

ARTICLES

Switching on and off fear by distinct neuronal circuits

Cyril Herry^{1*}, Stephane Ciocchi^{1*}, Verena Senn¹, Lynda Demmou¹, Christian Müller¹ & Andreas Lüthi¹

Switching between exploratory and defensive behaviour is fundamental to survival of many animals, but how this transition is achieved by specific neuronal circuits is not known. Here, using the converse behavioural states of fear extinction and its context-dependent renewal as a model in mice, we show that bi-directional transitions between states of high and low fear are triggered by a rapid switch in the balance of activity between two distinct populations of basal amygdala neurons. These two populations are integrated into discrete neuronal circuits differentially connected with the hippocampus and the medial prefrontal cortex. Targeted and reversible neuronal inactivation of the basal amygdala prevents behavioural changes without affecting memory or expression of behaviour. Our findings indicate that switching between distinct behavioural states can be triggered by selective activation of specific neuronal circuits integrating sensory and contextual information. These observations provide a new framework for understanding context-dependent changes of fear behaviour.

The amygdala is a key brain structure mediating defensive behaviour in states of fear and anxiety. Such states can be induced by classical auditory fear conditioning, in which an initially neutral auditory stimulus (the conditioned stimulus, CS) comes to elicit a fear response after pairing with an aversive foot shock (the unconditioned stimulus, US). Subsequent repetitive presentations of the CS alone induce a progressive decrease in the fear response, a phenomenon called extinction. Whereas firing of amygdala neurons is critical for the retrieval of conditioned fear memories^{1–6}, their firing after the extinction of conditioned fear is thought to be constrained by local inhibitory circuits activated by the medial prefrontal cortex (mPFC)^{6–11}. Converging evidence from animal studies indicates, however, that the basolateral complex of the amygdala (BLA), comprising the lateral (LA) and the basal (BA) nuclei, actively participates in fear extinction^{12–17}. Although fear extinction is an active learning process eventually leading to the formation of a consolidated extinction memory^{16,17}, it is a fragile behavioural state that is readily influenced by context^{18,19}. Changing context results in the immediate recovery of the previously conditioned fear response, a process known as fear renewal^{18,19}. *In vivo* pharmacological studies indicate that the hippocampus, which is reciprocally connected to the BLA²⁰, processes contextual information during fear conditioning, extinction and renewal^{19,21–23}. Thus, bi-directional changes in fear behaviour during extinction and context-dependent renewal are likely to be encoded within a distributed network containing the BLA, the mPFC and the hippocampus; however, the neuronal circuits mediating such behavioural transitions are not known. In particular, this raises the question of whether there are specialized circuits driving behavioural transitions in opposite directions.

To address this question, we used a combination of *in vivo* single-unit recordings and targeted pharmacological inactivation in behaving mice. Because the BA is strongly connected to the hippocampus²⁰ and to the mPFC^{24,25}, and because extinction has previously been shown to induce the expression of the activity-dependent immediate early gene product Fos in BA neurons²⁶, we focused our study on this sub-nucleus. Here we identify two distinct neuronal circuits differentially connected with the mPFC and the hippocampus, and show

that a rapid switch in the balance of activity between those circuits specifically drives behavioural transitions without being necessary for memory storage or behavioural expression.

Distinct BA neurons encode fear and extinction

To examine plasticity of spike firing of individual BA neurons, C57Bl/6 mice were implanted with chronic recording electrodes and trained in a discriminative fear-conditioning paradigm (Fig. 1a). During training, mice learned to discriminate two auditory CS of different frequencies. One CS (the CS⁺) was paired with an aversive foot shock (US), whereas the second CS (CS⁻) was not paired. Twenty-four hours after fear conditioning, mice ($n = 30$) exhibited a selective increase in fear behaviour (as measured by freezing) when exposed to the CS⁺ in a different context (Fig. 1c). Extinction of conditioned fear behaviour was induced by exposing mice to 24 CS⁺ presentations in the absence of any aversive stimuli. After extinction training, CS⁺-induced freezing behaviour was reduced back to pre-conditioning levels, and did not differ from CS⁻-induced freezing (Fig. 1c).

Analysis of changes in CS⁺- and CS⁻-evoked spike firing during extinction training revealed that BA neurons (259 recorded units; Fig. 1b) could be divided into distinct functional classes. Consistent with previous reports^{27,28}, we found a class of neurons ('fear neurons'; $n = 43$ neurons, 22 mice; 17% of recorded units) that exhibited a selective increase in CS⁺-evoked spike firing during and after fear conditioning (Fig. 1d, Supplementary Fig. 1 and Supplementary Table 1). Subsequent extinction completely abolished this increase and converted it into a CS⁺-evoked inhibition (Fig. 1d). On average, spontaneous activity of fear neurons was not affected by fear conditioning or extinction (Supplementary Table 1). Thus, fear-conditioning-induced behavioural discrimination between the CS⁺ and the CS⁻, and its reversal by extinction, was accurately reflected at the neuronal level by the discriminative and reversible activity of fear neurons.

During extinction training, another class of neurons emerged. In contrast to fear neurons, 'extinction neurons' ($n = 35$ neurons, 20 mice; 14% of recorded units) did not show any increase in CS-evoked

¹Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, CH-4058 Basel, Switzerland.

*These authors contributed equally to this work.

responses during or after fear conditioning, but instead showed a slight reduction (Fig. 1e). However, subsequent extinction training induced a marked and selective increase in CS⁺-evoked activity in these neurons (Fig. 1e), without any changes in spontaneous activity. Plotting extinction-induced changes in z-score for individual fear and extinction neurons revealed that the two populations were separated in a bi-modal distribution (Supplementary Fig. 2). The remaining neurons did not exhibit any changes in activity during extinction (Supplementary Table 1). Thus, changes in CS⁺-evoked firing of fear and extinction neurons were oppositely correlated with behavioural extinction.

Although these results demonstrate a specific activation of fear and extinction neurons by a given CS, they do not address the question of whether individual extinction neurons can function as fear neurons for another CS, or vice versa. We therefore trained mice in a discriminative

extinction paradigm (Fig. 2a). In this paradigm, two different CS (CS1 and CS2) were first fear-conditioned, followed by extinction of only one of them (CS1). At the end of extinction, mice exhibited selective freezing behaviour when exposed to the non-extinguished CS2 (Fig. 2b). Fear neurons and extinction neurons were identified during fear conditioning and extinction of CS1 according to the same criteria as described above, and CS1- and CS2-evoked spike firing was compared at the end of extinction. Whereas individual extinction neurons ($n = 9$ neurons, 3 mice) responded to the extinguished CS (CS1), but not to the non-extinguished CS (CS2), fear neurons ($n = 8$, 3 mice) only fired during CS2 exposure, but remained unresponsive to the CS1 (Fig. 2c, d). These observations confirm that individual fear neurons and extinction neurons represent functionally distinct classes of neurons that can discriminate between extinguished and non-extinguished stimuli.

In addition to the BA, we also recorded from 38 neurons in the LA, which represents the main target of sensory afferents from the thalamus and cortex⁵. In keeping with previous studies^{1,29}, we did not observe any LA neuron in which CS⁺-evoked firing increased during extinction. Although we cannot exclude the existence of such neurons in the LA, this may suggest that extinction neurons are specific for the BA, where they represent 14% of all recorded neurons.

Activity balance predicts behaviour

Comparing the averaged time courses of CS-evoked activity of fear and extinction neurons during the acquisition of behavioural extinction indicated that significant behavioural changes occurred after the activity scores of the two populations of neurons crossed over (Fig. 3a). The largest changes in CS-evoked activity for both fear and extinction neurons occurred between the third and the fourth blocks of extinction training, which are separated by 24 h, suggesting that an overnight consolidation process may be required. To investigate further the exact time point during extinction learning at which fear and extinction neurons displayed a significant change in activity, we applied a change-point analysis algorithm³⁰. Change-point analysis identifies the trial(s) exhibiting a significant change in neuronal activity or freezing behaviour relative to the preceding trials. This analysis confirmed that changes in neuronal activity precede behavioural changes, and revealed that the activity of extinction neurons started to increase one trial before the activity of fear neurons began to decline (Fig. 3b, c). Plotting activity changes of single fear and extinction neurons recorded in the same animal showed that the sequence of events is the same in an individual animal, and that such changes occur abruptly in an all-or-none manner (Fig. 3c). This is consistent with the idea that behavioural changes are driven by sequential switches in the activity of two distinct neuronal circuits.

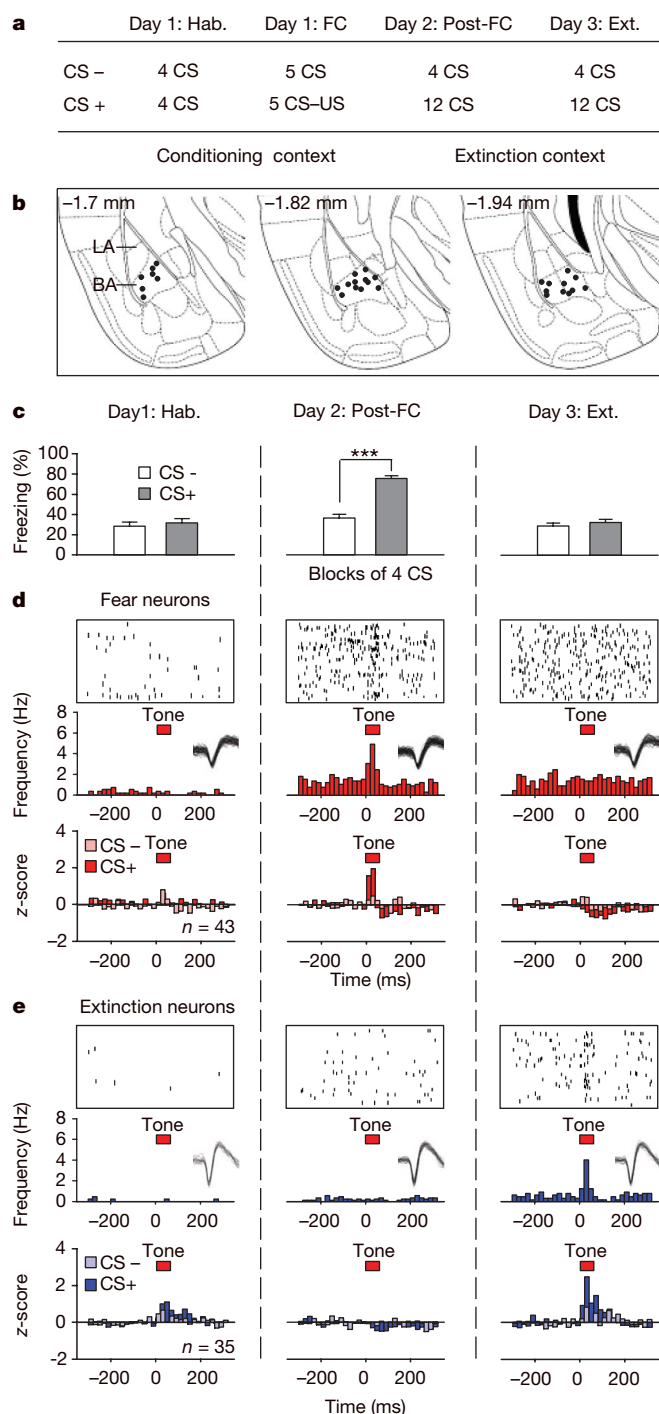


Figure 1 | Distinct populations of BA neurons encode fear conditioning and extinction. **a**, Experimental protocol. Ext., extinction; FC, fear conditioning; Hab., habituation. **b**, Coronal sections through the rostro-caudal extent of the amygdala, showing the location of the recording sites in the BA. **c**, Summary graphs illustrating behavioural data. During habituation, mice ($n = 30$) exhibited equally low freezing levels in response to CS⁺ and CS⁻ exposure. Twenty-four hours after fear conditioning, presentation of the CS⁺ (CS 1–4 on day 2), but not the CS⁻, evoked significantly higher freezing levels. After extinction, both CS⁺ (CS 9–12 on day 3) and CS⁻ elicited low freezing levels. Error bars indicate mean \pm s.e.m. **d**, **e**, Raster plots (top) and peristimulus time histograms (middle) illustrating selective changes in CS⁺-evoked firing of a representative fear (**d**) and an extinction (**e**) neuron. The duration of the auditory stimulus is indicated (red bar; tone). Insets show superimposed spike waveforms recorded during habituation, after fear conditioning and after extinction, respectively. Bottom: fear conditioning and extinction-induced changes in CS⁺-evoked firing of fear and extinction neurons. Fear neurons ($n = 43$ neurons from 22 mice) exhibited a selective increase in CS⁺-evoked firing after fear conditioning ($P < 0.001$ versus habituation or versus CS⁻), which was fully reversed on extinction. In contrast, CS⁺-evoked firing of extinction neurons ($n = 35$ neurons from 20 mice) was selectively increased after extinction ($P < 0.001$ versus post-FC or versus CS⁻). *** $P < 0.001$.

Rapid reversal of activity during fear renewal

To test whether the activity of fear and extinction neurons represents the same behavioural values in a different paradigm, we analysed renewal of extinguished fear behaviour and associated changes in CS-evoked spike firing. To make sure that extinction memory was stably consolidated, mice ($n = 15$) were tested for extinction memory 7 days after extinction training in the same context in which extinction training occurred (Fig. 4a). After successful recall of extinction memory (Fig. 4b), mice were transferred to the context in which they had been initially fear conditioned. Changing context resulted in a modest, but significant, increase in baseline freezing levels owing to contextual fear conditioning (Supplementary Fig. 3), and in a full renewal of the original cued fear memory (Fig. 4b).

During recall of extinction memory in the extinction context, presentation of the CS⁺ induced a selective activation of extinction neurons ($n = 14$, 8 mice) with no effect on fear neurons ($n = 19$, 9 mice; Fig. 4c, d). Thus, activation of extinction neurons by an extinguished CS is not a transient phenomenon, but remains stable for at least one week. After placing the animals in the fear-conditioning

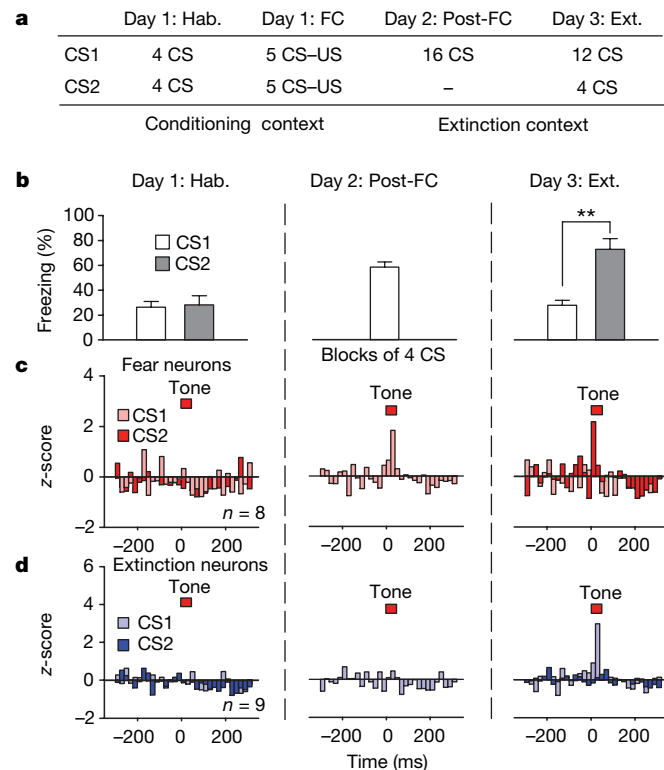


Figure 2 | Fear and extinction neurons discriminate stimuli with different emotional significance. **a**, Experimental design for discriminative extinction training. Initially, animals were fear conditioned to two distinct CS (CS1 and CS2). Both CS were paired with a US (CS-US). Subsequently, only one CS (CS1) was extinguished. **b**, Summary of behavioural data. During habituation, mice ($n = 6$) exhibited equally low freezing levels in response to CS1 and CS2 exposure. After fear conditioning, presentation of the CS1 (CS 1–4) evoked significantly increased freezing levels. After extinction to CS1, CS1 exposure (CS 9–12) elicited low freezing levels, whereas CS2-evoked freezing behaviour remains high. Error bars indicate mean \pm s.e.m. **c**, Fear conditioning- and extinction-induced changes in CS1- and CS2-evoked firing of fear neurons ($n = 8$ neurons from 3 mice). Twenty-four hours after fear conditioning (day 2), fear neurons exhibited increased firing in response to CS1 stimulation. After extinction of CS1, only CS2 stimulation elicited significant firing (day 3; $P < 0.05$ versus CS1). **d**, Fear conditioning- and extinction-induced changes in CS1- and CS2-evoked firing of extinction neurons ($n = 9$ neurons, 3 mice). After fear conditioning (day 2), extinction neurons did not respond to CS1 stimulation. After extinction of CS1, only CS1 stimulation elicited significant firing (day 3; $P < 0.05$ versus CS2). ** $P < 0.01$.

context, increased CS⁺-evoked freezing behaviour was associated with a complete reversal of spiking activity at the cellular level. Whereas extinction neurons stopped responding to CS⁺ stimulation, fear neurons exhibited a significant and selective increase in CS⁺-evoked spike firing (Fig. 4d). Extinction-resistant neurons were not significantly activated during renewal (not shown). Thus, a switch in the balance of activity between fear and extinction neurons not only

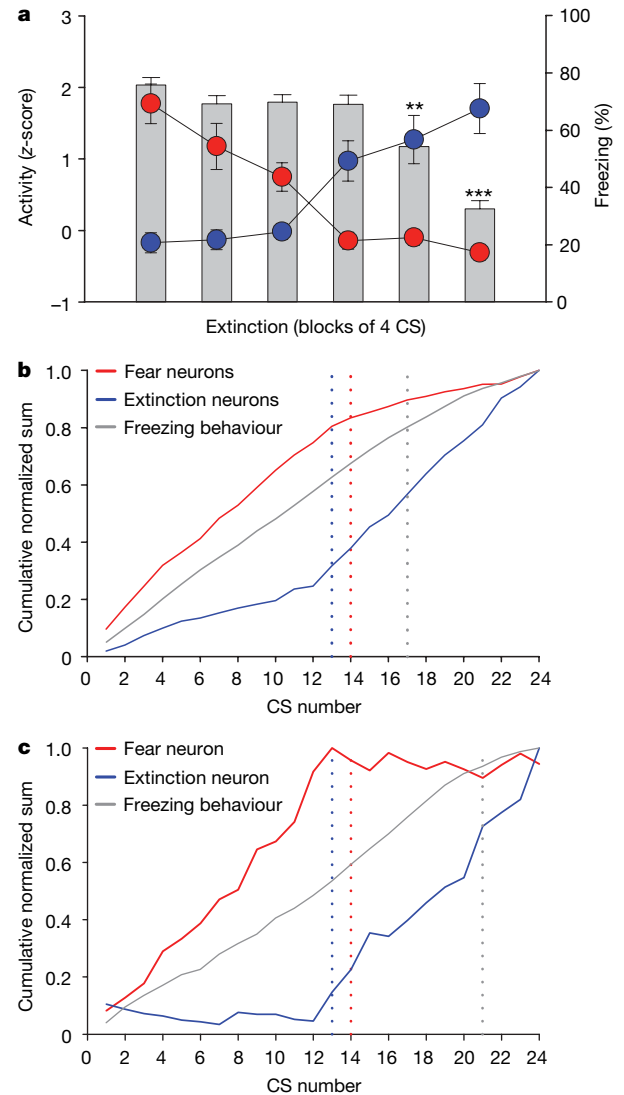


Figure 3 | Sequential switches in neuronal activity precede behavioural changes. **a**, Averaged time courses of freezing behaviour (grey bars; $n = 30$ mice) and neuronal activity (z-scores) of BA fear neurons (red circles; $n = 43$) and extinction neurons (blue circles; $n = 35$) during extinction training. Significant behavioural changes (that is, decreased freezing levels) occurred after activity scores of fear and extinction neurons have crossed over. Error bars indicate mean \pm s.e.m. **b**, Change-point analysis confirms that changes in neuronal activity preceded behavioural changes, and demonstrates that the activity of extinction neurons started to increase one trial before the activity of fear neurons changed. The plot represents the cumulative sums of the averaged and normalized z-scores of fear and extinction neurons, and freezing behaviour during extinction training. Change points are indicated by dotted lines. **c**, Normalized cumulative sums of the z-scores of a single fear neuron and a single extinction neuron recorded in the same animal together with the corresponding freezing behaviour during extinction training. Change-point analysis reveals that the extinction neuron abruptly switched on one trial before the fear neuron switched off. Changes in neuronal activity preceded behavioural changes. Change points are indicated by dotted lines. ** $P < 0.01$, *** $P < 0.001$.

reflects extinction but also parallels rapid context-dependent renewal of conditioned fear responses.

Differential long-range connectivity

We next addressed the question of whether fear neurons and extinction neurons are anatomically segregated. Comparing the location of electrolytic lesions made by the electrodes from which fear and extinction neurons were recorded did not provide any evidence for anatomical segregation (Supplementary Fig. 4). As a complementary approach, we compared the anatomical distribution of BA neurons activated during exposure to an extinguished or to a non-extinguished CS using the immediate early gene product Fos as an activity-marker. Given the similar numbers of extinction and fear neurons, one would

predict that an extinguished and a non-extinguished CS should induce Fos expression in an equal number of BA neurons with an overlapping anatomical distribution. Consistent with this, we found no difference in the density and anatomical distribution of Fos-positive neurons in animals exposed to an extinguished and a non-extinguished CS (Supplementary Fig. 4). Together, these results suggest that BA fear and extinction neurons are intermingled in a salt-and-pepper-like manner.

Converging evidence supports a role for the mPFC in the consolidation of extinction memory^{8,16,17,31}, and for the hippocampus in processing contextual information relevant for the expression and extinction of conditioned fear behaviour¹⁹. This raises the question as to how fear and extinction neurons in the BA communicate with the mPFC and the hippocampus during context-dependent behavioural transitions. We first addressed the possibility that fear neurons might be excitatory projection neurons whereas extinction neurons might be inhibitory interneurons. However, both fear and extinction neurons exhibited low spontaneous firing rates characteristic of BLA projection neurons³² (Supplementary Table 1). Consistent with this, analysis of cross-correlations between identified fear or extinction neurons and neighbouring BA neurons revealed short-latency excitatory interactions (Supplementary Fig. 5). To examine whether identified fear and extinction neurons project to, or receive input from, the mPFC and/or the hippocampus, we tested for antidromic activation of BA efferents and orthodromic activation of afferents by using extracellular stimulation electrodes in re-anaesthetized mice (Fig. 5a;

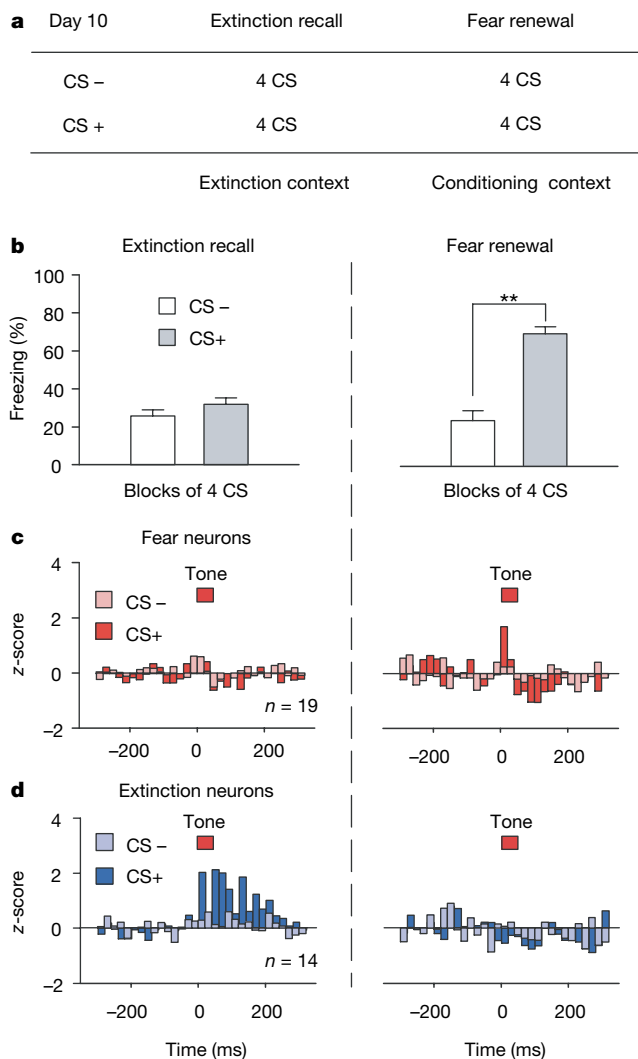


Figure 4 | Context-dependent fear renewal induces rapid reversal of neuronal activity patterns. **a**, Experimental protocol. **b**, Summary of behavioural data. Seven days after extinction, extinction memory was tested in the same context in which extinction training took place ($n = 15$ animals). Both CS⁺ and CS⁻ elicited low freezing behaviour. Subsequently, mice were placed back into the context in which fear conditioning took place. In this context, exposure to the CS⁺ evoked significantly more freezing than CS⁻ stimulation. Error bars indicate mean \pm s.e.m. **c**, Context-dependent changes in CS⁺-evoked firing of fear neurons ($n = 19$ neurons from 9 mice). Fear neurons exhibited a context-dependent increase in CS⁺-evoked firing in the fear-conditioning context where freezing levels were high ($P < 0.05$ versus extinction context and versus CS⁻). **d**, Extinction neurons ($n = 14$ neurons, 8 mice) showed the opposite pattern. Whereas CS⁺-exposure elicited strong firing in the extinction context ($P < 0.05$ versus fear-conditioning context and versus CS⁻), extinction neurons did not show any CS⁺-evoked responses in the fear-conditioning context. ** $P < 0.01$.

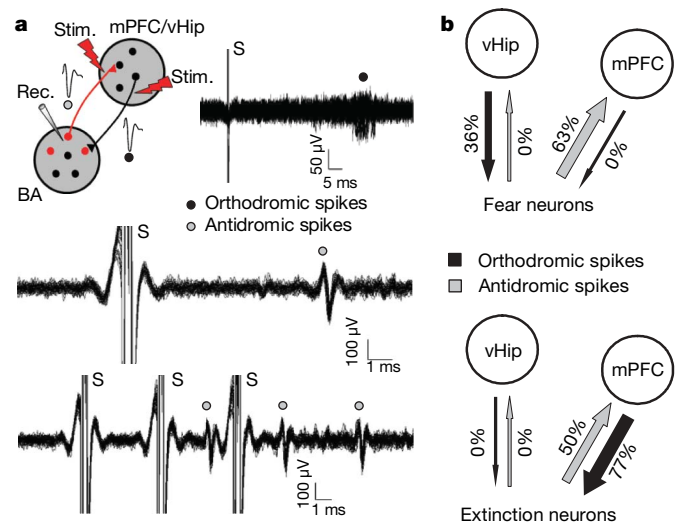


Figure 5 | Fear neurons and extinction neurons are part of distinct neuronal circuits. **a**, Use of extracellular stimulation in anaesthetized mice to identify orthodromic and antidromic connections between BA neurons and the mPFC or the hippocampus. Top left: schematic illustrating the placement of stimulating and recording electrodes. Rec., recording electrode; Stim., stimulation electrode. Top right: orthodromic spikes elicited in a BA fear neuron on stimulation of the ventral hippocampus. Orthodromic spikes exhibited a large temporal jitter and high failure rates. Middle: antidromic spikes recorded from a BA extinction neuron in response to mPFC stimulation. Antidromic spikes exhibited low temporal jitter, and followed high frequency (200 Hz) stimulation (bottom). **b**, Top: fear neurons project to the mPFC (5 out of 8 stimulated neurons) and receive input from the hippocampus (5 out of 14 stimulated neurons). vHip, ventral hippocampus. No antidromic responses from the hippocampus (0 out of 14 stimulated neurons) or orthodromic responses from the mPFC (0 out of 8 stimulated neurons) were observed. The graph depicts the percentage of all stimulation experiments in which a particular response was observed in identified fear neurons. Bottom: extinction neurons are reciprocally connected with the mPFC (antidromic responses, 3 out of 6 stimulated neurons; orthodromic responses, 7 out of 9 stimulated neurons, $P < 0.001$ versus fear neurons). No connections with the hippocampus were observed (0 out of 9 stimulated neurons, $P < 0.05$ versus fear neurons).

see Methods). These experiments revealed that fear neurons received input from the hippocampus, whereas no connections with the hippocampus were found for extinction neurons ($P < 0.05$ versus fear neurons; Fig. 5b). Although these findings cannot exclude that some extinction neurons might be contacted by hippocampal afferents, they demonstrate that the probability of receiving hippocampal input is significantly different for fear and extinction neurons. Likewise, fear and extinction neurons were differentially connected with the mPFC. Whereas extinction neurons were reciprocally connected, fear neurons projected to the mPFC, but we did not find any inputs ($P < 0.001$ versus extinction neurons; Fig. 5b). Extinction-resistant neurons were reciprocally connected to both the mPFC and to the hippocampus (Supplementary Fig. 6). Taken together, these findings indicate that fear and extinction neurons, although co-localized within the same nucleus, not only are functionally specialized but also form part of discrete neuronal circuits.

BA inactivation prevents behavioural transitions

The observed changes in CS⁺-evoked spike firing of fear and extinction neurons during the extinction and context-dependent renewal of conditioned fear responses could be necessary for the acquisition, storage and/or behavioural expression of the learned information. To distinguish between these possibilities, we used micro-iontophoresis of a fluorescently labelled GABA_A (γ -aminobutyric acid subtype A) receptor agonist (muscimol) to reversibly inactivate neuronal activity in the BA in a targeted and controlled manner (Fig. 6a). Simultaneous iontophoresis and multi-unit recording revealed that muscimol application silenced neuronal activity in the BA for more than 60 min (Fig. 6b). We first tested whether BA activity was necessary for the acquisition of extinction. Inactivation of the BA completely prevented the decrease in freezing behaviour normally observed during extinction training (Fig. 6c), with no effect on pre-CS freezing levels (not shown). Twenty-four hours later, after wash-out of muscimol, the same animals initially exhibited high freezing levels followed by normal fear extinction, demonstrating that BA inactivation did not merely interfere with the behavioural expression of extinction, nor irreversibly damage BA function (Fig. 6d). These results demonstrate that BA activity is necessary for the acquisition of extinction.

Next, we tested whether BA activity was necessary for the context-dependent renewal of previously extinguished fear responses. Mice exhibiting low freezing levels during recall of extinction memory one week after extinction training were injected with muscimol before renewal. In contrast to control animals injected with the fluorescent label only, muscimol-injected animals exhibited no increase in freezing levels when exposed to the CS⁺ in the fear-conditioning context (Fig. 6e). These results demonstrate that BA activity is necessary for context-dependent fear renewal.

Because muscimol unselectively silences all neurons in the targeted region, the high fear level observed in muscimol-injected mice during extinction learning cannot be accounted for by activity of fear neurons. Conversely, the low fear level displayed by muscimol-injected mice during context-dependent fear renewal cannot be dependent on the activation of extinction neurons. Thus, whereas animals with inactivated BA are able to express high and low fear states, possibly by activation of other parts of the amygdala and the mPFC, they exhibit emotional perseveration (that is, they remained in the emotional state they were in before BA inactivation). This suggests that the BA is unlikely to be associated with the storage, retrieval or expression of conditioned fear and extinction memories, but may instead mediate context-dependent behavioural transitions between low and high fear states.

Thus, silencing of BA activity should have no effect on the retrieval and expression of conditioned and extinguished fear memories when there is no need to change fear levels in a context-dependent manner. Consistent with this scenario, BA inactivation had no effect on the retrieval or expression of consolidated extinction memories (Fig. 6e).

Moreover, in animals that had been fear conditioned one week before, but that did not receive extinction training, muscimol had no effect on the retrieval and expression of the fear memory independently of the context in which they were tested (Fig. 6f).

Discussion

Our data show that the BA contains distinct populations of neurons for which activity is oppositely correlated with high and low fear behaviour—two converse behavioural states. Although fear and extinction neurons represent relatively small sub-populations within the BA, a rapid switch in the balance of their activity is essential for

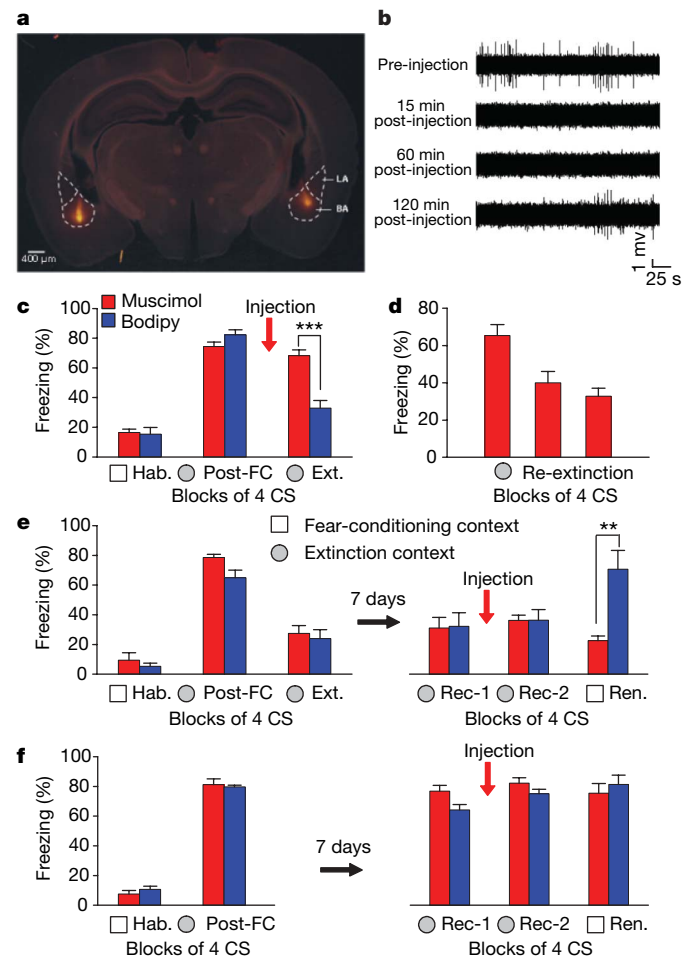


Figure 6 | Targeted inactivation of the BA prevents behavioural changes without affecting memory. **a**, Epifluorescence image illustrating bilateral targeting of the BA with muscimol covalently attached to a fluorescent tag (bodipy, dipyrromethene boron difluoride). **b**, Simultaneous multi-unit recordings revealed silencing of neuronal activity for up to two hours after muscimol iontophoresis. **c**, Inactivation of the BA before extinction training prevented the acquisition of extinction. Control mice injected with fluorophore only ($n = 5$; blue bars) exhibited significant reduction of freezing levels after extinction training. Muscimol-injected animals ($n = 11$; red bars) showed high freezing levels after extinction. **d**, Twenty-four hours later, in the absence of muscimol, the same animals showed normal acquisition of extinction ($P < 0.05$). Bars illustrate the progressive decrease in freezing behaviour during extinction learning. **e**, Inactivation of the BA prevented context-dependent renewal. Control mice injected with fluorophore only ($n = 5$) exhibited a significant increase in freezing levels on change of context ($P < 0.05$). Muscimol-injected animals ($n = 5$) did not show any context-dependent fear renewal ($P < 0.01$ versus control). **f**, In the absence of extinction training, BA inactivation did not affect fear memory retrieval. Fluorophore-injected mice ($n = 4$) and muscimol-injected mice ($n = 5$) exhibited equal freezing levels during CS⁺ exposure in the fear-conditioning context one week after fear conditioning. Error bars indicate mean \pm s.e.m. ** $P < 0.01$, *** $P < 0.001$.

triggering behavioural transitions during extinction and context-dependent fear renewal. Although intermingled within the BA, fear and extinction neurons are differentially connected with the hippocampus and the mPFC, two brain areas previously implicated in extinction and context-dependent renewal of conditioned fear responses. In keeping with the proposed role of the ventral hippocampus in mediating context-dependent renewal of fear behaviour in animals subjected to extinction²³, we found that hippocampal input to the BA selectively targets fear neurons over extinction neurons. Thus, hippocampal input to BA fear neurons may override the retrieval of extinction memory allowing for fear expression after a particular CS has undergone extinction. Extinction neurons, in turn, are bi-directionally connected with the mPFC and are switched on during extinction training. This indicates that they may be upstream of a previously identified population of mPFC neurons thought to mediate consolidation of extinction memory, because they are activated by an extinguished CS during recall, but not during the acquisition of extinction⁸.

Previous findings demonstrate that the BLA is not critical for triggering behavioural transitions during reversal learning in a two-odour-discrimination task^{33,34}. Nevertheless, abnormally persistent BLA activity induced by orbitofrontal cortex lesions³³ or repeated cocaine administration³⁴ interferes with reversal learning. This suggests that, whereas the BLA can only veto slow behavioural transitions during more complex reversal learning tasks, it is actively involved in situations requiring rapid context-dependent switching between two converse behavioural states.

How might activity of BA fear and extinction neurons mediate behavioural transitions? In keeping with a role for the amygdala in facilitating network function and memory formation in other parts of the brain^{35–37}, a possible interpretation is that BA fear and extinction neurons might drive or facilitate the induction of synaptic plasticity in their respective target areas. Moreover, whereas previous studies using pre-fear-conditioning lesions came to the conclusion that the BA does not contribute to the acquisition or the expression of conditioned fear^{38–41} (but see ref. 42), a recent analysis using post-fear-conditioning lesions indicates that the BA also contributes to the consolidation of long-term fear memories⁴¹. This suggests that repeated activity of BA fear neurons, over longer-time periods, may be required for fear memory consolidation.

Our findings are consistent with the idea that in mammals, as in invertebrates^{43,44}, switches between appropriate behavioural states can be driven by discrete neuronal circuits. It may be a general principle of the functional micro-architecture of the nervous system in diverse species, that circuits mediating switches between distinct behavioural states are located in close anatomical proximity thereby allowing for local interactions. However, it remains to be shown how fear and extinction neurons interact locally. Finally, our results also suggest that context-dependent recovery of extinguished fear behaviour in humans⁴⁵, which represents a major clinical obstacle for the therapy of certain anxiety disorders⁴⁶, might be modulated by tipping the balance of activity between specific neuronal circuits.

METHODS SUMMARY

Behaviour and pharmacological inactivations. Mice were submitted to a discriminative auditory fear-conditioning paradigm in which the CS⁺, but not the CS⁻, was paired with a US (mild foot shock). Extinction training was performed over two days in a different context¹⁴. One week later, mice were placed in the extinction context for recall of extinction, and in the original conditioning context for fear renewal. Freezing behaviour was quantified during each behavioural session using an automatic infrared beam detection system as described previously¹⁷. Bilateral inactivation of the BA was achieved using micro-iontophoretic injection of fluorescently labelled muscimol before extinction training or context-dependent fear renewal.

Electrophysiological recordings and analysis. Individual neurons were recorded extracellularly in freely behaving mice during fear conditioning, extinction, recall of extinction and context-dependent fear renewal. Spikes of individual neurons were sorted by time–amplitude window discrimination and

template matching as described previously^{47,48}. Cluster quality and unit stability were verified by quantifying the cluster separation and the stability of the average waveform shape over time^{48,49} (Supplementary Fig. 7). Unit isolation was verified using auto- and cross-correlation histograms. Spike rasters and histograms were constructed by aligning sweeps relative to the CS onset, and CS-evoked responses were normalized to baseline activity using a z-score transformation. Antidromic and orthodromic spikes evoked by extracellular stimulations of the mPFC or the ventral hippocampus were recorded in previously identified neurons in anaesthetized mice.

Immunohistochemistry. Mice were submitted to an auditory fear-conditioning paradigm in which only one CS was used. Two hours after the end of training, mice exposed to the CS only, extinguished and non-extinguished mice were transcardially perfused, and their brains collected for immunohistochemical staining of the immediate early gene product Fos as previously described²⁶.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 19 March; accepted 11 June 2008.

Published online 9 July 2008.

- Quirk, G. J., Repa, C. & LeDoux, J. E. Fear conditioning enhances short-latency auditory responses of lateral amygdala neurons: parallel recordings in the freely behaving rat. *Neuron* **15**, 1029–1039 (1995).
- Collins, D. R. & Paré, D. Differential fear conditioning induces reciprocal changes in the sensory responses of lateral amygdala neurons to the CS(+) and CS(-). *Learn. Mem.* **7**, 97–103 (2000).
- Rosenkranz, J. A. & Grace, A. A. Dopamine-mediated modulation of odour-evoked amygdala potentials during Pavlovian conditioning. *Nature* **417**, 282–287 (2002).
- Goossens, K. A., Hobin, J. A. & Maren, S. Auditory-evoked spike firing in the lateral amygdala and Pavlovian fear conditioning: mnemonic code or fear bias? *Neuron* **40**, 1013–1022 (2003).
- LeDoux, J. E. Emotion circuits in the brain. *Annu. Rev. Neurosci.* **23**, 155–184 (2000).
- Maren, S. & Quirk, G. J. Neuronal signaling of fear memory. *Nature Rev. Neurosci.* **5**, 844–852 (2004).
- Quirk, G. J., Russo, G. K., Barron, J. L. & Lebron, K. The role of ventromedial prefrontal cortex in the recovery of extinguished fear. *J. Neurosci.* **20**, 6225–6231 (2000).
- Milad, M. R. & Quirk, G. J. Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature* **420**, 70–74 (2002).
- Rosenkranz, J. A., Moore, H. & Grace, A. A. The prefrontal cortex regulates lateral amygdala neuronal plasticity and responses to previously conditioned stimuli. *J. Neurosci.* **23**, 11054–11064 (2003).
- Berretta, S., Pantazopoulos, H., Caldera, M., Pantazopoulos, P. & Paré, D. Infralimbic cortex activation increases c-Fos expression in intercalated neurons of the amygdala. *Neuroscience* **132**, 943–953 (2005).
- Paré, D., Quirk, G. J. & LeDoux, J. E. New vistas on amygdala networks in conditioned fear. *J. Neurophysiol.* **92**, 1–9 (2004).
- Falls, W. A., Miserendino, M. J. & Davis, M. Extinction of fear-potentiated startle: blockade by infusion of an NMDA antagonist into the amygdala. *J. Neurosci.* **12**, 854–863 (1992).
- Lu, K. T., Walker, D. L. & Davis, M. Mitogen-activated protein kinase cascade in the basolateral nucleus of amygdala is involved in extinction of fear-potentiated startle. *J. Neurosci.* **21**, RC162 (2001).
- Herry, C., Trifilieff, P., Micheau, J., Lüthi, A. & Mons, N. Extinction of auditory fear conditioning requires MAPK/ERK activation in the basolateral amygdala. *Eur. J. Neurosci.* **24**, 261–269 (2006).
- Sotres-Bayon, F., Bush, D. E. & LeDoux, J. E. Acquisition of fear extinction requires activation of NR2B-containing NMDA receptors in the lateral amygdala. *Neuropsychopharmacology* **32**, 1929–1940 (2007).
- Quirk, G. J., Garcia, R. & Gonzalez-Lima, F. Prefrontal mechanisms in extinction of conditioned fear. *Biol. Psychiatry* **60**, 337–343 (2006).
- Myers, K. M. & Davis, M. Mechanisms of fear extinction. *Mol. Psychiatry* **12**, 120–150 (2007).
- Bouton, M. E. Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biol. Psychiatry* **52**, 976–986 (2002).
- Bouton, M. E., Westbrook, R. F., Corcoran, K. A. & Maren, S. Contextual and temporal modulation of extinction: Behavioral and biological mechanisms. *Biol. Psychiatry* **60**, 352–360 (2006).
- Pitkänen, A., Pikkarainen, M., Nurminen, N. & Ylinen, A. Reciprocal connections between the amygdala and the hippocampal formation, perirhinal cortex, and postrhinal cortex in rat. A review. *Ann. NY Acad. Sci.* **911**, 369–391 (2000).
- Corcoran, K. A. & Maren, S. Hippocampal inactivation disrupts contextual retrieval of fear memory after extinction. *J. Neurosci.* **21**, 1720–1726 (2001).
- Corcoran, K. A., Desmond, T. J., Frey, K. A. & Maren, S. Hippocampal inactivation disrupts the acquisition and contextual encoding of fear extinction. *J. Neurosci.* **25**, 8978–8987 (2005).
- Hobin, J. A., Ji, J. & Maren, S. Ventral hippocampal muscimol disrupts context-specific fear memory retrieval after extinction in rats. *Hippocampus* **16**, 174–182 (2006).

24. McDonald, A. J. Organization of amygdaloid projections to the prefrontal cortex and associated striatum in the rat. *Neuroscience* **44**, 1–14 (1991).
25. McDonald, A. J., Mascagni, F. & Guo, L. Projections of the medial and lateral prefrontal cortices to the amygdala: a *Phaseolus vulgaris* leucoagglutinin study in the rat. *Neuroscience* **71**, 55–75 (1996).
26. Herry, C. & Mons, N. Resistance to extinction is associated with impaired immediate early gene induction in medial prefrontal cortex and amygdala. *Eur. J. Neurosci.* **20**, 781–790 (2004).
27. Muramoto, K., Ono, T., Nishijo, H. & Fukuda, M. Rat amygdaloid neuron responses during auditory discrimination. *Neuroscience* **52**, 621–636 (1993).
28. Maren, S., Poremba, A. & Gabriel, M. Basolateral amygdaloid multi-unit neuronal correlates of discriminative avoidance learning in rabbits. *Brain Res.* **549**, 311–316 (1991).
29. Repa, J. C. *et al.* Two different lateral amygdala cell populations contribute to the initiation and storage of memory. *Nature Neurosci.* **4**, 724–731 (2001).
30. Gallistel, C. R., Fairhurst, S. & Balsam, P. The learning curve: Implications of a quantitative analysis. *Proc. Natl Acad. Sci. USA* **101**, 13124–13131 (2004).
31. Herry, C. & Garcia, R. Prefrontal cortex long-term potentiation, but not long-term depression, is associated with the maintenance of extinction of learned fear in mice. *J. Neurosci.* **22**, 577–583 (2002).
32. Likhtik, E., Pelletier, J. G., Popescu, A. T. & Paré, D. Identification of basolateral amygdala projection cells and interneurons using extracellular recordings. *J. Neurophysiol.* **96**, 3257–3265 (2006).
33. Stalnaker, T. A., Franz, T. M., Singh, T. & Schoenbaum, G. Basolateral amygdala lesions abolish orbitofrontal-dependent reversal impairments. *Neuron* **54**, 51–58 (2007).
34. Stalnaker, T. A. *et al.* Cocaine-induced decision-making deficits are mediated by miscoding in basolateral amygdala. *Nature Neurosci.* **10**, 949–951 (2007).
35. Paré, D. Role of the basolateral amygdala in memory consolidation. *Prog. Neurobiol.* **70**, 409–420 (2003).
36. McGaugh, J. L. The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annu. Rev. Neurosci.* **27**, 1–28 (2004).
37. Paz, R., Pelletier, J. G., Bauer, E. P. & Paré, D. Emotional enhancement of memory via amygdala-driven facilitation of rhinal interactions. *Nature Neurosci.* **9**, 1321–1329 (2006).
38. Amorapanth, P., LeDoux, J. E. & Nader, K. Different lateral amygdala outputs mediate reactions and actions elicited by a fear-arousing stimulus. *Nature Neurosci.* **3**, 74–79 (2000).
39. Nader, K., Majidishad, P., Amorapanth, P. & LeDoux, J. E. Damage to the lateral and central, but not other, amygdaloid nuclei prevents the acquisition of auditory fear conditioning. *Learn. Mem.* **8**, 156–163 (2001).
40. Sotres-Bayon, F., Bush, D. E. & LeDoux, J. E. Emotional perseveration: an update on prefrontal-amygdala interactions in fear extinction. *Learn. Mem.* **11**, 525–535 (2004).
41. Anglada-Figueroa, D. & Quirk, G. J. Lesions of the basal amygdala block expression of conditioned fear but not extinction. *J. Neurosci.* **25**, 9680–9685 (2005).
42. Goossens, K. A. & Maren, S. Contextual and auditory fear conditioning are mediated by the lateral, basal, and central amygdaloid nuclei in rats. *Learn. Mem.* **8**, 148–155 (2001).
43. Jing, J. & Gillette, R. Escape swim network interneurons have diverse roles in behavioral switching and putative arousal in *Pleurobranchanea*. *J. Neurophysiol.* **83**, 1346–1355 (2000).
44. Yapici, N., Kim, Y.-J., Ribeiro, C. & Dickson, B. J. A receptor that mediates the post-mating switch in *Drosophila* reproductive behavior. *Nature* **451**, 33–37 (2007).
45. Milad, M. R., Orr, S. P., Pitman, R. K. & Rauch, S. L. Context modulation of memory for fear extinction in humans. *Psychophysiology* **42**, 456–464 (2005).
46. Rodriguez, B. I., Craske, M. G., Mineka, S. & Hladek, D. Context-specificity of relapse: effects of therapist and environmental context on return of fear. *Behav. Res. Ther.* **37**, 845–862 (1999).
47. Herry, C. *et al.* Processing of temporal unpredictability in human and animal amygdala. *J. Neurosci.* **27**, 5958–5966 (2007).
48. Nicolelis, M. A. L. *et al.* Chronic, multisite, multielectrode recordings in macaque monkeys. *Proc. Natl Acad. Sci. USA* **100**, 11041–11046 (2003).
49. Jackson, A. & Fetzi, E. E. Compact movable microwire array for long-term chronic unit recordings in cerebral cortex of primates. *J. Neurophysiol.* **98**, 3109–3118 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. Friedrich, A. Matus and all members of the Lüthi laboratory for discussions and critical comments on the manuscript and E. Oakeley and M. Stadler for advice on statistical analysis and programming. This work was supported by the Austrian Science Fund (FWF), the Novartis Institutes for Biomedical Research, and the Novartis Research Foundation.

Author Contributions C.H., S.C., V.S., L.D. and C.M. performed the experiments and analysed the data. C.H., S.C. and A.L. designed the experiments and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.L. (andreas.luthi@fmi.ch) or C.H. (cyril.herry@fmi.ch).

METHODS

Animals. Male C57BL/6J mice (3 months old; RCC Ltd) were individually housed for 7 days before all experiments, under a 12 h light/dark cycle, and provided with food and water *ad libitum*. All animal procedures were performed in accordance with institutional guidelines and were approved by the Veterinary Department of the Canton of Basel-Stadt.

Behaviour. Fear conditioning and extinction took place in two different contexts (context A and B). The conditioning and extinction boxes and the floor were cleaned with 70% ethanol or 1% acetic acid before and after each session, respectively. To score freezing behaviour, an automatic infrared beam detection system placed on the bottom of the experimental chambers (Coulbourn Instruments) was used. The animals were considered to be freezing if no movement was detected for 2 s. On day 1, mice were submitted to a habituation session in context A, in which they received 4 presentations of the CS⁺ and the CS⁻ (total CS duration of 30 s, consisting of 50-ms pips repeated at 0.9 Hz, 2-ms rise and fall; pip frequency: 7.5 kHz or white-noise, 80 dB sound pressure level). Discriminative fear conditioning was performed on the same day by pairing the CS⁺ with a US (1-s foot shock, 0.6 mA, 5 CS⁺/US pairings; inter-trial interval: 20–180 s). The onset of the US coincided with the offset of the CS⁺. The CS⁻ was presented after each CS⁺/US association but was never reinforced (5 CS⁻ presentations, inter-trial interval: 20–180 s). The frequencies used for CS⁺ and CS⁻ were counterbalanced across animals. On day 2 and day 3, conditioned mice were submitted to extinction training in context B, during which they received 4 and 12 presentations of the CS⁻ and the CS⁺, respectively. Recall of extinction and context-dependent fear renewal were tested 7 days later in context B and A, respectively, with 4 presentations of the CS⁻ and the CS⁺. Pharmacological experiments were performed using the same conditioning and extinction protocol except for one group of mice that was not submitted to extinction training but tested for conditioned fear with 4 CS⁻ and 4 CS⁺ presentations on day 2 in context B. Seven days later, mice were submitted to 2 sessions of extinction recall 5 h apart in context B (4 presentations of each CS for each session). Finally, 10 min after the second recall session, mice were submitted to 4 CS⁻ and 4 CS⁺ presentations in context A for context-dependent fear renewal.

For discriminative extinction, mice were habituated on day 1 to 4 presentations of two different CS in context A (total CS duration of 30 s, consisting of 50-ms pips repeated at 0.9 Hz, 2 ms rise and fall; pip frequency: 7.5 kHz or white-noise, 80 dB sound pressure level). Both CS were subsequently paired with a US (1-s foot shock, 0.6 mA, 5 CS/US pairings for each CS; inter-trial interval: 20–180 s). The onset of the US coincided with the offset of the CS. On days 3 and 4, only one of the two CS was extinguished by 16 and 12 presentations in context B, respectively. At the end of the second extinction session, mice were exposed to 4 presentations of the non-extinguished CS in context B.

Surgery and recordings. Mice were anesthetized with isoflurane (induction 5%, maintenance 2.5%) in O₂. Body temperature was maintained with a heating pad (CMA/150, CMA/Microdialysis). Mice were secured in a stereotaxic frame and unilaterally implanted in the amygdala with a multi-wire electrode aimed at the following coordinates⁵⁰: 1.7 mm posterior to bregma; ±3.1 mm lateral to midline; and 4 mm to 4.3 mm deep from the cortical surface. The electrodes consisted of 8 to 16 individually insulated nichrome wires (13- μ m inner diameter, impedance 1–3 M Ω ; California Fine Wire) contained in a 26 gauge stainless steel guide canula. The wires were attached to a 10 pin to 18 pin connector (Omnetics). The implant was secured using cyanoacrylate adhesive gel. After surgery mice were allowed to recover for 7 days. Analgesia was applied before, and during the 3 days after, surgery (Metacam, Boehringer). Electrodes were connected to a headstage (Plexon) containing eight to sixteen unity-gain operational amplifiers. The headstage was connected to a 16-channel computer-controlled preamplifier (gain 100 \times , bandpass filter from 150 Hz to 9 kHz, Plexon). Neuronal activity was digitized at 40 kHz and bandpass filtered from 250 Hz to 8 kHz, and was isolated by time–amplitude window discrimination and template matching using a Multichannel Acquisition Processor system (Plexon). At the conclusion of the experiment, recording sites were marked with electrolytic lesions before perfusion, and electrode locations were reconstructed with standard histological techniques.

Single-unit spike sorting and analysis. Single-unit spike sorting was performed using Off-Line Spike Sorter (OFSS, Plexon) as described^{47,48} (Supplementary Fig. 7). Principal component scores were calculated for unsorted waveforms and plotted on three-dimensional principal component spaces, and clusters containing similar valid waveforms were manually defined. A group of waveforms was considered to be generated from a single neuron if it defined a discrete cluster in principal component space that was distinct from clusters for other units and if it displayed a clear refractory period (>1 ms) in the auto-correlogram histograms. In addition, two parameters were used to quantify the overall separation between identified clusters in a particular channel. These parameters include the J3 statistic,

which corresponds to the ratio of between-cluster to within-cluster scatter, and the Davies–Bouldin validity index (DB), which reflects the ratio of the sum of within-cluster scatter to between-cluster separation⁴⁸. High values for the J3 and low values for the DB are indicative of good single-unit isolation (Supplementary Fig. 7). Control values for this statistics were obtained by artificially defining two clusters from the centred cloud of points in the principal component space from channels in which no units could be detected. Template waveforms were then calculated for well-separated clusters and stored for further analysis. Clusters of identified neurons were analysed offline for each recording session using principal component analysis and a template-matching algorithm. Only stable clusters of single units recorded over the time course of the entire behavioural training were considered. Long-term single-unit stability isolation was first evaluated using Wavetracker (Plexon) in which principal component space-cylinders were calculated from a 5 min segment of data spontaneously recorded before any training session. Straight cylinders suggest that the same set of single units was recorded during the entire training session (Supplementary Fig. 7).

Second, we quantitatively evaluated the similarity of waveform shape by calculating linear correlation (r) values between average waveforms obtained over training days⁴⁹ (Supplementary Fig. 7). As a control, we computed the r values from average waveforms of different neurons.

Third, for each unit we used correlation analysis to quantitatively compare the similarity of waveform shape during CS⁺-stimulation and during a 60 s period of spontaneous activity recorded before each behavioural session (Supplementary Fig. 8). To avoid analysis of the same neuron recorded on different channels, we computed cross-correlation histograms. If a target neuron presented a peak of activity at a time that the reference neuron fires, only one of the two neurons was considered for further analysis. CS-induced neural activity was calculated by comparing the firing rate after stimulus onset with the firing rate recorded during the 500 ms before stimulus onset (bin size, 20 ms; averaged over blocks of 4 CS presentations consisting of 108 individual sound pips in total) using a z -score transformation. z -score values were calculated by subtracting the average baseline firing rate established over the 500 ms preceding stimulus onset from individual raw values and by dividing the difference by the baseline standard deviation. Only CS-excited neurons were considered for analysis. Classification of units was performed by comparing the largest significant z -score values within 100 ms after CS-onset during post-fear conditioning and extinction sessions according to the freezing levels. For high-fear states, the entire post-fear conditioning session was analysed, whereas, for low-fear states, analysis was restricted to the block of 4 CS presentations during which the fear level was the lowest. A unit was classified as a fear neuron if it exhibited a significant z -score value after fear conditioning (when freezing levels were high), but no significant z -score value after extinction (when freezing levels were low), and vice versa for extinction neurons. Finally, units were classified as extinction-resistant neurons if they displayed a significant z -score value during both post-fear conditioning and extinction sessions, independently of freezing levels. For statistical analysis, z -score comparisons were performed using the average z -score value calculated during the 40 ms after CS-onset. In cases in which shorter or longer CS-evoked activity was observed, the average z -score was calculated during the 20 ms and 80 ms after CS-onset, respectively. To identify the trial in which individual neurons changed their CS-evoked responses during fear conditioning and extinction, we applied a change point analysis algorithm³⁰. Change point analysis identifies the trial(s) exhibiting a significant change in neuronal activity or freezing behaviour relative to the preceding trials. Change points are graphically represented by a change in the slope of a plot showing the cumulative sums of the averaged and normalized z -score and freezing values. Statistical analyses were performed using paired Student's t -tests *post hoc* comparisons at the $P < 0.05$ level of significance unless indicated otherwise. Results are presented as mean \pm s.e.m.

Extracellular stimulation. To determine the connectivity of recorded neurons, we used extracellular stimulation of the mPFC and the vHip in a subset of animals. At the end of the training procedure, animals were anesthetized using urethane (1.4 g kg⁻¹), and concentric stimulating electrodes (FHC) were lowered in the mPFC (2 mm anterior to bregma; ±0.3 mm lateral to midline; and 1.6 mm to 2 mm deep from the cortical surface) and the ventral hippocampus (3.6 mm posterior to bregma; ±3.1 mm lateral to midline; and 4 mm to 4.2 mm deep from the cortical surface). During the experiments, the stimulation electrodes were advanced in steps of 5 μ m by a motorized micromanipulator (David Kopf Instruments), and BA-evoked responses were recorded. Stimulation-induced and spontaneous spikes were sorted using principal component analysis and template matching. The similarity of stimulation-induced spike waveforms was quantitatively compared to the waveforms of units previously identified in the awake animal and recorded on the same wire using correlation analysis (Supplementary Fig. 9). To be classified as antidromic, evoked responses had to meet at least two out of three criteria: (1) stable latency (<0.3 ms jitter), (2)

collision with spontaneously occurring spikes, and (3) the ability to follow high-frequency stimulation (200 Hz). At the end of the experiments, stimulating sites were marked with electrolytic lesions before perfusion, and electrode locations were reconstructed with standard histological techniques. For each stimulation site, orthodromic and antidromic response probabilities of fear and extinction neurons were analysed using binomial statistics, with $P < 0.05$ indicating non-random connectivity.

Muscimol iontophoresis. Muscimol micro-iontophoresis injection was performed in chronically implanted animals. Single-barrel micropipettes with a tip diameter of 10 to 15 μm were cut at 1 cm length and filled with a solution containing muscimol covalently coupled to a fluorophore (Muscimol-Bodipy-TMR conjugated, Invitrogen; 5 mM in phosphate buffered saline (PBS) 0.1 M, DMSO 40%) or with bodipy alone (Invitrogen; 5 mM in PBS 0.1 M, DMSO 40%). Mice were bilaterally implanted at the following coordinates (according to ref. 50): 1.7 mm posterior to bregma; ± 3.1 mm lateral to midline; and 4 mm to 4.3 mm deep from the cortical surface. Chlorided silver wires were inserted in each micropipette and attached to a connector. A third silver wire screwed onto the skull and attached to the connector served as a reference electrode. The entire miniature was secured using cyanoacrylate adhesive gel. After surgery, mice were allowed to recover for 2 days. On the injection day, iontophoretic applications were performed by means of cationic current (+12 μA to +15 μA) for 15 min per side using a precision current source device (Stoelting). Mice were submitted to the behavioural procedure 5 min after the end of iontophoretic injections and were immediately perfused at the end of the experiments. Brains were collected for further histological analysis. Serial slices containing the amygdala were imaged at $\times 5$ using an epifluorescence stereo microscope (Leica), and the location and the extent of the injections were controlled. Mice were included in the analysis only if they presented a bilateral injection targeting exclusively the BA and if the targeted injections cover at least 25% of the BA. Statistical analyses were performed using paired and unpaired Student's *t*-tests *post hoc* comparisons at the $P < 0.05$ level of significance. Results are presented as mean \pm s.e.m.

Immunohistochemistry. Mice were transcardially perfused with ice-cold 4% paraformaldehyde in 0.1 M PBS 120 min after the onset of the training session²⁶. Brains were prepared for immunohistochemistry using primary polyclonal rabbit anti-c-Fos antibody (Calbiochem; anti-c-Fos, Ab-5, 4-17, rabbit pAb, PC38; 1:20,000 dilution). A fluorescent-dye-coupled goat anti-rabbit antibody (Invitrogen; Alexa-Fluor 633; 1:1,000 in PBS) was used as secondary antibody. Stained slices were imaged at $\times 40$ using an LSM 510 Meta confocal microscope (Carl Zeiss Inc.). Quantitative analysis of c-Fos-positive nuclei was performed using a computerized image analysis system (Imaris 4.2, Bitplane). Structures were defined according to ref. 50. Immunoreactive neurons were counted bilaterally using a minimum of three sections per hemisphere per animal. Statistical analyses were performed using unpaired Student's *t*-tests at the $P < 0.05$ level of significance. Results are presented as mean \pm s.e.m.

50. Franklin, K. J. B. & Paxinos, G. *The Mouse Brain in Stereotaxic Coordinates* (Academic, 1997).

Switching on and off fear by distinct neuronal circuits

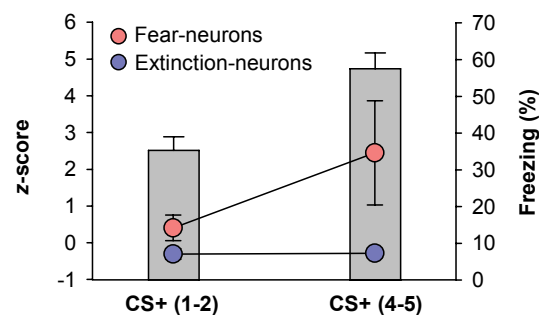
Cyril Herry, Stéphane Ciocchi, Verena Senn, Lynda Demmou, Christian Müller
and Andreas Lüthi

Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland

	Habituation			Post-FC			Extinction			Percent of units (%)
	z-score		Spont. Freq. (Hz)	z-score		Spont. Freq. (Hz)	z-score		Spont. Freq. (Hz)	
	CS-	CS+		CS-	CS+		CS-	CS+		
Fear neurons	0.46±0.44	0.38±0.25	2.2±0.5	0.37±0.21	1.78±0.28^{***}	1.8±0.3	0.42±0.21	-0.30±0.10	1.6±0.3	16.6
Extinction neurons	0.61±0.41	0.57±0.20	1.5±0.3	-0.04±0.13	-0.17±0.14	1.1±0.3	0.51±0.18	1.71±0.35^{***}	1.5±0.3	13.5
Extinction-resistant neurons	1.49±0.37	0.84±0.19	2.3±0.5	1.29±0.24	2.64±0.45^{**}	3.6±0.7^{***}	1.54±0.30	2.49±0.44^{**}	4.2±0.8^{**}	25.5

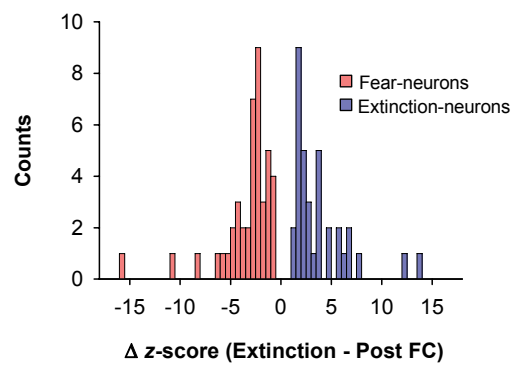
Supplementary Table 1. Summary of units recorded in BA.

This table summarizes changes in CS-induced neuronal activity (z-scores) and in spontaneous activity across behavioral sessions. Post-fear conditioning (post-FC) values were obtained using the first 4 CS⁺ presentations on day 2. Post-extinction (extinction) values were obtained using the last 4 CS⁺ presentations on day 3. Spontaneous activity was measured during the 500 ms preceding CS stimulation. Statistical comparisons: z-scores, CS⁺ vs. CS⁻ within each behavioral session; spontaneous activity, post-FC and extinction vs. habituation. **P* < 0.05, ***P* < 0.01, ****P* < 0.001.



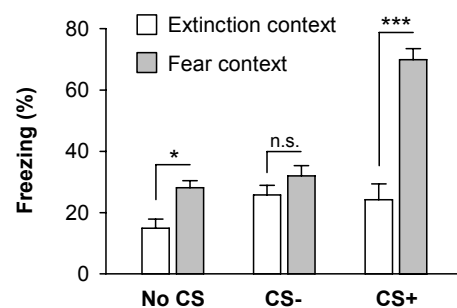
Supplementary Figure 1. Changes in CS-evoked activity during fear conditioning.

Summary graph illustrating changes in freezing behavior (grey bars), and CS-evoked activity of fear-neurons (red circles) and extinction-neurons (blue circles). Comparing the first two CSs (CS 1-2) with the last two CSs (CS 4-5) reveals that increased freezing behavior (CS 1-2: $35 \pm 4\%$ of time; CS 4-5: $58 \pm 4\%$ of time) was associated with enhanced CS-evoked activity in fear-neurons ($n = 43$ neurons from 22 mice, z-score, CS 1-2: 0.41 ± 0.35 ; CS 4-5: 2.45 ± 1.42), but not in extinction neurons ($n = 35$ neurons from 20 mice, z-score, CS 1-2: -0.31 ± 0.15 ; CS 4-5: -0.29 ± 0.11).



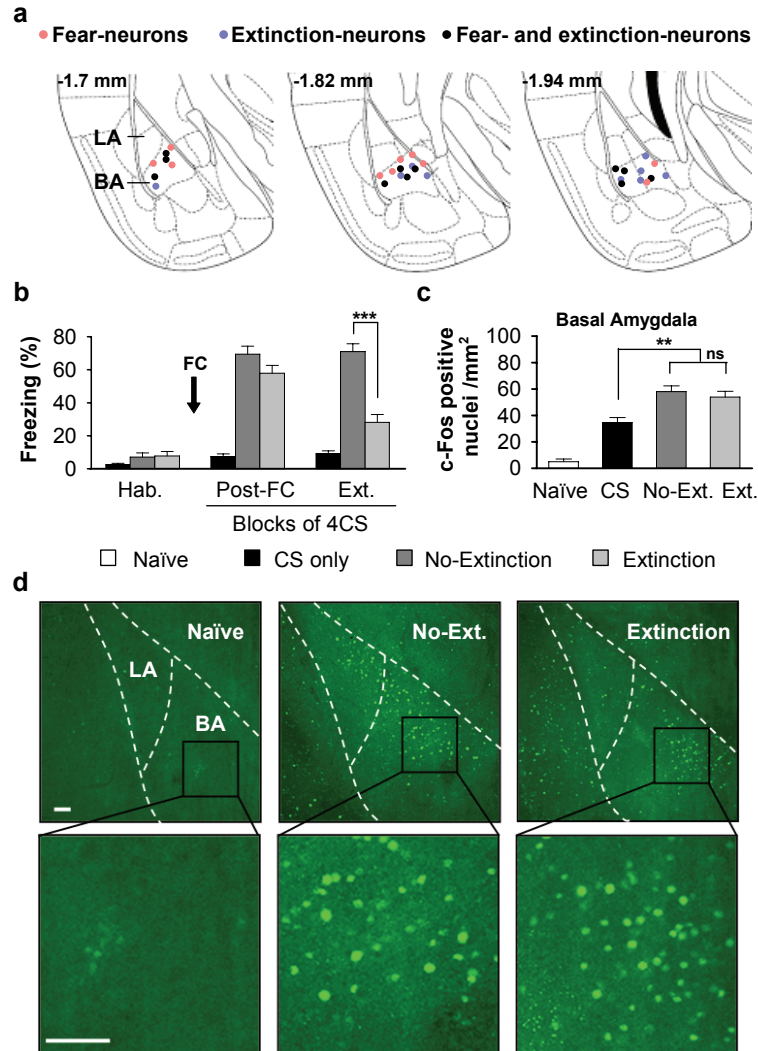
Supplementary Figure 2. Extinction-induced changes in CS-evoked activity reveal a bimodal distribution of fear- and extinction-neurons.

Histogram representing the extinction-induced changes in the CS⁺-evoked neuronal activity (z-score) of individual fear-neurons ($n = 43$) and extinction-neurons ($n = 35$). A negative Δz -score value indicates a preferential activation after fear conditioning, whereas a positive Δz -score value indicates a preferential activation after extinction. Fear- and extinction-neurons formed two well-separated populations.



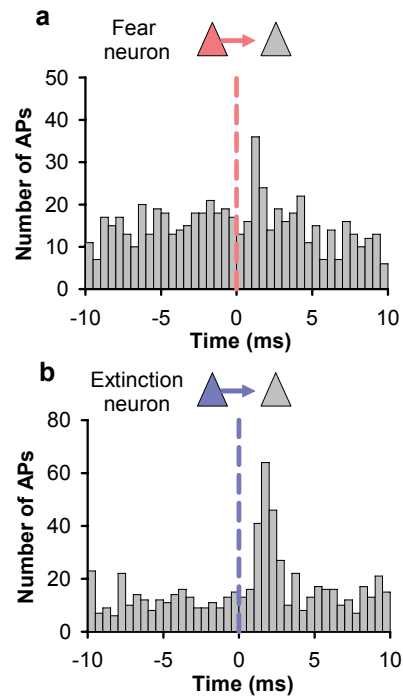
Supplementary Figure 3. Context-dependent freezing during fear renewal.

Seven days after extinction, mice ($n = 15$ animals) were exposed to the CS⁺ and to the CS⁻ in the extinction context and in the context in which fear conditioning took place. In the extinction context, both the CS⁺ and the CS⁻ elicited low freezing behavior (CS⁻: $26 \pm 3\%$ of time; CS⁺: $32 \pm 3\%$, $P = 0.128$ vs. CS⁻, $P = 0.513$ vs. extinction; same data as shown in figure 4). In the fear conditioning context, mice exhibited a modest, but significant increase in baseline freezing levels due to contextual fear conditioning (extinction context: $15 \pm 3\%$ of time; fear conditioning context: $28 \pm 2\%$, $P < 0.05$), which was not significantly different from CS⁻-induced freezing. In this context, exposure to the CS⁺ evoked significantly more freezing than CS⁻ stimulation (CS⁻: $24 \pm 5\%$ of time; CS⁺: $70 \pm 4\%$, $P < 0.01$ vs. CS⁻, $P < 0.001$ vs. extinction recall; same data as shown in figure 4). * $P < 0.05$, *** $P < 0.001$.



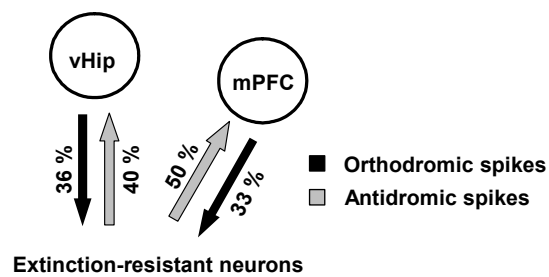
Supplementary Figure 4. Fear and extinction neurons are intermingled within BA.

a, Coronal sections through the rostrocaudal extent of the amygdala showing the location of the recording wires in the BA from which activity of fear and extinction neuron was recorded. BA: basal nucleus of the amygdala; LA: lateral nucleus of the amygdala. **b**, Naïve mice ($n = 7$) and control animals ($n = 21$) exposed to the CS and to the context exhibited low freezing levels throughout the experiment. Fear conditioned animals showed high freezing levels at both time points. In mice subjected to extinction training, freezing levels were significantly reduced (Day 3, no-extinction: $71 \pm 5\%$ of time, $n = 16$; extinction: $28 \pm 5\%$ of time, $n = 13$, $P < 0.001$, two-tailed unpaired t -test). **c**, Averaged data illustrating that even though freezing behavior was significantly different, equal numbers of c-Fos expressing neurons were detected in the BA of mice exposed to an extinguished or to a non-extinguished CS (No-extinction: 58 ± 5 cells per mm^2 ; extinction: 54 ± 4 cells per mm^2 , $P = 0.533$; two-tailed unpaired t -test). **d**, Examples of c-Fos expression in BA neurons of a naïve, non-extinguished and extinguished mice. $**P < 0.01$, $***P < 0.001$, scale bar $100 \mu\text{m}$



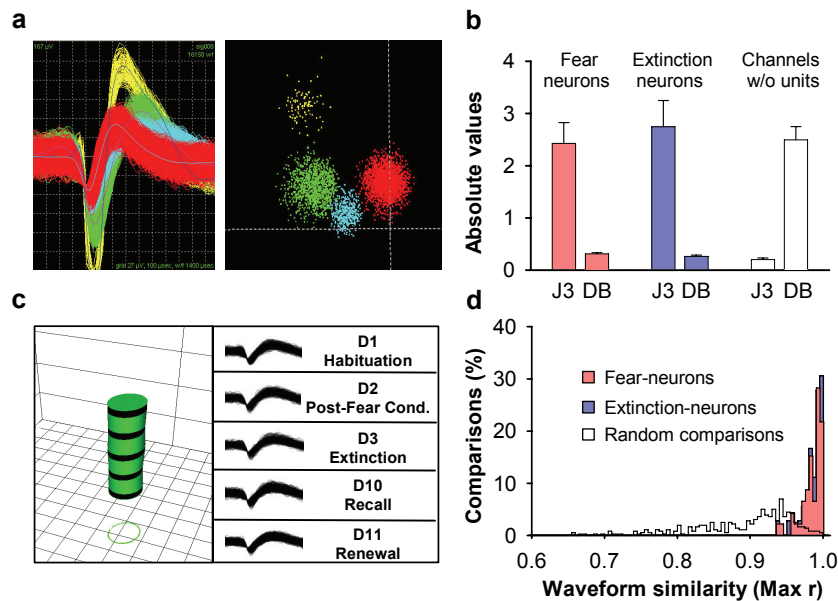
Supplementary Figure 5. Cross-correlation analysis.

Consistent with the extracellular stimulation experiments, analysis of cross-correlations between identified fear- or extinction-neurons and neighboring BA neurons indicate that fear- and extinction-neurons are projection neurons. **a**, Cross-correlation between a fear-neuron and a non-identified neuron showing a short-latency, monosynaptic, excitatory interaction. Reference event is the spike of the fear neuron (dotted line at time 0). **b**, Cross-correlation between an extinction-neuron and a non-identified neuron showing a short-latency, monosynaptic, excitatory interaction. Reference event is the spike of the extinction neuron (dotted line at time 0).



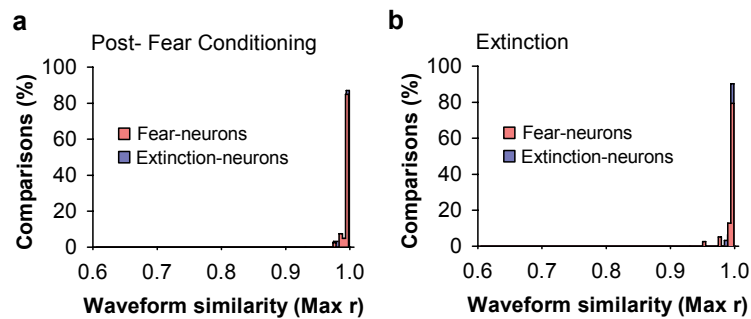
Supplementary Figure 6. Connectivity of extinction-resistant neurons.

Extinction-resistant neurons are reciprocally connected to the mPFC (orthodromic responses: 3 out of 9 stimulated neurons; antidromic responses: 6 out of 12 neurons) and to the hippocampus (orthodromic responses: 4 out of 11 stimulated neurons; antidromic responses: 2 out of 5 neurons). The graph depicts the percentage of all stimulation experiments in which a particular response was observed in identified extinction-resistant neurons.



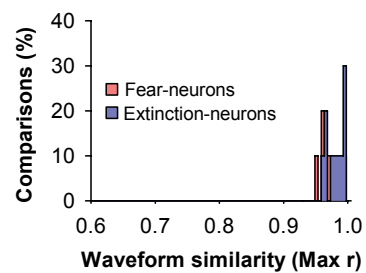
Supplementary Figure 7. Stability of chronic single unit recordings from mouse amygdala.

a, Top left: Superimposed waveforms recorded from four different units. Top right: Spikes originating from individual units were sorted using 3D-principal component analysis. **b**, Quantitative J3 and Davies Bouldin validity index (DB) statistics calculated for fear and extinction neurons. Controls values were obtained using two clusters defined from the centered cloud of points from channels in which no units could be detected. High values for the J3 and low values for the DB are indicative of good single unit isolation. **c**, Left: Stability of clustered waveforms during long-term recordings was assessed by calculating principal component (PC) space cylinders. Straight cylinders suggest that the same set of single units was recorded during the entire training session. Right: Superimposed waveforms used to calculate the PC space cylinder recorded before habituation, extinction, recall and renewal sessions. **d**, In addition, to quantitatively evaluate similarity of different spike shapes recorded on different days, linear correlation values between time-shifted average waveforms were calculated for fear and extinction neurons. As a control we computed the r values from average waveforms of different neurons. The maximum r value across time shifts was used to quantify similarity ($r = 1$ would indicate identical spike shapes). These calculations revealed that 94.4% of extinction neurons and 95.65% of fear neurons had an r value above 0.95, compared with only 17.9% of similarity scores calculated between waveforms of different cells.



Supplementary Figure 8. Quantitative comparisons of waveforms across periods of spontaneous activity and sensory stimulation.

a, For each identified fear- and extinction-neuron we calculated linear correlation values between time-shifted average waveforms obtained during a 60 s period of spontaneous activity recorded prior to each behavioral session and during CS stimulation. The maximum r value across time shifts was used to quantify similarity ($r = 1$ would indicate identical spike shapes). These calculations revealed r values above 0.95 for 100% of all units. **b**, Same plot for all units recorded before and during the extinction session.



Supplementary Figure 9. Identification of units activated by extracellular stimulation.

Similarity of stimulation-induced spike waveforms was quantitatively compared to the waveforms of fear- and extinction-neurons previously identified in the awake animal and recorded on the same wire using correlation analysis. For each unit we calculated linear correlation values between time-shifted average waveforms obtained during the extinction session and during extracellular stimulation in the anaesthetized animal. The maximum r value across time shifts was used to quantify similarity ($r = 1$ would indicate identical spike shapes). These calculations revealed r values above 0.95 for 100% of all units.