nature neuroscience

Dopamine transients are sufficient and necessary for acquisition of model-based associations

Melissa J Sharpe[1,2], Chun Yun Chang[1], Melissa A Liu[1], Hannah M Batchelor[1], Lauren E Mueller[1], Joshua L Jones[1], Yael Niv[2] & Geoffrey Schoenbaum[1,3,4]

Associative learning is driven by prediction errors. Dopamine transients correlate with these errors, which current interpretations limit to endowing cues with a scalar quantity reflecting the value of future rewards. We tested whether dopamine might act more broadly to support learning of an associative model of the environment. Using sensory preconditioning, we show that prediction errors underlying stimulus–stimulus learning can be blocked behaviorally and reinstated by optogenetically activating dopamine neurons. We further show that suppressing the firing of these neurons across the transition prevents normal stimulus–stimulus learning. These results establish that the acquisition of model-based information about transitions between nonrewarding events is also driven by prediction errors and that, contrary to existing canon, dopamine transients are both sufficient and necessary to support this type of learning. Our findings open new possibilities for how these biological signals might support associative learning in the mammalian brain in these and other contexts.

The discovery that midbrain dopamine neurons emit a teaching signal when an unexpected reward or reward-predicting cue occurs has transformed how we conceptualize dopamine function[1]. The response to unpredicted rewards, initially large, wanes as the subject comes to anticipate the rewarding event, transferring instead to antecedent stimuli that reliably predict future reward. This finding has been influential because transient changes in dopamine are so like the prediction errors proposed as driving learning in reinforcement learning models[2–5]. Indeed, the dopaminergic prediction error has become almost synonymous with the reward prediction error defined in these models. However, these errors are thought to support only a relatively limited form of learning, in which predictive cues are endowed with a scalar quantity that reflects the rewarding value of future events at the time of learning. This cached or model-free value does not capture any specific information about the identity of those future events, even in more expansive recent proposals that incorporate elements of reward structure[4]. As a result, the behaviors supported by these values are relatively inflexible, since they cannot reflect information about the predicted events other than their general value at the time of learning.

Yet much behavior reflects specific information about predicted events, rewarding or otherwise[6]. Such behavior reveals the existence of a rich and navigable associative representation or model of the structure of the environment. For instance, when walking into your favorite neighborhood restaurant, you expect not only a good meal but also one that consists of sushi, not pasta. Because this prediction contains specific information beyond value, it supports flexible and adaptive behavior[7–11]. You might love Japanese and Italian food equally, but if you become pregnant and are instructed to avoid raw fish, you can adjust your choice of restaurant without additional direct experience. Can dopaminergic prediction errors support the formation of these model-based associations, or do they only support learning of model-free associations that contain scalar values? Although optogenetic studies have confirmed that dopamine transients can function as errors to support associative learning[12–18], this critical question remains unaddressed, since in each of these experiments the resultant behavior could be accounted for by model-free learning mechanisms.

Here we directly address this question using sensory preconditioning[19–22] in rats. Sensory preconditioning entails presenting subjects with two neutral cues, for example, C and X, in close succession, such that a predictive relationship C→X can form between them. Notably, in this preconditioning phase, no rewards are delivered, and consequently no new behavioral responses or scalar values are learned. However, the contents of what is learned in preconditioning can be revealed if the second cue is subsequently paired with an unconditional stimulus, for instance, a reward (i.e., X→US). Subsequently, both C and X will elicit robust conditioned responses. Since C was never paired with reward, the response to C demonstrates the existence of an associative link between C and X. The use of this C→X association to support responding for the reward is a classic example of model-based behavior.

We used this behavioral approach in two experiments. The first was designed to test whether a dopamine transient is sufficient to support the formation of the associative representations underlying model-based behavior. For this, we combined sensory preconditioning with blocking[23], a procedure developed to show that associative learning

[1]NIDA Intramural Research Program, Baltimore, Maryland, USA. [2]Department of Psychology and Neuroscience Institute, Princeton University, Princeton, New Jersey, USA. [3]Departments of Anatomy and of Neurobiology and Psychiatry, University of Maryland School of Medicine, Baltimore, Maryland, USA. [4]Solomon H. Snyder Department of Neuroscience, The Johns Hopkins University, Baltimore, Maryland, USA. Correspondence or requests for material should be addressed to M.J.S. (melissa.sharpe@nih.gov) or G.S. (geoffrey.schoenbaum@nih.gov).

Received 3 June 2016; accepted 28 February 2017; published online 3 April 2017; corrected online 10 April 2017 (details online); corrected after print 5 May 2017; corrected after print 17 July 2018; doi:10.1038/nn.4538

depends on the presence of a prediction error. While blocking has previously been shown only in the context of learning about a valuable reward[23], we hypothesized that learning associations that do not involve reward or value should also be regulated by an error mechanism. To test this, we applied the same logic used in reward blocking to reduce acquisition of the C→X relationship during preconditioning. In particular, we first paired a different cue, A, with X (A→X). Then, during preconditioning, A was presented in compound with C, followed by X (AC→X). Because A already predicts X, if learning of the stimulus–stimulus association was driven by errors in prediction, the presence of A should diminish or block the formation of any association between C and X. Indeed, we observed such blocking in pilot testing (**Supplementary Fig. 1**), confirming that initial learning in sensory preconditioning was driven by prediction errors (termed 'state prediction errors' in current computational models[8]), even though there was no reward or value present.

Against this background, we attempted to reinstate learning of the C→X association by briefly activating the dopamine neurons at the start of the X cue in the AC→X trials, using parameters designed to evoke firing similar to that sometimes observed for rewards[13,24–28] or even neutral cues[24,29,30]. We reasoned that if dopamine transients can support learning of associations between the neural representations of events in the environment, as opposed to being restricted to the addition or subtraction of value, then this manipulation should restore normal sensory preconditioning of C. In a second experiment, we tested the necessity of dopamine for this learning process by suppressing the dopamine neurons across the transition between the cues during a standard sensory preconditioning task. The results of the two experiments show that dopamine transients were both sufficient and likely necessary to support the acquisition of the associative structures underlying model-based behavior.

RESULTS

Dopamine transients are sufficient for the formation of model-based associations

Prior to training, all rats underwent surgery to infuse virus and implant fiber optics targeting the ventral tegmental area (VTA; **Fig. 1**). We infused AAV5-EF1α-DIO-ChR2-eYFP (channelrhodopsin-2 (ChR2) experimental group; n = 18) or AAV5-EF1α-DIO-eYFP (eYFP control group; n = 19) into the VTA of rats expressing Cre recombinase under the control of the tyrosine hydroxylase (TH) promoter[31]. After surgery and recovery, rats were food-restricted until their body weight reached 85% of baseline and training commenced. Training began with 2 d of preconditioning. On the first day, the rats received a total of 16 pairings of two 10-s neutral cues (A→X). On the second day, the rats continued to receive pairings of the same two neutral cues (A→X; 8 trials). In addition, on other trials, the first cue was presented together with a second, novel neutral cue (either AC→X or AD→X; 8 trials each). On AC trials, blue light (473 nm, 20 Hz, 16–18 mW output; Shanghai Laser & Optics Century Co., Ltd) was delivered for 1 s at the start of X to activate VTA dopamine neurons. As a temporal control for nonspecific effects, the same light pattern was delivered on AD trials in the intertrial interval, 120–180 s after termination of X. Finally, to verify that sensory preconditioning could be obtained with compound cues, the rats also received pairings of two novel 10-s cues with X (EF→X; 8 trials). As expected, since training did not involve pairing with rewards, rats in both groups (ChR2 and eYFP controls) exhibited little response at the food cup during any of the cues on either day of training (**Fig. 2a**); a two-factor ANOVA on food cup entries during cue presentations (cue × group) revealed no main effect ($F_{4,140}$ = 1.52, P = 0.2) nor any interaction with group ($F_{4,140}$ = 0.276, P = 0.893).

Following preconditioning, the rats began conditioning, which continued for 4 d. Each day, the rats received 24 trials in which X was presented followed by delivery of two 45-mg sucrose pellets (X→2US). Rats in both groups acquired a conditioned response to X. This was evident as an increase in the number of times they entered the food cup to look for sucrose pellets during X, across days of conditioning (**Fig. 2b**). Notably, acquisition of this conditioned response was similar in the two groups; a two-factor ANOVA (group × day) revealed a main effect of day ($F_{3,105}$ = 39.71, P < 0.0001) but neither main effect ($F_{1,35}$ = 0.553, P = 0.46) nor any interaction with group ($F_{3,105}$ = 0.13, P = 0.94). Thus, the introduction of a dopamine transient at the start of X did not produce any lasting effect on subsequent processing of or learning about X.

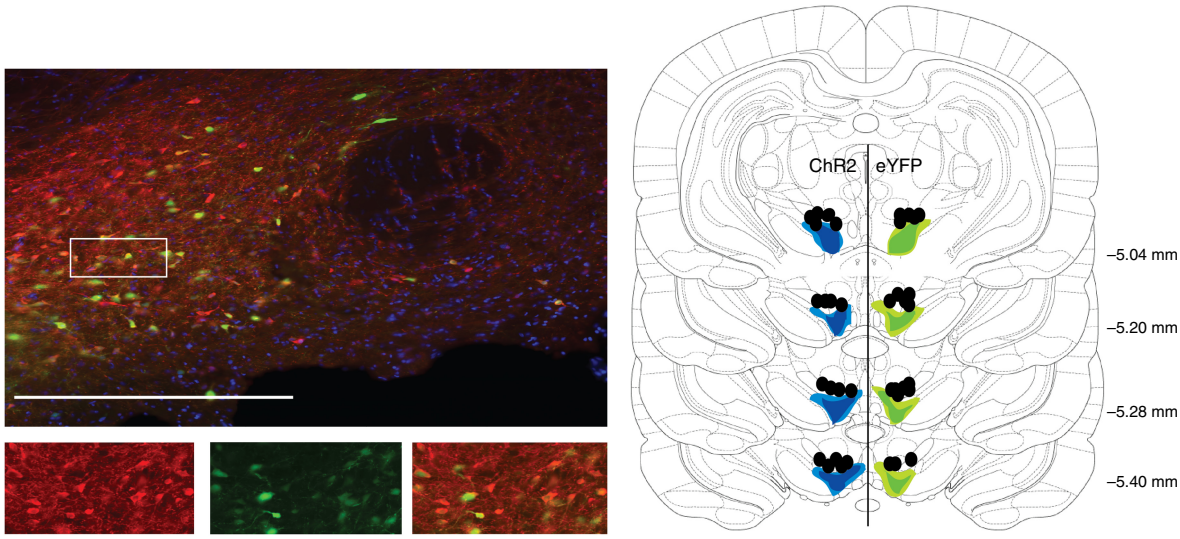

Figure 1 Immunohistochemical verification of Cre-dependent ChR2 and eYFP expression in TH+ neurons and fiber placements in the VTA. Left: 90% of YFP-expressing neurons (green) also expressed TH (red). Bottom left: expansions of the region boxed at top. Right: Unilateral representation of the bilateral fiber placements and virus expression in each group. Fiber implants (black circles) were localized in the vicinity of eYFP (green) and ChR2 (blue) expression in VTA. Light shading represents the maximal and dark shading indicates the minimal spread of expression at each level. Scale bar, = 20 μm.

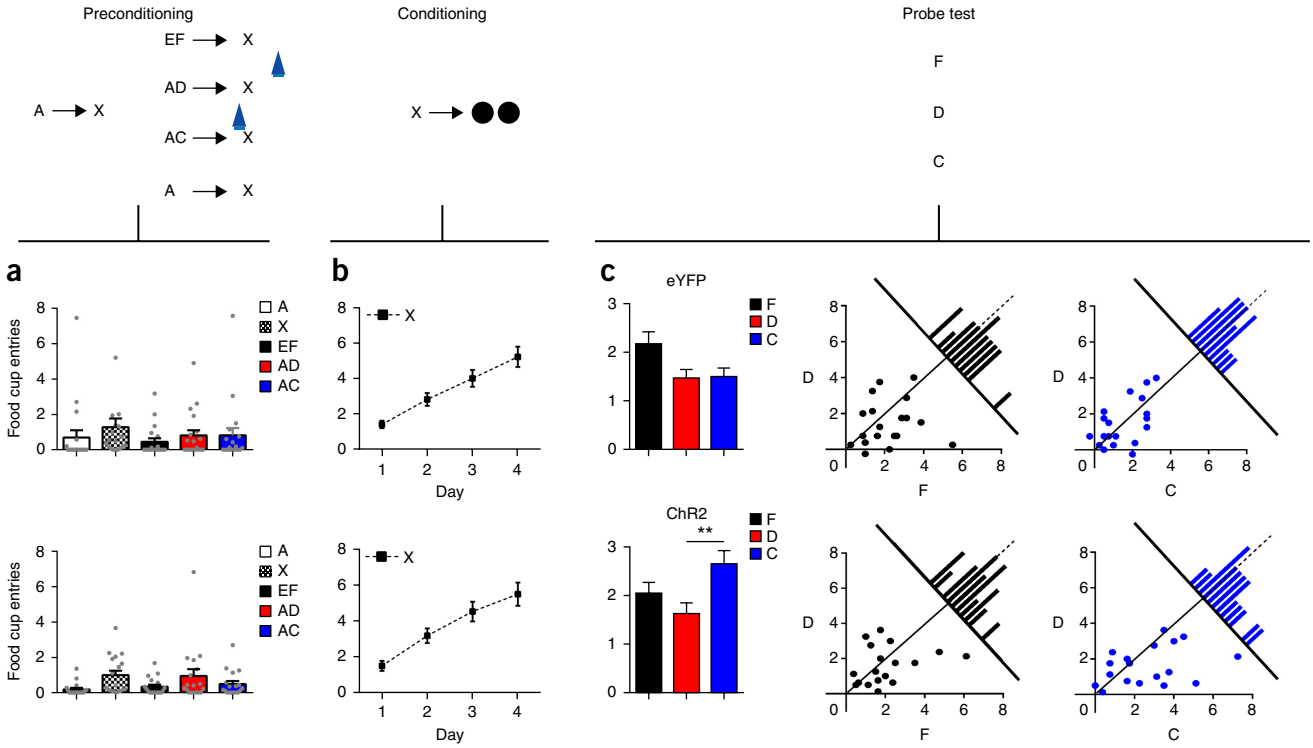

Figure 2 Brief optogenetic activation of VTA dopamine neurons strengthens associations between cues. Top: VTA dopamine neurons were activated by light delivery (blue symbol) at the beginning of presentations of X when preceded by audiovisual compound AC and during the intertrial interval on AD trials. Double dots represent food pellets. (**a**–**c**) Plots show number of food cup entries occurring during cue presentation across all phases of the blocking of sensory preconditioning task: (**a**) preconditioning, (**b**) conditioning and (**c**) the probe test. Probe test data are represented as the mean level of entries (left) or as individual rats' responses to F and D (middle), or C and D (right). In each panel, top graph shows data from the eYFP control group ($n = 19$); bottom graph shows data from the experimental ChR2 group ($n = 18$). To the extent that responding to F and C are equal to D in scatterplots represented in **c**, points should congregate around the diagonal. Histograms along the diagonal reveal the frequency (subject counts) of difference scores in responding to the cues that fall within a particular range. A two-factor ANOVA on food cup entries during cue presentations (cue × group) revealed no main effect ($F_{4,140} = 1.52$, $P = 0.2$) nor any interaction with group ($F_{4,140} = 0.276$, $P = 0.893$). A two-factor ANOVA (group × day) on responding during conditioning (**b**) revealed a main effect of day ($F_{3,105} = 39.71$, $P < 0.0001$) but neither main effect ($F_{1,35} = 0.553$, $P = 0.46$) nor any interaction with group ($F_{3,105} = 0.13$, $P = 0.94$). A two-factor ANOVA (cue × group) on responding during presentation of cues F and D revealed a main effect of cue ($F_{1,35} = 4.372$, $P = 0.044$) but no main effect ($F_{1,35} = 0.001$, $P = 0.982$) or interaction with group ($F_{1,35} = 0.287$, $P = 0.595$). A two-factor ANOVA (cue × group) on responding to C and D revealed a main effect of cue ($F_{1,35} = 4.599$, $P = 0.039$) and a significant interaction with group ($F_{1,35} = 4.154$, $P = 0.049$). This interaction was due to a significant difference between responding to C and D in the ChR2 group ($F_{1,35} = 8.52$, $P = 0.006$) but not in the eYFP group ($F_{1,35} = 0.006$, $P = 0.940$). $^{**}P < 0.01$. Error bars, = s.e.m. Please see Online Methods for comment on response measures and **Supplementary Figure 4** for further details on responding during individual sessions in preconditioning.

Finally, the rats received a probe test in which each of the critical test cues (C, D, F) were presented 4 times each, in an interleaved and counterbalanced order, alone and without reward. This probe test was designed to assess whether these preconditioned cues had acquired the ability to predict sucrose pellet delivery. As expected from studies of normal sensory preconditioning, rats in both groups demonstrated frequent responses to F, suggesting that, despite the use of a compound cue, they learned that F predicted X and used that relationship in the probe test to infer that F predicted sucrose pellets (**Fig. 2c**). Rats in both ChR2 and eYFP groups also demonstrated infrequent responses to D (as in our pilot study; **Supplementary Fig. 1**), indicating that the presence of A and its ability to predict X had blocked D from becoming associated with X (**Fig. 2c**). Notably, this occurred despite transient activation of the VTA dopamine neurons during the intertrial interval following AD trials. A two-factor ANOVA (cue × group) on responding during presentation of cues F and D revealed a main effect of cue ($F_{1,35} = 4.372$, $P = 0.044$) but no main effect ($F_{1,35} = 0.001$, $P = 0.982$) or interaction with group ($F_{1,35} = 0.287$, $P = 0.595$). Thus, both groups exhibited identical blocking of sensory preconditioning, as indexed by a significant difference between F and D.

When delivered at the start of X on the AC trials, however, transient activation of the dopamine neurons unblocked learning, so that responses to C were more common than responses to D in the ChR2 group but not in the eYFP controls (**Fig. 2c**). A two-factor ANOVA (cue × group) on responding to C and D revealed a main effect of cue ($F_{1,35} = 4.599$, $P = 0.039$) and a significant interaction with group ($F_{1,35} = 4.154$, $P = 0.049$). This interaction was due to a significant difference between responding to C and D in the ChR2 group ($F_{1,35} = 8.52$, $P = 0.006$) but not in the eYFP group ($F_{1,35} = 0.006$, $P = 0.940$). In addition, responding to D did not differ between groups ($F_{1,35} = 0.153$, $P = 0.698$), whereas responding to C was significantly more common in the ChR2 rats than in the eYFP controls ($F_{1,35} = 5.277$, $P = 0.028$). Thus, transient activation of the VTA dopamine neurons at the start of X on AC trials reversed the blocking effect, as indexed by the significant increase in responding to C only in the ChR2 rats.

But is the learning supported by transient activation of dopamine neurons the same as what is normally learned during sensory preconditioning? That is, did the rats in the ChR2 group respond to C because it evoked a prediction that sucrose pellets would be delivered

to the food cup? To test this, we assessed the effect of devaluating the sucrose pellets on responding to C in a subset of the ChR2 rats that had been trained on the blocking of sensory preconditioning task. We divided the rats into two groups with equal responding to C ($F_{1,8}$ = 0.028, P = 0.871). After reminder training (X→2US; 12 trials; $F_{1,8}$ = 2.802, P = 0.133), rats in each group received sucrose pellets and lithium chloride injections to induce nausea (LiCl; 10 ml/kg 0.15 M) on three successive days. For one group (devalued group; n = 5), sucrose pellets were presented immediately before induction of illness. For the other group (nondevalued group; n = 5), sucrose pellets were presented ~6 h after the induction of illness. Two days after the final LiCl injection, the rats received a probe test in which C was presented as before, alone and without reward. In this test, rats in the devalued group responded significantly less to C than rats in the nondevalued group (12 trials; **Fig. 3a**; $F_{1,8}$ = 6.777, P = 0.031). Devalued rats also consumed fewer sucrose pellets during a subsequent consumption test (**Fig. 3b**; $F_{1,8}$ = 13.425, P = 0.006), confirming a reduced desire for the pellets. The effect of devaluation on responding to C in the ChR2 rats was the same as what has been previously reported for a normally preconditioned cue[20,22], suggesting that activating dopamine neurons transiently at the start of X on the AC trials restored normal acquisition of the predictive relationship between C and X, effectively leading to anticipation of sucrose pellets upon presentation of C.

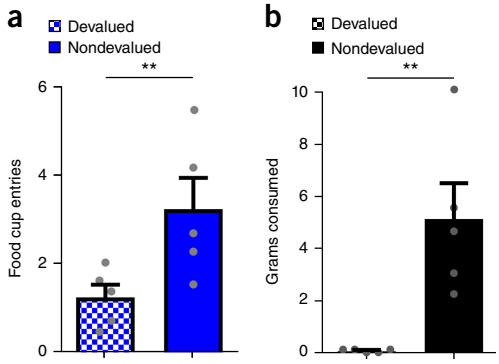

Figure 3 Conditioned responding resulting from learning, supported by brief activation of VTA dopamine neurons, is sensitive to devaluation of the predicted reward. (**a**) Food cup entries during presentation of C in the probe test following illness-induced devaluation of the predicted sucrose pellet reward. (**b**) Grams of sucrose pellets consumed in subsequent consumption test. A one-way ANOVA revealed a significant difference between responding to cue C ($F_{1,8}$ = 6.777, P = 0.031) and consumption of the sucrose food pellets ($F_{1,8}$ = 13.425, P = 0.006) in the devalued group relative to the nondevalued group. $^{**}P < 0.05$. Error bars, = s.e.m.

Dopamine transients are necessary for the formation of model-based associations

The above shows that transient activation of VTA dopamine neurons was sufficient to drive the formation of an association between two sensory representations. This association can then support model-based behavior, with rats responding to C as if it predicts food through its association with X. This is important because we know that dopamine neurons exhibit transient increases in firing in the context of unexpected reward. The results described above suggest that the dopamine transient at the time of an unexpected reward should result in an association between the cue and the sensory features of the reward that could later be used to support devaluation-sensitive behavior or even economic decision-making.

Of course, the finding above does not address whether transient activation of these neurons normally contributed to sensory preconditioning or stimulus–stimulus learning in the absence of reward. Although the timing and duration of the optogenetic activation we used was designed based on the dopamine responses to reward[13,24–28,32–34], its duration was longer than the peak response typically observed in unit studies. Further, while dopamine neurons have been shown to fire in response to neutral cues[24,29,30], such activity is weaker than that in response to unexpected rewards. Therefore it is not clear how similar the signal that our stimulation generated was to that caused by unexpected sensory input in the absence of reward. Further, idiosyncrasies governing viral expression and light penetration dictate that no pattern of optogenetic activation is likely to reproduce what happens normally, either here or in other similar work.

To address whether dopamine transients are necessary for model-based learning in the absence of reward, we ontogenetically suppressed activity in VTA dopamine neurons across the critical transition between the sensory cues in the first phase of a standard sensory preconditioning task. Rats were presented with two pairs of neutral cues in close succession (i.e., A→X; B→Y). Dopamine neurons were prevented from firing during the transition between B and Y but were free to fire between A and X. Subsequently, X and Y were paired directly with reward (X→US; Y→US). We reasoned that if dopamine transients were necessary for learning associations between nonrewarding events in the environment, then suppressing the firing of dopamine neurons across this transition would disrupt normal sensory preconditioning of B.

Prior to training, all rats underwent surgery to infuse virus and implant fiber optics targeting the VTA (**Fig. 4**). We infused AAV5-EF1α-DIO-eNpHR3.0-eYFP (NpHR experimental group; n = 17) or AAV5-EF1α-DIO-eYFP (eYFP control group; n = 24) into the VTA of rats expressing Cre recombinase under the control of the TH promoter[31]. Note that, because reward was provided much more often in this experiment versus the first experiment (approximately twice as often), the nature of the conditioned response was different in this experiment. Rather than checking briefly many times for reward, the rats made fewer entries and spent more time inside the food cup. As a result, although we observed similar effects on both measures, here we plotted conditioned responding as the amount of time spent in the food cup rather than number of entries (see comment on response measures in Online Methods and **Supplementary Fig. 2** for more details).

After surgery and recovery, rats were food restricted until their body weight reached 85% of baseline. Training began with a day of preconditioning. Rats received a total of 12 pairings of two 10-s neutral cues (B→Y). On B→Y trials, continuous green light (532 nm, 16–18 mW output; Shanghai Laser & Optics Century Co., Ltd) was delivered for 2.5 s beginning 500 ms before the termination of B and continuing across the start of Y for 2 s in order to inactivate VTA dopamine neurons across a time window that would prevent any transient increase in activity of these neurons at the beginning of X. As a positive control, the rats also received 12 pairings of two other novel 10-s cues during this phase (A→X; 12 trials). No light was delivered across A→ X pairings. As no rewards were delivered during this phase of training, rats in both groups (eYFP and NpHR) exhibited very little responding at the food cup during cue presentation (**Fig. 5a**); a two-factor ANOVA on food cup responding during cue presentations (cue × group) revealed no main effect ($F_{1,39}$ = 1.88, P = 0.177) nor any interaction with group ($F_{3,117}$ = 0.425, P = 0.736).

Following preconditioning, the rats began conditioning, which continued for 4 d. Each day, the rats received 24 trials in which X and Y were both presented followed by delivery of two 45-mg sucrose pellets

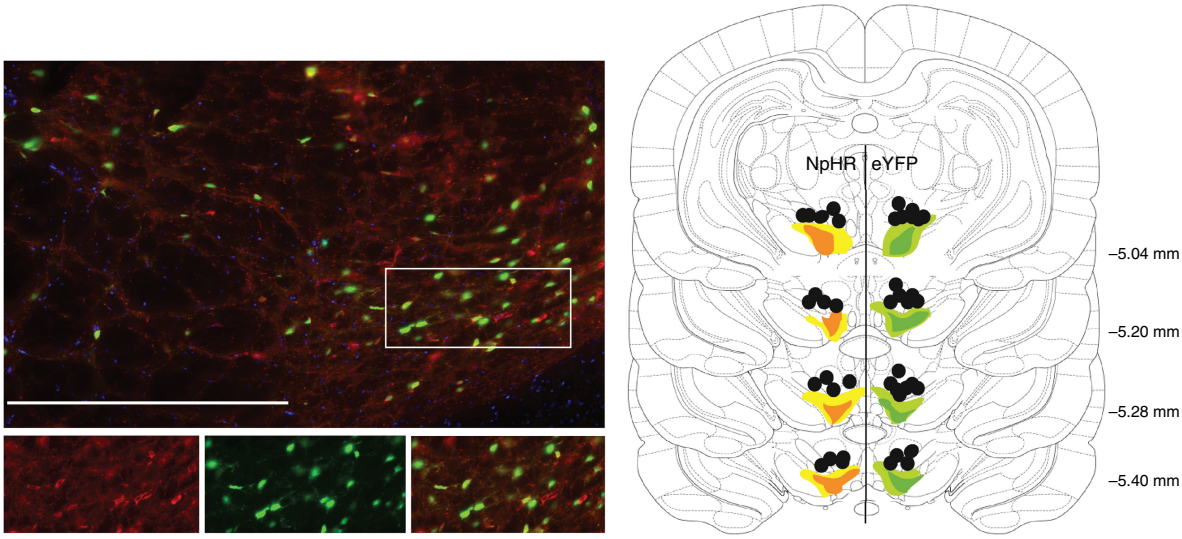

Figure 4 Immunohistochemical verification of Cre-dependent NpHR and eYFP expression in TH+ neurons and fiber placements in the VTA. Left: 90% of YFP-expressing neurons (green) also expressed TH (red). Bottom left: expansions of the region boxed at top. Right: Unilateral representation of the bilateral fiber placements and virus expression in each group. Fiber implants (black circles) were localized in the vicinity of eYFP (green) and NpHR (orange) expression in VTA. The light shading represents the maximal and the dark shading indicates the minimal spread of expression at each level. Scale bar, = 20 µm.

(X→2US). X was paired with one flavor of sucrose pellet, whereas Y was paired with another flavor (banana or grape, counterbalanced). Rats in both groups acquired a conditioned response to X and Y, as evident from the increase in the percentage of time they spent in the food cup during X and Y in expectation of sucrose pellets across days of conditioning (**Fig. 5b**). The acquisition of this conditioned response was similar in the two groups and for both cues; a three-factor ANOVA (cue × group × day) revealed a main effect of day ($F_{3,105}$ = 43.181, $P < 0.0001$) but no main effect of cue ($F_{1,39} = 0.008$, $P = 0.927$), group ($F_{1,39} = 0.094$, $P = 0.761$) or any cue × group interaction ($F_{1,39}$ = 1.113, $P = 0.298$). Thus, suppression of dopaminergic activity across the transition between B and Y did not produce lasting effects on processing of or learning about Y.

Lastly, the rats received a probe test in which each of the critical test cues (A and B) were presented six times each, in an interleaved and counterbalanced order, alone and without reward. As expected, rats in the eYFP group exhibited equally high rates of conditioned responding to both A and B (**Fig. 5c**), showing that regardless of light delivery, they learned the predictive relationship between both cue pairs and used them in the probe test to predict the delivery of sucrose pellets. By contrast, rats in the NpHR group exhibited significantly lower conditioned responding to B, the cue at the end of which we suppressed the dopamine neurons, than to A, the control cue. A two-factor ANOVA (cue × group) revealed a main effect of cue ($F_{1,39} = 5.94$, $P = 0.019$) and a significant cue by group interaction ($F_{1,39} = 4.68$, $P = 0.037$). Subsequent comparisons showed that this interaction was due to a significant difference in responding to cues A and B in the NpHR group ($F_{1,39} = 4.952$, $P = 0.012$), which was not present in the eYFP group ($F_{1,39} = 0.742$, $P = 0.483$). Notably, this within-subject difference could not be explained by the slightly (but not significantly) increased responding to A in the NpHR group ($F_{1,15} = 1.9$, $P = 0.189$, cue × response level in NpHR group; see **Supplementary Fig. 3** for additional information on high versus low responders). This difference is consistent with the proposal that, by preventing transient activation of the dopamine neurons at the B→Y transition, we prevented formation of the normal association between these two cues.

DISCUSSION

We have shown that activity in VTA dopamine neurons was sufficient and necessary for the formation of associative structures that underlie model-based behavior. In our first experiment, we demonstrated that transient activation of dopaminergic neurons, with a timing and duration designed to mimic a prediction error, unblocked stimulus–stimulus learning in a sensory preconditioning task, resulting in later responding that reflected a prediction of sucrose pellet delivery that could not have been directly acquired under the influence of the artificial dopamine transient we induced. In the second experiment, we demonstrated that suppressing dopamine neurons, with a timing and duration designed to interfere with any dopamine transients, blocked stimulus–stimulus learning in a sensory preconditioning task.

Current conceptualizations of dopamine transients as the reward prediction errors postulated by model-free reinforcement learning algorithms cannot explain these data. This is because the error signal in these models functions only to endow the predictive cue with a scalar quantity that reflects the value of future events; the resultant associative representation does not incorporate or link to specific information about the identity of these events beyond their value at the time of learning. As a result, this type of learning cannot explain why the rats in the first experiment searched in the food cup for sucrose pellets when C was presented in the probe test, and then stopped doing so when the pellets were no longer desirable. Even if the dopamine transients endowed C with cached value (as reinforcement-learning models propose), and this was the reason for the food cup responses, such effects on behavior would generalize beyond the specific reward and therefore be insensitive to its devaluation[35]. Indeed, if we stimulated dopamine to unblock learning when food was present, as has been done[12], these models predict that resultant responding would be insensitive to devaluation. Likewise, responding to C in our experiment also could not have reflected direct reinforcement of the motor response by the dopamine transient, since, in contrast to even the most well-controlled prior studies[12,14], this response was not present when dopamine neuron activity was manipulated. Of course, such nonspecific responding would also be insensitive to devaluation

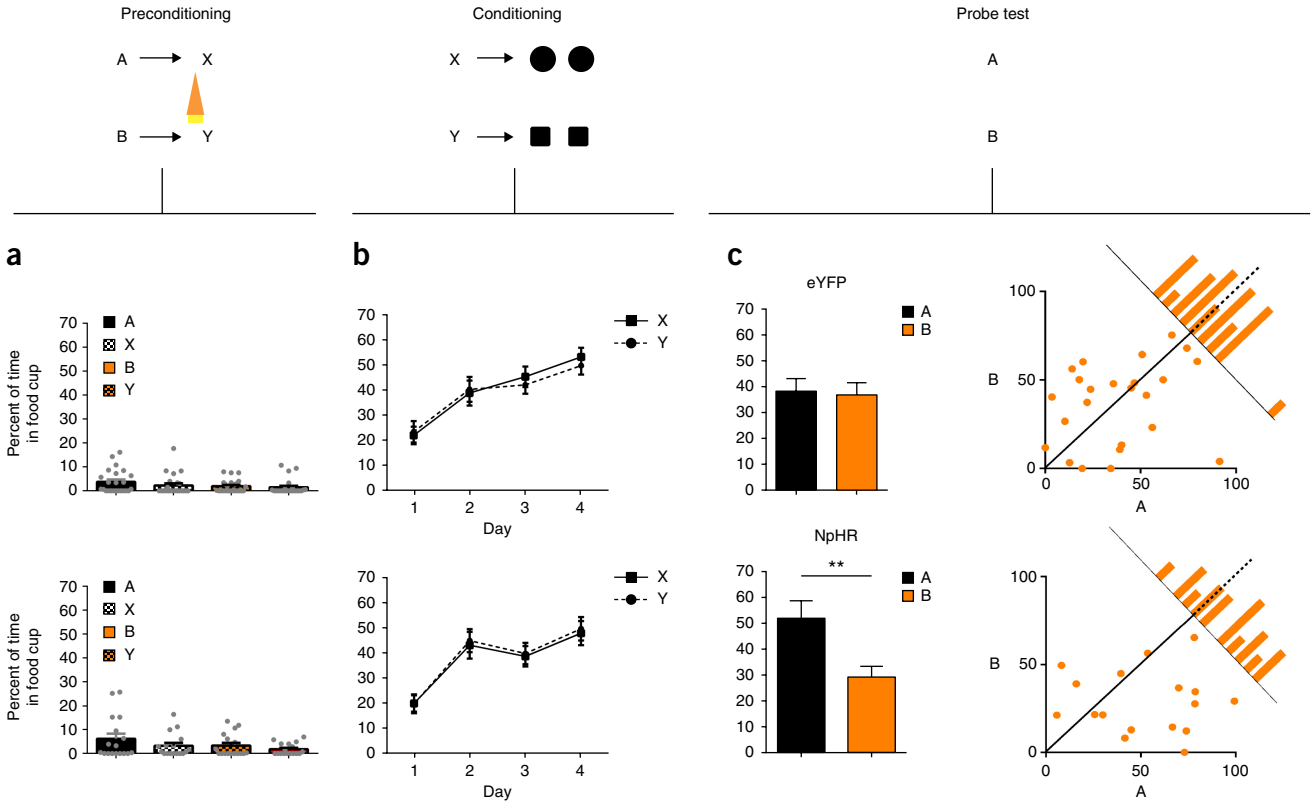

Figure 5 Brief optogenetic inhibition of dopamine neurons reduces the strength of associations between cues. Top: VTA dopamine neurons were inhibited by light delivery (orange symbol) in the 500 ms before the offset of A and carried through the first 2 s of X. Double dots and squares represent flavors of food pellets. (**a**–**c**) Plots show the percentage of time spent in the food cup during cue presentation across all phases of the sensory preconditioning task: (**a**) preconditioning, (**b**) conditioning and (**c**) the probe test. In each panel, top graph shows data from the eYFP control group ($n = 24$); bottom graph shows data from the experimental NpHR group ($n = 17$). To the extent that responding to A is equal to B in scatterplots represented in **c**, points should congregate around the diagonal. Histograms along the diagonal reveal the frequency (subject counts) of difference scores in responding to the cues that fall within a particular range. A two-factor ANOVA on food cup responding during cue presentations (cue × group) in preconditioning (**a**) revealed no main effect ($F_{1,39} = 1.88$, $P = 0.177$) nor any interaction with group ($F_{3,117} = 0.425$, $P = 0.736$). A three-factor ANOVA (cue × group × day) on data from conditioning (**b**) revealed a main effect of day ($F_{3,105} = 43.181$, $P < 0.0001$) but no main effect of cue ($F_{1,39} = 0.008$, $P = 0.927$), group ($F_{1,39} = 0.094$, $P = 0.761$) or any cue × group interaction ($F_{1,39} = 1.113$, $P = 0.298$). A two-factor ANOVA (cue × group) revealed a main effect of cue ($F_{1,39} = 5.94$, $P = 0.019$) and a significant cue × group interaction ($F_{1,39} = 4.68$, $P = 0.037$). Subsequent comparisons showed that this interaction was due to a significant difference in responding to cues A and B in the NpHR group ($F_{1,39} = 4.952$, $P = 0.012$), which was not present in the eYFP group ($F_{1,39} = 0.742$, $P = 0.483$). $^{**}P < 0.012$. Error bars, = s.e.m. Please see Online Methods for comment on response measures and **Supplementary Figure 4** for further details on responding during individual sessions in preconditioning.

of the food reward[36], contrary to our results. Thus our results go far beyond what can be explained by a cached-value prediction error.

Nor could the results from either experiment have reflected changes in salience or associability caused by manipulation of the dopamine neurons, either directly or via the addition or subtraction of cached value. While such effects have been reported following optogenetic activation of dopamine terminals in medial prefrontal cortex[37], we saw no evidence of this in either of our experiments involving manipulation of the cell bodies. For example, while increasing the salience or associability of X on the AC trials in our first experiment might have indirectly allowed X to enter into an association more readily with C, all theoretical accounts of which we are aware[38–40] would also predict lasting effects on processing and associability of X. These effects would facilitate learning for X in other parts of our task, but we did not observe any evidence of increased learning about X in other trials in the ChR2 rats. In particular, the ChR2 rats did not respond more to D than controls, nor did they show more rapid conditioning to X in the second phase of training. The same is true for our second experiment, in which we saw no changes in learning about Y during conditioning, indicating that suppressing dopamine neurons did not alter the salience or value of Y. It is also worth noting that direct effects on salience would be inconsistent with evidence that activation of VTA dopamine neurons diminishes extinction learning while inhibition of these neurons facilitates it[12,14]. These effects, achieved using the same optogenetic approaches applied here, are the opposite of what would be expected if manipulating these neurons directly altered salience.

Instead, the most parsimonious explanation of our results is that dopamine transients played a role in the formation of associative links between the neural representations of external events—whether rewarding or not—linking representations of neutral cues during preconditioning and representations of neutral cues with representations of rewards in other settings. Notably, this interpretation holds whether the ultimate behavior in the probe test reflected inference (i.e., if A→X and X→US, then A→US) or mediated learning during the conditioning phase (i.e., X evoked a memory of A that became directly associated with the US, so that later A→US; **Supplementary Fig. 1**). In either case, dopamine must be influencing the association between the cues in the first phase of training. While this proposal does not negate

a role for dopamine in learning about cached values, it does represent a substantial expansion of the kind of learning that dopaminergic prediction errors are thought to support. Along with recent data showing that these prediction-error signals can reflect value predictions derived from model-based associative structures[11,41–45], our results show that dopaminergic error signals are potentially richer, more complex and more capable than previously envisioned.

This is good news, given how difficult it has been to find plausible candidate neural substrates to signal these other types of prediction errors; the dopamine neurons appear relatively unique in the strength of their error signaling[46]. Of course, our experimental approach affected a general population of VTA dopamine neurons that likely projects broadly to multiple target regions. The neurons activated were determined somewhat at random, based on viral expression and light penetration. In this sense, our manipulations—both the activation as well as the suppression—were not, strictly speaking, physiological. This caveat is important to keep in mind when evaluating the importance of this or any other similar study. One way to view the ability of these manipulations to produce principled results is that the relatively simple and highly constrained behavioral designs allowed us to see real effects despite our poor ability to truly reproduce real-world patterns of activity. We speculate that in normal settings, the precise sort of associative information that is acquired under the influence of dopaminergic error signals will presumably reflect subtle variations in the content of the signal[33,47] combined with specialization of the downstream region or circuit[21,48,49].

Finally, it is worth noting that our results represent the first demonstration of which we are aware that learning about neutral cues is regulated by prediction errors. That is, in our blocking of sensory preconditioning procedure, we found that prior learning of the association between A and X blocked the ability of animals to learn that D predicts X. This shows that learning to associate neutral cues reflected contingency and not just contiguity between the two cues, matching previous demonstrations of blocking for cues predictive of motivationally consequential outcomes[23]. That dopamine transients were both sufficient and necessary for this type of learning is in accord with observations that dopamine neurons exhibit error-like responses to novel or unexpected neutral cues under some conditions[24,29,30]. Rather than reflecting a 'novelty bonus', such responses may reflect the informational prediction errors available in these circumstances to drive the sort of learning we have isolated here. Viewed from this perspective, the classic reward prediction errors normally observed in the firing of individual dopamine neurons might be a special (and especially strong) example of a more general function played by dopaminergic ensembles in signaling errors in event prediction. To determine whether this is true, it will be necessary to interrogate dopaminergic activity in more complex behavioral paradigms, in which the source of the errors can be manipulated independent of value. In addition, it will likely be important to monitor groups of dopamine neurons in real time, using approaches such as calcium imaging[33] to identify information represented across neurons, as has been done effectively to understand the functions of other brain regions[50].

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

ACKNOWLEDGMENTS

The authors thank K. Deisseroth and the Gene Therapy Center at the University of North Carolina at Chapel Hill for providing viral reagents and G. Stuber for technical advice on their use. We also thank B. Harvey and the NIDA Optogenetic and Transgenic Core, M. Morales and the NIDA Histology Core for their assistance, and P. Dayan and N. Daw for their comments. This work was supported by R01-MH098861 (to Y.N.) and by the Intramural Research Program at NIDA ZIA-DA000587 (to G.S.). The opinions expressed in this article are the authors' own and do not reflect the view of the NIH/DHHS.

AUTHOR CONTRIBUTIONS

M.J.S. and G.S. designed the experiments; M.J.S., M.A.L., H.M.B. and L.E.M. collected the data with technical advice and assistance from C.Y.C. and J.L.J. M.J.S. and G.S. analyzed the data. M.J.S., Y.N. and G.S. interpreted the data and wrote the manuscript with input from all authors.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Schultz, W. Dopamine neurons and their role in reward mechanisms. *Curr. Opin. Neurobiol.* **7**, 191–197 (1997).
2. Schultz, W., Dayan, P. & Montague, P.R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
3. Sutton, R.S. & Barto, A.G. Toward a modern theory of adaptive networks: expectation and prediction. *Psychol. Rev.* **88**, 135–170 (1981).
4. Nakahara, H. Multiplexing signals in reinforcement learning with internal models and dopamine. *Curr. Opin. Neurobiol.* **25**, 123–129 (2014).
5. Schultz, W. Dopamine reward prediction-error signalling: a two-component response. *Nat. Rev. Neurosci.* **17**, 183–195 (2016).
6. Tolman, E.C. Cognitive maps in rats and men. *Psychol. Rev.* **55**, 189–208 (1948).
7. Daw, N.D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* **8**, 1704–1711 (2005).
8. Gläscher, J., Daw, N., Dayan, P. & O'Doherty, J.P. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* **66**, 585–595 (2010).
9. Colwill, R.M. An associative analysis of instrumental learning. *Curr. Dir. Psychol. Sci.* **2**, 111–116 (1993).
10. Hollland, P.C. & Rescorla, R.A. The effect of two ways of devaluing the unconditioned stimulus after first- and second-order appetitive conditioning. *J. Exp. Psychol. Anim. Behav. Process.* **1**, 355–363 (1975).
11. Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P. & Dolan, R.J. Model-based influences on humans' choices and striatal prediction errors. *Neuron* **69**, 1204–1215 (2011).
12. Steinberg, E.E. *et al.* A causal link between prediction errors, dopamine neurons and learning. *Nat. Neurosci.* **16**, 966–973 (2013).
13. Eshel, N. *et al.* Arithmetic and local circuitry underlying dopamine prediction errors. *Nature* **525**, 243–246 (2015).
14. Chang, C.Y. *et al.* Brief optogenetic inhibition of dopamine neurons mimics endogenous negative prediction errors. *Nat. Neurosci.* **19**, 111–116 (2016).
15. Tsai, H.C. *et al.* Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science* **324**, 1080–1084 (2009).
16. Adamantidis, A.R. *et al.* Optogenetic interrogation of dopaminergic modulation of the multiple phases of reward-seeking behavior. *J. Neurosci.* **31**, 10829–10835 (2011).
17. Ilango, A. *et al.* Similar roles of substantia nigra and ventral tegmental dopamine neurons in reward and aversion. *J. Neurosci.* **34**, 817–822 (2014).
18. Stopper, C.M., Tse, M.T., Montes, D.R., Wiedman, C.R. & Floresco, S.B. Overriding phasic dopamine signals redirects action selection during risk/reward decision making. *Neuron* **84**, 177–189 (2014).
19. Brogden, W.J. Sensory pre-conditioning. *J. Exp. Psychol.* **25**, 323–332 (1939).
20. Blundell, P., Hall, G. & Killcross, S. Preserved sensitivity to outcome value after lesions of the basolateral amygdala. *J. Neurosci.* **23**, 7702–7709 (2003).
21. Jones, J.L. *et al.* Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science* **338**, 953–956 (2012).
22. Rizley, R.C. & Rescorla, R.A. Associations in second-order conditioning and sensory preconditioning. *J Comp Physiol Psychol* **81**, 1–11 (1972).
23. Kamin, L.J. "Attention-like" processes in classical conditioning. in *Miami Symposium on the Prediction of Behavior, 1967: Aversive Stimulation* (ed. M.R. Jones) 9–31 (University of Miami Press, 1968).
24. Tobler, P.N., Dickinson, A. & Schultz, W. Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *J. Neurosci.* **23**, 10402–10410 (2003).
25. Pan, W.-X., Schmidt, R., Wickens, J.R. & Hyland, B.I. Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *J. Neurosci.* **25**, 6235–6242 (2005).
26. Hollerman, J.R. & Schultz, W. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* **1**, 304–309 (1998).
27. Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B. & Uchida, N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* **482**, 85–88 (2012).

28. Takahashi, Y.K. *et al.* The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron* **62**, 269–280 (2009).
29. Kakade, S. & Dayan, P. Dopamine: generalization and bonuses. *Neural Netw.* **15**, 549–559 (2002).
30. Horvitz, J.C., Stewart, T. & Jacobs, B.L. Burst activity of ventral tegmental dopamine neurons is elicited by sensory stimuli in the awake cat. *Brain Res.* **759**, 251–258 (1997).
31. Witten, I.B. *et al.* Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron* **72**, 721–733 (2011).
32. D'Ardenne, K., McClure, S.M., Nystrom, L.E. & Cohen, J.D. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science* **319**, 1264–1267 (2008).
33. Parker, N.F. *et al.* Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. *Nat. Neurosci.* **19**, 845–854 (2016).
34. Day, J.J., Roitman, M.F., Wightman, R.M. & Carelli, R.M. Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat. Neurosci.* **10**, 1020–1028 (2007).
35. Holland, P.C. Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *J. Exp. Psychol. Anim. Behav. Process.* **30**, 104–117 (2004).
36. Dickinson, A. & Balleine, B.W. Motivational control of goal-directed action. *Anim. Learn. Behav.* **22**, 1–18 (1994).
37. Popescu, A.T., Zhou, M.R. & Poo, M.-M. Phasic dopamine release in the medial prefrontal cortex enhances stimulus discrimination. *Proc. Natl. Acad. Sci. USA* **113**, E3169–E3176 (2016).
38. Mackintosh, N.J. A theory of attention: variations in the associability of stimuli with reinforcement. *Psychol. Rev.* **82**, 276–298 (1975).
39. Pearce, J.M. & Hall, G. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* **87**, 532–552 (1980).
40. Esber, G.R. & Haselgrove, M. Reconciling the influence of predictiveness and uncertainty on stimulus salience: a model of attention in associative learning. *Proceedings of the Royal Society of London B: Biological Sciences* http://dx.doi.org/10.1098/rspb.2011.0836 (2011).
41. Sadacca, B.F., Jones, J.L. & Schoenbaum, G. Midbrain dopamine neurons compute inferred and cached value prediction errors in a common framework. *eLife* **5**, e13665 (2016).
42. Cone, J.J. *et al.* Physiological state gates acquisition and expression of mesolimbic reward prediction signals. *Proc. Natl. Acad. Sci. USA* **113**, 1943–1948 (2016).
43. Bromberg-Martin, E.S., Matsumoto, M., Hong, S. & Hikosaka, O. A pallidus-habenula-dopamine pathway signals inferred stimulus values. *J. Neurophysiol.* **104**, 1068–1076 (2010).
44. Aitken, T.J., Greenfield, V.Y. & Wassum, K.M. Nucleus accumbens core dopamine signaling tracks the need-based motivational value of food-paired cues. *J. Neurochem.* **136**, 1026–1036 (2016).
45. Deserno, L. *et al.* Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. *Proc. Natl. Acad. Sci. USA* **112**, 1595–1600 (2015).
46. Eshel, N., Tian, J., Bukwich, M. & Uchida, N. Dopamine neurons share common response function for reward prediction error. *Nat. Neurosci.* **19**, 479–486 (2016).
47. Lammel, S. *et al.* Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron* **57**, 760–773 (2008).
48. Wimmer, G.E. & Shohamy, D. Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science* **338**, 270–273 (2012).
49. Robinson, S. *et al.* Chemogenetic silencing of neurons in retrosplenial cortex disrupts sensory preconditioning. *J. Neurosci.* **34**, 10982–10988 (2014).
50. Johnson, A., Fenton, A.A., Kentros, C. & Redish, A.D. Looking for cognition in the structure within the noise. *Trends Cogn. Sci.* **13**, 55–64 (2009).

ONLINE METHODS

Surgical procedures. Rats received bilateral infusions of 1.2 µL AAV5-EF1α–DIO-ChR2-eYFP (n = 18), AAV5-EF1α-DIO (n = 17), eNpHR3.0-eYFP (n = 17) or AAV5-EF1α-DIO-eYFP (n = 43) into the VTA at the following coordinates relative to bregma: AP: −5.3 mm; ML: ± 0.7 mm; DV: −6.5 mm and −7.7 (females) or −7.0 mm and −8.2 mm (males). Virus was obtained from the Vector Core at University of North Carolina at Chapel Hill (UNC Vector Core). During surgery, optic fibers were implanted bilaterally (200-µm diameter, Precision Fiber Products, CA) at the following coordinates relative to bregma: AP: −5.3 mm; ML: ± 2.61 mm and DV: −7.05 mm (female) or −7.55 mm (male) at an angle of 15° pointed toward the midline.

Apparatus. Training was conducted in eight standard behavioral chambers (Coulbourn Instruments; Allentown, PA), which were individually housed in light- and sound-attenuating boxes (Jim Garmon, JHU Psychology Machine Shop). Each chamber was equipped with a pellet dispenser that delivered 45-mg pellets into a recessed magazine (food cup) when activated. Access to the food cup was detected by means of infrared detectors mounted across the opening of the recess. Two differently shaped panel lights were located on the right wall of the chamber above the food cup. The chambers contained a speaker connected to white noise and tone generators and a relay that delivered a 5-kHz click stimulus. A computer equipped with GS3 software (Coulbourn Instruments, Allentown, PA) controlled the equipment and recorded the responses. Raw data were output to and processed in Matlab (Mathworks, Natick, MA) to extract relevant response measures, which were analyzed in SPSS software (IBM analytics, Sydney, Australia).

Housing. Rats were housed singly and maintained on a 12-h light–dark cycle; all behavioral experiments took place during the light cycle. Rats had *ad libitum* access to food and water unless undergoing the behavioral experiment, during which they received sufficient chow to maintain them at ~85% of their free-feeding body weight. All experimental procedures were conducted in accordance with the NIDA-IRP Institutional Animal Care and Use Committee of the US National Institute of Health guidelines.

General behavioral procedures. Trials consisted of 10-s cues as described below. Trial types were interleaved in miniblocks, with the specific order unique to each rat and counterbalanced across groups. Intertrial intervals varied around a 6-min mean. Unless otherwise noted, daily training was divided into a morning (AM) and afternoon (PM) session. Inclusion of AM/PM as a factor in our analyses found no significant interactions with group, so we collapsed this factor in the analyses in the main text.

Response measures. We measured entry into the food cup to assess conditioned responding. Food cup entries were registered when the rat broke a light beam placed across the opening of the food cup. This simple measure allowed us to calculate a variety of metrics including response latency after cue onset, number of entries to the food cup during the cue and the overall percentage of time spent in the food cup during the cue. These metrics were generally correlated during conditioning, and all reflect to some extent the expectation of food delivery at the end of the cue in a task such as that used here[51,52]. One exception was the relationship between number of entries and time spent in the food cup, which can vary depending on reward density in a task or behavioral setting[52–56]. This tendency prompted us to focus on two different metrics in the two experiments presented in this paper. Specifically, in Experiment 1, the overall rate of reward was relatively low, since only one cue was rewarded and it was only rewarded in a handful of trials in one phase of training. As a result, the rats exhibited relatively brief entries into the food cup, so number of entries was the most reliable measure of conditioning. By contrast, in Experiment 2, the overall rate of reward was much higher, since two cues were rewarded, with two different rewards, in a larger number of trials. As a result, the rats spent much more time in the food cup each time they entered, so the percentage of time in the food cup was the most reliable measure of conditioning consistent with previous research[52–56]. We note that in each case, the other measure yielded a qualitatively similar pattern of responding, but we have chosen to present the more reliable one for each experiment.

Histology and immunohistochemistry. All rats were euthanized with an overdose of carbon dioxide and perfused with phosphate buffered saline (PBS) followed by 4% paraformaldehyde (Santa Cruz Biotechnology Inc., CA). Fixed brains were cut in 40-µm sections to examine fiber tip position under a fluorescence microscope (Olympus Microscopy, Japan). For immunohistochemistry, the brain slices were first blocked in 10% goat serum made in 0.1% Triton X-100/1× PBS and then incubated in anti-TH antisera (MAB318, 1:600, EMD Millipore, Billerica, Massachusetts) followed by Alexa Fluor 568 secondary antisera (A11031, 1:1,000, Invitrogen, Carlsbad, CA). Images of these brain slices were acquired by a fluorescence Virtual Slide microscope (Olympus America, NY) and later analyzed in Adobe Photoshop. We then counted the proportion of cells expressing eYFP that also co-stained for TH within the VTA of 4 subjects using sections taken from AP −5.0 mm to −6.0 mm. Positive staining was defined as a signal 2.5× baseline intensity, with a cell diameter larger than 5 µm, co-localized within cells reactive to DAPI staining. This encompassed the area likely to achieve good light penetration from our fibers.

Statistical analyses. All statistics were conducted using the SPSS 24 IBM statistics package. Generally, analyses were conducted using a mixed-design repeated-measures ANOVA, with the exception of the data represented in **Figure 3**, for which we conducted one-way between-subjects ANOVA as appropriate. All analyses of simple main effects were planned and orthogonal and therefore did not necessitate controlling for multiple comparisons. Data distribution was assumed to be normal, but homoscedasticity was not formally tested. Except for histological analysis, data collection and analyses were not performed blind to the conditions of the experiments. A **Supplementary Methods Checklist** is available.

Data availability. The data that support the findings of this study, and any associated custom programs used for its acquisition, are available from the corresponding authors upon reasonable request.

Experiment 1. *Subjects*. Thirty-seven experimentally naive male (n = 19) and female (n = 18) Long-Evans transgenic rats of approximately 4 months of age at surgery and carrying a TH-dependent Cre expressing system (NIDA Animals Breeding Facility) were used in this study. Sample sizes were chosen based on similar prior experiments that have elicited significant results with a similar number of rats. No formal power analyses were conducted. Rats were randomly assigned to groups and distributed equally by age, gender and weight. Prior to data analysis, six rats were removed from the experiment due to illness, virus or cannula misplacement.

Blocking of sensory preconditioning. Training used a total of six different stimuli, drawn from stock equipment available from Coulbourn and included four auditory (tone, siren, clicker, white noise) and two visual stimuli (flashing light, steady light). Assignment of these stimuli to the cues depicted in **Figure 2** and described in the text was counterbalanced across rats within each modality (A and E were visual while C, D, F and X were auditory).

Training began with 2 d of preconditioning. On the first day, the rats received 16 presentations of A→X, in which a 10-s presentation of A was immediately followed by a 10-s presentation of X. On the second day, the rats received 8 presentations of A→X alone, as well as 8 presentations each of three 10-s compound cues (EF, AD, AC) followed by X (i.e., EF→X; AD→X; AC→X). On AC trials, light (473 nm, 16–18 mW output; Shanghai Laser & Optics Century Co., Ltd) was delivered into the VTA for 1 s at a rate of 20 Hz at the beginning of X; on AD trials, the same light pattern was delivered during the intertrial interval, 120–180 s after termination of X. Following preconditioning, rats underwent 4 d of conditioning in which X was presented 24 times each day and was followed immediately by delivery of two 45-mg sucrose pellets (5TUT; TestDiet, MO). Finally, rats received a probe test in which each of the critical test cues (C, D, F) was presented four times, alone and without reward.

Devaluation. A subset of the rats in the experimental ChR2 group (n = 10) underwent additional training after the probe test described above. These rats received reminder training, in which X was again presented 12 times with reward, and then they were divided into two equal, performance-matched groups. Subsequently they received 30 min of access to 10 g of the sucrose pellet reward to habituate them to receiving pellets outside of the training chamber, after which they began 3 d of training to devalue the sucrose pellet reward. Each day, one group (devalued; n = 5) received access to the sucrose pellets for 30 min, followed immediately by an intraperitoneal injection of a 0.15-M solution of lithium chloride (LiCl; Sigma-Aldrich, MO) to induce nausea; the other group (nondevalued; n = 5)

received the injections and were given a yoked amount of sucrose pellets approximately 6 h later. Forty-eight hours after the third LiCl injection, all rats were given a final probe test in which C was presented 12 times, alone and without reward, followed by a final consumption test in which all rats were received access to 10 g of the sucrose pellets for 30 min.

Experiment 2. *Subjects*. Forty-one experimentally-naive male ($n = 33$) and female ($n = 8$) Long-Evans transgenic rats of approximately 4 months of age at surgery and carrying a TH-dependent Cre expressing system (NIDA animal breeding facility) were used in this study. Sample sizes were chosen based on similar prior experiments that elicited significant results with a similar number of rats. No formal power analyses were conducted. Rats were randomly assigned to groups and distributed equally by age, gender and weight. Prior to data analysis, two rats were removed from the experiment due to illness, virus or cannula misplacement.

Sensory preconditioning. Training used a total of four different auditory stimuli, drawn from stock equipment available from Coulbourn, which included tone, siren, clicker and white noise. Assignment of these stimuli to the cues depicted in **Figure 5** and described in the text was counterbalanced across rats. Training began with 1 d of preconditioning, in which where rats received 12 presentations of the A→X serial compound and 12 trials of the B→Y serial compound. Following preconditioning, rats began conditioning, in which they received 24 trials of X and 24 trials of Y each paired with a different reinforcer (either banana or grape pellets). Following 4 d of this training, rats received a probe test in which cues A and B were each presented six times in the absence of any reinforcement.

51. Holland, P.C. Conditioned stimulus as a determinant of the form of the Pavlovian conditioned response. *J. Exp. Psychol. Anim. Behav. Process.* **3**, 77–104 (1977).
52. McDannald, M.A., Lucantonio, F., Burke, K.A., Niv, Y. & Schoenbaum, G. Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. *J. Neurosci.* **31**, 2700–2705 (2011).
53. Holland, P.C. & Gallagher, M. Effects of amygdala central nucleus lesions on blocking and unblocking. *Behav. Neurosci.* **107**, 235–245 (1993).
54. Holland, P.C. & Kenmuir, C. Variations in unconditioned stimulus processing in unblocking. *J. Exp. Psychol. Anim. Behav. Process.* **31**, 155–171 (2005).
55. Sharpe, M.J. & Killcross, S. The prelimbic cortex contributes to the down-regulation of attention toward redundant cues. *Cereb. Cortex* **24**, 1066–1074 (2014).
56. Burke, K.A., Franz, T.M., Miller, D.N. & Schoenbaum, G. The role of the orbitofrontal cortex in the pursuit of happiness and more specific rewards. *Nature* **454**, 340–344 (2008).

doi:10.1038/nn.4538

nature neuroscience

Corresponding Author: Geoffrey Schoenbaum

Manuscript Number: NN-A56755-A (RE-REVISION)

Manuscript Type: Article

Main Figures: 5

Supplementary Figures: 4

Supplementary Tables: 0

Supplementary Videos: 0

Reporting Checklist for Nature Neuroscience

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. For more information, please read Reporting Life Sciences Research.

Please note that in the event of publication, it is mandatory that authors include all relevant methodological and statistical information in the manuscript.

Statistics reporting, by figure

- Please specify the following information for each panel reporting quantitative data, and where each item is reported (section, e.g. Results, & paragraph number).
- Each figure legend should ideally contain an exact sample size (n) for each experimental group/condition, where n is an exact number and not a range, a clear definition of how n is defined (for example x cells from x slices from x animals from x litters, collected over x days), a description of the statistical test used, the results of the tests, any descriptive statistics and clearly defined error bars if applicable.
- For any experiments using custom statistics, please indicate the test used and stats obtained for each experiment.
- Each figure legend should include a statement of how many times the experiment shown was replicated in the lab; the details of sample collection should be sufficiently clear so that the replicability of the experiment is obvious to the reader.
- For experiments reported in the text but not in the figures, please use the paragraph number instead of the figure number.

Note: Mean and standard deviation are not appropriate on small samples, and plotting independent data points is usually more informative. When technical replicates are reported, error and significance measures reflect the experimental variability and not the variability of the biological process; it is misleading not to state this clearly.

		TEST USED		n			DESCRIPTIVE STATS (AVERAGE, VARIANCE)		P VALUE		DEGREES OF FREEDOM & F/t/z/R/ETC VALUE	
	FIGURE NUMBER	WHICH TEST?	SECTION & PARAGRAPH #	EXACT VALUE	DEFINED?	SECTION & PARAGRAPH #	REPORTED?	SECTION & PARAGRAPH #	EXACT VALUE	SECTION & PARAGRAPH #	VALUE	SECTION & PARAGRAPH #
example	1a	one-way ANOVA	Fig. legend	9, 9, 10, 15	mice from at least 3 litters/group	Methods para 8	error bars are mean +/- SEM	Fig. legend	p = 0.044	Fig. legend	F(3, 36) = 2.97	Fig. legend
example	results, para 6	unpaired t-test	Results para 6	15	slices from 10 mice	Results para 6	error bars are mean +/- SEM	Results para 6	p = 0.0006	Results para 6	t(28) = 2.808	Results para 6

Nature Neuroscience: doi:10.1038/nn.4538

		TEST USED		n			DESCRIPTIVE STATS (AVERAGE, VARIANCE)		P VALUE		DEGREES OF FREEDOM & F/t/z/R/ETC VALUE	
	FIGURE NUMBER	WHICH TEST?	SECTION & PARAGRAPH #	EXACT VALUE	DEFINED?	SECTION & PARAGRAPH #	REPORTED?	SECTION & PARAGRAPH #	EXACT VALUE	SECTION & PARAGRAPH #	VALUE	SECTION & PARAGRAPH #
+ -	2a	ANOVA	description of first experiment	37	rats in two groups	start of description of first experiment	data and SEM plotted in fig 2	fig 2a	multiple, please see text for relevant p's	results of first expt	multiple, depending on effect/ interaction in multifactor factor anova, please see text	results of first expt
+ -	2b	ANOVA and posthocs	description of first experiment	37	rats in two groups	start of description of first experiment	data and SEM plotted in fig 2	fig 2b	multiple, please see text for relevant p's	results of first expt	multiple, depending on effect/ interaction in multifactor factor anova, please see text	results of first expt
+ -	2c	ANOVA and posthocs	description of first experiment	37	rats in two groups	start of description of first experiment	data and SEM plotted in fig 2	fig 2c	multiple, please see text for relevant p's	results of first expt	multiple, depending on effect/ interaction in multifactor factor anova, please see text	results of first expt
+ -	fig 3	ANOVAs	description of first experiment	10	rats in two groups	description of devaluation experiment	data and SEM plotted in fig 3	fig 3	multiple, please see text for relevant p's	results of first expt	multiple, depending on effect/ interaction in multifactor factor anova, please see text	results of first expt
+ -	5a	ANOVAs	description of second experiment	41	rats in two groups	start of description of second experiment	data and SEM plotted in fig 5	fig 5a	multiple, please see text for relevant p's	results of second expt	multiple, depending on effect/ interaction in multifactor factor anova, please see text	results of second expt
+ -	5b	ANOVAs	description of second experiment	41	rats in two groups	start of description of second experiment	data and SEM plotted in fig 5	fig 5b	multiple, please see text for relevant p's	results of second expt	multiple, depending on effect/ interaction in multifactor factor anova, please see text	results of second expt
+ -	5c	ANOVAs	description of second experiment	41	rats in two groups	start of description of second experiment	data and SEM plotted in fig 5	fig 5c	multiple, please see text for relevant p's	results of second expt	multiple, depending on effect/ interaction in multifactor factor anova, please see text	results of second expt

Nature Neuroscience: doi:10.1038/nn.4538

Representative figures

1. Are any representative images shown (including Western blots and immunohistochemistry/staining) in the paper?

   If so, what figure(s)?

   Yes fig 1 and 4

2. For each representative image, is there a clear statement of how many times this experiment was successfully repeated and a discussion of any limitations in repeatability?

   If so, where is this reported (section, paragraph #)?

   the figures show histology showing expression. the methods describe the number of rats and extent of expression is in the results.

Statistics and general methods

1. Is there a justification of the sample size?

   If so, how was it justified?

   Where (section, paragraph #)?

   Even if no sample size calculation was performed, authors should report why the sample size is adequate to measure their effect size.

   sample sizes are based on past experience, and is justified with references to prior papers showing the effect we are replicating here.

2. Are statistical tests justified as appropriate for every figure?

   Where (section, paragraph #)?

   yes. statistical tests are described throughout the results. typically we use ANOVA's with interactions between groups or main effects, with posthocs. In response to some of the new reviewer comments, we also use MANOVA as described in supplemental.

   a. If there is a section summarizing the statistical methods in the methods, is the statistical test for each experiment clearly defined?

      yes there are clear descriptions of the statistics for each experiment in the text

   b. Do the data meet the assumptions of the specific statistical test you chose (e.g. normality for a parametric test)?

      Where is this described (section, paragraph #)?

      yes. the statistics are described and enumerated throughout the results.

   c. Is there any estimate of variance within each group of data?

      Is the variance similar between groups that are being statistically compared?

      Where is this described (section, paragraph #)?

      we assume the variance is normal, but we have not formally tested it. we say this in the methods.

   d. Are tests specified as one- or two-sided?

      tests are two sided.

   e. Are there adjustments for multiple comparisons?

      we do not do multiple comparisons. data from each phase are analyzed via ANOVA based on a priori predictions . only with appropriate main effects or interactions do we do posthocs.

Nature Neuroscience: doi:10.1038/nn.4538

3. To promote transparency, *Nature Neuroscience* has stopped allowing bar graphs to report statistics in the papers it publishes. If you have bar graphs in your paper, please make sure to switch them to dot-plots (with central and dispersion statistics displayed) or to box-and-whisker plots to show data distributions.	done
4. Are criteria for excluding data points reported? Was this criterion established prior to data collection? Where is this described (section, paragraph #)?	we did not exclude any data points. several rats were excluded during the experiments for illness or virus/cannula misplacement. their data were not analyzed.
5. Define the method of randomization used to assign subjects (or samples) to the experimental groups and to collect and process data. If no randomization was used, state so. Where does this appear (section, paragraph #)?	we did not randomize subjects. they were assigned to control or expt groups based on whether they received eYFP or ChR2 or NpHR infusions. all data collection was by computer. identity of cues was counterbalanced. all of this is described in our methods.
6. Is a statement of the extent to which investigator knew the group allocation during the experiment and in assessing outcome included? If no blinding was done, state so. Where (section, paragraph #)?	all data collection was by computer.
7. For experiments in live vertebrates, is a statement of compliance with ethical guidelines/regulations included? Where (section, paragraph #)?	yes in the first paragraph of the methods.
8. Is the species of the animals used reported? Where (section, paragraph #)?	yes in the first paragraph of methods.
9. Is the strain of the animals (including background strains of KO/ transgenic animals used) reported? Where (section, paragraph #)?	yes in the first paragraph of methods and in main text.
10. Is the sex of the animals/subjects used reported? Where (section, paragraph #)?	yes in the first paragraph of methods.
11. Is the age of the animals/subjects reported? Where (section, paragraph #)?	yes in the first paragraph of methods
12. For animals housed in a vivarium, is the light/dark cycle reported? Where (section, paragraph #)?	yes in the first paragraph of methods
13. For animals housed in a vivarium, is the housing group (i.e. number of animals per cage) reported? Where (section, paragraph #)?	yes in the first paragraph of methods

Nature Neuroscience: doi:10.1038/nn.4538

14. For behavioral experiments, is the time of day reported (e.g. light or dark cycle)?

    Where (section, paragraph #)?

yes in the first paragraph of methods

15. Is the previous history of the animals/subjects (e.g. prior drug administration, surgery, behavioral testing) reported?

    Where (section, paragraph #)?

no prior history

    a. If multiple behavioral tests were conducted in the same group of animals, is this reported?

       Where (section, paragraph #)?

yes this is described clearly in the entire paper. it is two experiments basically, each conducted on a new set of rats.

16. If any animals/subjects were excluded from analysis, is this reported?

    Where (section, paragraph #)?

we excluded 10 subjects who failed to show viral expression within VTA or whose fibers were not located within the VTA or who could not enter the food cup with the fibers on their heads to respond. exclusions were made prior to data analysis, and thus they are not reported in the manuscript.

    a. How were the criteria for exclusion defined?

       Where is this described (section, paragraph #)?

criterion for inclusion were that fibers were located within the target area in VTA and that there was visible viral expression in VTA.

    b. Specify reasons for any discrepancy between the number of animals at the beginning and end of the study.

       Where is this described (section, paragraph #)?

none

Reagents

1. Have antibodies been validated for use in the system under study (assay and species)?

no

    a. Is antibody catalog number given?

       Where does this appear (section, paragraph #)?

    b. Where were the validation data reported (citation, supplementary information, Antibodypedia)?

       Where does this appear (section, paragraph #)?

2. Cell line identity

    a. Are any cell lines used in this paper listed in the database of commonly misidentified cell lines maintained by ICLAC and NCBI Biosample?

       Where (section, paragraph #)?

no

Nature Neuroscience: doi:10.1038/nn.4538

b. If yes, include in the Methods section a scientific justification of their use--indicate here in which section and paragraph the justification can be found.

c. For each cell line, include in the Methods section a statement that specifies:
- the source of the cell lines
- have the cell lines been authenticated? If so, by which method?
- have the cell lines been tested for mycoplasma contamination?

Where (section, paragraph #)?

Data availability

Provide a Data availability statement in the Methods section under "Data availability", which should include, where applicable:
• Accession codes for deposited data
• Other unique identifiers (such as DOIs and hyperlinks for any other datasets)
• At a minimum, a statement confirming that all relevant data are available from the authors
• Formal citations of datasets that are assigned DOIs
• A statement regarding data available in the manuscript as source data
• A statement regarding data available with restrictions

See our data availability and data citations policy page for more information.

Data deposition in a public repository is mandatory for:

a. Protein, DNA and RNA sequences
b. Macromolecular structures
c. Crystallographic data for small molecules
d. Microarray data

Deposition is strongly recommended for many other datasets for which structured public repositories exist; more details on our data policy are available here. We encourage the provision of other source data in supplementary information or in unstructured repositories such as Figshare and Dryad.

We encourage publication of Data Descriptors (see Scientific Data) to maximize data reuse.

Where is the Data Availability statement provided (section, paragraph #)?

Data is available from the authors as per NIH policy.

Nature Neuroscience: doi:10.1038/nn.4538

Computer code/software

Any custom algorithm/software that is central to the methods must be supplied by the authors in a usable and readable form for readers at the time of publication. However, referees may ask for this information at any time during the review process.

1. Identify all custom software or scripts that were required to conduct the study and where in the procedures each was used.

Computer programs were written and anlayses were run using commercially available software.

2. If computer code was used to generate results that are central to the paper's conclusions, include a statement in the Methods section under "**Code availability**" to indicate whether and how the code can be accessed. Include version information as necessary and any restrictions on availability.

All information about the computer programs used is in the methods .

Human subjects

1. Which IRB approved the protocol?

   Where is this stated (section, paragraph #)?

2. Is demographic information on all subjects provided?

   Where (section, paragraph #)?

3. Is the number of human subjects, their age and sex clearly defined?

   Where (section, paragraph #)?

4. Are the inclusion and exclusion criteria (if any) clearly specified?

   Where (section, paragraph #)?

5. How well were the groups matched?

   Where is this information described (section, paragraph #)?

6. Is a statement included confirming that informed consent was obtained from all subjects?

   Where (section, paragraph #)?

7. For publication of patient photos, is a statement included confirming that consent to publish was obtained?

   Where (section, paragraph #)?

Nature Neuroscience: doi:10.1038/nn.4538

fMRI studies

For papers reporting functional imaging (fMRI) results please ensure that these minimal reporting guidelines are met and that all this information is clearly provided in the methods:

1. Were any subjects scanned but then rejected for the analysis after the data was collected?

    a. If yes, is the number rejected and reasons for rejection described?

    Where (section, paragraph #)?

2. Is the number of blocks, trials or experimental units per session and/or subjects specified?

    Where (section, paragraph #)?

3. Is the length of each trial and interval between trials specified?

4. Is a blocked, event-related, or mixed design being used? If applicable, please specify the block length or how the event-related or mixed design was optimized.

5. Is the task design clearly described?

    Where (section, paragraph #)?

6. How was behavioral performance measured?

7. Is an ANOVA or factorial design being used?

8. For data acquisition, is a whole brain scan used?

    If not, state area of acquisition.

    a. How was this region determined?

9. Is the field strength (in Tesla) of the MRI system stated?

    a. Is the pulse sequence type (gradient/spin echo, EPI/spiral) stated?

    b. Are the field-of-view, matrix size, slice thickness, and TE/TR/flip angle clearly stated?

10. Are the software and specific parameters (model/functions, smoothing kernel size if applicable, etc.) used for data processing and pre-processing clearly stated?

Nature Neuroscience: doi:10.1038/nn.4538

11. Is the coordinate space for the anatomical/functional imaging data clearly defined as subject/native space or standardized stereotaxic space, e.g., original Talairach, MNI305, ICBM152, etc? Where (section, paragraph #)?

12. If there was data normalization/standardization to a specific space template, are the type of transformation (linear vs. nonlinear) used and image types being transformed clearly described? Where (section, paragraph #)?

13. How were anatomical locations determined, e.g., via an automated labeling algorithm (AAL), standardized coordinate database (Talairach daemon), probabilistic atlases, etc.?

14. Were any additional regressors (behavioral covariates, motion etc) used?

15. Is the contrast construction clearly defined?

16. Is a mixed/random effects or fixed inference used?

    a. If fixed effects inference used, is this justified?

17. Were repeated measures used (multiple measurements per subject)?

    a. If so, are the method to account for within subject correlation and the assumptions made about variance clearly stated?

18. If the threshold used for inference and visualization in figures varies, is this clearly stated?

19. Are statistical inferences corrected for multiple comparisons?

    a. If not, is this labeled as uncorrected?

20. Are the results based on an ROI (region of interest) analysis?

    a. If so, is the rationale clearly described?

    b. How were the ROI's defined (functional vs anatomical localization)?

21. Is there correction for multiple comparisons within each voxel?

22. For cluster-wise significance, is the cluster-defining threshold and the corrected significance level defined?

Nature Neuroscience: doi:10.1038/nn.4538

Additional comments

Additional Comments

Nature Neuroscience: doi:10.1038/nn.4538

Corrigendum: Dopamine transients are sufficient and necessary for acquisition of model-based associations

Melissa J Sharpe, Chun Yun Chang, Melissa A Liu, Hannah M Batchelor, Lauren E Mueller, Joshua L Jones, Yael Niv & Geoffrey Schoenbaum

***Nat. Neurosci.*; doi:10.1038/nn.4538; corrected online 10 April 2017**

In the version of this article initially published online, the checkered and filled boxes were reversed in the keys to Figures 3a and 3b. The error has been corrected in the print, PDF and HTML versions of this article.

Corrigendum: Dopamine transients are sufficient and necessary for acquisition of model-based associations

Melissa J Sharpe, Chun Yun Chang, Melissa A Liu, Hannah M Batchelor, Lauren E Mueller, Joshua L Jones, Yael Niv & Geoffrey Schoenbaum

***Nat. Neurosci.* 20, 735–742 (2017); published online 3 April 2017; corrected online 10 April 2017; corrected after print 5 May 2017**

In the version of this article initially published, the histogram in Figure 2c, center top graph, was duplicated from the panel below, and the remaining histograms accompanying the scatter plots in Figures 2c and 5c were slightly mis-scaled and misaligned relative to the scatterplots. The histograms, as well as the vertical scaling of Figure 5c, bottom right graph, have been adjusted. Also, one data point from the scatterplot in the top right panel of Figure 2c had originally been transformed from a negative value on the vertical axis to its absolute value. The errors have been corrected in the PDF and HTML versions of this article.

nature neuroscience

CORRECTION

https://doi.org/10.1038/s41593-018-0202-5

Author Correction: Dopamine transients are sufficient and necessary for acquisition of model-based associations

Melissa J Sharpe, Chun Yun Chang, Melissa A Liu, Hannah M Batchelor, Lauren E Mueller, Joshua L Jones, Yael Niv and Geoffrey Schoenbaum

Correction to: *Nature Neuroscience* https://doi.org/10.1038/nn.4538, published online 3 April 2017.

In the version of this article initially published, the laser activation at the start of cue X in experiment 1 was described in the first paragraph of the Results and in the third paragraph of the Experiment 1 section of the Methods as lasting 2 s; in fact, it lasted only 1 s. The error has been corrected in the HTML and PDF versions of the article.

Published online: 17 July 2018
https://doi.org/10.1038/s41593-018-0202-5