HHS Public Access

Author manuscript

Nat Neurosci. Author manuscript; available in PMC 2023 October 30.

Published in final edited form as:

Nat Neurosci. 2022 September ; 25(9): 1213–1224. doi:10.1038/s41593-022-01148-9.

A molecularly integrated amygdalo-fronto-striatal network coordinates flexible learning and memory

Dan C. Li[1,2,3], **Niharika M. Dighe**[2,3], **Britton R. Barbee**[2,3], **Elizabeth G. Pitts**[2,3], **Brik Kochoian**[3], **Sarah A. Blumenthal**[3], **Janet Figueroa**[2], **Traci Leong**[2], **Shannon L. Gourley**[2,3]

[1]Medical Scientist Training Program, Emory University School of Medicine, Atlanta, GA, USA.

[2]Department of Pediatrics, Children's Healthcare of Atlanta, Emory University School of Medicine, Atlanta, GA, USA.

[3]Emory National Primate Research Center, Emory University, Atlanta, GA, USA.

Abstract

Behavioral flexibility—that is, the ability to deviate from established behavioral sequences—is critical for navigating dynamic environments and requires the durable encoding and retrieval of new memories to guide future choice. The orbitofrontal cortex (OFC) supports outcome-guided behaviors. However, the coordinated neural circuitry and cellular mechanisms by which OFC connections sustain flexible learning and memory remain elusive. Here we demonstrate in mice that basolateral amygdala (BLA)→OFC projections bidirectionally control memory formation when familiar behaviors are unexpectedly not rewarded, whereas OFC→dorsomedial striatum (DMS) projections facilitate memory retrieval. OFC neuronal ensembles store a memory trace for newly learned information, which appears to be facilitated by circuit-specific dendritic spine plasticity and neurotrophin signaling within defined BLA–OFC–DMS connections and obstructed by cocaine. Thus, we describe the directional transmission of information within an integrated amygdalo-fronto-striatal circuit across time, whereby novel memories are encoded by BLA→OFC inputs, represented within OFC ensembles and retrieved via OFC→DMS outputs during future choice.

Correspondence and requests for materials should be addressed to Shannon L. Gourley. shannon.l.gourley@emory.edu.

Author contributions

Conceptualization: D.C.L. and S.L.G. Methodology: D.C.L. and S.L.G. Investigation (surgical preparation, behavioral testing and microscopy experiments): D.C.L., N.M.D., B.R.B., E.G.P., B.K. and S.A.B. Formal analysis (including statistical analyses): D.C.L., J.F. and T.L. Writing: D.C.L. and S.L.G. Supervision: S.L.G.

Competing interests

The authors declare no competing interests.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41593-022-01148-9.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41593-022-01148-9.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Behavioral flexibility—that is, deviating from established behavioral sequences—requires the encoding and retrieval of new reward-related memories to shape future choice. For instance, if one encounters road construction while driving a familiar route to work, then one would ideally learn that new information and later apply it to select a different path on subsequent days. The capacity to encode, stably represent, update and retrieve reward-related expectations likely requires the concerted integration of numerous cognitive processes across multiple time scales and brain regions—processes that are not yet fully understood. Flexible learning and memory are critical adaptations for navigating dynamic environments, and failures of these processes are thought to represent a core feature of many neuropsychiatric disorders, including addiction[1].

The orbitofrontal cortex (OFC) supports outcome-guided decision-making across a diverse array of behavioral tasks[2,3] and appears to be especially important for integrating novel information into established behavioral strategies[4-6]. A unified model posits that the OFC supports this varied set of computations by encoding a cognitive map of task space[3,7,8]—an abstract representation of the associative and causal structure formed by task-relevant variables within an environment[9]. This schematic organization allows for efficient integration of new learning into existing knowledge, thus enabling adaptive modification of action strategies when familiar expectancies are violated or change[10]. However, understanding the coordinated neural circuitry and cellular mechanisms by which specific decision variables are conveyed to, converge within and are ultimately accessed by OFC networks are still being defined. Studying these processes is particularly challenging given the functional heterogeneity of the OFC both across topographical subregions[2] and between intermingled neuronal populations[11,12]. Thus, lesions or bulk inactivation may fail to capture more intricate neural computations within the OFC, which appear to be highly dependent upon circuit connectivity[13-15].

Results

In this study, we sought to identify OFC circuits that mediate the learning and memory processes by which flexible behaviors are established and sustained in mice. Accordingly, we measured the ability of mice to adapt their behaviors when familiar actions were unexpectedly not rewarded (Fig. 1a). Mice were first trained across several sessions to respond in two distinct nose-poke apertures, each reinforced with delivery of a food pellet into a separate magazine, such that mice responded equivalently on both apertures after training (Extended Data Fig. 1). Next, mice were given isolated access to one aperture, wherein responding remained reinforced, followed by access to the other aperture, wherein responding was no longer reinforced, and, instead, pellets were delivered non-contingently (that is, for free) at a rate equal to that of the prior, reinforced session. Thus, for one nose-poking behavior, established reinforcement conditions were maintained, whereas, for the other, those expectations were violated. Notably, the number of food pellets associated with each response and aperture was equivalent. After training with ratio schedules of reinforcement, which generate strong associations between actions and outcomes, nose-poking is instrumental in nature (Extended Data Fig. 2), as opposed to a Pavlovian response based on the stimulus properties of the nose-poke aperture. As such, during a brief choice test, conducted in extinction, typical mice preferentially engage the aperture where nose-

poking had remained reinforced, thus demonstrating sustained, flexible deviation from the equivalent responding at both apertures, as during training.

BLA→OFC projections are necessary for memory encoding

We first investigated the potential role of inputs to the OFC in sustained response flexibility. The OFC is a topographically expansive and functionally heterogeneous structure[2]. Accordingly, we focused on the anterior aspect of the ventral OFC, which appears to be sensitive to changes in reward availability and updating response strategies according to action–outcome associations compared to more lateral or posterior subregions[5] (Extended Data Fig. 3). A prominent source of afferent OFC projections, particularly those terminating in the anterior OFC, originates in the basolateral amygdala (BLA)[16], lesions of which alter reward-related firing of OFC neurons[17,18]. Behavioral flexibility in our task requires both (1) encoding new memories about changing reward expectancies and (2) later retrieving that information during subsequent choice epochs to guide response strategies (Fig. 1a). Accordingly, we first examined whether BLA→OFC projections contribute to memory encoding by selectively expressing inhibitory (Gi-coupled) chemogenetic receptors on BLA→OFC projecting neurons (Fig. 1b-d and Supplementary Fig. 1). We administered the chemogenetic ligand clozapine-*N*-oxide (CNO) immediately after a session when a familiar expectation was violated (Fig. 1e), thus inactivating BLA→OFC projections during the putative memory-encoding period but without altering responding during that non-reinforced session itself (Extended Data Fig. 4). Choice behavior was first assessed after fixed ratio training (Fig. 1f). Silencing BLA→OFC projections during memory encoding with a 1.0 mg kg^{-1} CNO dose disrupted subsequent choice behavior, such that mice did not favor the reinforced behavior (Fig. 1g). Meanwhile, control mice and mice administered a lower, behaviorally sub-threshold CNO dose (0.1 mg kg^{-1}) preferred the action that had remained reinforced, thus displaying flexible choice behavior. Throughout, response patterns were stable during choice tests (Supplementary Fig. 2).

We then further trained mice using random interval schedules of reinforcement whereby some interval of time is inserted between reinforcer deliveries, during which reinforcement is not available. Interval schedules, with time, weaken the associative link between actions and outcomes[19,20], causing animals to persist in a given behavioral sequence, even if it is not explicitly reinforced (Extended Data Fig. 5). Neither moderate interval training nor sub-threshold inactivation of BLA→OFC projections alone were sufficient to disrupt memory encoding, but these manipulations exhibited an additive effect, such that mice experiencing both were then unable to flexibly deviate from established action strategies (Fig. 1h-j). This pattern suggests that the memories being encoded by BLA→OFC projections pertain to action–outcome associations—that is, that an action will (or will not) produce an outcome. Of course, another possibility is that mice assign the negative association from the unexpected absence of a food pellet reward directly onto the nose-poke action itself, thus circumventing the need to update the specific predictive relationship between action and outcome per se. However, the frequency of these unexpectedly non-reinforced actions did not predict subsequent choice behavior (Supplementary Fig. 3 and Extended Data Fig. 6), so we think that this possibility is unlikely.

As a control, we confirmed that extended interval training occluded response flexibility in all mice, as would be expected with prolonged interval training (Fig. 1i-j). Notably, we also confirmed that neither CNO nor its metabolite clozapine affected choice behavior in the absence of chemogenetic receptor expression (Fig. 1k and Supplementary Fig. 4). Additionally, flexible choice behavior remained intact when BLA→OFC inactivation was delayed by 6 hours (Fig. 1l-o), indicating that new memory encoding occurs within a narrow time window after exposure to new task parameters. Altogether, these findings demonstrate that BLA→OFC projections are necessary for encoding new memories when familiar expectancies are violated.

We next examined whether BLA→OFC projections are necessary during choice epochs when mice must retrieve previously encoded memories to guide flexible action selection. Here, inactivating these projections did not alter choice behavior in any training condition (Fig. 1p-s). So, whereas the encoding of new reward-related memories requires BLA→OFC projections, the subsequent retrieval does not. Finally, we also confirmed that inactivation of BLA neuron terminals in the OFC via intracranial CNO infusion (Fig. 2a-d) disrupted flexible memory encoding (Fig. 2e,f), as with systemic CNO administration.

Bidirectional control of new action memory encoding

We next asked whether BLA→OFC projection activation can augment memory encoding. To do so, we interfered with flexible decision-making behavior using cocaine, exposure to which disrupts goal-seeking action strategies in mice and humans alike[21-23]. We expressed excitatory (Gq-coupled) chemogenetic receptors on BLA→OFC-projecting neurons in cocaine-exposed mice and then stimulated these projections with CNO administration during the memory-encoding period (Fig. 3a-d). Notably, prior cocaine exposure did not disrupt nose-poke responding during training (Fig. 3e) or pre-synaptic marker content in the OFC (Supplementary Fig. 5), suggesting that any disruptions in later response flexibility are unlikely attributable to gross deficits in the acquisition of instrumental learning or pre-synaptic inputs. As anticipated, prior cocaine disrupted the capacity of mice to modify response strategies when established expectancies change (Fig. 3f-h). BLA→OFC projection stimulation during memory encoding reinstated flexible expectancy updating in these mice (Fig. 3f,g). Thus, BLA→OFC projections appear to exert bidirectional control over memory encoding, such that silencing and stimulating these neurons respectively disrupts and enhances the formation of new memory necessary for later flexible action.

Surprisingly, BLA→OFC projection stimulation disrupted flexible choice in cocaine-naive mice (Fig. 3f-h). Specifically, response preference ratios revealed an interaction among training schedule, cocaine and BLA→OFC projection stimulation (Fig. 3i). Thus, the relationship between BLA→OFC neuron activity and expectancy updating appears to follow a non-linear, inverted-U-shaped relationship, with ‘too little’ or ‘too much’ activity hindering memory encoding, a shared motif across many prefrontal cortical functions[24]. Finally, the inability of BLA→OFC projection stimulation to generate flexible responding after extended interval training (Fig. 3h) suggests that the mechanisms underlying cocaine-induced deficits in response flexibility are separable from those induced by extended experience with a behavioral routine—specifically, that cocaine might weaken networks

controlling new learning, which depend on BLA→OFC connections, whereas extended training might strengthen habit-promoting networks, which do not.

OFC→DMS projections facilitate memory encoding and retrieval

Given that BLA inputs to the OFC mediate memory encoding but are no longer required during retrieval, we next hypothesized that outputs from the OFC may access previously learned information to guide behavior during choice epochs. A major target of efferent OFC projections is the dorsomedial striatum (DMS)[25], whose interactions with the OFC support the online arbitration between goals and habits[26], thus raising the possibility that these projections may also access learned expectancies during choice epochs. We again used projection-selective chemogenetic inhibition strategies, but of OFC→DMS projections (Fig. 4a-d), and tested whether these projections were required for retrieving or encoding new memories. Here, silencing OFC→DMS projections during either encoding or retrieval periods (Fig. 4e) impaired flexible choice (Fig. 4f-h); these deficits were not attributable to repeated chemogenetic manipulation of the same mice (Extended Data Fig. 7). Thus, OFC→DMS projections are necessary for multiple learning and memory phases necessary for sustained response flexibility.

Another prominent target of OFC projections is the BLA[27]. For choice behaviors guided by Pavlovian associations, cue-related and reward-related activity within the BLA is dysregulated by OFC lesions[28,29]. However, we observed no effects of inhibiting OFC→BLA projections during either the encoding or retrieval of outcome expectancy memories here (Fig. 4i-k).

Neural ensembles in the OFC store a memory trace for new action learning

We found that new expectancy information appears to be first conveyed to, and later retrieved from, the OFC to sustain response flexibility. Thus, we further hypothesized that a stable representation of new information (that is, a memory trace) is stored within OFC neuronal ensembles, which is later accessed to guide flexible choice. To test this possibility, we used *Fos*$^{2A\text{-}iCreER}$ mice to manipulate OFC neurons that were activated during a precise temporal window via administration of 4-hydroxytamoxifen (4OHT; Fig. 5a,b), which we used to drive inhibitory chemogenetic receptor expression among active OFC neuronal populations, with minimal 4OHT-independent expression (Fig. 5c-f). We administered 4OHT after a session when familiar expectations were violated, thus allowing selective chemogenetic manipulation of OFC neurons that were activated during new memory encoding (Fig. 5g). When these encoding-activated OFC neurons were silenced during a later choice test, the ability of mice to execute flexible choice was abolished (Fig. 5h-j). This indicates that neuronal ensembles in the OFC indeed store a memory trace for new outcome expectancy information.

OFC neurons respond to a diverse array of potentially task-related variables[2]. We hypothesized that this memory trace is specifically tuned for novel task parameters—in this case, that a familiar behavior is no longer rewarded. Accordingly, we performed a parallel control experiment in which we administered 4OHT after a session when responding remained reinforced (Fig. 5k). Interestingly, we detected a significant amount of activity-

dependent labeling in response to these familiar reinforcement conditions (Fig. 5f), but silencing these OFC neurons during subsequent choice tests had no effect (Fig. 5l-n). This suggests that, although these neurons might process certain features of the task environment, they are not necessary for shaping future action selection itself. Notably, the number of labeled OFC neurons in neither novel nor familiar conditions predicted choice subsequent behavior (Extended Data Fig. 8), suggesting that our findings reflect the functions of specifically recruited neuronal populations rather than a non-specific effect of OFC inactivation more generally. Thus, the allocation of neurons into a persistent memory trace necessary for future response flexibility is gated by novel, and not familiar, information, and it is the reactivation of this specific neuronal ensemble that is ultimately required to guide future choice.

New learning triggers spine plasticity on defined neurons

Of course, long-term storage of action variables by neural circuits can occur only if new experiences trigger lasting cellular changes that allow prior learning to be later accessed. Thus, we sought to uncover neuronal plasticity-related mechanisms that may durably sustain cohesive network function across time in support of flexible decision-making. Accordingly, we examined the density and structure of dendritic spines, the principal sites of excitatory neurotransmission and plasticity in the brain[30] on deep-layer principal excitatory neurons in the OFC, which both send and receive the bulk of long-range cortical projections[31]. To visualize OFC neurons that specifically link BLA→OFC to OFC→DMS circuits, we developed a combinatorial viral-mediated method for trans-synaptic anterograde relay mapping (ARM) by infusing an anterograde trans-synaptically transduced Cre-recombinase in the BLA, alongside a retrogradely transported Cre-dependent fluorescent reporter in the DMS (Fig. 6a), thus labeling monosynaptic ‘relay’ neurons in the OFC that both receive input from the BLA and send direct outputs to the DMS (Fig. 6b). We performed ARM in *Thy1*-YFP mice, thus enabling high-fidelity reconstruction of three-dimensional dendritic spine morphologies specifically among this population of connectivity-defined OFC neurons (Fig. 6b-d and Extended Data Fig. 9).

We compared mice that underwent instrumental response training to mice that received food pellets at a rate yoked to that of a trained cage mate, independent of nose-poke actions (Fig. 6e). Thus, mice had identical experiences, but one group lacked the opportunity for instrumental reward-related learning and memory (Fig. 6f). We found that instrumental learning was associated with overall dendritic spine pruning on BLA→OFC→DMS relay neurons (Fig. 6g). Stratification of spines by morphological subtype revealed that this effect was driven by selective reduction of thin-shaped spines (Fig. 6h-j), resulting in an enhanced proportion of mushroom-shaped spines (Fig. 6k), which contain more mature, stable synaptic connections, over thin-shaped spines, which are more labile and functionally variable[30]. Among mushroom-shaped spines, instrumental learning also triggered spine head enlargement (Fig. 6l), an anatomical substrate for synaptic potentiation necessary for memory formation[32,33]. Learning-induced enhancements in both the size and relative proportion of mushroom-shaped spines on BLA→OFC→DMS relay neurons is consistent with augmented signal-to-noise processing, potentially priming OFC→DMS projections for reactivation during memory retrieval.

We hypothesized that mice that are not capable of flexible expectancy updating will not display the same pattern of dendritic spine plasticity. To that end, we next visualized BLA→OFC→DMS relay neurons in cocaine-exposed mice, which could not update their behavioral response strategies after novel reinforcement conditions (Fig. 6f). We found that cocaine exposure caused a general loss of dendritic spines, principally driven by decreased mushroom-shaped spine densities (Fig. 6g,h). Notably, cocaine attenuated the learning-induced patterns of dendritic spine plasticity observed in typical mice, by blunting thin spine elimination (Fig. 6i), preventing enhancement of mushroom spine proportions (Fig. 6k) and mushroom spine head enlargement (Fig. 6l). Thus, cocaine disrupted learning-induced dendritic spine plasticity, such that OFC neurons of cocaine-exposed mice grossly resembled those from mice that entirely lacked instrumental learning. Furthermore, dendritic spine densities and morphological parameters robustly predicted individual choice behavior across cocaine-exposed and naive animals (Fig. 6m,n and Extended Data Fig. 10). Thus, we posit that new learning triggers OFC neurons to prioritize mature, mushroom-shaped spines, increasing their size and relative proportions, to provide a cellular substrate for flexible expectancy updating in support of sustained flexible action, and that the disruption of typical learning-related dendritic spine plasticity may contribute to the emergence of cocaine-induced behavioral sequelae.

Neurotrophin tone in defined circuits supports new learning

Our findings suggest that processes canonically associated with synaptic and dendritic spine plasticity should be necessary within the OFC for mice to update expectancies for sustained response flexibility. Thus, we tested whether brain-derived neurotrophic factor (BDNF), a potent and ubiquitously expressed facilitator of synaptic and dendritic spine plasticity[34], may be essential for BLA→OFC→DMS circuit function. Specifically, we paired unilateral *Bdnf* knockdown in the OFC with unilateral chemogenetic inhibition of the ipsilateral or contralateral BLA or DMS during memory encoding (Fig. 7a). Notably, BLA–OFC and OFC–DMS projections are overwhelmingly ipsilateral[16,27]. Thus, mice with contralateral infusions will have one hemisphere in which a functionally intact BLA or DMS interacts with a *Bdnf*-deficient OFC and another hemisphere in which a typical OFC interacts with a chemogenetically silenced BLA or DMS. If the mechanism by which BLA–OFC or OFC–DMS connections support the memory-encoding process depends on BDNF tone in the OFC, then (1) mice with contralateral infusions will display impaired expectancy updating, whereas (2) mice with ipsilateral infusions will retain the capacity for flexible choice, as these mice still possess one hemisphere in which a functionally intact BLA or DMS interacts with a typical *Bdnf*-expressing OFC.

We first tested whether BLA–OFC interactions during memory encoding are BDNF dependent (Fig. 7b). Here, mice with contralateral, but not ipsilateral or control, infusions displayed inflexible choice behavior (Fig. 7c,d), indicating that the mechanism by which BLA–OFC interactions support memory encoding is dependent on BDNF in the OFC. We next tested whether OFC–DMS interactions are BDNF dependent (Fig. 7e). Again, only mice with contralateral infusions were inflexible (Fig. 7f,g), demonstrating that the function of OFC–DMS connections in supporting memory encoding is also BDNF dependent. Given that (1) BLA→OFC, but not OFC→BLA, projections are involved in memory encoding

(Fig. 1), and (2) OFC–DMS connections are unidirectional (that is, OFC→DMS)[25], our findings suggest that circuit-specific neurotrophin signaling in the OFC is required for adaptive expectancy updating, likely by sustaining plasticity within a BLA→OFC→DMS circuit.

Discussion

In this study, we identified an extended amygdalo-fronto-striatal circuit that controls separable features of reward-related learning and memory necessary to sustain response flexibility, thus enabling adaptive decision-making. BLA→OFC projections mediate the encoding of new memories when familiar, previously rewarded behaviors are unexpectedly not reinforced. Then, during future choice epochs, OFC→DMS projections facilitate the retrieval of those memories, to prioritize actions that remained likely to be reinforced. Thus, learned expectancy information is transmitted in a directionally defined manner across a discrete BLA→OFC→DMS circuit, which functions across time to enable flexible choice. Expectancy updating triggers dendritic spine plasticity on excitatory OFC neurons that serves as monosynaptic relays within this circuit and requires circuit-specific neurotrophin signaling. These neural circuit and cellular plasticity mechanisms likely coordinate the formation and persistence of a memory trace for learned expectancy information, which we demonstrate is stably represented by novelty-gated, encoding-activated neuronal ensembles in the OFC, and whose reactivation is necessary for subsequent memory retrieval. The OFC, therefore, serves as the critical locus within this distributed amygdalo-fronto-striatal network, providing the temporal link between memory encoding and retrieval, thereby bridging the initial learning of new information with its future application.

The OFC performs multiple cognitive functions in support of reward-seeking behavior, from processing outcome-predictive cues[17,18], to tracking reward value[15,35,36], to facilitating causal attribution[37,38]. A common feature that emerges across various learning modalities points to the OFC as being especially important for flexibly integrating novel or unexpected information into established behavioral strategies[4-6]. For example, the OFC is necessary for accessing outcome value information to guide choice but only after shifts in familiar instrumental contingencies and not in situations when those contingencies remained stable[5]. Thus, OFC networks appear to be recruited in the service of action updating, as opposed to initial learning[39,40], which relies on the flexible modulation of existing behavioral strategies. In our task, mice were first trained to generate two actions for food reward, conditions that remained stable across many sessions and produced non-preferential performance of either action. Then, the likelihood that one action would be rewarded was unexpectedly changed, a situation that typically modulates future choice, such that mice later favor the more consistently rewarded behavior. We demonstrate that BLA→OFC and OFC→DMS projections are required for encoding new memories, which are later retrieved via OFC→DMS connections for response strategy updating.

The OFC's capacity to use new learning to modify established behavioral strategies may be due to its ability to construct cognitive maps of task space[3,7,8]—schematic representations of the associative and causal structures formed by specific task demands[10]. The OFC is capable of forming and maintaining long-term memories for many different types of

associations[14,15,41,42] but likely prioritizes those variables most relevant to obtaining a given desired outcome[9]. Consistent with this notion, we found that experiencing novel task conditions—that is, that a familiar behavior was unexpectedly not reinforced—triggered a greater degree of OFC neuronal activation than did familiar conditions. And notably, only suppression of novelty-activated, but not familiarity-activated, neurons obstructed later choice adaptation, indicating that only certain OFC neurons are allocated toward a memory trace for new learning, which are ultimately necessary for guiding future choices. Thus, this memory trace is likely both (1) sparsely encoded within the OFC, meaning reactivation of only a relatively small number of cells is necessary for memory retrieval, consistent with prior observations for other cortical memory systems[43,44], and (2) selectively recruited based, at least in part, on the novelty of the actions variables within the task environment.

We demonstrate that BLA→OFC projections are required for the formation of these OFC-dependent memories. One possibility is that these afferent projections transmit information regarding action–outcome links, consistent with the role of BLA→OFC projections in processing positive outcome-related information during reinforcement learning[13]. They may also conceivably converge with other inputs, such as those from the hippocampus, which are required for contingency memory updating and training state representations in the OFC[45,46]. In this case, BLA→OFC inputs might organize multiplexed input signals by biasing the allocation of certain OFC neurons into a memory trace. The BLA also supports memory consolidation by broadcasting salience signals that modulate attentional bias and facilitate information prioritization[47], which is also a prominent means by which decision variables are allocated into task space representations[48,49].

We also aimed to determine whether stimulation of BLA→OFC projections could enhance memory encoding in situations where response flexibility is weakened. Accordingly, we examined the effect of cocaine, not only because cocaine reliably disrupts OFC-dependent behaviors in mice but also because inflexible decision-making is a hallmark deficit among patients with cocaine use disorder[21-23]. Among the effects of cocaine on excitatory OFC neurons is the loss of dendritic spines, the primary sites of excitatory plasticity in the brain[30], distinguishing these neurons from those in other brain regions that, instead, undergo spinogensis[50,51]. Loss of dendritic spine densities would alter the ability of OFC neurons to receive and process synaptic inputs, but, of course, cocaine does not eliminate all dendritic spines, and we found that stimulating BLA→OFC connections restored response flexibility after cocaine, potentially reflecting a generalizable attention-biasing mechanism that augments the strength of new memories[52,53].

Given that afferent BLA→OFC projections bidirectionally control the encoding of new outcome-expectancy memories but become dispensable during future choice, and that OFC ensembles are a key locus for stable representation and later retrieval of newly learned reinforcement information, we next identified an efferent OFC→DMS pathway that facilitates memory retrieval during choice epochs. Our findings raise questions about how OFC→DMS outputs translate OFC-dependent task representations into adaptive action selection during memory retrieval. OFC-dependent cognitive maps are crucial for performing tasks in which animals must inferentially apply knowledge to guide behavior[54-56]—that is, when optimal action selection cannot be determined based solely

on observable features of the task environment. Although the striatum is capable of independently accessing task state information to guide choice, these representations are bound to observable stimuli, and, as such, the striatum's ability to infer non-observable or partially observable task states requires OFC input[57]. Notably, because choice tests in our task are performed in extinction (that is, without reinforcer delivery), action selection is effectively unbound from observable outcomes. Thus, our finding that OFC→DMS projections are indispensable for memory retrieval may reflect a role for these projections in linking previously encoded learning with current task demands to guide action selection.

Long-term storage of action variables by neural circuits likely requires lasting structural changes that allow prior learning to be later accessed. Accordingly, we found that new learning required neurotrophin tone and was associated with dendritic spine plasticity on excitatory OFC neurons, specifically within a BLA→OFC→DMS circuit. In particular, new learning reduced thin-shaped spine densities, enhancing the relative proportion of mushroom-shaped spines, which contain more mature, stable synaptic connections[30]. Among mushroom-shaped spines, new learning also triggered spine head enlargement, which serves as an anatomical substrate for synaptic potentiation[33] and appears necessary for the stability of long-term memories[32,58]. Learning-induced enhancements in both the size and relative proportion of mushroom-shaped spines on OFC neurons receiving input from the BLA and projecting to the DMS is consistent with augmented signal-to-noise processing, potentially resulting in the prioritization or priming of OFC→DMS projection reactivation during memory retrieval.

Altogether, we delineated specific OFC connections that control distinct phases of learning and memory that support expectancy updating—from the encoding of new information when expectations are violated to the retrieval of these newly formed memories—and we have shown that OFC neurons form stable memory ensembles. These processes are accompanied by dendritic spine plasticity on neurons defined by both afferent and efferent connections. Finally, we also discovered that cocaine causes dendritic spine loss on these OFC relay neurons and that it obstructs the typical pattern of learning-induced plasticity. Thus, a key question is whether neuron structural dynamics specifically among connectivity-defined populations of dendritic spines in the OFC are causally linked with reward-seeking behavior. This possibility seems likely given that pharmacological interventions that restore response flexibility after cocaine require dendritic spine plasticity in the OFC[23]. Currently, bioengineered actin cytoskeletal modulators can be leveraged to allow for optical control of spine dynamics in genetically defined neuronal populations[59]. As technologies such as these are further refined to enable experimental manipulation of neuronal structural dynamics at single-spine resolution in vivo—that is, to selectively target spines that receive pre-synaptic inputs from defined projections, such as from the BLA, while sparring those that do not—it may become feasible to causally link connectivity-defined dendritic spine plasticity with behavioral outcomes, thus closing the gap between the mesoscale identification of unitary circuits that underlie specific behaviors and the cellular and molecular mechanisms that govern the assembly, coordination and potential dysfunction among coherent brain networks over time.

Methods

Animals

Mice were bred from Jackson Laboratory stock. (1) Unless noted, experiments used C57BL/6J mice (no. 000664). (2) Dendritic spine imaging was performed using B6.Cg-Tg(*Thy1*-YFP)HJrs/J mice (no. 003782)[60], back-crossed onto a C57BL/6 background. (3) Memory trace experiments were performed using Fos$^{tm2.1(iCre/ERT2)Luo}$/J mice (Fos$^{2A\text{-}iCreER}$; no. 030323)[43]. (4) BDNF-dependent functional disconnection experiments used Bdnf^{tm3Jae}/J mice (*Bdnf*-flox; no. 004339)[61]. Male mice were used throughout this study, except in memory trace experiments in which male and female Fos$^{2A\text{-}iCreER}$ mice were evenly distributed for each group, with no effects of sex detected.

Mice were group-housed on a 12-hour light cycle and provided food and water ad libitum before testing, with temperature and humidity ranges of 17.8–26.1 °C and 30–70%. Throughout instrumental conditioning, mice were food-restricted until body weights were reduced to 90–93% of baseline to motivate food-reinforced responding. For all food-reinforced experiments, instrumental training began between postnatal day 60 (P60) and P70. Procedures were performed in accordance with National Institutes of Health Guidelines for the Care and Use of Laboratory Animals and approved by the Emory University Institutional Animal Care and Use Committee.

Instrumental response training

Behavioral testing was conducted using operant conditioning chambers equipped with two distinct nose-poke apertures (left versus right), with a separate magazine for food pellet delivery (Med Associates). For all experiments, each mouse performed all tasks in the same chamber. Mice were trained to nose-poke for delivery of a 20-mg grain pellet food reinforcer (Bio-Serv). Initial training proceeded under a fixed ratio 1 (FR1) schedule of reinforcement whereby each response resulted in a single reinforcer. Mice underwent daily training sessions, which lasted either 70 minutes or until mice earned 30 reinforcers from responding on each of the two apertures (60 total reinforced responses), whichever occurred first. Mice required between five and seven training sessions to acquire 60 pellets within one session, at which point response training was considered completed.

For experiments with multiple training conditions, we used 30-second and 60-second random interval (RI30 and RI60) schedules of reinforcement. During RI training sessions, each time a nose-poke response resulted in reinforcer delivery, that aperture would become inactive for a random interval of time (30 or 60 seconds on average throughout the session), during which time nose-poke responses on that aperture produced no effect. These sessions also ended at either 70 minutes or until 60 reinforcers were earned, whichever occurred first. Mice underwent a predetermined number of RI sessions based on experimental requirements.

Test for response flexibility

After response training, we assessed the capacity of mice to modify response strategies based on reinforcer likelihood. This procedure consisted of two 25-minute sessions

occurring on consecutive days. On the first day, one nose-poke aperture was occluded, while responding on the other, available aperture remained reinforced. On the second day, the opposite aperture was occluded, but responding on the available aperture was no longer reinforced. Instead, pellets were delivered into the magazine at a rate equal to what each mouse experienced during the previous day's reinforced session, independent of nose-poking responses. Thus, for one aperture, the association between nose-poking and pellet delivery was maintained, whereas, for the other aperture, that expectation was violated. Which aperture (left versus right) was designated to be reinforced versus non-reinforced was counterbalanced within and between groups.

Mice were returned to the same chambers on a third day for a 15-minute choice test when both apertures were once again available, and we recorded nose-poke responses on either aperture. For experiments when mice underwent multiple choice tests after different reinforcement schedules, the assignment of which aperture was to be reinforced versus non-reinforced was reversed for each round of testing so that mice could not use information from prior tests to inform their subsequent response strategies.

During non-reinforced sessions, some pellet deliveries occurred by chance simultaneously with, or shortly after, a nose-poke response. These pellet deliveries could, therefore, appear to be a result of the preceding nose-poke. We defined these 'apparently' reinforced nose-pokes as those followed within 5 seconds by pellet delivery. We then computed the standard contingency measure (ΔP) to index the net apparent reinforcement probability for each non-reinforced session:

$$\Delta P = P(R \mid \mathrm{NP}) - \ - P(R \mid \ \sim \mathrm{NP})$$

where ΔP is defined as the difference between the conditional probability of receiving a food pellet reward (R) given a nose-poke (NP) action and the conditional probability of receiving a food pellet in the absence of a nose-poke action[62,63].

Instrumental omission

Omission experiments were conducted using the same equipment and parameters as above, except that only one nose-poke aperture was available. After response training using an FR1 schedule of reinforcement, mice underwent a 30-minute omission procedure described previously[64]. Here, a pellet was delivered every 20 seconds, but each nose-poke response reset the counter and delayed reinforcer delivery. Thus, mice must inhibit a behavior to be reinforced. If nose-poking behavior is governed by action–outcome relationships, mice will refrain from nose-poking[64]. As a replication, mice were then re-trained to press a previously recessed lever, again using an FR1 schedule of reinforcement. These mice then underwent the same omission procedure as above.

Projection-selective chemogenetic manipulations

Intracranial surgeries were performed 2 weeks before training to allow for expression of viral vectors. Mice were anesthetized with ketamine (80 mg kg^{-1}) and dexmedetomidine (0.5 mg kg^{-1}) and placed in a digitized stereotaxic frame (Stoelting). Surgeries were

performed under aseptic conditions. A mid-sagittal incision exposed the skull, and a craniotomy was performed to allow for intracranial viral vector infusion. Infusion volumes and coordinates relative to bregma were: OFC (0.5 μ per hemisphere; ML: ±1.40 mm, AP: +2.50 mm, DV: −2.80 mm), BLA (0.25 μl per hemisphere; ML: ±3.00 mm, AP: −1.50 mm, DV: −4.90 mm) and DMS (0.5 μl per hemisphere; ML: ±1.50 mm, AP: +0.50 mm, DV: −3.00 mm). Viruses were infused over 5 minutes using a microliter syringe (Hamilton), left in place for an additional 10 minutes before retraction to restrict off-target viral vector spread. To achieve projection-selective chemogenetic receptor expression, a retrogradely transported Cre-recombinase construct (AAVrg-hSyn-HI-eGFP-Cre-WPRE-SV40; Addgene, 105540; deposited by James Wilson) was infused into the projection target region alongside an anterogradely transported Cre-dependent chemogenetic receptor construct in the projection source region (AAV5-hSyn-DIO-hM4D(Gi)-mCherry; Addgene, 44362, and AAV5-hSyn-DIO-hM3D(Gq)-mCherry; Addgene, 44361, both deposited by Bryan Roth)[65]. After testing, mice were euthanized, and brains were fixed and prepared as described below. Green fluorescent protein (GFP) or mCherry expression was examined to determine the extent of Cre and hM4D(Gi) or hM3D(Gq) expression, respectively.

The chemogenetic receptor ligand CNO (0.1 mg kg^{-1} or 1.0 mg kg^{-1} in 2% DMSO and saline; RTI International) was administered intraperitoneally (i.p.) in a volume of 1 ml per 100 g immediately after the non-reinforced session (encoding), 6 hours after the non-reinforced session (delayed) or 30 minutes before a choice test (retrieval). For experiments when mice underwent multiple choice tests, mice received the same CNO or vehicle dose throughout the entire experiment. We also examined potential off-target effects of both CNO (1.0 mg kg^{-1}) and its metabolite, clozapine (0.1 mg kg^{-1} in 2% DMSO; RTI International), on our behavioral measures in the absence of chemogenetic receptor expression due to concerns regarding CNO back-metabolism to clozapine[66].

Chemogenetic projection terminal inactivation

A constitutively expressed, inhibitory chemogenetic receptor (AAV5-CamKIIα-hM4Di-mCherry; Addgene, 50477, deposited by Bryan Roth) or control construct (AAV5-CamKIIα-mCherry; Addgene, 114469, deposited by Karl Deisseroth) was infused in the BLA bilaterally (0.25 μl per hemisphere; ML: ±3.00 mm, AP: −1.50 mm, DV: −4.90 mm) before behavioral testing. After instrumental response training, bilateral guide cannula (P1 Technologies) targeting the OFC were implanted (ML: ±1.35 mm, AP: +2.50 mm, DV: −2.25 mm). Mice were individually housed after surgery and underwent additional response training to habituate them to nose-poking with the cannula in place. Response flexibility was tested as described above except that, immediately after the non-reinforced session, mice were lightly anesthetized with isoflurane, and internal cannula were lowered (DV: −2.75 mm). CNO was infused bilaterally over a 5-minute period (300 nl; 300 μM in 2% DMSO and saline), with mice free-roaming in their home cage.

Histology

Mice were anesthetized with ketamine (100 mg kg^{-1}) and xylazine (10 mg kg^{-1}) and trans-cardially perfused with ice-cold PBS and 4% paraformaldehyde (PFA). Brains completed fixation in 4% PFA over 48 hours and were then transferred to 30% w/v sucrose. Fixed

brains were sectioned into 50-μm coronal sections using a freezing microtome. Brain sections were visualized using a Leica DM5500 microscope (Leica Microsystems), and images were collected using a spinning disk confocal microscope (VisiTech International) by a blinded investigator.

For experiments using projection-specific chemogenetic manipulations, we confirmed viral vector expression in the infused brain regions. Notably, BLA output neurons that project to various brain regions show sparse collateralization and largely represent non-overlapping cell populations[67]. However, OFC projection neurons do commonly exhibit axonal collateralization[68,69]. Thus, we also examined other brain regions involved in reward-related learning and memory and that might be expected to receive OFC axon collaterals, including the medial prefrontal cortex, nucleus accumbens, hippocampus, pallidum, ventral tegmental area and hypothalamus. We did not detect viral-mediated fluorescence in these regions, suggesting that axon collateral density is quite low relative to the robust expression observed in targeted, infused brain regions.

Cocaine administration

Cocaine (10 mg kg^{-1} in saline; Sigma-Aldrich) was administered i.p. daily from P31 to P35 in a volume of 1 ml per 100 g.

Protein quantification

Mice were anesthetized with isoflurane and rapidly decapitated. Brains were immediately extracted, flash-frozen and stored at −80 °C. Frozen brains were sectioned into 1-mm coronal slices and dissected using a 1-mm tissue core. Tissue was homogenized and protein concentrations were determined using a Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). Then, 15 μg of total protein for each sample was separated by SDS-PAGE on 4–20% gradient Tris-glycine gels (Bio-Rad). Protein was transferred to a PVDF membrane (Bio-Rad), blocked with 5% non-fat dry milk and incubated overnight at 4 °C in the following primary antibodies: rabbit anti-synaptophysin (1:10,000, Abcam) and rabbit anti-HSP-70 (1:1,000, Cell Signaling Technology). Membranes were incubated in horseradish peroxidase–conjugated goat anti-rabbit secondary antibody (1:10,000, Vector Laboratories) for 1 hour at room temperature. Immunoreactivity was assessed using Pierce ECL chemiluminescence substrate (Thermo Fisher Scientific) and measured using a ChemiDoc XRS+ Imaging System (Bio-Rad). Densitometry values were analyzed using Image Lab Software (Bio-Rad, version 5.0) and were normalized to each sample's respective loading control (HSP-70). All samples were processed and quantified in triplicate by experimenters blinded to group. Protein levels were expressed as a fold change compared to control samples from each membrane replicate.

Memory trace inactivation

$Fos^{2A\text{-}iCreER}$ mice underwent intracranial surgeries to bilaterally infuse a Cre-dependent inhibitory chemogenetic construct (AAV5-hSyn-DIO-hM4D(Gi)-mCherry) into the OFC, as described above. Mice were singly housed in a dark, quiet environment for 1 hour before each training and test session and placed back in the same cage for another hour once the session was completed, before eventually being returned to their respective home cages. For

the final 2–3 training sessions before the non-reinforced session, mice were given i.p. saline injections immediately after the conclusion of each training session to habituate mice to potential i.p. injection stress. Together, these measures were implemented to limit extraneous neuronal activation in the OFC unrelated to processing outcome expectancies and task variables.

4OHT (40 mg kg^{-1} in 2% Tween 80, 5% DMSO and saline; Sigma-Aldrich) was administered i.p. in a volume of 2 ml per 100 g immediately after either the reinforced or non-reinforced session, thus inducing permanent recombination of Cre-dependent constructs by triggering nuclear localization of iCre-ER^{T2}, whose expression is driven by activity-dependent *Fos* promoter and enhancer elements. Mice were then left undisturbed in their home cages for 1 week to allow for expression of the newly recombined construct. Finally, CNO (1.0 mg kg^{-1}, i.p.) was administered 30 minutes before a choice test.

After testing, mice were euthanized, and brains were fixed and prepared as described in the 'Dendritic spine imaging and reconstruction' section below. mCherry-expressing OFC neurons were imaged to quantify the number of activity-dependently labeled cells. A 600-μm^2 (676 μm × 887 μm) region of interest (ROI) was imaged with a ×10 objective. For each animal, three ROIs were analyzed from coronal sections corresponding to the AP coordinates of viral infusions (+2.50 mm from bregma), which were determined based on anatomical features, per reference mouse brain atlas[70]. For each coronal section, an ROI was imaged from the center of fluorescence expression within the ventrolateral OFC. Cell counting was performed using Fiji image analysis software version 2.1.0 (ref. [71]). Background subtraction and binary thresholding was applied to enhance contrast between fluorescence signal arising from neuronal cell bodies versus neurites or imaging artifacts. An average cell density was computed for each animal across three sampled ROIs, and statistical analysis was performed on a per-animal basis. Investigators were blinded for all cell-counting analyses.

Trans-synaptic ARM

To selectively label OFC neurons that both receive projections from the BLA and send projections to the DMS, we infused an anterograde trans-synaptically transduced[72] Cre-recombinase construct in the BLA (AAV1-hSyn-Cre-WPRE-hGH; Addgene, 105553; deposited by James Wilson) alongside a retrogradely transported Cre-dependent fluorophore in the DMS (AAVrg-hSyn-DIO-mCherry; Addgene; deposited by Bryan Roth). Thus, only neurons transduced with the Cre-recombinase from pre-synaptic terminals originating in the BLA and the Cre-dependent fluorophore from its own projections to the DMS will express the mCherry reporter. Notably, the AAV1 serotype exhibits retrograde transport and, thus, will also transduce top-down OFC→BLA projections[72]. However, this 'off-target' expression is eliminated here by limiting recombinase expression to OFC neurons projecting to a downstream target (that is, OFC→DMS).

We tested the effect of instrumental conditioning on dendritic spine density and morphology on connectivity-defined OFC neurons. Accordingly, after surgery, pairs of cage mates were randomly assigned to undergo either training as above or an identical series of sessions, except that pellets were delivered into the chambers at a rate yoked to each mouse's paired

cage mate, independent of nose-poking responses. Yoked mice were food-restricted to the same degree and placed in behavioral chambers daily for the same duration as their cage mate pairs. Thus, paired mice had identical experiences during all phases of testing, except that one mouse lacked the opportunity for new instrumental learning.

Dendritic spine imaging and reconstruction

Brains were extracted, sectioned and prepared for confocal microscopy as above. Basal dendritic segments located 50–150 μm from the soma on deep-layer, YFP-expressing OFC neurons were identified, and Z-stack series containing dendritic segments were collected using a ×100 1.4 NA oil-immersion objective and a 0.1-μm step size. Three-dimensional reconstructions of dendritic spine morphologies were performed using the FilamentTracer module in Imaris software (Oxford Instruments, version 8). Dendritic spines along an approximately 20-μm dendritic segment were identified using a semi-automated auto-depth function. Morphological classification of dendritic spines was performed using established parameters for pyramidal prefrontal cortex neurons[73]. For spines with a head:neck diameter ratio ≥1.1, those with head diameter ≥0.4 μm were classified as mushroom-type or otherwise considered thin-type. For spines with a head:neck diameter ratio <1.1, those with a spine length:neck diameter ratio ≥2.5 were also classified as thin-type or otherwise considered stubby-type. One dendrite was analyzed per cell, and six dendrites were analyzed per animal by a blinded rater.

Molecular–functional circuit disconnection

Bdnf-floxed mice underwent intracranial surgeries and received a unilateral infusion of a lentiviral construct expressing either Cre-recombinase or a GFP control (LV-CMV-Cre, LV-CMV-gFP; Emory University Viral Vector Core) in the OFC and a unilateral infusion of an inhibitory chemogenetic receptor construct (AAV5-CaMKIIα-hM4D(Gi)-mCherry; Addgene, 504777; deposited by Bryan Roth) in either the ipsilateral or contralateral BLA or DMS. The infusion sides (left versus right hemisphere) were randomly assigned. Mice underwent behavioral testing as above and administered CNO (1.0 mg kg^{-1}, i.p.) immediately after the non-reinforced session, thus assessing whether OFC BDNF is necessary for the function of BLA–OFC or OFC–DMS connections during memory encoding.

After behavioral testing, expression of the chemogenetic construct in the BLA or DMS, or of the control virus in the OFC, was confirmed by examining fluorescence from the mCherry or GFP reporters, respectively, by a blinded investigator. Expression of the Cre-recombinase constructs in the OFC were confirmed using immunohistochemistry. Free-floating brain sections, prepared as above, were washed three times for 5 minutes each in PBS and then blocked in a 0.01% Triton X-100 and 5% normal goat serum (NGS) solution for 1 hour at room temperature. Tissues were then incubated in primary antibody against Cre-recombinase (Cell Signaling Technology, 5036; 1:1,000 in 0.1% Triton X-100 and 1% NGS) overnight at 4 °C. Tissues were washed three times for 10 minutes each in PBS and then incubated in a fluorescence-conjugated (Alexa Fluor 488 goat anti-rabbit; Jackson ImmunoResearch, 111-545-144) secondary antibody (1:500 in 0.1% Triton X-100 and 1% NGS) for 2.5 hours at room temperature.

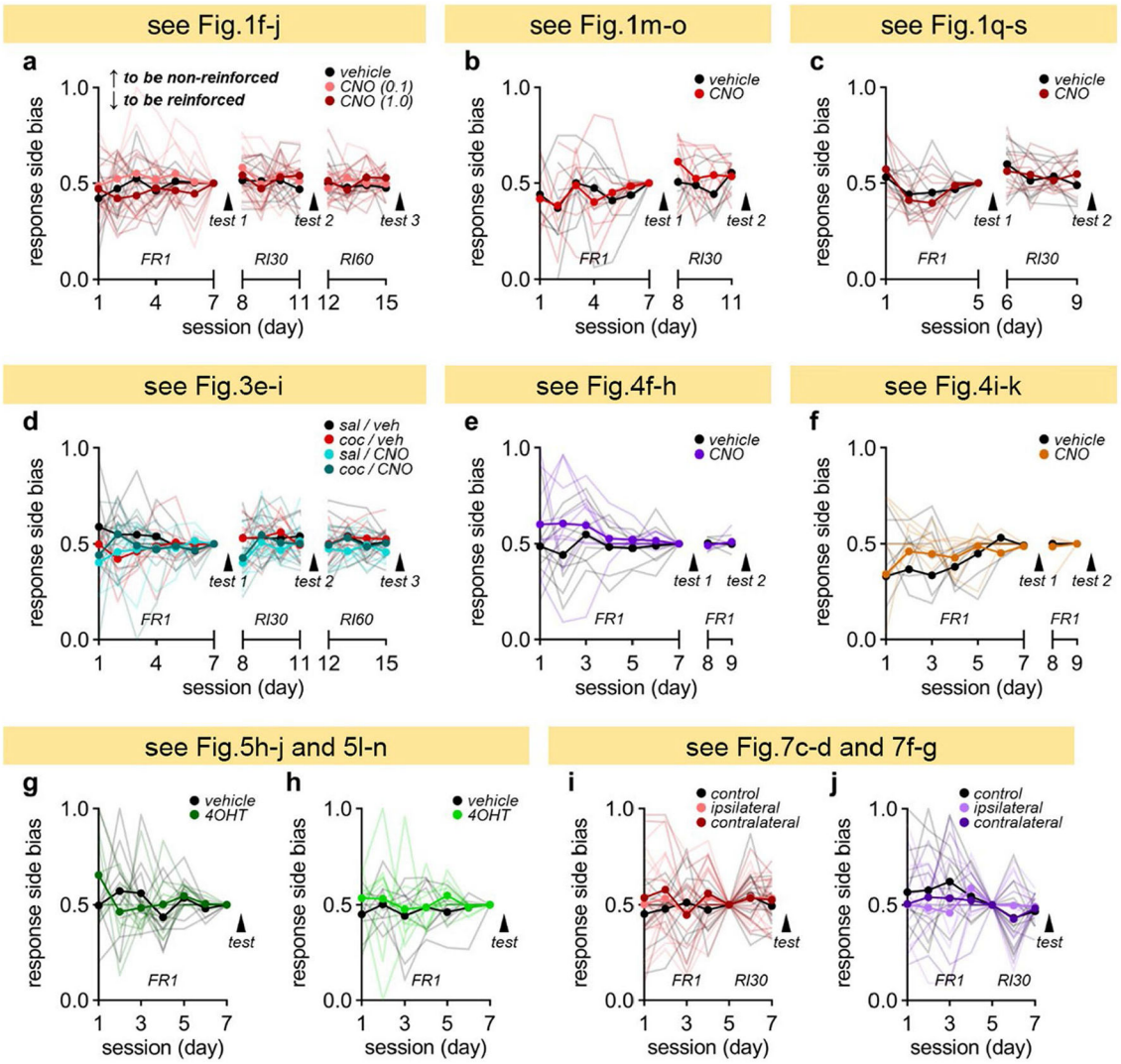
Statistical analysis

Statistical analysis was performed using SPSS software version 27 (IBM), with $\alpha < 0.05$. For behavioral experiments, response rates were compared using ANOVA. Several factors were analyzed as repeating measures, including session number during instrumental response training and reinforcement condition (that is, reinforced versus non-reinforced) during choice tests. Post hoc comparisons were performed after significant interaction effects or main effects distinguishing >2 conditions. For all image quantification data, each animal contributed one value for analysis.

For dendritic spine analysis, we accounted for hierarchical data structure (six animals per group, six dendrites per animal) using linear mixed model (LMM) analysis. To assess whether each individual dendrite measurement ($n = 144$) could be treated as independent measurements, or if adjustment for clustering within mice was needed, we computed intra-class correlation coefficients (ICCs) for each experimental group using random-intercept mixed models for each mouse. Across the board, ICC values indicated the need to adjust for subject-wise clustering. Accordingly, we then performed a two-factor (cocaine × training) LMM with an interaction term to test for differences in each dendritic spine variable across both factors. Thus, this approach enables the analysis of all individual dendrite measurements but avoids oversampling by accounting for correlated error due to subject-wise (that is, per-mouse) clustering. Post hoc analysis after interaction effects was performed in the same way as a one-factor LMM for pair-wise comparisons.

Sample sizes were calculated based on power analyses using data from prior, similar experiments. Data were not subject to exclusion except in cases of viral vector misplacement. Data distribution was assumed to be normal, but this was not uniformly tested. All experiments were replicated at least once. Mice were randomized to groups. Comparisons were two-tailed throughout, except for the use of one-tailed *t*-tests in Supplementary Fig. 5.

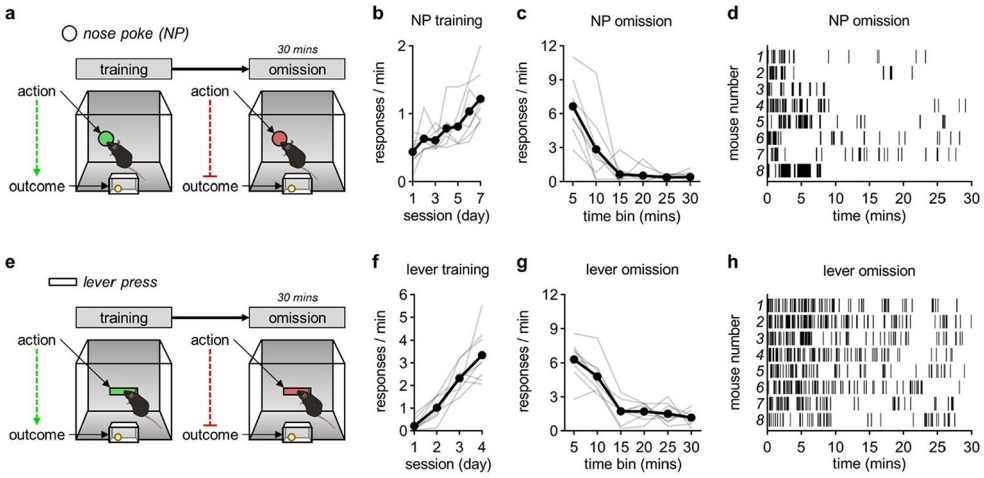
Extended Data



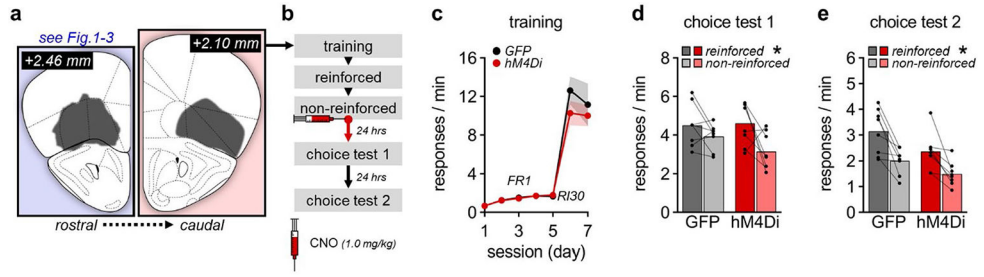
Extended Data Fig. 1 |. Mice do not display preference for one nose-poke aperture during training.

Response side bias (responses on aperture to be non-reinforced / total responses) during training sessions. **(a)** BLA→OFC inactivation: memory encoding (session: $F_{14,378} = 1.22$, 0.256; session × CNO: $F_{28,378} < 1$), **(b)** delayed memory encoding (session: $F_{10,80} = 3.18$, $p = 0.002$; session × CNO: $F_{10,80} < 1$), or **(c)** memory retrieval (session: $F_{8,128} = 4.86$, $p < 0.001$; session × CNO: $F_{8,128} < 1$). **(d)** BLA→OFC stimulation (session: $F_{14,392} < 1$; session × cocaine: $F_{14,392} < 1$; session × CNO: $F_{14,392} = 1.30$, $p = 0.205$; session × cocaine × CNO: $F_{14,392} < 1$). **(e)** OFC→DMS inactivation (session: $F_{8,144} = 1.05$, $p = 0.404$; session× CNO: $F_{8,144} = 1.21$, $p = 0.295$). **(f)** OFC→BLA inactivation (session: $F_{8,112} = 1.04$, $p = 0.205$; session × CNO: $F_{8,112} < 1$). **(g)** OFC memory trace inactivation: novel (session: $F_{6,96} = 1.03$, $p = 0.414$; session × 4OHT: $F_{6,96} = 1.73$, $p = 0.122$) or **(h)** familiar reinforcement conditions (session: $F_{6,108} < 1$; session × 4OHT: $F_{6,108} < 1$). **(i)** BDNF-dependent circuit function: BLA-OFC (session: $F_{6,204} = 1.10$ $p = 0.365$; session × lateralization: $F_{12,204} < 1$)

or **(j)** OFC-DMS disconnections (session: $F_{6,138} = 1.44$, $p = 0.204$; session × lateralization: $F_{12,138} < 1$). Data presented as individual points (semi-transparent) and group means (solid). Correspondence to main figures noted.

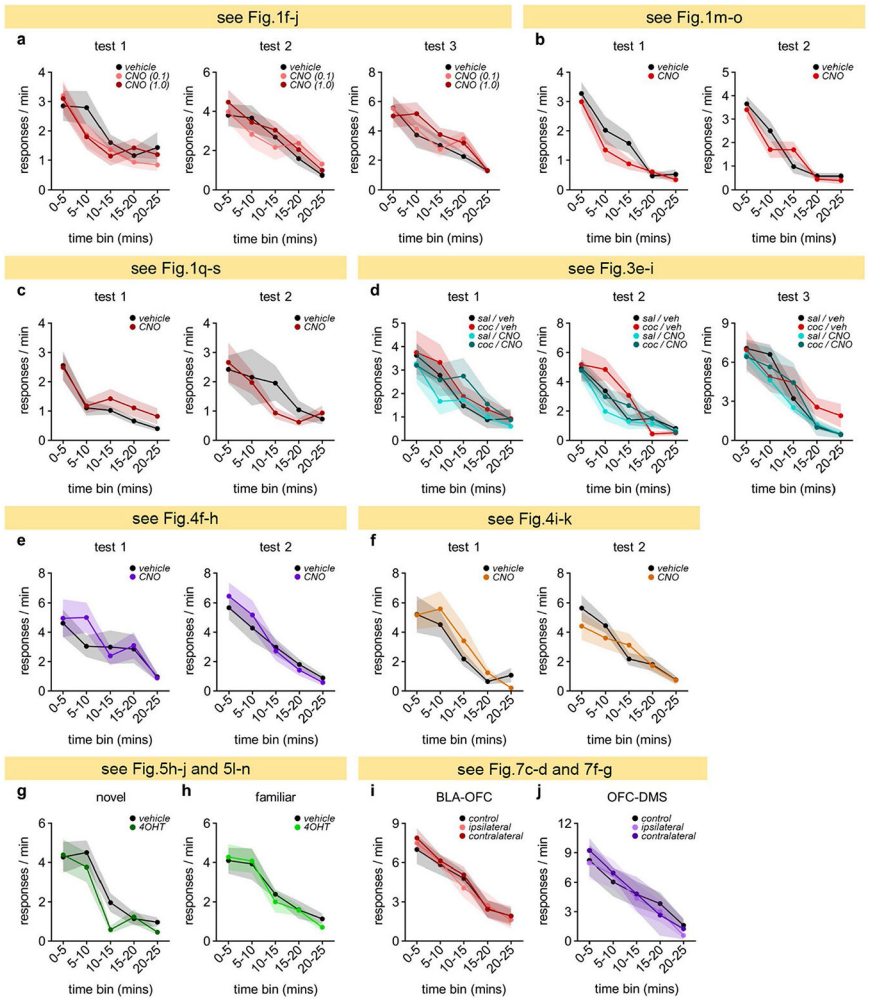


Extended Data Fig. 2 |. Nose-poking and lever-pressing actions are instrumental in nature.
(a,e) Behavioral procedure used to assess sensitivity to instrumental omission for nose-poking or lever-pressing. **(b, c)** Nose-poke responses across training ($F_{6,42} = 10.1$, $p < 0.001$) and during omission session ($F_{5,35} = 16.9$, $p < 0.001$). **(d)** Raster plot of nose-poking responses for each animal throughout the omission session. **(f, g)** Lever-pressing across training ($F_{3,21} = 29.1$, $p < 0.001$) and during omission session ($F_{5,35} = 35.0$, $p < 0.001$). **(h)** Raster plot of lever-pressing responses for each animal throughout the omission session. Data presented as individual points (semi-transparent) and group means (solid). Repeating measures ANOVA was applied, 2-sided, with no adjustment for multiple comparisons required.



Extended Data Fig. 3 |. Inactivation of posterolateral OFC does not disrupt flexible memory encoding.
(a) *Left.* Chemogenetic receptor expression in the anterior ventrolateral OFC from experiments described in Figs.1-3 of main text. *Right.* Extent of inhibitory chemogenetic receptor expression in the posterolateral OFC. Anterior-posterior (A-P) distance from bregma noted. **(b)** Timing of CNO administration for posterolateral OFC inactivation during memory encoding. **(c)** Responses across training (session: $F_{6,84} = 73.9$, $p < 0.001$; session × virus: $F_{6,84} < 1$). **(d**, **e)** Responses during first (reinforcement: $F_{1,14} = 8.49$, $p = 0.011$; reinforcement × virus: $F_{1,14} = 1.59$, $p = 0.228$) and second choice tests (reinforcement:

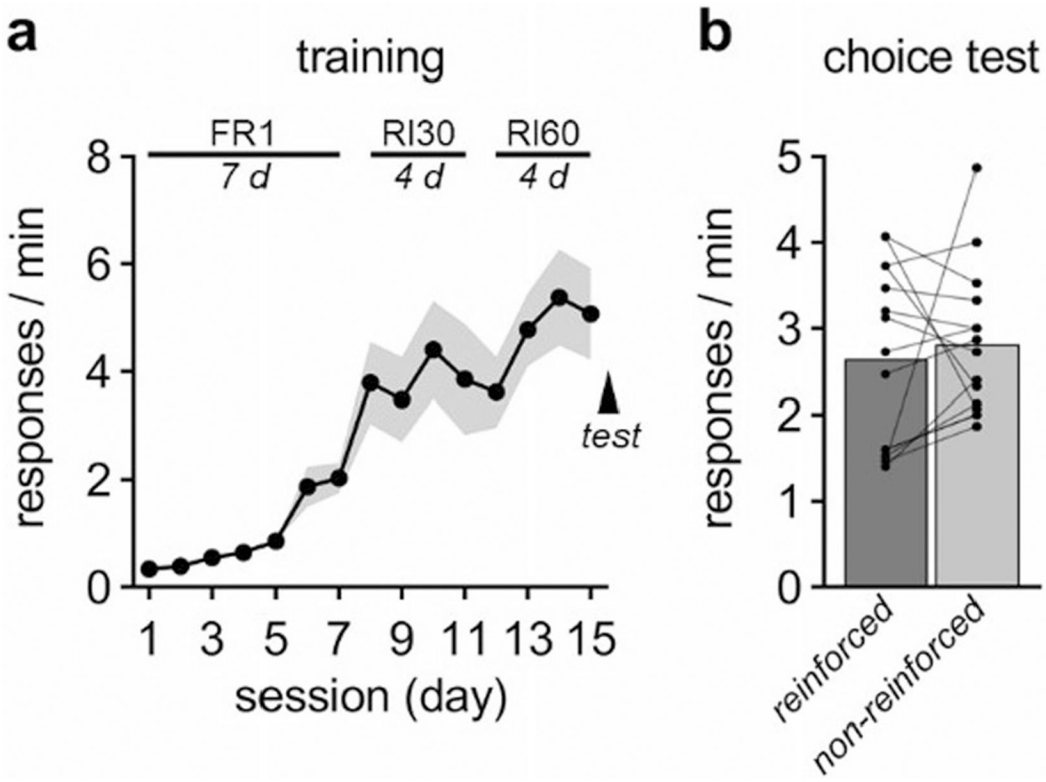
$F_{1,14}$ = 26.9, p < 0.001; reinforcement × virus: $F_{1,14}$ < 1). Choice tests were performed on sequential days. Data presented as individual points or mean ± S.E.M. *p < 0.05 (main effect). *n* = 8 GFP, 8 hM4Di mice. Correspondence to main figures noted. Analyses were performed by ANOVA (2-sided) with repeating measures when appropriate; no adjustments for multiple comparisons required.



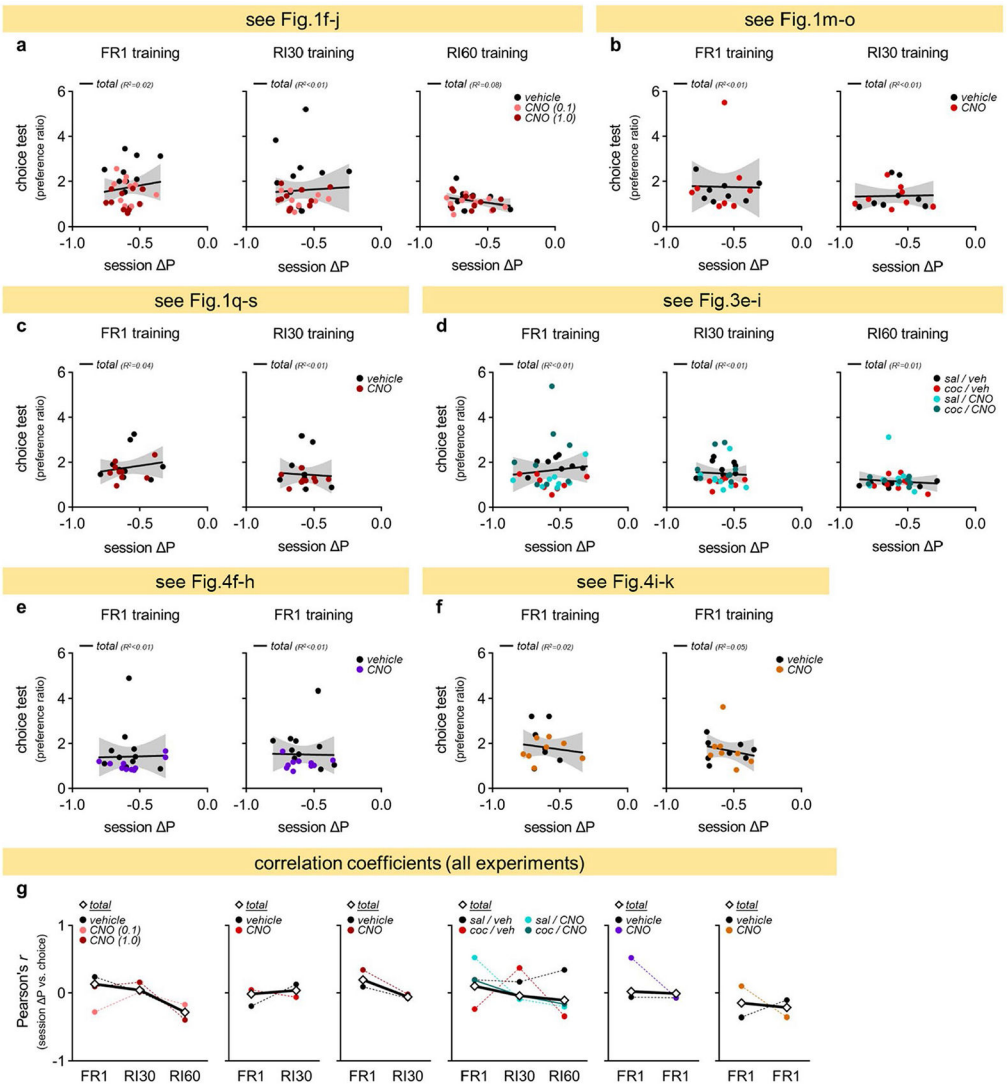
Extended Data Fig. 4 |. Responding during non-reinforced sessions did not differ between groups prior to choice tests.

All non-reinforced sessions were performed drug- and manipulation-free. **(a)** BLA→OFC inactivation (memory encoding): test 1 (time: $F_{4,108}$ = 16.5, p < 0.001; time × CNO: $F_{8,108}$ < 1), test 2 (time: $F_{4,108}$ = 17.9, p < 0.001; time × CNO: $F_{8,108}$ < 1), or test 3 (time: $F_{4,108}$ = 16.2, p < 0.001; time × CNO: $F_{8,108}$ < 1). **(b)** BLA→OFC inactivation (delayed memory encoding): test 1 (time: $F_{4,56}$ = 29.9, p < 0.001; time × CNO: $F_{4,56}$ < 1) or test 2 (time: $F_{4,56}$ = 35.2, p < 0.001; time × CNO: $F_{4,56}$ = 1.63, p = 0.179). **(c)** BLA→OFC inactivation (memory retrieval): test 1 (time: $F_{4,64}$ = 12.2, p < 0.001; time × CNO: $F_{4,64}$ < 1) or test 2 (time: $F_{4,64}$ = 7.17, p < 0.001; time × CNO: $F_{4,64}$ < 1). **(d)** BLA→OFC stimulation: test 1 (time: $F_{4,112}$ = 16.3, p < 0.001; time × cocaine: $F_{4,112}$ < 1; time × CNO: $F_{4,112}$ = 1.18, p = 0.324; time × cocaine × CNO: $F_{4,112}$ < 1), test 2 (time: $F_{4,112}$ = 47.0, p < 0.001; time ×

cocaine: $F_{4,112} = 2.61$, $p = 0.056$; time × CNO: $F_{4,112} = 2.19$, $p = 0.075$; time × cocaine × CNO: $F_{4,112} < 1$), or test 3 (time: $F_{4,112} = 55.2$, $p < 0.001$; time × cocaine: $F_{4,112} = 1.27$, $p = 0.284$; time × CNO: $F_{4,112} < 1$; time × cocaine × CNO: $F_{4,112} = 1.74$, $p = 0.147$). **(e)** OFC→DMS inactivation: test 1 (time: $F_{4,72} = 8.30$, $p < 0.001$; time × CNO: $F_{4,72} < 1$) or test 2 (time: $F_{4,72} = 40.5$, $p < 0.001$; time × CNO: $F_{4,72} < 1$). **(f)** OFC→BLA inactivation: test 1 (time: $F_{4,56} = 20.8$, $p < 0.001$; time × CNO: $F_{4,56} < 1$) or test 2 (time: $F_{4,56} = 27.1$, $p < 0.001$; time × CNO: $F_{4,56} = 1.62$, $p = 0.183$). **(g-h)** OFC memory trace inactivation: novel (time: $F_{4,84} = 26.1$, $p < 0.001$; time × 4OHT: $F_{4,84} < 1$) or familiar reinforcement conditions (time: $F_{4,72} = 19.3$, $p < 0.001$; time × 4OHT: $F_{4,72} < 1$). **(i-j)** BDNF-dependent circuit function: BLA-OFC (time: $F_{4,136} = 61.6$, $p < 0.001$; time × lateralization: $F_{8,136} < 1$) or OFC-DMS disconnections (time: $F_{4,92} = 31.7$, $p < 0.001$; time × lateralization: $F_{8,92} < 1$). Data presented as mean ± S.E.M. Correspondence to main figures noted.

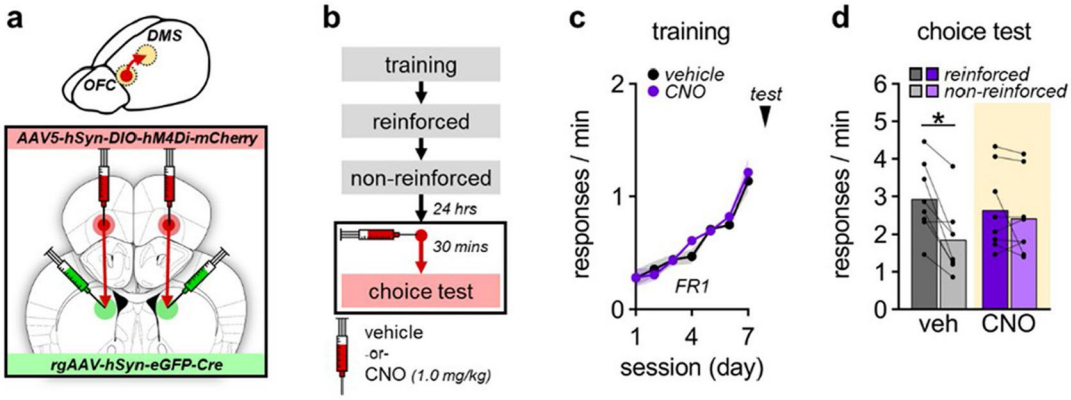


Extended Data Fig. 5 |. Extended interval training prompts inflexible choice behavior.
(a) Responses across training ($F_{14,196} = 16.9$, $p < 0.001$). **(b)** Choice test responses ($t_{14} < 1$). Data presented as individual points or mean ± S.E.M. $n = 15$ mice. Analyses were performed by ANOVA with repeating measures, and paired *t*-test (2-sided).



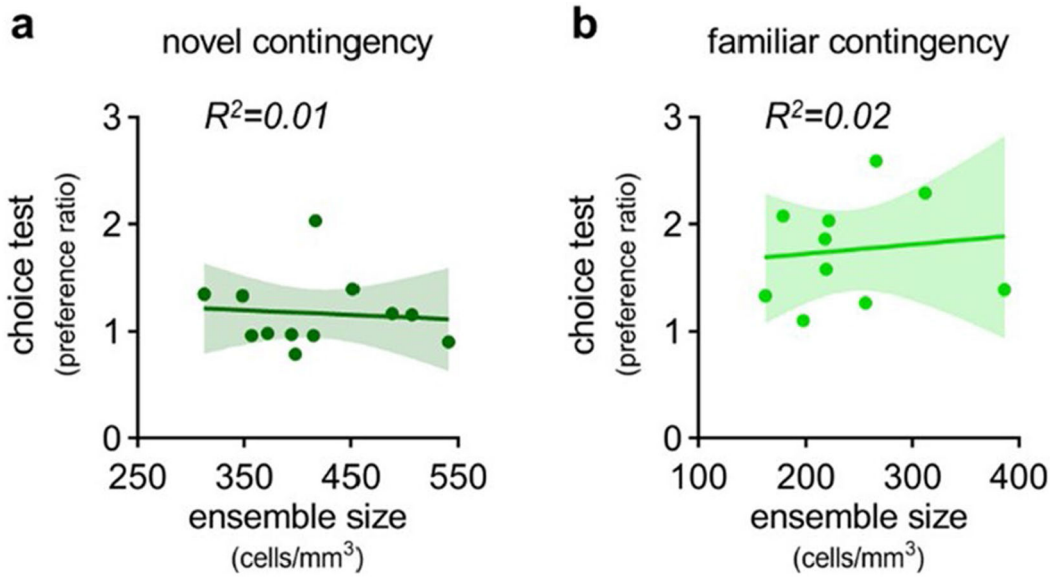
Extended Data Fig. 6 | Correlations between choice behavior and relative experience frequency of reinforced *vs.* non-reinforced nose pokes.

Correlation between individual choice test preference ratios (reinforced / non-reinforced) and the standard contingency measure (ΔP; see Methods) for each 25-minute non-reinforced session. **(a)** BLA→OFC inactivation: memory encoding (FR1: $F_{1,28} < 1$; RI30: $F_{1,28} < 1$; RI60: $F_{1,28} = 2.40$, $p = 0.132$), **(b)** delayed memory encoding (all $F_{1,14} < 1$), or **(c)** memory retrieval (all $F_{1,16} < 1$). **(d)** BLA→OFC stimulation (all $F_{1,30} < 1$). **(e)** OFC→DMS inactivation (all $F_{1,18} < 1$). **(f)** OFC→BLA inactivation (all $F_{1,14} < 1$). **(g)** Correlation coefficients (Pearson's *r*) between session ΔP and choice test preference ratios for all experiments in panels **a-f** (in order). Data presented as individual points or group means. 95% confidence interval (grey shading). Correspondence to main figures noted.



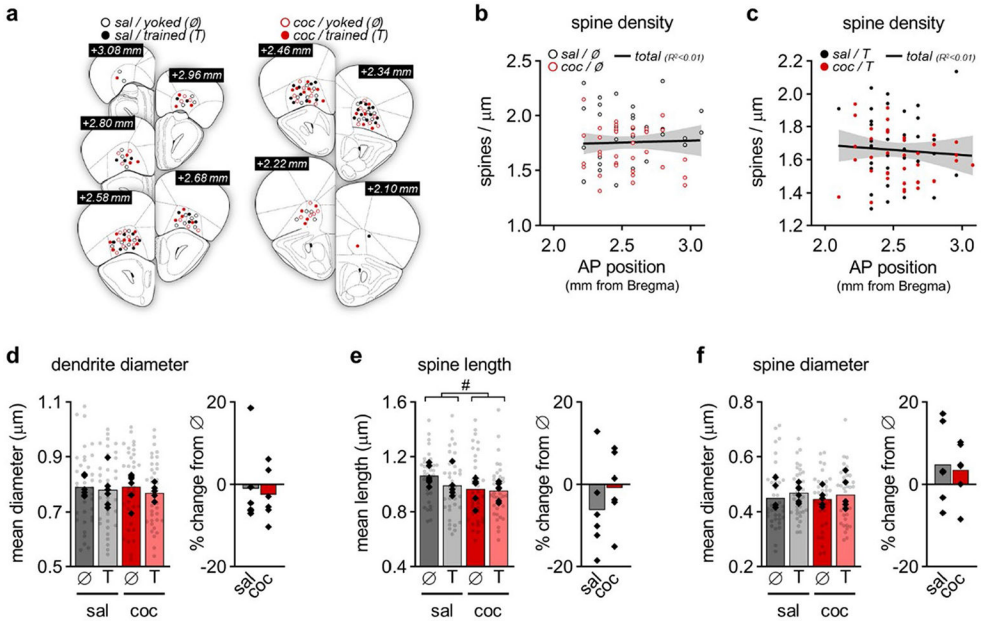
Extended Data Fig. 7 | Chemogenetic inactivation of OFC→DMS projections disrupts memory retrieval independent of repeated testing.

(a) Combinatorial viral targeting of OFC→DMS projections. **(b)** Timing of CNO administration for OFC→DMS projection inactivation during memory retrieval. **(c)** Responses across training (session: $F_{6,48} = 33.3$, $p < 0.001$; session × CNO: $F_{6,48} < 1$). **(d)** Choice test responses (reinforcement: $F_{1,14} = 15.2$, $p = 0.002$; reinforcement × CNO: $F_{1,14} = 6.74$, $p = 0.021$). Data resented as mean ± S.E.M. *$p < 0.05$ (post-hoc). $n = 8$ veh, 8 CNO mice. Experiments were replicated at least once, with concordant results.



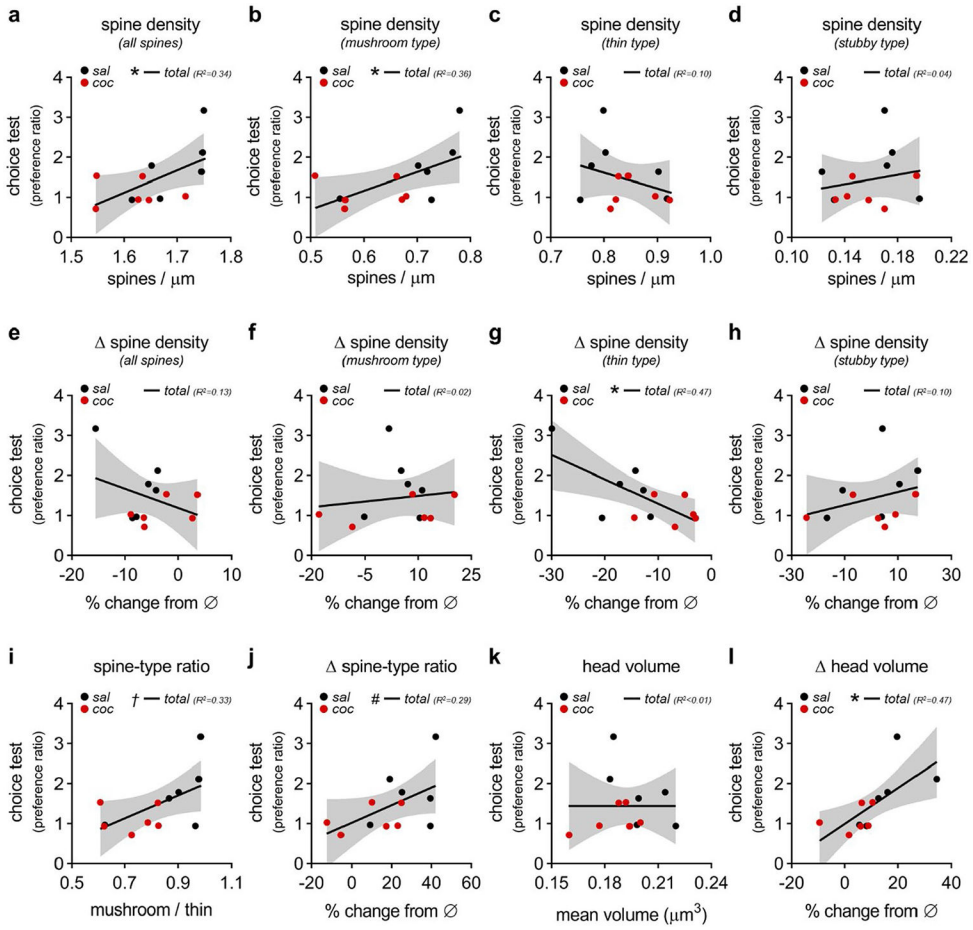
Extended Data Fig. 8 | Size of chemogenetically inactivated OFC neuronal ensembles does not predict choice behavior.

(a, b) Correlation between number of chemogenetically inactivated OFC neurons and choice test preference ratios (reinforced / non-reinforced) for OFC ensembles labelled following exposure to novel ($F_{1,10} < 1$) or familiar reinforcement conditions ($F_{1,8} < 1$). Data presented as individual points. 95% confidence interval (shading). Centre lines indicate regression.



Extended Data Fig. 9 |. Additional dendritic spine parameters among BLA→OFC→DMS relay neurons.

(a) Location of all sampled dendrites from trained (T; filled circles) and yoked (∅; open circles) mice by anterior-posterior (A-P) distance from bregma. **(b, c)** Dendritic spine density across A-P extent of the ventrolateral OFC for yoked ($F_{1,70} < 1$) and trained mice ($F_{1,70} < 1$). Centre lines indicate regression. **(d–f)** *Left panels.* Dendrite diameter (cocaine: $F_{1,20} < 1$; training: $F_{1,20} < 1$; cocaine × training: $F_{1,20} < 1$), dendritic spine length (cocaine: $F_{1,20} = 3.80$, $p = 0.053$; training: $F_{1,20} = 1.47$, $p = 0.227$; cocaine × training: $F_{1,20} < 1$) and dendritic spine diameter (cocaine: $F_{1,20} < 1$; training: $F_{1,20} < 1$; cocaine × training: $F_{1,20} < 1$). Right panels. Percent change (trained mouse vs. yoked cage mate) in dendrite diameter ($t_{10} < 1$), dendritic spine length ($t_{10} < 1$), and dendritic spine diameter ($t_{10} < 1$). Data presented as individual points (solid = per animal; transparent=per dendrite). $^{\#}p = 0.053$ (main effect). $n = 6$ sal ∅, 6 coc ∅, 6 sal T, 6 coc T mice.



Extended Data Fig. 10 | Correlations between BLA→OFC→DMS circuit-defined dendritic spine plasticity and choice behavior.

Correlation between individual choice test preference ratios (reinforced / non-reinforced) and dendritic spine parameters from Fig.6. **(a–d)** Dendritic spine density for all spines ($F_{1,10}$ = 5.12, p = 0.047), and by mushroom- ($F_{1,10}$ = 5.71, p = 0.038), thin- ($F_{1,10}$ = 1.14, p = 0.311), and stubby-type spines ($F_{1,10}$ < 1). **(e–h)** Percent change (trained mouse vs. yoked [∅] cage mate) in dendritic spine density for all spines ($F_{1,10}$ = 1.43, p = 0.259), and by mushroom- ($F_{1,10}$ < 1), thin- ($F_{1,10}$ = 8.79, p = 0.014), and stubby-type spines ($F_{1,10}$ = 1.10, p = 0.319). **(i-j)** Mushroom-to-thin spine-type ratio ($F_{1,10}$ = 4.88, p = 0.052). Percent change ($F_{1,10}$ = 4.03, p = 0.073). **(k, l)** Head volume of mushroom-type spines ($F_{1,10}$ < 1). Percent change ($F_{1,10}$ = 8.82, p = 0.014). Data presented as individual points. 95% confidence interval (grey shading). *p < 0.05. †p = 0.052. #p = 0.073. Panels **g** and **l** reproduced in Fig.6. Centre lines indicate regression.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank R. A. Davies for technical assistance and laboratory members for feedback on the manuscript. This work was supported by National Institutes of Health grants F30MH117873 (D.C.L.), R01MH117103 (S.L.G.) and R01DA044297 (S.L.G.). The Emory Viral Vector Core is supported by National Institute of Neurological Disorders and Stroke Core Facilities grant P30NS055077. The Emory National Primate Research Center is supported by Office of Research Infrastructure Programs grant P51OD011132. Research reported in this publication was also supported, in part, by the Emory University Integrated Cellular Imaging Core and Children's Healthcare of Atlanta.

Data availability

Individual data points are represented throughout. More detailed datasets are available from the corresponding author upon reasonable request.

References

1. Everitt BJ & Robbins TW Drug addiction: updating actions to habits to compulsions ten years on. Annu. Rev. Psychol 67, 23–50 (2016). [PubMed: 26253543]
2. Izquierdo A. Functional heterogeneity within rat orbitofrontal cortex in reward learning and decision making. J. Neurosci 37, 10529–10540 (2017). [PubMed: 29093055]
3. Stalnaker TA, Cooch NK & Schoenbaum G What the orbitofrontal cortex does not do. Nat. Neurosci 18, 620–627 (2015). [PubMed: 25919962]
4. Gardner MPH, Conroy JC, Sanchez DC, Zhou J & Schoenbaum G Real-time value integration during economic choice is regulated by orbitofrontal cortex. Curr. Biol 29, 4315–4322 (2019). [PubMed: 31813612]
5. Parkes SL et al. Insular and ventrolateral orbitofrontal cortices differentially contribute to goal-directed behavior in rodents. Cereb. Cortex 28, 2313–2325 (2018). [PubMed: 28541407]
6. Banerjee A. et al. Value-guided remapping of sensory cortex by lateral orbitofrontal cortex. Nature 585, 245–250 (2020). [PubMed: 32884146]
7. Schuck NW, Cai MB, Wilson RC & Niv Y Human orbitofrontal cortex represents a cognitive map of state space. Neuron 91, 1402–1412 (2016). [PubMed: 27657452]
8. Wilson RC, Takahashi YK, Schoenbaum G & Niv Y Orbitofrontal cortex as a cognitive map of task space. Neuron 81, 267–279 (2014). [PubMed: 24462094]
9. Niv Y. Learning task-state representations. Nat. Neurosci 22, 1544–1553 (2019). [PubMed: 31551597]
10. Behrens TEJ et al. What is a cognitive map? Organizing knowledge for flexible behavior. Neuron 100, 490–509 (2018). [PubMed: 30359611]
11. Hirokawa J, Vaughan A, Masset P, Ott T & Kepecs A Frontal cortex neuron types categorically encode single decision variables. Nature 576, 446–451 (2019). [PubMed: 31801999]
12. Zhou J. et al. Rat orbitofrontal ensemble activity contains multiplexed but dissociable representations of value and task structure in an odor sequence task. Curr. Biol 29, 897–907 (2019). [PubMed: 30827919]
13. Groman SM et al. Orbitofrontal circuits control multiple reinforcement-learning processes. Neuron 103, 734–746 (2019). [PubMed: 31253468]
14. Namboodiri VMK et al. Single-cell activity tracking reveals that orbitofrontal neurons acquire and maintain a long-term memory to guide behavioral adaptation. Nat. Neurosci 22, 1110–1121 (2019). [PubMed: 31160741]
15. Malvaez M, Shieh C, Murphy MD, Greenfield VY & Wassum KM Distinct cortical–amygdala projections drive reward value encoding and retrieval. Nat. Neurosci 22, 762–769 (2019). [PubMed: 30962632]
16. Barreiros IV, Panayi MC & Walton ME Organization of afferents along the anterior–posterior and medial–lateral axes of the rat orbitofrontal cortex. Neuroscience 460, 53–68 (2021). [PubMed: 33609638]

17. Schoenbaum G, Setlow B, Saddoris MP & Gallagher M Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. Neuron 39, 855–867 (2003). [PubMed: 12948451]
18. Rudebeck PH, Mitz AR, Chacko RV & Murray EA Effects of amygdala lesions on reward–value coding in orbital and medial prefrontal cortex. Neuron 80, 1519–1531 (2013). [PubMed: 24360550]
19. de Wit S, Ostlund SB, Balleine BW & Dickinson A Resolution of conflict between goal-directed actions: outcome encoding and neural control processes. J. Exp. Psychol. Anim. Behav. Process 35, 382–393 (2009). [PubMed: 19594283]
20. Dickinson A, Nicholas DJ & Adams AD The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. Q. J. Exp. Psychol. B 35, 35–51 (1983).
21. Lucantonio F, Stalnaker TA, Shaham Y, Niv Y & Schoenbaum G The impact of orbitofrontal dysfunction on cocaine addiction. Nat. Neurosci 15, 358–366 (2012). [PubMed: 22267164]
22. Ersche KD et al. Carrots and sticks fail to change behavior in cocaine addiction. Science 352, 1468–1471 (2016). [PubMed: 27313048]
23. DePoy LM, Zimmermann KS, Marvar PJ & Gourley SL Induction and blockade of adolescent cocaine-induced habits. Biol. Psychiatry 81, 595–605 (2017). [PubMed: 27871669]
24. Northoff G & Tumati S ‘Average is good, extremes are bad’—non-linear inverted U-shaped relationship between neural mechanisms and functionality of mental features. Neurosci. Biobehav. Rev 104, 11–25 (2019). [PubMed: 31251964]
25. Pan WX, Mao T & Dudman JT Inputs to the dorsal striatum of the mouse reflect the parallel circuit architecture of the forebrain. Front. Neuroanat 4, 147 (2010). [PubMed: 21212837]
26. Gremel CM & Costa RM Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. Nat. Commun 4, 2264 (2013). [PubMed: 23921250]
27. Zimmermann KS, Yamin JA, Rainnie DG, Ressler KJ & Gourley SL Connections of the mouse orbitofrontal cortex and regulation of goal-directed action selection by brain-derived neurotrophic factor. Biol. Psychiatry 81, 366–377 (2017). [PubMed: 26786312]
28. Saddoris MP, Gallagher M & Schoenbaum G Rapid associative encoding in basolateral amygdala depends on connections with orbitofrontal cortex. Neuron 46, 321–331 (2005). [PubMed: 15848809]
29. Lucantonio F. et al. Neural estimates of imagined outcomes in basolateral amygdala depend on orbitofrontal cortex. J. Neurosci 35, 16521–16530 (2015). [PubMed: 26674876]
30. Berry KP & Nedivi E Spine dynamics: are they all the same? Neuron 96, 43–55 (2017). [PubMed: 28957675]
31. DeNardo LA, Berns DS, DeLoach K & Luo L Connectivity of mouse somatosensory and prefrontal cortex examined with trans-synaptic tracing. Nat. Neurosci 18, 1687–1697 (2015). [PubMed: 26457553]
32. Yang G, Pan F & Gan WB Stably maintained dendritic spines are associated with lifelong memories. Nature 462, 920–924 (2009). [PubMed: 19946265]
33. Matsuzaki M, Honkura N, Ellis-Davies GC & Kasai H Structural basis of long-term potentiation in single dendritic spines. Nature 429, 761–766 (2004). [PubMed: 15190253]
34. Chao MV Neurotrophins and their receptors: a convergence point for many signalling pathways. Nat. Rev. Neurosci 4, 299–309 (2003). [PubMed: 12671646]
35. Gallagher M, McMahan RW & Schoenbaum G Orbitofrontal cortex and representation of incentive value in associative learning. J. Neurosci 19, 6610–6614 (1999). [PubMed: 10414988]
36. Pickens CL et al. Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. J. Neurosci 23, 11078–11084 (2003). [PubMed: 14657165]
37. Noonan MP, Chau BKH, Rushworth MFS & Fellows LK Contrasting effects of medial and lateral orbitofrontal cortex lesions on credit assignment and decision-making in humans. J. Neurosci 37, 7023–7035 (2017). [PubMed: 28630257]
38. Jocham G. et al. Reward-guided learning with and without causal attribution. Neuron 90, 177–190 (2016). [PubMed: 26971947]

39. Panayi MC & Killcross S Functional heterogeneity within the rodent lateral orbitofrontal cortex dissociates outcome devaluation and reversal learning deficits. eLife 7, e37357 (2018). [PubMed: 30044220]
40. Ostlund SB & Balleine BW Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. J. Neurosci 27, 4819–4825 (2007). [PubMed: 17475789]
41. Ramirez-Lugo L, Penas-Rincon A, Angeles-Duran S & Sotres-Bayon F Choice behavior guided by learned, but not innate, taste aversion recruits the orbitofrontal cortex. J. Neurosci 36, 10574–10583 (2016). [PubMed: 27733609]
42. Farovik A. et al. Orbitofrontal cortex encodes memories within value-based schemas and represents contexts that guide memory retrieval. J. Neurosci 35, 8333–8344 (2015). [PubMed: 26019346]
43. DeNardo LA et al. Temporal evolution of cortical ensembles promoting remote memory retrieval. Nat. Neurosci 22, 460–469 (2019). [PubMed: 30692687]
44. Josselyn SA & Tonegawa S Memory engrams: recalling the past and imagining the future. Science 367, eaaw4325 (2020). [PubMed: 31896692]
45. Wikenheiser AM, Marrero-Garcia Y & Schoenbaum G Suppression of ventral hippocampal output impairs integrated orbitofrontal encoding of task structure. Neuron 95, 1197–1207 (2017). [PubMed: 28823726]
46. Zhou J et al. Complementary task structure representations in hippocampus and orbitofrontal cortex during an odor sequence task. Curr. Biol 29, 3402–3409 (2019). [PubMed: 31588004]
47. McGaugh JL The amygdala modulates the consolidation of memories of emotionally arousing experiences. Annu. Rev. Neurosci 27, 1–28 (2004). [PubMed: 15217324]
48. Leong YC, Radulescu A, Daniel R, DeWoskin V & Niv Y Dynamic interaction between reinforcement learning and attention in multidimensional environments. Neuron 93, 451–463 (2017). [PubMed: 28103483]
49. Niv Y. et al. Reinforcement learning in multidimensional environments relies on attention mechanisms. J. Neurosci 35, 8145–8157 (2015). [PubMed: 26019331]
50. DePoy LM & Gourley SL Synaptic cytoskeletal plasticity in the prefrontal cortex following psychostimulant exposure. Traffic 16, 919–940 (2015). [PubMed: 25951902]
51. Wright WJ et al. Silent synapses dictate cocaine memory destabilization and reconsolidation. Nat. Neurosci 23, 32–46 (2020). [PubMed: 31792465]
52. Kanta V, Pare D & Headley DB Closed-loop control of gamma oscillations in the amygdala demonstrates their role in spatial memory consolidation. Nat. Commun 10, 3970 (2019). [PubMed: 31481701]
53. Huff ML, Miller RL, Deisseroth K, Moorman DE & LaLumiere RT Posttraining optogenetic manipulations of basolateral amygdala activity modulate consolidation of inhibitory avoidance memory in rats. Proc. Natl Acad. Sci. USA 110, 3597–3602 (2013). [PubMed: 23401523]
54. Jones JL et al. Orbitofrontal cortex supports behavior and learning using inferred but not cached values. Science 338, 953–956 (2012). [PubMed: 23162000]
55. Vertechi P. et al. Inference-based decisions in a hidden state foraging task: differential contributions of prefrontal cortical areas. Neuron 106, 166–176 (2020). [PubMed: 32048995]
56. Stalnaker TA et al. Orbitofrontal neurons infer the value and identity of predicted outcomes. Nat. Commun 5, 3926 (2014). [PubMed: 24894805]
57. Takahashi YK et al. Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. Nat. Neurosci 14, 1590–1597 (2011). [PubMed: 22037501]
58. Hayashi-Takagi A. et al. Labelling and optical erasure of synaptic memory traces in the motor cortex. Nature 525, 333–338 (2015). [PubMed: 26352471]
59. Wu YI et al. A genetically encoded photoactivatable Rac controls the motility of living cells. Nature 461, 104–108 (2009). [PubMed: 19693014]
60. Feng G. et al. Imaging neuronal subsets in transgenic mice expressing multiple spectral variants of GFP. Neuron 28, 41–51 (2000). [PubMed: 11086982]
61. Rios M. et al. Conditional deletion of brain-derived neurotrophic factor in the postnatal brain leads to obesity and hyperactivity. Mol. Endocrinol 15, 1748–1757 (2001). [PubMed: 11579207]

62. Dickinson A & Balleine B Motivational control of goal-directed action. Anim. Learn. Behav 22, 1–18 (1994).

63. Vaghi MM et al. Action-outcome knowledge dissociates from behavior in obsessive-compulsive disorder following contingency degradation. Biol. Psychiatry Cogn. Neurosci. Neuroimaging 4, 200–209 (2019). [PubMed: 30545754]

64. Rossi MA & Yin HH Methods for studying habitual behavior in mice. Curr. Protoc. Neurosci 60, 8.29.1–8.29.9 (2012).

65. Krashes MJ et al. Rapid, reversible activation of AgRP neurons drives feeding behavior in mice. J. Clin. Invest 121, 1424–1428 (2011). [PubMed: 21364278]

66. Gomez JL et al. Chemogenetics revealed: DREADD occupancy and activation via converted clozapine. Science 357, 503–507 (2017). [PubMed: 28774929]

67. Huang L. et al. Organizational principles of amygdalar input-output neuronal circuits. Mol. Psychiatry 26, 7118–7129 (2021). [PubMed: 34400771]

68. Wang J et al. Divergent projection patterns revealed by reconstruction of individual neurons in orbitofrontal cortex. Neurosci. Bull 37, 461–477 (2021). [PubMed: 33373031]

69. Shinonaga Y, Takada M & Mizuno N Topographic organization of collateral projections from the basolateral amygdaloid nucleus to both the prefrontal cortex and nucleus accumbens in the rat. Neuroscience 58, 389–397 (1994). [PubMed: 8152545]

70. Paxinos G & Franklin KBJ The Mouse Brain in Stereotaxic Coordinates 2nd edn (Academic Press, 2001).

71. Schindelin J. et al. Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682 (2012). [PubMed: 22743772]

72. Zingg B. et al. AAV-mediated anterograde transsynaptic tagging: mapping corticocollicular input-defined neural pathways for defense behaviors. Neuron 93, 33–47 (2017). [PubMed: 27989459]

73. Radley JJ, Anderson RM, Hamilton BA, Alcock JA & Romig-Martin SA Chronic stress-induced alterations of dendritic spine subtypes predict functional decrements in an hypothalamo-pituitary-adrenal-inhibitory prefrontal circuit. J. Neurosci 33, 14379–14391 (2013). [PubMed: 24005291]

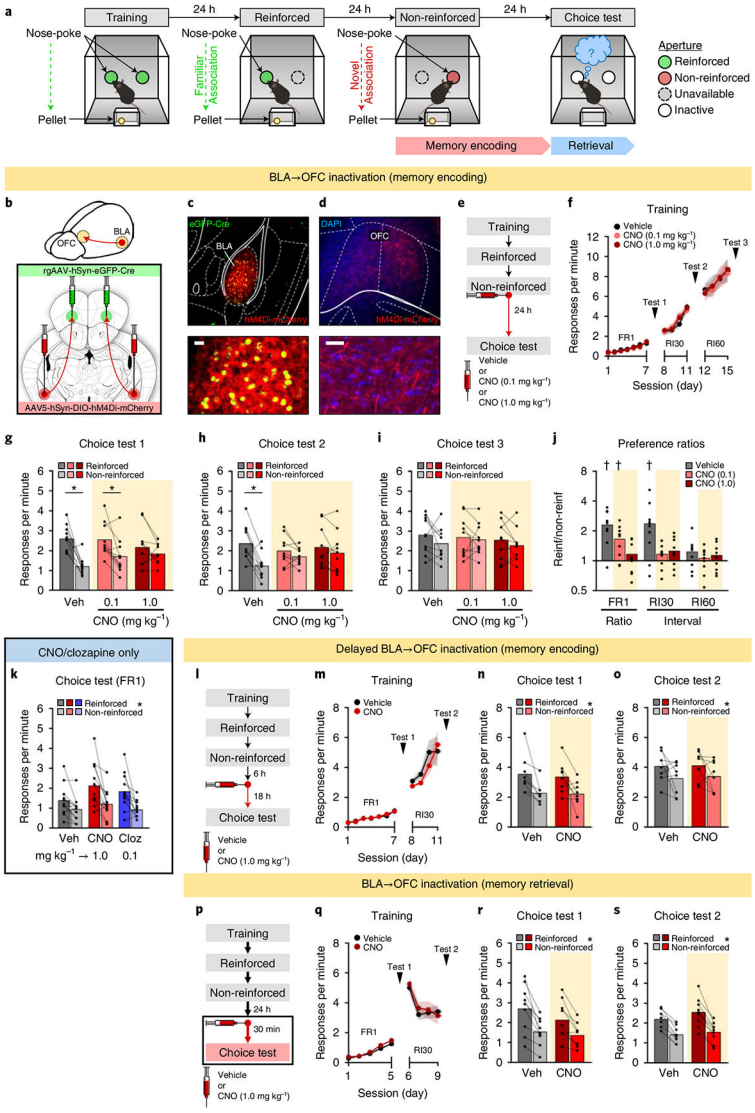


Fig. 1 | BLA→OFC projections are necessary for encoding, but not retrieving, new action memories for sustained response flexibility.

a, Mice were trained to generate two food-reinforced nose-poke responses (training/reinforced), and then the food associated with one response was delivered independent of nose-poking (non-reinforced), triggering new memory encoding necessary for adaptive choice the next day (choice test). **b**, Combinatorial viral targeting of BLA→OFC projections. **c**,**d**, Retrogradely transported eGFP-Cre driving hM4Di-mCherry expression in the BLA (**c**) with axon terminals detectable in the ventrolateral OFC (**d**). Scale bar, 25 μm. **e**, Timing of CNO administration for BLA→OFC projection inactivation during memory encoding. **f**, Responses across training (session: $F_{14,378} = 123$; P<0.001; session × CNO: $F_{28,378} < 1$). **g**–**i**, Choice test responses after ratio (choice test 1 (**g**); reinforcement: $F_{1,27} = 31.2$; $P < 0.001$; reinforcement × CNO: $F_{2,27} = 4.07$; $P = 0.029$), moderate interval training (choice test 2 (**h**); reinforcement: $F_{1,27} = 18.2$; $P < 0.001$; reinforcement × CNO: $P_{2,27} = 4.40$; $P = 0.022$) or extended interval training (choice test 3 (**i**); reinforcement: $F_{1,27} = 3.54$; $P = 0.071$; reinforcement × CNO: $F_{2,27} < 1$). **j**, Choice test response preference

ratios (training schedule: $F_{2,54} = 8.43$; $P = 0.001$, training schedule × CNO: $F_{4,54} = 3.73$; $P = 0.009$). **k**, CNO (1.0 mg kg^{-1}) and clozapine (0.1 mg kg^{-1}) administration during memory-encoding absent chemogenetic receptor expression (reinforcement: $F_{1,35} = 32.3$, $P < 0.001$; reinforcement × drug: $F_{2,35} = 1.33$, $P = 0.278$). **l**, Timing of CNO administration for BLA→OFC projection inactivation 6 hours after non-reinforced session. **m**, Responses across training (session: $F_{10,80} = 51.9$, $P < 0.001$; session × CNO: $F_{10,80} < 1$). **n**,**o**, Choice test responses after ratio (choice test 1 (**n**); reinforcement: $F_{1,14} = 16.9$, $P = 0.001$; reinforcement × CNO: $F_{1,14} < 1$) or moderate interval training (choice test 2 (**o**); reinforcement: $F_{1,14} = 5.26$, $P = 0.038$; reinforcement × CNO: $F_{1,14} < 1$). **p**, Timing of CNO administration for BLA→OFC projection inactivation during memory retrieval. **q**, Responses across training (session: $F_{8,128} = 71.5$, $P < 0.001$; session × CNO: $F_{8,128} < 1$). **r**,**s**, Choice test responses after ratio (choice test 1 (**r**); reinforcement: $F_{1,16} = 34.5$; $P < 0.001$; reinforcement × CNO: $F_{1,16} = 1.44$; $P = 0.247$) or moderate interval training (choice test 2 (**s**); reinforcement: $F_{1,16} = 129$; $P < 0.001$; reinforcement × CNO: $F_{1,16} = 2.01$; $P = 0.176$). Data are presented as individual points or mean ± s.e.m. $^{*}P < 0.05$ (main effect or post hoc). $^{\dagger}P < 0.05$ (one-sample test versus 1). Experiments were replicated at least once, with concordant results. See Supplementary Table 1 for complete statistics.

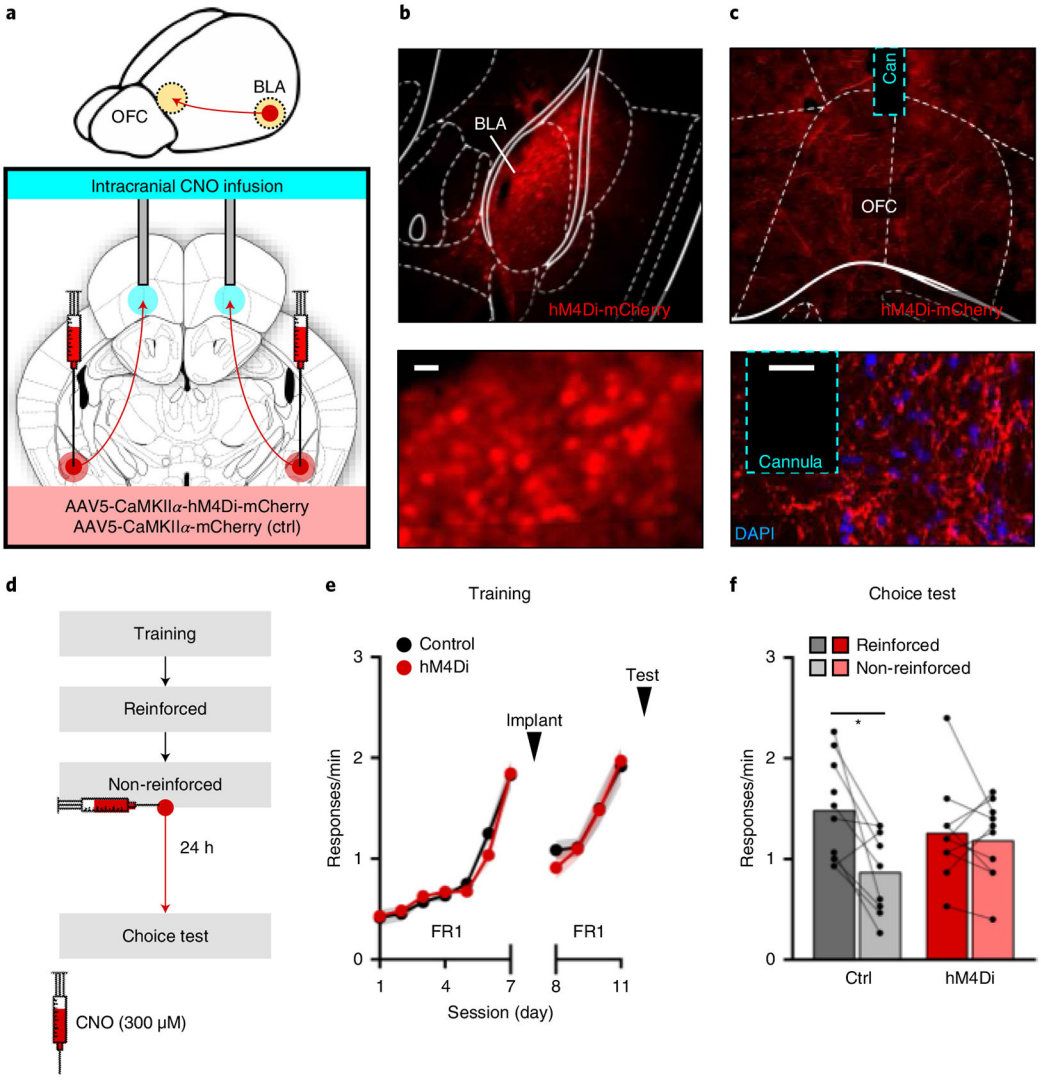


Fig. 2 | Selective inactivation of BLA→OFC axon terminals is sufficient to disrupt flexible memory encoding.

a, Intracranial targeting of BLA→OFC axon terminals. **b**,**c**, Cannula placement within the OFC (**c**) targeting axon terminals from hM4Di-mCherry-expressing BLA projection neurons (**b**). Scale bar, 25 μm. **d**, Timing of intracranial CNO infusions for BLA→OFC projection inactivation during memory encoding. **e**, Responses across training (session: $F_{9,63} = 18.7$, $P < 0.001$; session × virus: $F_{9,63} < 1$). **f**, Choice test responses (reinforcement: $F_{1,17} = 7.30$, $P = 0.015$; reinforcement × virus: $F_{1,17} = 4.72$, $P = 0.034$). Data are presented as individual points or mean ± s.e.m., with groups compared by two-factor ANOVA with repeating measures. $^*P < 0.05$ (post hoc). $n = 9$ control and 10 hM4Di mice. Experiments were replicated at least once, with concordant results.

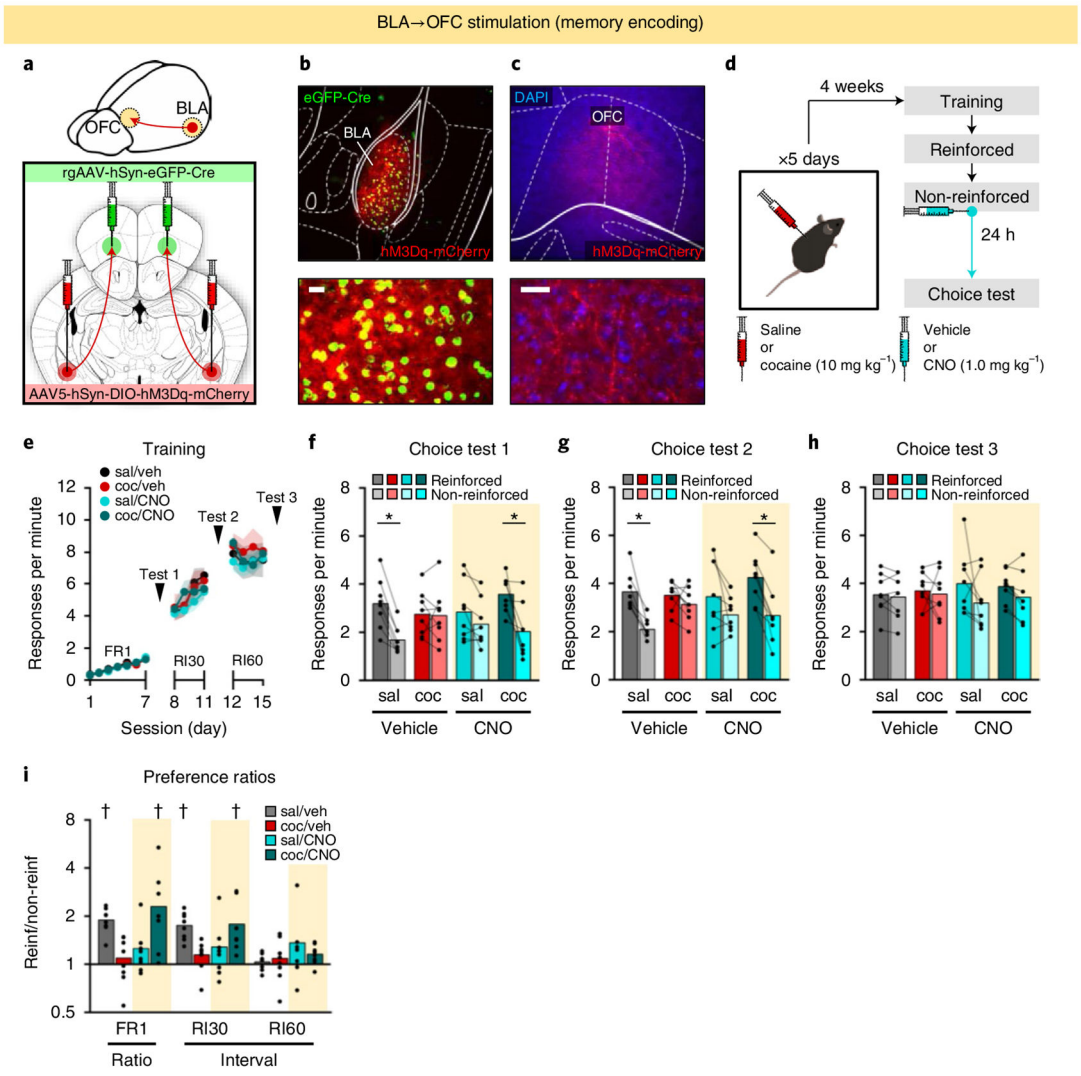


Fig. 3 | Stimulating BLA→OFC projections reinstates flexible action memory encoding after cocaine.

a, Combinatorial viral targeting of BLA→OFC projections. **b**,**c**, Retrogradely transported eGFP-Cre driving hM4Di-mCherry expression in the BLA (**b**) with axon terminals detectable in the ventrolateral OFC (**c**). Scale bar, 25 μm. **d**, Cocaine administration before behavioral testing and timing of CNO administration for BLA→OFC projection stimulation during memory encoding. **e**, Responses across training (session: $F_{14,392} = 17.23$; $P < 0.001$; session × cocaine × CNO: $F_{14,392} < 1$). **f**–**h**, Choice test responses after ratio (choice test 1 (**f**); reinforcement: $F_{1,28} = 31.3$; $P < 0.001$; reinforcement × cocaine × CNO: $F_{1,28} = 15.0$; $P < 0.001$), moderate interval training (choice test 2 (**g**); reinforcement: $F_{1,28} = 45.3$; $P < 0.001$; reinforcement × cocaine × CNO: $F_{1,28} = 10.0$; $P = 0.004$) or extended interval training (choice test 3 (**h**); reinforcement: $F_{1,28} = 3.83$; $P = 0.061$; reinforcement × cocaine × CNO: $F_{1,28} < 1$). **i**, Choice test response preference ratios (training schedule: $F_{2,56} = 5.28$; $P = 0.008$, training schedule × cocaine × CNO: $F_{2,56} = 6.39$; $P = 0.003$). Data are presented as individual points or mean ± s.e.m. $^{*}P < 0.05$ (main effect or post hoc). $^{\dagger}P <$

0.05 (one-sample test versus 1). Experiments were replicated at least once, with concordant results. See Supplementary Table 2 for complete statistics.

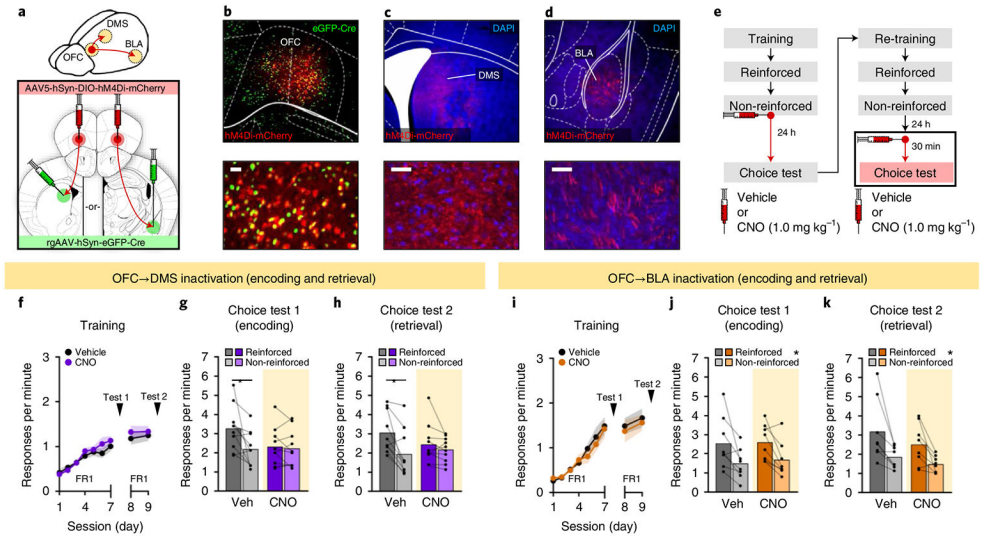


Fig. 4 | OFC→DMS, but not OFC→BLA, projections are necessary for the encoding and retrieval of new action memories for sustained response flexibility.

a, Combinatorial viral targeting of OFC→DMS or OFC→BLA projections. **b**–**d**, Retrogradely transported eGFP-Cre driving hM4Di-mCherry expression in the ventrolateral OFC (**b**) with axon terminals detectable in the DMS (**c**) or BLA (**d**). Scale bar, 25 μm. **e**, Timing of CNO administration for OFC→DMS or OFC→BLA projection inactivation during memory encoding or retrieval. **f**, Responses across training (session: $F_{8,144} = 29.2$; $P < 0.001$; session × CNO: $F_{8,144} < 1$). **g**,**h**, Choice test responses for memory-encoding (**g**) (reinforcement: $F_{1,18} = 6.92$; $P = 0.017$; reinforcement × CNO $F_{1,18} = 4.79$; $P = 0.042$) and retrieval-targeted (**h**) OFC→DMS projection inactivation (reinforcement: $F_{1,18} = 13.9$; $P = 0.002$; reinforcement × CNO $F_{1,18} = 5.17$; $P = 0.036$). **i**, Responses across training (session: $F_{8,112} = 37.5$; $P < 0.001$; session × CNO: $F_{8,112} < 1$). **j**,**k**, Choice test responses for encoding-targeted (**j**) (reinforcement: $F_{1,14} = 17.0$; $P < 0.001$; reinforcement × CNO $F_{1,148} < 1$) or retrieval-targeted (**k**) OFC→BLA projection inactivation (reinforcement: $F_{1,14} = 18.3$; $P = 0.001$; reinforcement × CNO $F_{1,14} < 1$). Data are presented as individual points or mean ± s.e.m. *$P < 0.05$ (main effect or post hoc). Experiments were replicated at least once, with concordant results. See Supplementary Table 3 for complete statistics.

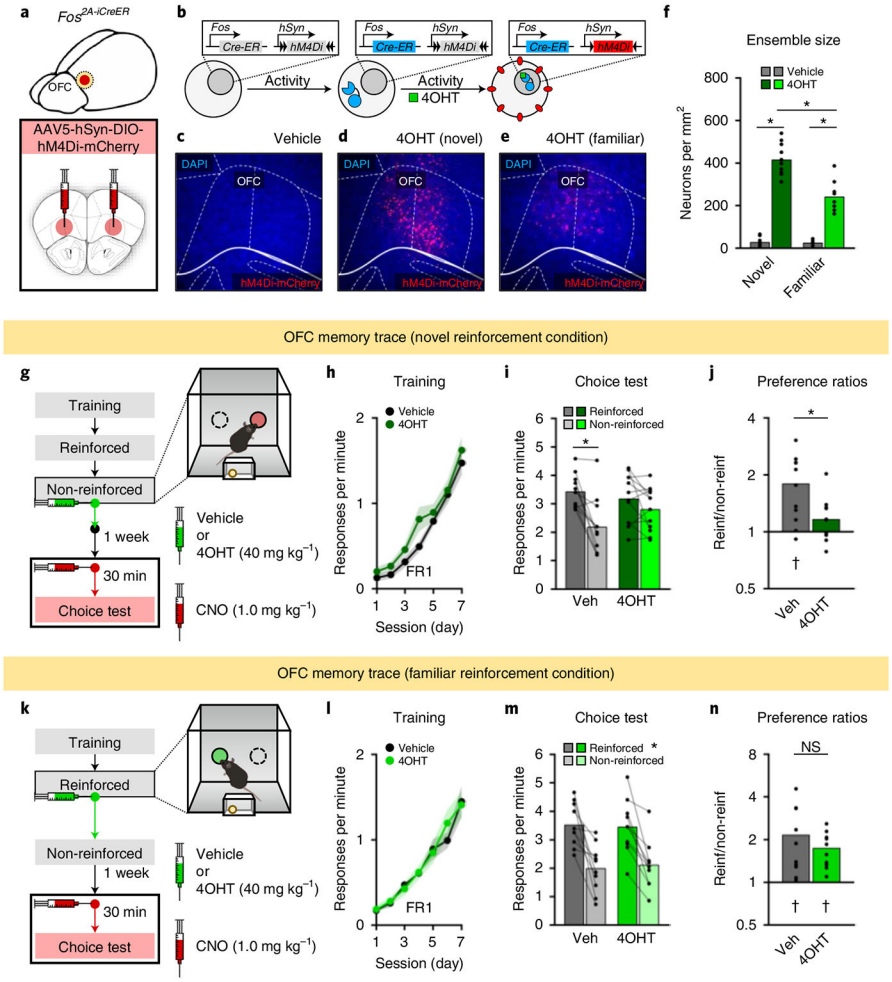


Fig. 5 | Encoding-activated neuronal ensembles in the OFC form a memory trace for later response flexibility.

a,**b**, Induction of hM4Di-mCherry expression among virally targeted OFC neurons (**a**). Activity-dependent Fos promotor and enhancer elements drive expression of Cre-ER, which is only trafficked to the nucleus to catalyze recombination after 4OHT binding (**b**). **c**–**f**, Quantification of activity-dependent (**c**) and 4OHT-dependent hM4Di-mCherry expression in the ventrolateral OFC after exposure to novel (**d**) or familiar (**e**) reinforcement conditions (novelty: $F_{1,39} = 34.3$, $P < 0.001$; 4OHT: $F_{1,39} = 393$, $P < 0.001$; novelty × 4OHT: $F_{1,39} = 37.7$, $P < 0.001$) (**f**). **g**, Timing of 4OHT and CNO administration for inactivation of novelty-responsive OFC neurons during choice test. **h**, Responses across training (session: $F_{6.96} = 50.3$, $P < 0.001$; session × 4OHT: $F_{6,96} = 1.04$, $P = 0.406$). **i**, Choice test responses (reinforcement: $F_{1,21} = 20.6$, $P < 0.001$; reinforcement × 4OHT: $F_{1,21} = 5.90$, $P = 0.024$). **j**, Choice test response preference ratios ($t_{21} = 2.81$, $P = 0.010$). **k**, Timing of 4OHT and CNO administration for inactivation of OFC neurons responsive to familiar reinforcement conditions. **l**, Responses across training (session: $F_{6.108} = 63.3$, $P < 0.001$; session × 4OHT: $F_{6,108} < 1$). **m**, Choice test responses (reinforcement: $F_{1,18} = 40.2$, $P < 0.001$; reinforcement × 4OHT: $F_{1,18} < 1$). **n**, Choice test response preference ratios ($t_{18} = 1.00$, $P = 0.330$). Data are presented as individual points or mean ± s.e.m. *$P < 0.05$ (main effect or post hoc). **P

< 0.05 (one-sample test versus 1). NS, not significant. Experiments were replicated at least once, with concordant results. See Supplementary Table 4 for complete statistics.

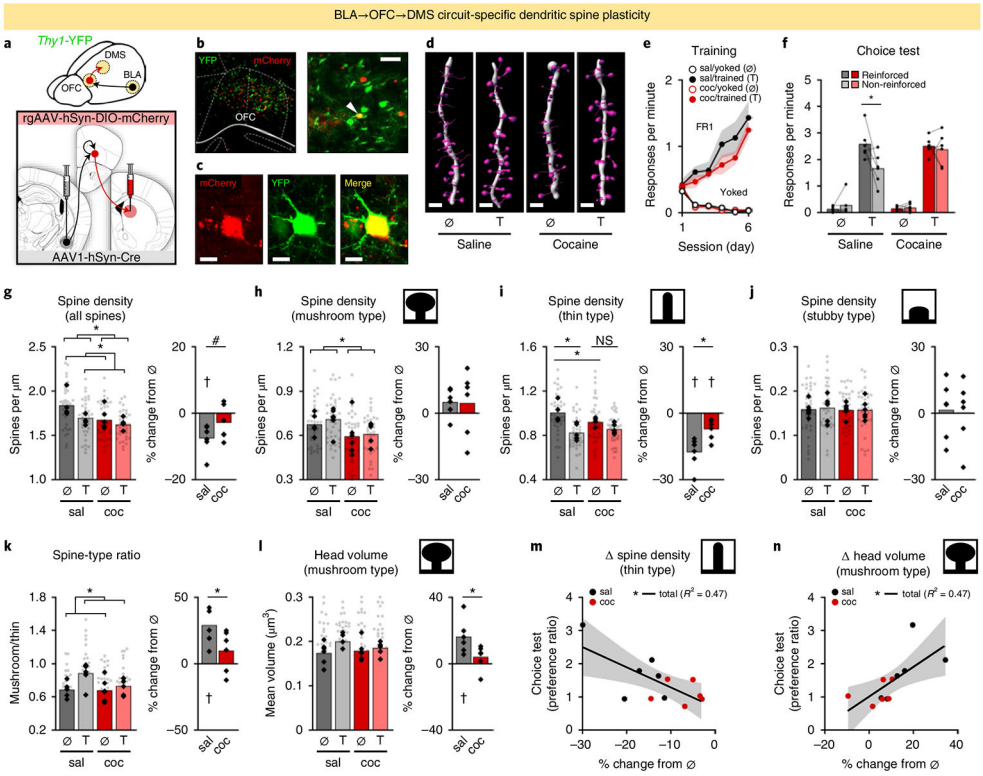


Fig. 6 | . New action learning triggers dendritic spine plasticity within a di-synaptic BLA→OFC→DMS circuit.

a, Viral-mediated trans-synaptic ARM for labeling ‘relay’ neurons with a BLA→OFC→DMS circuit. **b**,**c**, Connectivity-defined ventrolateral OFC neurons (**b**) identified by ARM for imaging using transgenic *Thy1*-driven YFP fluorescence (**c**). Top right: scale bar, 50µm; bottom: scale bar, 10µm. **d**, Representative three-dimensional dendritic spine reconstructions. Scale bar, 2µm. **e**, Responses across training for trained (T) mice and their yoked (∅) cage mates (session: $F_{5,100} = 12.6$, $P < 0.001$; session × cocaine × training: $F_{5,100} < 1$). **f**, Choice test responses (reinforcement: $F_{1,20} = 3.33$, $P = 0.083$; reinforcement × cocaine × training: $F_{1,20} = 3.68$, $P = 0.069$). **g**–**j**, Left panels: dendritic spine density for all spines (**g**) (cocaine: $F_{1,20} = 9.89$, $P = 0.005$; training: $F_{1,20} = 6.46$, $P = 0.019$; cocaine × training: $F_{1,20} = 1.37$, $P = 0.255$) and stratified by mushroom-type spines (**h**) (cocaine: $F_{1,20} = 6.11$, $P = 0.023$; training: $F_{1,20} < 1$; cocaine × training: $F_{1,20} < 1$), thin-type spines (**i**) (cocaine: $F_{1,20} = 1.01$, $P = 0.316$; training: $F_{1,20} = 20.3$, $P < 0.001$; cocaine × training: $F_{1,20} = 4.23$, $P = 0.041$) or stubby-type spines (**j**) (cocaine: $F_{1,20} < 1$; training: $F_{1,20} < 1$; cocaine × training: $F_{1,20} < 1$). Right panels: percent change (trained mouse versus yoked cage mate). **k**,**l**, Left panels: mushroom-to-thin spine-type ratio (**k**) (cocaine: $F_{1,20} = 2.89$, $P = 0.105$; training: $F_{1,20} = 7.10$, $P = 0.015$; cocaine × training: $F_{1,20} = 2.40$, $P = 0.137$) and head volume of mushroom-type spines (**l**) (cocaine: $F_{1,20} < 1$; training: $F_{1,20} = 2.78$, $P = 0.098$; cocaine × training: $F_{1,20} = 1.07$, $P = 0.303$). Right panels: percent change. **m**,**n**, Correlation between individual choice behavior and percent change in thin-type dendritic spine density (**m**) ($F_{1,10} = 8.79$, $P = 0.014$) and in head volume of mushroom-type spines (**n**) ($F_{1,10} = 8.82$, $P = 0.014$). 95% confidence interval (gray shading). Data are presented as individual points (solid, per animal; semi-transparent, per dendrite) or mean ± s.e.m.

*$P < 0.05$ (main effect, post hoc or planned comparison in **f**). **$P = 0.12$. ***$P < 0.05$ (one-sample test versus 0). NS, not significant. Experiments were replicated at least once, with concordant results. See Supplementary Table 5 for complete statistics.

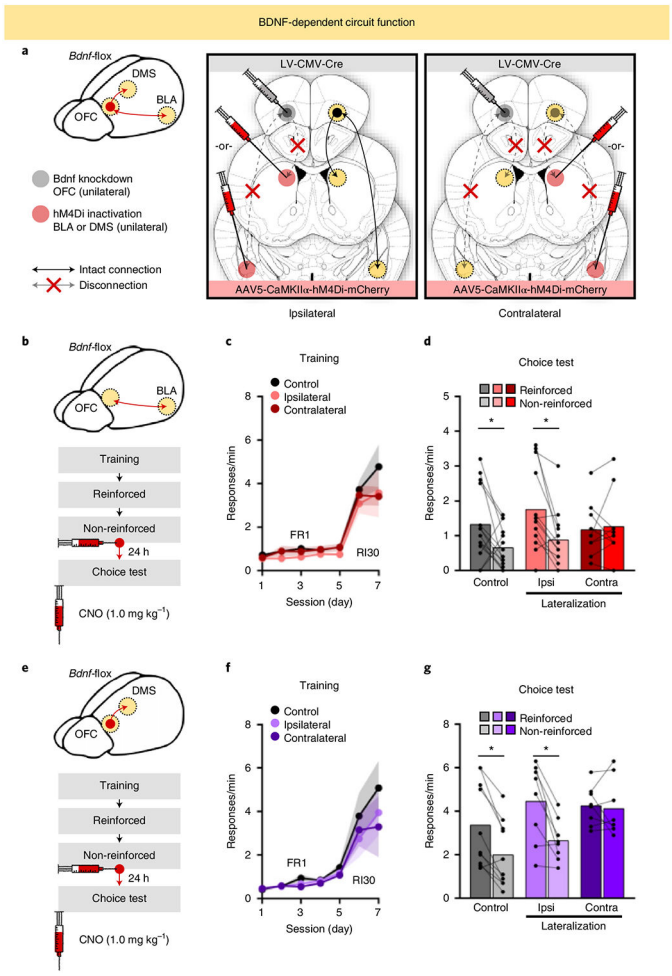


Fig. 7 | Circuit-specific neurotrophin tone in the OFC is required for new learning for sustained response flexibility.

a, Multiplexed viral-mediated method for assessing molecular–functional circuit interactions and circuit-specific BDNF function. **b**, Timing of CNO administration for unilateral BLA inactivation during memory encoding. **c**, Responses across training (session: $F_{6,204} = 37.1$, $P < 0.001$; session × lateralization: $F_{12,204} < 1$). **d**, Choice test responses (reinforcement: $F_{1,34} = 10.3$, $P = 0.003$; reinforcement × lateralization: $F_{2,34} = 3.43$, $P = 0.044$). **e**, Timing of CNO administration for unilateral DMS inactivation during memory encoding. **f**, Responses across training (session: $F_{6,138} = 22.2$, $P < 0.001$; session × lateralization: $F_{12,138} < 1$). **g**, Choice test responses (reinforcement: $F_{1,23} = 28.3$, $P < 0.001$; reinforcement × lateralization: $F_{2,23} = 5.46$, $P = 0.011$). Data are presented as individual points or mean ± s.e.m. *$P <$ 0.05 (post hoc). Experiments were replicated at least once, with concordant results. See Supplementary Table 6 for complete statistics.