bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Distinct cortical-amygdala projections drive reward value encoding and retrieval

Melissa Malvaez[1], Christine Shieh[1], Michael D. Murphy[1], Venuz Y. Greenfield[1], Kate M. Wassum[1, 2]

1. Dept. of Psychology, UCLA, Los Angeles, CA 90095. 2. Brain Research Institute, UCLA, Los Angeles, CA 90095, USA.

Correspondence:
Kate Wassum: kwassum@ucla.edu
Dept. of Psychology, UCLA
1285 Franz Hall
Box 951563
Los Angeles, CA 90095-1563

Short title: OFC-BLA mediates value encoding and retrieval

Key words: reward, encoding, learning, retrieval, memory, decision making, instrumental conditioning, incentive learning, basolateral amygdala, orbitofrontal cortex

Number of:
Words
Abstract: 161
Article body: 2714
Figures: 5
Tables: 0
Supplemental Information:
Materials and Methods: 1
Tables: 1
Figures: 18

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

The value of an anticipated rewarding event is crucial information in the decision to engage in its pursuit. The networks responsible for encoding and retrieval this value are largely unknown. Using glutamate biosensors and pharmacological manipulations, we found that basolateral amygdala (BLA) glutamatergic activity tracks and mediates both the encoding and retrieval of the hunger-state-dependent value of a palatable food reward. Projection-specific chemogenetic and optogenetic manipulations revealed and it wathe orbitofrontal cortex (OFC) supports the BLA in these processes. Critically, the function of the ventrolateral (lOFC) and medial (mOFC) OFC→BLA projections was found to be doubly dissociable. Whereas activity in lOFC→BLA projections is necessary for and sufficient to drive encoding of a positive change in the value of a reward, mOFC→BLA projections are necessary and sufficient for retrieving this value from memory to guide reward pursuit. These data reveal a new circuit for adaptive reward valuation and pursuit and provide mechanistic insight into the dysfunction in these processes that characterizes myriad psychiatric diseases.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Prospective consideration of the outcomes of potential action choices is crucial to adaptive decision making. Chief among these considerations is the value of anticipated rewarding events. This incentive information is state-dependent; e.g., a food reward is more valuable when hungry than when sated. It is also learned. The value of a specific reward is encoded during its experience in a relevant state [1]. Retrieval of the previously-encoded value of an anticipated reward allows adaptive reward pursuit decisions. Dysfunction in either the value encoding or retrieval process will lead to the aberrant reward pursuit and ill-informed decision making─ cognitive symptoms that characterize myriad psychiatric diseases. Despite importance to understanding adaptive and maladaptive behavior, little is known of the neural circuits that support reward value encoding and retrieval.

The basolateral amygdala (BLA) has long been known to mediate emotional learning [2-4]. Accordingly, this structure is necessary for reward value encoding [5-13]. But the circuitry supporting the BLA in this function is unknown. Whether the BLA participates in retrieving reward value is less clear and has been disputed [7,10,11] and its contribution, if any, to active decision making is uncertain.

Given the BLA is densely innervated by glutamatergic projections from regions themselves implicated in reward learning and decision making [14,15], we sought to begin to fill these gaps in knowledge by using electroenzymatic biosensors to characterize BLA glutamate release during performance of behavioral tests that allow us to experimentally isolate reward value encoding and retrieval (Fig. 1a) [5,6,16]. These biosensors allow sub-second, spatially-precise, sensitive, and selective measurement of neuronally-released glutamate (Fig. S1) [17-19]. Rats were trained while relatively sated (4 hr food deprived) on a self-paced, 2-lever action sequence to earn a sucrose reward wherein pressing a ‘seeking’ lever introduced a ‘taking’ lever, a press on which retracted this lever and triggered reward delivery. In the sated state, the sucrose reward has a low value and supports a low rate of lever pressing. Once stable baseline performance was achieved, rats

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

were re-exposed to the sucrose in either the familiar sated state or in a hungry, 20-hr food-deprived state. Because rats had never before experienced the sucrose while hungry, the latter provided an *incentive learning* opportunity to encode the high value of the sucrose reward in the hungry state. Re-exposure was non-contingent and conducted 'offline' (i.e., without the levers present) to isolate reward value encoding from reinforcement-related confounds and to prevent caching of value to the seeking and taking actions themselves. The effect of this incentive learning opportunity on rats' reward pursuit was then tested the following day in a brief lever-pressing probe test. No rewards were delivered during this test to force the retrieval of reward value from memory and to avoid online reinforcement confounds. Seeking presses were the primary measure because they have been shown to be selectively sensitive to learned changes in the value of an anticipated reward and relatively immune to more general motivational processes [5,6,20-22]. All rats were hungry for this test, but only those rats that had previously experienced the sucrose reward in the hungry state escalated their reward-seeking actions (Figs. 1c, S2; t_{10}=2.50, p=0.03). This result is consistent with the interpretation that the rats retrieved from memory the encoded higher value of the anticipated sucrose reward and used this information to increase their reward pursuit vigor.

BLA glutamate release was found to track reward value encoding. During the re-exposure, reward consumption triggered a transient increase in BLA glutamate concentration, but only if a new value was being encoded (i.e., re-exposure hungry; Figs. 1d-e, S3; Time: $F_{2,20}$=5.04, P=0.02; Deprivation: $F_{1,10}$=6.67, P=0.03; Time x Deprivation: $F_{2,20}$=4.99, P=0.02). This response was largest early in re-exposure (Fig. S4), when incentive learning should be the greatest. There was no BLA glutamate response to reward in the absence of incentive learning either in the familiar sated state (Fig. 1d-e) or in a familiar hungry state (Fig. S5).

BLA glutamate release was also found to track reward value retrieval. In the subsequent lever-pressing test, BLA glutamate transients preceded the initiation of bouts (Table S1) of reward-seeking presses, but only if rats had prior experience with the reward in the hungry state and

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

could, therefore, retrieve the current value of the anticipated reward to guide their reward pursuit actions (Figs. 1f-g, S7; Time: $F_{2,20}$=1.87, P=0.18; Deprivation: $F_{1,10}$=3.90, P=0.08; Time x Deprivation: $F_{2,20}$=4.31, P=0.03). BLA glutamate transients selectively preceded the *initiation* of reward-seeking activity and did not occur prior to subsequent lever presses within a bout (Fig. S3d, S6), suggesting these signals might relate to the considerations driving reward pursuit. This was further supported by evidence that, across groups, the magnitude of pre-bout-initiation BLA glutamate release positively correlated with the number of seeking presses in and duration of the subsequent bout (presses: r_{88}=0.23, p=0.03, duration: r_{88}=0.21, p=0.05); longer bouts of reward seeking were preceded by larger amplitude glutamate transients. In the group that received incentive learning, glutamate release magnitude significantly predicted future reward-seeking activity in the seconds prior to, but not following the initiation of reward seeking (Fig. 1h).

We next assessed whether BLA glutamate activity is necessary for the encoding and/or retrieval of reward value by blocking BLA glutamate receptors during either reward re-exposure (encoding) or the post-re-exposure lever-pressing test (retrieval) (Fig. 2). Following training in the sated state, all rats were provided the incentive learning opportunity (reward re-exposure hungry; Fig. 2a). During this re-exposure, inactivation of neither NMDA, with infenprodil [7,23], nor AMPA, with NBQX [24], receptors in the BLA receptors altered reward-checking behavior (Fig. 2c; $F_{2,23}$=0.81, P=0.46) or reward palatability responses (Fig. 2d; $F_{2,21}$=0.12, P=0.88). Inactivation of BLA NMDA, but not AMPA receptors did, however, prevent the subsequent upshift in reward seeking that would have otherwise occurred when animals were tested in the hungry state drug-free the next day (Figs. 2e, S7; $F_{2,23}$=4.48, P=0.03), indicating BLA NMDA receptors are necessary for assigning positive value to a reward. All rats were then given the incentive learning opportunity drug-free, and were tested again for their lever pressing in the hungry state on drug (Fig. 2f). In this case, either BLA AMPA or NMDA receptor inactivation prevented the increase in value-guided reward seeking that should have occurred following incentive learning (Fig. 2g, S8; $F_{2,19}$=7.22,

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

P=0.005). Therefore, BLA glutamate signaling tracks and is necessary for both reward value encoding and value-guided reward pursuit.

The results from the glutamate recordings suggested that an excitatory input to the BLA might facilitate its function in reward value encoding and retrieval. The orbitofrontal cortex (OFC) is a prime candidate for this because it sends dense glutamatergic innervation to the BLA [14,15] and is itself implicated in reward processing and decision making [25-34]. Therefore, we next used a chemogenetic approach and the same behavioral task to ask whether OFC→BLA projections are necessary for reward value encoding and/or retrieval (Fig. 3). Recent evidence indicates the lateral (lOFC) and medial (mOFC) OFC subdivisions are anatomically distinct [35]. Data from human and non-human primates suggests they are also functionally distinct [36-38], though whether this is true in rodents is unknown [35]. We identified projections to the BLA from both the mOFC and lOFC (Fig. S9). Nothing is known of the unique or similar function of these projections. Therefore, we assessed function of both lOFC→BLA and mOFC→BLA projections in reward value encoding and retrieval.

Rats expressing the inhibitory designer receptor *human M4 muscarinic receptor* (hM4Di) in excitatory cells of either the lOFC or mOFC showed robust expression in terminals in the BLA in the vicinity of implanted guide cannula (Fig. 3b-c). Clozapine N-oxide (CNO) was infused into the BLA to inactivate these terminals (Fig. S10) [39] during the reward re-exposure, incentive learning opportunity and lever-pressing activity was assessed the following day drug-free (Fig. 3a). Neither manipulation altered reward-checking behavior (Fig. 3d; $F_{2,26}$=0.54, *P*=0.59) or reward palatability responses (Fig. 3e; $F_{2,26}$=1.33, *P*=0.28) online during the re-exposure. Inhibition of lOFC, but not mOFC terminals in the BLA did, however, prevent the subsequent upshift in reward seeking that would have otherwise occurred (Figs. 3f, S11; $F_{2,26}$=5.06, *P*=0.014). These data suggest that activity in lOFC→BLA, but not mOFC→BLA projections is necessary for assigning positive value to a reward.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To determine whether OFC→BLA projections are necessary for reward value retrieval, we allowed all rats to encode the high reward value in the hungry state drug-free and then evaluated their lever-pressing activity in the hungry state following intra-BLA vehicle or CNO infusion (Fig. 3g). In this case, inhibition of mOFC, but not lOFC terminals in the BLA attenuated reward-seeking activity (Figs. 3h, S12; $F_{2,25}$=9.81, P=0.0007), without altering performance of other indices of motivated behavior (Fig. S12). Inactivation of mOFC→BLA projections was without effect if reward value was not being retrieved from memory either because it had not been learned or because it was observable to the subject and could, therefore, be held in working memory at test (Fig. S13). These data indicate the necessity of activity in mOFC→BLA, but not lOFC→BLA projections in retrieving the value of an anticipated reward.

That lOFC→BLA projections were necessary for positive reward value encoding, suggests that activity in these projections might drive such encoding. To test this, we optically stimulated lOFC terminals in the BLA (Fig. S10) concurrent with reward experience under conditions in which incentive learning would not normally occur: the familiar sated state (Fig. 4a). We restricted optical stimulation (473 nm, 20Hz, 10mW, 5 s) to the reward consumption periods during non-contingent reward re-exposure to match the timing of glutamate release detected during incentive learning (Fig. 1d). Rats expressing the excitatory opsin *channelrhodopsin-2* (ChR2) in excitatory cells of the lOFC showed robust expression in terminals in the BLA in the vicinity of implanted optical fibers (Fig. 4b-c). Stimulation of lOFC terminals in the BLA concurrent with reward consumption in the familiar sated state did not significantly alter reward-checking behavior (Fig. 4d; t_{16}=0.20, p=0.84) or reward palatability responses (Fig. 4e; t_{16}=0.25, p=0.80) online. It did, however, cause a dramatic increase in reward-seeking presses in the test conducted in that same sated state manipulation-free the following day (Figs. 4f, S14; $F_{2,24}$=9.25, P=0.001), mimicking the effect of hunger-induced incentive learning (Fig. S15). This did not occur under otherwise identical circumstances with stimulation paired with a task-irrelevant rewarding event (food pellet), ruling

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

out the confounding possibility of enhanced context salience or other factors unrelated to motivation to obtain the specific anticipated sucrose reward (Fig. 4f). lOFC→BLA stimulation also amplified normal, hunger-induced incentive learning (Fig. S15). lOFC→BLA projection activity is, therefore, both necessary for and sufficient to drive the assignment of positive value to a reward that is later retrieved to guide pursuit of that specific reward.

That mOFC→BLA projections were necessary for reward value retrieval suggests activity in these projections might facilitate the retrieval of anticipated reward value information. If this is true, then optically stimulating mOFC→BLA projections during a lever-pressing test should enhance reward seeking following an incentive learning opportunity that would not itself support an upshift in reward pursuit. To test this, we expressed ChR2 in the mOFC and, following reward re-exposure in a moderate, 8hr food-deprived hunger state, optically stimulated mOFC terminals in the BLA during a lever-pressing test in that state (Fig. 5a-c). In controls, sucrose exposure 8hr food-deprived was not sufficient to drive an increase in reward pursuit when tested the following day in this same hunger state, confirming subthreshold incentive learning (Fig. 5d). Stimulation of mOFC terminals in the BLA (473 nm, 20Hz, 10mW, 3 s, once/min) promoted reward-seeking activity under these conditions (Figs. 5d, S16; t_{15}=3.62, p=0.003). Stimulation did not increase reward seeking when tested in the well-learned low-value state, or following effective incentive learning in the high-value hungry state (Fig. S17). mOFC→BLA stimulation was also without effect under otherwise identical circumstances in the absence of the subthreshold incentive learning opportunity (Fig. S18), isolating its effect to reward value retrieval. Together, these data demonstrate that activity in mOFC→BLA projections is both necessary for and sufficient to enhance reward value retrieval.

These data provide evidence for the BLA as a crucial locus for not only learning about the value of primary rewarding events, but also for retrieving this information to guide adaptive reward

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

pursuit, revealing it as a critical contributor to value-based decision making. These value encoding and retrieval functions are supported via doubly dissociable contributions of excitatory input from the lOFC and mOFC. Whereas activity in lOFC→BLA projections during reward experience is necessary for and sufficient to drive the encoding of a positive shift in a reward's value, activity in mOFC→BLA projections is necessary and sufficient for retrieving this value from memory to guide reward pursuit decisions.

These data accord well with previous evidence of BLA necessity for reward value updating [5,7,8,10,12], but differ from data suggesting that the BLA is not required for retrieving reward value to guide reward-seeking activity [7,10,11]. In these latter experiments, the value shift was negative, temporary, and occurred immediately prior to test. Our value learning was positive, permanent, and occurred at least 24 hr prior to test. We suggest, therefore, that the BLA facilitates the encoding and retrieval of long-term, need-state-dependent reward value memories, and, as such, is a critical contributor to value-based decision making. This interpretation is consistent with evidence from humans and non-human primates that BLA neuronal activity can encode value [40], prospectively reflect goal plans [41], and predict behavioral choices [42], and with evidence of temporally-specific BLA inactivation disrupting choice behavior [43].

The demonstrated dissociable function of lOFC→BLA and mOFC→BLA projections in encoding and retrieving, respectively, reward value is consistent with evidence of broad OFC encoding of reward value [44-51]. These results also translate recent evidence from primates of similar dissociable function of these OFC subregions in credit assignment and value-guided decision making [36-38] to rodents and, using bi-directional, projection-specific manipulations, suggest these functions are achieved, at least in part, via projections to the BLA. This is consistent with evidence of cooperative OFC and amygdala function in reward learning and choice [52,53]. Both the lOFC and mOFC have been proposed to be involved in representing and using information about the current and anticipated states, or situations, to guide adaptive behavior when the information defining those states is 'hidden', i.e., not externally observable [27-29,54,55]. Adaptive

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

behavior in our task relies on such representation. Although there has been no perceptual change (i.e., same context, levers, etc.), following incentive learning, the state is nonetheless different: the anticipated reward is now more valuable. The critical elements defining this state— internal need and the reward itself— are not externally perceptible. Our data, therefore, indicate that lOFC→BLA and mOFC→BLA projections mediate the encoding and retrieval, respectively, of the state-dependent value of a specific anticipated reward.

This function of lOFC→BLA projections is in line with evidence of both reward identity coding [56] and the sensitivity of these responses to reward value shifts [57] in human lOFC. Activation of BLA NMDA receptors is known to mediate BLA synaptic plasticity [58-60] and to be necessary for establishing long-term, BLA-dependent memories [61,62]. That activity at both glutamatergic lOFC terminals and NMDA receptors in the BLA is necessary for reward value encoding, suggests, therefore, that lOFC→BLA projections might direct the encoding of the reward value in the BLA.

mOFC→BLA projections were found to mediate the retrieval of this value memory to ensure reward pursuit that is adaptive in the current state. This is consistent with evidence that the mOFC itself mediates aspects of reward-related decision making [63,64] and effort allocation according to anticipated reward value [65]. Stimulation of mOFC→BLA manipulations only augmented reward pursuit if two conditions were met: 1) a state-dependent reward value had been encoded and 2) the internal state was not sufficiently discriminable to increase reward seeking following incentive learning on its own. mOFC→BLA inactivation was without effect when rewards were present at test. In accordance with mOFC function in representing hidden, but not observable states [27], we speculate these data may indicate a function of mOFC→BLA activity in discriminating reward values between different hidden states, in this case, internal need states.

OFC-BLA circuitry is known to become dysfunctional in patients diagnosed with addiction [66,67], anxiety [68,69], depression [70] and schizophrenia [71]. These conditions are also marked by maladaptive motivation and poor decision making. The current data, therefore, provide some mechanistic insight into how cortical-amygdala dysfunction might contribute to these and other psychiatric

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

diseases characterized by maladaptive reward valuation and poor reward-related decision making.

MATERIALS AND METHODS

Subjects.

Male, Long Evans rats (aged 8-10 weeks at the start of the experiment; Charles River Laboratories, Wilmington, MA) were group housed and handled for 3-5 days prior to the onset of the experiment. Unless otherwise noted, separate groups of naïve rats were used for each experiment. Rats were provided with water provided *ad libitum* in the home cage and were maintained on food-restriction for a certain amount of time each day, as described below. Experiments were performed during the dark phase of the 12:12 hr reverse dark/light cycle. All procedures were conducted in accordance with the NIH Guide for the Care and Use of Laboratory Animals and were approved by the UCLA Institutional Animal Care and Use Committee.

Surgery.

Standard surgical procedures described previously [19] were used for all surgeries. Rats were anesthetized with isoflurane (4–5% induction, 1–2% maintenance) and a nonsteroidal anti-inflammatory agent was administered pre- and post-operatively to minimize pain and discomfort. Following surgery rats were individually housed.

Electroenzymatic glutamate recordings. Following training to stable performance, rats were implanted with a unilateral, pre-calibrated glutamate biosensor into the BLA (AP -3.0 mm, ML + 5.1, DV -8.0) and a Ag/AgCl reference electrode into the contralateral cortex. Biosensor placements were verified using standard histological procedures (Fig. 1B).

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

BLA glutamate receptor inactivation. Following training to stable performance, rats were implanted with guide cannula (22-gauge stainless steel; Plastics One, Roanoke, VA) targeted bilaterally 1 mm above the BLA (AP -3.0 mm, ML ± 5.1, DV -7.0). Cannula placements were verified using standard histological procedures (Fig. 1B) and subjects were removed from the study if placements were off target (*N*=1).

Chemogenetic manipulation of OFC→BLA projections. Prior to onset of behavioral training, rats were randomly assigned to a viral group, anesthetized using isoflurane and infused bilaterally with adeno-associated virus (AAV) expressing the inhibitory designer receptor *human M4 muscarinic receptor* (hM4D(Gi); AAV8-CaMKIIa-HA-hM4D(Gi)-IRES-mCitrine). Virus (0.30 μl) was infused at a rate of 6 μl/hr via an infusion needle positioned in the ventrolateral orbitofrontal cortex (lOFC; AP: +3.2 mm; ML: ± 2.4; DV: -5.4) or medial OFC (mOFC; AP: +4.0; ML: ± 0.5; DV: -5.2). Bilateral guide cannula (22-gauge stainless steel; Plastics One) were implanted 1 mm above the BLA (AP -3.0 mm, ML ± 5.1, DV -7.0). Testing commenced 8 weeks post-surgery to ensure axonal transport and expression in lOFC or mOFC terminals in the BLA. Restriction of expression to the lOFC or mOFC was verified with immunofluorescence using an antibody to recognize the HA tag. Cannula placements in the terminal expression region were verified using standard histological procedures. Subjects were removed from the study due to lack of expression or if cannula were misplaced outside the BLA (lOFC, *N*=0, mOFC, *N*=2).

Optogenetic manipulation of OFC→BLA projections. Prior to onset of behavioral training, rats were randomly assigned to viral group, anesthetized using isoflurane and infused bilaterally with AAV expressing excitatory opsin channelrhodopsin-2 (ChR2; AAV5-CaMKIIa-hChR2(H134R)-eYFP) or the empty vector (EV) control (AAV8-CaMKIIa-eYFP). Virus (0.30 μl) was infused at a rate of 6 μl/hr via an infusion needle positioned in the lOFC or mOFC. Bilateral optical fibers (200

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

μm core, numerical aperture 0.66; Prizmatix, Southfield, MI) held in ferrules (Kientec Systems Inc., Stuart, FL) were implanted 0.3 mm above the BLA (AP -3.0 mm, ML ± 5.1, DV -7.7). Testing commenced 8 weeks post-surgery to ensure axonal transport and expression in lOFC or mOFC terminals in the BLA. Restriction of virus to either the lOFC or mOFC was verified with immunofluorescence using an antibody against eYFP and optical fiber placements in vicinity of terminal expression were verified using standard histological procedures. Subjects were removed from the study due to lack of expression or if optical fibers were misplaced outside the BLA (lOFC, *N*=1, mOFC, *N*=1).

Validation of chemogenetic and optogenetic manipulation of OFC→BLA projections. In a separate group of rats, hM4d(Gi) and ChR2 were co-expressed by infusing both AAV8-CaMKIIa-HA-hM4D(Gi)-IRES-mCitrine and AAV5-CaMKIIa-hChR2(H134R)-eYFP bilaterally in the lOFC (AP: +3.2 mm; ML: ± 2.4; DV: -5.4) or mOFC (AP: +4.0; ML: ± 0.5; DV: -5.2). 8-weeks post viral infusion, rats were anesthetized and a pre-calibrated microelectrode array (MEA) glutamate biosensor affixed to an optical fiber and guide cannula was acutely implanted into the BLA (AP -3.0 mm, ML + 5.1, DV -8.0), and a Ag/AgCl reference electrode placed in the contralateral cortex. The optical fiber was affixed behind the MEA (to reduce photovoltaic artifacts) and the optical fiber tip terminated 0.3 mm above the glutamate sensing electrodes. The guide cannula (Plastics One) terminated 6.5 mm above the MEA tip to avoid tissue damage and was positioned such that, when inserted, the injector (Plastics One) would protrude 6.2 mm and end within 100 μm from the microelectrodes. The injector was inserted after the biosensor/optical fiber probe was lowered into the BLA to further minimize tissue damage. Level of anesthesia was kept constant throughout recordings by maintaining breaths per minute (bpm) constant (1 bpm) by adjusting isoflurane level (1–1.5%). Viral expression was verified using immunofluorescence and biosensor placements were verified using standard histological procedures (Fig. S10).

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Electroenzymatic glutamate biosensors.

Biosensor fabrication. MEA probes were fabricated in the Nanoelectronics Research Facility at UCLA and modified for glutamate detection as described previously [17-19]. Briefly, these biosensors use glutamate oxidase (GluOx) as the biological recognition element and rely on electro-oxidation, via constant-potential amperometry (0.7 V versus a Ag/AgCl reference electrode), of enzymatically-generated hydrogen peroxide reporter molecule to provide a current signal. This current output is recorded and converted to glutamate concentration using a calibration factor determined *in vitro*. Enzyme immobilization was accomplished by chemical crosslinking using a solution consisting of GluOx, bovine serum albumin (BSA), and glutaraldehyde. Interference from both electroactive anions and cations is effectively excluded from the amperometric recordings, while still maintaining a subsecond response time, by electropolymerization of polypyrrole (PPY) or poly o-phenylenediamine (PPD), as well as dip-coat application of Nafion to the electrode sites prior to enzyme immobilization [17-19]. Each MEA had two non-enzyme-coated sentinel electrodes for the removal of correlated noise from the glutamate sensing electrodes by signal subtraction, as described previously [18,19]. These electrodes were prepared identically with the exception that the BSA/glutaraldehyde solution did not contain GluOx. The average *in vivo* limit of glutamate detection of the sensors used in this study was 0.36 μM (sem=0.03, range 0.13-0.67 μM).

Reagents. Nafion (5 wt.% solution in lower aliphatic alcohols/H_2O mix), bovine serum albumin (BSA, min 96%), glutaraldehyde (25% in water), pyrrole (98%), p-Phenylenediammine (98%), L-glutamic acid, L-ascorbic acid, 3-hydroxytyramine (dopamine) were purchased from Aldrich Chemical Co. (Milwaukee, WI, USA). L-Glutamate oxidase (GluOx) from *Streptomyces* Sp. X119-6, with a rated activity of 24.9 units per mg protein (U mg^{-1}, Lowry's method), produced by Yamasa Corporation (Chiba, Japan), was purchased from US Biological (Massachusetts MA). Phosphate buffered saline (PBS) was composed of 50 mM Na_2HPO_4 with 100 mM NaCl (pH 7.4). Ultrapure

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

water generated using a Millipore Milli-Q Water System (resistivity = 18 MΩ cm) was used for preparation of all solutions used in this work.

Instrumentation. Electrochemical preparation of the sensors was performed using a Versatile Multichannel Potentiostat (model VMP3) equipped with the 'p' low current option and low current N' stat box (Bio-Logic USA, LLC, Knoxville, TN). *In vitro* and *in vivo* measurements were conducted using a low-noise multichannel Fast-16 mkIII potentiostat (Quanteon LLC, Nicholasville, KY), with reference electrodes consisting of a glass-enclosed Ag/AgCl wire in 3 M NaCl solution (Bioanalytical Systems, Inc., West Lafayette, IN) or a 200 μm diameter Ag/AgCl wire, respectively. All potentials are reported versus the Ag/AgCl reference electrode. Oxidative current was recorded at 80 kHz and averaged over 0.25-s intervals.

In Vitro Biosensor Characterization. All biosensors were calibrated *in vitro* to test for sensitivity and selectivity of glutamate measurement prior to implantation. A constant potential of 0.7 V was applied to the working electrodes against a Ag/AgCl reference electrode in 40 mL of stirred PBS at pH 7.4 and 37ºC within a Faraday cage. After the current detected at the electrodes equilibrated (~30-45 min), aliquots of glutamate were added to the beaker to reach final glutamate concentrations in the range 5 – 60 μM. A calibration factor based on these response was calculated for each GluOx-coated electrode. The average calibration factor for the sensors used in these studies was 135.98 μM/nA. Control electrodes, coated with PPy or PPD, Nafion, and BSA/glutaraldehyde, but not GluOx, showed no detectable response to glutamate. Aliquots of ascorbic acid (250 μM final concentration) and dopamine (5-10 μM final concentration) were added to the beaker as representative examples of readily oxidizable potential anionic and cationic interferent neurochemicals, respectively, to confirm selectivity for glutamate (Fig. S1). For the sensors used in these studies no current changes above the level of the noise were detected

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

to the addition of cationic or anionic interferents, as reported previously [17-19]. To assess uniformity of H_2O_2 sensitivity across control and GluOx-coated electrodes, aliquots of H_2O_2 (10 μM) were also added to the beaker. There was less than a 10%, statistically insignificant (t_{42}=0.32, p=0.75) difference in the H_2O_2 sensitivity on control electrode sites relative to enzyme-coated sites, indicating that any changes detected *in vivo* on the enzyme-coated biosensor sites following control channel signal subtraction could not be attributed to endogenous H_2O_2.

In vivo validation of chemogenetic and optogenetic manipulation of OFC→BLA projections. Glutamate biosensors were used to validate optogenetic stimulation and chemogenetic inhibition, respectively, of OFC terminals in the BLA. Animals expressing ChR2 and hM4di in either the lOFC or mOFC were anesthetized and implanted with a pre-calibrated MEA-fiber-cannula probe into the BLA, as described above. Experiments were conducted inside a Faraday cage. Following sensor implant, an injector was inserted into the cannula. A constant potential of 0.7 V was applied to the working electrodes against the Ag/AgCl reference electrode implanted in the contralateral hemisphere. The detected current was allowed to equilibrate (~30-45 min). Baseline spontaneous glutamate release events (i.e., glutamate transients) were measured for 2 min prior to infusion of vehicle. Spontaneous transients were then monitored for 15 min post-infusion. Following this, glutamate release was optically evoked by delivery of blue light pulses (473 nm, 5-20 mW, 20Hz Hz, 5 s or 3 s) to stimulate lOFC or mOFC terminals in the BLA. Each stimulation parameter was repeated 3x, with at least 60 s in between stimulations. Rats then received an infusion of CNO (1 mM, 0.5 μl) into the extracellular space surrounding the MEA. Spontaneous glutamate transients were monitored 2 min before (baseline) and 15 min following CNO infusion. The light delivery protocol was then repeated to assess CNO:hM4D(Gi) attenuation of optically-evoked glutamate release from OFC terminals in the BLA. As an iterative control, in a subset of subjects, the applied potential was lowered to 0. 2 V, below the H_2O_2 oxidizing potential, and recordings of spontaneous and optically-evoked glutamate release were made following CNO infusion.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Optical stimulation.

Light was delivered to the OFC terminals in the BLA using a laser (Dragon Lasers, ChangChun, JiLin, China) connected through a ceramic mating sleeve (Thorlabs, Newton, NJ) to the ferrule implanted on the rat. We used a 473 nm laser to activate ChR2-transfected projection neurons, or a 589 nm laser (which is largely outside the range of the ChR2 sensitivity range [72]) as a control for the effects of construct expression and light delivery in ChR2-transfected projection neurons. For optical stimulation, light pulses (25 msec pulse) were delivered at 20 Hz. This was based on previous studies showing reward-induced firing rates of OFC neurons that range from 6-40 spikes/second [73-76]. We also found this stimulation frequency to effectively stimulate glutamate release from OFC terminals in the BLA *in vivo* (Fig. S10).

Drug Administration.

Ifenprodil (Tocris Bioscience, Bristol, UK) and NBQX (2,3-Dioxo-6-nitro-1,2,3,4-tetrahydrobenzo[*f*]quinoxaline-7-sulfonamide disodium salt; Tocris Bioscience, Bristol, UK) were dissolved in sterile saline vehicle. CNO (Tocris Bioscience, Bristol, UK) was dissolved in aCSF to 1mM. Drugs were infused bilaterally into the BLA in a volume of 0.5 µl over 1 min via injectors inserted into the guide cannula fabricated to protrude 1 mm ventral to the cannula tip using a microinfusion pump. Injectors were left in place for at least 1 additional min to ensure full infusion. Rats were placed in the operant chamber 5 min after infusion to allow sufficient time for the drug to become effective. The dose of ifenprodil (1.67 µg/side), an N-methyl-D-aspartate (NMDA) receptor antagonist with selective targeting of receptors that contain the NR2B subunit [23], was selected because it has been shown to impair value-based decision making [7]. The alpha-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate (AMPA) receptor antagonist, NBQX, at a dose of 1.0 µg/side, was selected based on our previous evidence of its effectiveness in reward-related tasks [19,77]. CNO dose was selected based on our previous demonstration of the efficacy and duration

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of action of this dose and our evidence showing effective inhibition of glutamate release from OFC terminals in the BLA with this dose (Fig. S10) [39].

Behavioral Procedures.

General training and testing.

Apparatus. Training took place in Med Associates operant chambers (East Fairfield, VT) housed within sound- and light-attenuating boxes, described previously [19]. For *in vivo* glutamate measurements all testing was conducted in a single Med Associates operant chamber housed within a continuously-connected, copper mesh-lined sound attenuating chamber and outfitted with an electrical swivel (Crist Instrument Co, Hagerstown, MD) connecting a headstage tether that extended within the operant chamber to the potentiostat recording unit (Fast-16 mkIII, Quanteon, LLC) positioned outside the operant chamber. For optogenetic experiments, testing was conducted in Med associates operant chambers outfitted with an Intensity Division Fiberoptic Rotary Joint (Doric Lenses, Quebec, QC, Canada) connecting the output fiberoptic patchcords to a laser (Dragon Lasers, ChangChun, JiLin, China) positioned outside the operant chamber.

All chambers contained 2 retractable levers that could be inserted to the left and right of a recessed food-delivery port in the front wall. A photobeam entry detector was positioned at the entry to the food port to provide a goal approach measure. The chambers were equipped with syringe pump to deliver 20% sucrose solution in 0.1ml increments through a stainless steel tube, or a pellet dispenser that delivered a single 45-mg pellet (Bio-Serv, Frenchtown, NJ), into a custom-designed electrically-isolated Acetal plastic well in the food port. A lickometer circuit (Med Associates), connecting the grid floor of the boxes and stainless steel sucrose-delivery tubes, with the circuit closed by the rats' tongue allowed recording of the lick frequency when rats consumed each sucrose delivery. A 3-watt, 24-volt house light mounted on the top of the back wall opposite the food-delivery port provided illumination.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Training. Each experiment followed the same general structure. Rats were trained on a self-paced, 2-lever, action sequence to earn a delivery of 0.1ml 20% sucrose reward. Training procedures were similar to those we have described previously [5,6,16]. Except where noted, rats were deprived of food for 4 hr prior to each training session. Each session began with the illumination of the houselight and insertion of the lever, where appropriate, and ended with the retraction of the lever and turning off of the houselight. Rats were given only one training session/day. Rats received 3 d of magazine training in which they were exposed to non-contingent sucrose or water deliveries (30 outcomes over 35 min) in the operant chamber with the levers retracted, to learn where to receive rewards. This was followed by daily instrumental training sessions in which sucrose rewards could be earned by lever pressing. Rats were first given 3 d of single-action, taking lever, instrumental training on the lever to the right (i.e., 'taking' lever) of the food-delivery port with the sucrose delivered on a continuous reinforcement schedule. Each session lasted until 20 outcomes had been earned or 30 min elapsed. Following single-action instrumental training, the 'seeking' lever (i.e., the lever to the left of the food-delivery port) was introduced into the chamber. Rats were allowed to press on the seeking lever to gain access to the taking lever, a single press on which delivered the sucrose solution and retracted this lever. The seeking lever remained present during the entire session. Rats were trained on this self-paced, 2-lever, action sequence for a total of 12-18 days: 3 days in which a press on the 'seeking' lever was continuously reinforced with the taking lever, 2-4 days in which the seeking lever was reinforced on a random ratio 2 (RR-2) schedule, 3-5 days in which the seeking lever was reinforced on a RR-5 schedule, and 4-6 days in which the seeking lever was reinforced on the final RR-10 schedule until stable responding was established. The taking lever was always continuously reinforced. Each session lasted until 20 outcomes had been earned or 40 min elapsed.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Incentive learning opportunity and test. Following training to stable response rates, rats received non-contingent re-exposure to the sucrose outcome (30 exposures/35 min) in the operant box with the levers retracted. Unless otherwise noted, food-port entries and lickometer palatability measures [78-81] were collected during this phase of the experiment. These non-contingent sucrose deliveries provided an incentive learning opportunity wherein the value of the sucrose reward may be updated (see specific experimental procedures). The next day lever-press behavior was measured during a brief, 5-min, non-reinforced probe test to assess the effects of the previous day's incentive learning opportunity on reward-seeking actions.

Online, near-real time glutamate detection during reward exposure or reward seeking. Following training on the self-paced action sequence in the sated (4-hr food-deprived) state and surgery (see Fig. 1a), testing commenced. Prior to each test, rats were placed in the recording operant chamber and the biosensor was tethered to the potentiostat via the electrical swivel for application of the 0.7 V potential. The recorded amperometric signal was allowed to stabilize prior to session onset (~30-45 min). First, rats received a single day of instrumental re-training similar to the training described above, but with the ratio requirement progressively increasing from a fixed-ratio-1 to RR-10 after each 5th outcome earned to re-establish lever pressing post-surgery. The next day, rats were non-contingently exposed to the sucrose in the familiar sated state (4 hr food-deprived) or in a hungry (20 hr food-deprived) state. For hunger state group assignment, subjects were counterbalanced based on average lever-press rate during the last 2 instrumental training sessions. The next day, all rats were tested hungry. A separate group of rats were maintained hungry throughout training and test (Fig. S5). To prevent electrical interference, lickometers were not connected during recording sessions.

BLA AMPA and NMDA glutamate receptor inactivation during reward re-exposure or post-re-exposure lever-pressing test. Following training in the sated state as described above, drug

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

groups were counterbalanced based on lever-press rate during the two final instrumental training sessions. On 2 of the instrumental training days immediately prior to the first incentive learning opportunity rats were given mock infusions to habituate them to the infusion procedures; injectors were inserted into the cannula, but no fluid was infused. All rats then received the non-contingent re-exposure to the sucrose in the 20 hr food-deprived hungry state. Prior to this incentive learning opportunity, rats received intra-BLA infusions of vehicle, Ifenprodil, or NBQX. The next day all rats received a drug-free, non-reinforced, lever-pressing probe test in the hungry state (see Fig. 2a). Following 2 days to reestablish satiety, rats received two sessions of retraining (1/day) on the action sequence in the 4-hr food-deprived state. They were then given another round of re-exposure and a lever-pressing test. In this case, non-contingent exposure to the sucrose in the hungry state was conducted drug-free. To ensure value encoding and to equate the number of incentive learning opportunities with intact glutamate receptor activity, rats previously assigned to the vehicle group received 2 drug-free re-exposure sessions, while the rats previously assigned to Ifenprodil or NBQX groups received 3 drug-free re-exposure sessions. The day following the last day of re-exposure, all rats received a non-reinforced, lever-pressing probe test in the hungry state. Prior to this test, rats received an infusion of vehicle, Ifenprodil, or NBQX (see Fig. 2f). Drug group assignment for this test was counterbalanced with respect to previous drug treatment.

Chemogenetic inactivation of lOFC→BLA or mOFC→BLA projections during reward re-exposure or post-re-exposure lever-pressing test. Training and test was identical to that for the BLA glutamate receptor inactivation experiments, except that rats expressing hM4D(Gi) in the lOFC or mOFC received infusion of either vehicle or CNO prior to the first non-contingent re-exposure session in the hungry state (see Fig. 3a) or prior to the second post-re-exposure, non-reinforced, lever-pressing probe test in the hungry state (see Fig. 3g). There were no significant differences in reward-seeking lever presses between vehicle-treated subjects expressing hM4D(Gi) in the

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

lOFC or mOFC during either the first (t_{11}=2.00, p=0.07) or second test (t_9=0.20, p=0.85), and therefore, these groups were collapsed to serve as a single control group.

To evaluate the effect of mOFC→BLA projection inactivation on reward seeking in the absence of reward value retrieval, a separate group of rats expressing hM4D(Gi) in the mOFC was trained sated and received intra-BLA infusions of vehicle or CNO prior to a non-reinforced, lever-pressing probe test in the hungry state as above, but without prior non-contingent re-exposure to the sucrose in the hungry state (i.e., without a reward value encoding opportunity; Fig. S13). Each rat was given 2 non-reinforced probe tests, one each following vehicle or CNO infusion for a within-subject drug comparison (test order counterbalanced). Two days after the last non-reinforced robe test, rats were retrained sated for two days, given a drug-free incentive learning opportunity in the hungry state, and then received intra-BLA infusions of Vehicle or CNO prior to a reinforced lever-pressing test (Fig. S13). In this test, the presence of the reward made retrieval of the reward's value from memory unnecessary. Each rat was given 2 reinforced tests, one each following vehicle or CNO infusion to allow a within-subject drug comparison (test order counterbalanced).

Optogenetic activation of lOFC→BLA projections during reward re-exposure. Rats expressing ChR2, or the empty vector (EV) control in the lOFC with optical fibers above the BLA were trained sated as described above (Fig. 4a). On the last two days of instrumental training, rats were tethered to the patchcord, but no light was delivered to allow habituation to the optical tether. At test, rats were maintained in the familiar 4-hr food-deprived sated state and received non-contingent re-exposure to the sucrose or to a task-irrelevant food-pellet reward. During this non-contingent reward exposure blue light (473 nm, 20Hz, 10mW, 5 s) was delivered for optical activation of lOFC terminals, in the BLA in ChR2-expressing subjects, during consumption of the reward. The laser was triggered by the first lick following sucrose delivery or the first food-port entry following pellet delivery. Optical stimulation timing was based on evidence that BLA

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

glutamate release occurred in response to reward consumption during incentive learning and peaked on average 2.79 s (s.e.m.=0.67; range = 0.63-6.1 s) post reward (Fig. 1d) and evidence that rats finish reward consumption and exited the food delivery port ~5-10 s following reward collection. A subset of rats expressing ChR2 in the lOFC received 589 nm light delivery (outside the range of ChR2 sensitivity [72]) in the BLA. The next day, all rats received a non-reinforced probe test in the familiar sated state while tethered, but without light delivery. This sequence of re-exposure and testing was repeated twice, first in a novel, moderate hunger state (8-hr food-deprived) and then in a novel hungry (20-hr food-deprived) state (Fig. S15). Rats were given 2 days off and retrained in the 4-hr food-deprived state for two days in between each test set. In no case did reward-seeking lever-press activity significantly differ between ChR2-expressing rats that received 589 nm optical activation and EV controls receiving 473 nm optical activation (t_6=0.10-0.95, p=0.38-0.93) and, thus, these controls groups were collapsed to serve as a single control group for each test.

Optogenetic activation of mOFC→BLA projections during lever-pressing test. Rats expressing ChR2, or EV in the mOFC with optical fibers above the BLA received training, non-contingent sucrose exposure, and testing as the described for optogenetic activation of lOFC, except light (473 nm, 20Hz, 10mW, 3 s) was delivered during each of the non-reinforced lever-pressing tests to, in ChR2-expressing subjects, activate mOFC terminals in the BLA. Light was delivered 1/minute, for a total of 10 light deliveries throughout the 10-minute test. The first light delivery occurred 30 s after test onset. The duration of optical stimulation was based on the finding that glutamate release preceded the initiation of reward seeking and the rise time to peak glutamate release prior to reward-seeking bouts was on average 1.95 s (s.e.m.=0.43; range = 0.40-3.0 s; Fig. 1f). As above, a subset of ChR2-expressing subjects received 589 nm light delivery. Tests were conducted 4-, 8-, and 20-hr food-deprived, as above, with each pressing test preceded by non-contingent sucrose reward re-exposure in the absence of light delivery. The moderate 8-hr

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

food-deprived state provided a subthreshold incentive learning opportunity that was, on its own, not sufficiently discriminable to induce an upshift in reward seeking. Reward-seeking presses did not significantly differ between ChR2-expressing rats that received 589 nm light delivery and EV controls receiving 473 nm light delivery (t_6=0.30-2.44, p=0.051-0.77) and, thus, these groups were collapsed to serve as a single control group for each test.

To examine the effect of mOFC→BLA projection activation on reward seeking in the moderate food-deprivation state, but in the absence of incentive learning, a separate group of rats expressing ChR2 in the mOFC was trained while sated, and received light delivery during a non-reinforced probe test in the moderate 8-hr food-deprived state as above, but without prior re-exposure to the sucrose in the 8-hr state (i.e., without the subthreshold incentive learning opportunity; Fig. S18). Each rat was given 2 non-reinforced probe tests, one each with either 473 nm (for ChR2 activation) or 589 nm (control wavelength) light delivery, to allow within-subject comparison (test order counterbalanced).

Histology.

Rats were transcardially perfused at the conclusion of behavioral testing with PBS followed by 10% formalin. The brains were removed, post-fixed in formalin, then cryoprotected, cut with a cryostat at a thickness of 30 µm, and collected in PBS. eYFP fluorescence was used to verify ChR2 expression. To verify hM4D(Gi) expression, immunohistochemical analysis was performed as described previously [82-84]. Briefly, floating coronal sections were blocked for 1 hr at room temperature in 8% normal goat serum (NGS, Jackson ImmunoResearch Laboratories) with 0.3% Triton X-100 in PBS and then incubated overnight at 4°C in 2% NGS, 0.3% Triton X-100 in PBS with primary antibody (anti-HA, 1:500, Biolegend, San Diego, CA, cat. no. 901501). The sections were then incubated for 2 hr at room temperature with goat anti-mouse IgG, Alexa 594 conjugate (1:1000, Invitrogen, cat. no. A11005). All sections were washed 3 times for 5 min each in PBS before and after each incubation step and mounted on slides using ProLong Gold antifade reagent

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

with DAPI (Invitrogen). All images were acquired using a Keyence (BZ-X710) microscope with a 4X or 20X objective (CFI Plan Apo), CCD camera, and BZ-X Analyze software. Biosensor and cannula placements in non-AAV subjects, were verified using standard histological procedures.

Data analysis

Behavioral analysis. Seeking and taking lever presses and/or food-port entries collected continuously for each training and test session. Seeking lever presses were normalized to baseline response rate averaged across the last 2 training sessions prior to test to control for pre-test response variability and allow comparison across tests conducted in different deprivation states (see [5,6,85,86]). Raw press rates data are presented in the supplemental materials. Lickometer measurements were made during sucrose consumption during the non-contingent re-exposure sessions.

Chemogenetic and optogenetic manipulation of glutamate release. Analysis details and characterization of glutamate release events have been described previously [18,19]. Electrochemical data were baseline-subtracted. Detected current was averaged across the first 10 s of the 2-min, pre-infusion, baseline period and this baseline was subtracted from current output at each time point. Current changes from baseline on the PPY(or PPD)/Nafion-coated sentinel electrode were then subtracted from current changes on the PPY(or PPD)/Nafion/GluOx glutamate biosensor electrode to remove correlated noise. This signal was then converted to glutamate concentration using an electrode-specific calibration factor obtained *in vitro*. Mini Analysis (Synaptosoft, Decatur, GA) was used to determine the frequency and amplitude of spontaneous glutamate transient release events. A fluctuation in the glutamate trace was deemed a glutamate transient if it was at >2.5x the RMS noise sampled from the pre-test baseline period. To determine transient amplitude, a baseline was taken by averaging 3 sample bins around the first minima located 0.5-5 s before the peak and this baseline was subtracted from the peak

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

amplitude. If one peak followed another within 5 s the baseline was taken after the first peak to distinguish these events. Peaks with a total duration below 0.5 s or with an immediately preceding or following negative deflection greater than half the peak amplitude were considered noise spikes and were omitted from the analysis. To evaluate optically-evoked glutamate release, we isolated the 5-s or 3-s period prior to, during, and following light delivery. The average glutamate concentration change in the 5-s or 3-s optical stimulation period was subtracted from that during an equivalent period immediately prior to optical stimulation. This was averaged across each of the 3 replicates for each parameter. There were no statistically-significant main effects of OFC subregion (mOFC v. lOFC; $F_{1,4}$=2.09, P=0.22; Treatment: $F_{1,4}$=8.78, P=0.04; Brain region x Treatment: $F_{1,4}$=0.01, P=0.91)) and, thus, these data were collapsed.

Temporal relationship between glutamate release and behavior. As above, electrochemical data were baseline-subtracted. Detected current was averaged across the 10 s baseline period 2-min prior to test and this baseline was subtracted from current output at each time point. We evaluated the temporal relationship between glutamate release and behavioral events as described previously [18,19]. For the sucrose reward re-exposure, we isolated glutamate concentration changes in the 5 s prior to and 10 s following the first food-port entry following each reward delivery (i.e., reward collection). This period was chosen to give an adequate pre-reward baseline and based on evidence that rats disengaged from the food port ~5-10 s following reward collection. The average glutamate concentration in the 1-s period 5 s prior to reward collection served as the baseline and this was subtracted from each data point in the peri-reward glutamate concentration v. time trace. To quantify the reward-evoked glutamate concentration change, for each trial the average glutamate concentration change in the 10-s post-reward period was averaged across trials and this was compared to average glutamate concentration change in the 5-s prior to reward collection and to equivalent analysis of glutamate concentration changes in 5-s periods in the absence of reward or reward-checking behavior.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

During the non-reinforced, lever-pressing probe test, because rats tended to organize their reward-seeking lever presses into bouts, we focused on those presses that initiated bouts of reward-seeking activity (i.e., 'initiating presses'), excluding presses that occurred within a pressing bout, as we have described previously [19]. An 'initiating seeking press' was defined as the first press after completion of an action-sequence or, because rats often disengaged from the lever and then reinitiated reward seeking, the first press after ≥6 s pause in pressing. Similar definitions of initiation of reward seeking and instrumental bouts defined by pauses in activity have been described previously [19,87-89]. See Table S1 for seeking bout information. We evaluated glutamate concentration changes in the 5 s prior to and following each initiating reward-seeking press. The average glutamate concentration in the 1-s period, 5 s prior to each initiating press served as the baseline. This analysis window was selected to avoid contaminating events (e.g., termination of a previous bout, food-port entries, etc.). Average glutamate concentration change for each initiating press was quantified in the 3-s period immediately prior to and after each initiating press and this was compared to equivalent analysis of glutamate concentration changes in the absence of lever pressing. Data were averaged across trials. We quantified glutamate concentration around all intra-bout seeking presses similarly (Fig. S6). Pearson correlations were used to assess the relationship between glutamate fluctuations around bout initiation and the number of presses and duration of subsequent bouts.

Palatability analysis. A lickometer circuit (Med Associates), connecting the grid floor of the box and the stainless steel sucrose-delivery tubes, with the circuit closed by the rats' tongue, allowed recording of individual lick events. Lickometer measures were amplified and fed through an interface to a PC programmed to record the time of each lick to the nearest 1 msec. Based on previous reports [5,78,85,90-92], we used licking frequency as a measure of sucrose palatability. This measure of licking microstructure during consumption provides a similar analysis of palatability

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

changes as those assessing taste reactivity following oral infusions [80]. These data were analyzed with custom-written python-based code.

Statistical analysis.

Datasets were analyzed by two-sided, Student's *t* tests, one- or two-way repeated-measures analysis of variance (ANOVA), as appropriate. Bonferroni corrected *post hoc* tests were performed to clarify all main effects and interactions. Two-tailed, paired t-tests were used for *a priori* planned comparisons, as advised by [93] based on a logical extension of Fisher's protected least significant difference (PLSD) procedure for controlling familywise Type I error rates. All datasets met equal covariance assumptions, justifying ANOVA interpretation [94]. Alpha levels were set at $P<0.05$.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

REFERENCES

1 Dickinson, A. & Balleine, B. W. Motivational control over goal-directed action. *Animal Learning and Behavior* **22**, 1-18 (1994).

2 LeDoux, J. E. Emotional memory systems in the brain. *Behav Brain Res* **58**, 69-79 (1993).

3 Janak, P. H. & Tye, K. M. From circuits to behaviour in the amygdala. *Nature* **517**, 284-292, doi:10.1038/nature14188 (2015).

4 Wassum, K. M. & Izquierdo, A. The basolateral amygdala in reward learning and addiction. *Neurosci Biobehav Rev* **57**, 271-283, doi:10.1016/j.neubiorev.2015.08.017 (2015).

5 Wassum, K. M., Ostlund, S. B., Maidment, N. T. & Balleine, B. W. Distinct opioid circuits determine the palatability and the desirability of rewarding events. *Proc Natl Acad Sci U S A* **106**, 12512-12517 (2009).

6 Wassum, K. M., Cely, I. C., Balleine, B. W. & Maidment, N. T. Mu opioid receptor activation in the basolateral amygdala mediates the learning of increases but not decreases in the incentive value of a food reward. *Journal of Neuroscience* **31**, 1583-1599 (2011).

7 Parkes, S. L. & Balleine, B. W. Incentive Memory: Evidence the Basolateral Amygdala Encodes and the Insular Cortex Retrieves Outcome Values to Guide Choice between Goal-Directed Actions. *J Neurosci* **33**, 8753-8763, doi:10.1523/JNEUROSCI.5071-12.2013 (2013).

8 Hatfield, T., Han, J. S., Conley, M., Gallagher, M. & Holland, P. Neurotoxic lesions of basolateral, but not central, amygdala interfere with Pavlovian second-order conditioning and reinforcer devaluation effects. *J Neurosci* **16**, 5256-5265 (1996).

9 Johnson, A. W., Gallagher, M. & Holland, P. C. The basolateral amygdala is critical to the expression of pavlovian and instrumental outcome-specific reinforcer devaluation effects. *J Neurosci* **29**, 696-704 (2009).

10 West, E. A. *et al.* Transient inactivation of basolateral amygdala during selective satiation disrupts reinforcer devaluation in rats. *Behav Neurosci* **126**, 563-574, doi:10.1037/a0029080 (2012).

11 Wellman, L. L., Gale, K. & Malkova, L. GABAA-mediated inhibition of basolateral amygdala blocks reward devaluation in macaques. *J Neurosci* **25**, 4577-4586 (2005).

12 Málková, L., Gaffan, D. & Murray, E. A. Excitotoxic lesions of the amygdala fail to produce impairment in visual learning for auditory secondary reinforcement but interfere with reinforcer devaluation effects in rhesus monkeys. *J Neurosci* **17**, 6011-6020 (1997).

13 Salinas, J. A., Packard, M. G. & McGaugh, J. L. Amygdala modulates memory for changes in reward magnitude: reversible post-training inactivation with lidocaine attenuates the response to a reduction in reward. *Behav Brain Res* **59**, 153-159 (1993).

14 Carmichael, S. T. & Price, J. L. Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J Comp Neurol* **363**, 615-641, doi:10.1002/cne.903630408 (1995).

15 Price, J. L. Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions. *Ann N Y Acad Sci* **1121**, 54-71, doi:10.1196/annals.1401.008 (2007).

16 Wassum, K. M., Greenfield, V. Y., Linker, K. E., Maidment, N. T. & Ostlund, S. B. Inflated reward value in early opiate withdrawal. *Addict Biol* **21**, 221-233, doi:10.1111/adb.12172 (2016).

17 Wassum, K. M. *et al.* Silicon Wafer-Based Platinum Microelectrode Array Biosensor for Near Real-Time Measurement of Glutamate In Vivo. *Sensors* **8**, 5023-5036 (2008).

18 Wassum, K. M. *et al.* Transient Extracellular Glutamate Events in the Basolateral Amygdala Track Reward-Seeking Actions. *J Neurosci* **32**, 2734-2746, doi:32/8/2734 [pii]10.1523/JNEUROSCI.5780-11.2012 (2012).

19 Malvaez, M. *et al.* Basolateral amygdala rapid glutamate release encodes an outcome-specific representation vital for reward-predictive cues to selectively invigorate reward-seeking actions. S*ci Rep* 5**,** 12511, doi:10.1038/srep12511 (2015).

20 Balleine, B. W., Garner, C., Gonzalez, F. & Dickinson, A. Motivational control of heterogeneous instrumental chains. J *Exp Psychol Anim Behav Process* 2**1,** 203-217 (1995).

21 Corbit, L. H. & Balleine, B. W. Instrumental and Pavlovian incentive processes have dissociable effects on components of a heterogeneous instrumental chain. J *Exp Psychol Anim Behav Process* 2**9,** 99-106 (2003).

22 Balleine, B., Paredes-Olay, C. & Dickinson, A. Effects of Outcome Devaluation on the Performance of a Heterogeneous Instrumental Chain. I*nternational Journal of Comparative Psychology* 1**8,** 257-272 (2005).

23 Williams, K. Ifenprodil discriminates subtypes of the N-methyl-D-aspartate receptor: selectivity and mechanisms at recombinant heteromeric receptors. M*ol Pharmacol* 4**4,** 851-859 (1993).

24 Sheardown, M. J., Nielsen, E. O., Hansen, A. J., Jacobsen, P. & Honoré, T. 2,3-Dihydroxy-6-nitro-7-sulfamoyl-benzo(F)quinoxaline: a neuroprotectant for cerebral ischemia. S*cience* 2**47,** 571-574 (1990).

25 Ostlund, S. B. & Balleine, B. W. Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. J *Neurosci* 2**7,** 4819-4825 (2007).

26 Ostlund, S. B. & Balleine, B. W. The contribution of orbitofrontal cortex to action selection. A*nn N Y Acad Sci* 1**121,** 174-192 (2007).

27 Bradfield, L. A., Dezfouli, A., van Holstein, M., Chieng, B. & Balleine, B. W. Medial Orbitofrontal Cortex Mediates Outcome Retrieval in Partially Observable Task Situations. N*euron* 8**8,** 1268-1280, doi:10.1016/j.neuron.2015.10.044 (2015).

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

28 Wilson, R. C., Takahashi, Y. K., Schoenbaum, G. & Niv, Y. Orbitofrontal cortex as a cognitive map of task space. N*euron* 8**1,** 267-279, doi:10.1016/j.neuron.2013.11.005 (2014).

29 Schuck, N. W., Cai, M. B., Wilson, R. C. & Niv, Y. Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. N*euron* 9**1,** 1402-1412, doi:10.1016/j.neuron.2016.08.019 (2016).

30 Izquierdo, A. & Murray, E. A. Combined unilateral lesions of the amygdala and orbital prefrontal cortex impair affective processing in rhesus monkeys. J *Neurophysiol* 9**1,** 2023-2039, doi:10.1152/jn.00968.2003 (2004).

31 Murray, E. A. & Izquierdo, A. Orbitofrontal cortex and amygdala contributions to affect and action in primates. A*nn N Y Acad Sci* 1**121,** 273-296, doi:10.1196/annals.1401.021 (2007).

32 Pickens, C. L. *et al.* Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. J *Neurosci* 2**3,** 11078-11084 (2003).

33 Schoenbaum, G., Chang, C. Y., Lucantonio, F. & Takahashi, Y. K. Thinking Outside the Box: Orbitofrontal Cortex, Imagination, and How We Can Treat Addiction. N*europsychopharmacology* 4**1,** 2966-2976, doi:10.1038/npp.2016.147 (2016).

34 Sharpe, M. J. & Schoenbaum, G. Back to basics: Making predictions in the orbitofrontal-amygdala circuit. N*eurobiol Learn Mem* 1**31,** 201-206, doi:10.1016/j.nlm.2016.04.009 (2016).

35 Izquierdo, A. Functional Heterogeneity within Rat Orbitofrontal Cortex in Reward Learning and Decision Making. J *Neurosci* 3**7,** 10529-10540, doi:10.1523/JNEUROSCI.1678-17.2017 (2017).

36 Noonan, M. P., Chau, B., Rushworth, M. F. & Fellows, L. K. Contrasting effects of medial and lateral orbitofrontal cortex lesions on credit assignment and decision making in humans. J *Neurosci,* doi:10.1523/JNEUROSCI.0692-17.2017 (2017).

37 Noonan, M. P. *et al.* Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. P*roc Natl Acad Sci U S A* 1**07,** 20547-20552, doi:10.1073/pnas.1012246107 (2010).

38 Rudebeck, P. H. & Murray, E. A. Balkanizing the primate orbitofrontal cortex: distinct subregions for comparing and contrasting values. A*nn N Y Acad Sci* 1**239,** 1-13, doi:10.1111/j.1749-6632.2011.06267.x (2011).

39 Lichtenberg, N. T. *et al.* Basolateral amygdala to orbitofrontal cortex projections enable cue-triggered reward expectations. J *Neurosci,* doi:10.1523/JNEUROSCI.0486-17.2017 (2017).

40 Jenison, R. L., Rangel, A., Oya, H., Kawasaki, H. & Howard, M. A. Value encoding in single neurons in the human amygdala during decision making. J *Neurosci* 3**1,** 331-338, doi:10.1523/jneurosci.4461-10.2011 (2011).

41 Hernádi, I., Grabenhorst, F. & Schultz, W. Planning activity for internally generated reward goals in monkey amygdala neurons. N*at Neurosci* 1**8,** 461-469, doi:10.1038/nn.3925 (2015).

42 Grabenhorst, F., Hernádi, I. & Schultz, W. Prediction of economic choice by primate amygdala neurons. P*roc Natl Acad Sci U S A* 1**09,** 18950-18955, doi:10.1073/pnas.1212706109 (2012).

43 Orsini, C. A. *et al.* Optogenetic Inhibition Reveals Distinct Roles for Basolateral Amygdala Activity at Discrete Time Points during Risky Decision Making. J *Neurosci* 3**7,** 11537-11548, doi:10.1523/JNEUROSCI.2344-17.2017 (2017).

44 Wallis, J. D. & Miller, E. K. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. E*ur J Neurosci* 1**8,** 2069-2081 (2003).

45 Roesch, M. R. & Olson, C. R. Neuronal activity related to reward value and motivation in primate frontal cortex. S*cience* 3**04,** 307-310, doi:10.1126/science.1093223 (2004).

46 Tremblay, L. & Schultz, W. Relative reward preference in primate orbitofrontal cortex. N*ature* 3**98,** 704-708 (1999).

47 Hosokawa, T., Kato, K., Inoue, M. & Mikami, A. Neurons in the macaque orbitofrontal cortex code relative preference of both rewarding and aversive outcomes. N*eurosci Res* 5**7,** 434-445, doi:10.1016/j.neures.2006.12.003 (2007).

48 Padoa-Schioppa, C. & Assad, J. A. The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. N*at Neurosci* 1**1,** 95-102, doi:10.1038/nn2020 (2008).

49 Wallis, J. D. Orbitofrontal cortex and its contribution to decision-making. A*nnu Rev Neurosci* 3**0,** 31-56, doi:10.1146/annurev.neuro.30.051606.094334 (2007).

50 Gottfried, J. A., O'Doherty, J. & Dolan, R. J. Encoding predictive reward value in human amygdala and orbitofrontal cortex. S*cience* 3**01,** 1104-1107, doi:10.1126/science.1087919 (2003).

51 Rudebeck, P. H., Saunders, R. C., Lundgren, D. A. & Murray, E. A. Specialized Representations of Value in the Orbital and Ventrolateral Prefrontal Cortex: Desirability versus Availability of Outcomes. N*euron* 9**5,** 1208-1220.e1205, doi:10.1016/j.neuron.2017.07.042 (2017).

52 Zeeb, F. D. & Winstanley, C. A. Functional disconnection of the orbitofrontal cortex and basolateral amygdala impairs acquisition of a rat gambling task and disrupts animals' ability to alter decision-making behavior after reinforcer devaluation. J *Neurosci* 3**3,** 6434-6443, doi:10.1523/JNEUROSCI.3971-12.2013 (2013).

53 Fiuzat, E. C., Rhodes, S. E. & Murray, E. A. The role of orbitofrontal-amygdala interactions in updating action-outcome valuations in macaques. J *Neurosci,* doi:10.1523/JNEUROSCI.1839-16.2017 (2017).

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

54 Sharpe, M. J., Wikenheiser, A. M., Niv, Y. & Schoenbaum, G. The State of the Orbitofrontal Cortex. N*euron* 8**8,** 1075-1077, doi:10.1016/j.neuron.2015.12.004 (2015).

55 Jones, J. L. *et al.* Orbitofrontal cortex supports behavior and learning using inferred but not cached values. S*cience* 3**38,** 953-956, doi:10.1126/science.1227489 (2012).

56 Howard, J. D., Gottfried, J. A., Tobler, P. N. & Kahnt, T. Identity-specific coding of future rewards in the human orbitofrontal cortex. P*roc Natl Acad Sci U S A* 1**12,** 5195-5200, doi:10.1073/pnas.1503550112 (2015).

57 Howard, J. D. & Kahnt, T. Identity-Specific Reward Representations in Orbitofrontal Cortex Are Modulated by Selective Devaluation. J *Neurosci* 3**7,** 2627-2638, doi:10.1523/JNEUROSCI.3473-16.2017 (2017).

58 Huang, Y. Y. & Kandel, E. R. Postsynaptic induction and PKA-dependent expression of LTP in the lateral amygdala. N*euron* 2**1,** 169-178 (1998).

59 Bauer, E. P., Schafe, G. E. & LeDoux, J. E. NMDA receptors and L-type voltage-gated calcium channels contribute to long-term potentiation and different components of fear memory formation in the lateral amygdala. J *Neurosci* 2**2,** 5239-5249 (2002).

60 Müller, T., Albrecht, D. & Gebhardt, C. Both NR2A and NR2B subunits of the NMDA receptor are critical for long-term potentiation and long-term depression in the lateral amygdala of horizontal slices of adult mice. L*earn Mem* 1**6,** 395-405, doi:10.1101/lm.1398709 (2009).

61 Riedel, G., Platt, B. & Micheau, J. Glutamate receptor function in learning and memory. B*ehav Brain Res* 1**40,** 1-47 (2003).

62 Rodrigues, S. M., Schafe, G. E. & LeDoux, J. E. Intra-amygdala blockade of the NR2B subunit of the NMDA receptor disrupts the acquisition but not the expression of fear conditioning. J *Neurosci* 2**1,** 6889-6896 (2001).

63 Stopper, C. M., Green, E. B. & Floresco, S. B. Selective involvement by the medial orbitofrontal cortex in biasing risky, but not impulsive, choice. C*ereb Cortex* 2**4,** 154-162, doi:10.1093/cercor/bhs297 (2014).

64 Dalton, G. L., Wang, N. Y., Phillips, A. G. & Floresco, S. B. Multifaceted Contributions by Different Regions of the Orbitofrontal and Medial Prefrontal Cortex to Probabilistic Reversal Learning. J *Neurosci* 3**6,** 1996-2006, doi:10.1523/JNEUROSCI.3366-15.2016 (2016).

65 Gourley, S. L., Zimmermann, K. S., Allen, A. G. & Taylor, J. R. The Medial Orbitofrontal Cortex Regulates Sensitivity to Outcome Value. J *Neurosci* 3**6,** 4600-4613, doi:10.1523/JNEUROSCI.4253-15.2016 (2016).

66 Volkow, N. D. & Fowler, J. S. Addiction, a disease of compulsion and drive: involvement of the orbitofrontal cortex. C*ereb Cortex* 1**0,** 318-325 (2000).

67 Goldstein, R. Z. & Volkow, N. D. Dysfunction of the prefrontal cortex in addiction: neuroimaging findings and clinical implications. N*at Rev Neurosci* 1**2,** 652-669, doi:10.1038/nrn3119 (2011).

68 Ressler, K. J. & Mayberg, H. S. Targeting abnormal neural circuits in mood and anxiety disorders: from the laboratory to the clinic. N*at Neurosci* 1**0,** 1116-1124, doi:10.1038/nn1944 (2007).

69 Sladky, R. *et al.* Disrupted effective connectivity between the amygdala and orbitofrontal cortex in social anxiety disorder during emotion discrimination revealed by dynamic causal modeling for FMRI. C*ereb Cortex* 2**5,** 895-903, doi:10.1093/cercor/bht279 (2015).

70 Price, J. L. & Drevets, W. C. Neurocircuitry of mood disorders. N*europsychopharmacology* 3**5,** 192-216, doi:10.1038/npp.2009.104 (2010).

71 Liu, H. *et al.* Differentiating patterns of amygdala-frontal functional connectivity in schizophrenia and bipolar disorder. S*chizophr Bull* 4**0,** 469-477, doi:10.1093/schbul/sbt044 (2014).

72 Zhang, F. *et al.* Multimodal fast optical interrogation of neural circuitry. N*ature* 4**46,** 633-639, doi:10.1038/nature05744 (2007).

73 Schoenbaum, G., Chiba, A. A. & Gallagher, M. Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. J *Neurosci* 1**9,** 1876-1884 (1999).

74 Schoenbaum, G., Chiba, A. A. & Gallagher, M. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. N*at Neurosci* 1**,** 155-159, doi:10.1038/407 (1998).

75 Roesch, M. R., Taylor, A. R. & Schoenbaum, G. Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. N*euron* 5**1,** 509-520, doi:S0896-6273(06)00507-1 [pii]10.1016/j.neuron.2006.06.027 (2006).

76 van Duuren, E. e*t al. N*eural coding of reward magnitude in the orbitofrontal cortex of the rat during a five-odor olfactory discrimination task. L*earn Mem 1*4**,** 446-456, doi:10.1101/lm.546207 (2007).

77 Feltenstein, M. W. & See, R. E. NMDA receptor blockade in the basolateral amygdala disrupts consolidation of stimulus-reward memory and extinction learning during reinstatement of cocaine-seeking in an animal model of relapse. N*eurobiol Learn Mem 8*8**,** 435-444, doi:10.1016/j.nlm.2007.05.006 (2007).

78 Davis, J. D. & Smith, G. P. Analysis of the microstructure of the rhythmic tongue movements of rats ingesting maltose and sucrose solutions. B*ehav Neurosci 1*0**6,** 217-228 (1992).

79 Davis, J. D. & Smith, G. P. Analysis of lick rate measures the positive and negative feedback effects of carbohydrates on eating. Ap*petite 1*1**,** 229-238 (1988).

80 Davis, J. D. & Perez, M. C. Food deprivation- and palatability-induced microstructural changes in ingestive behavior. Am *J Physiol 2*6**4,** R97-103 (1993).

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

81 Berridge, K. C. Modulation of taste affect by hunger, caloric satiety, and sensory-specific satiety in the rat. Ap*petite 1*6**,** 103-120 (1991).

82 Malvaez, M. e*t al. H*DAC3-selective inhibitor enhances extinction of cocaine-seeking behavior in a persistent manner. Pr*oc Natl Acad Sci U S A 11***0,** 2647-2652, doi:10.1073/pnas.1213364110 (2013).

83 Malvaez, M., Mhillaj, E., Matheos, D. P., Palmery, M. & Wood, M. A. CBP in the nucleus accumbens regulates cocaine-induced histone acetylation and is critical for cocaine-associated behaviors. J *Neurosci 3*1**,** 16941-16948, doi:10.1523/JNEUROSCI.2747-11.2011 (2011).

84 Malvaez, M., Sanchis-Segura, C., Vo, D., Lattal, K. M. & Wood, M. A. Modulation of chromatin modification facilitates extinction of cocaine-induced conditioned place preference. Bio*l Psychiatry 6*7**,** 36-43, doi:10.1016/j.biopsych.2009.07.032 (2010).

85 Wassum, K. M., Ostlund, S. B., Balleine, B. W. & Maidment, N. T. Differential dependence of Pavlovian incentive motivation and instrumental incentive learning processes on dopamine signaling. Le*arn Mem 1*8**,** 475-483, doi:18/7/475 [pii]10.1101/lm.2229311 (2011).

86 Wassum, K. M., Greenfield, V. Y., Linker, K. E., Maidment, N. T. & Ostlund, S. B. Inflated reward value in early opiate withdrawal. Add*ict Biol, d*oi:10.1111/adb.12172 (2014).

87 Wassum, K. M., Ostlund, S. B., Loewinger, G. C. & Maidment, N. T. Phasic Mesolimbic Dopamine Release Tracks Reward Seeking During Expression of Pavlovian-to-Instrumental Transfer. Bio*l Psychiatry 73*, **7**47-755, doi:10.1016/j.biopsych.2012.12.005 (2013).

88 Shull, R. L., Gaynor, S. T. & Grimes, J. A. Response rate viewed as engagement bouts: resistance to extinction. J E*xp Anal Behav 77*, **2**11-231, doi:10.1901/jeab.2002.77-211 (2002).

89 Mellgren, R. L. & Elsmore, T. F. Extinction of operant behavior: An analysis based on foraging considerations Ani*mal Learning & Behavior 19*, **3**17-325 (1991).

90 Kaplan, J. M., Roitman, M. F. & Grill, H. J. Ingestive taste reactivity as licking behavior. Neu*rosci Biobehav Rev 19*, **8**9-98 (1995).

91 Baird, J. P. et *al. Ef*fects of melanin-concentrating hormone on licking microstructure and brief-access taste responses. Am *J Physiol Regul Integr Comp Physiol 29*1**, R**1265-1274 (2006).

92 Thornton-Jones, Z. D., Kennett, G. A., Vickers, S. P. & Clifton, P. G. A comparison of the effects of the CB(1) receptor antagonist SR141716A, pre-feeding and changed palatability on the microstructure of ingestive behaviour. Psy*chopharmacology (Berl) 19*3**, 1**-9 (2007).

93 Levin, J. R., Serlin, R.C., Seaman & M.A. A controlled powerful multiple-comparison strategy for several situations . Psy*chological Bulletin 11*5**, 1**53-159 (1994).

94 Tabachnick, B. G., Fidell, L. S. & Osterlind, S. J. Using multivariate statistics. (2001).

95 Wassum, K. M. et *al. Si*licon Wafer-Based Platinum Microelectrode Array Biosensor for Near Real-Time Measurement of Glutamate in Vivo. Sen*sors (Basel) 8,* **5**023-5036, doi:10.3390/s8085023 (2008).

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

DATA AVAILABILITY

All data that support the findings of this study are available from the corresponding author.

ACKNOWLEDGEMENTS

This research was supported by NIH grant DA035443, MH106972, and NS087494 to KMW and NIH grant DA038942 and DA024635 to MM. We would like to acknowledge the helpful feedback from Nina Lichtenberg and Dr. Alicia Izquierdo on these data and this manuscript.

AUTHOR CONTRIBUTIONS

MM and KMW designed the research, analyzed, and interpreted the data. MM conducted the research with assistance from VYG, CS, and MDM. MM and KMW wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no biomedical financial interests or potential conflicts of interest.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURES

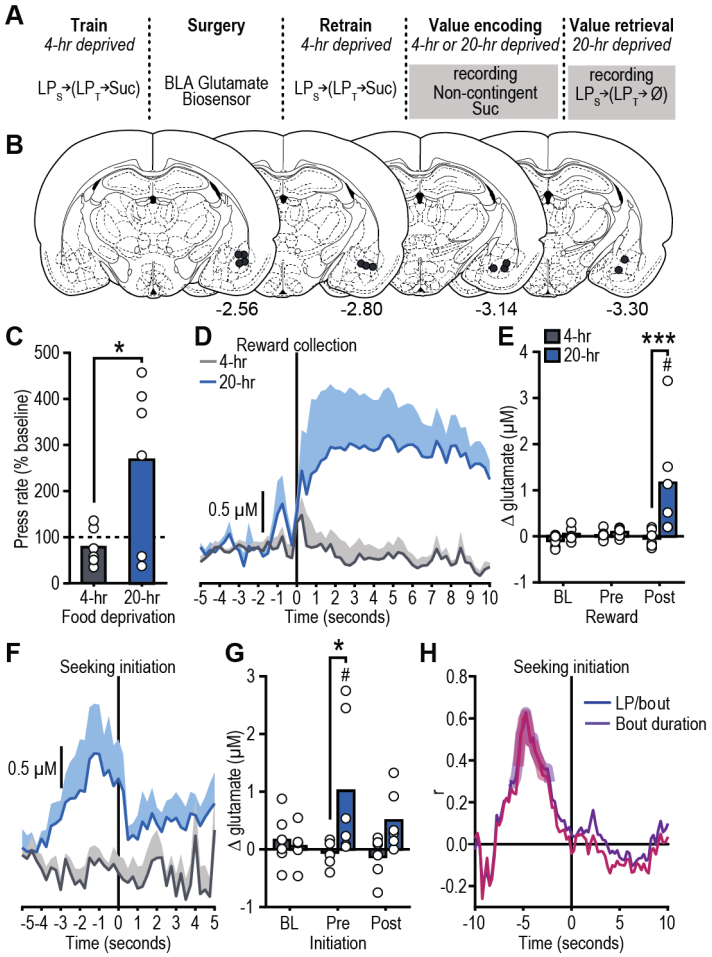


Figure 1. BLA glutamate release tracks reward value encoding and retrieval. (**a**) Procedure schematic (LPs, seeking lever press; LPt, taking lever press; Suc, sucrose; Ø, no reward). (**b**) Representation of biosensor tip placements. Numbers represent anterior-posterior distance (mm) from bregma. (**c**) Reward-seeking press rate (seeking presses/min), normalized to baseline press rate (average of last-two training sessions in 4-hr food-deprived state prior to test; dashed line), during lever-pressing probe test in the hungry state for rats given prior non-contingent sucrose exposure in control sated (4-hr food-deprived; no value encoding) or hungry (20-hr deprived; value encoding opportunity) state. (**d**) Trial-averaged BLA glutamate concentration v. time trace (shading reflects between-subject s.e.m) and (**e**) quantification (mean + scatter) of average glutamate change prior to (pre) and following (post) reward collection/consumption (occurring at time 0 s), or equivalent baseline periods (BL) during non-contingent sucrose re-exposure in sated or hungry. (**f**) Trial-averaged BLA glutamate concentration v. time trace and (**g**) quantification of average glutamate change around bout-initiating reward-seeking presses during the lever-pressing probe test in the hungry state. (**h**) Correlation coefficient between glutamate concentration at each time point around and reward-seeking bout initiation and either total seeking presses in or the duration of the subsequent bout. Shaded region indicates significant at at least $P<0.05$. ($N=6$/group) $^{*}P<0.05$, $^{**}P<0.01$, between groups; $^{\#}P<0.05$, relative to baseline.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

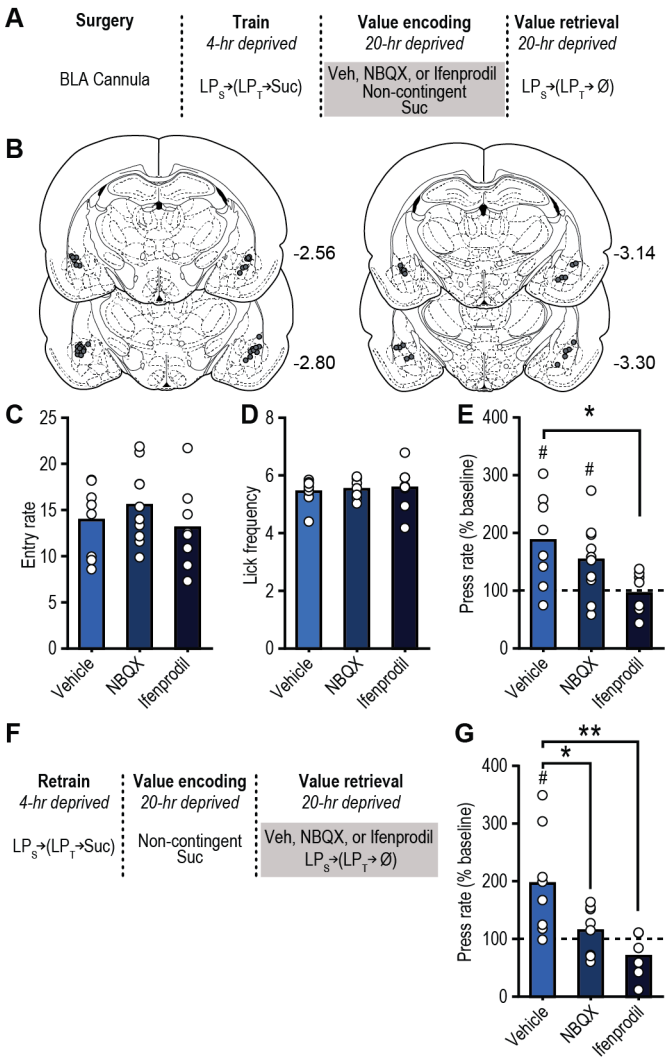


Figure 2. BLA glutamate receptor activity is necessary for reward value encoding and retrieval. (**a**) Procedure schematic (LPs, seeking lever press; LPt, taking lever press; Suc, sucrose; Ø, no reward; Veh, Vehicle; NBQX, AMPA antagonist; Ifenprodil, NMDA antagonist). (**b**) Microinfusion injector tip placements. Numbers represent anterior-posterior distance (mm) from bregma. (**c**) Food-port entry rate (entries/min) and (**d**) palatability responses (lick frequency) during non-contingent sucrose re-exposure in hungry state (20-hr deprived; value encoding opportunity) following intra-BLA infusion of Vehicle (*N*=8), AMPA (*N*=10), or NMDA (*N*=9) antagonist. (**e**) Reward-seeking press rate (seeking presses/min), relative to baseline press rate (dashed line), during drug-free, lever-pressing probe test in hungry state. (**f**) Procedure schematic. (**g**) Following off-drug sucrose re-exposure in hungry state, reward-seeking press rate, relative to baseline, during the on-drug (intra-BLA Vehicle (*N*=8), AMPA (*N*=8), or NMDA (*N*=7) antagonist) lever-pressing probe test in the hungry state. **P*<0.05, ***P*<0.01 between groups; #*P*<0.05 relative to baseline. Data presented as mean + scatter.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

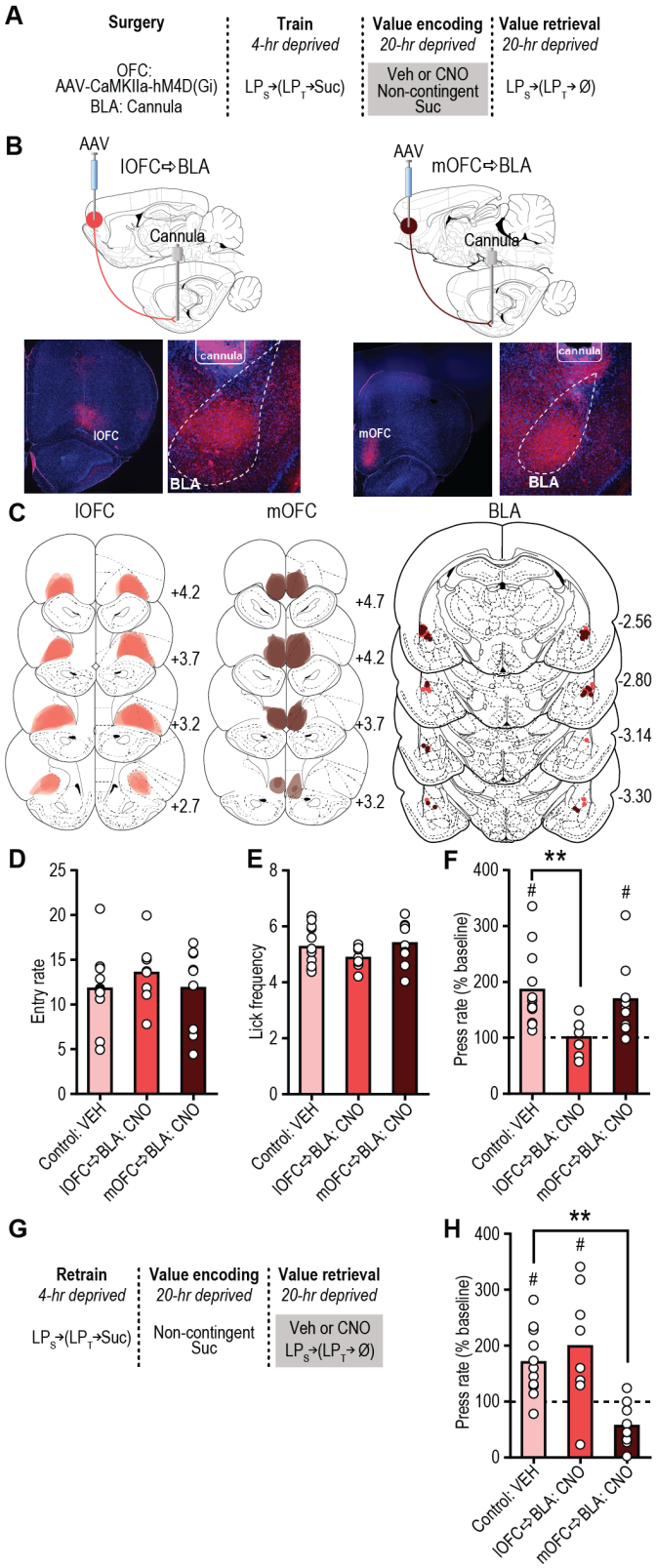


Figure 3. IOFC→BLA and mOFC→BLA projections are necessary for reward value encoding and retrieval, respectively. (**a**) Procedure schematic (LPs, seeking lever press; LPt, taking lever press; Suc, sucrose; Ø, no reward; Veh, Vehicle; CNO, Clozapine N-oxide). (**b**) *Top,* schematic of

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

chemogenetic approach for inactivation of lOFC (*left*) or mOFC (*right*) terminals in the BLA. *Bottom*, representative immunofluorescent images of HA-tagged hM4D(Gi) expression in lOFC (*left*) or mOFC (*right*) and cannula above terminal expression in the BLA. (**c**) Schematic representation of hM4D(Gi) expression in lOFC or mOFC and placement of microinfusion injector tips in the BLA for all subjects. Numbers represent anterior-posterior distance (mm) from bregma. (**d**) Food-port entry rate (entries/min) and (**e**) palatability responses (lick frequency) during non-contingent sucrose re-exposure in hungry state (20-hr deprived; value encoding opportunity) following intra-BLA infusion of Vehicle (N=12) or CNO (lOFC→BLA:CNO, N=8; mOFC→BLA:CNO, N=9). (**f**) Reward-seeking press rate (seeking presses/min), relative to baseline press rate (dashed line), during drug-free, lever-pressing probe test in hungry state. (**g**) Procedure schematic. (**h**) Following off-drug sucrose re-exposure in hungry state, reward-seeking press rate, relative to baseline, during the on-drug test (intra-BLA Vehicle (N=11) or CNO (lOFC→BLA:CNO, N=8; mOFC→BLA:CNO, N=9)). $^{**}P<0.01$, between groups; $^{\#}P<0.05$ relative to baseline. Data presented as mean + scatter.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

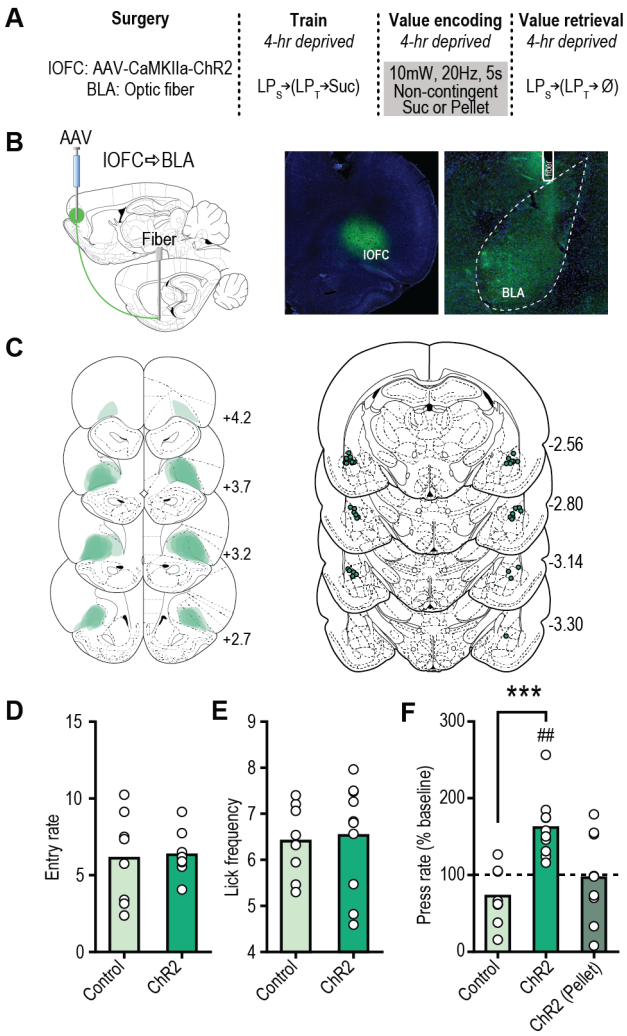


Figure 4. Optical stimulation of lOFC terminals in BLA concurrent with reward experience is sufficient to drive positive value assignment. (**a**) Procedure schematic (LPs, seeking lever press; LPt, taking lever press; Suc, sucrose; Ø, no reward). (**b**) *Left,* schematic of optogenetic approach for stimulation of lOFC terminals in the BLA. *Right*, representative fluorescent images of ChR2-eYFP expression in lOFC and BLA terminal field. (**c**) Schematic representation of ChR2 expression in lOFC and placement of optical fiber tips in BLA for all subjects. Numbers represent anterior-posterior distance (mm) from bregma. (**d**) Food-port entry rate (entries/min) and (**e**) palatability responses (lick frequency) during non-contingent sucrose re-exposure in control 4-hr food-deprived sated state. Light (10mW, 20Hz, 5 s) was delivered concurrent with each sucrose collection. Control group consisted of half eYFP-only + 473 nm and half ChR2 + 589 nm light delivery. (**f**) Reward-seeking press rate (seeking presses/min), relative to baseline press rate (dashed line), during manipulation-free, lever-pressing probe test in sated state. Control, N=8; ChR2, N=10; ChR2 (Pellet), N=9. ***P<0.001, between groups; $^{\#\#}P$<0.01, relative to baseline. Data presented as mean + scatter.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

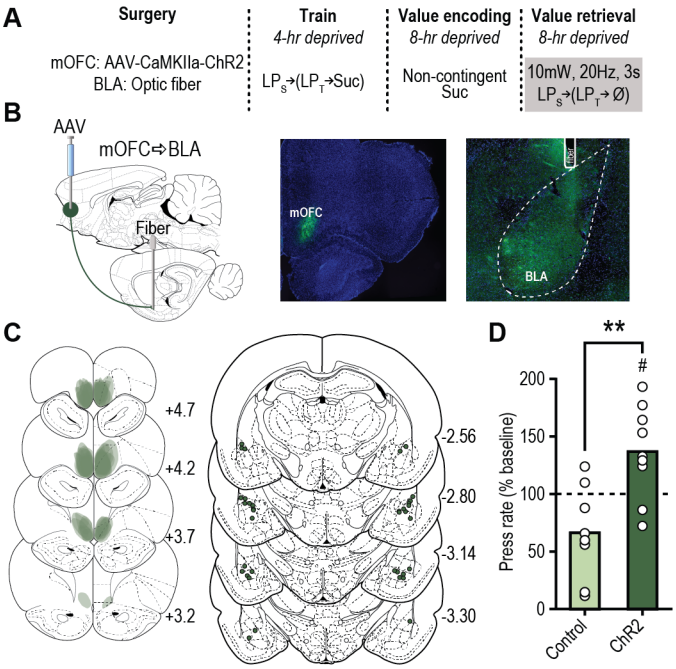


Figure 5. Optical stimulation of mOFC→BLA projections is sufficient to enhance reward value retrieval. (**a**) Procedure schematic (LPs, seeking lever press; LPt, taking lever press; Suc, sucrose; Ø, no reward). (**b**) *Left,* schematic of optogenetic approach for stimulation of mOFC terminals in the BLA. *Right*, representative fluorescent images of ChR2-eYFP expression in mOFC and BLA terminal field. (**c**) Schematic representation of ChR2 expression in mOFC and placement of fiber tips in BLA for all subjects. Numbers represent anterior-posterior distance (mm) from bregma. (**d**) Reward-seeking press rate (seeking presses/min), relative to baseline press rate (dashed line), during lever-pressing probe test in moderate (8-hr food-deprived) hunger state following sucrose re-exposure in 8-hr food-deprived state (a sub-threshold incentive learning opportunity). Light (10mW, 20Hz, 3 s, once/min) was delivered during this test. Control group consisted of half eYFP-only + 473 nm and half ChR2 + 589 nm light delivery. Control, N=8; ChR2, N=9. $^{**}P<0.01$, between groups; $^{\#}P<0.05$, relative to baseline. Data presented as mean + scatter.

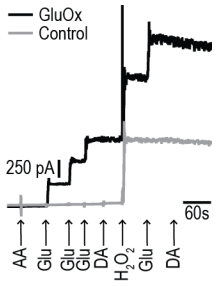
bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplemental Figures

Exposure Group	Number of bouts (per 5-min test)	Average presses per bout	Average bout duration (s)
4-hr	9.33±1.26	2.35±0.33	3.24±0.76
20-hr	9.67±1.69	5.86±2.35	6.10±2.41

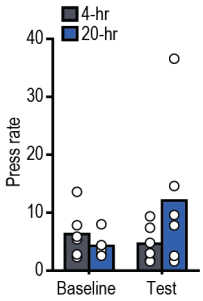
Table S1. Summary of seeking lever press bouts during non-reinforced, lever-pressing probe test. Data presented as average ± s.e.m.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



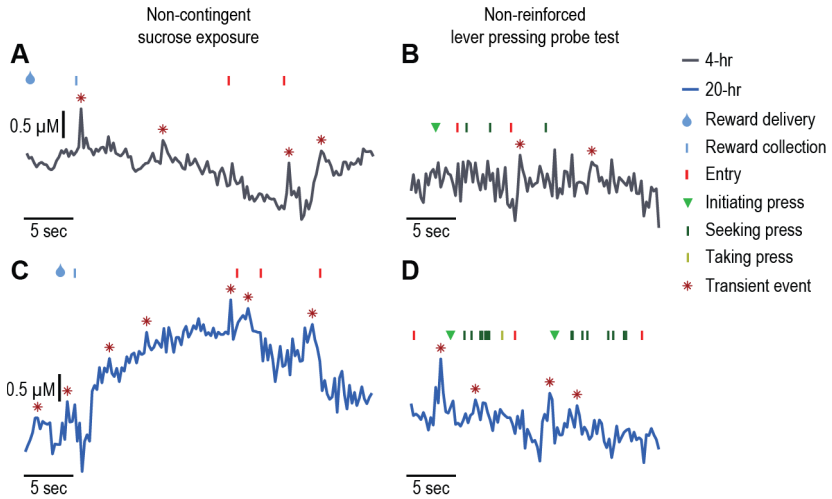
Supplement Figure 1. Representative calibration of a microelectrode array glutamate biosensor. Silicon-wafer-based platinum microelectrode array (MEA) probes were modified for glutamate detection as we have described previously [18,19,95]. Glutamate oxidase (GluOx) serves as the biological recognition. Electro-oxidation, by constant potential amperommetry, of the enzymatically-generated hydrogen peroxide reporter molecule provides the signal. Selectively against both cations and anions is achieved by the addition of polymer coatings (see Methods). Control electrodes are identically coated with the exception that GluOx is omitted. These sensors have a subsecond response time [18,19,95]. To test for sensitivity and selectivity of glutamate measurement, all biosensors were calibrated *in vitro* by sequential addition of ascorbic acid (AA; 250μM), glutamate (Glu; 20 μM), dopamine (DA; 5 μM), hydrogen peroxide (H_2O_2; 20 μM), Glu (40 μM), and DA (10 μM), in stirred PBS at 37 ^{o}C. The *in vitro* glutamate current response was used to determine the electrode-specific calibration factor, which averaged 135.98 μM/nA for the sensors used in these studies. The sensitivity to peroxide between glutamate oxidase coated (GluOx) and control sites did not differ more than 10% (t_{42}=0.32, p=0.75). The average *in vivo* limit of glutamate detection of the sensors used in this study was 0.36 μM (sem=0.03, range 0.13-0.67 μM).

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



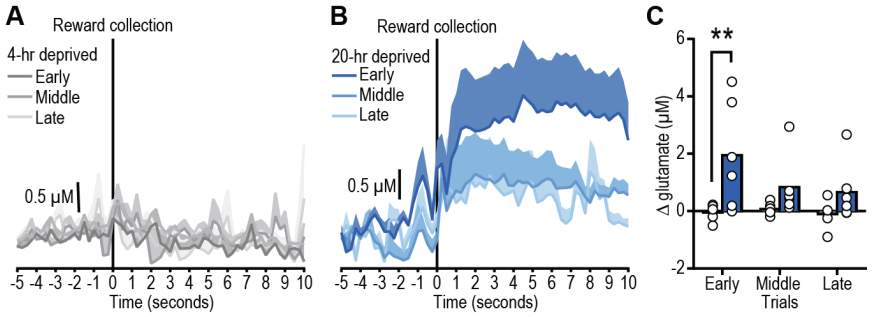
Supplement Figure 2. Effect of incentive learning on reward seeking- raw press rates. Reward-seeking press rate (seeking presses/min) during baseline (average of last-two training sessions in 4-hr food-deprived state prior to test) and non-reinforced, lever-pressing probe test in the hungry state (Test: $F_{1,10}$=3.1.577, P=0.24; Deprivation: $F_{1,10}$=0.71, P=0.42; Test x Deprivation: $F_{1,10}$=3.73, P=0.08) for rats prior non-contingent sucrose exposure in control sated (4-hr food-deprived; no value encoding) or hungry (20-hr deprived; value encoding opportunity) state. Data presented as mean + scatter.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



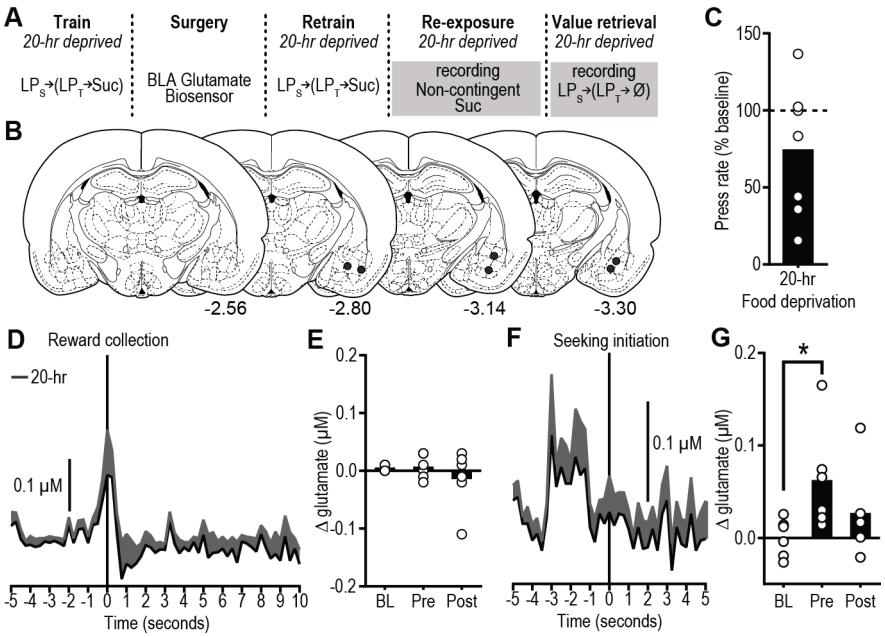
Supplement Figure 3. Representative BLA glutamate v. time traces during reward value encoding and retrieval. Representative, single-trial BLA glutamate concentration v. time traces from a rat that received non-contingent sucrose re-exposure in (**a-b**) the control, familiar sated (4-hr), or (**c-d**) the novel hungry (20-hr; positive value encoding opportunity) state around (**a, c**) reward collection during the non-contingent re-exposure and (**b, d**) the subsequent lever-pressing activity in the non-reinforced probe test conducted in the hungry state.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



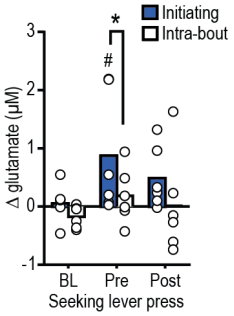
Supplement Figure 4. BLA glutamate release during non-contingent reward re-exposure - binned. (**a-b**) Trial-averaged, BLA glutamate v. time trace (shading reflects between-subject s.e.m) around reward collection and (**c**) quantification (mean + scatter) of average glutamate immediately post reward consumption during early (1-10), middle (11-20), or late (21-30) reward-delivery trials (**a**) in the familiar sated (4-hr food deprived) state or (**b**) novel hungry (20-hr food-deprived, incentive learning opportunity) state (Trial bin: $F_{2,20}$=3.70, P=0.04; Deprivation: $F_{1,10}$=5.52, P=0.04; Trial bin x Deprivation: $F_{2,20}$=3.81, P=0.04). Reward-evoked glutamate release is largest early in the re-exposure session, when incentive learning is expected to be the highest.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



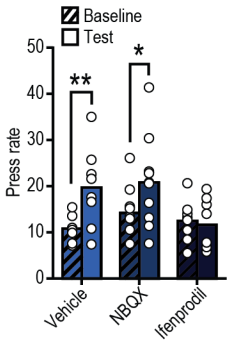
Supplement Figure 5. BLA glutamate release during reward value encoding and retrieval in familiar hungry state. (**a**) Procedure schematic (LPs, seeking lever press; LPt, taking lever press; Suc, sucrose; Ø, no reward). Rats were trained while hungry (20-hr food-deprived) to press on the seeking-taking chain to earn sucrose reward. Biosensor glutamate recordings were made during non-contingent re-exposure to the sucrose in the familiar hungry state and during a lever-pressing probe test also in the hungry state. (**b**) Placement of the microelectrode array biosensor tips in BLA. Numbers represent anterior-posterior distance (mm) from bregma. (**c**) Reward-seeking press rate (seeking presses/min), relative to baseline press rate (dashed line), during non-reinforced, lever-pressing probe test in the hungry (20-hr food-deprived) state. (**d**) Trial-averaged BLA glutamate concentration v. time trace (shading reflects between-subject s.e.m.) and (**e**) quantification (mean + scatter) of average glutamate concentration change prior to (pre) and following (post) reward collection/consumption (occurring at time 0 s), or equivalent baseline periods (BL) during non-contingent sucrose re-exposure in familiar hungry state (N=6; $F_{2,10}$=0.86, P=0.409). (**f**) Trial-averaged BLA glutamate concentration v. time trace and (**g**) quantification of average glutamate concentration change around bout-initiating reward-seeking presses during the lever-pressing probe test in the hungry state ($F_{2,10}$=4.13, P=0.049). *$p < 0.05$, relative to baseline. Reward experience in the hungry state does not increase BLA glutamate concentration in the absence of encoding new information about the value of the reward in that state.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



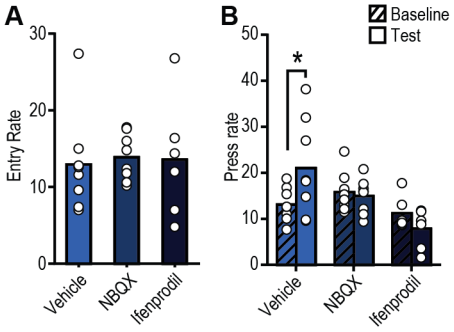
Supplement Figure 6. BLA glutamate concentration around all reward-seeking presses. Quantification of average glutamate concentration change around all seeking presses, separating intra-bout presses from bout-initiating seeking presses, during the lever-pressing probe test in the hungry state for subjects that had prior incentive learning experience with the sucrose in the hungry state (Time: $F_{2,10}$=3.07, P=0.09; Press type: $F_{1,5}$=8.15, P=0.04; Time x Press type: $F_{2,10}$=0.96, P=0.42). *P<0.05, between groups; $^{\#}P$<0.05, relative to baseline. Glutamate transients do not precede each individual press but rather only precede bout-initiating reward-seeking presses. Data presented as mean + scatter.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



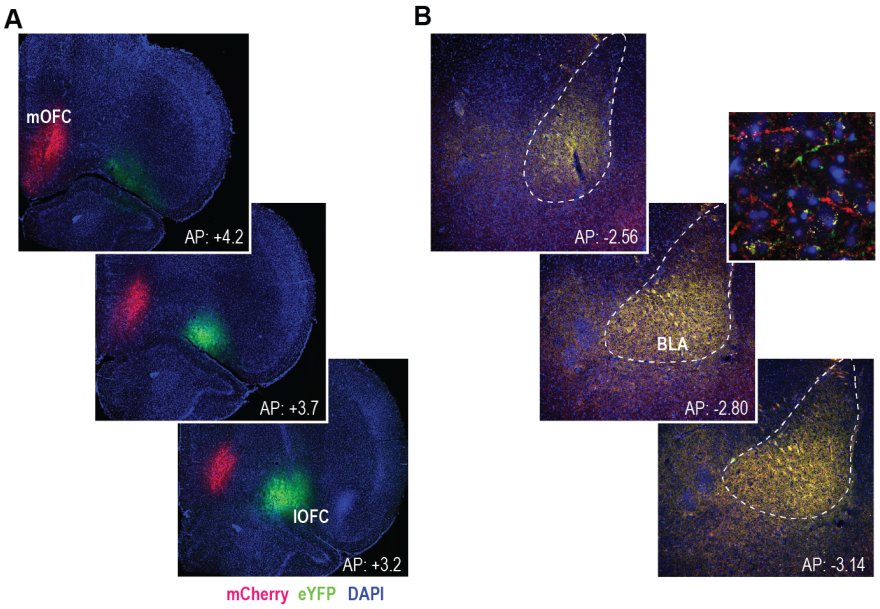
Supplement Figure 7. Effect of glutamate receptor antagonist on value encoding - raw press rates. Reward-seeking press rate (seeking presses/min) during baseline and drug-free non-reinforced lever-pressing probe test in the hungry state (Test: $F_{1,23}$=12.57, P=0.002; Treatment: $F_{2,23}$=2.01, P=0.16; Test x Treatment: $F_{2,23}$=4.31, P=0.03) in rats that received BLA microinfusion of vehicle, AMPA, or NMDA antagonist during prior non-contingent sucrose exposure in hungry (20-hr) state. *P<0.05, **P<0.01 compared to baseline. Data presented as mean + scatter.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



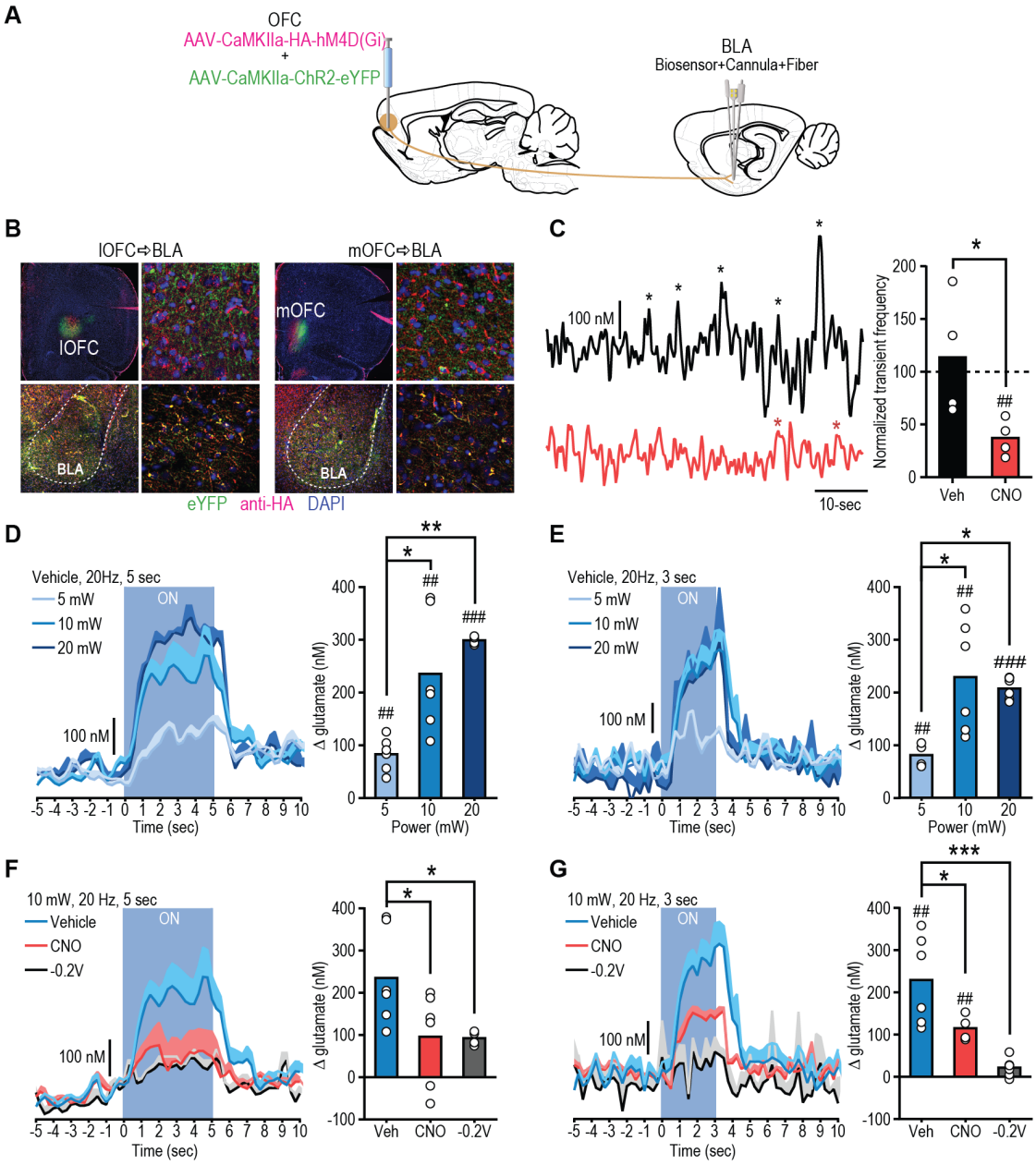
Supplement Figure 8. Effect of glutamate receptor antagonists on reward checking and reward seeking - raw entry/press rates. Following non-contingent sucrose exposure in hungry (20-hr food-deprived) state, rats received intra-BLA of Vehicle, AMPA, or NMDA antagonist prior to a non-reinforced, lever-pressing probe test in the hungry state. (**a**) Food-port entry rate (entries/min; $F_{2,19}$=0.06, P=0.95) during this test. Neither treatment affected this reward-checking measure. (**b**) Reward-seeking press rate (seeking presses/min) during baseline and the on-drug post-re-exposure, non-reinforced, lever-pressing probe test (Test: $F_{1,19}$=0.69, P=0.42; Treatment: $F_{2,19}$=4.95, P=0.02; Test x Treatment: $F_{2,19}$=5.44, P=0.01). *P<0.05, relative to baseline. Data presented as mean + scatter.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplement Figure 9. lOFC and mOFC projections to the BLA. AAV5-CaMKIIa-mCherry was infused in the mOFC and AAV5-CaMKIIa-eYFP was infused into the lOFC and allowed to express for 8 weeks, to ensure terminal expression, prior to histological assessment. (**a**) Representative expression of mCherry and eYFP in the mOFC and lOFC, respectively. (**b**) Terminal expression of mCherry and eYFP in the BLA. These data provide anatomical evidence of projections from both the ventrolateral OFC and medial OFC to the BLA.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

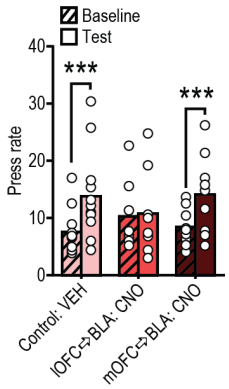


Supplement Figure 10. Chemogenetic and optogenetic manipulation of OFC terminals in BLA. (**a**) Procedure schematic. hM4d(Gi) and ChR2 were co-expressed in either the lOFC or mOFC. Following 8 weeks for terminal expression, we measured spontaneous and optically-evoked glutamate release events in the BLA of anesthetized rats prior to and following CNO infusion. (**b**) Representative immunofluorescent images of HA-tagged hM4D(Gi) and eYFP-tagged ChR2 expression in lOFC (*left*) mOFC (*right*) and BLA terminal expression (*bottom*). Data from lOFC and mOFC subjects was collapsed following evidence of no statistically significant differences between these groups. (**c**) Representative glutamate concentration v. time trace showing spontaneous, transient glutamate release events following vehicle or CNO treatment and quantification of glutamate transient frequency normalized to pre-infusion baseline frequency (dashed line) (t_6=2.54, p=0.04). (**d-e**) Optically-evoked BLA glutamate concentration v. time trace (shading reflects s.e.m.) and quantification (mean + scatter) of optically-evoked glutamate concentration changes. Blue light delivery for (**d**) 5-sec ($F_{2,13}$=11.65, P=0.001) or (**e**) 3-sec ($F_{2,13}$=6.34, P=0.01) over OFC

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

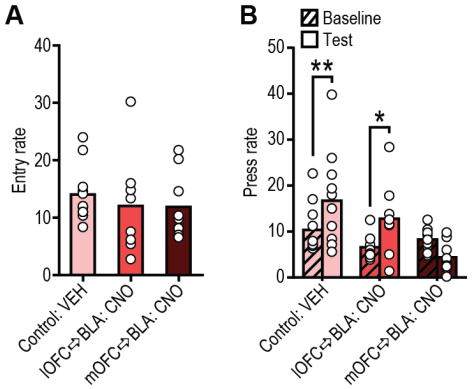
terminals in the BLA power-dependently evoked a glutamate concentration change. (**f-g**) Glutamate concentration v. time trace around (**f**) 5-sec ($F_{2,13}$=3.77, P=0.05) or (**g**) 3-sec ($F_{2,13}$=13.80, P>0.001) optical stimulation of OFC terminals in BLA following intra-BLA Vehicle or CNO infusion and quantification. Optically-evoked response following CNO did not differ from current changes detected below the H_2O_2 (glutamate reporter molecule) oxidizing potential (0.2 V). *P<0.05, ** P<0.01, *** P<0.001, between groups; $^{\#}P$<0.05, $^{\#\#}P$<0.01 relative to baseline. See also [39] for additional validation of OFC→BLA chemogenetic and optogenetic terminal manipulations.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



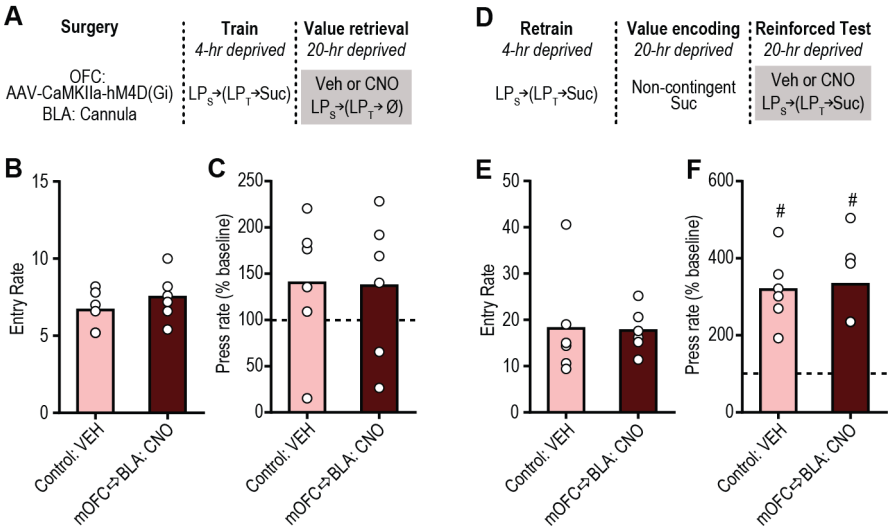
Supplement Figure 11. Effect of inactivation of lOFC or mOFC terminals in the BLA on reward value encoding - raw press rates. Reward-seeking press rate (seeking presses/min) during baseline and drug-free, non-reinforced lever-pressing probe test in the hungry state (Test: $F_{1,26}$=22.94, P<0.0001; Treatment: $F_{2,26}$=0.04, P=0.96; Test x Treatment: $F_{2,26}$=4.21, P=0.03) for rats that received BLA microinfusion of Vehicle or CNO during the non-contingent sucrose re-exposure in the hungry (20-hr food-deprived) state. **P<0.01, ***P<0.001, relative to baseline. Data presented as mean + scatter.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



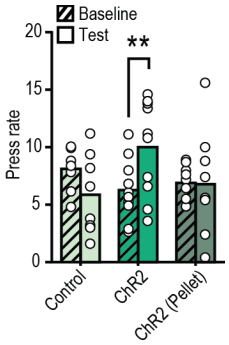
Supplement Figure 12. Effect of inactivation of lOFC or mOFC terminals in the BLA on reward-checking and reward seeking - raw entry/press rates. Following non-contingent sucrose exposure in hungry (20-hr food-deprived) state, rats received BLA microinfusion of vehicle or CNO prior to a non-reinforced lever-pressing probe test in the hungry state. (**a**) Food-port entry rate (entries/min) was not altered by inactivation of either lOFC or mOFC terminals in the BLA during this test ($F_{2,25}$=0.36, P=0.70). (**b**) Reward-seeking press rate (seeking presses/min) during baseline and the on-drug post-re-exposure, non-reinforced, lever-pressing probe test (Test: $F_{1,25}$=6.54, P=0.02; Treatment: $F_{2,25}$=4.30, P=0.02; Test x Treatment: $F_{2,25}$=8.94, P=0.001). *P<0.05, **P<0.01, relative to baseline. Data presented as mean + scatter.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



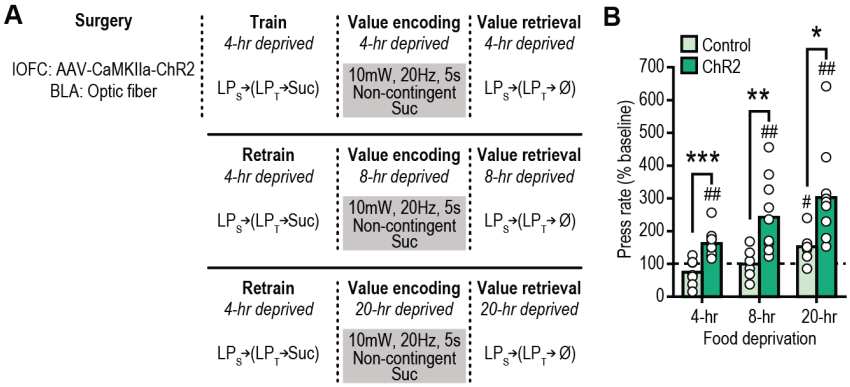
Supplement Figure 13. Inactivation of mOFC terminals in the BLA does not disrupt reward seeking when reward value is not being retrieved from memory. (**a**) Procedure schematic. Rats were trained while sated to lever press on the seeking-taking chain to earn sucrose reward. Following training, they were given two non-reinforced, lever-pressing, probe tests in the hungry state- one each following intra-BLA vehicle or CNO infusion. (LPs, seeking lever press; LPt, taking lever press; Suc, sucrose; Ø, no reward; Veh, vehicle; CNO; Clozapine N-oxide) (**b**) Food-port entry rate (entries/min) (t_5=1.01, p=0.36) and (**c**) reward-seeking press rate (seeking presses/min), normalized to baseline press rate (dashed line), during the non-reinforced lever-pressing probe test in the hungry state following BLA microinfusion of Vehicle or CNO (N=6). mOFC→BLA terminal inactivation was ineffective at altering reward-seeking activity in the absence of prior hunger-induced incentive learning (t_5=0.09, p=0.93). (**d**) Procedure schematic. Following retraining in the sated state, rats were given non-contingent re-exposure to the sucrose in the hungry state (the incentive learning opportunity) and then were given two reinforced lever-pressing tests in the hungry state, one each following BLA vehicle or CNO infusion. (**e**) Food-port entry rate (entries/min) (t_5=0.15, p=0.89) and (**f**) reward-seeking press rate (seeking presses/min), relative to baseline press rate (dashed line). (N=6) mOFC→BLA terminal inactivation was ineffective at altering reward-seeking activity if reward value had been encoded, but did not have to be retrieved because the reward was present at test (t_5=0.34, p=0.75). $^{\#}P$<0.05, relative to baseline. Data presented as mean + scatter.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



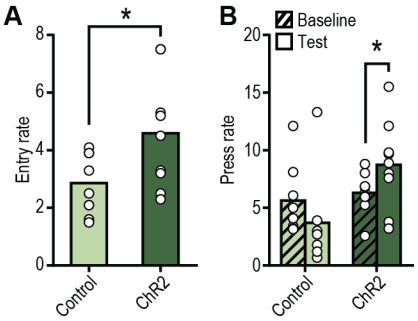
Supplement Figure 14. Effect of optical stimulation of lOFC→BLA projections on value encoding - raw press rates. Reward-seeking press rate (seeking presses/min) during baseline and the post-re-exposure, manipulation-free, non-reinforced, lever-pressing probe test in the sated state (Test: $F_{1,24}$=0.54, P=0.47; Group: $F_{2,24}$=0.60, P=0.55; Test x Group: $F_{2,24}$=7.89, P=0.002) in rats that received prior non-contingent sucrose or task-irrelevant (Pellet) exposure concurrent with light delivery during in sated state. **P<0.01, relative to baseline. Data presented as mean + scatter.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



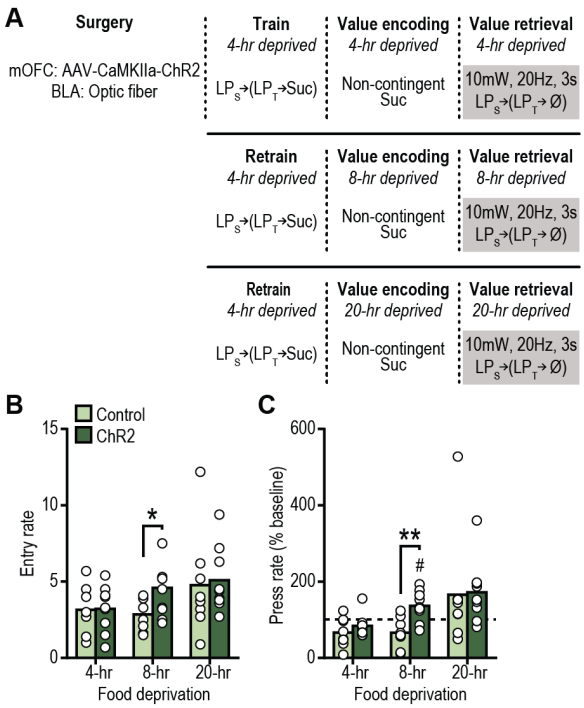
Supplement Figure 15. Activation of lOFC to BLA projections concurrent with sucrose experience is sufficient to enhance value assignment across escalating food-deprivation states. (**a**) Procedure schematic. Rats received 3 test sets in which they first received non-contingent re-exposure to the sucrose with concurrent optical activation of lOFC terminals in BLA (ChR2 + 473 nm, 10mW, 20Hz, 5 s) or control light delivery (control group consisted of half eYFP + 473 nm and half ChR2 + 589 nm light delivery), and then, the next day, received a non-reinforced, lever-pressing probe test in the same deprivation state. (**b**) Reward-seeking press rate (seeking presses/min), relative to baseline press rate (dashed line), during the non-reinforced, lever-pressing probe test conducted the day following non-contingent sucrose re-exposure. At each deprivation state tested, activation of lOFC terminals in the BLA concurrent with non-contingent reward-experience caused a subsequent upshift in reward-seeking activity (Group: $F_{1,15}$=20.74, P=0.0004; deprivation: $F_{2,30}$=7.46, P=0.002; Group x deprivation: $F_{2,30}$=0.73, P=0.49). Planned comparisons: *P<0.05, **P<0.01, ***P<0.001, between groups; $^{\#}P$<0.05 $^{\#\#}P$<0.01, relative to baseline. Data presented as mean + scatter.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



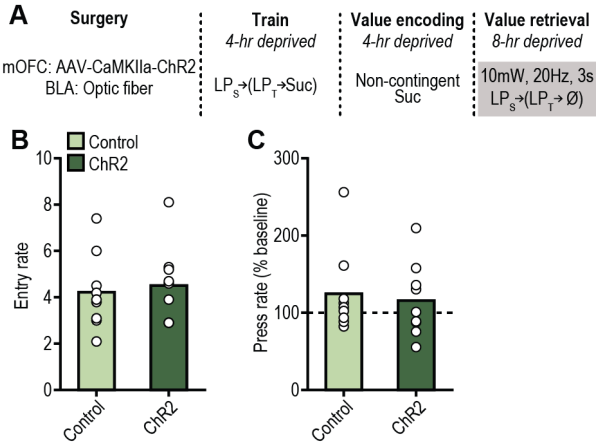
Supplement Figure 16. Effect of optical stimulation of mOFC terminals in the BLA on reward checking and reward seeking - raw entry/press rates. Following non-contingent sucrose exposure in moderate hunger (8-hr food deprived) state, rats received optical stimulation of mOFC terminals in BLA during a non-reinforced, lever-pressing probe test in that moderate hunger state. (**a**) Food-port entry rate (entries/min; t_{15}=2.20, p=0.04) and (**b**) reward-seeking press rate (seeking presses/min) during baseline and the 8-hr food-deprived non-reinforced lever-pressing probe test with optical stimulation of mOFC terminals in BLA (Control, N=8; ChR2, N=9; Test: $F_{1,15}$=0.12, P=0.73; Group: $F_{1,15}$=3.96, P=0.075; Test x Group: $F_{1,15}$=9.74, P=0.007). *P<0.05, between groups (entries) or relative to baseline (presses). Data presented as mean + scatter.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplement Figure 17. Activation of mOFC terminals in the BLA during reward-seeking tests is only sufficient to enhance reward seeking in moderate deprivation state. (**a**) Procedure schematic. Rats received 3 test sets in which they first received non-contingent re-exposure to the sucrose and then, the next day, received a non-reinforced, lever-pressing probe test in the same deprivation state. Light (473 nm, 10mW, 20Hz, 3 s, once/min) was delivered during this test. The control group consisted of half eYFP + 473 nm and half ChR2 + 589 nm light delivery. Rats were tested at escalating food-deprivation levels. The 8-hr food-deprived state provided a sub-threshold incentive learning opportunity. (**b**) Food-port entry rate (entries/min; Group: $F_{1,15}$=0.99, P=0.34; Deprivation: $F_{2,30}$=4.19, P=0.03; Group x Deprivation: $F_{2,30}$=1.07, P=0.36) during the lever-pressing test. (**c**) Reward-seeking press rate (seeking presses/min), relative to baseline press rate (dashed line), during the probe test with optical activation of mOFC terminals in the BLA across escalating food deprivation states (Group: $F_{1,15}$=1.83, P=0.20; deprivation: $F_{2,30}$=7.81, P=0.002; Group x deprivation: $F_{2,30}$=0.99, P=0.38). Planned comparisons: *P<0.05, **P<0.01, between groups. Data presented as mean + scatter. Optical stimulation of mOFC→BLA projections only enhanced reward-seeking activity following subthreshold incentive learning.

bioRxiv preprint doi: https://doi.org/10.1101/299958; this version posted April 11, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplement Figure 18. Optical stimulation of mOFC terminals in BLA projections does not alter reward-seeking without prior value-encoding opportunity. (**a**) Procedure schematic. Rats were trained while sated to lever press on the seeking-taking chain to earn sucrose reward. Following training, they were given re-exposure to the sucrose in the familiar sated state. They were then given a non-reinforced lever-pressing probe test in the 8-hr food-deprived moderate hunger state. Light (473 nm, 10mW, 20Hz, 3 s, once/min) was delivered during this test. The control group consisted of half eYFP + 473 nm and half ChR2 + 589 nm light delivery. (LPs, seeking lever press; LPt, taking lever press; Suc, sucrose; Ø, no reward) (**b**) Food-port entry rate (entries/min) (t_8=0.47, p=0.65) and (**c**) reward-seeking press rate (seeking presses/min), relative to baseline press rate (dashed line), during the probe test in the moderate (8-hr food-deprived) hunger state. mOFC→BLA stimulation was without effect on reward-seeking activity in the moderate hunger state in the absence of prior to reward experience in that state (t_8=0.67, p=0.52). Data presented as mean + scatter.