



Published in final edited form as:

Neuron. 2013 October 16; 80(2): . doi:10.1016/j.neuron.2013.08.008.

Neural estimates of imagined outcomes in the orbitofrontal cortex drive behavior and learning

Yuji K. Takahashi^{1,*†}, Chun Yun Chang^{1,†}, Federica Lucantonio², Richard Z. Haney², Benjamin A. Berg¹, Hau-Jie Yau¹, Antonello Bonci^{1,3,4}, and Geoffrey Schoenbaum^{1,2,3,*}

¹National Institute on Drug Abuse Intramural Research Program, Cellular Neurobiology Research Branch, Behavioral Neurophysiology Research Section, Baltimore, MD 21224

²Department of Anatomy and Neurobiology, University of Maryland, School of Medicine, Baltimore, MD 21201

³Solomon H. Snyder Department of Neuroscience, The Johns Hopkins University, Baltimore, MD 21287

⁴Department of Psychiatry, The Johns Hopkins University, Baltimore, MD 21287

SUMMARY

Imagination, defined as the ability to interpret reality in ways that diverge from past experience, is fundamental to adaptive behavior. This can be seen at a simple level in our capacity to predict novel outcomes in new situations. The ability to anticipate outcomes never before received can also influence learning if those imagined outcomes are not received. The orbitofrontal cortex is a key candidate for where the process of imagining likely outcomes occurs; however its precise role in generating these estimates and applying them to learning remain open questions. Here we address these questions by showing that single-unit activity in orbitofrontal cortex reflects novel outcome estimates. The strength of these neural correlates predicted both behavior and learning, learning which was abolished by temporally-specific inhibition of orbitofrontal neurons. These results are consistent with the proposal that the orbitofrontal cortex is critical for integrating information to imagine future outcomes.

Keywords

orbitofrontal; learning; imagination; single unit; rat

INTRODUCTION

Imagination, defined as the ability to interpret reality in ways that diverge from past experience, is fundamental to normal, adaptive behavior. This can be seen at a very simple level in our capacity to predict novel outcomes in new situations, unbound from our past experience with any particular static element or feature. This ability to imagine new outcomes - to expect or anticipate outcomes that have never before been received - can also

*Correspondence and requests for material should be addressed to: Y.T. (yuji.takahashi@nih.gov) or G.S. (geoffrey.schoenbaum@nih.gov).

[†]These authors contributed equally to this work.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

facilitate learning if those imagined or estimated outcomes turn out to be incorrect. Indeed this is an implicit and distinguishing feature of modern learning theories, in which expectations for reward take into account all predictors that are present even if they have never been encountered together previously (Hall and Pearce, 1982; Lepelley, 2004; Rescorla and Wagner, 1972; Sutton, 1988). The orbitofrontal cortex (OFC) is a key candidate for where the process of imagining likely outcomes occurs (Schoenbaum and Esber, 2010); however its precise role in generating these novel estimates and also its involvement in the application of this information to learning remain unresolved.

To address these questions, we recorded single-unit activity from the OFC during performance of a Pavlovian over-expectation task (Rescorla, 1970). This task consists of three phases: simple conditioning, compound training, and extinction testing. In simple conditioning, rats are trained that several cues predict reward. Subsequently, in compound training, two of the cues are presented together, still followed by the same reward. Typically this results in increased responding to the compound cue. This increased responding – termed summation – is thought to reflect a heightened expectation for reward. Importantly this heightened expectation represents a novel prediction. The rats have never before experienced the cues compounded and have never received a double reward, and yet even on the very first exposure to the compound cue, the rats respond more. This behavior is particularly counterintuitive since the compounded cues each predict the same food pellets, in the same number, delivered in the same location. Thus it is not immediately apparent, based on past experience, that the food pellets should be larger or more plentiful when both cues are presented. Indeed to the extent the compound cue is perceived as a new thing, one would predict less rather than more responding. And while it might seem reasonable for the rats to infer that the food pellets are more likely to appear when both cues are present, the pellets have always come in the past, even when only 1 cue was presented, so increased certainty would not seem to explain the increase in responding. Yet summation does occur, suggesting that the rats jump to the conclusion that the compound cue will be followed by a larger reward. Furthermore, not only is this novel estimate evident in their behavior, it also supports error-based learning when it goes unmet. This learning is evident in the extinction test, when the previously compounded cues are presented separately and without reward. Rats that have shown summation during compound training suddenly respond less to the cues when they are separated.

Prior work has shown that inactivation of the OFC prevents both summation and the resultant extinction learning (Takahashi et al., 2009). These data are consistent with an involvement of the OFC in generating the novel estimates upon which summation and learning depend; however they do not require this. Instead they could reflect the OFC's contribution to signaling the associative strength or learned value of the individual cues based on past experience, with neural summation occurring downstream. Additionally, there are reports that the OFC directly signals reward prediction errors (Sul et al., 2010; Tobler et al., 2006), which could provide an independent explanation for why OFC inactivation during compound training affects learning.

To resolve these accounts, we recorded single-unit activity in the OFC during training in a version of the above task. We reasoned that if the OFC were only representing the associative history or value of the prior cues, then firing to the cues should develop with learning and change during extinction in the probe test, however it should not change substantially at the transition points where novel estimates must be generated, specifically at the point of compounding and perhaps again when the cues are separated. On the other hand, if OFC is involved in generating these novel estimates then some population of neurons in the OFC should increase firing spontaneously in concert with the sudden changes in

behavior at these two transition points. Indeed the firing of these neurons might even predict the resultant summation and learning.

RESULTS

We recorded single-unit activity from OFC in fifteen rats during training on a modified version of the Pavlovian over-expectation task (Fig. 1a). The results to be presented below came from 37 rounds of training in which we observed evidence of over-expectation; data from a handful of sessions in which we did not observe evidence of over-expectation (i.e. in which rats presumably adopted a different strategy) are analyzed separately (see Supplemental Material). The Pavlovian over-expectation task was identical to that used in prior inactivation studies (Haney et al., 2010; Takahashi et al., 2009), except that the transition points between simple conditioning and compound training and between compound training and extinction testing were compressed into two “probe” sessions. This was done to allow us to examine firing in single-units across these critical transition points, without any question as to whether we were recording from the same neurons. All other data come from sessions separated by at least a day; we will not make any claims about whether we are recording the same neurons across days (see Table 1 for a full accounting of the numbers of neurons recorded in different phases).

Electrodes were implanted prior to any training (Fig. 1b). After recovery from surgery, rats were food deprived and underwent simple conditioning, during which cues were paired with flavored sucrose pellets (banana and grape, designated as O1 and O2, counterbalanced). We have shown elsewhere that these flavored pellets are equally preferred but discriminable (Burke et al., 2008). Three unique auditory cues (tone, white noise and clicker, designated A1, A2, and A3, counterbalanced) were the primary cues of interest. A1 served as the “over-expected cue” and was associated with three pellets of O1. A2 served as a control cue and was associated with three pellets of O2. A3 was associated with no reward and thus served as a CS-. Rats were also trained to associate a visual cue (cue light, V) with three pellets of O1. V was to be paired with A1 in the compound phase to induce over-expectation; therefore a non-auditory cue was used in order to discourage the formation of compound representations.

As expected, rats developed conditioned responding and phasic neural responses to the cues predictive of reward across sessions (Fig. 2a). A 2-factor ANOVA (session X cue) of conditioned responding during cue presentation demonstrated significant main effects of both factors as well as a significant interaction (p values < 0.01). Post-hoc testing showed that there were no differences in responding to A1 and A2 at any point in training (p values > 0.68).

This increase in conditioned responding to the cues paired with reward was paralleled by an increase in the proportion of single units responding to the cues (Fig. 2b and c). Cue-evoked activity was present in 46% of OFC neurons recorded in the first two sessions of conditioning. This included 28% that increased firing to at least one of 4 cues and 18% that suppressed firing. The proportion of neurons that showed a phasic increase in firing grew steadily across conditioning, reaching 55% by the last two conditioning sessions. Interestingly, the proportion of neurons that suppressed firing did not change substantially (Fig. 2b). Thus, all subsequent analyses of associative encoding were conducted on the population of neurons that showed excitatory phasic responses to the cues.

Firing of cue-responsive OFC neurons increases spontaneously when two cues are presented in compound and then declines with further training

After simple conditioning, the rats were trained in a compound probe session (CP in Fig. 1a). This single session consisted of additional conditioning (CP 1/2) followed by compound training (CP 2/2), in which A1 and V were presented concurrently (A1/V) followed by the same reward as initial conditioning. A2, A3 and V were presented throughout. As expected, rats showed a significant increase in responding to A1 when it was presented in compound with V (Fig. 3a, inset, ANOVA, $F(1,27) = 4.26$, $p < 0.05$). Responding to A2 control cue did not change between two phases (Fig. 3a, inset, ANOVA, $F(1,27) = 1.10$, $p = 0.30$).

We recorded 130 neurons during these compound probe sessions, 70 of which exhibited an excitatory response to at least one of the cues. Consistent with the hypothesis that the OFC signals the novel estimates regarding expected outcomes in a setting like over-expectation, summation at the start of compound training was accompanied by a sudden increase in neural activity to the compound cue. This was evident in the population response, which was similar for A1, A2, and V during the conditioning phase, but increased selectively to A1/V at the start of compound training (Fig. 3b). This increase was evident over the entire session, and also when only the first trial of compound training was considered (Fig. 3b, insets and Fig. 3c). A 2-factor ANOVA (cue X phase) comparing firing on the first trial of A1/V versus A2 revealed significant main effects of both cue ($F(2,138) = 16.5$, $p < 0.01$) and phase ($F(1,69) = 4.82$, $p = 0.03$) and a significant interaction between them ($F(2,138) = 13.3$, $p < 0.01$) (Fig. 3c). Direct comparisons showed that firing to A1/V in compound phase was significantly greater than that to A1 in conditioning phase ($F(1,69) = 48.1$, $p < 0.01$), whereas firing to A2 and V did not change (A2: $F(1,69) = 1.21$, $p = 0.27$; V: $F(1,69) = 3.01$, $p = 0.09$) (Fig. 3c).

The effect of compounding the two cues was also evident in the summation index scores, comparing neural activity in each cue-responsive neuron to A1/V, A2 and V during conditioning and compound training (Fig. 3d–g). The distribution of these summation index scores shifted significantly above zero for A1/V (Fig. 3d and e; Wilcoxon signed-rank tests, p 's < 0.01), but not for A2 (Fig. 3g; $p > 0.05$) or V (Fig. 3f; $p > 0.05$). In addition, the distribution of the summation index scores was significantly different between A1/V and either A2 or V (Mann-Whitney U tests, p 's < 0.01). Indeed, the increase in firing to the compound was evident in both A1 and V preferring neurons (Fig. 3i; $p < 0.01$). In fact, activity to the very first presentation of the compound cue at the start of compound training was larger than the sum of the activity to the two individual cues at the end of conditioning (Fig. 3h; $p < 0.01$). In addition, the shift in firing to the A1/V compound cue was directly correlated with the shift in conditioned responding shown by the rat in that session (Fig. 3j). Thus neural summation in OFC predicted behavioral summation.

Importantly, the spontaneous increase in firing to the A1/V compound was not simply a reflection of the increased sensory input associated with the sudden combination of the two cues, but rather seemed to reflect the elevated expectations of reward. This was evident in a trial-by-trial analysis of activity in response to A1 and A2 within the first compound session; while activity to A2 was stable across trials (Fig. 3k, inset, t-test, $p = 0.53$), activity to A1 was highest on the first trial and then declined (Fig. 3k, inset, t-test, $p = 0.025$). A similar pattern was evident in a comparison of the activity to A1 and A2 in OFC neurons recorded in the compound probe test versus that in neurons recorded in the same locations in later compound sessions (CP2 – CP4; see Fig. 3k for n 's). The ratio of activity to A1 versus A2 during conditioning (CP 1/2) was approximately 1, indicating that OFC neurons fired equally to these cues. This ratio increased significantly in the compound phase of the probe (CP 2/2) when A1 and V were presented together (Fig. 3k, ANOVA, $p < 0.01$). However rather than being maintained in subsequent compound sessions, as would be expected if it

were a sensory phenomenon, the ratio gradually decreased (Fig. 3k, ANOVA, $p < 0.01$), returning to near unity by the last compound session.

Firing of cue-responsive OFC neurons decreases spontaneously when a previously compounded cue is presented alone

After compound training, the rats were trained in an extinction probe session (PB in Fig. 1a). This single session consisted of additional compound training (PB 1/2) followed by extinction training, in which A1 and the other auditory cues were presented alone and unreinforced. During the compound training, the rats continued to exhibit elevated responding to the cues predictive of reward (PB 1/2 in Fig. 4a); at this point, responding to A1/V and A2 did not differ statistically (ANOVA, $F(1,27) = 0.33$, $p = 0.57$). However, when A1 was separated from V at the start of extinction, rats showed a sudden and selective decline in responding to A1, which persisted throughout extinction (Fig. 4a). A 2-factor ANOVA (cue X trial) comparing conditioned responding to the cues during extinction revealed significant main effects of cue ($F(2,54) = 114.7$, $p < 0.01$) and trial ($F(7,189) = 37.8$, $p < 0.01$), and a significant interaction ($F(14,378) = 12.3$, $p < 0.01$). Post-hoc comparisons revealed significantly less responding to A1 than A2 ($F(1,27) = 93.6$, $p < 0.01$).

We recorded 140 neurons in these extinction probe sessions, 61 of which exhibited an excitatory phasic response to at least one of the cues. Firing in response to A1/V and A2 in these neurons was similar during the compound phase, (PB 1/2, Fig. 4b and c) but then spontaneously declined to A1, but not A2, at the start of extinction training (PB 1T, Fig. 4b and c). A 2-factor ANOVA comparing average firing to A1 and A2 (cue X phase) revealed significant main effects of both cue ($F(1,60) = 9.95$, $p < 0.01$) and phase ($F(1,60) = 20.5$, $p < 0.01$), and a significant interaction between them ($F(1,60) = 27.1$, $p < 0.01$) (Fig. 4c). Direct comparisons revealed a significant reduction of firing on the 1st trial of the probe phase compared to firing in the compound phase for A1 ($F(1,60) = 51.9$, $p < 0.01$), but not for A2 ($F(1,60) = 0.26$, $p = 0.61$).

Similar effects were evident in the distribution of index scores comparing firing of each neuron to A1 and A2 at the end of compound training versus the 1st trial in extinction. The distribution of these scores was shifted significantly below zero for A1 (Fig. 4d; Wilcoxon signed-rank test, $p < 0.01$), but not for A2 (Fig. 4e; $p = 0.97$), and the distribution of these scores differed significantly between A1 and A2 (Mann-Whitney U test, $p < 0.01$). Interestingly, firing to A1/V at the end of compound training remained larger than the sum of the activity to the two individual cues presented at that same time (Fig. 4f; $p < 0.01$).

Consistent with the hypothesis that this activity is important to behavior, the shift in firing in OFC to the A1 cue on the 1st trial of extinction was directly correlated with reduced responding shown by the rat in that session (Fig. 4g, left). Furthermore, reduced behavioral responding to A1 was inversely correlated with neural summation measured earlier, in the first compound training session (Fig. 4g, right). In other words, the stronger the signaling of novel summed expectancies for reward during compound training in a given rat, the weaker responding to the A1 cue was at the start of extinction training. Thus neural estimates of outcomes in OFC were predictive of both behavior and learning.

Suppression of neural activity in OFC during presentation of the compound cue prevents learning

The neural data described above suggests that elevated activity in OFC to the compound cue is critical for learning. This is consistent with earlier data, in which we showed that pharmacological inactivation of OFC during compound training prevented learning, assessed later during the probe test. However as noted earlier, this work is also consistent

with other explanations, since activity within OFC is suppressed throughout compound training in a non-specific manner. In order to provide a more specific causal test of this hypothesis, we next used optogenetic methods to inhibit activity of OFC neurons just at the time of presentation of the compound cue.

Rats received bilateral infusions of either AAV-CaMKIIa-eNpHR3.0-eYFP (halo, $n = 11$ including 9 that underwent behavioral testing and 2 additional rats used for *ex vivo* recording) or AAV-CaMKIIa-eYFP (control, $n = 9$) into OFC at the same location as our recording work; expression was verified histologically post-mortem (Fig. 5a–c). Lightdependent inhibition of OFC neurons was tested using *ex vivo* recording in 2 rats (Fig. 5d). The remaining rats ($n = 9$) received fiber optic assemblies immediately over the injection sites. Three weeks after surgery, these rats began training in the same over-expectation task described above, except that light was delivered into OFC bilaterally during the presentation of the compound cue (Fig. 5e). While there were neither main effects nor any interactions of group on conditioned responding across either conditioning (F 's < 0.91 ; p 's > 0.61) or during the compound sessions (F 's < 2.41 ; p 's > 0.08) (Supplemental Figure 5), there were significant differences during the subsequent probe test. Specifically NpHR rats in whom light was delivered during the compound cue failed to show any difference in conditioned responding to the A1 versus A2 cues in the subsequent probe test (Fig. 5f), whereas eYFP rats that received the same treatment responded much less to A1 than to A2 (Fig. 5g), particularly on the very first trial of the extinction probe test. This impression was confirmed by a 2-factor ANOVA (cue X group) comparing responding to A1 versus A2 on the first trial, which revealed a significant main effect of group ($F(1,16) = 9.68$, $p < 0.01$) and a significant interaction between cue and group ($F(1,16) = 19.33$, $p < 0.01$). Post-hoc testing showed that this interaction was due to a difference in responding between groups to the A1 but not the A2 cue (p 's < 0.05). As a further control, the same rats were then retrained and over-expectation was repeated (as was done in the recording study), except this time light was delivered not during the compound cue but instead during the inter-trial interval period after each compound. This treatment had no effect on later learning; both groups exhibited lower responding to A1 than to A2 in the probe test (Fig. 5h and i; F 's > 6.57 ; p 's < 0.03).

DISCUSSION

These results distinguish several explanations for the involvement of the OFC in Pavlovian over-expectation and, by extension, other behaviors such as reinforcer devaluation. With regard to over-expectation, we have previously shown that inactivation of the OFC during compound training, via the local infusion of GABA agonists, selectively blocks both behavioral summation, assessed during these sessions, and learning, assessed in drug-free animals during subsequent probe tests (Takahashi et al., 2009). Here we show that neural activity in OFC at the time of summation increases suddenly, on the very first presentation of the compound cue, and then declines, as the heightened expectations of the compound cue go unmet. Activity also suddenly declines again, at the start of extinction training, when the cues are separated. And the neural summation evident on the first trial of compound training predicts both behavior and learning. This pattern of results cannot be easily explained by the reinforcement history of the individual cues, which does not change on the first trial of compound training, nor can it be explained by sensory input, which remains constant during compound training, or even salience or the perception of novelty, which should increase both at the start of compound training and extinction and, moreover, would be anti-correlated with conditioned responding. Instead neural activity to the cues in OFC seems to be best described as reflecting the spontaneous or real-time integration of outcome expectations derived from the individual cues.

That neural activity in OFC reflected the spontaneous integration of outcome expectations in our modified version of the Pavlovian over-expectation task strongly supports a role of OFC in actually estimating the new outcome. While these observations do not by themselves preclude a role in also signaling the significance of the individual cues, this role cannot be unique to the OFC, since inactivation or damage of this area does not generally affect Pavlovian conditioned responding or even discrimination learning where performance can be based on these individual histories (Gallagher et al., 1999; Hornak et al., 2004; Izquierdo et al., 2004; Schoenbaum et al., 2002). Indeed OFC-lesioned rats that were impaired at extinction by over-expectation showed no deficits in extinction by reward omission (see supplemental for Takahashi et al., 2009). These two forms of learning are distinguished only by their requirement for integration of expectancies. This suggests that OFC is not critical either to signaling individual reinforcement histories or, in fact, the actual prediction errors, an inference corroborated by our failure to observe any evidence of error signaling in single-unit activity either here (see Supplemental Material) or previously (Takahashi et al., 2009). The critical role for neural summation in OFC is further supported by observations that, in the current experiment, when rats failed to show evidence of learning as a result of summation, OFC neurons fired normally in most regards except they failed to show neural summation (see Supplemental Material).

Our results here also favor a similar interpretation of the importance of OFC to changes in learned behaviors after reinforcer devaluation (Critchley and Rolls, 1996; Gallagher et al., 1999; Gottfried et al., 2003; Izquierdo and Murray, 2000; Machado and Bachevalier, 2007). Changing performance of a learned response spontaneously after devaluation of the predicted outcome (ie without further contact with the reinforcer) requires the subject to integrate across independently acquired associative structures to imagine what is essentially a novel outcome (Holland and Rescorla, 1975). Work in both monkeys and rats has shown that this change in behavior requires the OFC to be online at the time of responding (Pickens et al., 2005; West et al., 2011). The current data suggests this reflects an involvement of OFC in generating this novel prediction during the decision process, rather than a role in simply storing the various associations or the new value of the outcome.

Of course our data alone do not require that integration happen within OFC; it might occur upstream and simply be transmitted through OFC. However, major afferent areas to the OFC (Groenewegen et al., 1990; Kahnt et al., 2012; Ongur and Price, 2000; Price, 2007), such as amygdala, medial temporal lobe, or even other prefrontal areas, typically do not have OFC's broad involvement in tasks that require integration and novel expectancies. For example, rhinal and hippocampal areas are not required for reinforcer devaluation effects (Chudasama et al., 2008; Thornton et al., 1998), and while the basolateral amygdala is important for reinforcer devaluation (Hatfield et al., 1996; Malkova et al., 1997), it appears to be preferentially involved in the learning rather than the performance phase (Pickens et al., 2003). This suggests a more fundamental role for such afferent regions in acquiring the individual associations and perhaps allowing them to be represented in a way that is accessible later rather than in integrating them in novel ways at the time a decision is made. Accordingly, the basolateral amygdala is not necessary for either over-expectation (Haney et al., 2010) or, typically, for closely-related phenomena such as extinction and reversal learning (Izquierdo and Murray, 2005, 2007; Schoenbaum et al., 2003). Indeed in some recent work, removing the amygdala can facilitate reversal learning (Rudebeck and Murray, 2008).

Of course we do not mean to dismiss the possibility that areas upstream from OFC may contribute to or even accomplish in parallel this sort of integration process. As noted above, there are several reports that the basolateral amygdala is necessary for the expression of devaluation effects, particularly when they are reinforcer-specific (Johnson et al., 2009;

Wellmann et al., 2005). And hippocampus appears to be necessary for tasks involving mediated learning or inference that appear to share this property of imaging and integrating outcomes (Bunsey and Eichenbaum, 1996; Wimmer and Shohamy, 2012). Overall the current evidence shows that the OFC plays a critical role for integrating past reward histories, but other areas – including less well-explored cortical regions - may also contribute to this process.

More broadly, our results might also have implications for proposals that OFC represents value in a common neural currency (Camille et al., 2011; Levy and Glimcher, 2011, 2012; Montague and Berns, 2002; Padoa-Schioppa, 2011; Padoa-Schioppa and Assad, 2006, 2008; Plassmann et al., 2007). If activity in the OFC were signaling value in a common neural currency, then one might expect to see neural summation. Indeed in a cartoon version of this idea, neural activity on the first presentation of the compound cue should be equal to the sum of activity on the last presentation of each individual cue. In other words, 1 plus 1 should equal 2. Yet this is not the case; instead, at both the start (Fig. 3h) and the end of compound training (Fig. 4f), the neural response to the compound cue was actually greater than the sum of the response to its constituent parts. This result is inconsistent with the straightforward addition of the respective values of the two cues. If anything, one might expect some non-linearity in encoding that would reduce or suppress firing to the combined value of the compound cue, since OFC neurons have been shown to adapt to the range of rewards historically available in a given situation (Padoa-Schioppa, 2009; Tremblay and Schultz, 1999). This would predict an initial ceiling effect in coding the value of the compound cue, yet the neural summation shows the opposite property. The increased activity is also at odds with other explanations such as any novelty or encoding of the conjunction between the two cues, since it is present even after several sessions of training, when any novelty should have worn off, and it is correlated with behavior and learning, which would not be the case if higher activity reflected the perception of a new sensory construct. Rather the most parsimonious interpretation of neural supra-summation is that it represents a novel expectation of something never before received. Notably this idea would be somewhat similar to signaling of hypothetical outcomes previously reported in monkey OFC neurons (Abe and Lee, 2011), however in this case the OFC neurons are signaling an outcome that has never previously been received.

In fact, none of the evidence here or in any other study or which we are aware requires that what is represented in the OFC be value at all. Rather in each case, the OFC might be said to contribute information about the path to the outcome and its specific attributes. That signal might include a value attribute or the value attribute might be added elsewhere. Indeed, one perspective on the last 20 years of research on this area is that the OFC's function is orthogonal to a common sense definition of value, since OFC can be shown to be required for behaviors when value is held constant and not for behaviors when value is manipulated directly (Jones et al., 2012; McDannald et al., 2011). What determines the involvement of the OFC in value-guided behavior is the need to infer the path to value. Accordingly much neural activity in the OFC seems to reflect this path in different task variants as much as it does the final good and its scalar value (Luk and Wallis, 2013). Here we show that the fundamental involvement of OFC in inferring that path is the ability to integrate across the individual reinforcement histories of cues in the environment to *imagine* the outcomes. When this occurs in previously experienced settings, this would appear as simple representation of the experiential knowledge, however in a novel setting, as we have employed here, the signal in the OFC clearly is able to represent a novel or imagined outcome. Though we have studied this in a rudimentary way here in rats, we would suggest that this ability to interpret rather than be bound by reality and ones past experiences is likely to be deeply important to what distinguishes the most interesting and the most puzzling aspects of behavior.

EXPERIMENTAL PROCEDURES

Recording Experiment

Subjects—Fifteen male Long-Evans rats (Charles Rivers, 275–300 m on arrival) were housed individually and placed on a 12 h light/dark schedule. All rats were given ad libitum access to food except during testing periods. During testing, rats were food deprived to 85% of their baseline weight. All testing was conducted at the University of Maryland School of Medicine in accordance with the University of Maryland School of Medicine Animal Care and Use Committee and US National Institutes of Health guidelines.

Surgery and Histology—Drivable bundles of 10 25- μ m diameter FeNiCr recording electrodes (Stablohm 675, California Fine Wire, Grover Beach, CA) were surgically implanted under stereotaxic guidance in unilateral OFC (3.0 mm anterior and 3.2 mm lateral to bregma, 4.2 mm ventral to the brain surface). At the end of the study, the final electrode position was marked by the passage of a current through each microwire to create a small iron deposit. The rats were then perfused with 4% PFA and potassium ferrocyanide solution to visualize the iron deposit. The brains were removed from the skulls and processed for histology using standard techniques.

Pavlovian Over-expectation Training—Training and recording were conducted in aluminum chambers approximately 18 inches on each side with sloping walls narrowing to an area of 12 \times 12 inches at the bottom. A food cup was recessed in the center of one end wall. Entries were monitored by photobeam. Two food dispensers containing 45 mg sucrose pellets (Banana or grape-flavored; Bio-serv, Frenchtown, NJ) allowed delivery of pellets in the food cup. White noise or a tone, each measuring approximately 76 dB, was delivered via a wall speaker. A clicker (2 Hz) and a 6W bulb were also mounted on that wall.

Rats were shaped to retrieve food pellets, and then underwent 12 conditioning sessions. In each session, the rats received eight 30 s presentations of three different auditory stimuli (A1, A2 and A3) and one visual stimulus (V) (Coulbourn Instruments). Each session consisted of 8 blocks, and each block consisted of 4 presentations of a cue; inter-trial intervals (periods between cues) ranged from 120–150 s. The order of cue-blocks was counterbalanced and randomized. For all conditioning, V consisted of a cue light, and A1, A2 and A3 consisted of a tone, clicker or white noise, respectively (counterbalanced). Two differently flavored sucrose pellets (banana and grape, designated as O1 and O2, counterbalanced) were used as rewards. A1 and V terminated with delivery of three pellets of O1, and A2 terminated with delivery of three pellets of O2. A3 was paired with no food. After completion of the 12 days of conditioning, rats received a single session of compound probe (CP). During the 1st half of the session, the simple conditioning continued, with 6 trials each of 4 cues, in a blocked design, with order counterbalanced. During the 2nd half of the session, compound training began with 6 trials of concurrent A1 and V presentation, followed by delivery of the same reward as during initial conditioning. A2, A3 and V continued to be presented as in simple conditioning, with 6 trials each stimulus. These cues were also presented in a blocked design with order counterbalanced. After the compound probe, rats received 3 days of compound training sessions (CP2 – CP4) with 12 presentations of A1/V, A2, A3 and V. One day after the last compound training, rats received a single session of extinction probe (PB). During the 1st half of the session, the compound training continued with 6 presentations of A1/V, A2, A3 and V. During 2nd half of the session, rats received eight non-reinforced presentations of A1, A2 and A3, with the order mixed and counterbalanced. In some rats ($n = 11/15$), the electrode was then moved to a new location, and the rats repeated days 11 and 12 of conditioning and then underwent additional rounds of over-expectation training in order to acquire additional data. Neural

data from the initial compound and extinction days (n 's = 25 and 21) were not statistically different from data gathered in later rounds of training (n 's = 45 and 40) and thus these neurons are analyzed together in the text. However separate analyses of the main results are presented in the Supplemental Material.

Response Measures—The primary measure of conditioning to cues was the percentage of time that each rat spent with its head in the food cue during the last 20 s of conditioned stimulus (CS) presentation, as indicated by disruption of the photobeam. We also measured the percentage of time that each rat showed rearing behavior during the last 20 s of the CS period. To correct for time spent rearing, the percentage of responding during the last 20 s of the CS was calculate as follows: % of responding = $100 * (\% \text{ of time in food cup}) / [100 - (\% \text{ of time of rearing})]$.

Single-Unit Recording—Neural activity was recorded using two identical Plexon Multichannel Acquisition Processor Systems (Dallas, TX), interfaced with training chambers described above. After amplification and filtering, waveforms ($> 2.5:1$ signal-to-noise) were extracted from active channels and recorded to disk by an associated workstation with event timestamps. Units were stored using Offline Sorter software from Plexon Inc (Dallas, TX), using a template matching algorithm. Sorted files were processed in Neuroexplorer to extract unit timestamps and relevant event markers and analyzed in Matlab (Natick, MA).

Prior to each session, wires were screened for activity. Active wires were selected for recording, and the session was begun. If fewer than 4/8 wires were active, then the electrode assembly was advanced 40 or 80 μm at the end of the session. Otherwise the electrode was kept in the same position between sessions within a single round of over-expectation training. After the probe test, ending a round of training, the electrode assembly was advanced 80 μm regardless of the number of active wires in order to acquire activity from a new group of neurons in any subsequent training.

Neural data analysis—Firing activity in the last 20 s of each CS was compared to activity in the last 20 s of the pre-CS period by t-test ($p < 0.05$). Neurons with significantly higher activity during at least one of the 4 cues were defined as “cue-responsive” as described in the main text. Normalized firing rate was calculated by dividing the average firing rate during the last 20 sec of CS by the average firing rate in the last 20 sec of pre-CS period.

Optogenetic Experiment

Subjects—Twenty male Long-Evans rats (Charles Rivers, 275–300 m on arrival) were housed individually and placed on a 12 h light/dark schedule. All rats were given ad libitum access to food except during testing periods. During testing, rats were food deprived to 85% of their baseline weight. All testing was conducted at the NIDA-IRP in accordance with the NIDA-IRP Animal Care and Use Committee and US National Institutes of Health guidelines.

Surgery, Histology—AAV-CaMKIIa-eNpHR3.0-eYFP or AAV-CaMKIIa-eYFP (from Gene Therapy Center at University of North Carolina at Chapel Hill, courtesy of Dr. Karl Deisseroth) was injected bilaterally in OFC under stereotaxic guidance at AP -3.0 mm, ML ± 3.2 mm, and DV 4.4 and 4.5 mm from the brain surface. A total 1–1.2 μl of virus (titer $\sim 10^{12}$) per hemisphere was delivered at the rate of ~ 0.1 $\mu\text{l}/\text{min}$ by Picospritzer microinjection system (Parker, Hollins, NH). Two rats that received eNpHR3.0 transgene were saved for later slice work; the remaining rats designated for behavioral testing had

optic fibers (200 μm in core diameter; Thorlab, Newton, NJ) implanted bilaterally at AP -3.0 mm, ML ± 3.2 mm, and DV 4.2 mm. At the end of the study, these rats were perfused with phosphate buffer saline and then 4 % PFA. The brains were then immersed in 30% sucrose/PFA for at least 24 hr. The brains were sliced at 40 μm with microtome. The brain slices were then stained with DAPI (through Vectashield-DAPI, Vector Lab, Burlingame, CA) or NeuroTrace (Invitrogen, Carlsbad, CA) and mounted to slides with Vectashield (in the case of staining with NeuroTrace) mounting media. The location of the fiber tip and NpHR-eYFP or eYFP expression was verified using an Olympus confocal microscope. The Z-stack images were merged and processed in Image J (National Institute of Health).

Ex vivo electrophysiology—Approximately 2 months after surgery, 2 rats that had received AAV-CaMKIIa-eNpHR3.0-eYFP injection were anesthetized with isoflurane and perfused transcardially with ~ 40 ml ice-cold NMDG-based artificial CSF (aCSF) solution containing (in mM) 92 NMDG, 20 HEPES, 2.5 KCl, 1.2 NaH_2PO_4 , 10 MgSO_4 , 0.5 CaCl_2 , 30 NaHCO_3 , 25 glucose, 2 thiourea, 5 Na-ascorbate, 3 Na-pyruvate and 12 N-acetyl-L-cysteine (300–310 mOsm, pH 7.3 ~ 7.4). After perfusion, the brain was immediately removed and 300 μm coronal brain slices containing the OFC were made using a Vibratome (Leica, Nussloch, Germany). The brain slices were recovered for less than 15 min at 32°C in NMDG-based aCSF and then transferred and stored for at least 1 hour in HEPES-based aCSF containing (in mM) 92 NaCl, 20 HEPES, 2.5 KCl, 1.2 NaH_2PO_4 , 1 MgSO_4 , 2 CaCl_2 , 30 NaHCO_3 , 25 glucose, 2 thiourea, 5 Na-ascorbate, 3 Na-pyruvate and 12 N-acetyl-L-cysteine (300–310 mOsm, pH 7.3 ~ 7.4, room temperature). During the recording, the brain slices were superfused with standard aCSF constituted (in mM) of 125 NaCl, 2.5 KCl, 1.25 NaH_2PO_4 , 1 MgCl_2 , 2.4 CaCl_2 , 26 NaHCO_3 , 11 glucose, 0.1 picrotoxin, and 2 kynurenic acid, and was saturated with 95% O_2 , and 5 % CO_2 at 32–34°C. Glass pipette (pipette resistance 2.8–4.0 M Ω , King Precision Glass, Claremont, CA) with K^+ -based internal solution (in mM: 140 KMeSO_4 , 5 KCl, 0.05 EGTA, 2 MgCl_2 , 2 Na_2ATP , 0.4 NaGTP , 10 HEPES and 0.05 Alexa Fluor 594 (Invitrogen, Carlsbad, CA), pH 7.3, 290 mOsm) was used throughout the experiment. Whole-cell configuration was made using MultiClamp 700B amplifier (Molecular Devices, Sunnyvale, CA). To verify the functional expression of NpHR in the patched neurons, a 800 ms pulse of green light (532 nm) was delivered at the intensity of 4.6–5.8 mW via a optic fiber that was positioned right above the slice. NpHR expression was confirmed by a significant membrane hyperpolarization under current clamp, or an outward current under voltage clamp upon light stimulation. To examine the effect of light-induced hyperpolarization on neuron excitability, a series of step current injections (100 pA increment up to 1000 pA) was delivered for 1 s in the presence or absence of light (1.5 s, starting 0.5 s prior to step current injection). Throughout the recording, series resistance (10–30 M Ω) was continually monitored on-line with a 20 pA, 300 ms current injection after every current injection step. If the series resistance changed for more than 20%, the cell was excluded. Signal was sampled at 20k Hz and filtered at 10k Hz. Data was acquired in Clampex 10.3 (Molecular Devices, Foster city, CA), and was analyzed off-line in Clampfit 10.3 (Molecular Devices) and IGOR Pro 6.0 (WaveMetrics, Lake Oswego, OR).

Pavlovian Over-Expectation Training and Response Measures—Training began approximately 3 weeks after viral injection and fiber implantation. All procedures and response measures were as described for the recording experiment, except that: 1) training was conducted in behavioral chambers and using Graphic State 3 software provided by Coulbourn Instruments, 2) the initial conditioning was somewhat longer, consisting of 18–22 sessions, due to scheduling issues that did not differ between groups, 3) throughout training, rats were attached to fiberoptic patch cables coupled to a solid state laser (532 nm; Laser Century, Shanghai, China) via an optic commutator (Doric Lenses, Quebec, Canada), and 4) light (532 nm, 10–12 milliWatt) was delivered into OFC bilaterally during each

compound session during the compound cue or the intertrial interval after the compound cue. In some rats (5 NpHR and 5 eYFP), light was delivered only during the 30s compound cue. In other rats (4 NpHR rats and 4 eYFP) light was delivered during the compound cue and also for 30s prior, in order to maximize the light-dependent inhibition of OFC. Whether light was delivered only during the compound cue or also prior to it had no effect on behavioral responses during compound training or the probe test, so the groups were pooled. After retraining, all rats received light for 30s during the intertrial interval after each compound cue, starting 30s after each compound cue.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Intramural Research Program at the National Institute on Drug Abuse. The authors would like to thank Dr Karl Deisseroth and the Gene Therapy Center at the University of North Carolina at Chapel Hill core for providing viral reagents, and Dr Garret Stuber for technical advice on their use. The opinions expressed in this article are the authors' own and do not reflect the view of the NIH/DHHS.

REFERENCES

- Abe H, Lee D. Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron*. 2011; 70:731–741. [PubMed: 21609828]
- Bunsey M, Eichenbaum E. Conservation of hippocampal memory function in rats and humans. *Nature*. 1996; 379:255–257. [PubMed: 8538790]
- Burke KA, Franz TM, Miller DN, Schoenbaum G. The role of the orbitofrontal cortex in the pursuit of happiness and more specific rewards. *Nature*. 2008; 454:340–344. [PubMed: 18563088]
- Camille N, Griffiths CA, Vo K, Fellows LK, Kable JW. Ventromedial frontal lobe damage disrupts value maximization in humans. *Journal of Neuroscience*. 2011; 31:7527–7532. [PubMed: 21593337]
- Chudasama Y, Wright KS, Murray EA. Hippocampal lesions in rhesus monkeys disrupt emotional responses but not reinforcer devaluation effects. *Biological Psychiatry* epub. 2008
- Critchley HD, Rolls ET. Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex. *Journal of Neurophysiology*. 1996; 75:1673–1686. [PubMed: 8727405]
- Gallagher M, McMahan RW, Schoenbaum G. Orbitofrontal cortex and representation of incentive value in associative learning. *Journal of Neuroscience*. 1999; 19:6610–6614. [PubMed: 10414988]
- Gottfried JA, O'Doherty J, Dolan RJ. Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science*. 2003; 301:1104–1107. [PubMed: 12934011]
- Groenewegen HJ, Berendse HW, Wolters JG, Lohman AHM. The anatomical relationship of the prefrontal cortex with the striatopallidal system, the thalamus and the amygdala: evidence for a parallel organization. *Progress in Brain Research*. 1990; 85:95–118. [PubMed: 2094917]
- Hall, G.; Pearce, JM. Changes in stimulus associability during conditioning: implications for theories of acquisition. In: Commons, ML.; Herrnstein, RJ.; Wagner, AR., editors. *Quantitative Analyses of Behavior*. Cambridge, MA: Ballinger; 1982. p. 221-239.
- Haney RZ, Calu DJ, Takahashi YK, Hughes BW, Schoenbaum G. Inactivation of the central but not the basolateral nucleus of the amygdala disrupts learning in response to over-expectation of reward. *Journal of Neuroscience*. 2010; 30:2911–2917. [PubMed: 20181588]
- Hatfield T, Han JS, Conley M, Gallagher M, Holland P. Neurotoxic lesions of basolateral, but not central, amygdala interfere with Pavlovian second-order conditioning and reinforcer devaluation effects. *Journal of Neuroscience*. 1996; 16:5256–5265. [PubMed: 8756453]

- Holland PC, Rescorla RA. The effects of two ways of devaluing the unconditioned stimulus after first and second-order appetitive conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*. 1975; 1:355–363. [PubMed: 1202141]
- Hornak J, O'Doherty J, Bramham J, Rolls ET, Morris RG, Bullock PR, Polkey CE. Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *Journal of Cognitive Neuroscience*. 2004; 16:463–478. [PubMed: 15072681]
- Izquierdo AD, Murray EA. Bilateral orbital prefrontal cortex lesions disrupt reinforcer devaluation effects in rhesus monkeys. *Society for Neuroscience Abstracts*. 2000; 26:978.
- Izquierdo AD, Murray EA. Opposing effects of amygdala and orbital prefrontal cortex lesions on the extinction of instrumental responding in macaque monkeys. *European Journal of Neuroscience*. 2005; 22:2341–2346. [PubMed: 16262672]
- Izquierdo AD, Murray EA. Selective bilateral amygdala lesions in rhesus monkeys fail to disrupt object reversal learning. *Journal of Neuroscience*. 2007; 27:1054–1062. [PubMed: 17267559]
- Izquierdo AD, Suda RK, Murray EA. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *Journal of Neuroscience*. 2004; 24:7540–7548. [PubMed: 15329401]
- Johnson AW, Gallagher M, Holland PC. The basolateral amygdala is critical to the expression of Pavlovian and instrumental outcome-specific reinforcer devaluation effects. *Journal of Neuroscience*. 2009; 29:696–704. [PubMed: 19158296]
- Jones JL, Esber GR, McDannald MA, Gruber AJ, Hernandez G, Mirenzi A, Schoenbaum G. Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science*. 2012; 338:953–956. [PubMed: 23162000]
- Kahnt T, Chang LJ, Park SQ, Heinzle J, Haynes J-D. Connectivity-based parcellation of the human orbitofrontal cortex. *Journal of Neuroscience*. 2012; 32:6240–6250. [PubMed: 22553030]
- Lepelley ME. The role of associative history in models of associative learning: a selective review and a hybrid model. *Quarterly Journal of Experimental Psychology*. 2004; 57:193–243.
- Levy DJ, Glimcher PW. Comparing apples and oranges: Using reward-specific and reward-general subjective value representation in the brain. *Journal of Neuroscience*. 2011; 31:14693–14707. [PubMed: 21994386]
- Levy DJ, Glimcher PW. The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology* epub ahead of print. 2012
- Luk C-H, Wallis JD. Choice coding in frontal cortex during stimulus-guided or action-guided decision-making. *Journal of Neuroscience*. 2013; 33:1864–1871. [PubMed: 23365226]
- Machado CJ, Bachevalier J. The effects of selective amygdala, orbital frontal cortex or hippocampal formation lesions on reward assessment in nonhuman primates. *European Journal of Neuroscience*. 2007; 25:2885–2904. [PubMed: 17561849]
- Malkova L, Gaffan D, Murray EA. 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. *Journal of Neuroscience*. 1997; 17:6011–6020. [PubMed: 9221797]
- McDannald MA, Lucantonio F, Burke KA, Niv Y, Schoenbaum G. Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. *Journal of Neuroscience*. 2011; 31:2700–2705. [PubMed: 21325538]
- Montague PR, Berns GS. Neural economics and the biological substrates of valuation. *Neuron*. 2002; 36:265–284. [PubMed: 12383781]
- Ongur D, Price JL. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cerebral Cortex*. 2000; 10:206–219. [PubMed: 10731217]
- Padoa-Schioppa C. Range-adapting representation of economic value in the orbitofrontal cortex. *Journal of Neuroscience*. 2009; 29:14004–14014. [PubMed: 19890010]
- Padoa-Schioppa C. Neurobiology of economic choice: a goods-based model. *Annual Review of Neuroscience*. 2011; 34:333–359.
- Padoa-Schioppa C, Assad JA. Neurons in orbitofrontal cortex encode economic value. *Nature*. 2006; 441:223–226. [PubMed: 16633341]
- Padoa-Schioppa C, Assad JA. The representation of economic value in the orbitofrontal cortex is invariant for changes in menu. *Nature Neuroscience*. 2008; 11:95–102.

- Pickens CL, Saddoris MP, Gallagher M, Holland PC. Orbitofrontal lesions impair use of cue-outcome associations in a devaluation task. *Behavioral Neuroscience*. 2005; 119:317–322. [PubMed: 15727536]
- Pickens CL, Setlow B, Saddoris MP, Gallagher M, Holland PC, Schoenbaum G. Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *Journal of Neuroscience*. 2003; 23:11078–11084. [PubMed: 14657165]
- Plassmann H, O'Doherty J, Rangel A. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *Journal of Neuroscience*. 2007; 27:9984–9988. [PubMed: 17855612]
- Price JL. Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions. *Annals of the New York Academy of Science*. 2007; 1121:54–71.
- Rescorla RA. Reduction in the effectiveness of reinforcement after prior excitatory conditioning. *Learning & Motivation*. 1970; 1:372–381.
- Rescorla, RA.; Wagner, AR. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black, AH.; Prokasy, WF., editors. *Classical Conditioning II: Current Research and Theory*. New York: Appleton-Century-Crofts; 1972. p. 64-99.
- Rudebeck PH, Murray EA. Amygdala and orbitofrontal cortex lesions differentially influence choices during object reversal learning. *Journal of Neuroscience*. 2008; 28:8338–8343. [PubMed: 18701696]
- Schoenbaum G, Esber G. How do you (estimate you will) like them apples? Integration as a defining trait of orbitofrontal function. *Current Opinion in Neurobiology*. 2010; 20:205–211. [PubMed: 20206497]
- Schoenbaum G, Nugent S, Saddoris MP, Setlow B. Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. *Neuroreport*. 2002; 13:885–890. [PubMed: 11997707]
- Schoenbaum G, Setlow B, Nugent SL, Saddoris MP, Gallagher M. Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. *Learning and Memory*. 2003; 10:129–140. [PubMed: 12663751]
- Sul JH, Kim H, Huh N, Lee D, Jung MW. Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron*. 2010; 66:449–460. [PubMed: 20471357]
- Sutton RS. Learning to predict by the method of temporal difference. *Machine Learning*. 1988; 3:9–44.
- Takahashi Y, Roesch MR, Stalnaker TA, Haney RZ, Calu DJ, Taylor AR, Burke KA, Schoenbaum G. The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron*. 2009; 62:269–280. [PubMed: 19409271]
- Thornton JA, Malkova L, Murray EA. Rhinal cortex ablations fail to disrupt reinforcer devaluation effects in rhesus monkeys (*Macaca mulatta*). *Behavioral Neuroscience*. 1998; 112:1020–1025. [PubMed: 9733208]
- Tobler PN, O'Doherty J, Dolan RJ, Schultz W. Human neural learning depends on reward prediction errors in the blocking paradigm. *Journal of Neurophysiology*. 2006; 95:301–310. [PubMed: 16192329]
- Tremblay L, Schultz W. Relative reward preference in primate orbitofrontal cortex. *Nature*. 1999; 398:704–708. [PubMed: 10227292]
- Wellmann LL, Gale K, Malkova L. GABAA-mediated inhibition of basolateral amygdala blocks reward devaluation in macaques. *Journal of Neuroscience*. 2005; 25:4577–4586. [PubMed: 15872105]
- West EA, DesJardin JT, Gale K, Malkova L. Transient inactivation of orbitofrontal cortex blocks reinforcer devaluation in macaques. *Journal of Neuroscience*. 2011; 31:15128–15135. [PubMed: 22016546]
- Wimmer GE, Shohamy D. Preference by association: How memory mechanisms in the hippocampus bias decisions. *Science*. 2012; 338:270–273. [PubMed: 23066083]

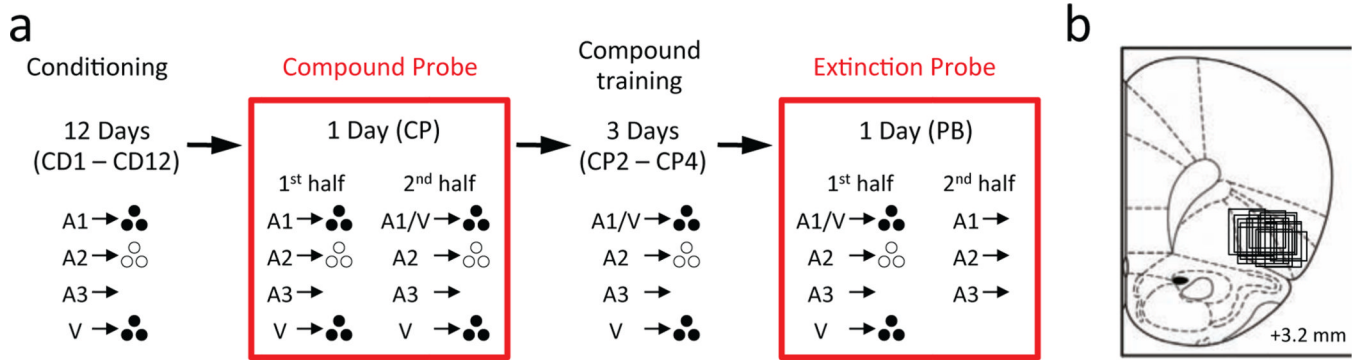


Figure 1. Task design and recording sites

a. Shown is the task design and experimental timeline. A1, A2 and A3 are auditory cues (tone, white noise and clicker, counterbalanced). V is a visual cue (a cue light). Two differently flavored sucrose pellets were used as reward (banana- or grape-flavored sucrose pellets, represented by solid or empty circles, counterbalanced). Training began with 12 conditioning sessions (CD1 – CD12) in which each cue was presented 8 times. A1 and V cues were paired with the same reward (3 pellets), and A2 were paired with the other reward (3 pellets). A3 was paired with no reward. After completion of the last conditioning session, rats underwent a single compound probe session (CP1) followed by 3 compound training sessions (CP2 – CP4). During the 1st half of the compound probe session (CP 1/2), rats continued to receive simple conditioning. During the 2nd half (CP 2/2), rats began compound training in which A1 and V were presented together as a compound (A1/V), followed by delivery of the same reward (3 pellets). A2, A3 and V continued to be presented as in simple conditioning. During the compound training sessions (CP2 – CP4), rats received presentations of A1/V, A2, A3 and V. After the completion of last compound training session, rats underwent a single extinction probe session (PB). The 1st half of the session (PB 1/2) consisted of further compound training. During the 2nd half of the session (PB 2/2), rats received eight non-reinforced presentations of A1, A2 and A3 with the order mixed and counterbalanced. **b.** Location of recording sites in OFC. Boxes indicate approximate location of recording sites in each rat, taking into account any vertical distance traveled during training and the approximate lateral spread of the electrode bundle.

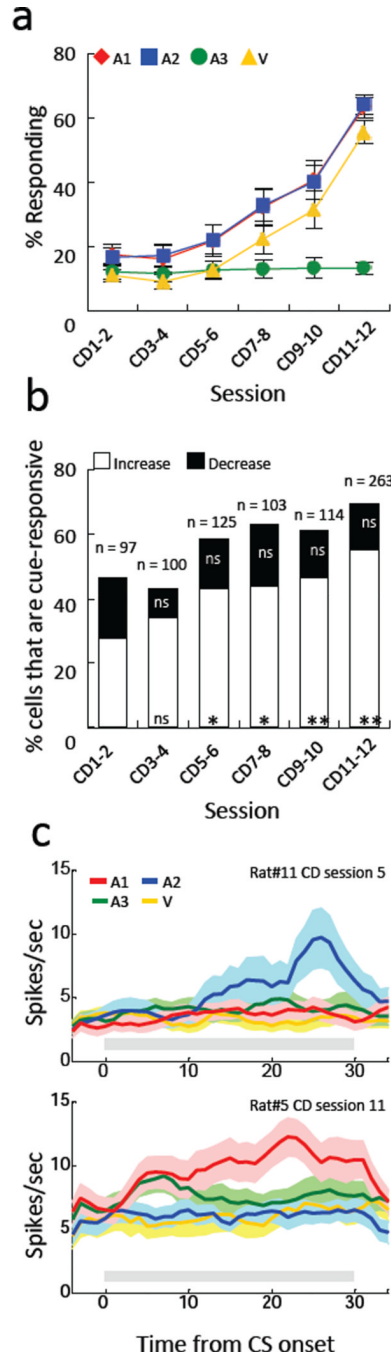


Figure 2. Conditioned responding and cue-evoked activity increased during simple conditioning
a. Plot illustrating increase in conditioned responding as a percentage of time in the food cup during each of the 4 cues across sessions. Red diamond; A1, blue square; A2, green circle; A3, yellow triangle; V. **b.** Proportions of neurons that were significantly responsive to any of the 4 cues, shown for each pair of sessions and separated by those that increased (white) or decreased (black) firing rate compared to baseline. The proportion of neurons that increased firing grew significantly across conditioning (chi-square test compared to proportion in the first pair of sessions), whereas the proportion of neurons that decreased firing did not change. ** $p < 0.01$, * $p < 0.05$. **c.** Examples of single units showing cue-

evoked responses. Top and bottom units were recorded from Rat#11 in conditioning day 5 and from Rat#5 in conditioning day 11, respectively. Activity shown is synchronized to the onset of the 30 s cues. Red, blue, green and yellow lines indicate A1, A2, A3 and V, respectively. Gray bars indicate a period of cue presentation. Bin size; 1 sec.

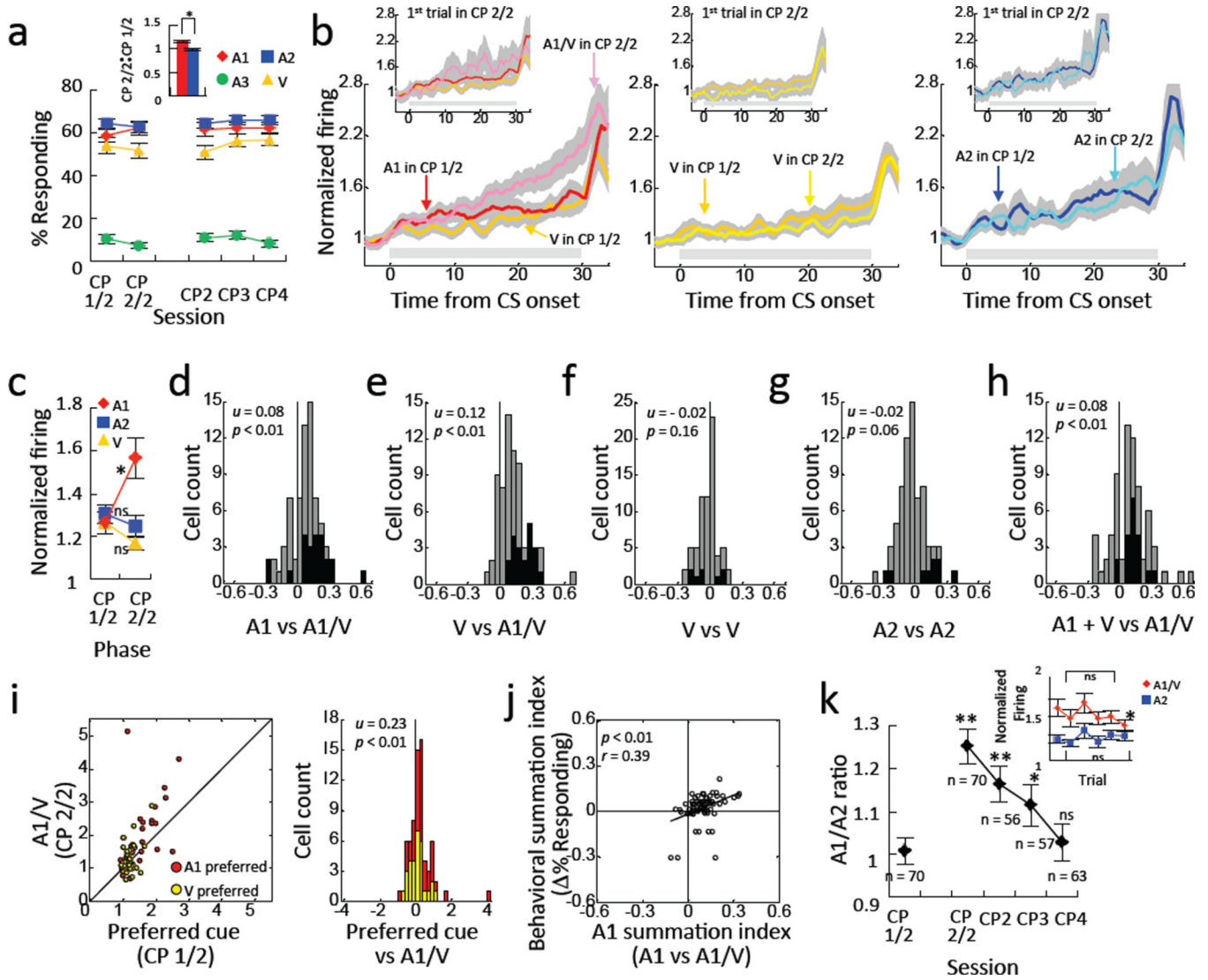


Figure 3. Conditioned responding and cue-evoked activity summates at the start of compound training

a. Conditioned responding as a percentage of time in the food cup during each of the 4 cues during the compound probe (CP) and 3 days of compound training (CP2 – CP4). Red diamonds indicate A1 in CP 1/2 phase, and A1/V in CP 2/2 and CP2 – CP4 phases. Blue squares, green circles and yellow triangles indicate A2, A3 and V, respectively. Red and blue bars in the inset indicate the change in responding to A1 (red) and A2 (blue) from the 1st half to the 2nd half of CP. * $p < 0.05$. Error bars = S.E.M. **b.** Population responses of all 70 cue-responsive neurons, with firing normalized by neuron, to A1 (left), V (middle) and A2 (right) during 28 compound probe sessions. Dark and light red indicate population response to A1 in the 1st half of the session and population response to A1/V in the 2nd half, respectively. Dark and light yellow indicate population response to V in the 1st half and 2nd half of the session, respectively. Dark and light blue indicate population responses to A2 in the 1st half and 2nd half of the session, respectively. Small insets in each panel indicate population response to each cue in the 1st half of the session and population response on the 1st trial in the 2nd half of the session. Gray shadings indicate S.E.M. Gray bars indicate a period of cue presentation. **c.** Average normalized firing to A1 (red), A2 (blue), and V

(yellow) in the 1st and 2nd half of the compound probe session. Average normalized activity was calculated by dividing average firing during the last 20 sec by average firing during the last 20 sec of pre CS period **d - g**. Distributions of summation index scores for firing to A1 (**d**), V (**e and f**) and A2 (**g**) in the compound probe. Each summation index compares firing on the first trial of the second half of the compound probe (CP 2/2) against firing in the first half of the compound probe (CP 1/2), using the following formula: $(2^{\text{nd}} \text{FR} - 1^{\text{st}} \text{FR}) / (2^{\text{nd}} \text{FR} + 1^{\text{st}} \text{FR})$, where FR represents average normalized firing for each condition. **h**. Distribution of compound index in the compound probe session. The compound index compares firing to the compound cue (A1/V) in the first trial of the second half of the session against the sum of firing to A1 and V in the first half of the session, using the following formula: $(2^{\text{nd}} \text{FR A1/V} - (1^{\text{st}} \text{FR A1} + 1^{\text{st}} \text{FR V})) / (2^{\text{nd}} \text{FR A1/V} + (1^{\text{st}} \text{FR A1} + 1^{\text{st}} \text{FR V}))$, where FR represents average normalized firing for each condition. Black bars represent neurons in which the difference in firing was statistically significant. The numbers in each panel indicate results of a Wilcoxon signed-rank test (p) on the distribution and the average summation index (u). **i**. Scatter plot in left represents relationship between average normalized firing of each neuron to preferred cue in the 1st half and average normalized firing to A1/V on the 1st trial in the 2nd half of the session. Distribution plot in right represents summation index calculated by average normalized firing to preferred cue in the 1st half and average normalized firing to A1/V on the 1st trial in the 2nd half of the session. **j**. Correlation between neural summation index scores and behavioral summation index scores during the compound probe session. The behavioral summation index compares conditioned responding to A1/V in the first trial of the second half of the session against that to A1 during the first half of the session, using the following formula: $(2^{\text{nd}} \text{CR A1/V1} - 1^{\text{st}} \text{CR A1}) / (2^{\text{nd}} \text{CR A1/V1} + 1^{\text{st}} \text{CR A1})$, where CR represents average percent of time spent in the food cup during each condition. **k**. Line plot indicates the ratio between normalized firing to A1/V and A2 during each compound training session (CP – CP4). N's indicate number of cue-responsive neurons in each session. A1/A2 ratio increased significantly in the compound phase of the probe, and then gradually decreased (ANOVA, **p < 0.01, *p < 0.05). Line plot in inset indicates normalized firing to A1/V and A2 across 6 trials in the 2nd half of the compound probe session, with red diamonds for A1 and blue squares for A2. Error bars = S.E.M. See also Figures S1, S2, S4.

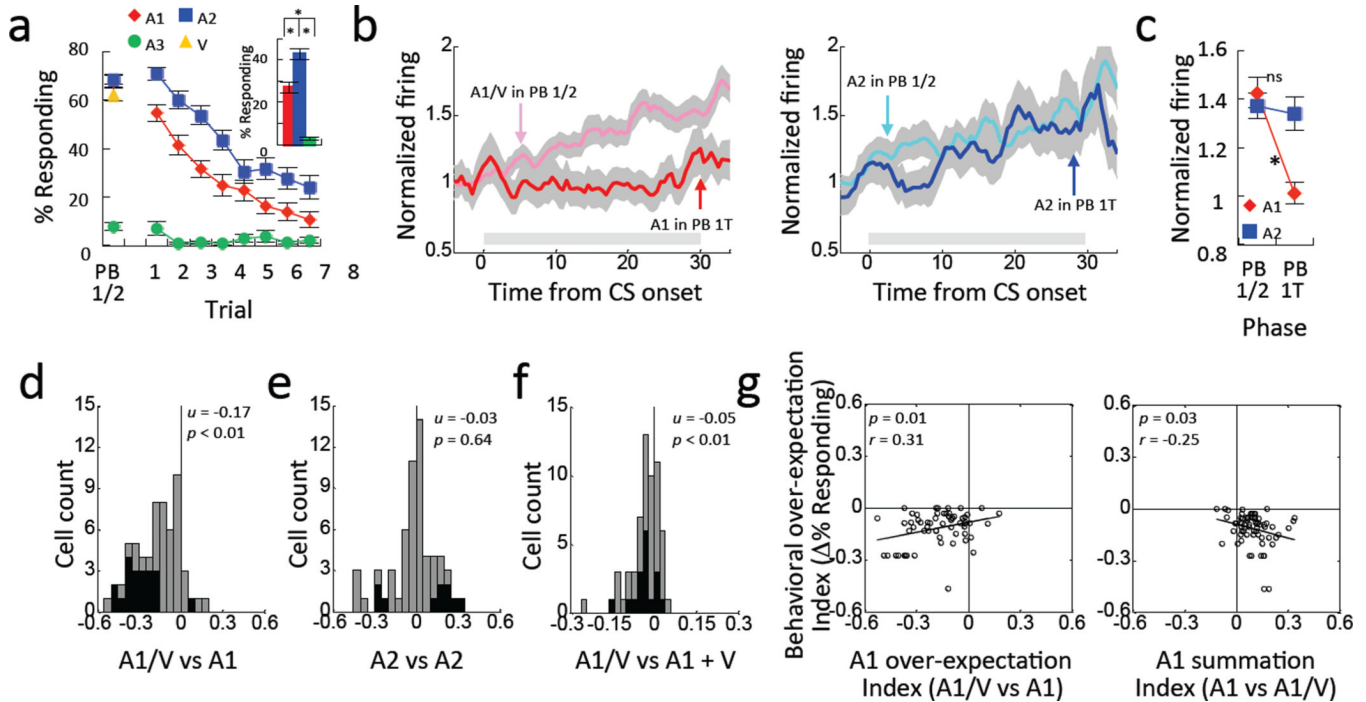


Figure 4. Conditioned responding and cue-evoked activity spontaneously declines at the start of extinction training

a. Conditioned responding as a percentage of time in the food cup during each of the 4 cues during the extinction probe (PB). Bar graph shows average responding during extinction trials only. Red indicates A1/V in PB 1/2, and A1 in the line plot and bar graph. Blue, green and yellow indicate A2, A3 and V, respectively. * $p < 0.01$. Error bars = S.E.M. **b.** Population responses of all 61 cue-responsive neurons, with firing normalized by neuron, to A1 (left) and A2 (right) during 28 extinction probe sessions. Light and dark red indicate population response to A1/V in the 1st half of the session and population response to A1 on the 1st trial in the 2nd half, respectively. Light and dark blue indicate population responses to A2 in the 1st half and population response on the 1st trial in the 2nd half of the session, respectively. Gray shadings indicate S.E.M. Gray bars indicate a period of cue presentation. **c.** Average normalized firing rate to A1 (red) and A2 (blue) in the extinction probe session. Average normalized activity was calculated by dividing average firing during the last 20 sec by average firing during last 20 sec of pre CS period. **d and e.** Distribution of over-expectation index scores for firing to A1 (**d**) and A2 (**e**) in the extinction probe. Each over-expectation index compares firing on the first trial of the second half of the probe (PB 2/2) against firing in the first half of the probe (PB 1/2), using the following formula: $(2^{\text{nd}} \text{FR} - 1^{\text{st}} \text{FR}) / (2^{\text{nd}} \text{FR} + 1^{\text{st}} \text{FR})$, where FR represents average normalized firing for each condition. **f.** Distribution of compound index in the extinction probe session. The compound index compares firing to the compound cue (A1/V) in the first half of the session against the sum of firing to V in the first half of the session and A1 in the first trial of the second half of the session, using the following formula: $((2^{\text{nd}} \text{FR A1} + 1^{\text{st}} \text{FR V}) - 1^{\text{st}} \text{FR A1/V}) / ((2^{\text{nd}} \text{FR A1} + 1^{\text{st}} \text{FR V}) + 1^{\text{st}} \text{FR A1/V})$, where FR represents average normalized firing for each condition. Black bars represent neurons in which the difference in firing was statistically significant. The numbers in each panel indicate results of a Wilcoxon signed-rank test (p) on the distribution and the average over-expectation index (u). **g.** Correlation between behavioral over-expectation and neural over-expectation, and between behavioral over-expectation and neural summation. The neural summation index was A1 index, computed as in Fig. 3 (i.e. from the compound probe session). The neural over-expectation index was

computed as in Fig. 4d. The behavioral over-expectation index compares conditioned responding to A1 in the first trial of the second half of the session against that to A1/V1 during the first half of the session, using the following formula: $(2^{\text{nd}} \text{ CR A1} - 1^{\text{st}} \text{ CR A1/V1}) / (2^{\text{nd}} \text{ CR A1} + 1^{\text{st}} \text{ CR A1/V1})$, where CR represents average percent of time spent in the food cup during each condition. See also Figures S3 and S4.

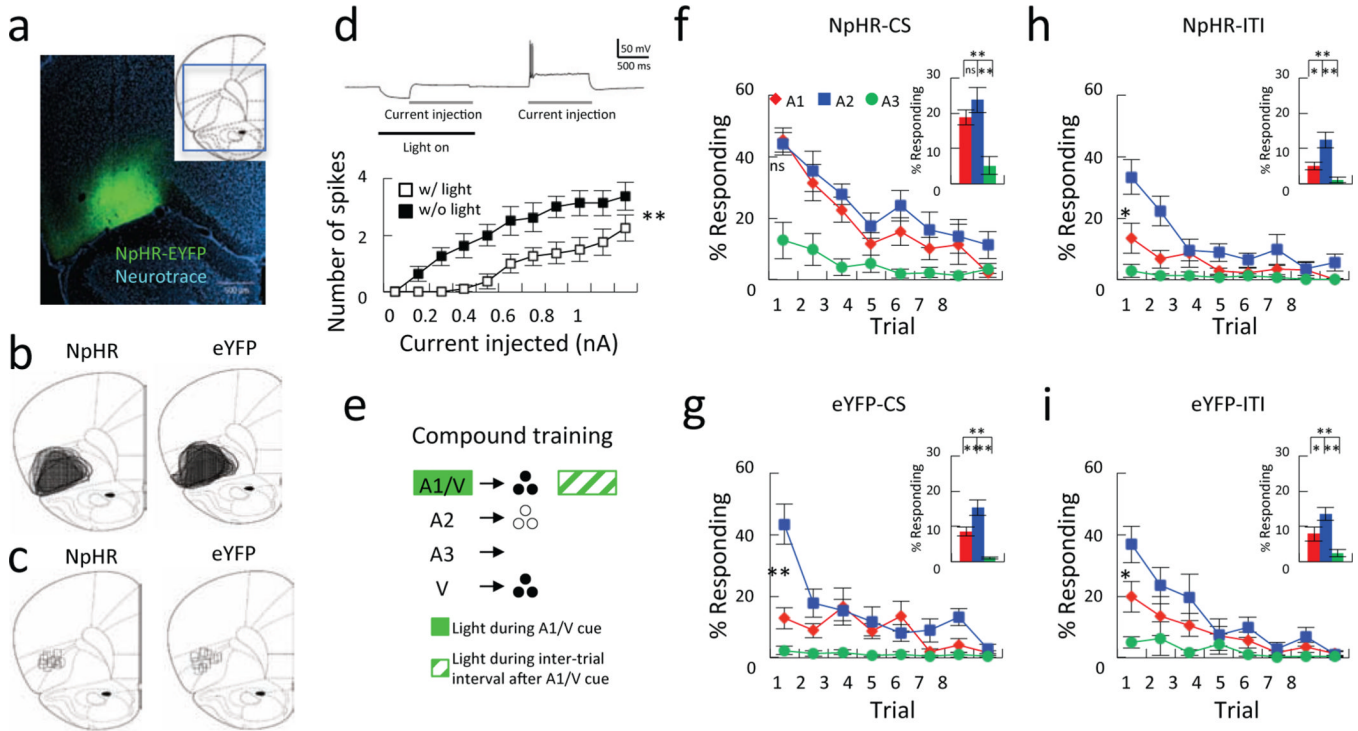


Figure 5. Optogenetic inhibition of OFC neurons prevents spontaneous decline in conditioned responding at the start of extinction training

a. Representative coronal brain slice showing expression of NpHR-eYFP (green) after virus injection into OFC. Blue, fluorescent Nissl staining with NeuroTracer. **b.** Traces showing the expression of NpHR-eYFP (left) and eYFP (right). **c.** Locations of fiber tips in NpHR-eYFP (left) and eYFP (right) groups. **d.** NpHR transgene reduced OFC neural excitability. The top panel represents an example trace of NpHR-eYFP-expressing OFC neuron firing pattern in the presence and absence of light. Gray bars; current injection period (300 pA in this case), black bar, light on period. The line plot at the bottom represents neuron excitability comparison of NpHR-eYFP-expressing OFC neurons ($n = 8$) in the presence (open square) or absence of light (solid square). NpHR-eYFP-expressing OFC neurons generate fewer evoked spikes during light-on conditions compared to light-off conditions ($F(1,14) = 8.94, p < 0.01$). **e.** Optical stimulation was delivered during presentation of A1/V (NpHR-CS and eYFP-CS groups) or during inter-trial interval 30 s after A1/V presentation (NpHR-ITI and eYFP-ITI group). **f-i.** Conditioned responding as a percentage of time in the food cup during each of 3 cues during the extinction probe in NpHR-CS (**f**), eYFP-CS (**g**), NpHR-ITI (**h**) and eYFP-ITI (**i**) groups. The line plots show responding across 8 trials, and bar graphs show average responding of 8 trials. Red, blue and yellow indicate A1, A2 and A3, respectively. * $p < 0.05$. ** $p < 0.01$. Error bars = S.E.M. See also Figure S5.

Table 1

Number of cells recorded in each training phase (see also Table S1)

	session	Learned		
		all	increase	decrease
Conditioning	CD1-2	97	27	18
	CD3-4	100	34	9
	CD5-6	125	54	19
	CD7-8	103	45	20
	CD9-10	114	53	17
	CD11-12	263	145	38
Compound probe	CP	130	70	20
Compound training	CP2	116	55	13
	CP3	121	57	19
	CP4	122	63	21
Extinction probe	PB	140	61	20