Shared yet dissociable neural codes across eye gaze, valence and expectation

Shared yet dissociable neural codes across eye-gaze and valence-expectation

Raviv Pryluk[1], Yosef Shohat[1], Anna Morozov[1], Dafna Friedman[1] , Aryeh H. Taub[1] , Rony Paz[1]

[1] Department of Neurobiology, Weizmann Institute of Science, Israel

Correspondence should be addressed to R.P. (rony.paz@weizmann.ac.il)

Abstract

The eye-gaze of others is a prominent social cue in primates and crucial for communication[1-10]. Although gaze can signal threat and elicit anxiety[6, 11, 12], it remains unclear if it shares neural circuitry with stimulus-value. Importantly, gaze not only has valence, but can also serve as predictor for the outcome of a social encounter: negative or positive[2, 11, 12]. Here we show that neural codes overlap for gaze and valence through two different mechanisms: one for the outcome, and another for its expectation. Monkeys participated in the human-intruder-test[12, 13] that included direct and averted gaze, interleaved with blocks of aversive and appetitive conditioning[14]. We find that single-neurons in the amygdala encode gaze[15], whereas neurons in the anterior-cingulate-cortex(ACC) encode social context[16], but not gaze. We identify a shared amygdala population where neural responses to direct and averted gaze parallel the responses to aversive and appetitive stimulus, correspondingly. Further, we distinguish between two mechanisms: an overall-activity scheme that is used for gaze and the unconditioned-stimulus(US), and a correlated-selectivity scheme that is used for gaze and the conditioned-stimulus(CS). The findings suggest new insights on the origins of the neural mechanisms underlying social and valence computations, and might shed light on social-anxiety and the comorbidity between anxiety and impaired social interactions.

Acknowledgments: We thank Dr. Yoav Kfir for scientific and technical advice; Drs. Eilat Kahana and Nir Samuel for medical and surgical procedures; Daniel Goldin for engineering design, Dr. Edna Furman-Haran and Fanny Attar for MRI procedures. This work was supported by ISF #2352/19 and ERC-2016-CoG #724910 grants to R. Paz.

Main text statements: R.Pryluk and R.Paz designed the study. R.Pryluk performed all experiments. Y.S, A.H.T and A.M contributed to experiments. R. Pryluk developed methods and analyzed the data. A.M. and D.F. contributed to data analysis and editing of the manuscript. R. Pryluk and R. Paz wrote the manuscript.

Data availability: All data supporting the findings of this study are available from the corresponding author upon reasonable request.

Code availability: Custom code for behavioral and electrophysiological tests is available from the corresponding author upon reasonable request.

Competing interests: The authors declare no competing interests.

Main text

Recognizing and learning about potentially harmful or beneficial stimuli is crucial for survival of all organisms. In humans and primates in general, facial expressions, and in particular the eye-gaze of others, is a prominent and highly instructive signal [2-4, 11, 17]. Specifically, averted or direct gaze is a social signal that can indicate submissive vs. aggressive interactions, correspondingly. In agreement with this, gaze was shown to elicit anxiety in primates [6, 11, 12], and evokes responses in the amygdala [15, 18-20] – a brain region that serves as a hub for emotional responses in general and threat and anxiety in particular [20]. Moreover, gaze processing is disrupted in several neurodevelopmental and social-disorders [1, 5, 7, 8] where abnormal activity of the amygdala is linked to gaze avoidance [9, 10]. Importantly, gaze is not only a valence-signal by itself, but can also serve as a predictor for future outcomes: aversive if an intruder maintains direct eye-contact (stare), or potentially rewarding if the other avoids eye-contact. This is in line with the amygdala playing a role not only in signaling outcome-valence (appetitive-aversive), but also in learning via conditioning [14, 21, 22] and signaling expectation for the outcome, namely exhibit responses to a conditioned-stimulus (CS) [14, 22]. However, it remains unknown whether similar mechanisms are used for coding of valence and eye-gaze; and moreover, whether there exists a shared coding for eye-gaze and outcome-expectation. Toward this end, we adapted the human intruder test (HIT) [12, 13]. HIT is widely used for assessing anxiety and defensive behaviors in non-human-primates, similar to the 'stranger test' in human infants [23]. We recorded the activity of single neurons in the Amygdala and the ACC during live interactions in a modified HIT paradigm that includes averted vs. direct gaze of the intruder and combined with an affective conditioning paradigm. We first validated previous results and show that here as well, both the ACC and the amygdala code for valence[21], but only the amygdala codes for gaze[15]. In line with our hypothesis, we demonstrate that in amygdala networks, valence of both outcome and its expectation are coded in the same population that also codes for the gaze of others, but via two different population codes.

Two monkeys participated in a modified version of the human intruder test (HIT) (Fig. 1a). Each HIT block consisted of 18 interactions with a human intruder that is seated behind an LCD shutter (<1ms RT), and when the shutter opens gazes directly at the monkey's eyes (eye-contact, EC), or away from the monkey (averted-gaze / no-eye-contact, NEC). These HIT blocks were interleaved with conditioning blocks of either appetitive or aversive trials (>=8 trials in a block, Fig.1b,c), where the shutter opening serves as the conditioned-stimulus(CS) and is followed after one second delay by the outcome/unconditioned-stimulus(US), liquid-reward or airpuff in appetitive/aversive blocks correspondingly. We tracked the eye-position of the monkeys and extracted four regions of interest (ROI, Fig.1d): 1. the eye-region of the intruder; 2. the face-region of the intruder; 3. the whole shutter region; and 4. outside the shutter region. Oculomotor behavior revealed distinct patterns (Fig.1d-g; Extended.Fig.1): shutter opening in the HIT blocks induced more interest in the eyes ROI compared to the conditioning blocks (Fig.1d, Kolmogorov-smirnov, p<1e-8, n-trials= 3108/2090 in HIT/conditioning trials; 49 sessions, 24/25 per monkey). After exploring the eye of the human intruder, the monkeys continues to look more to the eyes/face ROI in blocks of direct-gaze (Fig.1g, EC vs. NEC, 2 , p<1e-3). We further aligned each trial separately according to the first time the monkey gazed at the intruder eyes (Interquartile range: 180-700ms) and found similar results (Extended.Fig.1, $\chi 2$, p<1e-6).

We quantified elicited facial expressions and find that monkeys produced more facial expressions when the intruders made eye-contact (Fig.1h,i, $\chi 2$, p<1e-2; Extended.Fig.2), in agreement with the stressful,

threatening, and defensive responses that are traditionally induced by direct-gaze of a human intruder [12]. Heart-rate and heart-rate-variability (HRV) further suggest anxiety-related responses (Fig.1j,k, t-test, $p<0.05$). In the conditioning blocks, the monkeys quickly learned to distinguish and anticipate the different outcomes (appetitive/aversive) after shutter opening in each specific block (Fig.1l, $\chi 2$, $p<1e-3$). In aversive-airpuff blocks, they closed the eyes after shutter opening both in preparation for the airpuff as well as immediately after its delivery (Fig.1l, $\chi 2$,$p<1e-3$). In addition, they withheld inhale before the expected airpuff, but not before reward (Fig.1m, t-test, $p<1e-3$). The HR and HRV in conditioning blocks were different between airpuff and reward (Fig.1j,k, t-test, $p<1e-3$) and showing the same direction of modulation as in the human intruder (airpuff /EC is higher than reward /NEC, correspondingly). Therefore, there was a clear differential behavioral response in HIT sessions between eye-contact of the intruder and no-eye-contact, and there was a clear differential response between appetitive and aversive blocks, for both the US/outcome, and the CS/expectation.

To examine and compare neural responses, we recorded single-units from the basolateral-complex of the amygdala (BLA) and the anterior-cingulate-cortex (ACC) (Fig.2a, n=24/25 sessions per monkey, n= 356/203 neurons in the ACC/Amygdala, 224/103 and 132/100 per monkey). We define two epochs in the conditioning blocks: a preparatory/expectation epoch (CS-related, after the shutter opening but before US delivery), and an outcome epoch (US-related, following delivery of airpuff/reward) (Fig.2b). Confirming previous studies, we find that neurons in the amygdala and the ACC respond to the appetitive CS (Amy: 35/203, ACC: 43/356, $\chi 2$,$p<1e-3$ for both), respond to the aversive CS (Amy: 36/203, ACC: 57/356, $\chi 2$,$p<1e-3$ for both), and also discriminate between valence (Amy: 37/203, ACC: 71/356, $\chi 2$,$p<1e-3$ for both). Moreover, similar proportions of cells were responsive in the two regions (Fig.2c, 2 , $p>0.09$ for all). Similarly, neurons in the ACC and in the amygdala responded to the appetitive US (Amy: 25/203, ACC: 49/356, $p<1e-2$ for both), and aversive US (Amy: 73/203, ACC: 106/356, $p<1e-3$ for both) again with similar proportions in both regions ($\chi 2$, $p>0.1$). However, more amygdala neurons discriminated valence between appetitive and aversive outcome (Fig.2d, Amy: 90/203, ACC:114/356, $\chi 2$, $p<1e-2$).

In the HIT blocks (Fig.2e), neural responses were computed from the time when the monkey first looks at the eyes-ROI, as this is the first time that the monkey can differentiate between EC and NEC (Interquartile range: 180-700ms, Fig.1d). There were more responsive neurons in the amygdala than in the ACC during HIT blocks (Fig.2f, Amy: 58/203, ACC: 50/356, 2 , $p<1e-3$), and more amygdala neurons discriminate between EC and NEC of the intruder (Fig.2f, Amy: 21/203, ACC: 17/356, $\chi 2$, $p<0.05$). The number of ACC neurons that coded for the intruder gaze was not different than chance (Binomial test, $p>0.1$). We tested for overlap in responses and found that the proportion of neurons that responded to both gaze and valence was not different than chance, both in the amygdala and in the ACC (Fig.2g Binomial test, $p>0.1$).
We noticed that the proportion of amygdala neurons that code for gaze is low compared to the proportion of neurons that code for valence, both for CS-related responses and for US-related responses (Fig.2c,d,f, $\chi 2$, CS: $p<0.05$, US: $p<1e-3$), in line with previous studies in both monkeys and humans [15, 24]. Nevertheless, because a neuron can contribute at the population level even if it does not exhibit a significant response by itself, we further tested whether the combined ensemble of recorded neurons holds information about the eye-gaze of others by training a linear decoder on population vectors. In accordance with the single-cell analyses, population activity in the amygdala and the ACC could discriminate between appetitive and aversive trials, both using CS-related and using US-related activity (Fig.2h,

bootstrap analysis, CI 95%). However, only the amygdala population could discriminate between EC and NEC trials, whereas the ACC population did not exceed chance-level (Fig.2h, bootstrap analysis, CI 95%).

We conclude that in the current paradigm, similar to previous findings, the amygdala and the ACC code for valence[21], but only the amygdala codes for the eye-gaze of the intruder[15]. It is true both at the single-cell and at the population level.

The finding that the amygdala holds information about valence as well as eye-gaze of others within the same circuitry suggests that there might be a shared population code in the neural ensembles. In order to test this hypothesis of shared coding for valence and gaze, we used the decoder approach again, but this time we trained on one type of trials and tested on another. If discrimination accuracy is above chance-level, this would mean that the population uses similar mechanisms to hold information for one situation - appetitive vs. aversive, as for the other - EC vs. NEC. We therefore trained a linear decoder to distinguish between trials of EC and NEC and tested it on distinguishing between trials of aversive and appetitive. Importantly, this was done separately for the CS-related and the US-related responses.

In agreement with the aforementioned finding that the ACC does not hold information about eye-gaze, the decoding performance in the ACC was not different than chance in both CS and US related activity (Fig.3a,b top insets, bootstrap analysis, CI 95%). In contrast, decoding performance was significantly above chance level when using amygdala population, and moreover, it was the case when using either CS-related activity or US-related activity (Fig.3a,b, Extended.Fig.3, bootstrap analysis, CI 95%). Performance was approximately linear in the number of neurons, starting from chance-level and rising to more than 80% accuracy when using all available amygdala neurons (CS: 82.5%, US: 80%, n=203, $p<0.001$ for both; Fig.3a,b bottom insets). This suggests that the shared coding of valence and eye-gaze is not due to the few neurons that had significant responses to both contexts (Fig.2g), a notion that was further supported by the finding that accuracy remained similar when dropping these few neurons (CS: 81%, n=201; US: 79%, n=198). These findings demonstrate that a shared population code is used by amygdala neurons, because the decoder was trained only on gaze discrimination, yet successfully tested on valence discrimination.

In general, there could be two shared activity schemes that would allow training on one context and decoding of the other. In the first, termed here *correlated-selectivity*, neurons respond similarly to gaze and valence (Fig.3c). This means that the neurons respond in the same direction and with similar proportion (decrease/increase firing rates proportionally) for NEC vs. EC as for appetitive vs. aversive. Namely, a neuron's response is correlated along eye-gaze and valence. Alternatively, in the second option termed *overall-activity*, a population of neurons respond only in the same direction, high or low overall average firing-rate, to gaze and valence, yet individual neurons are not correlated across the contexts (Fig.3d). We therefore tested which scheme applies for the amygdala population, and is it different between CS-related responses and US-related activity. To do so, we applied several approaches.

We first examined the activity at the single cell level. Each neuron was assigned a selectivity index for gaze (SIG, -1 to 1, NEC to EC) and a selectivity index for valence (SIV, from -1 to 1, appetitive to aversive). The joint distribution of indices with the same direction of modulation was high in CS period (Fig.3e,f; $p<1e-3$ $\chi 2$ compared to chance-level; US: $p>0.1$), in opposite to the joint distribution with only positive modulation which was high in the US period (Fig.3e,f; $p<1e-3$ $\chi 2$; CS: $p>0.1$). Moreover, the two indices were correlated across the whole population only during CS activity ($r=0.26$, $p<0.01$ taking

only neurons with SI>1/3; r=0.2, p<0.01 and for the whole population; US: p>0.4; t-tests), and even when taking positive indices only, demonstrating that the correlation is beyond sign (Fig.3g-right, r=0.3,p<0.02; US: p>0.4).

Next, we used linear-regression on the individual responses for gaze, comparing EC to NEC, and separately on the responses for valence, comparing aversive to appetitive. We obtained and compared two separate coefficients: $\beta_{valence}$ that represents the difference in firing rate between airpuff and reward, and β_{gaze} that represents the difference in firing rate between EC and NEC. If the two coefficients are similar for individual neurons, it means that neurons code valence and eye-gaze not only along the same direction, but also with similar modulation. We found that the two coefficients were positively correlated in amygdala neurons, but only when using CS-related activity and not when using US-related activity (Fig.4a-c, Pearson correlation, amygdala: CS: r=0.4, p<1e-8, US: r=0.03 ,p>0.5; ACC: CS: r=-0.1, p<1e-2, US: r=0.03 ,p>0.5). This observation supports a *correlated-selectivity* scheme between valence and gaze for the CS epoch, yet an *overall-activity* for the US epoch. The *overall-activity* scheme for the US is further supported by direct examination of overall increases/decreases in firing-rates for direct gaze and US valence (Extended.Fig.4, Z-test p<1e-3).

This finding was further validated by examining the scalar product between the two coefficients ($\vec{\beta}_{gaze}$ and $\vec{\beta}_{valence}$). If more neurons respond in similar proportion (*correlated-selectivity)*, then the scalar-product would be positive; otherwise, the scalar product will be close to zero if neurons respond in random order (or negative if in opposite directions). In the amygdala, using CS-related activity outperforms a shuffling test (Fig.4d, bootstrap), yet using US-related activity does not (Fig.4e, bootstrap). In the ACC, neurons were similar or lower than the shuffled test (Fig.4f, bootstrap). In addition, the mean value for the US-related shuffled activity is higher than for the CS-related shuffled activity (Fig.4d,e, CS=0.1, US=1.8, bootstrap, p<0.05). This is because more neurons both in gaze and in US-valence increase their firing rate, resulting in a higher positive scalar product for shuffled neurons, further supporting the *overall-activity* scheme. In contrast, for the CS the similarity in the response increases the scalar product in the real neurons but not in the shuffled population.

To test the two schemes at the population level, we computed the angles between the decision boundaries of two linear decoders: one boundary that separates EC from NEC and one that separates aversive from appetitive. When computed over the US epoch, or using ACC population, the decision boundaries of valence and gaze are closer to being perpendicular to one another (dot-product not significantly different from zero), whereas only using CS activity from the amygdala population shows a significant difference from perpendicular decision boundaries (Fig.4g, bootstrap, CI 95%).
Finally, we trained the decoder on gaze and tested on valence while shuffling the order of neurons. This approach is used to test if it is the specific ensemble of neurons that matters, or just an overall increase in firing rate. In line with the previous results, performance using amygdala activity from the CS epoch was decreased dramatically from actual to shuffled neurons (Fig.4h, Extended.Fig.5 bootstrap analysis with CI 95%), whereas using US activity it even slightly increased (Fig.4h,i,j, Extended.Fig.5 bootstrap analysis with CI 95%), further supporting the two different shared coding schemes: *correlated-selectivity* between gaze and CS-valence, and *overall-activity* between gaze and US-valence.

The eyes of others became a prominent signal along evolution due to anatomical changes in facial morphology that forced a shift in salience from the shape of the face to the eyes [2]. The importance of the

amygdala in the processing of eye-gaze was shown in humans and in macaques [9, 15, 18]. Here, we recorded neural activity in the amygdala and the ACC during live interactions in a modified version of the human intruder test (HIT)[12] that included also a conditioning paradigm. Whereas both regions differentiated between valence in their CS-related and US-related responses [21], only the amygdala differentiated between averted vs. direct gaze of an intruder. This finding is in-line with multidimensional selectivity found in amygdala neurons[25, 26] and increased robustness compared to the ACC[27]. Importantly, we find that in the amygdala, both CS-related and US-related responses are shared with the eye-gaze of the intruder and in a valence-specific manner, namely aversive (airpuff) to appetitive (reward) parallel direct to averted eye-gaze. Our results, obtained in live-interactions, comparing aversive-to-appetitive with natural eye-gaze, suggest that social value evolved from, or in parallel to, primary-reinforcer value. Together with recent findings[28], the results further support the theory that processing of social stimuli and specifically eye-gaze does not occur in separate dedicated neural circuits [2, 28, 29].

The naturalistic paradigm we employed enables live social interactions and therefore important for the interpretation of natural behaviors, yet it also imposes some constraints on the possible contributors. To address this, we validated that our findings cannot be explained by differences that accompany direct vs. averted eye-gaze, such as vocalizations (of any type, Extended.Fig.6a), self-motor activity (Extended.Fig.6b), facial expressions (Extended.Fig.6c-e), saccades (Extended.Fig.1b), and stimulus saliency (Extended.Figs.7,8,9,10). The fact that we identified two different coding schemes argues against the possibility that the shared code reflects a general saliency and/or category code (Extended.Figs.7,8). This was further supported by control experiments showing that amygdala neurons code for species-differences[19], but this code was not shared with the outcome expectation (Extended.Fig.9); and additional experiments demonstrating that direct and averted gaze have different value compared to neutral trials (Extended.Fig.10).

We identified two different coding schemes that allow decoding of value based on responses to eye-gaze. The *overall-activity* scheme that is shared across gaze and outcome (US) occurs by an overall increase in firing rate, and suggests a simpler mechanism that points to origins within the same circuitry, where an aversive outcome is similar in value to a predator gaze[11] or to a threat by a peer. It is also in line with the findings of the human-intruder-test where gaze elicits anxiety[6, 12, 30]. The coding of expectation, namely the learned CS, was also shared with eye-gaze responses; but it was shared via a *correlated-selectivity* scheme that requires the responses to be correlated at a single-neuron level (rather than only on average over the population). Because *correlated-selectivity* might require more specific wiring design, and because the amygdala has evolved in parallel to the development of social interactions [31, 32], we suggest that *correlated-selectivity* could have facilitated the later evolution of other complex social processes such as learning by observation [8, 33] and social-based decision-making in extended circuits [16, 17]. Specifically, it can be used by the animal to anticipate social outcomes based on context - a direct prolonged gaze likely calls for a challenge and predicts a confrontation that entails dangerous outcome; whereas an averted gaze usually predicts a submissive and permissive encounter and potentially rewarding (mating, food sharing/offering). There are very few contexts in which a prolonged gaze is positive (e.g. mother-baby interactions), and it would be interesting to test if amygdala ensembles reverse coding direction, or rather contribute to integration of cues in down-stream circuits. Overall, our findings suggest new insights into coding schemes in the primate amygdala that underlie social-interactions, valence, and outcome-expectancy, and provide a new framework to understand social-anxiety and the comorbidity of anxiety, depression, and impaired social interactions[34].

References

1. Jones, W. & Klin, A. Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature* **504**, 427-431 (2013).
2. Emery, N.J. The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience and biobehavioral reviews* **24**, 581-604 (2000).
3. Gobel, M.S., Kim, H.S. & Richardson, D.C. The dual function of social gaze. *Cognition* **136**, 359-364 (2015).
4. Adolphs, R. Neural systems for recognizing emotion. *Curr Opin Neurobiol* **12**, 169-177 (2002).
5. Zhou, Y.*, et al.* Atypical behaviour and connectivity in SHANK3-mutant macaques. *Nature* **570**, 326-331 (2019).
6. Schneier, F.R., Kent, J.M., Star, A. & Hirsch, J. Neural circuitry of submissive behavior in social anxiety disorder: A preliminary study of response to direct eye gaze. *Psychiatry Res* **173**, 248-250 (2009).
7. Schneier, F.R., Rodebaugh, T.L., Blanco, C., Lewin, H. & Liebowitz, M.R. Fear and avoidance of eye contact in social anxiety disorder. *Comprehensive psychiatry* **52**, 81-87 (2011).
8. Gariepy, J.F.*, et al.* Social learning in humans and other animals. *Front Neurosci* **8**, 58 (2014).
9. Rutishauser, U.*, et al.* Single-neuron correlates of atypical face processing in autism. *Neuron* **80**, 887-899 (2013).
10. Hadjikhani, N.*, et al.* Look me in the eyes: constraining gaze in the eye-region provokes abnormally high subcortical activation in autism. *Scientific Reports* **7**, 3163 (2017).
11. Shepherd, S.V. Following gaze: gaze-following behavior as a window into social cognition. *Frontiers in integrative neuroscience* **4**, 5-5 (2010).
12. Kalin, N.H. & Shelton, S.E. Defensive behaviors in infant rhesus monkeys: environmental cues and neurochemical regulation. *Science (New York, N.Y.)* **243**, 1718-1721 (1989).
13. Oler, J.A.*, et al.* Amygdalar and hippocampal substrates of anxious temperament differ in their heritability. *Nature* **466**, 864-868 (2010).
14. Janak, P.H. & Tye, K.M. From circuits to behaviour in the amygdala. *Nature* **517**, 284-292 (2015).
15. Mosher, C.P., Zimmerman, P.E. & Gothard, K.M. Neurons in the monkey amygdala detect eye contact during naturalistic social interactions. *Curr Biol* **24**, 2459-2464 (2014).
16. Haroush, K. & Williams, Ziv M. Neuronal Prediction of Opponent's Behavior during Cooperative Social Interchange in Primates. *Cell* **160**, 1233-1245 (2015).
17. Ballesta, S. & Duhamel, J.-R. *Rudimentary empathy in macaques' social decision-making* (2015).
18. Gamer, M. & Büchel, C. Amygdala Activation Predicts Gaze toward Fearful Eyes. *The Journal of Neuroscience* **29**, 9123-9126 (2009).
19. Gothard, K.M., Battaglia, F.P., Erickson, C.A., Spitler, K.M. & Amaral, D.G. Neural responses to facial expression and face identity in the monkey amygdala. *Journal of neurophysiology* **97**, 1671-1683 (2007).
20. Adolphs, R. What does the amygdala contribute to social cognition? *Ann N Y Acad Sci* **1191**, 42-61 (2010).
21. Tovote, P., Fadok, J.P. & Luthi, A. Neuronal circuits for fear and anxiety. *Nat Rev Neurosci* **16**, 317-331 (2015).
22. Duvarci, S. & Pare, D. Amygdala microcircuits controlling learned fear. *Neuron* **82**, 966-980 (2014).
23. Kagan, J., Reznick, J.S. & Snidman, N. Biological bases of childhood shyness. *Science (New York, N.Y.)* **240**, 167-171 (1988).

24. Mormann, F.*, et al.* Neurons in the human amygdala encode face identity, but not gaze direction. *Nat Neurosci* **18**, 1568-1570 (2015).
25. Putnam, P.T. & Gothard, K.M. Multidimensional Neural Selectivity in the Primate Amygdala. *eNeuro* **6** (2019).
26. Kyriazi, P., Headley, D.B. & Pare, D. Multi-dimensional Coding by Basolateral Amygdala Neurons. *Neuron* **99**, 1315-1328 e1315 (2018).
27. Pryluk, R., Kfir, Y., Gelbard-Sagiv, H., Fried, I. & Paz, R. A Tradeoff in the Neural Code across Regions and Species. *Cell* **176**, 597-609.e518 (2019).
28. Munuera, J., Rigotti, M. & Salzman, C.D. Shared neural coding for social hierarchy and reward value in primate amygdala. *Nat Neurosci* **21**, 415-423 (2018).
29. Dunbar, R.I.M. The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews* **6**, 178-190 (1998).
30. Myllyneva, A., Ranta, K. & Hietanen, J.K. Psychophysiological responses to eye contact in adolescents with social anxiety disorder. *Biol Psychol* **109**, 151-158 (2015).
31. Sallet, J.*, et al.* Social network size affects neural circuits in macaques. *Science* **334**, 697-700 (2011).
32. 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).
33. Allsop, S.A.*, et al.* Corticoamygdala Transfer of Socially Derived Information Gates Observational Learning. *Cell* **173**, 1329-1342 e1318 (2018).
34. Stein, M.B. & Stein, D.J. Social anxiety disorder. *Lancet* **371**, 1115-1125 (2008).
35. Li, D., Babcock, J. & Parkhurst, D.J. openEyes: a low-cost head-mounted eye-tracking solution. in *Proceedings of the 2006 symposium on Eye tracking research & applications* 95-100 (ACM, 2006).
36. Li, D., Winfield, D. & Parkhurst, D.J. Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches. in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops* 79-79 (IEEE, 2005).
37. Mitz, A.R., Chacko, R.V., Putnam, P.T., Rudebeck, P.H. & Murray, E.A. Using pupil size and heart rate to infer affective states during behavioral neurophysiology and neuropsychology experiments. *Journal of neuroscience methods* **279**, 1-12 (2017).
38. Electrophysiology Task Force of the European Society of Cardiology the North American Society of, P. Heart Rate Variability. *Circulation* **93**, 1043-1065 (1996).
39. 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. *Journal of neurophysiology* **100**, 1407-1419 (2008).

Figure legends:

Figure 1. Paradigm and behavior during the Human-Intruder-Test (HIT) and affective conditioning blocks.

a. Human-Intruder-Test: The shutter opens and closes 18 times, and the human intruder pseudorandomly alters between direct-gaze (eye-contact, EC) and averted eye-gaze (no-eye-contact, NEC).
b. Classical conditioning blocks of either appetitive (reward) and aversive (airpuff) stimulus. In these blocks, the shutter-open serves as a predictor (conditioned-stimulus, CS) to the appetitive/aversive outcome (unconditioned-stimulus, US).
c. An example of the pseudorandom order of blocks in one recording session, with at least 120 seconds between blocks.
d. Regions of interest (ROI) for the eye-tracking of the observer monkey (left): Only to the eyes of the human intruder (green); Only to the face of the intruder (pink); The whole shutter (white); The whole possible space (gray). Notice we report the same regions even when there is no intruder (the regions are similar across intruders because the faces are accurately aligned by positioning). Cumulative density function (right) of the first time after shutter opening that the monkeys look into the eyes ROI, separately for HIT and for conditioning blocks.
*** represents a significant difference (Kolmogorov-smirnov, p<1e-8, n-trials= 3108/2090 in HIT/conditioning trials; 49 sessions, 24/25 per monkey).
e. Example of one shutter opening in an HIT block. Filled-circles (red) mark the location of the monkey's gaze overlaid on an intruder real-position (schematic). Shown are three consecutive time windows (one before shutter opening and two after).
f. Density function of all eye locations in the HIT blocks in the same three consecutive time windows as in (E). Immediately after shutter opening and for few hundreds of milliseconds, the monkey looks mainly at the eyes of the human intruder.
g. Left: Proportion of looking to the eyes-ROI in EC and NEC trials (Mean +/- SEM). The monkey first looks to the eyes region of the intruder, and then immediately breaks fixation in NEC trials significantly more than in EC trials.
Right: The difference in proportion of the monkey's look towards the eye and to the face ROI between EC and NEC trials. Both are significantly positive indicating that in EC trials the monkey maintains fixation to the face/eyes of the intruder.
Upper black line represents a significant difference (p<0.05, $\chi 2$, n-trial=1480/1628 in NEC/EC).
h. Shown are schemes of typical facial expressions made by the monkeys in EC trials (middle, "aggressive"), in NEC trials (right,"interest"), compared to a neutral expression (left).
i. The overall change in the facial expression in EC and NEC (Mean +/- SEM). Shown is the Root-Mean-Square (RMS) of change in the image over the whole face (left) and only for the lower half of the face (right), compared to the neutral expression. Upper black line represents a significant difference (p<0.05, two-sided t-test, n-trial=1480/1628 in NEC/EC). See methods and Extended.Fig.2.
j. Differences in heart-rate (Mean +/- SEM) between EC and NEC trials and between reward and airpuff trials. * and *** represent a significant difference (t-test, two-sided, p<0.05, n-trials 1703/1765 in NEC/EC and p<e-3, n-trials 1352/712 in Reward/Airpuff).

k. Differences in heart-rate-variability (HRV, Mean +/- SEM) between EC and NEC trials and between reward and airpuff trials. * and *** represent a significant difference (t-test, two-sided, p<0.05, n-trials 1703/1765 in NEC/EC and p<e-3, n-trials 1352/712 in Reward/Airpuff).
l. Response (Mean +/- SEM) to aversive (airpuff) vs. appetitive (reward) in the oculomotor behavior. *** represent a significant difference ($\chi 2$,p<1e-3, n-trial=1375/715 in Reward/Airpuff)
m. Differences in respiratory-rate (Mean +/- SEM) after shutter-opens between EC and NEC trials and between reward and airpuff trials. n.s and *** represent a non-significant/significant difference (t-test, two-sided, p=0.82, n-trials 1703/1765 in NEC/EC and p<e-3, n-trials 1352/712 in Reward/Airpuff).

Figure 2. The amygdala codes for gaze and valence, and the ACC mainly codes valence

a. Recording locations: MRI with electrode directed into the BLA (AC=-3); Recording locations overlaid on a primate brain map (AC=0) and on an MRI scan (AC=0).
b. PSTHs and raster plots of two representative neurons in the ACC and two in the amygdala during conditioning block.
c. Proportion of neurons (Mean +/- SEM) in the ACC and the amygdala (n-neurons=356/203 respectively) that respond to the CS (shutter-open) in aversive trials (left), appetitive trials (middle), and discriminate between the two (right).
d. Proportion of neurons (Mean +/- SEM) in the ACC and in the amygdala (n-neurons=356/203 respectively) that respond after the US (outcome) to the airpuff (left), reward (middle), and discriminate between the two (right). ** represents a significant $\chi 2$, p<e-2.
e. PSTH and raster plot of two representative neurons in the amygdala during the human-intruder block (HIT).
f. The proportion of neurons (Mean +/- SEM) in the ACC and the amygdala (n-neurons=356/203 respectively) that respond significantly in the HIT blocks, and that discriminate between EC and NEC (gaze neurons). * and *** represent a significant $\chi 2$, p<0.05 and p<e-3 respectively.
g. Overlaps in the number of neurons that respond across the different tasks. The size of each area is proportional to the percentage of neurons. The numbers inside the Venn diagram represents the total number of each group (Gaze neurons, CS valence and US valence) whereas the numbers outside represent overlaps.
h. Population decoding accuracy for HIT vs. conditioning blocks. Discriminating appetitive from aversive with amygdala and ACC neurons is significant using both the CS and the US responses, whereas only the Amygdala can decode gaze i.e. eye-contact from no-eye-contact. Significant above chance was tested in bootstrap analysis (n_rep=1000, n_ACC=356, n_AMY=203) with CI 95%. In the violin diagram, red represents median and black the mean.

Figure 3. Shared coding for valence and gaze in amygdala neurons

a. Population decoding accuracy (Mean +/- STD, bootstrap CI=95%, n_rep=1000, n_AMY=203) when training on eye-gaze (EC vs. NEC) and testing on valence (aversive vs. appetitive), using CS-related activity.
Right-top inset: Similar format using ACC population (n_ACC=356).
Right-bottom inset: Peak decoding accuracy using increasing numbers of neurons.

b. Same as (A) but using US-related activity.

c. The *correlated-selectivity* scheme. Here, neurons respond similarly to gaze and valence, meaning their response is correlated across NEC-to-EC and Appetitive-to-Aversive. Shown is the optimal linear separator for the neural population (demonstrated here for three neurons) during the HIT trials (EC trials in circles, NEC in triangles). A similar presentation is shown for the Conditioning (airpuff trials in circles, reward in triangles), overlaid with the separating surface from the HIT. The similar surfaces allow correct decoding.

d. Same as (c) for the *overall-activity* scheme. Here, different neurons in the population respond with similar changes in firing rate to gaze and valence, but individual neurons are not correlated. Although the separating surfaces are different, neurons provide enough spikes overall to allow correct decoding.

e. Distribution of Selectivity-index for Gaze (SIG, blue), overlaid with neurons that have the same direction of modulation for SIG and SIV (red), for CS activity (left), and US activity (right).

f. Left: proportion of neurons (mean and SEM) with same direction of modulation is higher than chance in CS only ($p<1e-3$ $\chi 2$, n_AMY=203,n_ACC=356), but not in US (shown is also ACC for comparison, dashed-line is chance level).
Right: proportion of neurons (mean and SEM) with positive indices is higher than chance in US only ($p<1e-3$ $\chi 2$, n_AMY=203,n_ACC=356), but not in CS (shown is also ACC for comparison, dashed-line is chance level).

g. Selectivity-index for gaze (SIG) is correlated with the Selectivity-index for valence (SIV) across the whole population during CS activity only ($r=0.26$, $p<0.01$ taking only classically selective neurons with SI>1/3; $r=0.2$, $p<0.01$ for the whole population; US: $p>0.4$; t-tests).
Right: Considering only positive indices ($r=0.3$, $p<0.02$, Pearson; US: $p>0.4$; t-tests two sided, n=203), demonstrating that the correlation is beyond sign only.

Figure 4. An *overall-activity* coding for eye-gaze and US-valence and a *correlated-selectivity* coding for eye-gaze and CS-valence

a. Correlation (Pearson's, n_AMY=203 and n_ACC=356) between the linear-regression coefficients of gaze (eye-contact vs. no-eye-contact, x-axis) and of valence (aversive vs. appetitive, y-axis) using CS-related activity. All amygdala neurons are shown. The beta values are from the time epochs of the maximal decoding from Fig.3.
b. Same as (A) using US-related activity.
c. Same as (A) and (B) for ACC activity, CS-related (top) and US-related (bottom).
d. Neurons respond in the same direction for eye-gaze and valence using CS-related activity, as evident by the scalar-product between the coefficients of gaze and of valence for each neuron. Black asterisks represent data from real neurons and shaded-magenta is 95% confidence interval based on bootstrap shuffle.
e. Same as (D) using US-related activity.
f. Same as (C) and (D) for ACC activity, CS-related (top) and US-related (bottom).
g. The angle between the decision boundaries derived from the population-vector of gaze and valence separately (shown is the scalar product between the two vectors). In the Violin diagram red represents the median and black the mean. n_AMY=203 and n_ACC=356
h. Population decoding accuracy for real and shuffled neurons using CS-related activity.
i. Same as (H) for ACC activity.
j. Cumulative-distribution of the difference in decoding accuracy between real and shuffled neurons of the amygdala. *** represents a significant difference (Two-sample Kolmogorov-Smirnov, p<e-3, n_AMY=203).

Online Methods

Animals and surgical procedures

Two male macaca fascicularis (4–8 kg) were implanted with a round recording chamber above the amygdala and ACC covering both regions in both hemispheres. All procedures were approved and conducted in accordance with the regulations of the Weizmann Institute Animal Care and Use Committee (IACUC), following NIH regulations and with AAALAC accreditation.

MRI based electrode positioning scans were acquired twice, on a 3-Tesla MRI scanner (MAGNETOM Trio, Siemens) with a CP knee coil (Siemens) and using 0.53mm resolution. A first scan before surgery was used to align and refine anatomical maps for individual animals (relative location of the amygdala, ACC and anatomical markers such as the interaural line and the anterior commissure), and to guide the positioning of the chamber on the skull. After surgery, we performed scans with deep electrodes directed toward the amygdala and the ACC (see for example Fig.2a), and calculate the anatomical anterior–posterior and lateral-medial borders relative to the penetrations. The depth of the amygdala is calculated from the dura surface.

Electrophysiology recordings

Each day, 1-3 multichannel (16 contacts each) microelectrodes vector arrays (NeuroNexus) were lowered into the brain using an electrode-positioning-system (NAN, Israel). Vectors were moved independently into the amygdala and ACC while identifying electro-physiological markers tracking the known anatomical pathway. We allowed 30 min for the tissue and signal to stabilize before starting acquisition and behavioral protocol. Data is pre-amplified and stored at 22Khz for later processing. In real-time a 0.3Hz-6KHz band-pass filter and on-line spike sorting was performed using a template-based algorithm (Alpha Lab Pro, Alpha Omega). Off-line spike sorting was performed on the raw data for all sessions to improve unit isolation (offline sorter, Plexon Inc).

Behavioral Paradigms

Fast LCD shutter (307 × 407 mm) is placed between the monkey and the intruder (FOS-307 × 406-PSCT-LV; Liquid Crystal Technologies) to block visual site. Direct current (48v) through the LCD shutter turns it clear/transparent with an onset/offset rise time of <1ms. To enhance precision for neural activity we placed a photodiode (BPX65 Silicon PIN Photodiode) that can be detected with onset/offset of <1e-4ms. There are three types of blocks in each daily session: Human intruder; Airpuffs; Liquid rewards. The blocks are randomized along a session, with more than 120 seconds separating blocks (Fig.1).

Human Intruder: Each block includes 6*3 shutter openings, in which the human intruder alters between Eye-Contact (EC) and No-Eye-contact (NEC) in a pseudorandom order. In both EC and NEC the human maintains gaze direction for 6-9 secs independently of the monkeys' behavior. We generated a per-day pre-defined sequence of EC and NEC with 3 options of sequences that alter across sessions: seq1 (BlockA: EC,NEC,EC,NEC,EC,NEC; BlockB: EC,EC,NEC,NEC,EC,EC; BlockC: NEC,NEC,EC,EC,NEC,EC); seq2 (BlockB, BlockC, BlockA) and seq3 (BlockC, BlockA, BlockB). This was aimed to randomize and prevent learning of EC/NEC order, but also to provide across-days statistics for neural recordings. The human intruder face was filmed and all the trials were monitored to validate that the intruders indeed maintained constant gaze and followed the daily sequence.

Reward: Each block contains 10 trials with an inter-trial-interval of a pseudorandom 20-40 secs. In each trial the shutter opening serves as the conditioned-stimulus (CS) and was followed after 1sec delay by few drops of juice delivered to the monkey's mouth.

Airpuff - Each block contains 8 trials with an inter-trial-interval of a pseudorandom 20-40 secs. In each trial the shutter opening serves as the conditioned-stimulus (CS) and was followed after 1sec delay by air puff (5-15 Psi; located 5 cm from the face).

The monkeys had information about which block is about to start as the human intruder paradigm starts with 5 secs of pure sinus wave (300Hz) followed by the human intruder entering the room and sitting in front of the monkey, with closed shutter. The monkey could not see any part of the human unless the shutter is open.

Behavioral analysis

Eye tracking

A stationary monocular eye tracker was installed for the purpose of eye tracking and gaze estimation. The system included two cameras (Ximea_MQ013RG) – one for eye capturing of the monkey and one for intruders' monitoring and an infrared LED light bar (MetaBright Exolight ISO-14-IRN-24) for face illumination and corneal reflection (CR) production. The eye-recording camera efficiently captured the CR due to its near IR (infra-red) property.

Software implementation was based on the open source project 'OpenEyes' [35], which allows the estimation of subject's point of gaze (POG) on the field of view (FOV) projection. In our case, the FOV scene images were extracted from the video stream of the intruder monitoring camera. The 'OpenEyes' framework makes use of the Starburst algorithm [36] for finding the pupil contour, and assesses the POG by the means of pupil center and CR method . The conditions of our experimental setup (brightly lighted room, large CR of near-rectangular shape and brown sclera of the subject) required a slight modification of the original algorithm for pupil and CR detection. In our variation of the software, the shot noise reduction was skipped, and the CR wasn't removed from the image after its detection, due to its large size. To find the pupil center, we extended the Starburst algorithm. After finding the features candidates for pupil contour, instead of fitting ellipse using RANSAC (random sample consensus) paradigm, we used the "imfindcircles" Matlab function, which searches for circle-candidates applying Hough transform based algorithm. To generate the input for the function, edges image was produced by gradient magnitude calculation followed by binarization. This procedure resulted in a black image with white edges, and was passed to "imfindcircles" with object polarity parameter set to "dark" (specifying that the object – the pupil - is darker than its background). The function returns a list of candidate circles, ordered by circle strengths. Starting from the circle with the biggest strength, the list is searched for the first circle containing a predefined number of minimum feature points that were extracted by the Srtarburst algorithm. Finally, the pupil center is estimated by the center of the found circle. A standard calibration procedure was performed, whereby the monkeys sequentially fixated on 3X3 known grid points in the scene image (according to the original openEyes implementation). To cause the subject's fixation, the screen with the shutter closed, was consecutively illuminated by a laser pointer in the 9 locations. The exact frames of subject's fixation were detected and coordinated with the illumination timings (each time the laser is activated, it records the exact time in the system). The human intruders were filmed throughout all the interactions with the monkeys, and their faces and eyes were marked both

automatically and manually for validation. The 9 (3X3) fixation points were filmed by the same camera, allowing the projections of the fixation points and the intruders on the same plane. Each frame from the eyes of the monkeys therefore result in a point (x and y position) on this plane, allowing to calculate the gaze of the monkey in one of the four ROI's – eyes of the intruders, face of the intruders, shutter region and all the rest.

Facial expression

One Ximea_MQ013RG camera filmed the face region of the recoded monkey in 34Hz. For every recording session, the mean image during the 'alone' period was calculated (i.e. when the monkey was alone in the room with closed shutter). This mean image (See Extended.Fig.2) was subtracted from every frame taken during the Human Intruder interactions. Root Mean Square (RMS) of all the pixels in this subtracted frame is then calculated and the mean and SEM are presented for EC and NEC trials (Fig.1i). Additionally, each day we manually define 3 ROI's – upper face, lower face and ears (See Extended.Fig.2). The same analysis is repeated separately to each ROI and differences between EC and NEC were validated across both for upper and lower face (Fig.1i and Extended.Fig.2)

Heart Rate and Respiratory rate measures

Piezoelectric pulse transducer*:* The cardiac and respiratory traces (for measure of Heart-rate, Heart-rate-variability and Respiratory-rate) [37] were recorded using a piezoelectric pulse transducer (UFI, model 1010) in 2790Hz. We use an elastic belt about 23cm (9 inches) long and fasten extender belt to one end of transducer package using VELCRO™ closures all wrapped around the monkey's chest. We use a piezoelectric pulse transducer (UFI, model 1010) glued around the center allowing direct sensing the heart pulse.

For validation, the respiratory trace is recorded also using solid-state transducer which measures changes in chest or abdominal circumference due to respiration (UFI, model 1132) at 2790Hz. The signal from the piezo sensor also provides respiratory rate parameters, allowing two independent measures for comparison and calibration of parameters.

The piezo-electric signal was processed using a custom made Matlab software. A respiratory trace was extracted using a first order Butterworth filter, and smoothed with running windows. Respiratory peaks were then extracted using 'findpeaks' function. A cardiac trace was extracted by subtracting the filtered respiratory signal from the raw piezo-electric signal. The resulting signal was then processed for each day separately, using filtering and findpeaks parameters. The parameters of the day-specific processing were derived by comparing different sets of parameters to manually tagged cardiac peaks from each day. The resulting day-tailored processed signal was validated using manual inspection of all trials. In addition, the quality of each trial was manually rated, and noisy signal epochs were marked to validate that the result is not due to trials of insufficient quality.

Respiratory rate and heart rate measurements were calculated for each trial using a sliding window of 1 second and heart rate variability (HRV) using running window of 5 seconds, yielding a continuous signal for further analysis. The HRV measure is the standard deviation of normal-normal beat interval (SDNN), a well-established and frequently used measure [38]. Finally, we normalized the changes in each measure by subtracting the mean value from the closed shutter epoch before each trial, to obtain evoked responses.

Vocalizations

Vocalizations were recorded using a microphone (PGA81, Cardioid Condenser Instrument Microphone), situated in close proximity to the monkey. The signal was processed using custom made Matlab software implementing a first order Butterworth filter and smoothed with running a window. Threshold detection was implemented after subtracting the background noise. Several thresholds were tested (1 STD, 2STD, 3STD, 4STD) and the conclusions remain similar.

Movement detection

Two accelerometers were used in the experiment (EVAL-ADXL335Z, Analog Devices), one was attached to the monkeys' chair and one to the setup itself. This allowed to differentiate acceleration caused by self-motor movements from other environmental noise. Movements were recorded in 2790Hz and processed using a custom made Matlab software implementing a first order Butterworth filter and smoothed with a running window. Peaks were then extracted using 'findpeaks' function.

Comparing conditions

We implemented a control based on the 'thinning method', traditionally used to compare distributions from different sources. Here, we compared the distribution of facial expressions or eye-gaze in EC vs. NEC trials. We created similar distributions of facial expressions (eye-movements) for EC and NEC trials, and repeated the main analysis.

Neural activity analysis

Single neuron analysis

The analysis of the neural data focused on three time epochs. In the human intruder blocks, we focused on 400-700ms after shutter opening. This time was chosen because of the oculomotor behavior of the monkeys (Fig.1) showing that the first time they can identify whether this is an EC or an NEC trial has an interquartile range of 180-700ms (see Fig.1d for the full CDF). All analyses were repeated (see Extended.Fig.1, Extended.Fig.5) also when aligning each trial according to the actual time in that trial that the EC/NEC information is available (first gaze to eyes ROI). Such an alignment was done in order to focus on the differences between EC and NEC of the intruder and because fixation shapes neural activity [15]. In the affective (reward/aversive) conditioning blocks, the neural data was taken from 0-300ms after the conditioned-stimulus, termed CS-related activity; and from 0-300ms after reward/airpuff delivery (outcome), termed US-related activity.

Neural activity is normalized according to the baseline activity before the relevant block, using the same window length (300ms) to calculate the mean and standard deviation of the firing rate.
Therefore, the normalized (z-scored) firing rate is:

$$FR_{Normalized} = \frac{FR - mean(baseline)}{std(baseline)}$$

These z-scores were used to quantify the percentage of responsive neurons to the different stimuli. T-tests are used to compare valence (airpuff to reward) or gaze (EC to NEC), and chi-square or binomial tests are used to compare proportions of neurons.

Population decoding

Pseudo-simultaneous population response vector is used for the decoding analysis. The same procedure as reported in details elsewhere [39] is used. The population vector contains spike counts of each neuron in a specific time bin. Each brain area has its own vectors, and the number of vectors is defined by the number of available trials:

$$\overrightarrow{PV\ (t)} =< Neuron_1^C, Neuron_2^C, \dots, Neuron_N^C >$$

$\overrightarrow{PV(t)}$ is the response vector of a specific trial in condition C, in time bin (t), in a brain region that has N neurons. We use the same number of neurons in the amygdala and ACC, therefore we randomly discarded excess neurons in the ACC, resulting in 203 neurons in both.

There are four conditions, airpuff and reward that belong to the valence class and EC and NEC that belong to the gaze class. In the analysis that was conducted in Fig.2 we trained and tested within the same class, whereas in all other analyses we trained on one class and tested on the other class. If we change the order in the training, such that training for NEC yield airpuff and training for EC yield reward, the decoding accuracy is exactly (100-CorectDecoding, see Extended.Fig.3). For both the training and testing we used linear classifier based on maximization procedure of the SVM algorithm (fitSVM Matlab function). Each training set yields a boundary line (set of weights for every neuron) and a threshold that separates the two conditions under consideration. The same output from the training was then used to assess the accuracy in the test set.

For a given neuron and a given condition we used 80% of the trials for training and 20% for testing when done within the same class. When we trained on one class and tested on the other, we used all the available trials for training and testing. The accuracy of every decoder was estimated by pseudorandom resampling from the available trials 1,000 times.

In the analysis of Fig.4 we shuffled the neurons such that the index of each neuron in $\overrightarrow{PV}$ is randomly assigned. Therefore, the spike count of every neuron remains, but its position in the vector changes.

Decision boundary analysis

In order to estimate if the mechanism that allows decoding of one class based on the other is due to *correlated-selectivity* or *overall-activity*, we estimate the angle between the boundary lines. Every training sample yields a vector of weights:

$$\overrightarrow{Boundary_{Class}} =< W_1, W_2, \dots, W_N >$$

$\overrightarrow{Boundary_{Class}}$ is the decision boundary of one training sample in a brain region with N neurons. Every brain region has two boundaries, one for gaze and one for valence.

$$\cos\alpha = \frac{\overrightarrow{Boundary_{Valence}} \cdot \overrightarrow{Boundary_{Gaze}}}{|\overrightarrow{Boundary_{Valence}}| * |\overrightarrow{Boundary_{Gaze}}|}$$

Each of the boundaries is sampled 1,000 times to obtain a distribution of angles. The results are presented as $\cos(\alpha)$ and not α, so zero (0) values represent perpendicular boundaries.

Linear regression analysis

We estimated the tuning of the neurons to valence and gaze by linear regression analysis. The firing rate, FR, of every neuron is fitted during every time bin with one of the following equations:

$$FR_{Valence} = \beta_V^0 + \beta_V \cdot Valence$$

$$FR_{Gaze} = \beta_G^0 + \beta_G \cdot Gaze$$

$Valence$ is 1 for airpuff trials and -1 for reward trials, whereas $Gaze$ is 1 for EC and -1 for NEC. The regression analysis yield for every neuron two coefficients, β_V and β_G.

Scalar product of linear regression coefficients

We calculated the scalar product between $\vec{\beta}_{gaze}$ and $\vec{\beta}_{valence}$ where the vector sign indicates that it is a vector of all neurons in a certain brain region $\vec{\beta}_{gaze} =< \beta_{G_1}, \beta_{G_2}, \dots, \beta_{G_N} >$ and $\vec{\beta}_{valence} =< \beta_{V_1}, \beta_{V_2}, \dots, \beta_{V_N} >$. The intuition behind this scalar product is that if more neurons response in a similar direction, then the scalar product is expected to be positive and vice versa.

$$\vec{\beta}_{gaze} \cdot \vec{\beta}_{valence} = \sum_{i=1}^{N} (\beta_{G_i} \beta_{V_i})$$

We also calculated a shuffled version where a random index is used, and hence the multiplication of the coefficients is done across two different neurons. The shuffled scalar product is repeated 1,000 times.

Selectivity-Index

We calculated a selectivity index for each neuron in the amygdala and ACC for gaze (SIG) and for valence (SIV) in the following way:

$$SIG = \frac{FR_{Normalized_{EC}} - FR_{Normalized_{NEC}}}{|FR_{Normalized} FR_{EC}| + |FR_{Normalized} FR_{NEC}|}$$

$$SIV = \frac{FR_{Normalized_{PUFF}} - FR_{Normalized_{REWARD}}}{|FR_{Normalized} FR_{PUFF}| + |FR_{Normalized} FR_{REWARD}|}$$

We tested both the values of SIV and SIG separately, as well as the overlap between the two, and whether the selectivity is in the same direction ($SIG * SIV > 0$).

Extended Figures legends

Extended.Fig.1. Differential behavioral response to EC and NEC

a. Same format as Fig.1f but for all shutter ROI (and not just face ROI). As can be seen, the monkeys look at the face and eyes ROI mainly in the human intruder interactions. Left – the gaze density during all the sessions.
b. Same format as Fig.1g, but aligned to the first time the monkeys looked to the intruder's eyes ROI in each trial separately.
c. Same format as Fig.1g-right, separately for each monkey

Extended.Fig.2. Extracting differences in facial expression

a. Examples of three original frames with different expressions, corresponding to the scheme in Fig.1h.
b. For every recording session, we averaged over all frames from the baseline period resulting in the mean image (baseline was taken over the period before any trial when the monkey was alone in the room with a closed shutter).
c. An example of a frame during EC (eye contact) interaction.
d. The mean frame (b) is subtracted from the frame in (C) during the interaction, to obtain a 'diff'/delta image. Three ROIs are defined manually for every day – Upper, Ears and Lower.
e. Root Mean Square of every ROI is calculated (Mean +/- SEM). Shown are differences between EC and NEC in the Upper part (see main Fig.1 for other parts/ROI's). Upper black line represents a significant difference ($p<0.05$, t-test two-sided, n-trial=1480/1628 in NEC/EC).

Extended.Fig.3. Reversing valence directionality (NEC-EC to aversive-appetitive)

a. Same format as in main Fig.3a,b. Population decoding accuracy (Mean +/- STD, bootstrap CI=95%, n_rep=1000, n_AMY=203) but when training on eye-gaze (NEC vs. EC) and testing on valence (aversive vs. appetitive), using CS-related activity.
b. Same as (a) but using US-related activity.

Extended.Fig.4. Single-neurons activity across conditions.

a. If *overall-activity* drives the successful decoding in the US epoch, we expect to find an overall change in the firing rate (increase or decrease) for gaze and for US valence. Indeed, we find that there are more valence positive neurons (increased firing rate to airpuff) in the amygdala in the US epoch, and that there are more gaze positive neurons (increased firing rate to EC) in the amygdala. Inset represent the mean and SEM, *** represent a significant differences in Z-test, $p<e-3$, n_AMY=203 and n_ACC=356).
b. Decoding accuracy with and without neurons that code for gaze (n_AMY=203 and n_Rep=1000). Black and red lines represents the mean and median respectively.

Extended.Fig.5. Decoding with trial-based alignment to shutter opening.

a. Same format as in Fig.4h,i,j. Population decoding accuracy for real and shuffled amygdala neurons (n_AMY=203). Black and red lines represents the mean and median respectively.
b. Same as (a) for ACC activity (n_ACC=356).
c. Cumulative-distribution of the difference in decoding accuracy between real and shuffled neurons. *** represents a significant difference (Two-sample Kolmogorov-Smirnov, p<e-3, n_AMY=203).

Extended.Fig.6. Behavioral differences between EC and NEC do not underlie neural findings

a. An example of vocalizations during one trial of human intruder, measured using a microphone placed in close proximity to the monkey (see methods). Inset – the proportion of trials in which vocalizations occur. Notice there is a very small proportion of trials in which vocalization occur, and it was similar across EC and NEC trials ($\chi 2$, p=0.88, n=1738/1807 in NEC/EC). Due to the low number of vocalizations, we were not able to characterize different types of vocalizations. In addition, we repeated analyses after removing trials during which vocalizations occur, and the main results were unchanged.
b. An example of movement in one trial in response to the human intruder, measured using an accelerometer attached to the chair of the monkey (see methods). Here as well there is a very small proportion of trials, and it was similar across EC and NEC trials. In addition, we repeated analyses after removing these trials, and the main results were unchanged.
c. The overall change in facial expressions (Mean +/- SEM) between EC and NEC (as in Fig.1i). Shown is the Root-Mean-Square (RMS) of the change between the image over the whole face (main) and only for the lower half of the face (inset), compared to the neutral expression obtained from averaging over baseline period when the monkey was alone (see methods). There is a significant difference (t-test, two-sided, p<0.05, n-trials 1703/1765 in NEC/EC).
d. Same as in (c) but after applying the 'thinning method' (iteratively selecting trials to obtain a similar distribution of behavior across EC and NEC; see methods). We applied the same also for eye-movements.
e. Decoding accuracy using only trials with similar behavior across EC and NEC, taken after the 'thinning' as shown in (d). Results remain the same (compare to Fig.4h). Violin – red for median and black for mean (n_AMY=203,n_REP=1000).

Extended.Fig.7. Consistency across stimulus saliency (no within-day adaptation)

Decoding accuracy divided into first and second half of trials; Similar results are obtained.
The presentation is a merged format of Fig.4h and Fig.4i.
Using CS-related activity (a,b) or US-related activity (c,d) in the first-half of trials (a,c) and second-half (b,d). In the violin diagrams red represents the nedian and black the mean. n_AMY=203, n_ACC=356 and n_Rep=1000.

Extended.Fig.8. Neuronal modulation.

a. We divided the amygdala neurons into three groups: the first contains neurons that increase their firing rate (FR) to gaze and valence (61/203, positive betas in Fig.4a); the second group decrease FR to both gaze and valence (65/203, negative betas in Fig.4a); and the third group increase FR to one condition and decrease to the other (77/203).
   For the first two groups, the decoding accuracy of valence based on gaze (similar analysis as in Fig.4h for CS-related activity) was significantly higher than chance, indicating that the overall result reported in the main text is based on both increases and decreases in FR.
   Right: same but for ACC neurons.
   In the violin diagrams red represents the median and black the mean.
b. Amygdala neurons were sorted according to degree of modulation (magnitude of beta_gaze*beta_valence; red line), decoding accuracy (mean) and its variance for increasing group size (namely, 10 with highest modulation, 20 …, and so forth) was re-calculated . This is compared to randomly choosing groups of similar size (green inset, notice the linear increase).
   The decoding accuracy increases until reaching a group size of 120-130 neurons (see dashed line), namely the number of neurons that contain the first two groups from (a) - only increasing or only decreasing FR (but not mixed).
   The bottom part shows the proportion of neurons from the two groups. It can be seen that both groups contribute to the increased accuracy.
   These results further support the conclusion that the shared neural mechanisms are not due only to increased firing rate, as an indication of saliency or alertness.
   n_AMY=203.

Extended.Fig.9. Neurons code for species, but it is not shared with valence of CS

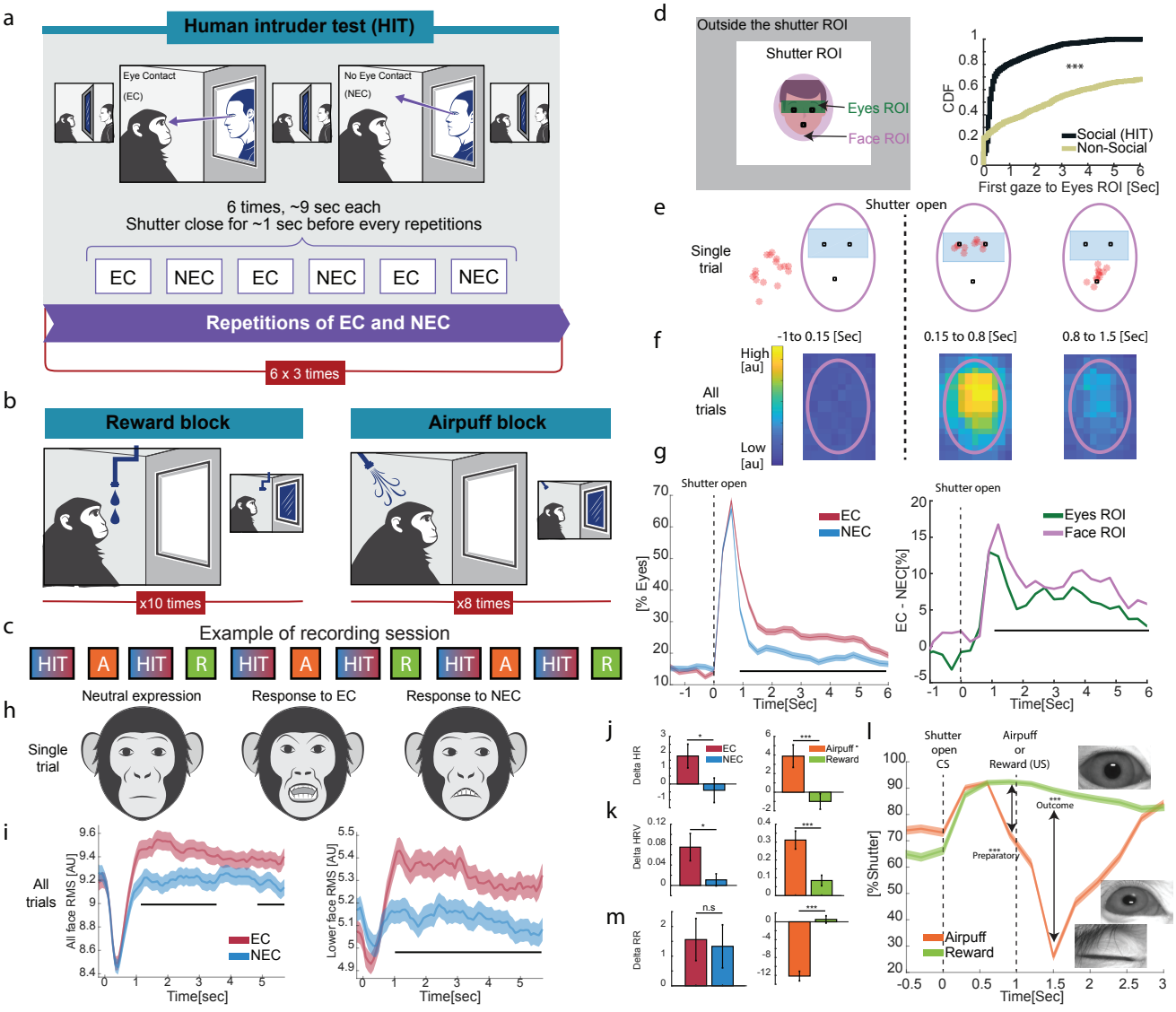
a. We introduced monkey-intruder blocks (top) in a similar way to the human-intruder trials (bottom). The same neurons reported in the main analysis were recorded during the monkey-monkey interaction as well. Each recording session, on average two (out of 6) monkeys served as intruders. All the monkeys lived together for several years.
b. Neurons in the amygdala (n=203), as well as in the ACC (n=356), code for species, namely differentiate human- from monkey-intruder (mean and SEM). Moreover, neurons differentiate between NEC-human and monkey-intruder.
c. In contrast to the findings in Fig.4a, there is no significant correlation (Pearson's correlation, r=0.05,p=0.45, n=203) between beta_species and beta_CS_valence, strongly arguing against a *correlated-selectivity* mechanism between species and CS.
d. Decoding accuracy of CS-valence (n_AMY=203 and n_REP=1000) after training the decoder to differentiate species, is not different than chance-level and significantly smaller than the decoding accuracy of CS-valence based on gaze. In the violin red/black represents the median/mean.
e. Differences in Heart Rate Variability (HRV) between monkey and NEC trials (as between EC and NEC trials, as also shown in Fig.1k). * represents a significant t-test two-sided, $p<0.05$, n-trial=1703/1765/1620 in NEC/EC/Monkey
f. Despite differences in HRV (e), the findings in (d) remain similar when using either only EC or only NEC trials of the human intruder (n_AMY=203). In the violin red/black represents the median/mean.

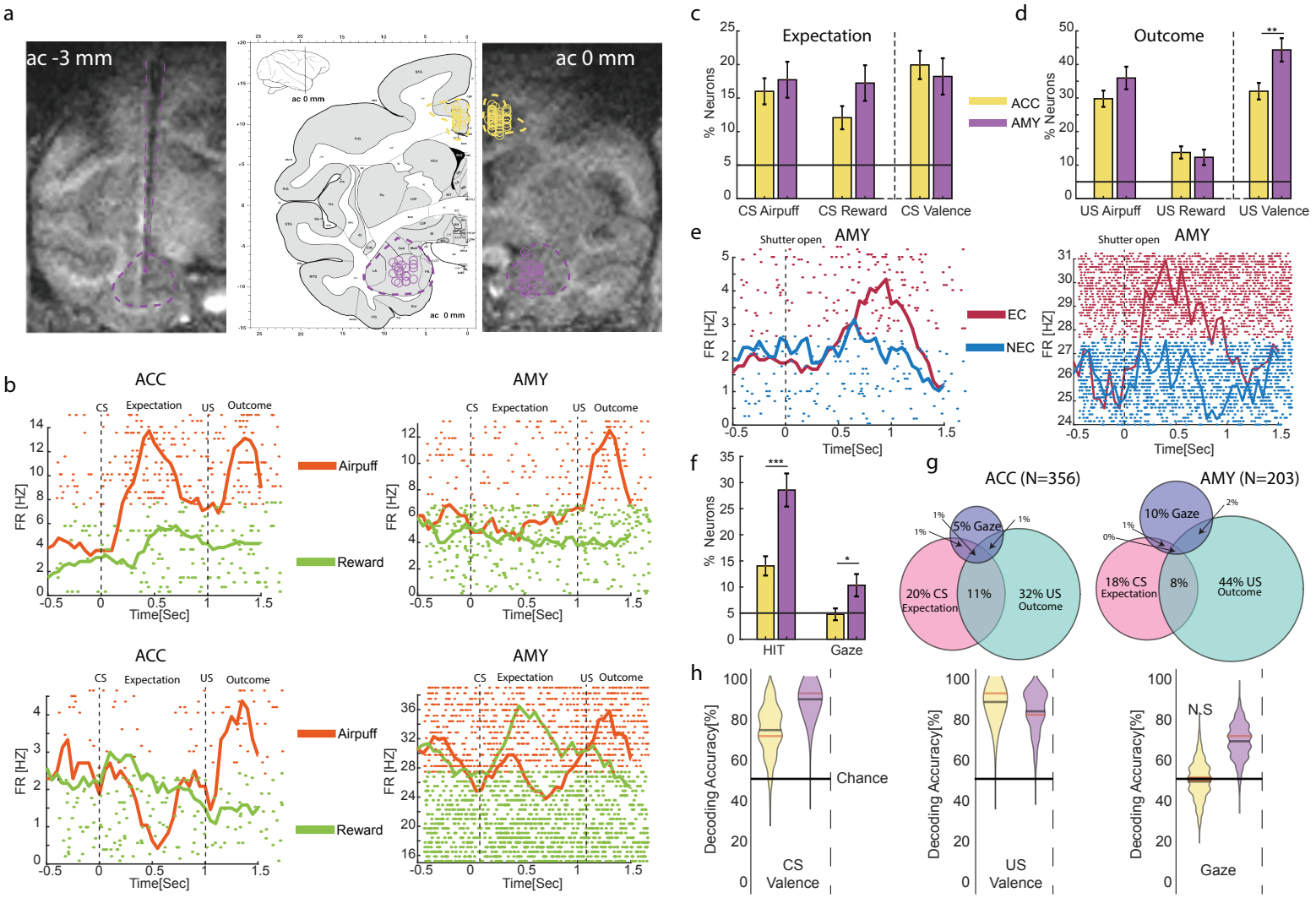
Extended.Fig.10. NEC trials are different than neutral trials

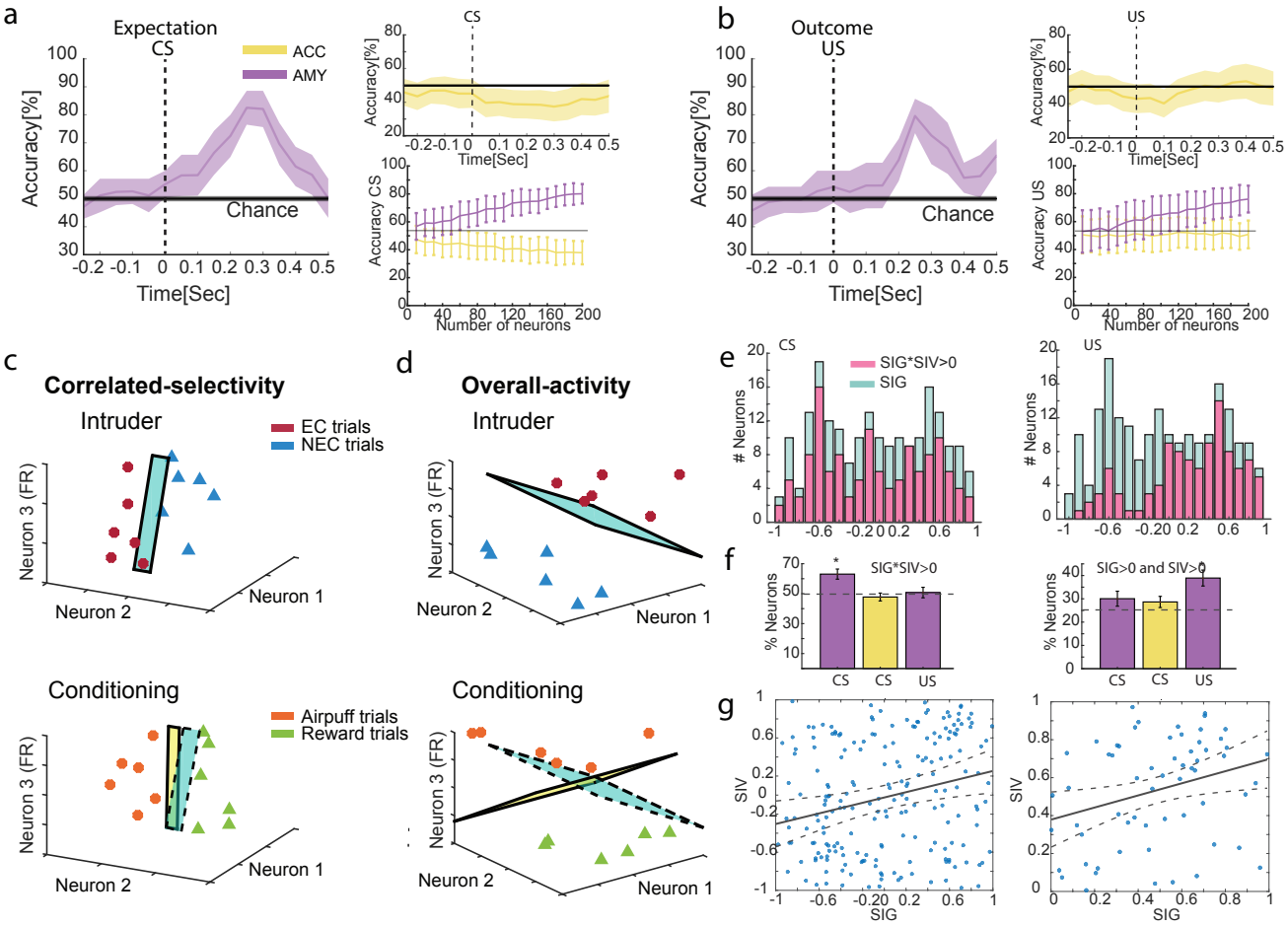
a. We introduced Neutral trials, where a shutter open (CS) is followed by nothing.
b. The Heart Rate is significantly lower in neutral trials compared to all others types, and specifically lower than NEC trials.
   Insets, Left: delta HR, same as in Fig1.j. ; Right: delta HR in the control days that included neutral trials, showing the same trend for all types, and no modulation for neutral trials.

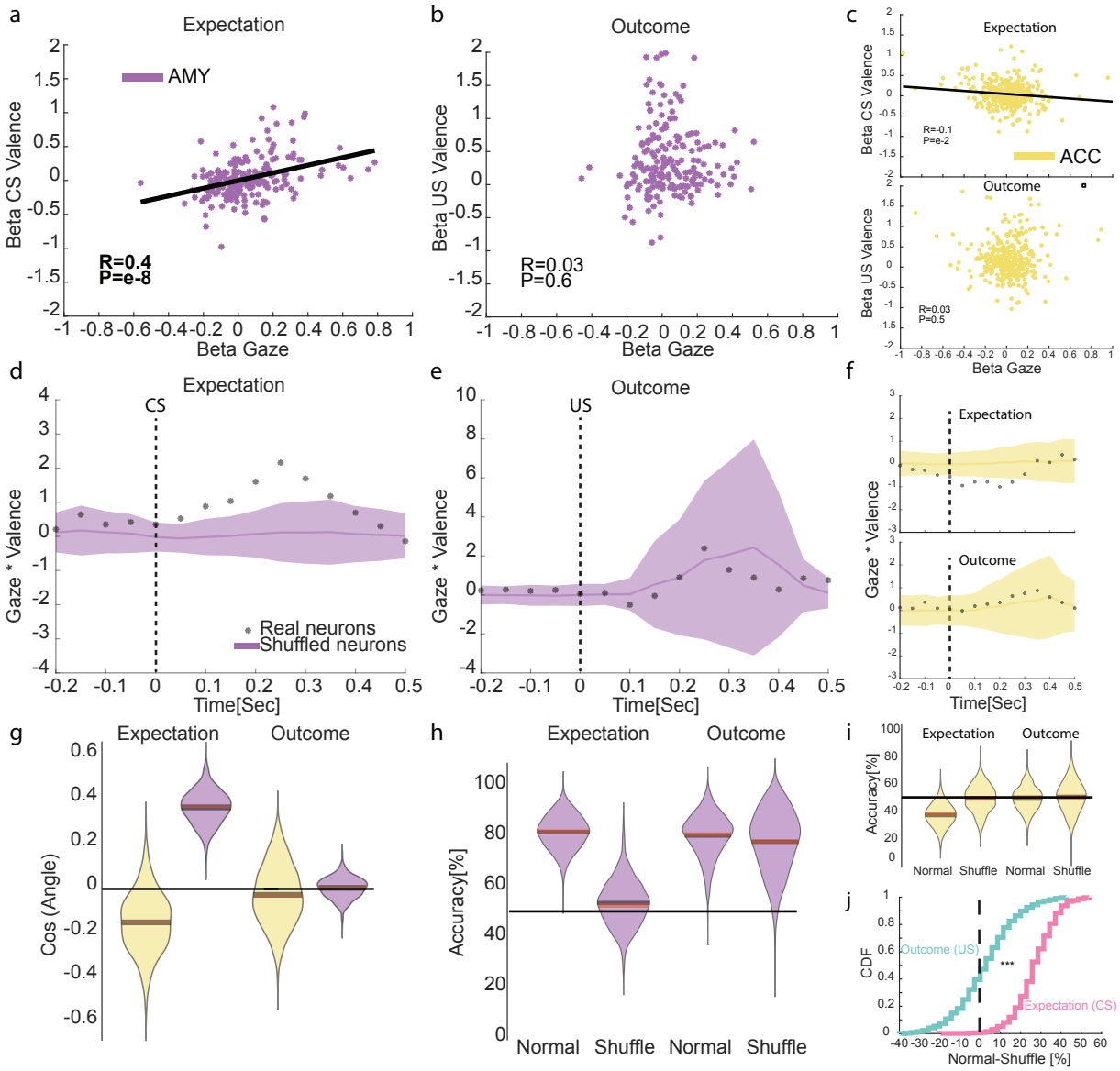
Together, this argues that the NEC trials are not salience-free, but rather highly salient in a different manner than the EC.

n-trial=1703/1765/1620/1352/712 in NEC/EC/Monkey/Reward/Airpuff respectively. ***represents a significant t-test two-sided, p<e-3. Bar plots represents mean and SEM.


a
Human intruder test (HIT)
Eye Contact
(EC)
No Eye Contact
(NEC)
6 times, ~9 sec each
Shutter close for ~1 sec before every repetitions
EC
NEC
EC
NEC
EC
NEC
Repetitions of EC and NEC
6 x 3 times
b
Reward block
Airpuff block
x10 times
x8 times
c
Example of recording session
HIT
A
HIT
R
HIT
A
HIT
R
HIT
A
HIT
R
d
Outside the shutter ROI
Shutter ROI
Eyes ROI
Face ROI
CDF
***
Social (HIT)
Non-Social
First gaze to Eyes ROI [Sec]
e
Shutter open
Single
trial
f
-1to 0.15 [Sec]
0.15 to 0.8 [Sec]
0.8 to 1.5 [Sec]
All
trials
High
[au]
Low
[au]
g
Shutter open
EC
NEC
[% Eyes]
Time[Sec]
Shutter open
Eyes ROI
Face ROI
EC - NEC[%]
Time[Sec]
h
Neutral expression
Response to EC
Response to NEC
Single
trial
i
All
trials
All face RMS [AU]
EC
NEC
Time[sec]
Lower face RMS [AU]
Time[sec]
j
Delta HR
EC
NEC
Airpuff
Reward
k
Delta HRV
m
Delta RR
n.s
l
Shutter
open
CS
Airpuff
or
Reward (US)
Outcome
Preparatory
[%Shutter]
Airpuff
Reward
Time[Sec]


a
ac -3 mm
ac 0 mm
ac 0 mm
ac 0 mm
b
ACC
AMY
CS
Expectation
US
Outcome
FR [HZ]
Time[Sec]
Airpuff
Reward
c
Expectation
% Neurons
ACC
AMY
CS Airpuff
CS Reward
CS Valence
d
Outcome
US Airpuff
US Reward
US Valence
e
Shutter open
AMY
EC
NEC
f
HIT
Gaze
g
ACC (N=356)
AMY (N=203)
5% Gaze
20% CS Expectation
11%
32% US Outcome
10% Gaze
18% CS Expectation
8%
44% US Outcome
h
Decoding Accuracy[%]
Chance
CS Valence
US Valence
N.S
Gaze


a
Expectation
CS
ACC
AMY
Accuracy[%]
Chance
Time[Sec]
CS
Time[Sec]
Accuracy CS
Number of neurons
b
Outcome
US
Accuracy[%]
Chance
Time[Sec]
US
Time[Sec]
Accuracy US
Number of neurons
c
Correlated-selectivity
Intruder
EC trials
NEC trials
Neuron 3 (FR)
Neuron 2
Neuron 1
Conditioning
Airpuff trials
Reward trials
Neuron 3 (FR)
Neuron 2
Neuron 1
d
Overall-activity
Intruder
Neuron 3 (FR)
Neuron 2
Neuron 1
Conditioning
Neuron 3 (FR)
Neuron 2
Neuron 1
e
CS
SIG*SIV>0
SIG
Neurons
US
Neurons
f
SIG*SIV>0
% Neurons
CS
CS
US
SIG>0 and SIV>0
% Neurons
CS
CS
US
g
SIV
SIG
SIV
SIG


a
Expectation
AMY
Beta CS Valence
R=0.4
P=e-8
Beta Gaze
b
Outcome
Beta US Valence
R=0.03
P=0.6
Beta Gaze
c
Expectation
Beta CS Valence
R=-0.1
P=e-2
ACC
Outcome
Beta US Valence
R=0.03
P=0.5
Beta Gaze
d
Expectation
CS
Gaze * Valence
Real neurons
Shuffled neurons
Time[Sec]
e
Outcome
US
Gaze * Valence
Time[Sec]
f
Expectation
Outcome
Gaze * Valence
Time[Sec]
g
Expectation
Outcome
Cos (Angle)
h
Expectation
Outcome
Accuracy[%]
Normal
Shuffle
Normal
Shuffle
i
Expectation
Outcome
Accuracy[%]
Normal
Shuffle
Normal
Shuffle
j
CDF
Outcome (US)
***
Expectation (CS)
Normal-Shuffle [%]