HHS Public Access

Author manuscript

Cell. Author manuscript; available in PMC 2019 May 31.

Published in final edited form as:

Cell. 2018 May 31; 173(6): 1329–1342.e18. doi:10.1016/j.cell.2018.04.004.

Corticoamygdala transfer of socially-derived information gates observational learning

Stephen A. Allsop[1,*], **Romy Wichmann**[1,*], **Fergil Mills**[1,*], **Anthony Burgos-Robles**[1,*], **Chia-Jung Chang**[1], **Ada C. Felix-Ortiz**[1], **Alienor Vienne**[1], **Anna Beyeler**[1], **Ehsan M. Izadmehr**[1], **Gordon Glober**[1], **Meghan I. Cum**[1], **Johanna Stergiadou**[1], **Kavitha K. Anandalingam**[1], **Kathryn Farris**[1], **Praneeth Namburi**[1], **Christopher A. Leppla**[1], **Javier C. Weddington**[1], **Edward H. Nieh**[1], **Anne C. Smith**[2], **Demba Ba**[1], **Emery H. Brown**[1], and **Kay M. Tye**[1,3]

[1]The Picower Institute for Learning and Memory, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[2]Evelyn F. McKnight Brain Institute, University of Arizona, Tucson, AZ 85724

Summary

Observational learning is a powerful survival tool, allowing individuals to learn about threat-predictive stimuli without directly experiencing the pairing of the predictive cue and punishment. This ability has been linked to the anterior cingulate cortex (ACC) and the basolateral amygdala (BLA). To investigate how information is encoded and transmitted through this circuit, we performed electrophysiological recordings in mice observing a demonstrator mouse undergo associative fear conditioning, and found that BLA-projecting ACC (ACC→BLA) neurons preferentially encode socially-derived aversive cue information. Inhibition of ACC→BLA alters real-time amygdala representation of the aversive cue during observational conditioning. Selective inhibition of the ACC→BLA projection impaired acquisition, but not expression, of observational fear conditioning. Together, we show that information derived from observation about the aversive value of the cue is transmitted from the ACC to the BLA and that this routing of information is critically instructive for observational fear conditioning.

In Brief

[3]Corresponding Author/Lead Contact: Kay M. Tye, PhD, Picower Institute for Learning and Memory, Department of Brain and Cognitive Sciences, 77 Massachusetts Ave, Bldg-Rm 46-6263, Massachusetts Institute of Technology, Cambridge, MA 01239, kaytye@mit.edu, twitter: @kaymtye.

*These authors contributed equally.

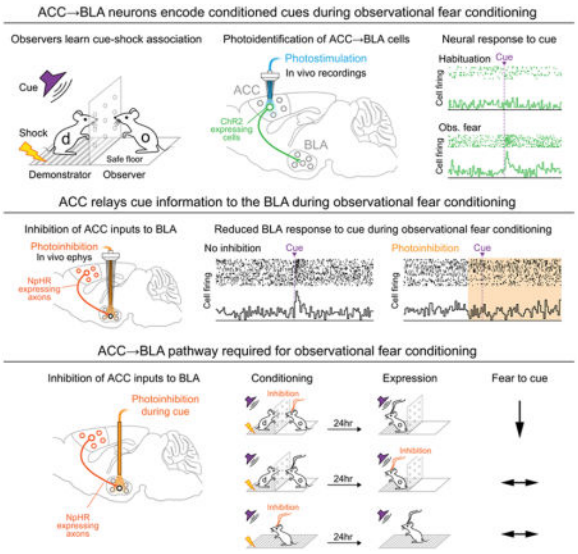
Declaration of Interests

The authors declare no competing interests.

Author Contributions

Conceptualization, SAA and KMT; Methodology, KMT, CJC and SAA; Investigation, SAA, ACFO, FM, ABR, RW, AV, CAL, EMI, GG, MIC, KF, JS and AB; Formal Analysis, SAA, EHN, DB, ACS, PN, CAL, EHB, CJC, ABR, RW, FM, EMI, GG, KA, KF, JCW, MIC and KMT. Writing – Original Draft, SSA, FM, RW and KMT; Writing – Review & Editing, SAA, KMT, RW, FM, ABR, CJC, ACFO, DB, ACS, and EHB. Visualization, SAA, KMT, RW, FM, ABR; Supervision, KMT, SAA, RW, FM, and ABR; Funding Acquisition, KMT.

For one individual to watch another's experience and learn from it, signals need to move from cortical neurons to the basolateral amygdala during detection and integration of the necessary social cues.



Introduction

Animals use direct sensory experiences to learn about aversive stimuli and the novel cues that predict them. However, learning aversive associations through direct experience (e.g., smells or colors predicting predation or poisonous food) can be life-threatening. Thus, the ability to learn through observing the experiences of others and to extract predictive information about potential threats is critical to evolutionary fitness.

Observational learning broadly describes any type of learning aided by observation of another individual. For humans, observational learning represents a critical means by which we learn about the world (Baeyens et al., 1996; Heyes and Dawson, 1990; Hopper et al., 2008; Meltzoff and Moore, 1977) and this highly conserved innate learning ability may form the basis for more complex behaviors such as empathy and altruism (Bastiaansen et al., 2009; Panksepp and Lahvis, 2011; Preston and de Waal, 2002).

Observational learning, emotional contagion, and other related behaviors have also been demonstrated experimentally in other animal species, including rodents. For example, rodents display defensive behaviors when in the presence of conspecifics undergoing aversive experiences (Atsak et al., 2011; Chen et al., 2009; Church, 1959; Kim et al., 2010; Pereira et al., 2012), and mice that observe demonstrators undergoing negative experiences show increased depression-like and anxiety-like behaviors (Warren et al., 2013). In addition, rodents are also capable of observational fear learning; the ability to learn about novel stimuli in the environment that are predictive of aversive consequences without directly experiencing them (Bruchey et al., 2010; Guzmán et al., 2009; Jeon et al., 2010; Kim et al., 2012; Twining et al., 2017; Yusufishaq and Rosenkranz, 2013).

Here, we define observational fear conditioning as the acquisition of an association between a conditioned stimulus and a punishment that may have each been directly experienced, but never temporally-paired in a contingent manner except through observation of another animal. This learning process is dependent on the detection and integration of social signals in order to adaptively change behavior (Bruchey et al., 2010; Chen et al., 2009; Dulac and Torello, 2003; Guzmán et al., 2009; Isogai et al., 2011; Twining et al., 2017; Yusufishaq and Rosenkranz, 2013).

In rodents, non-human primates, and humans, the ACC and the amygdala have been implicated in observational fear learning and social cognition (Adolphs et al., 1994; Chang et al., 2013, 2015; Haroush and Williams, 2015; Jeon et al., 2010; Kim et al., 2012; Olsson et al., 2007). In humans, both brain regions are recruited when subjects acquire fear responses to a novel cue through observation (Olsson et al., 2007). Likewise, mice show increased theta frequency synchronization between the ACC and the BLA during observational fear learning (Jeon et al., 2010), and pharmacological inhibition of either region inhibits observational learning (Jeon et al., 2010; Kim et al., 2012). The ACC and BLA also form reciprocal connections with each other (Bissiere et al., 2008; Cassell and Wright, 1986; Gabbott et al., 2005), raising the possibility that they work together during observational fear learning.

Nevertheless, although the ACC and BLA have been implicated in the acquisition of a fear response through observation, many questions remain open. For instance, what is the actual function of ACC→BLA pathways during observational fear learning? How do the ensemble dynamics in the ACC and BLA change across observational conditioning? What is the directionality of information flow during observational conditioning? Is the direct projection from either the ACC→BLA or BLA→ACC necessary for these functions?

To address these questions, we performed *in vivo* electrophysiological recordings of single-unit activity in the ACC or BLA of mice observing a demonstrator undergoing a fear-conditioning paradigm in which a cue predicted electric shock. Here, we show that neural populations in both the ACC and BLA of observer mice exhibit robust response correlates of observational fear acquisition. We also characterize neural ensemble dynamics across observational conditioning using state-space and neural trajectory analyses. We then utilize optogenetic-mediated photoidentification to demonstrate that ACC neurons directly projecting to the BLA exhibit an enhanced representation of cue information during observational learning. In addition, we show that a subset of BLA neurons require ACC input in order to encode conditioned cue information and that the ACC→BLA pathway is necessary for the acquisition of observational, but not classical, fear conditioning. Lastly, we demonstrate that the ACC→BLA pathway is also necessary for other ethologically-relevant social behaviors.

Together, our data support a model wherein ACC neurons represent socially-derived aversive cue information. This information is then transmitted to the BLA, where the association between the cue and the demonstrator's distress response is formed, thus endowing the cue with its predictive value without direct experience of the cue-shock pairing.

RESULTS

Defining the experimental parameters in observational fear conditioning

To study observational fear conditioning in a mouse model optimized for *in vivo* single-unit recordings, we needed a trial-based structure to compare neural responses to a cue across learning trials. We designed a behavioral paradigm in which a mouse observes a familiar demonstrator mouse undergo classical cued fear conditioning through a transparent, perforated divider allowing for the observation of auditory, visual, and olfactory information (Figure 1A). Mice termed "Experienced Observers" (EO) first explored the shock floor side of the chamber, received a single shock, and were immediately placed on the plastic "safe" floor side of the chamber. The "demonstrator" was then placed on the shock floor side and 30 cue-shock pairings were delivered. 24 hours later, observers were placed on the shock floor side to be tested for observational fear conditioning by measuring freezing responses to the cue in the absence of shock or demonstrator. EO mice demonstrated increased freezing in response to the cue during observational conditioning as well as on test day (Figure 1B–1D, S1A–S1B).

To confirm that conditioned responses by EO mice were due to the predictive value of the cue, we also examined an "Experienced Unpaired" (EU) group, in which shock-experienced mice observed demonstrators receiving shocks explicitly unpaired to the cue (Figure 1B). EU mice did not display conditioned freezing in response to the cue during conditioning or the test day (Figure 1B–1D, S1A–S1B). We also examined an "Experienced Solo" (ES) group, which received an initial shock experience but then observed the delivery of cues and shocks to an empty chamber (i.e., no demonstrator present). In contrast to EO mice, ES mice did not demonstrate conditioned responses to the cue during conditioning or the test day (Figure 1B–1D, S1A–S1B), confirming that the learning displayed by EO mice was not driven by nonspecific sensitization induced by the prior shock (Poulos et al., 2015).

We also examined a "Naïve Observer" (NO) group, which never directly experienced the shock, but observed demonstrator mice receive cue-shock pairings. Similar to the initial EO group, NO mice showed increased freezing to the cue during conditioning (Figure 1B–1C, S1A). However, NO mice did not show significantly higher freezing on test day than "Naïve Solo" (NS) mice, which never experienced the shock, observed cues and shocks been delivered to an empty chamber (i.e., no demonstrator present), and NO mice did not display cue-elicited freezing during conditioning or test (Figure 1B–1D, S1A–S1B). In addition to conditioned freezing responses, EO and NO mice also mimicked demonstrator mice displaying distress-related behaviors or interruption of grooming (Figure S1C, Supplementary Movie 1–2).

To confirm that EO mice learned the predictive value of the cue in a manner independent of contextual conditioning, we also examined mice that underwent observational conditioning and were then tested 24 hours later in a novel context (Figure 1E). These mice showed significantly greater freezing during the cue relative to baseline (Figure 1F), indicating context-independent memory of the cue-shock association. We also examined avoidance of the shock floor in the EO, ES, NO and NS groups before and after observational conditioning (Figure S1D–F), and found that EO mice showed increased avoidance of the

shock floor while NO mice did not. Additionally, ES mice showed shock floor avoidance, demonstrating that shock experience can drive place avoidance but is not sufficient to drive freezing during the cue as observed in the EO group (Figure S1D–S1F). Finally, we found that EO and NO, but not ES, mice showed increased affiliative social interactions with demonstrators after observational fear conditioning (Figure 1G–H). This suggested the increase in interaction was driven by the observation of distress in the demonstrator, which may reflect empathic processes (Burkett et al., 2016; Pisansky et al., 2017) or the social buffering of stress previously observed in rodents (Beery and Kaufer, 2015; Taylor, 1981).

Characterization of ACC and BLA neural dynamics during observational learning

We then performed single-unit recordings in EO (Paired) and EU (Unpaired) mice to examine how ACC and BLA neurons encode information during observational conditioning (Figure 2A, S1G–S1I). To allow for comparison of responses to the cue before, during, and after conditioning we added an initial block of trials (habituation) where cues were delivered in the absence of shock delivery prior to conditioning (Figure 2A). We examined responses in ACC and BLA neurons (Figure 2B–C, S2B) and classified neurons based on the direction and magnitude of responses to the cue during observational conditioning (Figure 2D,E). Surprisingly, the proportion of cue-responsive ACC neurons was similar in Paired (47.5%, n=112/236 neurons; N=16 mice) and Unpaired (46.7%, n=56/120; N=7) groups (Figure 2F), and the proportion of these cue-responsive neurons that showed conditioning-dependent cue responses was not detectably different between groups (Paired group: 41.1%, n=46/112; Unpaired group: 35.7%, n=20/56). However, we found that a significantly greater proportion of ACC neurons showed a potentiated change in response magnitude in the Paired group (74%, n=34/46) relative to the Unpaired group (40%, n=12/20).

In contrast, the proportion of BLA neurons that showed conditioning-dependent changes was significantly greater in the Paired group as compared to the Unpaired group (Figure 2G). While proportions of cue-responsive BLA neurons were again similar in the Paired (42.7%, n=63/110; N=7) and Unpaired group (36.9%, n=38/103; N=7), the proportion of neurons that showed conditioning-dependent changes in the BLA Paired group (38.3%, n=18/47) was significantly greater than the BLA Unpaired group (5.3%, n=2/38).

Taken together, BLA encoding of the cue across conditioning was tightly regulated by the predictive value of the cue. The ACC, however, showed similar proportions of neurons that encoded the cue across conditioning regardless of the temporal contiguity of the cue and shock to the demonstrator. Instead, differences in predictive value of the cue were reflected by changes in response magnitude (Figure 2F) and baseline firing rate changes (Figure S2F) in ACC neurons.

Temporal dynamics of neural ensembles across observational conditioning

To further characterize the changes in firing rate that occurred across observational conditioning, we used two different state-space approaches to reveal the precise temporal resolution of the neural dynamics in ACC and BLA both across (Figure 3A–D) and within (Figure 3F–I) trials. State-space models provide a general framework for solving a broad range of problems in dynamic systems (Churchland et al., 2012; Czanner et al., 2008; Smith

and Brown, 2003; Smith et al., 2010) including complex cognitive processes in the prefrontal cortex (Mante et al., 2013). Applied to neural data, it provides a means to estimate the underlying firing rate of neurons more accurately than other conventional approaches, and identify changes in firing rate across trials in a statistically robust manner (Barbieri et al., 2004; Czanner et al., 2008; Suzuki and Brown, 2005).

We applied this analysis to the firing rate of neurons which showed changes in cue responses that occurred during conditioning, which provided a probabilistic estimate of the trial at which neurons began to encode the learned meaning of the cue ("rate change trial" of the neuron) (Figure 3A–B, S3A–B) (Smith et al., 2010). The rate change trial was defined as the first trial during observational conditioning in which that trial and the subsequent trial had a >95% probability of having a greater or lesser firing rate (depending on whether the neuron was excited or inhibited to the cue, respectively) than the average "habituation" firing rate. We found that the ACC had a significantly earlier distribution of rate change trials when compared to the BLA (Figure 3C) and an earlier average rate change trial than the BLA (Figure 3D), which suggested that socially-derived information about the cue is first encoded in the ACC. This result led us to hypothesize that ACC cue responses underlie learning and that transmission of this information to the BLA could underpin observational fear conditioning. We also compared behavior for all mice in which we performed recordings from the ACC or BLA in Paired and Unpaired groups before and during the cue presentation (Figure 3E).

To further explore the neural dynamics within a given trial type in an unbiased manner, we used a second state-space approach wherein we used the entire ensemble (all neurons recorded during the session irrespective of any response characteristics) during baseline and cue presentations and plotted this ensemble activity in a state-space trajectory in reduced dimensionality space (Churchland et al., 2012; Cunningham and Yu, 2014; Mante et al., 2013). We then examined the neural trajectories in ACC and BLA for Paired and Unpaired groups during habituation and observational conditioning (Figure 3F–G). Qualitatively, ACC Paired and Unpaired trajectories shared a similar shape in habituation and observational conditioning trials, but these trajectories were repositioned in state-space (Figure 3F). We compared the distance between the habituation and observational conditioning trajectories in the Paired and Unpaired groups as a time series across these trial types, and found they were similar across the trial and did not change from baseline to cue (Figure 3H). This suggested that baseline changes could be driving the results seen in our rate-change trial analysis (Figure 3C–D, S2F).

In contrast, BLA trajectories revealed a distinct pattern of the trajectory for the neural activity in the BLA Paired, but not Unpaired, groups for observational conditioning trials, relative to the habituation trials (Figure 3G). The BLA Paired group had a greater change in the distance across trial types from baseline to cue presentation than did the trajectories from other groups (Figure 3I). Consistent with these results, the variance of baseline firing in ACC was significantly greater than in the BLA, for both Paired and Unpaired groups (Figure S3C).

To assess what memories were stored in the ACC and BLA, we also examined neural activity during the Test day, when no demonstrator was present and no shocks were delivered (Figure S2D). The proportions of ACC neurons that were cue-responsive in the Paired (30%) and Unpaired (29%) groups were similar (Figure S2D), but the proportion of BLA neurons that were cue-responsive on Test day was significantly greater in the Paired group (31%) than in the Unpaired group (17%).

Together, these data prompted us to revise our initial hypothesis that the ACC neurons were representing the association between the cue and shock to demonstrator earlier than BLA neurons. Instead, the baseline changes suggested that ACC neurons are rapidly changing basal firing rates in the face of demonstrator distress, thereby potentiating responses to salient stimuli, but that the acquisition of the predictive value of the cue is mediated by neurons in the BLA. Thus, ACC neurons transmit socially-derived information to the BLA *during* the cue presentation, allowing the BLA to form an association between the cue and shock to the demonstrator.

Optogenetic-mediated photoidentified ACC-BLA cells exhibit preferential cue encoding

To directly test the hypothesis that socially-derived information about the aversive cue is transmitted from the ACC to the BLA during observational conditioning, we used an intersectional viral strategy (Senn et al., 2014) to express Channelrhodopsin-2 (ChR2) in ACC neurons that monosynaptically project to the BLA (Figure 4A–B) (Burgos-Robles et al., 2017). *Ex vivo* recordings demonstrated that ChR2-expressing neurons fired action potentials in response to blue light while non-expressing neighbors showed no response (Figure 4C). Photoresponse latencies during patch-clamp recordings allowed us to establish a photoidentification latency threshold (Beyeler et al., 2016; Nieh et al., 2015) of 8 ms for ACC→BLA projectors (Figure 4C–D). We then performed *in vivo* recordings from ACC neurons during observational fear conditioning and subsequently photoidentified ACC neurons projecting to the BLA. (Figure 4E). Within the ACC, neurons were categorized as follows: short latency responses (<8ms; termed “ACC→BLA photoidentified neurons”), long-latency excitations (20–110ms) or inhibitions (termed *ACC→BLA excited* or *inhibited* network neurons, respectively), or no photoresponse (non-network neurons; Figure 4F, S4A–B). We found that of ACC→BLA photoidentified neurons, 62.5% showed phasic excitations while 0% were inhibited by the cue. In contrast, only 38% of non-network neurons responded to the cue (Figure 4G). Furthermore, 79% of *ACC→BLA excited* network neurons were excited in response to the cue while only 30% of *ACC→BLA inhibited* network neurons were excited to the cue (Figure 4G). Moreover, we found that the population z-score of cue responses of ACC→BLA photoidentified neurons was significantly greater than that of neurons outside the network (Figure 4H–J). The preferential cue-encoding within the ACC→BLA photoidentified population and its concomitant network provided further evidence suggesting that transfer of socially-derived cue information from the ACC to the BLA might be necessary for observational learning.

BLA encoding of cue information is dependent on ACC input during observational conditioning

If ACC-BLA information transfer is indeed necessary for observational learning, then one would expect a subset of cue responses in the BLA to be dependent on ACC input. To test this, we examined the effects of selectively inhibiting ACC inputs to the BLA during observational conditioning. We expressed halorhodopsin (NpHR) in ACC neurons of observer mice, placed an optrode in the BLA (Figure 5A, S5D–E), and inhibited ACC inputs to the BLA in a pseudo-random, interleaved subset of cue presentations during observational conditioning (Figure 5B). We found that BLA neurons showed differences in their cue responses when ACC inputs were inhibited (Figure 5C–5D). The population z-score of all neurons that were excited or inhibited to the cue was significantly attenuated by photoinhibition of ACC inputs (Figure 5E–F). Significantly fewer neurons were cue-responsive during the trials with inhibition of ACC input (29%, n=28/98) relative to trials without light (44%, n=43/98) (Figure 5G). This effect was not due to non-specific effects of illumination as the proportion of BLA neurons whose activity was modulated by photoinhibition of ACC input was significantly smaller during the baseline period (7%, n=7/98) than during the cue (28%, n=27/98) (Figure 5H), suggesting that ACC modulation of BLA activity is heightened during the cue.

ACC input to the BLA during conditioning is necessary for observational learning

The data from our photoinhibition experiments demonstrated that input from the ACC was required for BLA neurons to encode the cue during observational fear conditioning. We further predicted that the ACC→BLA projection underlies the acquisition of observational fear conditioning and that inhibiting this input during acquisition would lead to impairment in observational learning. To test this hypothesis, we expressed NpHR bilaterally in the ACC of observer mice and placed optical fibers over the BLA in both hemispheres (Figures 6A, S5, S6A–F) to inhibit ACC input to the BLA during cue presentations throughout observational fear conditioning (Figure 6B). Mice receiving this manipulation showed no differences in freezing during the conditioning session (Figure 6C, **left**), but when tested 24 hours later, in the absence of photoinhibition or a demonstrator, they showed significantly less cue-specific freezing as compared to eYFP control mice (Figure 6C, **right**). To assess whether these differences were due to deficits in attention, we analyzed the orientation and startle responses of mice and found no differences between NpHR and eYFP mice (Figure S6B). However, we *did* observe a significant reduction in mimicking behavior (Figure S6C), consistent with the notion that ACC→BLA input is important for transmitting socially-derived information.

We also performed a separate experiment wherein we inhibited ACC→BLA only upon expression of observational fear learning, during cue presentations on the test day (Figure 6D), but found no effect on cued freezing (Figure 6E). Inhibition of this circuit only during the delivery of shock to the demonstrator also had no detectable effect (Figure S6D–S6F), nor did 20 Hz stimulation of the ACC→BLA projection in mice expressing ChR2 (Figure S6G–S6K and S6N–S6P). Taken together, these experiments demonstrate that the transfer of socially-obtained cue information via the ACC→BLA circuit is necessary for observational fear conditioning but not for subsequent memory expression (Figure 6A–6E).

Importantly, when ACC input to the BLA was inhibited using the same approach during a classical fear conditioning paradigm – where associations are formed through direct experience – mice showed no differences in freezing during the cue presentation when compared to control mice (Figure 6F). This suggests that the ACC→BLA pathway is only required for observational learning, but not associative learning in general.

ACC input to the BLA is important for other ethologically-relevant social behaviors

We then sought to determine whether the same ACC→BLA circuit was involved in a more ethologically-relevant task – the avoidance of aggressive and potentially threatening animals (Hultman et al., 2016). Naïve observer mice were injected with NpHR or eYFP bilaterally in the ACC and implanted with optical fibers over the BLA, and the ACC→BLA projection was inhibited while the observer witnessed the interaction of a familiar demonstrator mouse and an aggressive CD-1 mouse for two 3-minute sessions (Figure 7A). While control mice showed avoidance of the CD-1 in a 3-chamber test following two sessions of observation, NpHR mice did not show avoidance (Figure 7B–D) suggesting that the ACC→BLA projection is also necessary for acquisition of avoidance following observation of social defeat.

We also examined whether the ACC→BLA projection was important for a more general role in processing social cues. We performed a resident-intruder task, in which a juvenile male mouse was introduced into the home cage of an observer and the time spent interacting was analyzed (Figure 7E). Inhibition of ACC inputs to the BLA led to a significant impairment in social interaction (Figure 7F–7G) but did not alter novel object exploration (Figure 7I–J). Optogenetic 20Hz stimulation of the ACC→BLA circuit had no effect on social interaction, novel object exploration, or place preference when comparing ChR2-expressing and eYFP control mice (Figure S6J–S6K, S6P). Optogenetic inhibition or excitation of this circuit had no effect on anxiety-related behavior or locomotion (Figure S6L–S6O). Thus, given that inhibition of this circuit impaired social interaction, this circuit may play a necessary role in observational learning because it is specialized for processing social cues to drive behavior (Apps et al., 2016). Because these regions have reciprocal connections, we also examined the role of the BLA→ACC projection (Fig S7) and found that that inhibition of this projection did not appear to affect observational fear conditioning, though further investigation of this pathway is warranted. Together, these results demonstrate that there is a causal relationship between the transfer of socially extracted information from the ACC to the BLA and the ability of mice to learn about dangerous stimuli in their environment through observation.

Discussion

The ACC→BLA circuit encodes observational learning

Here we show that the ACC→BLA circuit plays a critical role in mediating the acquisition of aversive memory through observational conditioning. We find that during observational learning, neurons within the ACC and BLA respond to a predictive cue and shock delivery to a familiar conspecific, and that within the ACC, BLA-projecting neurons show an enhanced representation of the predictive cue. Additionally, we found that optogenetic

inhibition of the ACC→BLA projection disrupts acquisition, but not expression, of observational fear conditioning.

Neurons in both the ACC and BLA show rapid conditioned responses to the cue during observational fear learning, as has previously been seen in the amygdala, auditory cortex and ACC during classical fear conditioning (Maren, 2000; Quirk et al., 1995; Steenland et al., 2012). However, we revealed several surprising features of neural dynamics in these regions during observational conditioning. In both the ACC and BLA, the overall proportion of cue responsive neurons alone was not indicative of learning the predictive value of the cue, as it was similar for the Paired and Unpaired groups, perhaps due to latent inhibition. Rather, the proportion of neurons undergoing significant changes in firing rate across learning (and the direction of those changes) proved to be a greater differentiator between the Paired and Unpaired groups. In addition, our state-space analysis of neural trajectories showed that ACC neurons show baseline changes in the context of the demonstrator's distress and that some BLA neurons were dependent on ACC input during the cue in order to respond appropriately to the cue. These data suggest that during observational conditioning, the ACC encodes the demonstrator's distress response, thereby enabling the acquisition of the aversive value of the cue by BLA neurons and subsequent behavioral output.

ACC→BLA circuit is necessary for observational, but not classical, fear conditioning

The results from behavioral experiments in which we inhibited the ACC→BLA projection provided further support for the importance of this circuit for observational conditioning and social processing in general. We found that optogenetic inhibition of the ACC→BLA pathway during the cue in observational conditioning led to an impairment of cue-elicited freezing on the test day, though freezing during conditioning was unaffected. The lack of a behavioral effect during conditioning is likely due to the presence of the demonstrator, which could impact the observer's behavior independent of learning the predictive value of the cue. We also found that ACC→BLA inhibition on the test day did not inhibit cue-induced freezing. This suggests that ACC input to the BLA is necessary during acquisition, but not expression, of freezing behavior once the associative memory of the cue's significance is consolidated in the BLA. Our finding that inhibition of the ACC→BLA circuit had no detectable effect on classical fear conditioning indicates that when mice have direct access to an unconditioned stimulus, the associative learning of the predictive value of a cue does not require the ACC→BLA projection. We posit that when the predictive value of the cue is learned through processing of social information (distress of the demonstrator), the necessary social information is routed through the ACC before being sent to the BLA to drive conditioned fear-related behaviors.

Although some have suggested a role for the rodent ACC in classical forms of conditioning (Bissiere et al., 2008; Buchanan and Powell, 1982; Steenland et al., 2012), the ACC is a large structure and different anatomical portions of the ACC may have different functions (Allman et al., 2001; Apps et al., 2016; Bussey et al., 1996; Devinsky et al., 1995; Jones et al., 2005). Differences in the targeted regions of the ACC between our and previous studies could lead to different behavioral effects. In agreement with this, experiments targeting the portion of ACC as we did also showed no effect on classical fear conditioning (Jeon et al.,

2010; Kim et al., 2012). It is likely that manipulations of specific ACC projections may reveal different behavioral effects than non-specific lesions, as has been seen in other brain regions such as the amygdala or ventral tegmental area (VTA) (Lammel et al., 2012; Tye et al., 2011).

The role of prior experience in observational learning

Various studies in rodents have raised some debate for the role of prior experience in learning through other's experiences. Some studies have found that rodents can undergo emotional contagion or observational learning without prior experience (Bruchey et al., 2010; Chen et al., 2009; Jeon et al., 2010; Kim et al., 2012; Twining et al., 2017; Yusufishaq and Rosenkranz, 2013), while others have found that prior experience is necessary (Atsak et al., 2011; Greene, 1969; Kim et al., 2010; Pereira et al., 2012; Sanders et al., 2013). Our findings show that prior experience is an important factor in observational conditioning, as mice with prior shock experience (EO group) demonstrated more robust acquisition of an association between the cue and the shock during observational conditioning than those without shock experience (NO group).

Experience may be particularly important when the punishment delivered to the demonstrator is not visible, audible or otherwise readily observed (e.g. delivery of foot shock, which may be imperceptible to an observer). Indeed, NO group mice did not learn to avoid the shock floor after watching demonstrator mice being shocked there, suggesting that they were unable to identify the shock floor as the source of the aversive event. Prior experience with punishments from an imperceptible source may equip animals with the added advantage of not only being able to identify predictive cues, but to also avoid the source of the aversive event (Masuda and Aou, 2009; Mineka and Cook, 1993; Sanders et al., 2013). With more naturalistic aversive stimuli (such as in our observational social defeat experiment) prior experience may be less important since the source of the punishment can be readily observed or inferred from similar experiences.

Conclusion

We demonstrate that the ACC→BLA circuit plays a critical role in routing socially-acquired information about environmental stimuli necessary for observational learning. The involvement of this circuit in social interaction suggests common evolutionarily-conserved circuits are used both to perform innate social processes and to infer environmental contingencies from other animals' experiences. Understanding how these circuits underlie fundamental aspects of social cognition may provide insight into psychiatric conditions, such as autism spectrum disorders, social anxiety disorder, and schizophrenia, where social cognition and behavior are impaired.

STAR METHODS

Contact for Reagent and Resources Sharing

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Kay Tye (kaytye@mit.edu).

Experimental Model and Subject Details

Test naïve, adult (8–12 weeks) wild-type male C57BL/6J mice from the Jackson Laboratory (RRID:IMSR_JAX:000664) were used for all experiments. All mice were pair-housed in the Picower Institute on a reverse 12 hour light-dark cycle with food and water *ad libitum*. Littermates were randomly assigned to experimental groups. All experiments were conducted during their light off period and in accordance with NIH guidelines and approval of the MIT Institutional Animal Care and Use Committee and the MIT Department of Comparative Medicine.

Method Details

Stereotactic surgery procedures—All surgeries were conducted under aseptic conditions using a digital small animal stereotaxic instrument (David Kopf Instruments, Tujunga, CA, USA). Mice were anaesthetized with isoflurane (5% for induction, 1%–2.5% for maintenance). Injections were performed using a beveled 33-gauge microinjection needle. A 10-μl microsyringe (nanofil; WPI, Sarasotam FL, USA) was used to deliver virus at a rate of 0.1 μl per min using a microsyringe pump (UMP3; WPI) and controller (Micro4; WPI). Mice were given a post-surgical recovery time of at least 7 days prior to start of any experimental procedures.

Surgery for *in vivo* recordings—To target the basolateral amygdala (BLA) for *in vivo* recordings, a craniotomy was made in the right hemisphere at anteroposterior (AP) −1.6 mm and mediolateral (ML) +3.35 mm. Two to three skull screws were implanted around the site of the craniotomy. One layer of adhesive cement (C&B Metabond; Parkell, Edgewood, NY, USA) followed by cranioplastic cement (Dental cement; Ortho-Jet, Lang Dental, Wheeling, IL, USA) was used to stabilize screws to the skull. A 16-channel multi array electrode (Innovative Neurophysiology) was then lowered at approximately 0.01 mm/s to −4.75 dorsoventral (DV) as measured from bregma. A ground wire was placed in the contralateral posterior hemisphere at an approximate depth of 1mm. An additional layer of cranioplastic cement (Ortho-Jet, Lang Dental, Wheeling, IL, USA) was applied to the skull as well as around the wires. The electrode was then lowered to −4.9 DV and stabilized with additional layers of cement.

In order to record from anterior cingulate cortex (ACC) neurons in a circuit-specific manner an adeno-associated virus serotype 5 carrying a construct for expression of channelrhodopsin-2 fused to enhanced yellow fluorescent protein, under the control of a double-inverted open reading frame expressing under the EF1α promoter (AAV5- EF1α-DIO-ChR2-eYFP) (1 μl) was injected into the ACC (AP: +1.0 mm, ML: −0.3 mm, DV: −2.1 mm) and 1 μl of the retrogradely traveling canine adenovirus carrying Cre-recombinase CAV2- Cre was injected into the BLA (AP: −1.6 mm, ML: 0.35 mm, DV: −4.9 mm). 5–8 weeks later, a second surgery was performed to implant an optrode (i.e., combination of recording electrode and optical fiber for light delivery). Using the same surgical conditions and anesthesia as previously described, one craniotomy was drilled over the ACC (AP: +1.0 mm, ML: −0.3 mm). The optrode was lowered at approximately 0.01 mm/s u to −1.9 mm DV. The ground wire was implanted at a depth of approximately 1 mm into the posterior ipsilateral hemisphere. Cranioplastic cement was placed around the optrode and ground wire

and the optrode was then lowered to −2.1 mm DV. Sham surgery was performed on the observer's cage mate. A craniotomy was made and mice were left under anesthesia for an equivalent amount of time as their observer but no electrode was implanted. A small amount of cranioplastic cement was placed on the skull to cover up the craniotomy. Cage mates were reunited directly after surgery.

Surgery for optogenetic experiments—In order to inhibit ACC→BLA input, 300μl of an AAV carrying the gene for a fusion protein comprised of enhanced Halorhodopsin and enhanced yellow fluorescent protein under the calmodulin kinase II promoter (AAV_5-CaMKIIα-eNpHR3.0-eYFP) or eYFP alone was bilaterally injected into the ACC (AP: +1 mm, ML: ±0.25 mm, DV: −2.1 mm). After waiting for 5 minutes the needle was raised to −2.0 mm and another 300 μl of virus was injected at the same rate. After waiting 5 additional minutes the needle was raised to −1.9 mm for 10 minutes before being slowly withdrawn. After 4–8 weeks mice underwent a second surgery in which two optical fibers were implanted bilaterally over the BLA (AP: −1.6 mm, ML: ±3.35 mm, DV: −4.5 mm). Fibers were lowered at approximately 0.01 mm/s and were secured using a thin layer of adhesive cement followed by dental cement. In a subset of mice instead of optical fibers an optrode was implanted in the BLA following the same surgical procedure as described above.

For optogenetic activation of ACC→BLA input, we injected the same amount of an AAV carrying the gene for a fusion protein comprised of Channelrhodopsin2 and enhanced yellow fluorescent protein under the CaMKIIα promoter (AAV_5-CaMKIIα-ChR2-eYFP) unilaterally into the right ACC, and an optical fiber unilaterally over the right BLA, as described above.

Behavioral Tasks

Observational conditioning task: Mice were placed in a soundproof conditioning chamber (Med Associates, St Albans, VT, USA) with a shock floor side and a plastic floor side separated by a transparent, perforated plastic divider. Mice in the Experienced Observers (EO) group received a "shock experience" by being placed on the shock floor side of the chamber and allowed to freely explore. After 5 minutes they received 1 unpredicted, un-cued, footshock (all mice greater than 30 g were shocked with 1.5 mA, while mice less than 30 g were shocked with 1mA) and were immediately transferred to the plastic floor side of the chamber. The cage mate of the observer was then placed into the shock side of the chamber as the demonstrator for "observational conditioning". After 5 minutes of habituation, demonstrators underwent 30 trials that occurred at random intervals (60, 90, 120, 150, 180 seconds) in which a 20 second compound cue (light and 10 kHz tone) predicted the delivery of a 2 second shock (1 mA – 1.5 mA) 10 seconds after the onset of the cue. Freezing was scored during the 20s cue and a 20s baseline period prior to cue onset. Directly after, mice were placed back into their home cages. 24 hrs later, observer mice were placed back into the shock side of the chamber and 30 cues were delivered to the chamber in the absence of shock. Mice in the Unpaired Observers (UO) group received a "shock experience", however during observational conditioning cues and shocks delivered to the demonstrator were explicitly unpaired. Experienced Solo mice (ES) also received a "shock

experience" but then no demonstrator was placed into the shock side of the chamber. Paired cues and shocks were delivered to the empty side of the chamber. Naïve Observers (NO) did not receive a "shock experience" but instead were placed on the plastic side of the chamber where they observed demonstrators in the same way as EO. Lastly, Naïve Solo mice (NS) also did not receive a "shock experience" but instead were placed on the plastic side of the chamber. Paired cues and shocks were delivered to the shock floor side of the chamber in the absence of a demonstrator.

Observational place preference task: On Day 1 mice were placed in the same soundproof conditioning chamber mentioned above except that there was no barrier between the shock floor and plastic floor so mice could freely explore both. After 30 minutes mice were removed. On Day 2, a barrier was placed in the chamber separating the two floors and mice underwent observational conditioning according to their behavioral groups as described above. On Day 3, the barrier was removed and mice were again allowed to freely explore for 30 minutes. After 5 minutes of exploration, cues were played.

Observational test in novel context: On Day 1, EO mice were observationally conditioned as described above. To test for cue learning in a novel context, on Day 2 observer mice were placed in a soundproof conditioning chamber that had a plastic floor covered with fresh bedding. 20 second cues (10 KHz tone and house light) were delivered to the box and time spent freezing during the baseline (20 seconds prior to the cue) and cue was measured.

Optogenetics during observational learning: Mice that expressed halorhodopsin, channelrhodopsin or eYFP in the ACC or BLA and had optical fibers over either BLA or ACC were tethered to a cable attached to a laser and placed in the same observational conditioning chamber described above. Mice received one footshock and were then placed on the plastic side of the chamber to observe their demonstrator. 30 cue-shock pairings were then delivered to the box in the same manner described for EO. Yellow (593 nm, 10mW) or blue (473nm, 15–20mW) light was delivered one second before every cue and the laser stayed on for one second after the cue offset. 24 hrs later mice were again tethered to the laser and placed back on the shock floor and 30 cues were delivered to the chamber in the absence of shock delivery and laser stimulation. All behavioral scoring for this experiment was performed in the manner described in the Quantitative and Statistical analysis section.

Optogenetics during classical fear conditioning: Mice that expressed halorhodopsin or eYFP bilaterally in the ACC or BLA and had bilateral optical fibers over either BLA or ACC were tethered to a cable attached to a laser and placed in a behavioral chamber with a shock floor. Mice received 12 trials in which a 20 second cue (2 kHz tone and houselight) was followed by delivery of a 0.5 ms footshock. The footshock occurred 10 seconds into the cue just as in observational conditioning. One second before every cue, yellow light (593 nm, 10mW) was emitted from the laser and stayed on until 1 second after the cue. 24 hrs later, mice were placed back in the chamber and 12 cues were played in the absence of shock delivery and laser stimulation. All behavioral scoring for this experiment was performed in the manner described in the Quantitative and Statistical analysis section.

Observational social defeat: The home cage of an aggressive CD-1 mouse (Charles River, RRID:IMSR_CRL:22) was modified by placing a transparent, perforated divider in the middle. Observer mice with viral vectors allowing for the expression of NpHR or eYFP bilaterally injected into the ACC and optical fibers over the BLA were tethered to a laser that would deliver yellow light (593nm, 10mW). They were then placed into the home cage of an aggressor CD-1, separated from the aggressor by the transparent divider. The light was turned on in order to inhibit ACC input to the BLA and the demonstrator was immediately placed into the side of the home cage with the CD-1 mouse for 3 minutes. After 3 minutes, the demonstrator was removed from the home cage and the light was turned off. Approximately one hour later, the demonstrator and observer were placed back into the CD-1 cage in an identical manner for 3 minutes. Immediately following this session, the observer was placed into an empty modified 3-chamber box for 5 minutes following which the same CD-1 aggressor as well as a novel object were placed in the left and right chamber (counterbalanced between mice) of the box for 5 minutes. The time spent in the zone of the CD-1 or novel object was recorded by The EthoVision XT video tracking system (Noldus, Wageningen, Netherlands). The 3-minute social defeat sessions were recorded with a video camera.

Resident-intruder assay: Social Interaction in the homecage was examined as follows. The cagemate was temporarily moved to a holding cage and the experimental mouse was allowed to explore its homecage freely for 1 min (habituation). A novel juvenile (3–4 weeks old) male C57BL/6 mouse was then introduced into the cage and allowed to interact freely for 3 min (test session). Each experimental mouse underwent two social interaction tests separated by 24 hrs, with one intruder paired with optical stimulation and one with no stimulation. Groups were counterbalanced for order of light stimulation. All behaviors were video recorded and analyzed by an experimenter blind to the testing condition using ODLog software (Macropod software). The overall score of social interaction was defined as any period of time in which the experimental mouse was actively investigating the juvenile intruder, including behaviors such as face or body sniffing, anogenital sniffing, direct contact, and close following (<1 cm). Nonsocial behaviors were also represented in an overall exploration score, which included walking, rearing, digging, and self-grooming. For the social interaction assay described in Figure 1G, the time spent engaging in social interaction by EO, ES, or NS mice was scored in the homecage for the 15 minutes prior to observational conditioning. Directly following observational conditioning, the demonstrator and observer mice were placed back into their homecage and time spent interacting was again scored for 15 minutes.

Novel object exploration: The novel object test was executed exactly like the resident-intruder assay. Instead of a juvenile intruder, a plastic object was introduced to the mice's home cage and total time spent investigating the object over 3 min was quantified. Objects were thoroughly cleaned with acidic acid in between tests. Each experimental mouse underwent two novel object investigation tests separated by 24 hrs, with one trial paired with optical stimulation and one with no stimulation, counterbalanced for order of light stimulation and object.

Open field test: The open field chamber was made of transparent plastic (53 × 53 cm) and divided into a central and a peripheral field. Individual mice were connected to the patch cables and placed in the center of the open field at the start of the session. The open field test consisted of a 9 min session with three 3 min epochs (OFF-ON-OFF) in which the mouse was permitted to freely investigate the chamber.

Elevated plus maze assay: The elevated plus maze was made of grey plastic and consisted of two open arms (30 × 5 cm) and two enclosed arms (30 × 5 × 30 cm) extending from a central platform (5 × 5 cm). The maze was elevated 75 cm from the floor. Individual mice were connected to the patch cable and allowed 2 min on the lid of the homecage for recovery from handling before the 9 min session was initiated. Each session was divided into three 3 min epochs with only the second epoch with light stimulation (OFF-ON-OFF).

Real-time place preference: Individual mice were placed in a transparent Plexiglas arena (57.15 × 22.5 × 30.5 cm; divided into left, right, and center compartments) and were allowed to freely move between compartments for 30 min. Entry into one half of the chamber resulted in photostimulation (20 Hz, 5 ms pulses). Stimulation and no stimulation sides were counterbalanced between mice. In between subjects, the behavioral chamber was thoroughly cleaned with 0.03% acetic acid diluted in water.

All behavioral tests were recorded by a video camera. The EthoVision XT video tracking system (Noldus, Wageningen, Netherlands) was used to track mouse location, velocity, and movement of head, body, and tail. All measurements displayed are relative to the center of the mouse body.

Electrophysiological Recordings

Ex vivo recordings and analysis: Six to eight weeks after surgery for ChR2 expression in ACC→BLA projectors, 4 mice were anesthetized with 90 mg/kg pentobarbital and perfused transcardially with 10 mL of modified artificial cerebrospinal fluid (ACSF, at ~4°C) containing (in mM): 75 sucrose, 87 NaCl, 2.5 KCl, 1.3 NaH_2PO_4, 7 $MgCl_2$, 0.5 $CaCl_2$, 25 $NaHCO_3$ and 5 ascorbic acid. The brain was then extracted and glued (Roti coll 1; Carh Roth GmbH, Karlsruhe, Germany) on the platform of a semiautomatic vibrating blade microtome (VT1200; Leica, Buffalo Grove, IL). The platform was then placed in the slicing chamber containing modified ACSF at 4°C. Coronal sections of 300 μm containing the ACC and BLA were collected in a holding chamber filled with ACSF saturated with 95% O_2 and 5% CO_2, containing (in mM): 126 NaCl, 2.5 KCl, 1.25 NaH_2PO_4, 1.0 $MgCl_2$, 2.4 $CaCl_2$, 26.0 $NaHCO_3$, 10 glucose. Recordings were started 1 h after slicing and the temperature was maintained at approximately 31°C both in the holding chamber and during the recordings.

The viral injection sites were checked and imaged with a camera (Hamatsu Photonics K.K., Japan) attached to the microscope (BX51; Olympus, Center Valley, PA). The slice images were registered to the mouse brain atlas (Paxinos and Watson) and the center of the injection was taken at the brightest point of the fluorescence. If the injection site was outside the ACC, data was not collected from that mouse.

Recordings were made from visually identified neurons expressing ChR2-eYFP (ChR2+) and non-expressing (ChR2−) neighboring cells. Patched cells were filled with Alexa Fluor (AF) 350 and biocytin. Voltage and current-clamp recordings of ACC→BLA projectors were made using glass microelectrodes (5–7 MΩ) shaped with a horizontal puller (P-1000, Sutter, CA) and filled with a solution containing (in mM): 125 potassium gluconate, 20 HEPES, 10 NaCl, 3 MgATP, 8 biocytin and 2 Alexa Fluor 350 (pH 7.25-7.4; 280-290 milliosmol). Recorded signals were amplified using Multiclamp 700B amplifier (Molecular Devices, Sunnyvale, CA). Analog signals were digitized at 10 kHz using a Digidata 1440 and recorded using the pClamp10 software (Molecular Devices, Sunnyvale, CA). Oxygenated ACSF was perfused onto the slice via a peristaltic pump (Minipuls3; Gilson, Middleton, WI) at ~3 mL/min. Cells were confirmed to be expressing ChR2 based on the constant inward current response to a 1 s constant blue light pulse in voltage clamp.

Off-line analysis was performed using Clampfit software (Molecular Devices, Sunnyvale, CA). Light evoked latencies of action potentials (AP) and excitatory postsynaptic potentials (EPSP) were measured for each cell for 20 pulses of a 1 Hz train with 5 ms pulses. Latencies were measured from the onset of the light pulse to the peak of the AP or EPSP.

In vivo recordings during observational learning: ACC or BLA electrode-implanted mice were connected to a head stage that plugged into a battery-operated commutator and pre-amplifier for multichannel spike acquisition (Tucker-Davis Technologies, Alachua, FL). Electrophysiological recordings were performed during either paired or unpaired observational training (i.e., EO or EU, respectively) as described above for the initial behavioral experiments. Recordings were also performed during an initial “habituation” session shortly prior to the observational conditioning phase in which only the cue was delivered at random intervals to allow for comparisons of neural cue responses before and during observational learning.

Mice that expressed halorhodopsin bilaterally in the ACC and had an optrode implanted in the right BLA were plugged into the spike acquisition system and an optical patch cord. Mice were then placed into the observational conditioning chamber described above, received an initial shock experience, and were then transferred to the plastic side of the chamber. The demonstrator was placed on the shock floor and 15 cue-shock pairings were delivered to the chamber. The next 30 trials had a subset of trials were the laser delivered yellow light (593 nm) 1 second before the onset of the cue and stayed on until 1 second after the cue. This was done in a pseudorandom order. Overall, there were 25 trials in which no laser stimulation was delivered and 20 trials where laser stimulation was delivered. Neural activity was recorded continuously throughout the experiment.

Immunohistochemistry and confocal microscopy—All mice were anesthetized with sodium pentobarbital and then transcardially perfused with ice-cold phosphate-buffered saline (PBS) followed by 4% paraformaldehyde (PFA) in PBS (pH 7.3). Extracted brains were post-fixed in 4% PFA overnight and then transferred to 30% sucrose in PBS until equilibration. 50–60 μm-thick coronal sections were sliced using a sliding microtome (HM430: Thermo Fisher Scientific, Waltham, MA) and stored in PBS at 4°C until processed for immunohistochemistry. Free-floating sections were incubated with a DNA specific

fluorescent probe (DAPI: 4′,6-Diamidino-2-Phenylindole, 1:50,000; Sigma-Aldrich) for 1 hour at room temperature. Sections were washed for 4 × 10 min with PBS followed by mounting on microscope slides with PVA-DABCO (Sigma-Aldrich). Fluorescence images were acquired using an Olympus FV1000 confocal laser scanning microscope using a 10x/0.40 NA or a 40x/1.30 NA oil-immersion objective. Mice without viral expression or mis-targeted fiber placements were excluded from further analysis.

Quantification and Statistical Analysis

Statistical Analysis—Statistical analyses were performed using either GraphPad Prism (GraphPad Software, Inc, La Jolla, CA, USA) or MATLAB (Mathworks, Natick, MA, USA). Group comparisons were made using either one-way or two-way analysis of variance (ANOVA) followed by Bonferroni post-hoc tests. Single variable comparisons were made with two-tailed unpaired Student's t tests while chi-square analyses were used to compare proportions. Non-parametric Wilcoxon signed-rank tests were used to compare in vivo neural firing rates across conditions, using an $\alpha = 0.01$. An $\alpha = 0.01$ was also used to determine whether z-score transformed peri-stimulus time histograms of neural data exhibited significant neural responses. Multiple comparisons were corrected when appropriate by adjusting *P* values using the Bonferroni method. Differences between behavioral groups were assessed using ANOVA tests followed by post-hoc tests when applicable using an $\alpha = 0.05$. The number of animals (N) and the number of neurons (n) recorded and is specified in the figures, the figure legends, and the text. All state space analyses were performed according to Smith et al. 2010.

Analysis of Observational conditioning—Behavioral performance was recorded by digital video cameras. All videos were manually analyzed offline by an experimenter blind to experimental conditions. Freezing behavior of the observer was scored on both conditioning and test day as the amount of freezing during the cue minus the amount of freezing in the 20 seconds directly preceding the cue (baseline). Freezing was defined as absence of movement, with the exception of respiration. Based on preliminary behavioral data, analysis of freezing behavior for the training day was performed on trials 5–20. Analysis of freezing behavior on test day was performed on the first 5 trials. Additional other stereotyped behaviors shown by the observer (e.g. grooming and escaping), were quantified by manual scoring. Observer mice were said to be mimicking when they displayed escape behaviors in direct response to demonstrators exhibiting escape behaviors during the cue.

Grooming and mimicking behaviors were quantified during the 20 s of the cue and compared to the baseline taken as the 20 s prior to the cue onset. The number of trials in which mice that were engaged in grooming stopped at cue onset was quantified as a percentage of all trials. This was averaged across all mice in the respective groups. The percentage of trials in which the Observers showed mimicking behaviors was also quantified and averaged across mice in a given group.

Analysis of Observational Place Preference Task—Observational Place Preference videos of all mice were recorded then viewed and analyzed by an experimenter blind to

experimental conditions. The percentage of time mice spent with all four limbs on the shock floor side of the chamber on Day 1 was scored and calculated. The percentage of time spent on the shock floor on Day 3 was also scored and calculated. A preference score ((% time on shock floor on Day 3) - (% time on shock floor on Day1)) was calculated for each mouse. This score was then divided by the average of the naïve solo group to give a normalized preference score ((% time on shock floor Day 3) - (%time on shock floor Day 1/average NS preference score)). The normalized preference score for the group was calculated as an average of all of the normalized preference scores for mice in that group.

Behavioral analysis of electrophysiology mice—Behavior of electrode-implanted observer mice was recorded throughout observational conditioning and test day, with videos manually scored offline as described above to identify bouts of freezing, grooming and escape behaviors. Raster plots of behavior were generated for each individual mouse and presented in 0.1s bins. Raster plots of average freezing behavior for each group were calculated and presented in 1s bins, indicating the average proportion of time (0%–100%) spent freezing during each 1s bin by all mice in each group.

Analysis of neural responses to cue delivery—The response of individual neurons to the cue was examined using the non-parametric Wilcoxon signed-rank test. Two signed-rank tests were performed per neuron to capture phasic and sustained responses to the cue. For phasic cue responses, neural activity was binned in 100 ms epochs, and statistical comparisons were made between the firing frequency within a baseline window of 1000 ms prior to cue onset and the firing frequency within an experimental window of 500 ms after cue onset. For sustained cue responses, neural activity was binned in 1 s epochs, and statistical comparisons were made between the firing frequency within a baseline window of 20 s prior to cue onset and the firing frequency within an experimental window of 9 s after cue onset. Both tests were performed with 1000 bootstraps, the significance threshold was set to $P < 0.01$, and Bonferroni corrections were performed to control for multiple comparisons (i.e., 0.01 divided by 2 tests, $P < 0.005$). Neurons were deemed as cue responsive if they exhibited statistical significance on either the phasic or sustained response test. Additional Bonferroni-corrected signed-rank tests were performed to determine whether cells exhibited training-induced changes in the magnitude of cue response by comparing firing frequencies during the habituation and observational conditioning phases. Since there was a higher number of trials during the observational conditioning phase, only a subset of 16 conditioning trials were used for comparisons against 15 habituation trials. Conditioning trials 5–20 were chosen for this analysis as they corresponded to the trials that we used to measure learning behaviorally.

Analysis of neural responses during shock—The response of neurons to shock delivery to demonstrator mice was also examined using the signed-rank test. Neural activity was binned in 100 ms epochs, and statistical comparisons were made between the firing frequency within a baseline window of 5 s prior to shock onset and the firing frequency within the 2 s of shock delivery to the demonstrators. To account for potential electrical contamination during shock delivery, statistical comparisons were also made between shock delivery with and without the presence of demonstrator mice, while observer mice freely

behaved on the safe plastic floor compartment. As for cue response analysis, shock response analysis was performed with 1000 bootstraps and Bonferroni corrections for multiple comparisons.

State-Space Generalized Linear Model—We propose a state-space model to define the within- and between-trial variation in our neural spiking data. The model framework is based on Czanner et al. 2008. We assume that the spike train can be represented as a point process (Brown et al. 2003). That is a time-series of random binary (0, 1) events that occur in continuous time. We designate spikes as 1's times at which no spikes occur as 0's. Given a trial interval $(0, T]$, we define $N(t)$ as the number of spikes recorded in interval $(0, t]$, for $t \in (0, T]$, in a trial. A point process model is completely defined by its conditional intensity function, $\lambda(t|H_t)$, which is

$$\lambda(t \mid H_t) = \lim_{\Delta \to 0} \frac{\Pr\left(N(t+\Delta) - N(t) = 1 \mid H_t\right)}{\Delta}, \quad \text{(A)}$$

where H_t is the spike history up to time t. Given the history up to time t, the approximate probability of one spike in a small interval $(t, t + \Delta]$ is $\lambda(t|H_t)\Delta$. The conditional intensity function generalizes the Poisson process rate function by allowing history dependence. The conditional intensity is the basis for constructing our model.

We assume that our experiment consists of K trials of the same task. Neural spiking activity is simultaneously recorded with the task execution. We index the trials as $k = 1, \ldots, K$. We define the observation interval as $(0, T]$, where T is the length of each trial. We develop a discrete time representation of the conditional intensity function by choosing L large and dividing the time interval T into subintervals of width $\Delta = TL^{-1}$. We choose L sufficiently large so that each subinterval contains at most one spike. We index the subintervals $\ell = 1, \ldots, L$ and define $n_{k,\ell}$ as the indicator, which is 1 if there is a spike in subinterval $((\ell - 1)\Delta, \ell\Delta]$ on trial k and is 0 otherwise. We take $n_k = \{n_{k,1}, \ldots, n_{k,L}\}$ as the spikes recorded on trial k and $N_{1:k} = \{n_1, \ldots, n_k\}$ as the spikes recorded from trials 1 to k. We let $H_{k,\ell}$ denote the spike history in trial k, up to time $\ell\Delta$.

We assume that the conditional intensity function for the spike trains in our experiment at time $\ell\Delta$ of trial k may be written as

$$\lambda_k(\ell\Delta \mid \theta_k, \gamma, H_{k,\ell}) = \lambda^S(\ell\Delta \mid \theta_k)\lambda^H(\ell\Delta \mid \gamma, H_{k,\ell}), \quad \text{(B)}$$

where the component $\lambda^S(\ell\Delta|\theta_k)$ describes the task effect on the neural spiking and the component defines the spike history effect on neural spiking. We model the task $\lambda^H(\ell\Delta|\gamma, H_{k,\ell})$ effect on the neural spiking activity by assuming that the first component of the conditional intensity function (B) has the form

$$\log \lambda^{S}(\ell\Delta \mid \theta_k) = \sum_{r=1}^{R} \theta_{k,r} g_r(\ell\Delta), \quad \text{(C)}$$

where the $g_r(\ell\Delta)$ are a set of R functions which model the within-trial task specific-effect on spiking activity parameterized by $\theta_k = \{\theta_{k,1},\ldots, \theta_{k,r},\ldots, \theta_{k,R}\}$ for each trial $k = 1,\ldots, K$. We use unit pulse functions to represent the $g_r(\ell\Delta)$'s (Czanner et al., 2008). That is, Eq. C defines for each trial how the task modulates the spiking activity. We assume a different θ_k for each trial. We use a state-space model to define the relation between the θ_k's on different trials. This dependence defines the between-trial dynamics.

To model the effect of spike history on current neural spiking activity, we assume that the second component on the right of Eq. B can be defined as

$$\log \lambda^{H}(\ell\Delta \mid \gamma, H_{k,\ell}) = \sum_{j=1}^{J} \gamma_j n_{k,\ell-j} \quad \text{(D)}$$

for $k = 1,\ldots, K$, where $\gamma = (\gamma_1,\ldots, \gamma_j,\ldots, \gamma_J)$ is the vector of parameters which define the dependence of the current spiking activity on recent spike history. In this model, we take the history to be $H_{k,\ell} = \{n_{k,\ell-J},\ldots, n_{k,\ell-1}\}$, which is the spiking activity during the preceding time intervals J prior to time $\ell\Delta$. Together Eqs. B–D define the conditional intensity function, which defines the observation equation of our state space model of neural activity. It may be written as

$$\lambda_k(\ell\Delta \mid \theta_k, \gamma, H_{k,\ell}) = \exp\left\{\sum_{r=1}^{R} \theta_{k,r} g_r(\ell\Delta)\right\} \exp\left\{\sum_{j=1}^{J} \gamma_j n_{k,\ell-j}\right\}, \quad \text{(E)}$$

for $k = 1,\ldots, K$ and $\ell = 1,\ldots, L$. Equation E defines a generalized linear model formulation of the conditional intensity function (Truccolo et al., 2005). The probability of a spike event (0 or 1) on trial k at time $\ell\Delta$ is

$$\Pr(n_{k,\ell} \mid H_{k,\ell}) = [\lambda_k(\ell\Delta \mid \theta_k, \gamma, H_{k,\ell})\Delta]^{n_{k,\ell}} \exp\{-\lambda_k(\ell\Delta \mid \theta_k, \gamma, H_{k,\ell})\Delta\}. \quad \text{(F)}$$

We complete the definition of our state-space model of neural activity by defining the state equation to specify the relation of spiking activity between trials. We assume stochastic dependence between the $\theta_k = \{\theta_{k,1},\ldots,\theta_{k,r},\ldots,\theta_{k,R}\}$ defined by the random walk model

$$\theta_k = \theta_{k-1} + \varepsilon_k, \quad \text{(G)}$$

for $k = 1,\ldots, K$, where ε_k is an R – dimensional Gaussian random vector with mean 0 and unknown covariance matrix Σ. We also assume that the initial vector θ_0 is unknown. The random walk model provides a straight forward way to constrain the coefficients of the task-effect functions to be different on different trials yet, keep the model estimation problem tractable.

Equations F and G define our state-space generalized linear model. We denote its unknown parameter as $\psi = (\gamma, \theta_0, \Sigma)$. Because the θ_k's are unobservable and ψ is an unknown parameter, we use an EM algorithm to compute their estimates by maximum likelihood (Czanner et al., 2008; Smith and Brown, 2003). We denote the maximum likelihood estimate of the model parameters ψ as $\hat{\psi} = (\hat{\gamma}, \hat{\theta}_0, \hat{\Sigma})$.

The log of the complete data log likelihood which is the basis of the EM algorithm is

$$\log p(N, \theta \mid \psi) = \sum_{k=1}^{K} \sum_{\ell=1}^{L} n_{k,\ell} \log [\lambda_k(\ell\Delta \mid \theta_k, \gamma, H_{k,\ell})\Delta] - \lambda_k(\ell\Delta \mid \theta_k, \gamma, H_{k,\ell})\Delta + K \log ((2\pi)^{-R/2} \mid \sum \mid^{-1/2}) - 2^{-1} \sum_{k=1}^{K} (\theta_k - \theta_{k-1})' \sum^{-1} (\theta_k - \theta_{k-1}).$$

(H)

Equation H is a penalized log likelihood function, where the regularization constraint is imposed by the random walk model (Eq. G). That is, the task-effect coefficients impose a stochastic regularization constraint on the between-trial component of the conditional intensity function (Czanner et al. 2008). The maximum likelihood estimate of Σ governs the smoothness of the spike rate function by determining the degree of smoothing that is most consistent with the data.

Calculation of rate change trial—Because we were most interested in assessing how the neuron's firing during the cue changes over the course of learning, we selected a trial response window from the observation interval $(0, T]$, and analyzed how the firing rate during that window evolved over time using the state space approach outlined above. In order to find the trial during observational conditioning at which the firing rate during the cue was significantly different than the firing rate observed during habituation, we modified the trial-to-trial comparison algorithm in Smith et al 2010 (Appendix C) in the following way. In Smith et al 2010, samples from the rate at trial i were compared with samples at trial j and the probability that these distributions were different was computed. Here, we compare samples from the mean of the 15 habituation trials (making use of the first 15 trials' covariance matrix) with samples from trial j (Smith et al., 2010). When the samples from

trials *j* and *j*+1 are both greater/less than the combined samples from trials 1–15 with 95% probability, we designate trial *j* as the rate change trial.

We performed these analyses on all neurons that were classified as either showing conditioning selective responses or conditioned responses after application of the statistical tests described above in the section “Electrophysiology Analysis of learned cue responses”. We chose response windows of 1s and 9s as these were close to the windows chosen for our Wilcoxon analyses. We chose a 1s response window instead of 500ms as it provided the model with more data and allowed estimations with greater confidence. We performed the analysis with 1 second and 9 second response windows and used the earliest trial estimate. The algorithm successfully provided estimates for 87% (n = 13/15) of BLA and 92% (n = 34/37) of ACC neurons. These estimates were used to generate the histogram shown in Fig 3G–3H.

The matlab code for state-space analysis of neural firing is downloadable from http://annecsmith.net/firingrates.html

Neural trajectory analysis—We computed peri-stimulus time histograms (PSTHs) for each unit to estimate its trial-averaged activity over time for each condition. We aligned trials to CS onset, and each trial was sampled from 2 seconds before the CS onset to 5 seconds after the CS onset. We binned spike trains with a 50-msec window, and convolved these with a 150-msec Gaussian kernel to smooth the PSTHs. For each condition, a high-dimensional neural activity space will be formed, where each axis will be the firing rate for each unit. PSTHs from all the units will form a trajectory evolving in an n-dimensional space (n = the number of units). Prior to dimensionality reduction, we removed neurons that had extremely low firing rates over the 7-second window, and neurons that had either near-zero variability in the PSTH or large variability caused by a sparse spiking with a low number of trials. We then normalized the PSTH for each unit by its maximum variance across conditions for that unit, so we could avoid being biased by neurons with higher firing-rates and ensure that each unit has similar overall variability across conditions (Ames et al., 2014). We then performed principal component analysis (PCA) on these PSTHs for each condition to find the representative features in the face of the heterogeneity in neural dynamics. We preserved the dimensions that represented more than 70% of the total variance of the original data. As the first two principal components (PCs) are the most salient features of the population, we projected the trajectory on a 2D plane formed by PC1 and PC2.

Distance calculation—We calculated Euclidian distance between the neural trajectories for the habituation and observational conditioning conditions at each time point, and plotted this distance across time (from 2 s before CS onset to 5 s after CS onset).

We then calculated the mean distance during baseline period before CS onset (2 s, 40 data points where each data point is separated by 50 ms), and we chose 2 seconds right after CS onset and calculated the mean distance after CS onset (2 s, 40 data points). We calculated their difference to quantify how far the neural state at the habituation and observational conditioning training phases were during the baseline and CS presentation periods.

Identification of ACC network neurons—In mice expressing ChR2 in ACC→BLA projector neurons, an optrode was implanted into the ACC as described above. EO animals underwent observational conditioning while we recorded neural activity using the spike acquisition system. Immediately after conditioning the demonstrator mouse was removed and blue light (473 nm, 15–20 mW) from a laser was delivered through the optrode (phototagging). Electrophysiological recording was not stopped between observational conditioning and phototagging. During phototagging different stimulation parameters were used: 1 s pulse, 5 ms 1 Hz pulses, 5 ms 10 Hz pulses, and 5 ms 20 Hz pulses. We then analyzed neural responses to photostimulation. Due to the wide range of latencies in response to light stimulation in recorded units, we first used custom-written MATLAB scripts to calculate the latency from photostimulation onset to the first 10 ms bin (within 500 ms) with a 4 standard deviation (SD) increase over the baseline firing rate (−0.5 to 0 s) using the data from the 1Hz stimulation. We then used the Wilcoxon signed-rank test to determine if the firing rate within an experimental window was significantly different than the baseline firing rate. For neurons with latencies lower than 10 ms, a baseline window of 50 ms and response window of 20 ms was used. For neurons with latencies above 10 ms, a baseline of 500 ms was used and a response window of 50 ms that started 20 ms before the calculated latency. Lastly neurons that showed inhibitory responses were analyzed using the data from the 1 second constant pulse of light with a baseline window of 1 s and a response window of 500 ms. Neurons were categorized as ACC network neurons if they showed a significant excitatory or inhibitory response to photostimulation. Neurons were classified as ACC→BLA projectors if they had latencies under 8 ms and a significant difference in the rank-sum analysis ($P < 0.01$). Excitatory network neurons were defined as those that were excited to light with a latency of 20–100 ms and significant difference in the rank-sum analysis ($P < 0.01$). Inhibitory network neurons were defined as those that were inhibited to light within a 500 ms window using the rank-sum analysis ($P < 0.01$).

Analysis of BLA neural responses—All of the non-laser stimulation trials and the laser stimulation trials were grouped together for statistical analysis. Responses of single units to the cue onset were deemed statistically significant as described above. A neuron's response to the cue was categorized as being modulated by laser stimulation if the Wilcoxon signed-rank test for cue response was significant ($P < 0.01$) in either the laser stimulation or non-laser stimulation condition but not the other condition, or if the response windows were significantly different from each other. Multiple comparisons were corrected for by adjusting p values.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Yi-Ning Leow, Alexandriya Emonds, Amna Magzoub, Rain Thomas, Noa Golan, and Katie Goodwin for their contributions to this project. We would also like to thank Gwendolyn Calhoon, Li-Huei Tsai, Guoping Feng, and Rebecca Saxe for helpful discussion. K.M.T. is a New York Stem Cell Foundation - Robertson Investigator and McKnight Scholar. This work was supported by funding from the JPB Foundation, PIIF, PNDRF, JFDP, Whitehall Foundation, Klingenstein Foundation, NARSAD Young Investigator Award, Alfred P Sloan Foundation, New York Stem Cell Foundation, McKnight Foundation, Whitehead Career Development Chair, R01-MH102441-01 (NIMH),

RF1-AG047661-01 (NIA), the NIH Director's New Innovator Award DP2-DK-102256-01 (NIDDK), and the NIH Director's Pioneer Award (NCCIH DPI-AT009925). S.A.A. was supported by the Jeffrey and Nancy Halis Fellowship, the Henry E. Singleton Fund and a NLM training grant. R.W. was supported by the Simons Center for the Social Brain, the Netherlands Organization for Scientific Research (NWO) RUBICON fellowship program, and by a NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation. A.C.F.O. was supported by National Research Service Award Institutional Research Training Grant (5T32GM007484-38). F.M. was supported by a Canadian Institutes of Health Research Fellowship. A.B.-R. was supported by a NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation and the NIMH (Research Supplement to Promote Diversity in Health-Related Sciences). A.B. was supported by a fellowship from the Swiss National Science Foundation and a NARSAD Young Investigator Award. E.H.N. and C.A.L. were supported by NSF Graduate Research Fellowships (NSF GRFP), the Integrative Neuronal Systems Training Fellowship (T32 GM007484), and the Training Program in the Neurobiology of Learning and Memory. A.C.S. was supported by 1-R01-AG-050548-01. P.N. was supported by Singleton, Leventhal and Whitaker fellowships. E.N.B. was supported by: NIH Awards DP1-OD003646 and R01-GM104948.

References

Adolphs R, Tranel D, Damasio H, Damasio A. 1994; Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. Nature. 372:669–672. [PubMed: 7990957]

Allman JM, Hakeem A, Erwin JM, Nimchinsky E, Hof P. 2001; The Anterior Cingulate Cortex. Ann N Y Acad Sci. 935:107–117. [PubMed: 11411161]

Ames KC, Ryu SI, Shenoy KV. 2014; Neural Dynamics of Reaching Following Incorrect or Absent Motor Preparation. Neuron. 81:438–451. [PubMed: 24462104]

Apps MAJ, Rushworth MFS, Chang SWC. 2016; The Anterior Cingulate Gyrus and Social Cognition: Tracking the Motivation of Others. Neuron. 90:692–707. [PubMed: 27196973]

Atsak P, Orre M, Bakker P, Cerliani L, Roozendaal B, Gazzola V, Moita M, Keysers C. 2011; Experience Modulates Vicarious Freezing in Rats: A Model for Empathy. PLoS ONE. 6:e21855. [PubMed: 21765921]

Baeyens F, Kaes B, Eelen P, Silverans P. 1996; Observational evaluative conditioning of an embedded stimulus element. Eur J Soc Psychol. 26:15–28.

Barbieri R, Frank LM, Nguyen DP, Quirk MC, Solo V, Wilson MA, Brown EN. 2004; Dynamic analyses of information encoding in neural ensembles. Neural Comput. 16:277–307. [PubMed: 15006097]

Bastiaansen, JaCJ; Thioux, M; Keysers, C. 2009; Evidence for mirror systems in emotions. Philos Trans R Soc B Biol Sci. 364:2391–2404.

Beery AK, Kaufer D. 2015; Stress, social behavior, and resilience: Insights from rodents. Neurobiol Stress. 1:116–127. [PubMed: 25562050]

Beyeler A, Namburi P, Glober GF, Simonnet C, Calhoon GG, Conyers GF, Luck R, Wildes CP, Tye KM. 2016; Divergent Routing of Positive and Negative Information from the Amygdala during Memory Retrieval. Neuron. 90:348–361. [PubMed: 27041499]

Bissiere S, Plachta N, Hoyer D, Olpe HR, Grace AA, Cryan JF. 2008; The Rostral Anterior Cingulate Cortex Modulates the Efficiency of Amygdala-Dependent Fear Learning. Biol Psychiatry. 63:821–831. [PubMed: 18155183]

Bruchey AK, Jones CE, Monfils MH. 2010; Fear conditioning by-proxy: Social transmission of fear during memory retrieval. Behav Brain Res. 214:80–84. [PubMed: 20441779]

Buchanan SL, Powell DA. 1982; Cingulate cortex: Its role in Pavlovian conditioning. J Comp Physiol Psychol. 96:755–774. [PubMed: 7142487]

Burgos-Robles A, Kimchi EY, Izadmehr EM, Porzenheim MJ, Ramos-Guasp WA, Nieh EH, Felix-Ortiz AC, Namburi P, Leppla CA, Presbrey KN, et al. 2017; Amygdala inputs to prefrontal cortex guide behavior amid conflicting cues of reward and punishment. Nat Neurosci. 20:824–835. [PubMed: 28436980]

Burkett JP, Andari E, Johnson ZV, Curry DC, de Waal FBM, Young LJ. 2016; Oxytocin-dependent consolation behavior in rodents. Science. 351:375–378. [PubMed: 26798013]

Bussey TJ, Muir JL, Everitt BJ, Robbins TW. 1996; Dissociable effects of anterior and posterior cingulate cortex lesions on the acquisition of a conditional visual discrimination: Facilitation of early learning vs. impairment of late learning. Behav Brain Res. 82:45–56. [PubMed: 9021069]

Cassell MD, Wright DJ. 1986; Topography of projections from the medial prefrontal cortex to the amygdala in the rat. Brain Res Bull. 17:321–333. [PubMed: 2429740]

Chang SWC, Gariépy JF, Platt ML. 2013; Neuronal reference frames for social decisions in primate frontal cortex. Nat Neurosci. 16:243–250. [PubMed: 23263442]

Chang SWC, Fagan NA, Toda K, Utevsky AV, Pearson JM, Platt ML. 2015; Neural mechanisms of social decision-making in the primate amygdala. Proc Natl Acad Sci. 112:16012–16017. [PubMed: 26668400]

Chen Q, Panksepp JB, Lahvis GP. 2009; Empathy Is Moderated by Genetic Background in Mice. PLoS ONE. 4:e4387. [PubMed: 19209221]

Church RM. 1959; Emotional reactions of rats to the pain of others. J Comp Physiol Psychol. 52:132–134. [PubMed: 13654562]

Churchland MM, Cunningham JP, Kaufman MT, Foster JD, Nuyujukian P, Ryu SI, Shenoy KV. 2012; Neural population dynamics during reaching. Nature. 487:51. [PubMed: 22722855]

Cunningham JP, Yu BM. 2014; Dimensionality reduction for large-scale neural recordings. Nat Neurosci. 17:1500–1509. [PubMed: 25151264]

Czanner G, Eden UT, Wirth S, Yanike M, Suzuki WA, Brown EN. 2008; Analysis of between-trial and within-trial neural spiking dynamics. J Neurophysiol. 99:2672–2693. [PubMed: 18216233]

Devinsky O, Morrell MJ, Vogt BA. 1995; Contributions of anterior cingulate cortex to behaviour. Brain J Neurol. 118(Pt 1):279–306.

Dulac C, Torello AT. 2003; Molecular detection of pheromone signals in mammals: from genes to behaviour. Nat Rev Neurosci. 4:551–562. [PubMed: 12838330]

Gabbott PLA, Warner TA, Jays PRL, Salway P, Busby SJ. 2005; Prefrontal cortex in the rat: Projections to subcortical autonomic, motor, and limbic centers. J Comp Neurol. 492:145–177. [PubMed: 16196030]

Greene JT. 1969; Altruistic behavior in the albino rat. Psychon Sci. 14:47–48.

Guzmán YF, Tronson NC, Guedea A, Huh KH, Gao C, Radulovic J. 2009; Social modeling of conditioned fear in mice by non-fearful conspecifics. Behav Brain Res. 201:173–178. [PubMed: 19428631]

Haroush K, Williams ZM. 2015; Neuronal Prediction of Opponent's Behavior during Cooperative Social Interchange in Primates. Cell. 160:1233–1245. [PubMed: 25728667]

Heyes CM, Dawson GR. 1990; A demonstration of observational learning in rats using a bidirectional control. Q J Exp Psychol B. 42:59–71. [PubMed: 2326494]

Hopper LM, Lambeth SP, Schapiro SJ, Whiten A. 2008; Observational learning in chimpanzees and children studied through "ghost" conditions. Proc Biol Sci. 275:835–840. [PubMed: 18182368]

Hultman R, Mague SD, Li Q, Katz BM, Michel N, Lin L, Wang J, David LK, Blount C, Chandy R, et al. 2016; Dysregulation of Prefrontal Cortex-Mediated Slow-Evolving Limbic Dynamics Drives Stress-Induced Emotional Pathology. Neuron. 91:439–452. [PubMed: 27346529]

Isogai Y, Si S, Pont-Lezica L, Tan T, Kapoor V, Murthy VN, Dulac C. 2011; Molecular Organization of Vomeronasal Chemoreception. Nature. 478:241–245. [PubMed: 21937988]

Jeon D, Kim S, Chetana M, Jo D, Ruley HE, Lin SY, Rabah D, Kinet JP, Shin HS. 2010; Observational fear learning involves affective pain system and Cav1.2 Ca2+ channels in ACC. Nat Neurosci. 13:482–488. [PubMed: 20190743]

Jones BF, Groenewegen HJ, Witter MP. 2005; Intrinsic connections of the cingulate cortex in the rat suggest the existence of multiple functionally segregated networks. Neuroscience. 133:193–207. [PubMed: 15893643]

Kim EJ, Kim ES, Covey E, Kim JJ. 2010; Social Transmission of Fear in Rats: The Role of 22-kHz Ultrasonic Distress Vocalization. PLoS ONE. 5:e15077. [PubMed: 21152023]

Kim S, Mátyás F, Lee S, Acsády L, Shin HS. 2012; Lateralization of observational fear learning at the cortical but not thalamic level in mice. Proc Natl Acad Sci U S A. 109:15497–15501. [PubMed: 22949656]

Lammel S, Lim BK, Ran C, Huang KW, Betley MJ, Tye KM, Deisseroth K, Malenka RC. 2012; Input-specific control of reward and aversion in the ventral tegmental area. Nature. 491:212–217. [PubMed: 23064228]

Mante V, Sussillo D, Shenoy KV, Newsome WT. 2013; Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature. 503:78. [PubMed: 24201281]

Maren S. 2000; Auditory fear conditioning increases CS-elicited spike firing in lateral amygdala neurons even after extensive overtraining. Eur J Neurosci. 12:4047–4054. [PubMed: 11069601]

Masuda A, Aou S. 2009; Social Transmission of Avoidance Behavior under Situational Change in Learned and Unlearned Rats. PLOS ONE. 4:e6794. [PubMed: 19710921]

Meltzoff AN, Moore MK. 1977; Imitation of Facial and Manual Gestures by Human Neonates. Science. 198:75–78. [PubMed: 17741897]

Mineka S, Cook M. 1993; Mechanisms involved in the observational conditioning of fear. J Exp Psychol Gen. 122:23–38. [PubMed: 8440976]

Nieh EH, Matthews GA, Allsop SA, Presbrey KN, Leppla CA, Wichmann R, Neve R, Wildes CP, Tye KM. 2015; Decoding Neural Circuits that Control Compulsive Sucrose Seeking. Cell. 160:528–541. [PubMed: 25635460]

Olsson A, Nearing KI, Phelps EA. 2007; Learning fears by observing others: the neural systems of social fear transmission. Soc Cogn Affect Neurosci. 2:3–11. [PubMed: 18985115]

Panksepp JB, Lahvis GP. 2011; Rodent empathy and affective neuroscience. Neurosci Biobehav Rev. 35:1864–1875. [PubMed: 21672550]

Pereira AG, Cruz A, Lima SQ, Moita MA. 2012; Silence resulting from the cessation of movement signals danger. Curr Biol. 22:R627–R628. [PubMed: 22917506]

Pisansky MT, Hanson LR, Gottesman II, Gewirtz JC. 2017; Oxytocin enhances observational fear in mice. Nat Commun. 8:2102. [PubMed: 29235461]

Poulos AM, Zhuravka I, Long V, Gannam C, Fanselow M. 2015; Sensitization of fear learning to mild unconditional stimuli in male and female rats. Behav Neurosci. 129:62–67. [PubMed: 25621793]

Preston SD, de Waal FBM. 2002; Empathy: Its ultimate and proximate bases. Behav Brain Sci. 25:1–20. [PubMed: 12625087]

Quirk GJ, Repa JC, LeDoux JE. 1995; Fear conditioning enhances short-latency auditory responses of lateral amygdala neurons: Parallel recordings in the freely behaving rat. Neuron. 15:1029–1039. [PubMed: 7576647]

Sanders J, Mayford M, Jeste D. 2013; Empathic Fear Responses in Mice Are Triggered by Recognition of a Shared Experience. PLOS ONE. 8:e74609. [PubMed: 24058601]

Senn V, Wolff SBE, Herry C, Grenier F, Ehrlich I, Gründemann J, Fadok JP, Müller C, Letzkus JJ, Lüthi A. 2014; Long-range connectivity defines behavioral specificity of amygdala neurons. Neuron. 81:428–437. [PubMed: 24462103]

Smith AC, Brown EN. 2003; Estimating a state-space model from point process observations. Neural Comput. 15:965–991. [PubMed: 12803953]

Smith AC, Scalon JD, Wirth S, Yanike M, Suzuki WA, Brown EN. 2010; State-Space Algorithms for Estimating Spike Rate Functions. Comput Intell Neurosci. 2010

Steenland HW, Li XY, Zhuo M. 2012; Predicting Aversive Events and Terminating Fear in the Mouse Anterior Cingulate Cortex during Trace Fear Conditioning. J Neurosci. 32:1082–1095. [PubMed: 22262906]

Suzuki WA, Brown EN. 2005; Behavioral and neurophysiological analyses of dynamic learning processes. Behav Cogn Neurosci Rev. 4:67–95. [PubMed: 16251726]

Taylor GT. 1981; Fear and affiliation in domesticated male rats. J Comp Physiol Psychol. 95:685–693.

Truccolo W, Eden UT, Fellows MR, Donoghue JP, Brown EN. 2005; A Point Process Framework for Relating Neural Spiking Activity to Spiking History, Neural Ensemble, and Extrinsic Covariate Effects. J Neurophysiol. 93:1074–1089. [PubMed: 15356183]

Twining RC, Vantrease JE, Love S, Padival M, Rosenkranz JA. 2017An intra-amygdala circuit specifically regulates social fear learning. Nat Neurosci.

Tye KM, Prakash R, Kim SY, Fenno LE, Grosenick L, Zarabi H, Thompson KR, Gradinaru V, Ramakrishnan C, Deisseroth K. 2011; Amygdala circuitry mediating reversible and bidirectional control of anxiety. Nature. 471:358–362. [PubMed: 21389985]

Warren BL, Vialou VF, Iñiguez SD, Alcantara LF, Wright KN, Feng J, Kennedy PJ, Laplant Q, Shen L, Nestler EJ, et al. 2013; Neurobiological sequelae of witnessing stressful events in adult mice. Biol Psychiatry. 73:7–14. [PubMed: 22795644]

Yusufishaq S, Rosenkranz JA. 2013; Post-weaning social isolation impairs observational fear conditioning. Behav Brain Res. 242:142–149. [PubMed: 23295398]

Author Manuscript

Author Manuscript

Author Manuscript

- Neurons in cortex and amygdala respond to cues that predict shock to another mouse
- Cortex → amygdala neurons preferentially represent socially-derived information
- Cortical input to amygdala instructs encoding of observationally-learned cues
- Corticoamygdala inhibition impairs observational learning and social interaction

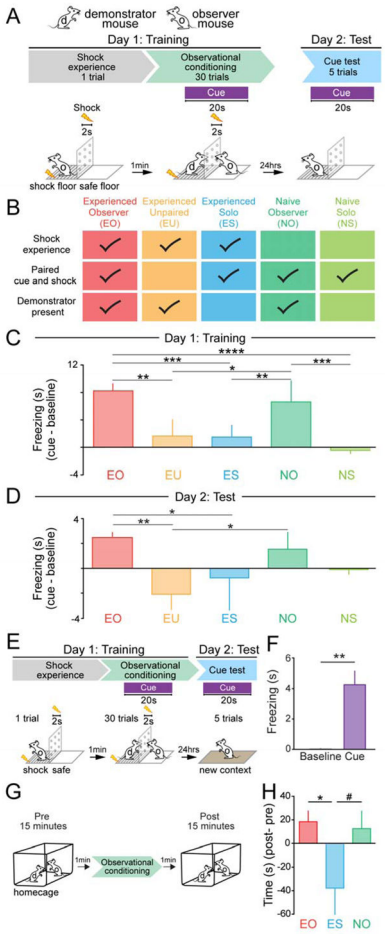


Figure 1. Parameters for mice to learn about a predictive cue via observational conditioning

A. Observational fear conditioning paradigm.

B. Conditioning paradigms for all behavioral groups (EO, N=7; EU, N=5; ES, N=6; NO, N=7; NS, N=4 mice).

C. EO and NO mice showed significantly higher freezing (cue-baseline) than EU, ES, and NS mice (one-way ANOVA, $F_{(4,24)}$=13.66, P< 0.0001, Bonferroni post-hoc analysis, *P<0.05, **P<0.01,***P<0.001, ****<0.0001).

D. On Day 2: Test, EO mice showed significantly higher freezing (cue-baseline) than EU and ES mice (one-way ANOVA, $F_{(4,24)}$=5.687, P=0.0023, Bonferroni post-hoc analysis, **P<0.01, *P<0.05).

E. Modified observational conditioning paradigm to test for context independent cue learning in EO mice.

F. EO mice showed significantly higher freezing during cue presentation in a novel context (two-tailed, unpaired student's t-test, t=4.535, df=10, **P=0.0011).

G. Testing for affiliative interactions between demonstrator and observer mice before and after observational conditioning.

H. Time interacting with the demonstrator after observational conditioning was statistically higher for EO mice (one-way ANOVA, $F_{(2,16)}$=3.779, P=0.0453, Bonferroni post-hoc

analysis, *P=0.0427), and showed a trend for NO mice (P=0.0859) when compared to ES mice. All error bars indicate ± SEM.

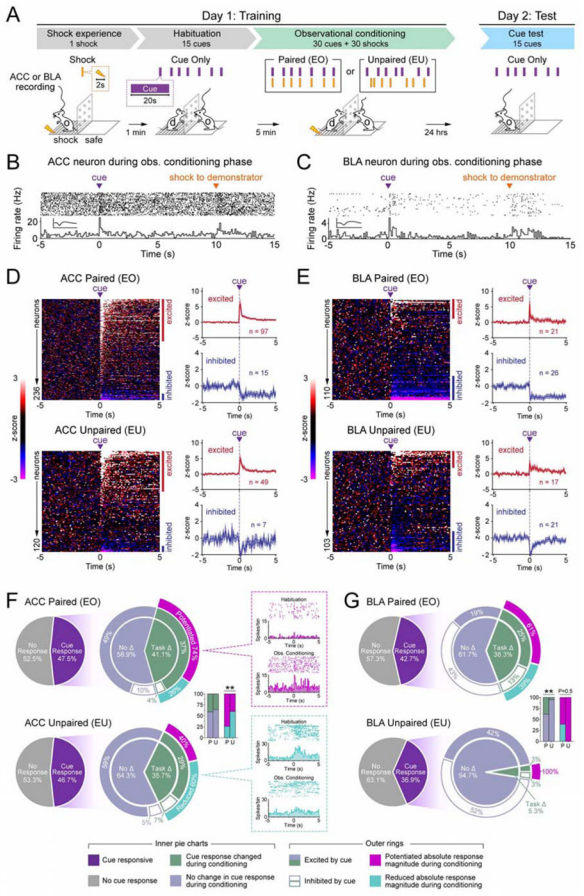


Figure 2. Encoding of observational conditioning in the ACC and BLA

A. Observational fear conditioning paradigm used for *in vivo* single-unit recordings in the ACC (paired group, N=16; unpaired group, N=7 mice) or BLA (paired group, N=6; unpaired group, N=6 mice).

B–C. Representative ACC and BLA neuron responses to cue and shock delivery during paired observational conditioning. Raster plots depict neural spikes (1 trial per row) and each peri-stimulus time histogram (PSTH) depicts the average firing frequency across all trials, relative to cue onset (100 ms bins). Insets show the average waveform recorded for each neuron (y-axis: 200 uV, x-axis: 1 ms).

D–E. Cue responses for paired and unpaired groups. Heatmap rows represent the z-score transformed average PSTH for individual neurons, columns represent time bins relative to cue onset (100 ms width). Blue and red bars indicate statistically significant cue-responsive cells. Plots to the right show average z-score responses for cue-excited and cue-inhibited cells.

F. Cue-responsive ACC subpopulations. Task-modulated neurons in the paired group showed a greater proportion of potentiated responses during conditioning than the unpaired group (bar graph inset; chi-square test: χ^2=6.93, ***P*=0.008). On the right, each PSTH shows example ACC neurons with training-induced potentiated or reduced cue responses.

G. Cue-responsive BLA subpopulations. A significantly greater proportion of cue-responsive neurons showed task-modulated responses in the paired group than the unpaired group (bar graph inset; chi-square test: χ^2=8.27, ** *P*=0.004).

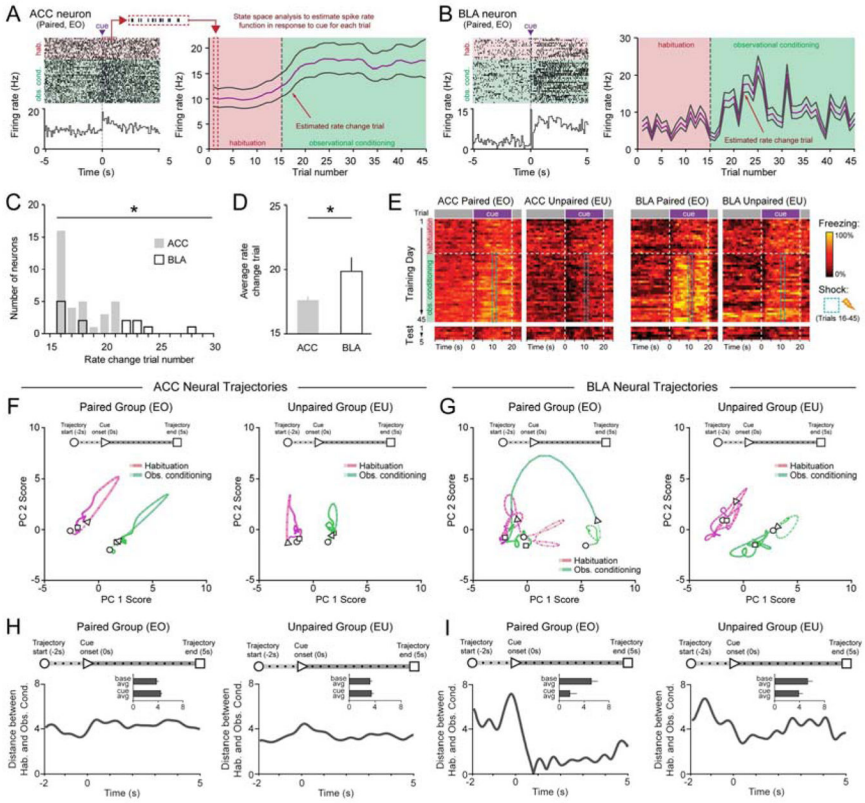


Figure 3. ACC and BLA contain neural correlates of observational learning

A–B. Raster and PSTH (100ms bins) of an ACC (**A**) and BLA (**B**) neuron identified as having a significant change in cue response during conditioning. State-space analysis provides a probabilistic estimate of the trial at which the neuron undergoes a rate change.

C. The distribution of rate change trials calculated by state-space analysis of conditioning-dependent neurons was significantly earlier in ACC than BLA neurons (Kolmogorov-Smirnov test, $*P<0.05$).

D. The average rate change trial of neurons in the ACC was earlier than those in the BLA (two-tailed, unpaired student's t-test, t=2.622, df=45, $*P=0.0119$).

E. Behavioral rasters (1s bins) of average freezing for all paired (EO) and unpaired (EU) mice across Day 1:Training and Day 2:Test in both ACC and BLA groups.

F–G. Neural ensemble dynamics in ACC and BLA across habituation and conditioning trials. Trial-averaged neural trajectories projected on a 2D space formed by first (PC1) and second (PC2) principle components for ACC neurons (**F**) (paired, n=201 neurons, N=12 mice; unpaired n=93 neurons, N=7 mice) and BLA neurons (paired, n=106, N=6 mice; unpaired n=97 neurons, N=6 mice). Dots on the trajectories represent timestamps (50 ms).

H–I. Calculated Euclidean distance between trajectory for habituation and trajectory for observational conditioning in paired and unpaired mice in the ACC (**H**) and BLA (**I**) plotted as distance across time (−2s to +5s from CS onset). Insets show averaged values for baseline and cue period. The distance between baseline and cue epochs in the BLA paired group was significantly different from other groups (Pearson's chi-square test: $\chi^2=4.953$, $*P=0.026$).

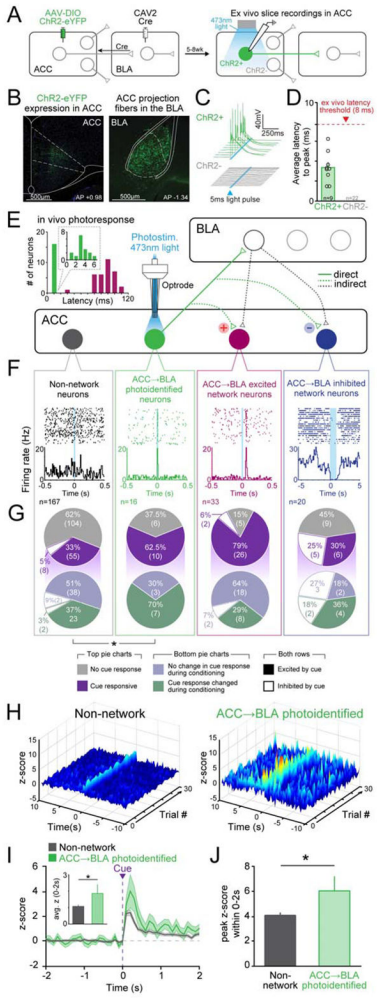


Figure 4. Photoidentified ACC→BLA projectors have an enhanced cue representation when compared to non-network ACC neurons

A. Schematic of intersectional viral approach. Retrograde virus CAV2-Cre was stereotaxically injected into the BLA and AAV-DIO-ChR2-eYFP into the ACC, resulting in ChR2 expression only in ACC neurons that monosynaptically project to the BLA.

B. Representative confocal images of ChR2 expression in the ACC and projection fibers in the BLA (blue=DAPI, green=eYFP).

C–D. *Ex vivo* electrophysiological recordings of ACC neurons in *ex vivo* in slices. (**C**) Voltage traces in response to light stimulation from ChR2+ (green) and ChR2- (grey) cells. (**D**) Average latency responses for all cells.

E. An optrode was placed into the ACC of mice (N=16) expressing ChR2 in ACC neurons projecting to the BLA. *In vivo* recordings during observational conditioning and subsequent phototagging were performed. Circuit model shows proposed ACC→BLA network connectivity of ACC neurons based on *in vivo* phototagging. Inset shows the range of photoresponse latencies seen during *in vivo* recordings (green bar = projectors (<8 ms), magenta bars = excited network neurons (20–120 ms)).

F. Example rasters and PSTH of non-network (10ms bins), ACC→BLA photoidentified (10ms bins), and ACC→BLA excited (10ms bins) or inhibited network neurons (100ms bins).

G. Proportions of non-network, ACC→BLA photoidentified, and ACC→BLA excited or inhibited network neurons that showed responses to the cue during observational conditioning. The ACC→BLA projector population exhibited a significantly greater proportion of cue-responsive neurons with 62.5% (n=10/16 neurons; N=16 mice) being excited and 0% (n=0/16) inhibited in response to the cue (Chi-square, χ^2=4.85, *df*=1 **P*=0.0276). In the ACC→BLA excited network, 79% (n=26/33) were excited and 6% (n=2/33) were inhibited to the cue.

H. Three-dimensional heat map (100ms bins) displaying the trial by trial z-score response of the ACC→BLA photo-identified and non-network neurons to the cue during observational conditioning.

I. Average z-score trace of cue responses in non-network (grey) or ACC→BLA phototagged (green) neurons. Inset: ACC→BLA projectors show a significantly greater average peak z-score response to the cue during conditioning compared (two-tailed, unpaired student's t-test, t=2.122, df=181, **P*=0.0352).

J. ACC→BLA projectors had a higher peak z-score response to the cue during observational conditioning when compared to non-network neurons (two-tailed, unpaired student's t-test, t=2.413, df=181, **P*=0.0168). All error bars indicate ± SEM.

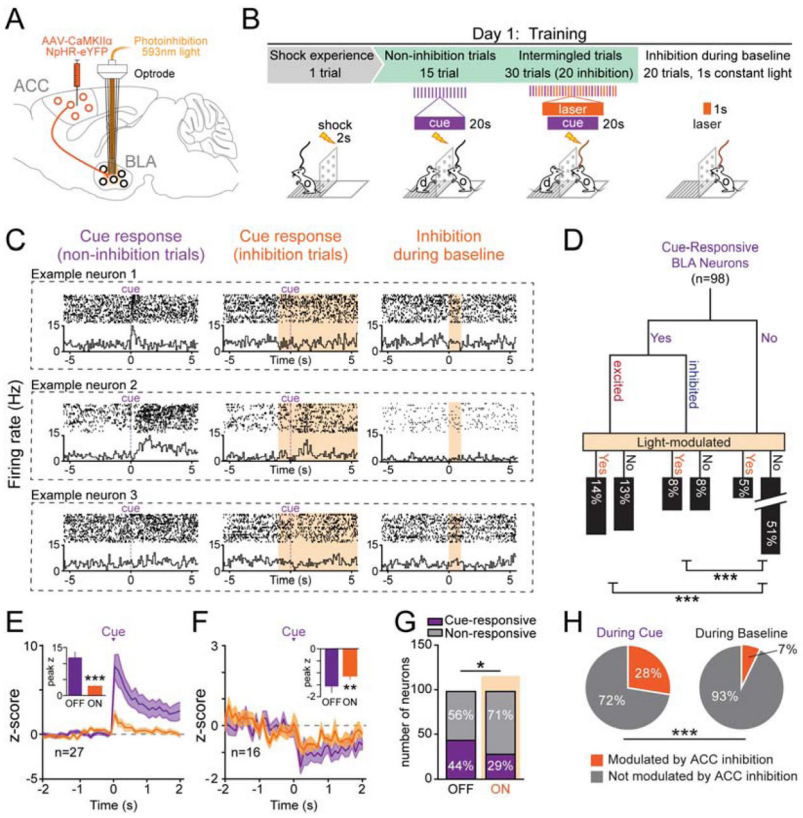


Figure 5. ACC input to the BLA governs cue-encoding during observational conditioning

A. Viral injection and optrode placement for selective inhibition of ACC→BLA input during individual trials of observational conditioning.

B. Behavioral paradigm during *in vivo* optrode recordings.

C. Rasters and PSTH (100ms bins) of example BLA neuron responses to the cue with and without optogenetic inhibition of ACC input to the BLA as well as inhibition during baseline.

D. Response tree of all BLA neurons (n=98; N=5 mice) with % of BLA neurons that were cue-responsive and whether ACC inhibition altered cue response. Cue responsive neurons showed greater modulation by ACC input inhibition than the non-responsive population (Chi-square test, excited: χ^2=18.60, df=1, ***P<0.0001; inhibited: χ^2=13.87, df=1, ***P=0.0002).

E. Average z-score trace of BLA neurons (n=27) that were excited in response to the cue with (orange) and without (purple) laser stimulation. Neurons excited by the cue showed a significantly reduced z-score response during laser stimulation. Inset shows average peak z-scores for the first 2s after the cue (N=5 mice, paired, two-tailed t-test, t=4.586, df=26, ***P=0.0001).

F. Average z-score trace of BLA neurons (n=16) that were inhibited in response to the cue was plotted with (orange) and without (purple) laser stimulation. Inset shows average peak z-scores for the first 2s after the cue (N=5 mice, paired, two-tailed t-test, t=3.01, df=15, ***P=0.0088).

G. A significantly smaller proportion of cells were cue responsive on trials where ACC input to the BLA was inhibited (Chi-square test, χ^2=4.969, df=1, **P*=0.0258).

H. Significantly more cells had firing rates that were modulated by light stimulation during cue presentation compared to baseline (Chi-square test, χ^2=12.10, df=1, ****P*=0.0005). All error bars indicate ± SEM.

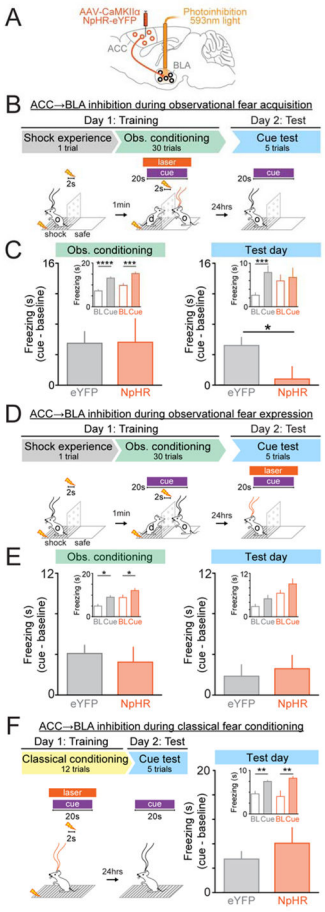


Figure 6. Photoinhibition of ACC→BLA impairs observational fear conditioning, but not classical fear conditioning

A. Viral injection and optic fiber placement for selective inhibition of ACC→BLA circuit.

B. Behavioral and light delivery paradigm for inhibition of ACC→BLA circuit during cue presentations during acquisition (Day 1) of observational conditioning.

C. During observational conditioning, there were no significant differences in freezing between NpHR (N=7) mice and eYFP (N=12) mice (unpaired, two-tailed t-test, t=0.0785, df=17, P =0.9383). However, on Test day, cue driven freezing was impaired in NpHR compared to eYFP mice (unpaired, two-tailed, t-test, t=2.378, df=17, *P =0.0294). Insets show cue and baseline (20s prior to cue onset) freezing values for observational conditioning and test day (BL=baseline; Observational conditioning: two-way ANOVA, group effect, $F_{(1,17)}$=8.286, P=0.0104, epoch effect, $F_{(1,17)}$=66.26, P<0.0001, group X epoch interaction, $F_{(1,17)}$= 0.0829, P0.7769; Bonferroni post-hoc analysis, ****P< 0.0001, ***P=0.0002; Test day: two-way ANOVA, group effect, $F_{(1,17)}$= 0.3596, P=0.5566, epoch effect, $F_{(1,17)}$=10.64, P=0.0046, group X epoch interaction, $F_{(1,17)}$=5.657, P=0.0294; Bonferroni post-hoc analysis, ***P=0.0005).

D. Behavioral and light delivery paradigm for inhibition of ACC→BLA circuit during cue presentations during expression (Day 2) of observational conditioning.

E. There were no significant differences in cue driven freezing between NpHR (N=9) and eYFP (N=8) mice during observational conditioning (unpaired, two-tailed, t-test, t=0.4916,

df=15, *P=0.6301) or Day2: Test (unpaired, two-tailed, t-test, t=0.5137, df=15, *P=0.6149). Insets show cue and baseline (20s prior to cue) freezing values during conditioning and test day (Observational conditioning: two-way ANOVA, group effect, $F_{(1,15)}$=10.46, P=0.0056, epoch effect, $F_{(1,15)}$=18.17, P=0.0007, group X epoch interaction, $F_{(1,15)}$=0.2416, P=0.6301; Bonferroni post-hoc analysis, *P<0.05; Test day: two-way ANOVA, group effect, $F_{(1,15)}$= 12.30, P=0.0032, epoch effect, $F_{(1,15)}$=6.778, P=0.02, group X epoch interaction, $F_{(1,15)}$=0.06837, P=0.7973; no significant Bonferroni post-hoc analysis).

F. Inhibition of ACC→BLA circuit during classical fear conditioning. No significant differences were detected between NpHR (N=7 mice) and eYFP (N=10) mice in cue driven freezing on test day (unpaired, two-tailed, t-test, t=1.02, df=15, *P=0.3237). Inset shows cue and baseline (20s prior to cue) freezing values (two-way ANOVA, group effect, $F_{(1,15)}$=0.0061, P=0.9389, epoch effect, $F_{(1,15)}$= 28.48, P<0.0001, group X epoch interaction, $F_{(1,15)}$=1.041, P=0.3237; Bonferroni post-hoc analysis, **P<0.01). All error bars indicate ± SEM.

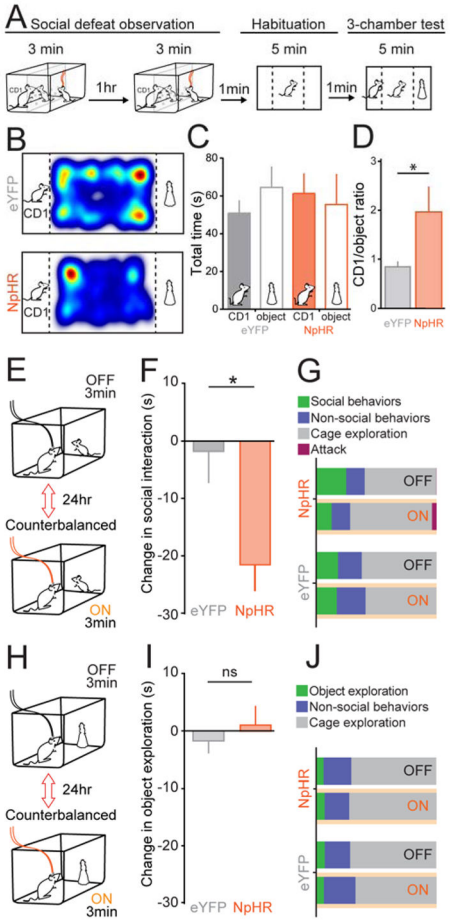


Figure 7. Photoinhibition of ACC→BLA also impairs other ethologically-relevant social behaviors

A. Social defeat observation paradigm. NpHR (N=9) and eYFP (N=9) mice naïve to a CD-1 mouse received 593 nm to inhibit the ACC→BLA circuit during two social defeat observation sessions (3 min each). Mice were then placed in a 3-chamber arena for habituation followed by conditioned avoidance.

B. Representative heat maps of time spent by NpHR and eYFP mice in the arena during the 3-chamber test.

C. Average of total time spent within the CD-1 or the object zone for NpHR and eYFP mice during the 3-chamber test.

D. NpHR mice had a higher ratio of time spent with the CD-1 instead of the object than eYFP mice (unpaired, Two-tailed t-test, t=2.147, df=16, **P*=0.0475).

E. Resident-intruder paradigm. A juvenile intruder was introduced to the homecage of a resident NpHR (N=7) or eYFP (N=11) mouse during light on and light off conditions separated by 24 hours and counterbalanced between mice.

F. Inhibition of ACC input to the BLA in the resident-intruder paradigm decreased social interaction time in NpHR compared to eYFP mice (unpaired, two-tailed t-test, t=2.609, df=16 **P*=0.019).

G. Summary of light-evoked changes in behavior during resident-intruder paradigm (Non-social behaviors = grooming and rearing). NpHR mice showed a decrease in social behaviors during ACC→BLA inhibition that was not evident in the eYFP group.

H. A novel object is introduced into the homecage of a resident NpHR or eYFP mouse during light on and light off conditions separated by 24 hours and counterbalanced between mice.

I. Novel object exploration was not altered by light-evoked inhibition of ACC input to the BLA (unpaired, two-tailed t-test, t=0.6952, df=15, P=0.4975).

J. Summary of light-evoked changes in behavior during novel object paradigm. NpHR and eYFP mice showed no change in object exploration during ACC-BLA inhibition. All error bars indicate ± SEM.