nature neuroscience

Fear generalization in the primate amygdala

Jennifer Resnik & Rony Paz

Broad generalization of negative memories is a potential etiology for anxiety disorders, yet the underlying mechanisms remain unknown. We developed a non-human primate model that replicates behavioral observations in humans and identifies specific changes in tuning properties of amygdala neurons: the width of auditory tuning increases with the distance of its center from the conditioned stimulus. This center-width relationship can account for better detection and at the same time explain the wide stimulus generalization.

We compared behavioral (n = 34+66 sessions in 2 monkeys) and amygdala single-unit responses (n = 107 + 193) to pure tones before and after one of the tones was conditioned (CS) with an aversive odor (US; **Fig. 1a**–**d**). The conditioning to a single stimulus mimics real-life exposure to aversive experiences (that is, without fine discrimination training) and promotes broad generalization, a candidate behavioral model for anxiety[1–4]. We quantified conditioned preparatory responses (CRs) as the increase in inhale volume in response to the tone (**Fig. 1c** and **Supplementary Fig. 1a**) and found wide generalization around the CS (P < 0.005 interaction effect of tones and phase, F = 2.82, degrees of freedom (df) = 8, two-way ANOVA; **Fig. 1e**,**f** and **Supplementary Fig. 1b**,**c**) when compared with tones that were paired with an appetitive odor (P < 0.05, df = 547, t test; **Fig. 1g** and **Supplementary Fig. 1d**) and resembling human thresholds following aversive conditioning (**Fig. 1h** and **Supplementary Figs. 1e** and **2**)[5,6].

Tuning curves of single cells in the basolateral amygdala (BLA; (**Supplementary Fig. 2**) exhibited a specific characteristic after conditioning: they were narrow when the preferred stimulus (PS, the tuning curve's mean) was close to the CS, but wide when it was farther away (**Fig. 2a**). To quantify this center-width relationship across the population, we binned neurons according to the distance of their post-conditioning PS from the CS, revealing a gradual increase in width (P < 0.0001, F = 18.33, df = 3, ANOVA; **Fig. 2b**,**c**) that was not observed

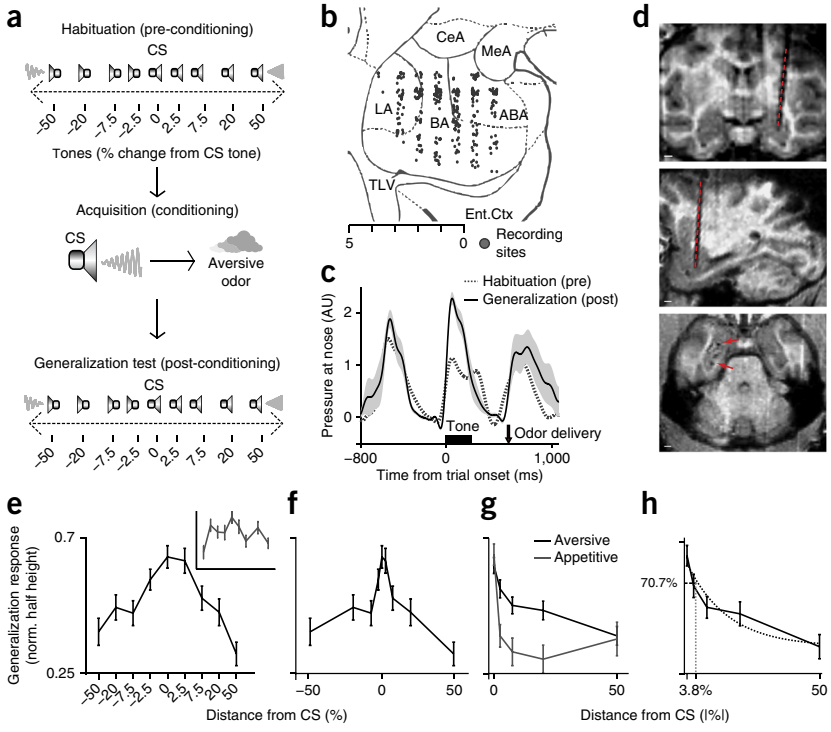

Figure 1 Behavioral generalization. (**a**) Nine pure tones were presented six times in a pseudorandom order (habituation/pre-conditioning), followed by partial conditioning of the center tone to a highly aversive odor (conditioning), and then the nine tones were presented again (generalization test/post-conditioning). A new set of tones was selected each day, with equal distances from the middle tone (–50%, –20, –7.5, –2.5, +7.5, +20, +50%). (**b**) Reconstruction of recording locations in the monkey BLA complex. (**c**) The CR is the change in the inhale that triggers the CS tone (that is, before the inhale that triggers odor release) compared with the habituation phase. (**d**) Structural magnetic resonance imaging with calibrating electrodes in the amygdala. Scale bars represent 5 mm. (**e**) Behavioral generalization: CR for the different tones (quantified as the half height time of the inhale in ms and normalized within session, mean ± s.e.m., n = 64). Inset shows the same during habituation. (**f**) Data are presented as in **e** with real x axis spacing. (**g**) Absolute generalization curves for aversive (black) and appetitive (gray) odors (mean ± s.e.m.). (**h**) Absolute generalization curve for aversive (solid line, same curve as in **g**), fitted with an exponential function (dotted black line). The dashed gray line indicates the equivalent of just noticeable difference (JND) computed in human psychophysics (equivalent to 70.7% correct choice). The JND here was 3.8%, similar to humans subjected to the same conditioning procedure[5].

Department of Neurobiology, Weizmann Institute of Science, Rehovot, Israel. Correspondence should be addressed to R.P. (rony.paz@weizmann.ac.il).

Received 6 August; accepted 16 November; published online 22 December 2014; doi:10.1038/nn.3900

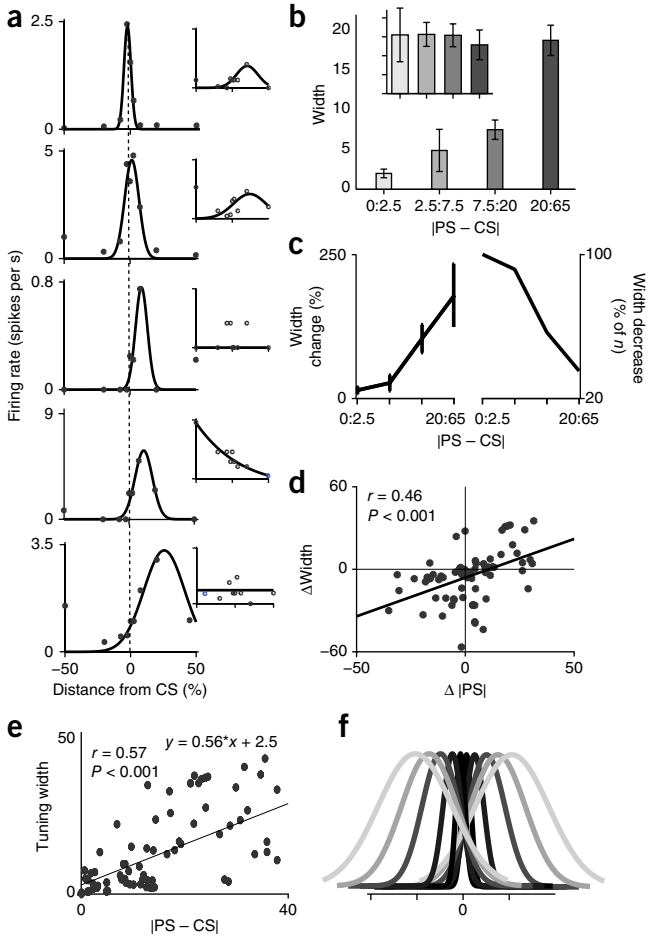

Figure 2 Tuning width changes with distance from the CS. (**a**) Single cells' tuning curves with Gaussian fit in the generalization test (post-conditioning; insets show the same cell during pre-conditioning). The further away the PS (the peak mean of the Gaussian) from the CS (the dashed line), the wider the tuning. (**b**) Tuning during generalization was wider with PS distance from the CS (mean ± s.e.m., ANOVA, $F = 4.65$, $P = 0.003$, df = 3). Inset shows tuning width taken from the habituation pre-conditioning ($P > 0.1$). (**c**) Percent change in width (left, mean ± s.e.m.) and proportion of neurons that reduced width (right), as a function of PS distance from CS. (**d**) Width change (post-pre) versus change in absolute PS ($r^2 = 0.21$, $P < 0.001$). (**e**) Linear regression between tuning width and PS distance from CS ($n = 100$, $P < 0.001$). (**f**) Illustration of the relationship between the PS and the tuning width.

pre-conditioning ($P > 0.9$, $F = 0.11$, df = 3, ANOVA; **Fig. 2b** and **Supplementary Fig. 3**). The width change was correlated with the individual shift in PS from pre- to post-conditioning ($P < 0.001$, $r^2 = 0.21$, Pearson coefficient; **Fig. 2d**), indicating that the tuning curves of neurons that shifted their PS toward the CS became narrower and that those that shifted their PS away from the CS became wider. There were no significant changes in other response properties ($P > 0.1$; **Supplementary Figs. 4–6**). When we quantified basic response properties using a traditional auditory procedure (4 octaves, 300–4,800 Hz, 40 evenly logarithmically spaced tones), we found typical one-peak-wide tuning (**Supplementary Fig. 7**), indicating that BLA neurons are broadly tuned in normal conditions. Notably, the center-width relationship was maintained 24 h after learning (**Supplementary Fig. 8**), suggesting that some BLA neurons hold a 24-h representation, which could underlie perceptual memory in humans[5].

To model this population characteristic of width change as a function of PS distance from the CS (center-width relationship), we fitted a linear regression to all neurons ($P < 0.001$, $r^2 = 0.32$; **Fig. 2e,f** and **Supplementary Fig. 9**) and used the coefficients slope and y intercept as parametric description of the population. In the experimental data, the width of behavioral generalization in individual sessions was indeed correlated with the slope ($P < 0.001$, df = 3, ANOVA, $F = 22.1$; **Fig. 3a**). Furthermore, this center-width relationship also accounted for the common finding that the BLA responds more robustly to the CS[7], as more neurons include the CS in their tuning curve and therefore generate more spikes (**Fig. 2f** and **Supplementary Fig. 10**). It can therefore support better CS detection.

We next used the parametric description of the population to generate surrogate data and test two complementary approaches: the first examines discrimination between two stimuli based on

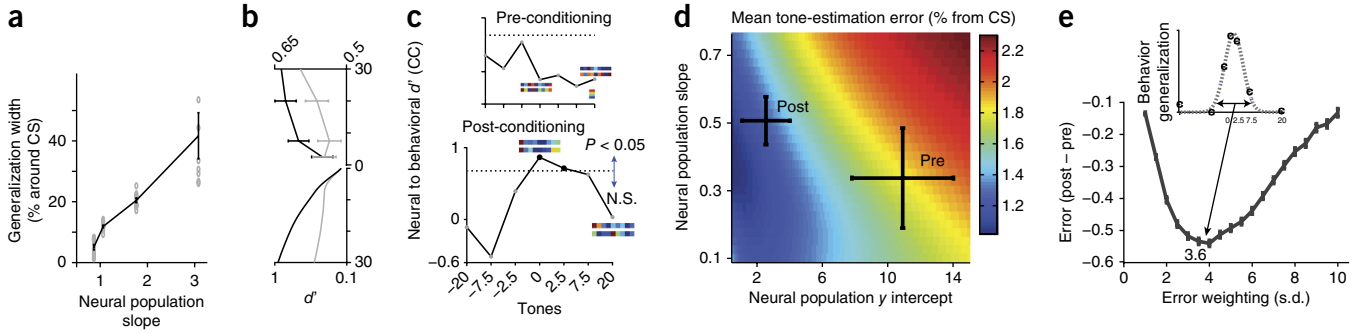

Figure 3 Neural organization match behavioral generalization. (**a**) Sessions were binned into quartiles according to the behavioral generalization width (the s.d. of the Gaussian fit to behavior in the session). All neurons recorded in the sessions from each quartile were pooled and the population slope was computed (similarly to **Fig. 2e**), and then plotted (*x* axis) against the generalization width in each session (*y* axis). (**b**) *d′* between the CS and other tones, averaged over all neurons (top: experimental data; bottom: surrogate data), during habituation (gray) and generalization-test (black). (**c**) Correlation between behavioral *d′* and neural *d′* per tone, for pre- (upper) and post-conditioning (lower). The *d′* between a tone and all other eight tones was based on behavioral responses or on neural responses (and averaged over sessions), resulting in two vectors (examples shown) that were correlated. N.S., not significant. (**d**) Mean error for tone estimation with optimal linear readout of neural activity. Error surface as a function of possible slopes and *y* intercepts created of the population relationship between width and PS distance from CS (as in **Fig. 2e**). The mean estimation error is represented by the color scale in percentage change from CS. Black crosses mark the experimental data from pre- and post-conditioning (mean ± s.e., bootstrap from **Fig. 2e**), showing the decrease in error after conditioning (pre: 1.84, post: 1.31, $P < 0.001$, *t* test). (**e**) Change in mean error (post – pre, as in **d**) as a function of different error surfaces created by different Gaussian weighting of the error across tones, that is, the importance of each tone is weighted by its distance from the CS. Inset shows the actual behavioral points (from **Fig. 1e,f**) fitted with a Gaussian ($\sigma = 3.64$, $r^2 = 0.9841$, $P < 0.001$). All error bars represent s.e.m.

neural responses and the second examines stimulus estimation via readout of neural activity. For discrimination, we calculated a sensitivity measure between all tones (d' between all combinations of two tones based on each neuron activity). There was an improvement in neural d' when comparing the CS with tones that were distant from it, but not when comparing with tones that were close to it (for both experimental and surrogate data; **Fig. 3b**). This was further evidenced when examining post- minus pre-conditioning d' for all stimuli combinations (**Supplementary Fig. 11a**,**b**): there was no improvement in d' in close proximity to the CS (~5%, comparable to human reports[5,6]). The reason for this lies in having more neurons with wide tuning and therefore shallow slope around the CS (**Supplementary Fig. 11c**). In other words, a downstream network that weighs all inputs equally would not improve or even deteriorate on discrimination around the CS, a factor that directly contributes to broader generalization. In addition, correlating behavioral d' to neural d' was significant only during generalization and only around the CS ($P < 0.05$; **Fig. 3c**), further linking amygdala population response to generalization behavior.

A reasonable assumption is that, before comparing stimuli, the brain first estimates the incoming stimulus. We therefore considered an optimal-linear readout solution. In this setup, a downstream network (such as the auditory cortex) reads out amygdalar inputs with the assumption that all stimuli are equally important for the animal, as before one tone is aversively conditioned. We found that the mean estimation error (ME) over all tones improved from the pre-conditioning population to the post-conditioning population (pre, 1.84; post, 1.31; $P < 0.001$, df = 99, t test; **Fig. 3d**). Furthermore, if we assume that error in the estimation of the CS is more important than errors in estimation of other tones (because it signals danger, a highly aversive odor that prevents normal breathing), we can re-calculate the ME across tones, but weighted by a Gaussian with width that represents the relative importance of tones around the CS (in other words, it should reflect the behavioral generalization). We therefore recalculated the improvement in the ME from pre- to post-conditioning as a function of the width of weighing the ME (**Fig. 3e**) and found the width that maximized the improvement. This width closely matched the behavioral generalization (σ that maximizes improvement in neural estimator = 4, σ fit to behavioral generalization = 3.6, $r^2 = 0.9841$, $P < 0.001$; **Fig. 3e**). In other words, the change in amygdala tuning is optimized for the observed behavioral generalization width.

We conclude that specific changes in BLA population code underlie the broad generalization and reduced discrimination seen following aversive conditioning[5,6,8,9], and suggest that such an architecture can potentially underlie the extreme broad generalization observed in anxiety disorders[1–3]. The complete circuit that mediates generalization of more complex real-life stimuli likely involves interactions between the BLA and the medial geniculate thalamus[10], auditory cortex[4,11,12], prefrontal cortex[13–16], hippocampus[17–19] and the periaqueductal gray[20], as well as microcircuits in the amygdala[7]. Our results suggest that even mild changes in BLA circuitry or in its inputs can disrupt the balance and result in exaggerated responses to similar stimuli, as in anxiety.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We are grateful to H. Sompolinsky for discussions and suggestions. We thank Y. Shohat for major contribution to the work and welfare of the animals, E. Kahana and G. Hecht for help with medical and surgical procedures, E. Furman-Haran and N. Stern for magnetic resonance imaging procedures. This work was supported by the Minerva Foundation and the Israel Science Foundation (#26613 and ERC-FP7-StG #281171, R.P.).

AUTHOR CONTRIBUTIONS

J.R. and R.P. designed the research. J.R. performed the experiments. J.R. analyzed the data. J.R. and R.P. wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Pitman, R.K. *et al. Nat. Rev. Neurosci.* **13**, 769–787 (2012).
2. Lissek, S. *Depress. Anxiety* **29**, 257–263 (2012).
3. Jovanovic, T. & Ressler, K.J. *Am. J. Psychiatry* **167**, 648–662 (2010).
4. Aizenberg, M. & Geffen, M.N. *Nat. Neurosci.* **16**, 994–996 (2013).
5. Resnik, J., Sobel, N. & Paz, R. *Nat. Neurosci.* **14**, 791–796 (2011).
6. Laufer, O. & Paz, R. *J. Neurosci.* **32**, 6304–6311 (2012).
7. Duvarci, S. & Pare, D. *Neuron* **82**, 966–980 (2014).
8. Armony, J.L., Servan-Schreiber, D., Romanski, L.M., Cohen, J.D. & LeDoux, J.E. *Cereb. Cortex* **7**, 157–165 (1997).
9. Shaban, H. *et al. Nat. Neurosci.* **9**, 1028–1035 (2006).
10. Han, J.H. *et al. Learn. Mem.* **15**, 443–453 (2008).
11. Quirk, G.J., Armony, J.L. & LeDoux, J.E. *Neuron* **19**, 613–624 (1997).
12. David, S.V., Fritz, J.B. & Shamma, S.A. *Proc. Natl. Acad. Sci. USA* **109**, 2144–2149 (2012).
13. Likhtik, E. & Stujenske, J.M. *Nat. Neurosci.* **17**, 106–113 (2014).
14. Chavez, C.M., McGaugh, J.L. & Weinberger, N.M. *Neurobiol. Learn. Mem.* **91**, 382–392 (2009).
15. Dunsmoor, J.E., Prince, S.E., Murty, V.P., Kragel, P.A. & LaBar, K.S. *Neuroimage* **55**, 1878–1888 (2011).
16. Courtin, J. *et al. Nature* **505**, 92–96 (2014).
17. Xu, W. & Sudhof, T.C. *Science* **339**, 1290–1295 (2013).
18. Bergado-Acosta, J.R. *et al. Learn. Mem.* **15**, 163–171 (2008).
19. Kaouane, N. *et al. Science* **335**, 1510–1513 (2012).
20. Johansen, J.P., Tarpley, J.W., LeDoux, J.E. & Blair, H.T. *Nat. Neurosci.* **13**, 979–986 (2010).

ONLINE METHODS

Animals. Two male *Macaca fascicularis* (4–7 kg, 3–4 years) were implanted with a recording chamber above the amygdala under full anesthesia and aseptic conditions. All surgical and experimental procedures were approved and conducted in accordance with the regulations of the Weizmann Institute Animal Care and Use Committee following NIH regulations and with AAALAC accreditation. Anatomical magnetic resonance imaging (MRI) scans with positioning electrodes were acquired before, during and after the recording period. Images were acquired on a 3-T MRI scanner (MAGNETOM Trio, Siemens) with a CP knee coil (Siemens). An MR scan was performed before surgery and used to align and guide the positioning of the chamber on the skull for each individual animal by calculating the relative location of the amygdala to anatomical markers of the interaural line, the anterior commissure and the Bregma. After surgery, we performed another scan with electrodes directed toward the amygdala and two to three observers separately inspected the images and calculated the anterior-posterior and lateral-medial borders of the amygdala relative to each of the electrode penetrations. The depth of the regions was calculated from the dura surface.

Recordings. Each day, three to six microelectrodes (0.5–1 MΩ glass/narylene-coated tungsten, Alpha Omega or We-Sense) were lowered inside individual metal guides (Gauge 25xxtw; outer diameter, 0.51 mm; inner diameter, 0.41 mm; Cadence) into the brain using a head tower and electrode-positioning system (Alpha Omega). The guide was lowered to penetrate and cross the dura and stopped at 2–5 mm in the cortex. Electrodes were then moved independently into the amygdala. We first performed 8–10 mapping sessions by moving slowly and identifying electrophysiological markers of firing properties tracking the anatomical pathway into the amygdala (a typical path is frontal cortex → white-matter → striatum → globus pallidus → white matter / Nucleus basalis of Meynert → Central nucleus → BLA). Locations of recording electrodes were assessed and later reconstructed by alignment to multiple MRI sessions performed with electrodes. Electrode signals were pre-amplified, 0.3–6-kHz band-pass filtered, and sampled at 25 kHz, and online spike sorting was performed using a template-based approach (Alpha Lab Pro, Alpha Omega). We allowed 30 min for the tissue and signal to stabilize before starting acquisition and behavioral protocol. At the end of the recording period, offline spike sorting was performed for all sessions to improve unit isolation (offline sorter, Plexon).

Behavior. Monkeys were seated in a chair with a custom-made nasal mask attached to their nose[21,22]. The mask was connected to two pressure sensors with different sensitivity range (1/4 inch and 1 inch H_2O pressure range, AllSensors) that enable continuous and real-time detection of breath onset.

Each day and session, a pure tone was chosen randomly from 800–2,400 Hz and assigned as the CS for this session. Eight other pure tones were constructed with fixed frequency differences from the daily CS (differences of +2.5, +7.5, +20, +50, −2.5, −7.5, −20 and −50%). Tones were presented for 250 ms with 5-ms onset and offset ramps, and delivered via Adam5 speaker (ADAM Audio GmbH), located 40 cm behind and to the center of the animal.

Sessions started with a habituation (pre-conditioning) phase: six randomly interleaved presentations of each of the nine tones followed by an acquisition and conditioning phase in which only the CS was presented and paired with an aversive odor (15 trials of partial reinforcement, 2:1 ratio). We used a 1:20 solution of propionic acid distilled in mineral oil (Sigma Aldrich), as propionic acid stimulates olfactory and trigeminal receptors in the nose and is highly aversive to animals[21–23]. To test that aversive conditioning indeed results in wider generalization curves compared with pleasant conditioning, we performed several separate behavioral sessions in which the CS was followed by an appetitive odor (banana extract distilled in mineral oil, Tevaoz). Because distinct populations and microcircuits in the amygdala support positive or negative value[24,25] and fear or safety[22,26–28], coding for appetitive and aversive stimuli can coexist while differing in generalization patterns.

Each tone was triggered by spontaneous breath onsets, and the odor (US) was released at the following breath onset (but not before 1 s elapsed)[21,22]. The last phase, the generalization test (post-conditioning), included the presentation of all 9 tones again in a randomly interleaved manner, similar to the habituation, but with the CS partially reinforced to avoid extinction and maintain natural generalization. The following day, extinction was performed before the new session started, to minimize crossover.

The preparatory CR was measured as the change in the inhale breath that triggered the tone. We used the half height time of this inhale, its total volume, and its peak as different behavioral measures to test for robustness of results. These measures were derived per trial and normalized within session (over spontaneous breaths, that is, breaths during the intertrial interval that were not followed by tones or odors) to account for day-to-day fluctuations in the placement of the mask and general state.

Behavioral generalization functions were computed separately pre- and post-conditioning by fitting responses to all 9 tones with a Gaussian using a nonlinear-least-squares optimization procedure (Matlab, Mathworks).

There were 100 aversive sessions (66/34 per monkey) combined with electrophysiology and 14 (9/5) appetitive control behavioral sessions. Out of the 100, 64 had an $r^2 > 0.4$ for the Gaussian fit to behavior (**Fig. 1e** and **Supplementary Fig. 1**).

For JND-like estimation (**Fig. 1h** and **Supplementary Fig. 1e**), an exponent was fitted to the one-sided (absolute) generalization curve and the point that corresponded to 70.7% of the full response was extracted. This is equivalent to the converging point of the adaptive procedure that was used in previous studies to extract the JND in humans[5,6].

Data analysis. Firing rate was taken from the first 500 ms following the tone; thus, it is a preparatory period that does not include direct sensory stimulus (after the tone and before the expected time of odor release). Tuning curves were computed separately pre- and post-conditioning by fitting responses to all nine tones with a Gaussian[29] using a nonlinear-least-squares optimization procedure (Matlab, Mathworks). Of 300 recorded neurons from the BLA complex (107 of 193 per monkey), 100 had goodness of fit of $r^2 > 0.4$ during post-conditioning and were considered tuned neurons for further analyses (but see **Supplementary Fig. 9** for robustness of results using more strict or more relaxed criteria), and 70 of them were tuned ($r^2 > 0.4$) both pre- and post-conditioning (**Fig. 2b**,**d**). The width of the tuning curve was defined as the sigma (σ) of the Gaussian-fit and the preferred-stimulus (PS) as its mean. This approach to estimate tuning width confirmed previous observations that BLA neurons are broadly tuned in normal conditions[30], and also confirmed that there were no significant changes in tonic activity[31] (**Supplementary Fig. 4**), in the overall distribution of PS (**Supplementary Fig. 5**) or appearance of a second peak around the CS (**Supplementary Fig. 6**).

In several sessions, we included a complete traditional auditory test before the procedure ($n = 24$). This included a sequence of 40 pure tones evenly spaced in logarithmic scale and spanning 4 octaves from 300–4,800 Hz. Each tone was repeated ten times in a pseudorandom order and lasted 200 ms with a rise/fall ramp of 10 ms. This test had two goals. First, it tests for neuronal response after 24 h. Second, it is a standard auditory test that allows homogenous unbiased sampling of auditory space. The firing rate during the passive listening of the tone sequence was calculated, and a tuning curve was computed if the firing rate was significantly above baseline by 2 s.d. for at least 1 of the 40 tones. The best frequency (BF), or PS, was determined as the frequency evoking maximal firing rate. The distance between the PS from the CS was calculated by using the CS from the previous day. Bandwidth was determined as the tuning width (expressed in octaves) at which the tuning curve falls below 50% of the maximal firing rate at BF[32].

Surrogate data was synthesized on the basis of the correlations we found between the generalization curves' PS and width (**Fig. 2e**). We generated neurons with Gaussian tuning curves of mean (PS) homogenously distributed between −100 to 100 (to avoid edge effects for the range we used in our procedure), and σ (width) matching the expected width from the linear relationship we found in the real data (**Fig. 2e**). For each neuron, we generated trials with firing rates from a Poisson distribution where λ is the expected mean from the Gaussian tuning curve.

To evaluate discrimination, we calculated the d', a non-parametric sensitivity index, for each neuron

$$d'(s_1, s_2) = \frac{|\mu(s_2) - \mu(s_1)|}{\sqrt{\frac{\sigma(s_2)^2 + \sigma(s_1)^2}{2}}}$$

where s_1 and s_2 are two different tones, μ is the mean firing rate and σ is its s.d. We computed it for real data and for the surrogate data. We then calculated the change in sensitivity by taking the difference in d' between generalization test phase to habituation phase (represented in the color scale; **Fig. 3c**).

Fisher Information (*FI*)[33], a measure of the amount of information about the stimulus encoded in the neural response, was calculated using the following equation (for the case of Gaussian distribution)

$$FI(s) = \sum_{a=1}^{N} \frac{(f_a'(s))^2}{f_a(s)}$$

where a is a specific neuron and $f(s)$ is the neuron's tuning function (Gaussian in this case). Hence, each neuron contributes to the FI proportional to the square of its tuning curve's slope and inversely proportional to the average firing rate for a particular stimulus. Assuming independence between neurons, the total FI is the sum of all the neurons' contribution (number of neurons was kept fixed for all purposes; **Fig. 3d**).

To see that the neural d' is indeed directly related to the behavior, we calculated d' for the behavioral data as well (using conditioned responses from the different trials). For each tone, we obtained d' between it and all other tones (a vector of length 8 per each tone), both for the behavioral data and for the neural data. These two vectors were correlated and averaged across sessions, resulting in a correlation-coefficient between neural and behavioral d' for each tone, and for pre- and post-conditioning (**Fig. 3e**).

For tone estimation, we used an optimal linear readout (OLE)[34,35] computed on all neurons and stimuli. An OLE is an estimator of the form

$$\hat{s} = \sum_{i=1}^{N} fr_i W_i$$

where fr_i is neuron's i response (its firing rate) in a specific trial and W_i is the weight given to this neuron. OLE is a set of linear weights W that minimizes the mean Euclidean error (ME) with a uniform prior $p(s)$ over all stimuli.

$$E = \int_{-\infty}^{\infty} \langle (\hat{s} - s)^2 \rangle_m ds\, p(s)$$

The set of optimal linear weights W is given by $C^{-1}U$, where $c_{ij} = f_i(s)f_j(s)_x + \delta_{ij} f_i(s)_s$ and $u_i = sf_i(s)_s$.

In the current case, a specific neuron's response depends on its Gaussian-like tuning curve, and hence a response to a particular stimulus s is

$$fr(\mathrm{s}) = fr\_max * e^{-\frac{(\mathrm{s}-\mathrm{PS})^2}{2*\sigma^2}}$$

where fr_max is the maximal firing rate, PS is the preferred-stimulus of the neuron, and σ is the width of the tuning curve.

In the surrogate data we synthesized, the relationship between PS and σ (width) depends on the regression line, we found in the real data; that is, on the slope and y intercept of this regression (**Fig. 2e**).

For each combination of slope and y-intercept we calculated the ME in tone estimation (represented in the color scale; **Fig. 3f**). The actual slope and y intercept from pre- and post-conditioning are marked, showing a decrease in mean error for the population characteristics in the generalization phase.

Next, the difference in ME between the pre-conditioning population and the post-conditioning one was re-calculated for different error surfaces, each created by using different priors for averaging across stimuli. Unlike the uniform prior used above (**Fig. 3f**), the different priors were Gaussians centered on the CS but with different widths, namely different σ.

$$p(s) = \frac{1}{\sqrt{(2*\pi*\sigma^2)}} * e^{-\frac{s^2}{2*\sigma^2}}$$

The underlying assumption here is that it is more important to have correct estimation for the CS, with a decaying function around it. We used σ ranging from 1 to 10 in steps of 0.5, and for each matching error surface, plotted the improvement in error from pre- to post-conditioning revealing a convex function of σ (minima = 4; **Fig. 3g**). To compare this neural measure to behavior, we separately fitted a Gaussian to the behavioral generalization data (**Figs. 1e**,**f** and **3g**).

Auditory experiments for basic tuning properties. Animals passively listened to a sequence of 40 pure tones evenly spaced in logarithmic scale and spanning four octaves from 300–4,800 Hz. The tones were delivered via Adam5 speaker (ADAM Audio GmbH), located 40 cm behind and to the center of the animal. Each tone was repeated 30 times in a pseudorandom order and lasted 200 ms with a rise/fall ramp of 10 ms. Inter-tone interval was of 200 ms.

Units were classified responsive if the firing rate during tones was significantly above inter-tone interval baseline by 2 s.d. for at least one of the tones. Spectro-temporal receptive fields (STRFs) and tuning curves were constructed from responses to the tonal stimuli by computing peristimulus time histograms. Suppressive effects of tones were very rare, possibly because of the low background firing rate of amygdalar units. Overall, 343 single units were recorded, 126 were classified responsive, and 99 possessed a tuning curve.

The BF was determined as the frequency that evoked the maximal firing rate. Bandwidth was determined as the tuning width (expressed in octaves) at which the tuning curve fell below 50% of the maximal firing rate[32]. The latency, that is, the initiation of the response to the BF tone, was calculated as the first bin after tone presentation in which the number of spikes per second was above 2 s.d. from baseline. To do so, the firing rate of each unit was Z scored by Z = (RA – RB)/SB, where RA is the firing rate in a given bin after the presentation of the tone, RB is the mean firing rate for the baseline period immediately before the presentation of the auditory stimulus, and SB is the s.d. of the firing rate during the baseline period. Next, Z scores for each unit were averaged across 30 trials for the BF and each time bin. Finally, the time course of the mean Z score was plotted as a function of time. After the response was initiated, the cessation of the response was calculated as the time in which the spike count was no longer statistically different from baseline.

A **Supplementary Methods Checklist** is available.

21. Livneh, U. & Paz, R. *J. Neurosci. Methods* **192**, 90–95 (2010).
22. Livneh, U. & Paz, R. *J. Neurosci.* **32**, 8598–8610 (2012).
23. Livneh, U. & Paz, R. *Neuron* **75**, 133–142 (2012).
24. Paton, J.J., Belova, M.A., Morrison, S.E. & Salzman, C.D. *Nature* **439**, 865–870 (2006).
25. Balleine, B.W. & Killcross, S. *Trends Neurosci.* **29**, 272–279 (2006).
26. Herry, C. *et al. Nature* **454**, 600–606 (2008).
27. Sangha, S., Chadick, J.Z. & Janak, P.H. *J. Neurosci.* **33**, 3744–3751 (2013).
28. Genud-Gabai, R., Klavir, O. & Paz, R. *J. Neurosci.* **33**, 17986–17994 (2013).
29. Moore, A.K. & Wehr, M. *J. Neurosci.* **33**, 13713–13723 (2013).
30. Bordi, F. & LeDoux, J. *J. Neurosci.* **12**, 2493–2503 (1992).
31. Ciocchi, S. *et al. Nature* **468**, 277–282 (2010).
32. Bartlett, E.L., Sadagopan, S. & Wang, X. *J. Neurophysiol.* **106**, 849–859 (2011).
33. Cover, T.T., Thomas, J.A. *Elements of Information Theory* (Wiley-Interscience, 1991).
34. Seung, H.S. & Sompolinsky, H. *Proc. Natl. Acad. Sci. USA* **90**, 10749–10753 (1993).
35. Salinas, E. & Abbott, L.F. *J. Comput. Neurosci.* **1**, 89–107 (1994).

doi:10.1038/nn.3900