

Title: The cognitive basis of intracranial self-stimulation of midbrain dopamine neurons.

One sentence summary: Dopamine neurons only support ICSS at supraphysiological frequencies and in a manner not reflecting dopamine's role in learning.

Authors: Samuel J. Millard¹, Ivy B. Hoang¹, Zara Greer¹, Shayna L. O'Connor^{2,3,4}, Kate M. Wassum¹, Morgan H. James^{2,3}, David J. Barker^{3,4}, Melissa J. Sharpe¹

Affiliations:

¹Department of Psychology, University of California, Los Angeles CA

²Department of Psychiatry, Robert Wood Johnson Medical School, Rutgers University, Piscataway NJ

³Brain Health Institute, Rutgers Biomedical Health Sciences, Rutgers University, Piscataway NJ

⁴Department of Psychology, Rutgers, The State University of New Jersey, Piscataway NJ

Correspondence to: melissa.j.sharpe@psych.ucla.edu

Abstract:

Recently there has been a reckoning in the dopamine field. This has suggested that the dopamine prediction error may function as a teaching signal, without endowing preceding events with value. We studied the cognitive basis of intracranial self-stimulation (ICSS), a setting where dopamine appears to be valuable. Physiological frequencies seen during reinforcement learning did not support robust ICSS or promote behavior that would indicate the stimulation was represented as a meaningful reward in a specific or general sense. This was despite demonstrating that this same physiologically-relevant signal could function as a teaching signal. However, supraphysiological frequencies supported robust ICSS where the stimulation was represented as a specific sensory event, which acted as a goal to motivate behavior. This demonstrates that dopamine neurons only support ICSS at supraphysiological frequencies, and in a manner that does not reflect our subjective experience with endogenous firing of dopamine neurons during reinforcement learning.

Main Text

For many decades intracranial self-stimulation (ICSS) has been used to investigate the neural substrates of motivation¹⁻⁹. The finding that many species, including humans, will work to receive stimulation of various brain regions has been taken as evidence that those regions encode natural rewards or pleasure^{5,6,8,10}. Regions that support ICSS predominantly rest across the medial forebrain bundle, but can also be found in many parts of the limbic and extrapyramidal motor systems^{1,10-13}. The use of this paradigm for studying the neural substrates of reward has increased considerably in the last decade, afforded in part by the optogenetic revolution that allows for incredibly precise control over neuronal activity in rodents¹⁴⁻¹⁸. This has extended our knowledge of the particular neuronal populations responsible for the ICSS effect, down to the specific projection targets that subserve these behaviors^{12,13,15,19-25}.

One of the primary neuronal populations that promote ICSS are midbrain dopamine neurons. Rodents, primates, and even humans will work vigorously to receive stimulation of dopamine neurons^{3,4,7,23,25-28}. This has been taken as strong evidence in favor of the value hypothesis of dopamine, which argues that phasic dopamine assigns scalar value to external stimuli or actions antecedent to dopamine firing to encourage motivated behavior directed towards rewards²⁹⁻³⁴. The value hypothesis of dopamine has been a long-standing dogma in the field. However, recent findings have suggested that this is not the whole story. Specifically, many published results now show that dopamine transients act as teaching signals to help us form associative maps of how stimuli relate to one another^{22,35-44}, without making those stimuli valuable³⁶. These data call into question the idea that phasic dopamine activity is solely dedicated to value assignment and require a more nuanced model, which is an endeavor reflected by several recent efforts in the field^{45,46}.

Apparent differences in the function of phasic dopamine may also result from the way in which dopamine neuronal activity is manipulated in different settings. For example, during ICSS, stimulation of dopamine neurons is delivered when the subject has completed an action. Yet the brief increases in the phasic firing of dopamine neurons that characterize the prediction error are generally seen at the unexpected transition between two events (e.g. cues and rewards, or actions and rewards)^{29,47-50}, and not at the completion of an inconsequential action. Effectively, the way a prediction error acts when couched in the context of learning to associate two things together in the physiological world is likely different to how it might function to produce ICSS³⁶. The second is that the quintessential dopamine prediction error signal, famously revealed by Schultz and colleagues in the nineties²⁹, is very brief and typically observed to be within the range of 10-20Hz in primates and rodents^{29,47,51}. However, ICSS stimulation parameters are generally much higher than the range seen during reinforcement learning; conventionally, dopamine stimulation with ICSS is around 50Hz^{4,7,52,53}, and studies often use frequencies above even this supraphysiological range^{2,4,28,52}.

To begin to understand why phasic activity of dopamine neurons subserve different functions in ICSS relative to reinforcement learning, we tested the cognitive basis of ICSS of dopamine neurons using both learning-relevant and supraphysiological frequencies of dopamine stimulation. This would allow us to simultaneously ask whether dopamine stimulation acts as a general value signal to produce ICSS, consistent with the value hypothesis²⁹, and if the role of dopamine in ICSS varies with stimulation parameters. Surprisingly, despite the thousands of manuscripts that have demonstrated ICSS, there are very few findings that have investigated the question of *how* this stimulation is represented in the brain. To test the nature of the representation supporting ICSS, we investigated how dopamine stimulation would interact with the well-studied

Pavlovian-to-Instrumental transfer (PIT) effect, which is seen in both rodents and humans, and allows for a differentiation between specific and general representations of rewards⁵⁴⁻⁵⁸. Specifically, PIT allows us to determine whether a reward is supporting learning because it acts as an internal representation of a goal, or whether it simply reinforces a response via a more general valence mechanism. The PIT procedure involves first teaching subjects that two cues lead to different types of rewards (e.g., grain pellets or sucrose). Then subjects learn separately that they can work to earn these two rewards by performing two different actions (e.g., a left or right lever press). Finally, subjects are given a test where the cues are played and two actions are available but do not produce their associated reward. This test is what allows us to probe the nature of the reward representation that contributed to learning in the earlier phases of the task. Specifically, if the reward was capable of driving learning by evoking a sensory-specific representation of itself, then presenting the cue should motivate the subject to perform the action associated with the same reward. For example, if the pellet-predictive cue is presented, it will make the subject think of the pellet, and motivate them to press the lever that is also associated with pellets. However, if learning was devoid of a specific representation of reward identity, and driven solely by a more general value mechanism, cue presentation will motivate lever-press responding in a non-specific way, increasing responding on both levers^{56,58}. Generally, food rewards will support selective PIT, indicating the presence of a specific association over and above a more general value mechanism^{56,58}. Armed with this knowledge, we redesigned this selective PIT design in rats to include dopamine stimulation as one of the rewards and compared this with a food reward. Given the value hypothesis argues that the dopamine error signal seen to unexpected food rewards reflects the magnitude of the reward's scalar value²⁹, this would allow us to directly compare how the signal functions as a reward in its own right relative to the reward it accompanies. We included two groups, one that received dopamine stimulation approaching physiological levels during reinforcement learning (20Hz)^{36,37,51}, and another at supraphysiological rates that are typically used to show ICSS (50Hz)^{4,7}. This allowed us to test how if dopamine stimulation is represented as a specific reward within the brain, how it compares to a natural food reward, and whether this differs for physiological and supraphysiological frequencies.

Prior to training, we infused 2ul of a Cre-dependent excitatory opsin channelrhodopsin (ChR2; pAAV5-Ef1a-DIO-hChR2(E123T/T159C)-eYFP) bilaterally into the ventral tegmental area (VTA) of male and female rats expressing Cre-recombinase under the control of the tyrosine hydroxylase (TH) promoter^{36,37} ($n=11$; Fig 1A-C; for power analyses see methods and Fig S1). During this surgery, we also implanted optic fibers bilaterally into the VTA. This allowed us to stimulate activity in dopamine neurons in the VTA via blue light delivery (473nm, 16mW, 1s)¹⁶. Four weeks after surgery, rats were placed on mild food restriction. Then, training began with presentations of two 2-min auditory cues (white noise or click) on separate trials, intermixed within a session, across 10 days. During presentation of these cues, the associated reward was delivered 4 times, randomly distributed through cue presentation. All rats received one cue paired with dopamine stimulation (20Hz or 50Hz) and the other with sucrose pellets, counterbalanced. We recorded locomotor activity as well as the number of entries rats made into the food port during cue presentation (Fig 1G and J). Locomotor activity increased across sessions for both the dopamine and food-paired cues (main effect, session: $F_{4,36}=8.595$, $p<0.0001$), with no difference in the degree of locomotor activity evoked by the cues (main effect, dopamine vs. food cues: $F_{1,9}=0.075$, $p=0.790$), or any between-group difference in locomotor activity (main effect, 20Hz vs. 50Hz groups: $F_{1,9}=0.482$, $p=0.505$; Fig 2A and D). Unsurprisingly, entries into the food port were only seen for the food-paired cue and not the dopamine-paired cue. A repeated-measures ANOVA on the food-port-entry data showed a significant main effect of cue (dopamine vs. food: $F_{1,9}=13.221$, $p=0.005$), with no interaction by group ($F_{1,9}=0.105$, $p=0.753$). Similarly, there was a main effect of session ($F_{4,36}=10.425$, $p=0.000$), and a session by cue interaction ($F_{4,36}=9.196$, $p=0.000$) with no interactions by group (session x group: $F_{4,36}=0.522$, $p=0.720$; session x cue x

group: $F_{4,36}=0.262$, $p=0.900$). Finally, there was no between-group difference in overall responding in the food port to the cues ($F_{1,9}=0.106$, $p=0.752$). This shows that locomotor activity increased across learning to both cues in both groups, and only the food-paired cue promoted entries into the food port.

After this training, rats moved onto the instrumental contingencies where they could press one lever to receive dopamine stimulation, and another to receive sucrose pellets. During instrumental training, one lever was presented at a time and rats could earn a maximum of 40 rewards, in line with other PIT studies^{59,60}. This also helped us to prevent competition between the two instrumental actions during learning. At first, it only took one lever press to earn these rewards (fixed ratio 1; FR1), then rats needed to increase the number of lever presses to an average of five (random ratio 10; RR5) and then ten (RR10) to earn the reward. This progressive schedule is typically used to establish a robust baseline of performance that will withstand the PIT test that takes place without reward deliveries^{59,60}. During instrumental training, significant differences between our 20Hz and 50Hz groups emerged (Fig 1H and K). Specifically, while our 20Hz group reached criterion on FR1 schedules earning the stimulation, their performance was not robust when we moved to leaner schedules. This was despite their maintenance of responding on these schedules for the food reward. However, rats in the 50Hz group continued to show robust performance for both rewards as instrumental training progressed. This was confirmed with statistical analyses, where a repeated-measures ANOVA demonstrated a significant main effect of reward (dopamine vs. food: $F_{1,9}=21.528$, $p=0.001$), and a reward x group interaction ($F_{1,9}=21.544$, $p=0.001$), owed to a significant difference in responding for dopamine stimulation between the 20Hz and 50Hz groups ($F_{1,9}=30.738$, $p=0.000$), which was not present when comparing responding for food rewards ($F_{1,9}=0.128$, $p=0.729$). This same pattern was also seen when considering the escalation of responding across sessions, which returned a main effect of session ($F_{7,63}=149.418$, $p=0.000$), a session by group interaction ($F_{7,63}=9.413$, $p=0.000$), owed to a difference in responding emerging between groups as instrumental training progressed (e.g., session 8: $F_{1,9}=7.629$, $p=0.022$). There was also a between-group difference in overall levels of responding ($F_{1,9}=15.647$, $p=0.003$). Thus, while 50Hz dopamine stimulation functioned to motivate vigorous responding that was comparable to food rewards, consistent with other reports^{4,7,17,18,25,52}, 20Hz of dopamine stimulation did not support robust instrumental responding beyond a continuously-reinforced schedule and was not comparable to a food reward.

Following training, rats received the PIT test where the dopamine- and food-paired cues were played and rats had the opportunity to press either lever. During this test, neither the cues or the levers produced their associated reward, which allowed us to investigate the nature of the reward representation that was present during learning in the earlier stages of the task, in the absence of reward feedback^{54,56}. In our 20Hz group, we found that the dopamine-paired cue did not produce specific PIT (Fig 1I). Specifically, rats did not elevate their responding on either lever when the dopamine-paired cue was presented, relative to baseline. However, when the food-paired cue was presented, these same rats selectively elevated responding on the lever that produced the food, demonstrating selective PIT. In contrast, in our 50Hz group, we found that the dopamine-paired cue supported robust selective PIT, which was comparable to that produced by the food-paired cue (Fig 1L). This was supported with statistical analyses. A repeated measures ANOVA comparing responding in our 20Hz group on the same and different levers, relative to the baseline period prior to cue presentation, revealed a cue by lever interaction (cue x lever: $F_{1,5}=15.383$, $p=0.011$), owed to a selective increase on the same lever in response to the food-paired cue (same vs. diff: $F_{1,5}=13.957$, $p=0.013$) that was not present for the dopamine cue ($F_{1,5}=3.759$, $p=0.110$). Importantly, there was also a significant difference when comparing responding on the same lever for the dopamine- vs. the food-paired cue ($F_{1,5}=7.885$, $p=0.038$), which was not seen for the different lever ($F_{1,5}=3.243$, $p=0.132$). This same analysis conducted

on responding exhibited by our 50Hz group revealed a main effect of lever (same vs. different: $F_{1,4}=20.193$, $p=0.011$), but no main effect of cue (dopamine vs. food: $F_{1,4}=0.585$, $p=0.487$), or any interaction between these factors ($F_{1,4}=0.505$, $p=0.516$). Thus, stimulation of dopamine at physiologically-relevant levels (i.e., 20Hz) did not function as a specific reward that was capable of supporting instrumental lever pressing in the same way that a natural reward could, and did not endow the cue with motivational significance that could support PIT. However, stimulating dopamine at higher ranges (i.e., 50Hz), typically used for ICSS studies^{4,52}, supported robust instrumental responding on leaner schedules and selective PIT. This demonstrated that stimulation of dopamine neurons at supraphysiological frequencies during ICSS produced a supraphysiological sensory event that was capable of acting as a representation of a specific reward to motivate behavior, over and above any role for this signal in endowing antecedent cues with general value.

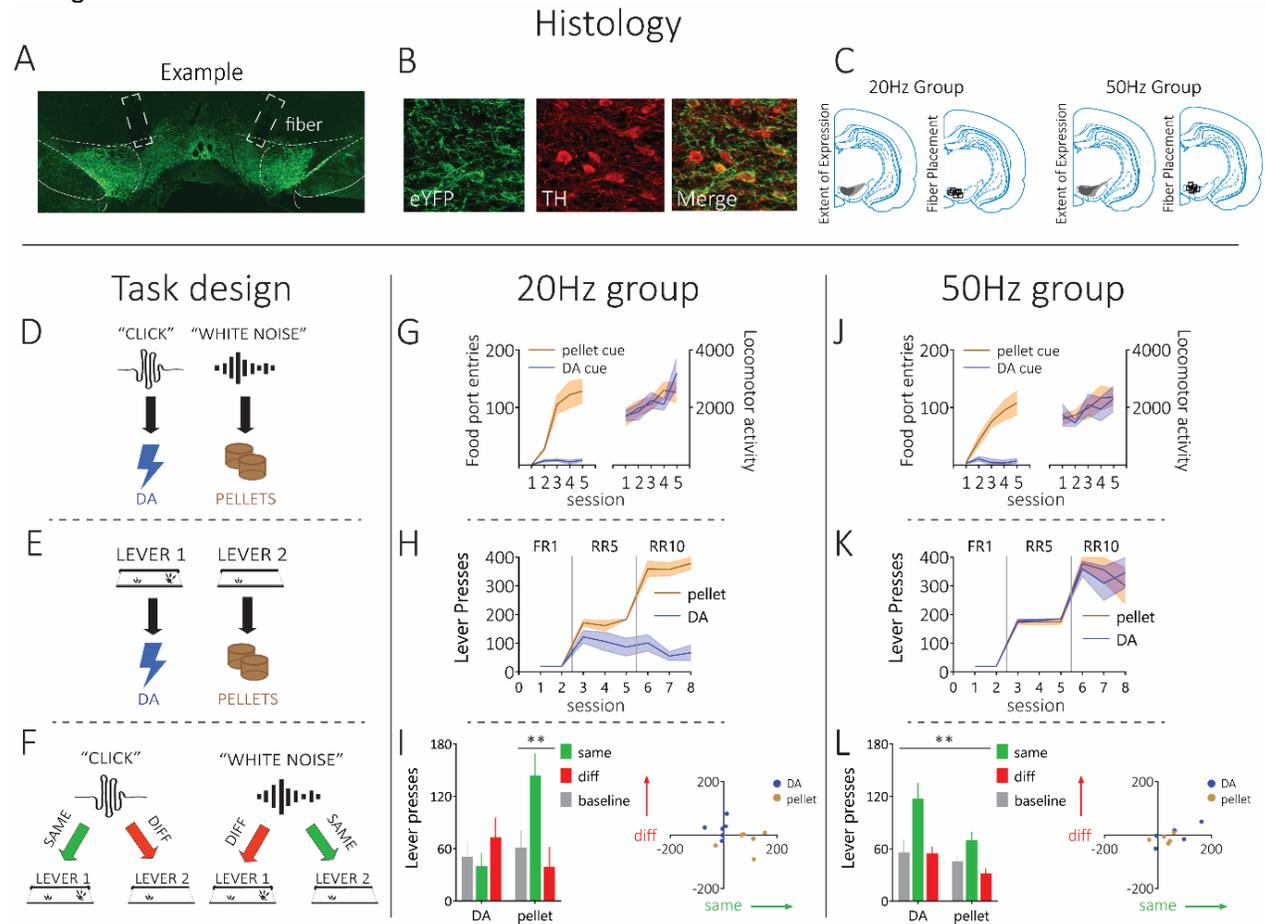


Figure 1. A physiologically-relevant frequency of dopamine stimulation (20Hz) does not function as a meaningful reward, however, high-frequency dopamine stimulation (50Hz) functions as a reward that is encoded as a specific sensory event. Top: Histological verification with **A**) bilateral Cre-dependent ChR2 expression in TH-Cre rats, **B**) colocalization of TH and virus expression approached ~90%, and **C**) schematic of minimum and maximum virus expression and fiber placement. Left column: Schematic illustrating the task design using one counterbalancing example, which consisted of **D**) Pavlovian conditioning, **E**) Instrumental Conditioning, and **F**) the PIT test. Rats first learnt that two auditory cues (e.g. click and white noise) led to two outcomes (e.g. dopamine stimulation and pellets), then they learnt to perform two lever presses that led to the two outcomes. Finally, rats were presented with the two auditory cues and given an opportunity to press either lever, without reward feedback. Middle column: **G**)

Rats in the 20Hz group ($n=6$) showed an increase in food-port entries during the pellet-paired stimulus, but not the dopamine-paired stimulus. These rats showed equivalent increases in locomotor activity across learning to both stimuli. **H**) During instrumental conditioning, where rats learned to make lever presses for the two outcomes, rats in the 20Hz group showed robust lever-pressing responses for the pellets, but not the dopamine stimulation. **I**) In the final PIT test, when the pellet-paired cue is presented, these rats showed significant elevations in responding on the pellet-paired lever, indicating specific PIT. However, they did not show PIT for the dopamine-paired cue. Right column: **J**) Rats in the 50Hz group ($n=5$) showed increases in food port entries during the pellet-paired stimulus, but not the dopamine-paired stimulus. Increases in locomotor activity across learning was similar for both the dopamine- and pellet-paired stimulus. **K**) During instrumental training, the 50Hz group showed robust lever pressing for both the dopamine stimulation and the pellets. **L**) In the final PIT test, the dopamine- and pellet-paired stimuli both produced robust specific PIT. Error bars =SEM.

Next, we wanted to confirm that our 20Hz stimulation of dopamine neurons was effective. To do this, we ran an additional experiment with a new cohort of rats to test whether our 20Hz stimulation of dopamine was capable of acting as a teaching signal to produce learning, as we and others have previously demonstrated^{22,36,37,61}. Here, we used the blocking procedure. We first trained rats that two light cues led to two different rewards (Figure 2 “acquisition”; e.g. house light → sucrose pellet, flash → grain pellet; counterbalanced). Then, we presented the lights in compound with novel auditory cues, which produced the same rewards (e.g. Figure 2 “blocking”; house light + white noise → sucrose pellet, flash + click → grain pellets; counterbalanced). Usually, subjects do not associate the novel auditory cues with the rewards, because the rewards are already predicted by the lights, termed blocking⁶². However, we stimulated dopamine neurons (1s 20Hz) as one of the rewards was delivered. This created an “unblocked cue”, which was the cue paired with dopamine stimulation during its associated reward. We could then compare with our “blocked cue”, which was the cue that was not paired with stimulation of dopamine neurons. All rats learnt to enter the port while the visual cues were presented, increasing above baseline across learning, which was unaffected by introduction of the auditory cues or dopamine stimulation (Figure 2C; time period: $F_{2,16}=11.072$, $p=0.001$, session: $F_{5,40}=1.389$, $p=0.249$; time period x session: $F_{10,80}=2.710$, $p=0.006$; unblocked vs. baseline: $p=0.006$; blocked vs. baseline: $p=0.003$, unblocked vs. blocked: $p=0.140$). This demonstrated that all rats learnt that the visual cues were predictive of reward, and that stimulating dopamine neurons did not interfere with the ability to continue to respond during the food-predictive cues.

We then tested responding to the auditory cues alone without reward to assess how much learning had accrued towards them (Figure 2 “probe test”). Here, we saw that rats spent more time in the food port during the unblocked cue relative to the blocked cue (Figure 2D; $F_{1,8}=3.614$, $p=0.047$). This showed that stimulating dopamine neurons was capable of driving additional learning to the stimulation-paired unblocked cue. To further probe the nature of the association that had developed via dopamine stimulation, we devalued the reward paired with the unblocked cue (Figure 2 “devaluation”). To do this, we allowed all rats to consume the reward paired with the unblocked cue outside the experimental chamber. Then, we immediately injected half the rats with lithium chloride (LiCl) to induce gastric malaise (devalued group). In the other rats, we injected LiCl 6 hours after consumption of the reward (non-devalued group). This would allow us to test whether the learning produced by dopamine stimulation involved a sensory-specific representation of the reward. Specifically, we could then test responding to the unblocked cue again and see if it was impacted by devaluation. During this test, we found that rats in the devalued group spent less time in the food port during the cue (Figure 2E; group: $F_{2,7}=5.15$, $p=0.029$). This was an effect not seen when we tested responding to the blocked cue, whose associated reward had not been devalued, confirming the sensory-specific nature of the effect (Fig S4, left; group:

$F_{2,7}=0.157$, $p=0.352$). Together, these data show that our physiologically-relevant stimulation parameters were capable of acting as a teaching signal to drive sensory-specific associations between events, despite not being able to function as a reward in itself.

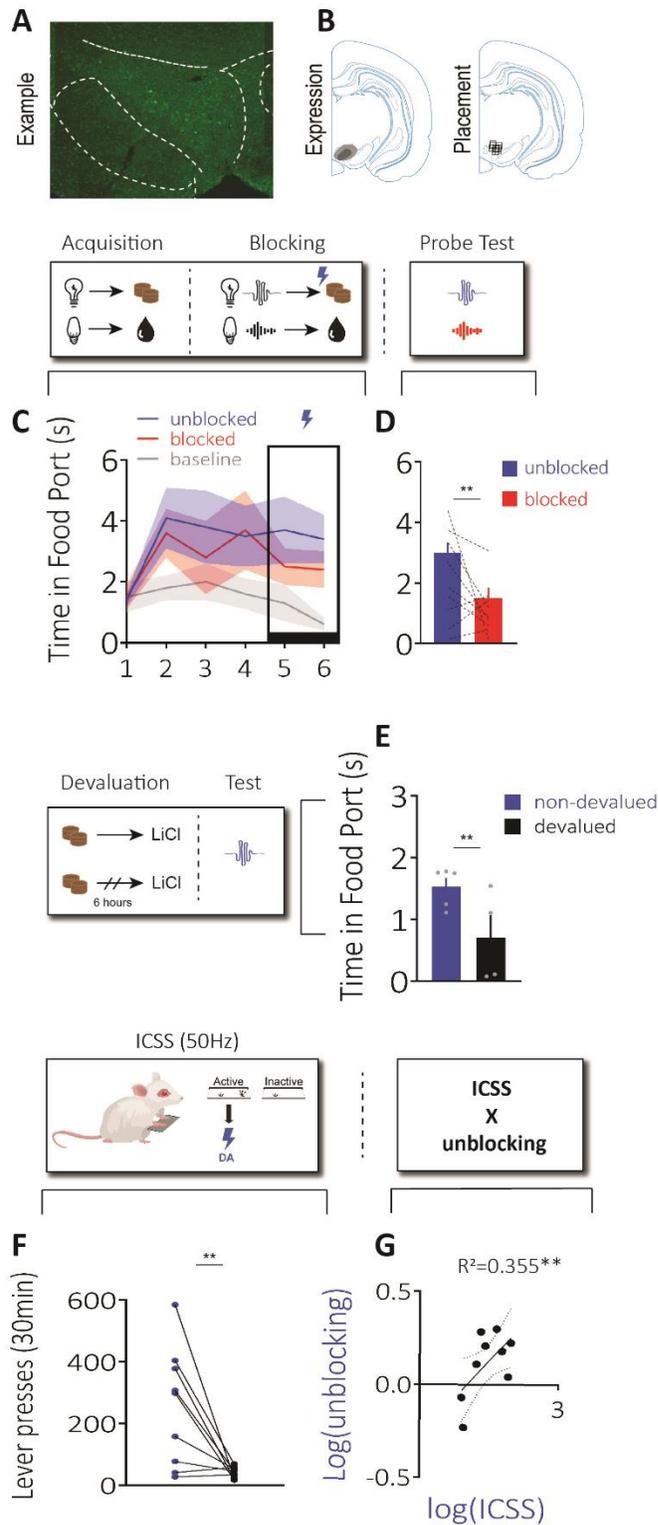


Figure 2. A physiologically-relevant frequency of dopamine stimulation (20Hz) functions as a teaching signal to drive sensory-specific learning. **A)** Unilateral example of bilateral ChR2 expression in dopamine neurons, **B)** Left: minimum and maximum expression of virus across rats, Right: placement of fiber tips across rats. Top schematic: design of our blocking task using one counterbalancing example, which consisted of acquisition, blocking, and a probe test. **C)** Rats ($n=9$) first learnt that two visual cues led to two distinct rewards (acquisition), then two novel auditory stimuli were introduced in compound with the visual cues and led to the same rewards (blocking). During blocking, we stimulated dopamine neurons (1s, 5ms, 20Hz) at reward delivery for one of the compounds (“unblocked” cue). Rats acquired the food-port response during the cues and maintained high levels of responding after introduction of the auditory cues and dopamine stimulation. **D)** We next gave rats a probe test and found they responded more to the unblocked cue, relative to blocked cue. Middle schematic: devaluation procedure used to devalue the reward paired with the unblocked cue. **E)** Rats showed a significant reduction in responding to the unblocked cue following devaluation, confirming dopamine stimulation unblocked the association between the cue and a sensory-specific representation of reward. Bottom schematics: design used for ICSS, rats were given access to a lever that produced 50Hz stimulation of dopamine neurons (active), or nothing (inactive). **F)** Rats pressed more on the active lever, **G)** This ICSS measure was positively correlated with the unblocking effect in B ($r=0.56$; $p=0.045$). Error = SEM.

In these same rats, we next wanted to ask whether the rewarding effects of high-frequency stimulation are related to the ability of these neurons to act as a teaching signal. To test this, we examined how much the rats would press to get optogenetic stimulation of dopamine neurons at 50Hz (Figure 3 “ICSS”). Unsurprisingly, we found that rats would press consistently for 50Hz of dopamine stimulation (Figure 3F: $F_{1,8}=10.42$, $p=0.006$). We hypothesized that the degree of lever presses for 50Hz stimulation would be positively correlated to the unblocking effect (Fig 2D). Indeed, when we compared the magnitude of the unblocking with the degree to which high-frequency dopamine stimulation would support ICSS, we found that these factors were strongly positive correlated (Fig 2G: Pearson $r = 0.596$, $R^2=0.355$, $p=0.045$). We did not see this same relationship with ICSS and the blocked cue (Fig S2, right: Pearson’s $r=0.148$, $R^2=0.022$; $p=0.352$). These data support a conclusion that the differential effects of 20Hz and 50Hz stimulation parameters are due to the change in frequency of firing of these neurons, and not because these different frequencies are tapping into different populations of dopamine neurons.

As a final experiment, we wanted to establish how VTA dopamine neurons were responding to our differential stimulation parameters. We infused a Cre-dependent ChR2 into the VTA of TH-rats. During this surgery, rats were implanted with a 16-channel microwire array with a central optic fiber into the VTA, which would afford simultaneous stimulation and recording of dopamine neurons. Similarly to prior results¹⁷, we found that the 20Hz stimulation produced firing rates of greater fidelity to the stimulation parameters than 50Hz (Figure 3A; $F_{1,6}=12.38$, $p=0.013$). Interestingly, the firing rates of the neurons did not significantly differ overall across the 1s stimulus train (Figure 3B). However, there was a significant difference in the rate of decline in fidelity between the stimulation parameters across the stimulus train. There was no difference in the spike probability as a result of a pulse within the first 5 light pulses between the 20Hz and 50Hz stimulation parameters (Figure 3C-D; $F_{1,6}=4.7$, $p=0.074$). However, a significant difference emerged towards the end of the stimulus train (Figure 3D; $F_{1,6}=11.61$, $p=0.014$). These data show that both stimulation parameters produce a similar number of action potentials, but that the 50Hz stimulation parameters produce them across a shorter period of time. This demonstrates that it is the higher frequency, and not the greater number of spikes *per se* that generate the difference in the ability of these stimulation parameters to drive reinforcement.

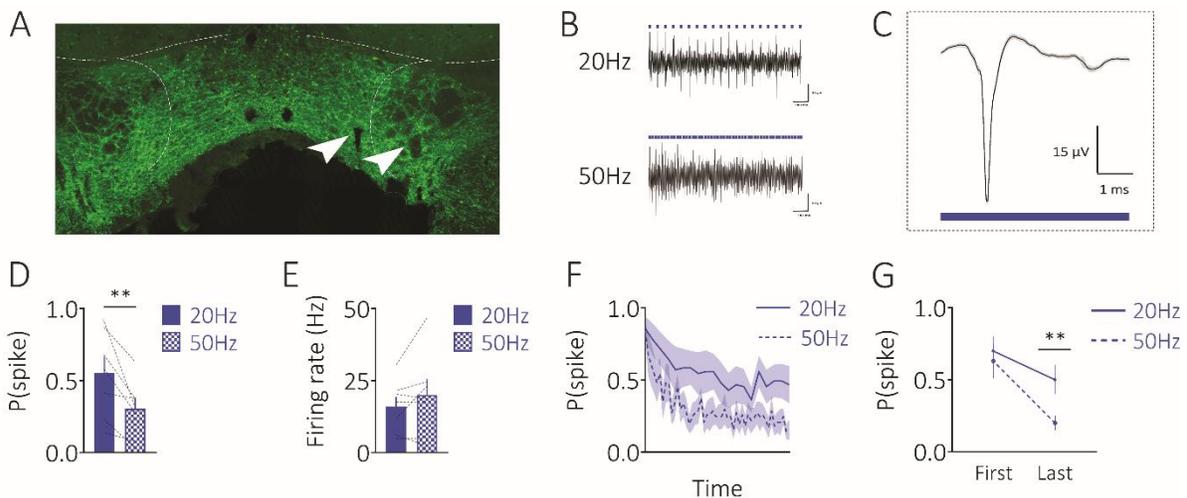


Figure 3. Optogenetic stimulation of VTA dopamine neurons at 20Hz and 50Hz results in the same number of action potentials across different timescales. Simultaneous recording and stimulation of VTA dopamine neurons reveal the firing dynamics produced by the two different stimulation trains. **A)** Example of bilateral Cre-dependent ChR2 expression (green) and

placement of microwires, **B)** example trace of 20Hz and 50Hz stimulation trains, **C)** typical triphasic DA extracellular spike, **D)** neuronal activity follows the 20Hz stimulus train more faithfully than 50Hz, **E)** there is no difference in the overall number of action potentials generated by the 20Hz and 50Hz stimulus train, **F)** the rate of neuronal firing in VTA dopamine neurons decreases more rapidly across the 50Hz stimulation train relative to 20Hz, **G)** the difference in spike probability between the stimulus trains emerges towards the end of the stimulus train.

Here, we investigated the psychological basis of ICSS with dopamine neurons. Somewhat surprisingly, there are very few studies that have asked this question, despite the high volume of manuscripts that have demonstrated ICSS throughout the brain. Using the PIT procedure, we asked whether stimulation of dopamine neurons functions to create a sensory-specific reward in its own right, much like a food reward, over and above the potential for this signal to endow the lever press or cue with general value. On this experimental backdrop, we also manipulated the frequency of stimulation between groups to see if a learning-relevant frequency of dopamine stimulation would function in the same way as high-frequency stimulation that is usually used to support ICSS. We found that 20Hz dopamine stimulation, relevant to physiological rates of phasic dopamine firing during learning^{36,37,47,51}, would not support robust ICSS or serve as a sensory-specific event that would support the PIT effect. This was despite these same rats showing a significant selective PIT effect for a food reward. In contrast, 50Hz dopamine stimulation did support robust ICSS, even when the stimulation was made harder to receive on our leaner reinforcement schedules. The cue paired with 50Hz dopamine stimulation also supported selective PIT, where the dopamine-paired cue selectively increased responding on the lever that produced dopamine stimulation.

It is worth noting that we did not see any evidence for general value accumulating to the dopamine-paired cue or instrumental response in either our 20Hz or 50Hz group. That is, if dopamine did endow the cues or actions with general value, and particularly if this was the dominant mechanism, we would have expected the dopamine-paired cue to elevate responding on either lever above baseline, indicating the ability of this cue to generally invigorate instrumental responding in non-specific ways⁵⁸. This is not what we saw. Instead, we saw in the 20Hz group that the dopamine-paired cue did not increase responding on either lever above baseline, and in our 50Hz group the dopamine-paired cue exclusively elevated lever-press responding on the same, congruent lever, and not on the different, incongruent lever. It is possible, however, that the presence of counterfactual information occluded any effect of general value in our PIT test⁶³. Specifically, it has been demonstrated that Pavlovian training also produces inhibitory associations between the current cue that is being presented, and the outcome that is not currently being predicted by that cue (e.g., presentation of the dopamine-paired cue could signal the absence of the food), which reduces the likelihood that rats will increase responding on the incongruent lever. If this was the case, then general value could have been present, but this was overridden by the specific associations promoting the congruent action, and inhibiting the incongruent action. Given that we did not find evidence for specific PIT in our 20Hz group (i.e. there was no increase on the same lever or suppression on the different lever relative to baselines pressing rates), this is an unlikely reason that we did not see evidence of general value in this group. While it is possible that the specific associations occluded the ability to see general value accumulating to the cue and action in our 50Hz group, what is clear is that general value was not the dominant mechanism driving learning in this task and could not overshadow the specific associations. At the very least, these results suggest that the ability of supraphysiological dopamine stimulation to support ICSS does not necessarily support the value hypothesis of dopamine, and provide no evidence for this hypothesis at a physiologically-relevant frequency.

Importantly, we performed a control experiment to show that our stimulation was functioning effectively. Specifically, we showed that 20Hz of dopamine stimulation could produce learning, which we and others have shown previously^{22,36,37,61}. In line with this, we found that 20Hz stimulation of dopamine was capable of producing learning about a sensory-specific cue-food associations when delivered as a prediction error during blocking. We showed this in both a novel cohort (Figure 2) as well as in the same rats that has received PIT (Figure S3). This demonstrated that our stimulation parameters were functioning effectively to produce a physiologically-relevant signal that could drive learning, and ineffective stimulation during PIT could not be a reason we found that 20Hz would not promote robust ICSS or the PIT effect. It is interesting to consider why 20Hz stimulation would produce any lever-press responding on a continuously-reinforced schedule, and also nose-poke responding in additional control experiments in line with other published findings¹⁴ (Fig S4), if it was not sufficient to support robust ICSS on leaner reinforcement schedules, or the PIT effect. This suggests that stimulation of dopamine neurons at 20Hz was reinforcing enough to promote some instrumental responding. Again, while we cannot rule out that this signal was generally valuable in some sense, what is clear is this value was not sufficient to support the kind of instrumental responding we see with other appetitive rewards, like food, water, or drugs of abuse. Indeed, we have explicitly demonstrated that a 20Hz signal delivered as a prediction error during learning will drive learning without making the antecedent cue valuable³⁶, consistent with these data and explicitly against the formal predictions of the value hypothesis of dopamine as formalized by the model-free reinforcement-learning algorithm in temporal-difference reinforcement learning (TDRL).

These results are important for two reasons. Firstly, they show that frequency of dopamine stimulation relevant to reinforcement learning (i.e., 20Hz) does not function as a reward in and of itself in that it will not support robust ICSS or specific PIT. We also did not see evidence that the dopamine-paired cue generally invigorated responding on either lever during our PIT test, consistent with an inability of this signal to endow cues with general value, mirroring our previous work³⁶. This is particularly pertinent to the field right now as we navigate new complex models to account for the seemingly discrepant function of dopamine neurons in different settings. These data now show that even in the context of ICSS, a prediction-error signal does not seem to function as a reward that possesses sufficient reinforcing properties to underlie that which is contained in other appetitive rewards, like food or drugs of abuse, necessary to uphold the value hypothesis. Secondly, these results also shed light on the psychological basis of ICSS using high frequencies of dopamine stimulation. Specifically, these data demonstrate that use of high-frequency dopamine stimulation in the context of ICSS functions to create a sensory event that acts as a specific reward in its own right. This does not have a basis in our everyday learning experience. Put simply, our learning experience does not contain representations of phasic dopamine as a rewarding sensory event. This questions the assumption that the ability of dopamine neurons to support ICSS at high frequencies indicates something about the region's role in reinforcement learning. More generally, it begs the question: is there any physiological experience that might relate to ICSS? Perhaps a circumstance where this may become relevant is drug seeking^{25,64}, where most drugs of abuse act in at least some way to increase phasic dopamine activity⁶⁵⁻⁶⁷. And if so, these data support an idea that people with substance use disorder seek out drugs of abuse to obtain a specific sensory experience, not because the actions or cues associated with the drug have become valuable.

References and notes

- 1 Carlezon Jr, W. A. & Chartoff, E. H. Intracranial self-stimulation (ICSS) in rodents to study the neurobiology of motivation. *Nature protocols* **2**, 2987 (2007).
- 2 Bishop, M., Elder, S. T. & Heath, R. G. Intracranial self-stimulation in man. *Science* **140**, 394-396 (1963).
- 3 Fibiger, H., LePiane, F., Jakubovic, A. & Phillips, A. The role of dopamine in intracranial self-stimulation of the ventral tegmental area. *Journal of Neuroscience* **7**, 3888-3896 (1987).
- 4 Corbett, D. & Wise, R. A. Intracranial self-stimulation in relation to the ascending dopaminergic systems of the midbrain: a moveable electrode mapping study. *Brain research* **185**, 1-15 (1980).
- 5 Olds, M. & Fobes, J. L. The central basis of motivation: intracranial self-stimulation studies. *Annual review of psychology* **32**, 523-574 (1981).
- 6 Margules, D. & Olds, J. Identical "feeding" and "rewarding" systems in the lateral hypothalamus of rats. *Science* **135**, 374-375 (1962).
- 7 Garris, P. A. *et al.* Dissociation of dopamine release in the nucleus accumbens from intracranial self-stimulation. *Nature* **398**, 67-69 (1999).
- 8 Olds, J. & Milner, P. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of comparative and physiological psychology* **47**, 419 (1954).
- 9 Hoebel, B. G. & Teitelbaum, P. Hypothalamic control of feeding and self-stimulation. *Science* **135**, 375-377 (1962).
- 10 Smith, K. S., Mahler, S. V., Peciña, S. & Berridge, K. C. Hedonic hotspots: Generating sensory pleasure in the brain. (2010).
- 11 Shizgal, P. & Murray, B. Neuronal basis of intracranial self-stimulation. (1989).
- 12 Britt, J. P. *et al.* Synaptic and behavioral profile of multiple glutamatergic inputs to the nucleus accumbens. *Neuron* **76**, 790-803 (2012).
- 13 Stuber, G. D. *et al.* Excitatory transmission from the amygdala to nucleus accumbens facilitates reward seeking. *nature* **475**, 377-380 (2011).
- 14 Witten, I. B. *et al.* Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron* **72**, 721-733 (2011).
- 15 Nieh, E. H. *et al.* Inhibitory input from the lateral hypothalamus to the ventral tegmental area disinhibits dopamine neurons and promotes behavioral activation. *Neuron* **90**, 1286-1298 (2016).
- 16 Deisseroth, K. Optogenetics. *Nature methods* **8**, 26-29 (2011).
- 17 Tsai, H.-C. *et al.* Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science* **324**, 1080-1084 (2009).
- 18 Bass, C. E. *et al.* Optogenetic control of striatal dopamine release in rats. *Journal of neurochemistry* **114**, 1344-1352 (2010).
- 19 Cole, S. L., Robinson, M. J. & Berridge, K. C. Optogenetic self-stimulation in the nucleus accumbens: D1 reward versus D2 ambivalence. *PLoS One* **13**, e0207694 (2018).
- 20 Urstadt, K. R. & Berridge, K. C. Optogenetic mapping of feeding and self-stimulation within the lateral hypothalamus of the rat. *Plos one* **15**, e0224301 (2020).
- 21 Rossi, M. A., Sukharnikova, T., Hayrapetyan, V. Y., Yang, L. & Yin, H. H. Operant self-stimulation of dopamine neurons in the substantia nigra. *PLoS One* **8**, e65799 (2013).
- 22 Keiflin, R., Pribut, H. J., Shah, N. B. & Janak, P. H. Ventral tegmental dopamine neurons participate in reward identity predictions. *Current Biology* **29**, 93-103. e103 (2019).
- 23 Ilango, A. *et al.* Similar roles of substantia nigra and ventral tegmental dopamine neurons in reward and aversion. *Journal of Neuroscience* **34**, 817-822 (2014).

- 24 Stuber, G. D. & Wise, R. A. Lateral hypothalamic circuits for feeding and reward. *Nature neuroscience* **19**, 198 (2016).
- 25 Pascoli, V., Terrier, J., Hiver, A. & Lüscher, C. Sufficiency of mesolimbic dopamine neuron stimulation for the progression to addiction. *Neuron* **88**, 1054-1066 (2015).
- 26 Arsenault, J. T., Rima, S., Stemmann, H. & Vanduffel, W. Role of the primate ventral tegmental area in reinforcement and motivation. *Current Biology* **24**, 1347-1353 (2014).
- 27 Owesson-White, C. A., Cheer, J. F., Beyene, M., Carelli, R. M. & Wightman, R. M. Dynamic changes in accumbens dopamine correlate with learning during intracranial self-stimulation. *Proceedings of the National Academy of Sciences* **105**, 11957-11962 (2008).
- 28 Heath, R. G. Electrical self-stimulation of the brain in man. *American Journal of Psychiatry* **120**, 571-577 (1963).
- 29 Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593-1599 (1997).
- 30 Hamid, A. A. *et al.* Mesolimbic dopamine signals the value of work. *Nature neuroscience* **19**, 117 (2016).
- 31 Schultz, W. Getting formal with dopamine and reward. *Neuron* **36**, 241-263 (2002).
- 32 Lak, A., Stauffer, W. R. & Schultz, W. Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proceedings of the National Academy of Sciences* **111**, 2343-2348 (2014).
- 33 Fiorillo, C. D. Two dimensions of value: dopamine neurons represent reward but not aversiveness. *Science* **341**, 546-549 (2013).
- 34 Dabney, W. *et al.* A distributional code for value in dopamine-based reinforcement learning. *Nature* **577**, 671-675 (2020).
- 35 Langdon, A. J., Sharpe, M. J., Schoenbaum, G. & Niv, Y. Model-based predictions for dopamine. *Current Opinion in Neurobiology* **49**, 1-7 (2018).
- 36 Sharpe, M. J. *et al.* Dopamine transients do not act as model-free prediction errors during associative learning. *Nature Communications* **11**, 1-10 (2020).
- 37 Sharpe, M. J. *et al.* Dopamine transients are sufficient and necessary for acquisition of model-based associations. *Nature Neuroscience* **20**, 735 (2017).
- 38 Sharpe, M. J. & Schoenbaum, G. Evaluation of the hypothesis that phasic dopamine constitutes a cached-value signal. *Neurobiology of learning and memory* **153**, 131-136 (2018).
- 39 Howard, J. D. & Kahnt, T. Identity prediction errors in the human midbrain update reward-identity expectations in the orbitofrontal cortex. *Nature communications* **9**, 1611 (2018).
- 40 Stalnaker, T. *et al.* Dopamine neuron ensembles signal the content of sensory prediction errors. *BioRxiv*, 723908 (2019).
- 41 Sharp, M. E., Foerde, K., Daw, N. D. & Shohamy, D. Dopamine selectively remediates 'model-based' reward learning: a computational approach. *Brain* **139**, 355-364 (2015).
- 42 Engelhard, B. *et al.* Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature*, 1 (2019).
- 43 Takahashi, Y. K. *et al.* Dopamine neurons respond to errors in the prediction of sensory features of expected rewards. *Neuron* **95**, 1395-1405. e1393 (2017).
- 44 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).
- 45 Gardner, M. P., Schoenbaum, G. & Gershman, S. J. Rethinking dopamine as generalized prediction error. *Proceedings of the Royal Society B* **285**, 20181645 (2018).
- 46 Lee, R. S., Engelhard, B., Witten, I. B. & Daw, N. D. A vector reward prediction error model explains dopaminergic heterogeneity. *bioRxiv* (2022).
- 47 Takahashi, Y. K. *et al.* The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron* **62**, 269-280 (2009).

- 48 Howard, J. D., Gottfried, J. A., Tobler, P. N. & Kahnt, T. Identity-specific coding of future rewards in the human orbitofrontal cortex. *Proceedings of the National Academy of Sciences* **112**, 5195-5200 (2015).
- 49 Kahnt, T. & Tobler, P. N. Dopamine modulates the functional organization of the orbitofrontal cortex. *Journal of Neuroscience* **37**, 1493-1504 (2017).
- 50 Waelti, P., Dickinson, A. & Schultz, W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* **412**, 43-48 (2001).
- 51 Roesch, M. R., Calu, D. J. & Schoenbaum, G. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nature neuroscience* **10**, 1615-1624 (2007).
- 52 Wise, R. A. Intracranial self-stimulation: mapping against the lateral boundaries of the dopaminergic cells of the substantia nigra. *Brain research* **213**, 190-194 (1981).
- 53 Yavich, L. & Tiihonen, J. In vivo voltammetry with removable carbon fibre electrodes in freely-moving mice: dopamine release during intracranial self-stimulation. *Journal of neuroscience methods* **104**, 55-63 (2000).
- 54 Holmes, N. M., Marchand, A. R. & Coutureau, E. Pavlovian to instrumental transfer: a neurobehavioural perspective. *Neuroscience & Biobehavioral Reviews* **34**, 1277-1295 (2010).
- 55 Nadler, N., Delgado, M. R. & Delamater, A. R. Pavlovian to instrumental transfer of control in a human learning task. *Emotion* **11**, 1112 (2011).
- 56 Cartoni, E., Balleine, B. & Baldassarre, G. Appetitive Pavlovian-instrumental transfer: a review. *Neuroscience & Biobehavioral Reviews* **71**, 829-848 (2016).
- 57 Prévost, C., Liljeholm, M., Tyszka, J. M. & O'Doherty, J. P. Neural correlates of specific and general Pavlovian-to-Instrumental Transfer within human amygdalar subregions: a high-resolution fMRI study. *Journal of Neuroscience* **32**, 8383-8390 (2012).
- 58 Corbit, L. H. & Balleine, B. W. Learning and motivational processes contributing to Pavlovian-instrumental transfer and their neural bases: dopamine and beyond. *Behavioral neuroscience of motivation*, 259-289 (2015).
- 59 Bradfield, L. A., Hart, G. & Balleine, B. W. Inferring action-dependent outcome representations depends on anterior but not posterior medial orbitofrontal cortex. *Neurobiology of learning and memory* **155**, 463-473 (2018).
- 60 Bradfield, L. A., Dezfouli, A., van Holstein, M., Chieng, B. & Balleine, B. W. Medial orbitofrontal cortex mediates outcome retrieval in partially observable task situations. *Neuron* **88**, 1268-1280 (2015).
- 61 Steinberg, E. E. *et al.* A causal link between prediction errors, dopamine neurons and learning. *Nature neuroscience* **16**, 966 (2013).
- 62 Kamin, L. J. Predictability, surprise, attention, and conditioning. (1967).
- 63 Laurent, V. & Balleine, B. W. Factual and counterfactual action-outcome mappings control choice between goal-directed actions in rats. *Current Biology* **25**, 1074-1079 (2015).
- 64 Wise, R. A. Addictive drugs and brain stimulation reward. *Annual review of neuroscience* **19**, 319-340 (1996).
- 65 Wanat, M. J., Willuhn, I., Clark, J. J. & Phillips, P. E. Phasic dopamine release in appetitive behaviors and drug addiction. *Current drug abuse reviews* **2**, 195-213 (2009).
- 66 Willuhn, I., Wanat, M. J., Clark, J. J. & Phillips, P. E. in *Behavioral neuroscience of drug addiction* 29-71 (Springer, 2010).
- 67 Phillips, P. E., Stuber, G. D., Heien, M. L., Wightman, R. M. & Carelli, R. M. Subsecond dopamine release promotes cocaine seeking. *Nature* **422**, 614-618 (2003).

Acknowledgements:

Funding: This work was supported by BBRF 30637, NSF2143910, MH126278 awarded to MJS, and DA043572 and DA045764, awarded to DJB and MHJ, respectively. Funding sources were not involved in study design, data collection, and interpretation, or in the decision to submit this research for publication. **Author contributions:** SM, KW, MJ, DB, and MJS designed the experiments. SM, IH, ZG, MHJ, and DJB conducted the experiments. SM and MJS wrote the paper with input from all authors. We thank Morgan Paladino for her assistance across the recording experiments. **Competing interests:** The authors have no competing interests to declare. **Data and materials availability:** Data are available upon reasonable request.

Supplementary Materials

Material and methods

Figs S1-S4

References

Supplementary Information

The cognitive basis of intracranial self-stimulation of midbrain dopamine neurons.

Authors: Samuel J. Millard¹, Ivy B. Hoang¹, Zara Greer¹, Shayna L. O'Connor^{2,3,4}, Kate M. Wassum¹, Morgan H. James^{2,3}, David J. Barker^{3,4}, Melissa J. Sharpe¹

Affiliations:

¹Department of Psychology, University of California, Los Angeles CA

²Department of Psychiatry, Robert Wood Johnson Medical School, Rutgers University, Piscataway NJ

³Brain Health Institute, Rutgers Biomedical Health Sciences, Rutgers University, Piscataway NJ

⁴Department of Psychology, Rutgers, The State University of New Jersey, Piscataway NJ

This PDF file includes:

Materials and Methods

Figs S1-S4

References

Materials and Methods

Subjects

Twenty-six experimentally naïve male and female Long-Evans transgenic rats carrying a Tyrosine hydroxylase (TH)-dependent Cre expressing system were used these experiments^{1,2}. Rats were approximately 4 months of age prior to surgical procedures. Rats were maintained on a 12-hour light-dark cycle, where 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 Animal Research Council at UCLA and/or Animal Care and Use Committee at Rutgers University.

Surgical Procedures

Surgical procedures have been described elsewhere^{1,2}. Briefly, rats were anesthetized with isoflurane (3%, 1-2% maintenance in 1L/m O₂) and secured in a stereotaxic device (David Kopf Instruments). Rats received two infusions of 1.0 μ L bilaterally (2.0 μ L per hemisphere) of AAV5-EF1 α -DiO-ChR2-eYFP (n=24) into the VTA at the following coordinates: AP: -5.3mm; ML: \pm 0.7mm; DV: -6.5 mm and -7.7 (females) or -7.0mm and -8.2mm (males). Virus was obtained from Addgene (MA, USA). For the behavioral experiments, optic fibers (Thorlabs, 200 μ m diameter) were implanted bilaterally at the following coordinates relative to bregma: AP: -5.3mm; ML: \pm 2.61mm and DV: -7.05mm (female) or -7.55mm (male) at an angle of 15° pointed toward the midline. For the recording experiment, a 16-channel microwire array with a central optic fiber was implanted targeting the VTA (-5 to -6.5 AP, +1.58 to 2.58 ML at 10° angle, -7.9 DV; Microprobes, Gaithersburg MD USA), per our previously published protocol³.

Apparatus

Training was conducted in 8 standard behavioral chambers, which were individually housed in light- and sound-attenuating boxes (Med Associates, VT). For behavioral experiments, the chambers were equipped with a pellet dispenser that delivered 45-mg sucrose pellets (5TUT; BioServ, Frenchtown, NJ) into a recessed magazine when activated. Two retractable levers could be inserted into the chambers either side of the recessed magazine. Access to the magazine was detected by means of infrared detectors mounted across the opening of the recess. A computer equipped with medPC software (Med Associates, VT) controlled the equipment and recorded the responses. Raw data was output and processed in medPC2XL (Med Associates, VT) to extract relevant response measures. The chambers contained a speaker connected to white noise generator and a relay that delivered a 5 kHz clicker stimulus, as well as a house light that could illuminate the chambers when programmed.

Behavioral Procedures

Pavlovian-to-Instrumental Transfer

The Pavlovian-to-Instrumental transfer experimental procedure was based on that previously described^{4,5}. Briefly, rats begin training with Pavlovian conditioning, which continued for 10 days. During the 60-min session, rats received four 2-min presentations of two cues (clicker or white noise, counterbalanced), each paired with one of two rewards

(dopamine stimulation or sucrose pellets, counterbalanced). Rewards were delivered at four points randomly throughout the 2-min cue. To stimulate dopamine neurons, light (473nm; 1s; 14-16mW) was delivered into the brain at either 20Hz or 50Hz (5ms pulses)^{1,6}. Inter-trial intervals (ITIs) averaged 5 minutes in duration. Locomotor activity was recorded throughout the session and analyzed via eztrack⁷. Following Pavlovian conditioning, rats were given 8 days of instrumental training, where they were trained to lever press for pellets or dopamine stimulation. Each session consisted of two 10-minute blocks on each lever separated by a 2.5-minute time-out period in which levers were retracted. If an animal pressed a lever more than 20 times during a 10-minute session, then the lever was retracted immediately, and the 2.5-minute time-out period began before the next lever was made available, earning a maximum of 40 rewards on each lever. For the first two days of instrumental training lever presses were continually reinforced (CRF). Rats were then moved on to three days of a random ratio schedule (RR) where each levers delivered a reward with a probability of 0.2 (i.e. RR5), then finally to an RR10, which delivered a reward with a probability of 0.1. During the critical PIT test, both levers were extended, and no rewards were delivered. To extinguish instrumental responding on both levers, rats first received 8 minutes of extinction prior to cue presentation. Then, each cue was presented four times in the following order: clicker-noise-noise-clicker-noise-clicker-clicker, with a fixed 3-min ITI between cues. Rats received three such sessions, where they receive one RR10 session in between these tests.

Blocking

Rats first received presentation of two light cues followed by delivery of two distinct rewards (e.g. flash→sucrose; house light →pellet). Rats received 8 sessions, consisting of 14 trials, separated by variable ITI of 4mins. Then, rats received the visual cues in compound with two novel auditory cues to create two novel audiovisual compounds, which were followed by the same rewards (e.g. flash + click→ sucrose; house light + white noise → pellets). During compound training, rats received stimulation of dopamine neurons that coincided with reward delivery after one of the compounds^{8,9} (473nm, 1s, 20Hz, 5ms pulses, 14-16mW), to mimic an endogenous prediction error¹⁰⁻¹². Rats then received two probe tests with the click and white noise alone without reward, consisting of 8 trials with an ITI averaging around 4 minutes. All stimuli were counterbalanced as was the order of their presentation.

Devaluation

Following blocking, we trained rats on instrumental contingencies for the two distinct rewards. We then tested if the cues would promote PIT. However, we did not see significant PIT in any direction, primarily because rats spent a lot of time in the food port during the unblocked cue (blocked (\pm SEM): 1.5s (0.2); unblocked (\pm SEM): 2.7s (0.2); $F_{1,8}=6.230$, $p=0.02$). Instead, we conducted a devaluation test to assess the nature of the association that had developed between the unblocked cue and the reward. To do this, we allowed rats to consume the reward paired with the unblocked cue in a novel context for 30mins. Immediately after this, half the rats received injections of lithium chloride (LiCl; 0.15M, 10mg/ml; devalued group). The other half were returned to the colony room and received injections 6 hours later (non-devalued group). We repeated this for 3 days. After

48 hours recovery, rats were again placed in the experimental chambers and given a final probe test, where the unblocked cue was presented without reward. We also tested responding to the blocked cue (Fig S3). Each cue was presented for a total of 4 trials, with an ITI averaging around 4 mins.

Intracranial self-stimulation

To examine whether the rats from our blocking experiment would perform ICSS for 50Hz stimulation of dopamine neurons, we allowed them access to a lever that would produce 1s of 50Hz stimulation (473nm; 1s; 14-16mW, 5ms pulse) on a CRF schedule, and another that produced nothing. This session lasted for 30mins and rats could press the levers as much as they wanted.

Recording Procedures

Traditional photo-tagging procedures have accounted for differences in the amount of light reaching proximal and distal microwires by titrating the light delivery for each wire^{13,14}. However, we opted against this approach, fixing the light intensity at 15mW, so as to provide an accurate model of neuronal firing in the tissue surrounding the optic fiber for the behavioral tasks used in the present paper. Accordingly, we observed a decrease in fidelity for neurons distal to the optic fiber, consistent with models of optogenetic stimulation. Neurons were classified as channelrhodopsin (ChR2)-responsive if they reliably exhibited responses within 15ms of the light onset¹³. To confirm the absence of photoelectric artifacts, we recorded 2 additional rats with wires dorsal to the VTA. In these rats, we found no evidence optically evoked activity, even at optical power levels above those used in the experimental preparation.

On the day of testing, rats were placed in a standard operant chamber and allowed to habituate for 10-15 minutes during which period neurons were identified and sorted online using Synapse software (TDT, Alachua, FL, USA). Single neurons were amplified and recorded using a digital headstage (ZD32, TDT, Alachua, FL, USA), preamplifier (PZ5), and bioamp (RZ2). Signals were then bandpass filtered at 300-5000Hz, and stored using Synapse (TDT, Alachua, FL, USA) at 24 kHz. Rats were first given 10 trials of 20 Hz stimulation (473 nm laser, 15mW, 5 ms pulse) with a 1-minute inter-trial interval, followed by 10 trials of 50 Hz stimulation (473 nm laser, 15mW, 5 ms pulse) with a 1-minute inter-trial interval.

Histology

Behavioral experiments

Rats were sacrificed by carbon dioxide and perfused with phosphate buffer followed by 4% paraformaldehyde in phosphate buffer. 20µm coronal sections were collected using a Cryostat (Leica Bio Systems), imaged and visualized for confirmation of virus expression and fiber tip placements using a Zeiss LSM 900 microscope with a 4x or 20x objective and ZEN Imaging software.

Recording experiments

Rats were terminally anesthetized and anodal current (50 μ A, 4 sec) was passed through each microwire to mark the tip location. Rats were then perfused with phosphate buffer (PB, 0.1M, pH 7.4) followed by 4% paraformaldehyde (PFA). The brain was then stored in 4% PFA overnight before storing in 18% sucrose at 4 °C until the brain equilibrated with the specific gravity of the sucrose solution. The VTA was then coronally sectioned at 30 μ m to validate ChR2 expression and localize each microwire, as described previously³. Briefly, the positions of the microwires were identified by their tracks in the brain and their lesions at the microwire tips.

Statistical Analyses

Statistical analyses were conducted using SPSS 24 IBM statistics package. Analyses were conducted using a mixed-design repeated-measures analysis of variance (ANOVA). One-tailed tests were used for results with an a priori directional hypothesis. The data from the PIT test were analyzed across all trials of all three sessions of the PIT tests, and the blocking data were analyzed across the first six trials of both probe tests, consistent with other published demonstrations^{8,15}. The ICSS and blocking data were correlated using Pearson's correlations on the log (x + 1) transformed data from the first 5 mins of the ICSS session and the ratio of responding to the blocked or unblocked cue relative to baseline (i.e. X - baseline/ X + baseline). For the recording studies, neuronal fidelity was calculated by examining the number of pulses within each trial where an action potential was recorded within 15 ms of the light onset. Fidelity was then further examined as a function of 1) each trial as a whole, 2) the pulse number in each train, and 3) the first 5 and last 5 pulses in each trial. The overall firing frequency across the 1 s train, irrespective of fidelity to individual pulses, was also analyzed. Data collection and analyses were not performed blind to the conditions of the experiments. Sample sizes were chosen based on similar prior experiments which have elicited significant results with a similar number of rats¹⁶. Further, we ran a formal post-hoc power analysis on the data elicited from these experiments to using G*Power 3.0 to estimate the power we had achieved using our sample sizes¹⁷. Specifically, we used the average of the partial η^2 (~0.8) from our analyses from the PIT test with the 20Hz and 50Hz groups to calculate the power (1- β) we had achieved. These analyses revealed an estimated power of 0.99 with the same sizes used in our study (Fig S1), with a type 1 error rate (α) below 0.05.

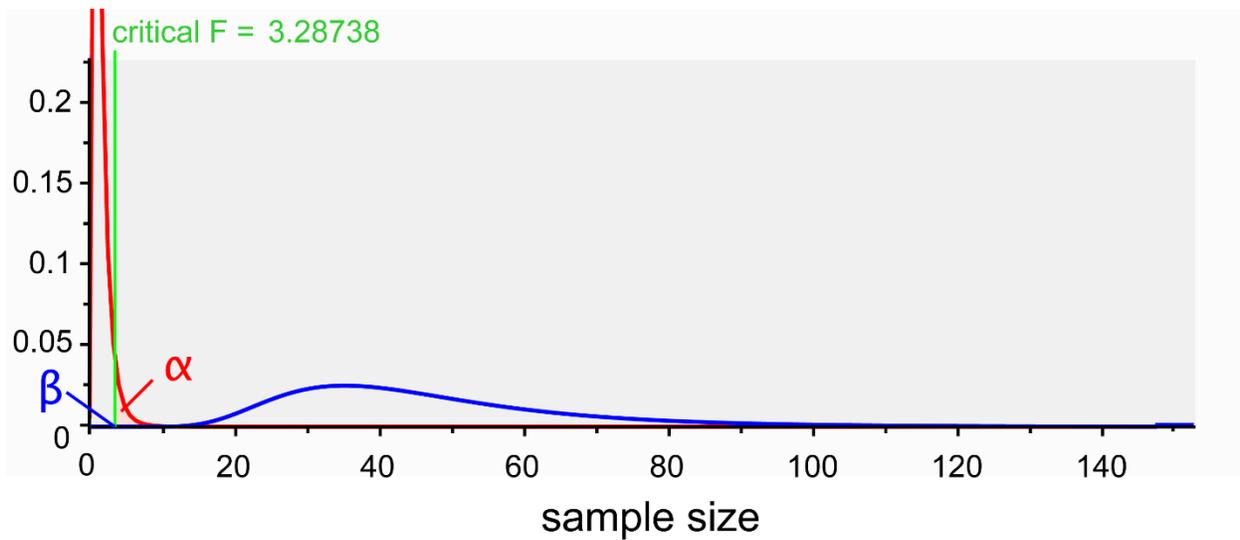


Fig. S1. Formal power analyses showed that we obtained a high degree of power with the sample sizes used in these studies. Using G*power 3.0¹⁷, we conducted formal post-hoc power analyses on the data elicited from our PIT tests. The average partial η^2 elicited from our critical effects in our 20Hz and 50Hz groups was ~ 0.8 , which resulted in a high degree of power ($1-\beta$; 0.99), and a type 1 error rate (α) below 0.05, with the sample sizes used in our study. This demonstrates that our sample sizes were sufficient to detect our critical effects with low likelihood of type 1 (α) or type 2 (β) errors.

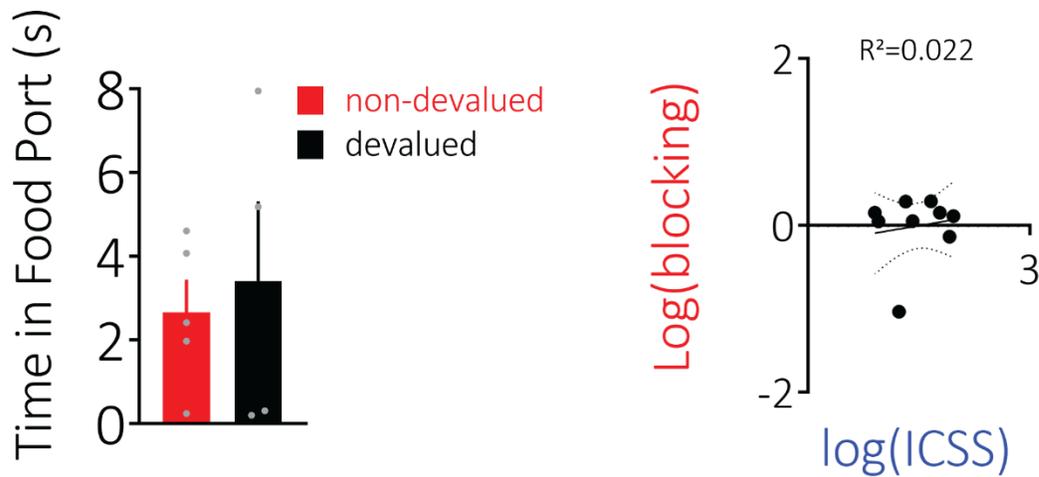


Fig S2. Responding to the blocked cue is not sensitive to devaluation and is not correlated with ICSS. Left: After devaluation of the reward paired with the unblocked cue, we tested responding to the blocked cue. We found that devaluation of the reward did not impact on responding to the blocked cue ($F_{1,7}=0.157$, $p=0.352$). Right: Responding to the blocked cue in the initial probe test illustrated in the main text in Figure 3B did not correlate with the degree to which rats would press the active lever to get stimulation of dopamine neurons (Pearson's $r=0.148$, $R^2=0.022$; $p=0.352$).

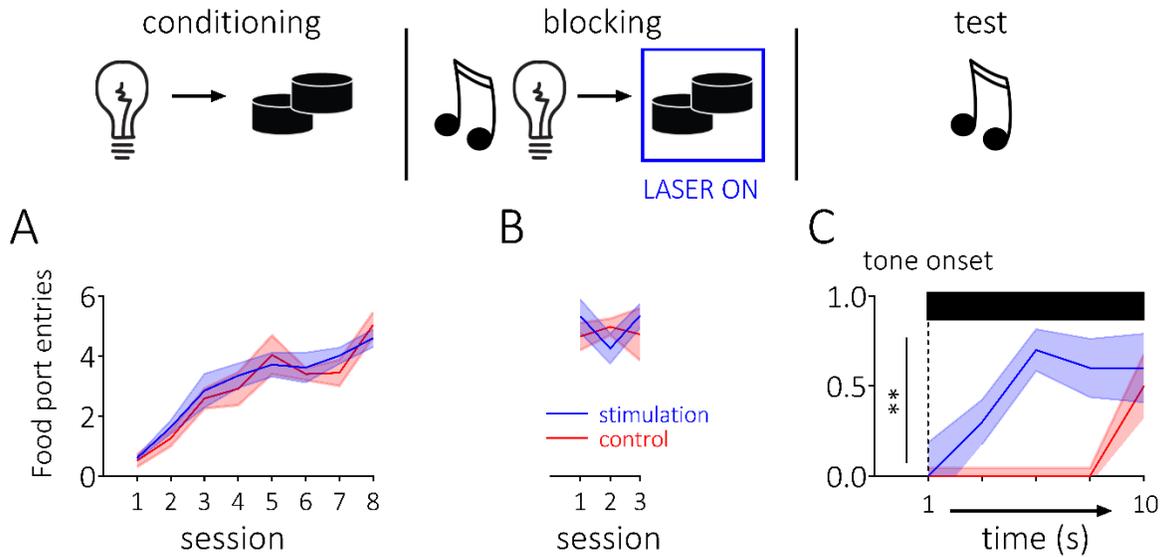


Fig S3. A physiologically-relevant frequency of dopamine stimulation (20Hz) functions as a prediction error to unblock learning of cue-reward associations. Top: schematic illustrating the task design. Bottom: food-port entries during conditioning (A), blocking (B), and the test (C). To ensure that our 20Hz stimulation was functioning to produce a physiologically-relevant signal, we tested whether these stimulation parameters would be sufficient to drive learning, as we and others have demonstrated previously. This experiment was conducted on the same rats that underwent the PIT experiment represented in Figure 2 of the main text. A) We trained rats ($n=10$) that a light led to delivery of two food pellets. All rats acquired the food-port entry response, with no difference between the groups (session: $F_{7,56}=14.369$, $p<0.0001$; session x group: $F_{7,56}=0.344$, $p=0.930$; group: $F_{1,8}=0.119$, $p=0.739$). B) Then, we presented a tone in compound with the light and the same food pellets. Usually, rats do not learn about the tone. However, when the food pellets were presented after the tone + light compound, we stimulated dopamine neurons at 20Hz in half the rats. All rats continued to exhibit high rates of responding at the food port, with no visible changes in these rates produced by dopamine stimulation (main effect, session: $F_{2,16}=1.098$, $p=0.357$; session x group interaction: $F_{2,16}=1.608$, $p=0.218$; group: $F_{1,8}=0.001$, $p=0.978$). We then tested responding to the tone alone and found that dopamine stimulation had successfully unblocked learning (time: $F_{4,32}=4.144$, $p=0.008$; time x group: $F_{4,32}=2.228$, $p=0.044$; group (early): $F_{1,8}=7.538$, $p=0.025$), confirming our stimulation parameters were producing a physiologically-relevant dopamine signal. Error bars = SEM.

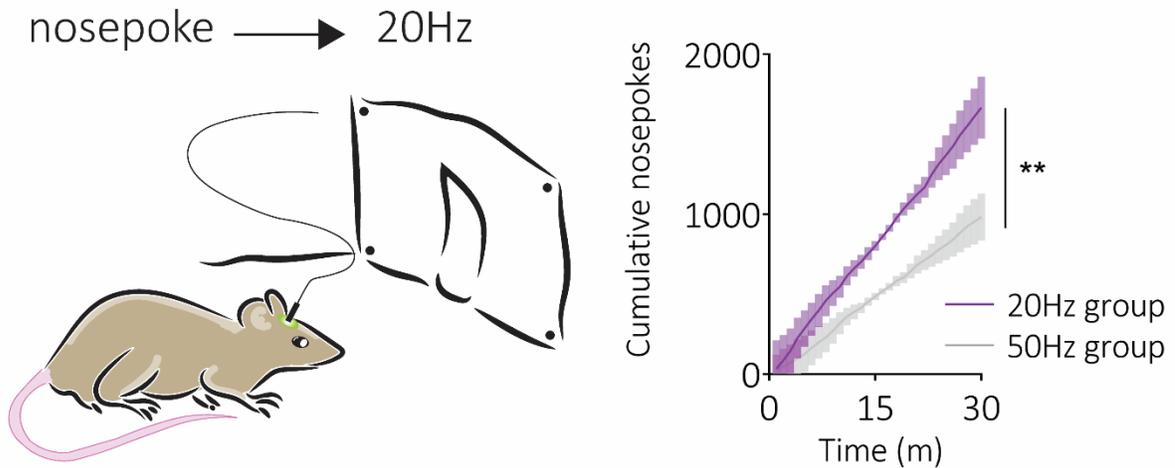


Fig. S4. Rats will nosepoke robustly for 20Hz stimulation of dopamine neurons on a continuously-reinforced schedule. We tested whether we could replicate the finding that rats will nosepoke at high rates for 20Hz dopamine stimulation¹⁸. Consistent with this data, we found that the rats that had previously received 20Hz of dopamine stimulation would nose poke robustly for 20Hz dopamine stimulation across the session. However, rats that had previously received 50Hz of dopamine stimulation would nose poke for 20Hz, but significantly less than the 20Hz group. This suggested that reducing the frequency of dopamine stimulation from 50 to 20Hz devalued the stimulation. This was confirmed with statistical analyses on the cumulative data, which showed a main effect of time ($F_{29,232}=30.830$, $p=0.000$), and an interaction by group ($F_{29,232}=1.798$, $p=0.010$), demonstrating that our 20Hz group increased responding across time at a greater rate than the 50Hz group.

References

- 1 Sharpe, M. J. *et al.* Dopamine transients do not act as model-free prediction errors during associative learning. *Nature Communications* **11**, 1-10 (2020).
- 2 Sharpe, M. J. *et al.* Lateral hypothalamic GABAergic neurons encode reward predictions that are relayed to the ventral tegmental area to regulate learning. *Current Biology* **27**, 2089-2100. e2085 (2017).
- 3 Barker, D. J., Root, D. H., Coffey, K. R., Ma, S. & West, M. O. A procedure for implanting organized arrays of microwires for single-unit recordings in awake, behaving animals. *JoVE (Journal of Visualized Experiments)*, e51004 (2014).
- 4 Bradfield, L. A., Dezfouli, A., van Holstein, M., Chieng, B. & Balleine, B. W. Medial orbitofrontal cortex mediates outcome retrieval in partially observable task situations. *Neuron* **88**, 1268-1280 (2015).
- 5 Bradfield, L. A., Hart, G. & Balleine, B. W. Inferring action-dependent outcome representations depends on anterior but not posterior medial orbitofrontal cortex. *Neurobiology of learning and memory* **155**, 463-473 (2018).
- 6 Sharpe, M. J. *et al.* Dopamine transients are sufficient and necessary for acquisition of model-based associations. *Nature Neuroscience* **20**, 735 (2017).
- 7 Pennington, Z. T. *et al.* ezTrack: An open-source video analysis pipeline for the investigation of animal behavior. *Scientific reports* **9**, 1-11 (2019).
- 8 Steinberg, E. E. *et al.* A causal link between prediction errors, dopamine neurons and learning. *Nature neuroscience* **16**, 966 (2013).
- 9 Keiflin, R., Pribut, H. J., Shah, N. B. & Janak, P. H. Ventral tegmental dopamine neurons participate in reward identity predictions. *Current Biology* **29**, 93-103. e103 (2019).
- 10 Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593-1599 (1997).
- 11 Roesch, M. R., Calu, D. J. & Schoenbaum, G. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nature neuroscience* **10**, 1615-1624 (2007).
- 12 Takahashi, Y. K. *et al.* The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron* **62**, 269-280 (2009).
- 13 Kravitz, A. V., Owen, S. F. & Kreitzer, A. C. Optogenetic identification of striatal projection neuron subtypes during in vivo recordings. *Brain research* **1511**, 21-32 (2013).
- 14 Root, D. H., Estrin, D. J. & Morales, M. Aversion or salience signaling by ventral tegmental area glutamate neurons. *Science* **2**, 51-62 (2018).
- 15 Sharpe, M. J. & Killcross, S. The prelimbic cortex contributes to the down-regulation of attention toward redundant cues. *Cerebral cortex* **24**, 1066-1074 (2012).
- 16 Sharpe, M., Clemens, K., Morris, M. & Westbrook, R. Daily exposure to sucrose impairs subsequent learning about food cues: a role for alterations in ghrelin signaling and dopamine D2 receptors. *Neuropsychopharmacology* **41**, 1357 (2016).

- 17 Faul, F., Erdfelder, E., Buchner, A. & Lang, A.-G. Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods* **41**, 1149-1160 (2009).
- 18 Witten, I. B. *et al.* Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron* **72**, 721-733 (2011).