkened, sound-attenuated and RF-shielded room. Throughout the session, the monkeys were seated in a primate chair with their head restrained. They viewed a 17-inch, 1,024 × 800 pixel CRT screen, refresh rate 100 Hz, placed 57 cm from their eyes. Eye movements were recorded at 1,000 Hz using an optometric system: Eyelink I system (SR Research). To measure licking behavior, a reward delivery spout was placed few millimeters in front of monkeys' mouths. Receipt of liquid reward required extrusion of the tongue to the delivery spout or the liquid reward would drop to the floor. Licking behavior was measured by using an infrared laser beam set between the monkey lips and the reward delivery spout. The interruption of this laser beam thereby demarcated anticipatory and consummatory approach behaviors to obtain reward.

Visual stimuli and the behavioral task were under the control of Expo software running on a dedicated Apple computer. Neuronal data, eye movements and licking behavior were stored using the Plexon Omniplex system (Plexon, Inc.). Synchronization between the Expo and Plexon system was achieved by timestamps sent from Expo during each trial to the Omniplex system. Data collection and analysis were not performed blind to the conditions of the experiments. However, all well-isolated neurons were kept, and we did not choose the type of neurons we recorded during data acquisition on the basis of their selectivity to the task events.

**Behavioral task.** Animals performed a trace-conditioning task where a conditioned stimulus (CS) predicts delivery of an unconditioned stimulus (US); the task required fixation, and actual reward receipt depended upon licking behavior. The two monkeys used in this study had previously been trained on a task using CS and US presentations, but the prior experience did not involve the presentation of social stimuli of any kind. The two monkeys experienced only two sessions on the specific task described below before recording sessions commenced.

In the present experiments, monkeys performed a trace conditioning task in which either the presentation of a fractal image or a monkey face image predicted the delivery of reward. Fractal image CS and monkey face CS trials were presented in separate blocks of trials. Monkeys initiated a trial by centering gaze on a fixation point (FP) at the center of the screen. The radius of the fixation window was 5 degrees. The average fixation position was not different between the fractal and social images, indicating that visual scanning of the images did not differ for the fractal vs. face images during fixation (Kruskal–Wallis test,  $P > 0.05$ ). Following 0.4 s of fixation, we presented an image (CS) under the FP; images were 15 degrees square. Monkeys were required to maintain fixation for an additional 0.4 s before FP disappearance. The CS remained on the screen for an additional 1 s, providing a free viewing epoch. After image extinction, a 0.75-s delay ensued before delivery of the unconditioned stimulus (US). The inter-trial interval (ITI) was 1 s for all trial types. Failure to engage or maintain fixation for the entire 0.8 s epoch led to the presentation of a black screen for a 1,000 ms timeout, followed by the ITI and then trial repetition. All trials were randomly interleaved within a block.

The CSs were either a fractal or social image. Three different fractal images were presented to the viewer monkeys, one associated with two drops of liquid reward, one with one drop of reward, and one with no reward. Successful fixation therefore resulted in reward delivery for two of the fractals. Social images were

pictures of the faces of eight 'viewed' rhesus monkeys (M1 to M8) who were all members of the same group of ten primates (the eight viewed monkeys and the two viewer monkeys who were the subject of electrophysiological recordings in this study) housed in the same room for more than 5 years. The position of the FP was between the superior lip and the nose of images. Face images were selected to be devoid of emotional expressions, such as lip smacking, aggressive open mouth, anxious bared-teeth grins or ear position. The gaze of the eyes of the face images was not directly toward the viewer monkeys. Successful fixation during trials in which social images were used as CSs resulted in one drop of reward. Note that reward delivery was not contingent on approach behavior to the reward tube, but that failure to approach the reward tube (for example, with licking) resulted in the reward falling to the floor. Each experiment was composed of a fractal block, followed by a social block, followed by a fractal block, with 5–20 trials per condition per block.

**Electrophysiological methods.** The activity of single neurons was recorded extracellularly using U- and V-probes (Plexon, Inc.) of 24 channels in amygdala and 16 channels each in OFC and ACC. All the probes had 100 μm spacing between each contact channel. The three probes were lowered simultaneously through stainless steel guide tubes by means of a motorized multi-electrode drive (NAN Instruments). Analog extracellular signals were amplified, bandpass filtered (250–8,000 Hz) and digitized (40,000 Hz) using the Plexon Omniplex system (Plexon, Inc.). Single units were isolated offline using the Plexon offline sorter 3.3 based on the three principal components of waveforms. Only well-isolated single cells with an adequate refractory period (1.5 ms) were kept. Cross-correlogram and selectivity analysis were performed through the different units to reject duplicate cells across contingent channels. 196 neurons were recorded in the basolateral amygdala (monkey F: 151 cells; monkey R: 45 cells), 134 in OFC (area 13 m and 13l; monkey F: 78 cells; monkey R: 56 cells) and 187 in ACC (monkey F: 117 cells; monkey R: 70 cells; 61%, 21% and 18% of ACC neurons were recorded in, respectively, the ventral bank of ACC, area 24c; the dorsal bank of ACC; and the ACC gyrus, area 32). Different types of neural modulation by social experiences have been previously reported in these different parts of ACC<sup>34</sup>, but we did not observe difference between these substructures. Recording site reconstruction is shown in Supplementary Fig. 4.

**Behavioral analysis. Scoring of social hierarchy by human observer.** A human observer scored the social status of each monkey of the colony based on the method described by Zumpfe and Michael<sup>35</sup>, which has successfully been applied in several studies<sup>33</sup>. Aggressive and submissive agonistic behaviors were recorded during a dozen series of at least 5- to 15-min observations, each of which were preceded by 5 min of habituation. These observations occurred before all the recording sessions with the two monkeys. Observation sessions were conducted either when two monkeys were in nearby cages such that they could interact socially or when two monkeys were placed in chairs positioned to face each other. We scored behaviors that included escaping, lip smacking, chasing and aggressive behaviors such as an open mouth or the shaking or slapping of a cage wall adjacent to the other monkey. A submissive behavior was recorded as a win for the other monkey. When two monkeys were seated in primate chairs facing each other, we also conducted a food competition test. A piece of food was placed two to four consecutive times on a small plate between the two monkeys. Each time one of the monkeys grabbed the food, he was scored as the winner<sup>31,32</sup>. Other agonistic behaviors (described above) were also scored during this test. If an aggressive behavior by one monkey induced a submissive behavior by another monkey, only the aggressive behavior was scored. Because grooming and mounting have no necessary relationships with dominance<sup>33</sup>, these behaviors were not classified as agonistic (Table 1 and Supplementary Table 1).

**Analysis of behavior during electrophysiology experiments.** All data analysis programs were written in Matlab (MathWorks Inc., Natick, MA) and used its statistical toolbox. Since preliminary analyses did not reveal substantial differences between the patterns of neuronal activity, the oculomotor behavior and the valuation of the social hierarchy between the two monkeys, we conducted all subsequent analyses on data pooled from all recording sessions in the two viewer monkeys.

**The assessment of social hierarchy when monkeys view face images.** A social status index was created that characterized how the two experimental monkeys assessed the social status of the eight monkeys whose faces they viewed. This index was computed using the oculomotor behavior of the two recorded ('viewer') monkeys. The index was constructed from four measures, which were weighted equally: (1) trial completion rate for each viewed face image (successful completion of fixation); (2) the proportion of time a viewer monkey spends looking at the image of a monkey's face during the last one-third of the free viewing interval; (3) the proportion of errors that are due to breaking fixation during viewing of a monkey face, as opposed to breaking fixation before image presentation or not engaging in fixation at all; and (4) the proportion of time during the last one-third of the free viewing interval that the monkey views the eyes of a face image divided by the time

the subject spends looking at other parts of the same image. The social status index was computed as

$$\text{Index} = \frac{A + B + C + (1-D)}{4}$$

where  $A$ ,  $B$ ,  $C$  and  $D$  are measures 1, 2, 3 and 4 described above.

In general, these measures of social status exploit the fact that monkeys tend to prefer viewing pictures of dominant monkeys<sup>40,54,55</sup> (measures 1 and 2), but that viewing of a dominant monkey face can be more arousing, leading to more frequent reflexive broken fixations (measure 3). These broken fixations could reflect arousal or other motivated behaviors (for example, gaze shifting) related to the social status of a viewed face. In support of this notion, measure 3 is strongly anticorrelated with propensity to re-engage fixation on the next trial after this type of error ( $r = -0.90$ ,  $P = 0.002$ , Pearson's linear correlation test (two-sided)), suggesting that these broken fixations may not be volitional and instead may simply be automatic responses due to arousal. Finally, we employ eye gaze directed toward the eyes as a measure because such gaze is a threat in the primate world, and therefore there is a tendency not to look at the eyes of a dominant monkey<sup>56</sup>.

**Analysis of oculomotor metrics.** We also analyzed oculomotor metrics that included the number of saccades or microsaccades and the amplitudes and orientations of these saccades during two time epochs: 100–400 ms after CS onset (when fixation was required) and 400–700 ms after CS onset (the first 300 ms of the free viewing epoch). Oculomotor parameters were examined for each of the eight viewed monkey face images to determine whether they varied as a function of the viewed monkey identity.

**Statistical methods.** No statistical methods were used to predetermine sample sizes, but our sample sizes are similar to those reported in previous publications (for example, ref. <sup>41</sup>). Data distribution was assumed to be normal, but this was not formally tested for all the analyses. The decoder does not assume any normality in the data distributions. Behavioral analyses were carried out using the Kruskal–Wallis test for main and interaction effects between groups and post hoc comparisons (Dunn's multiple comparison test) or Wilcoxon rank sum test, when appropriate. We used two-sided paired  $t$ -tests and Fisher's  $r$  to  $z$  transformation test to investigate statistical neural difference between conditions, and Pearson's linear correlation test (two-sided) to compute linear correlation coefficients (see below).

**Analysis of neurophysiological data.** For all analyses described below, data were combined across monkeys because all key features of the dataset were statistically indistinguishable between monkeys. In general, we required at least 5 trials per condition in order to consider a neuron for any analysis described below.

**Single-neuron analysis.** The analysis of neural data largely focused on a time epoch extending from 100–400 ms after CS onset or a time epoch spanning 400–700 ms after CS onset. The first time epoch was selected because of the known response latency of neurons in amygdala, OFC and ACC to visual stimuli of approximately 100 ms and because the fixation requirement terminated 400 ms after CS onset. The second time epoch was of the same length and corresponded to the first 300 ms of the free viewing epoch.

**Normalization of neural activity.** For some analyses, we normalized neural activity before averaging across neurons. In this case, for each session, the firing rate (FR) for different conditions was normalized with respect to baseline activity measure during the second half of the inter-trial interval (ITI), a time interval chosen because fixation point presentation can induce a visual response and a signal related to state value in the amygdala<sup>37</sup>:

$$FR_n = (FR_i - \mu) / \text{s.d.}$$

where  $FR_n$  represent the normalized FR for one condition in one epoch,  $FR_i$  is the absolute FR for that condition in the same epoch, and  $\mu$  and  $\text{s.d.}$  are, respectively, the mean and  $\text{s.d.}$  of the spike count across trials during the baseline interval.

**Linear regression analysis quantifying reward value and social hierarchy tuning.** Individual cells were characterized according to their tuning to reward value and social hierarchy in the following way. To quantify the tuning to the reward predicted by the fractals, the firing rate FR of each neuron during a time interval of interest was fitted to the value of the fractals  $V$  (the amount of reward predicted by a given fractal) in accordance to the linear regression:

$$FR = \beta_{\text{frac}}^0 + \beta_{\text{frac}} V$$

Notice that we obtain one regression coefficient  $\beta_{\text{frac}}$  per cell (and per considered time interval).

Analogously, we quantified neural tuning to social hierarchy by fitting the firing rate FR of each neuron with the linear regression

$$FR = \beta_{\text{soc}}^0 + \beta_{\text{soc}} R$$

where  $R$  indicates the rank of a given monkey face image in the social hierarchy as derived from the social status index (see above).

Correlations between regression coefficients (either of the same  $\beta$  computed in different time epochs or between  $\beta_{\text{frac}}$  and  $\beta_{\text{soc}}$  in the same time epoch) quantify the consistency of the neural tuning (across time epochs or between nonsocial and social value, respectively). This correlation was calculated by computing the Pearson's correlation coefficient  $r$  between  $\beta_{\text{frac}}$  and  $\beta_{\text{soc}}$  for the whole population of recorded cells (or for subpopulations of cells, as indicated in the text).

**Population decoding.** We used a standard linear classifier for the population decoding (see, for example, ref. <sup>41</sup>). The decoding algorithm was based on a population decoder trained on pseudo-simultaneous population response vectors<sup>58</sup>. The components of these vectors corresponded to the spike counts of the recorded neurons in specific time bins computed in the following way. The activity of each neuron in every trial was aligned to CS onset, the event that in the analysis defines time zero. Within each trial we then computed the spike count over time bins. All analyses were then carried out independently in every time bin as explained below. For every neuron and every time bin, we  $z$ -scored the distribution of spike counts across all trials and all conditions. The mean and  $\text{s.d.}$  used were always computed on the training data (see below). Given a task condition  $c$  (for instance, the fractal associated with large reward) and a time bin  $t$  (for instance, between 100 and 400 ms after CS presentation), we generated pseudo-simultaneous population response vectors by sampling, for every neuron  $i$ , the  $z$ -scored spike count in a randomly selected trial in condition  $c$ , which we indicate by  $n_i^c(t)$ . This procedure resulted in the single-trial population response vector  $n^c(t) = (n_1^c(t), n_2^c(t), \dots, n_N^c(t))$ , where  $N$  is the number of recorded neurons in the area under consideration.

**Training and testing of the population decoder.** The decoder analysis comprises a training phase, for which (unless otherwise stated) we used 75% of the trials, and a testing (or evaluation) phase done over the 25% of trials that were held out from training. Training and testing were repeated 1,000 times over different random partitions of the trials into training and test trials. We then report the mean performance and the associated confidence interval obtained from the resulting empirical distribution of decoding accuracy.

We discarded neurons with less than five repetitions per condition in the training stage of the decoding analysis. For standard decoding (training and testing on same type of stimuli) the two conditions tested could be either two fractal images (allowing the training of three binary classifiers, discriminating between large vs. no reward, medium vs. no reward, or large vs. medium reward, and combined in a three-way classifier) or two social images (resulting in 56 binary classifiers combined in an eight-way classifier). For some analyses, we trained the decoder to discriminate fractals associated with large vs. no reward and tested the decoder's ability to discriminate between any two monkey face images.

We performed the decoding analyses using an equal number of neurons in the three brain areas to reject a potential confound in the results due to differences in the number of neurons in the population of each area (Fig. 2). This was accomplished by randomly subsampling the recorded populations such that decoding was performed on 110 cells in each structure when the training procedure was performed on social images or 131 cells when the training occurred on fractal images; the number of neurons was equal to the number of OFC neurons passing our criterion for inclusion (see Fig. 2f). The random subsampling meant that we randomly rejected a new set of excess amygdala and ACC neurons at each iteration of the decoder. For all decoding analyses, similar results were obtained if we did not equate the number of neurons from each brain area and instead included all neurons.

When we examined decoding performance as a function of time across entire trials, we employed time bins of 300 ms stepped in 100-ms increments across the trial, with training and testing performed on each 300-ms time bin. We also performed decoding analyses focused on two time windows, 100–400 ms after CS onset (during fixation) and 400–700 ms after CS onset (the first 300 ms of free viewing) (see below). For these analyses, we used six 50-ms bins in each of the two 300-ms-long epochs. We then located the peak of decoding within these two 300-ms windows (the 50-ms epoch with decoding performance most different from chance), and this peak value is plotted in Fig. 4. In Fig. 4a–c, although all paired combinations of monkey face images have been tested (56 pairs), we present only half of the comparisons as the other half of comparisons have the same results with an opposite sign; for example, the value for the comparison M1 vs. M8 =  $-(M8 \text{ vs. } M1)$ .

**Justification of the neural windows selection (100–400 ms and 400–700 ms after image onset).** First, we selected the onset of the first window as 100 ms after image onset due to the visual response latencies that have been observed in these brain areas<sup>41</sup>. This effect can be seen in Fig. 2f, where the decoding performance starts to

increase approximately 100 ms after image presentation. The first time window is merely the remaining time in the fixation interval, which ends 400 ms after image onset. The second time window was set to 400–700 ms for several reasons: (1) to mirror the length of the first window; (2) to be able to analyze neural data in an interval of the free viewing epoch where we did not observe differences in oculomotor behavior as a function of the viewed monkey face images (the differences in oculomotor behavior occur later, during the last third of the free viewing epoch), thus avoiding a contribution of explicit concurrent behavioral differences in our interpretation of the neural response properties; and (3) because we observed a flip in decoding performance around FP extinction during the decoding analysis in which we trained on fractal images differing in reward amount and tested the decoder on monkey face images that differed substantially in social rank (Fig. 3f,i and Fig. 4a). This last analysis found that decoding different from chance occurred in a distinctive manner in each of these time windows, but that in later time intervals during free viewing, decoding performance was largely not different from chance.

Finally, as is evident in Fig. 2f, left, the peak of decoding in the three structures occurred between 300 and 700 ms after image onset (350, 500 and 650 ms in AMY, OFC and ACC respectively). This makes us confident that most of the neural signals of interest occur in the 100–700 ms interval. Indeed, no clear modulation happens later in the trial (Fig. 3).

**Statistical significance and estimate of confidence intervals.** As described above, on every iteration of the training and testing procedure, we partitioned the trials into training and testing sets by random sampling without replacement. The average decoding performance, its statistical significance and confidence intervals were estimated by repeating the partitioning procedure 1,000 times<sup>59</sup>. The confidence intervals shown throughout the population analyses in this work correspond to the 95th percentiles of the distributions obtained by repeated partitioning.

**Statistical methods summary for figures presenting neural data.** Fig. 2a–c shows PSTHs for individual cells. For Fig. 2d,e, we computed the percentage of neurons that had a statistically significant change in firing rate during the CS (fractal or face image) presentation during the time epoch spanning 100 to 400 ms after CS appearance (response significantly different from the ITI,  $P < 0.05$ , paired *t*-test). The percentage of neurons depicted in Fig. 2d corresponds to the percentage of neurons that exhibited a significant response to at least one of the eight viewed monkeys ('Monkey'), at least one of the fractals ('Fractal') or at least one of the monkeys and at least one of the fractals ('Both'). Fig. 2f depicts decoding performance when using a linear decoder (see above). All neurons are included in this analysis (with at least 5 trials of data per condition). We then equalized the number of neurons in the three areas by subsampling. In Fig. 2g, neurons were categorized as preferring dominant or submissive monkey face images if average response to M1 was either at least 1 Hz greater in response to M1 compared to M8 (dominant-preferring) or to M8 compared to M1 (submissive-preferring). Neural responses to M2–M7 during the time epoch 100–400 ms after CS onset were then normalized and averaged across all dominant- or submissive-preferring neurons. A correlation analysis was then performed on the average normalized responses to M2–M7.

Fig. 3 presents data from a linear decoder used to analyze all neurons in our dataset (see above).

Fig. 4a–c presents data from a linear decoder used to analyze all neurons in our dataset (see above). In Fig. 4d–i, correlation analyses are performed to determine whether there is a relationship for each viewed monkey face image between its social index score (see behavioral analysis methods) and the average decoding performance for distinguishing that face image from each of the other seven face images. This analysis is performed in two different time epochs: 100–400 ms after CS onset (Fig. 4d–f) and 400–700 ms after CS onset (Fig. 4g–i), with the peak decoding performance in a 50-ms bin presented (see above methods description for the decoding analyses). All neurons are included in the analysis.

Supplementary Fig. 5 depicts example PSTHs constructed from the response of two different individual neurons.

Supplementary Fig. 6 depicts a correlation analysis that characterizes the relationship between the difference in social rank of any two face images and the average decoding performance for distinguishing between those face images, an analysis that captures how decoding performance depends upon social 'distance'.

In Supplementary Fig. 7, we selected neurons with  $\beta$  coefficients significantly different from 0 in either of the two time epochs (100–400 ms or 400–700 ms after CS onset) when analyzing trials with fractals (Supplementary Fig. 7a) or face images (Supplementary Fig. 7b) ( $P < 0.05$  using the linear regression analysis for single cells described above). We then used a correlation analysis to characterize the relationship between  $\beta$  coefficients for the same type of image presented (fractal or face image) in the two time epochs.

In Supplementary Fig. 8, we categorized neurons as preferring dominant or submissive neurons by using the sign of the  $\beta$  coefficient from the linear regression analysis applied to data from the first time epoch, 100–400 ms after CS onset (see above). We did not impose a statistical threshold as an inclusion criterion in this analysis. We then asked whether  $\beta$  coefficients for fractal images and face images from each time epoch were correlated with each other.

In Supplementary Fig. 9, for the groups of neurons identified in Fig. 2g (low- or high-rank-preferring neurons), we calculated the average firing rate during fractal image presentation in the first analysis epoch (100–400 ms after CS onset).

In Supplementary Fig. 10a, we categorized neurons as preferring dominant or submissive neurons by using the sign of the  $\beta$  coefficient from the linear regression analysis for neural activity from the second time epoch, 400–700 ms after CS onset (see above). We did not impose a statistical threshold as an inclusion criterion in this analysis. We then asked whether  $\beta$  coefficients for fractal images and face images from each time epoch were correlated with each other. Supplementary Fig. 10b,c shows the average normalized activity during fractal image presentation in the second analysis epoch (400–700 ms) for neurons that prefer dominant or submissive monkey face images (as in Fig. 2g).

In Supplementary Fig. 11, to compute the population PSTHs during the fractal and social trials, we partitioned neurons into two groups based on the preference of each neuron for large- or no-reward fractal images (unpaired *t*-test,  $P < 0.05$ ) during the first epoch and second epoch.

Supplementary Table 2 represents the number and percentage of neurons with a significant regression factor ( $P < 0.05$ ) for social and fractal images in each of the two analyzed time epochs.

**Life Sciences Reporting Summary.** Further information on experimental design is available in the Life Sciences Reporting Summary.

**Data availability.** The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Code availability.** No custom code has been used. Analysis have been performed by using Matlab statistical toolboxes. The analysis code employed is available upon reasonable request.

## References

- Miller, R. E. & Murphy, J. V. Social interactions of rhesus monkeys: 1. Food-getting dominance as a dependent variable. *J. Soc. Psychol.* **44**, 249–255 (1956).
- Hamilton, C. L. Comparison of two methods of dominance testing in the monkey. *Psychol. Rep.* **6**, 247–250 (1960).
- Fedigan, L.M. *Primate Paradigms: Sex Roles and Social Bonds* (Eden Press, Montreal, 1982).
- Chang, S. W. et al. Neuroethology of primate social behavior. *Proc. Natl. Acad. Sci. USA* **110**, 10387–10394 (2013).
- McNelis, N. L. & Boatright-Horowitz, S. L. Social monitoring in a primate group: the relationship between visual attention and hierarchical ranks. *Anim. Cogn.* **1**, 65–69 (1998).
- de Waal, F. B. & Yoshihara, D. Reconciliation and redirected affection in rhesus monkeys. *Behaviour* **85**, 224–241 (1983).
- Belova, M. A., Paton, J. J. & Salzman, C. D. Moment-to-moment tracking of state value in the amygdala. *J. Neurosci.* **28**, 10023–10030 (2008).
- Meyers, E. M., Freedman, D. J., Kreiman, G., Miller, E. K. & Poggio, T. Dynamic population coding of category information in inferior temporal and prefrontal cortex. *J. Neurophysiol.* **100**, 1407–1419 (2008).
- Golland, P., Liang, F., Mukherjee, S. & Panchenko, D. Permutation tests for classification. *Lect. Notes Comput. Sci.* **3559**, 501–515 (2005).

## Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work we publish. This form is published with all life science papers and is intended to promote consistency and transparency in reporting. All life sciences submissions use this form; while some list items might not apply to an individual manuscript, all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### ▶ Experimental design

#### 1. Sample size

Describe how sample size was determined.

We did not run any statistical test to determine sample sizes a priori. We chose the sample size based on literature in the field. Also most statements regarding statistical significance were based on bootstrap tests involving either 1,000 permutations. Such numbers of permutations are commonly assumed to be large enough to support statistical claims.

#### 2. Data exclusions

Describe any data exclusions.

In general, we include in our analysis all units that we recorded and all behavioral sessions in which we acquired neural data. However we used one exclusion criteria that we report in the methods and main text: to include a neuron in an analysis, we ask at least 5 repetitions per cell for all the conditions used in this data analysis (minimization of the inter-trial variability). In Fig 2G, neurons were classified based on their neural preference to M1 or M8 ( $FR > 1\text{spk/sec.}$  for M1 or M8. If FR were equal then the neuron was not included in the analysis).

#### 3. Replication

Describe whether the experimental findings were reliably reproduced.

We reproduced the same pattern of results with the 2 different monkeys involved in our study (cf. Fig. 1)

#### 4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

As report in the methods section, during decoding analysis, we re-sampled at each iterations (1000) the trials we used for training vs. trials used for testing. Stimuli presentation was also randomized.

#### 5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Data collection and analysis were not performed blind to the conditions of the experiments. However we did not select the type of neurons we recorded during neural data acquisition. Moreover, during off-line sorting of the neurons, the task events and monkey behavior were hidden from the investigators.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

## 6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or the Methods section if additional space is needed).

- n/a Confirmed
- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
  - A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly.
  - A statement indicating how many times each experiment was replicated
  - The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
  - A description of any assumptions or corrections, such as an adjustment for multiple comparisons
  - The test results (e.g.  $p$  values) given as exact values whenever possible and with confidence intervals noted
  - A summary of the descriptive statistics, including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
  - Clearly defined error bars

See the web collection on [statistics for biologists](#) for further resources and guidance.

## ► Software

Policy information about [availability of computer code](#)

### 7. Software

Describe the software used to analyze the data in this study.

We used the statistical toolbox of Matlab (Matlab\_R2016b) and Plexon offline sorter (v3.3) for the off-line sorting of the neural data.

For all studies, we encourage code deposition in a community repository (e.g. GitHub). Authors must make computer code available to editors and reviewers upon request. The *Nature Methods* [guidance for providing algorithms and software for publication](#) may be useful for any submission.

## ► Materials and reagents

Policy information about [availability of materials](#)

### 8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

No unique materials were used.

### 9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

No antibodies were used.

### 10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

No eukaryotic cell lines were used.

b. Describe the method of cell line authentication used.

N/A

c. Report whether the cell lines were tested for mycoplasma contamination.

N/A

d. If any of the cell lines used in the paper are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

N/A

## ► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

### 11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

2 male Rhesus macaque (*macaca mulatta*). Monkey F was born on 10/17/03 and monkey R was born on 02/17/05.

## 12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

The study did not involve human research participants.