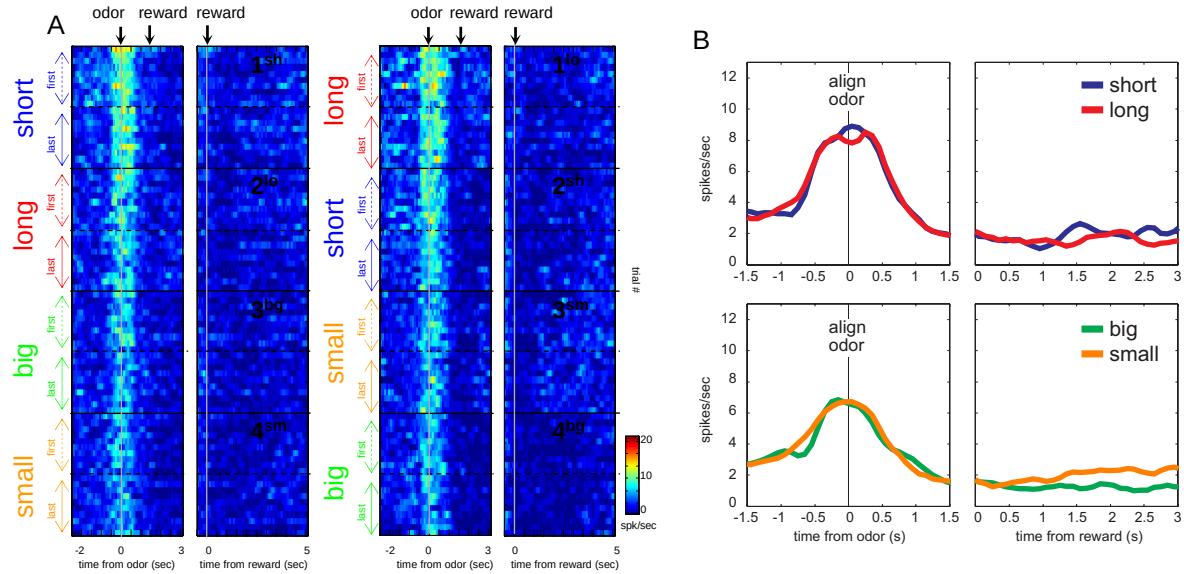
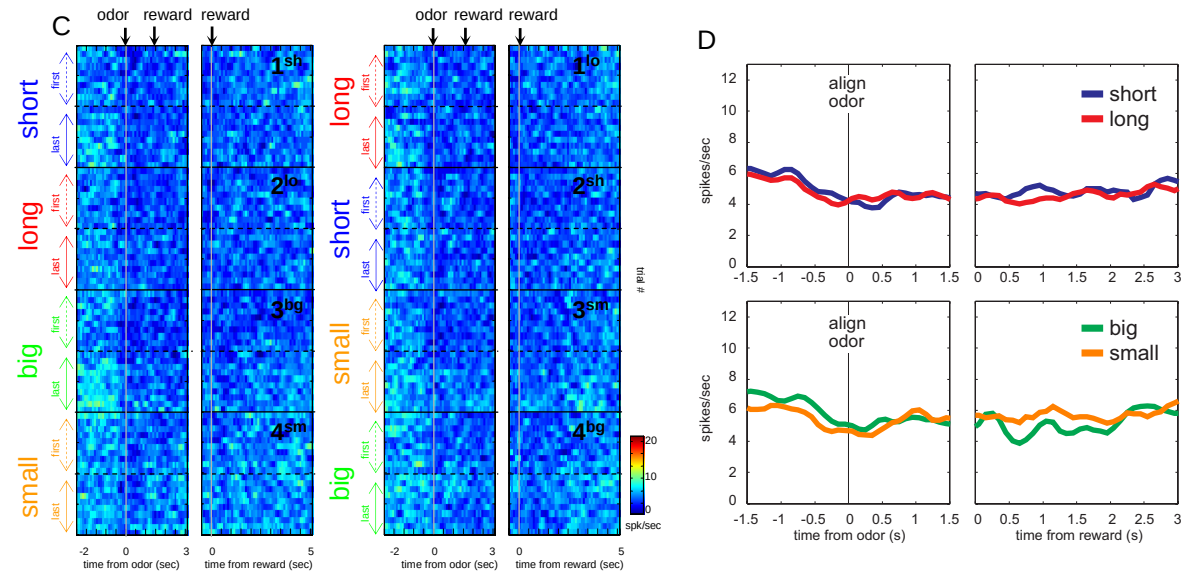
Putative DA cluster (VTA): Cue Responsive Only (n =14)

Excited during odor sampling (n = 6)



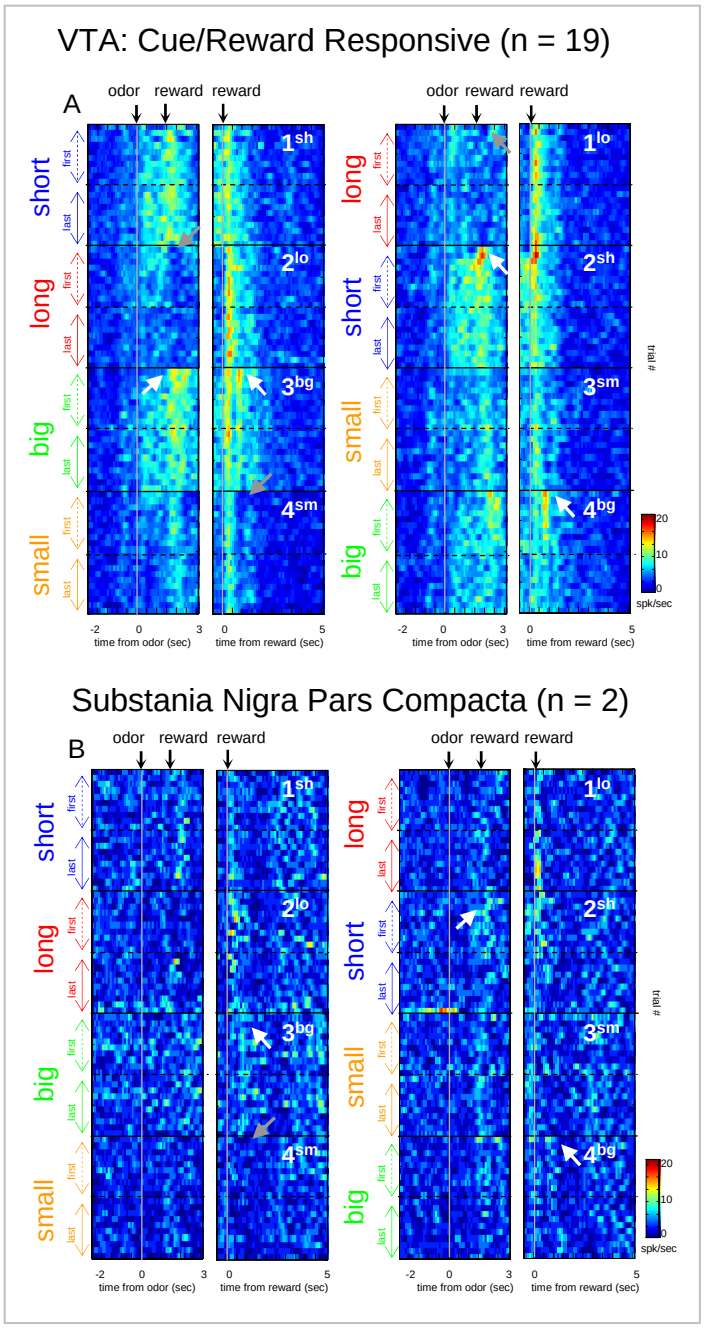
Inhibited during odor sampling (n = 8)



Supplemental Figure. 1

Figure S1. Activity in VTA dopamine neurons that were cue but not reward responsive. We classified waveforms based on spike duration (y-axis) and the amplitude ratio (x-axis) of the initial positive and negative segments as calculated in the in figure 2. This analysis identified 36 cells that met these previously established electrophysiological criteria. Of those recorded in VTA, 14 showed a significant increase (A-B) or decrease (C-D) above baseline during odor sampling. Heat plots showing of these two populations during the first and last twenty forced-choice trials (10 per direction) in each training block (Figure 1; Blocks 1-4). Activity is shown, aligned on odor onset ('align odor') and reward delivery ('align reward'). Blocks 1-4 are presented in the order that they were performed (top to bottom). Thus, during block 1, rats underwent a 'long' delay or a 'short' delay to receive reward. In block 2, the location of the 'short' delay and 'long' delay were reversed. In blocks 3-4, delays were held constant but the size of the reward ('big' or 'small') received varied. Line plots represent the average firing rate over the last 30 trials in each block. Blue: short; Red: long; Green: big; Orange: small

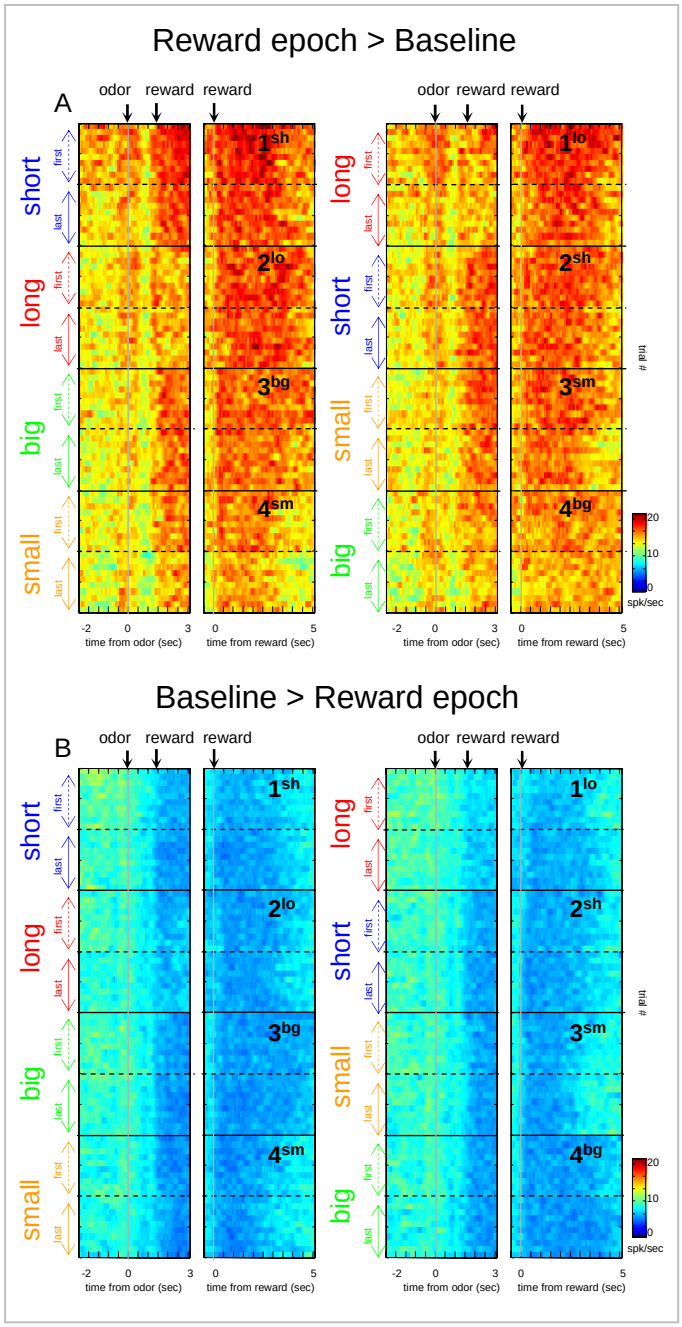
Putative DA cluster



Supplemental Figure. 2

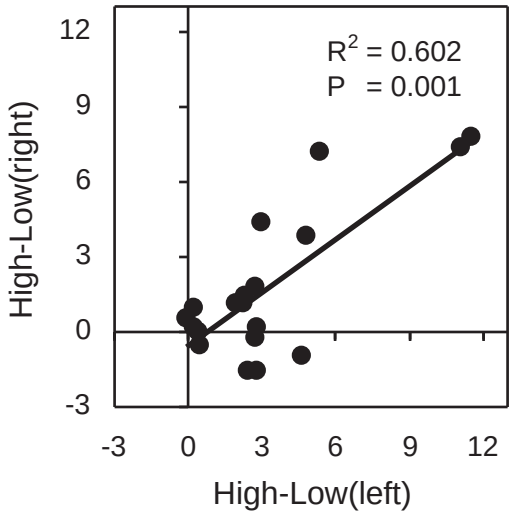
Figure S2. Activity in VTA dopamine neurons that were cue and reward responsive and also in reward-responsive SN neurons. The average firing rate of the 19 cue/reward responsive dopamine neurons that were shown in figure 4 in the main text are repeated in Figure S2, except here we show more activity after odor onset to allow one to appreciate the negative prediction error associated with the delayed reward in blocks 1 and 2 ('long'). As a result reward-related activity is plotted in both left and right panels for 'short', 'big' and 'small' conditions. Notably dips in activity related to negative prediction errors were often followed by a stronger increase in activity when the unexpected reward was finally predicted. This can be better appreciated in figure S6. Conventions are the same as in figure S1. B. Plots the average firing rate over 2 dopamine neurons recorded in SNpc. Although the sample is low, one can still appreciate hints of prediction error encoding as found in VTA, as noted by white (positive prediction errors) and gray (negative prediction errors) arrows.

Putative non-DA cluster



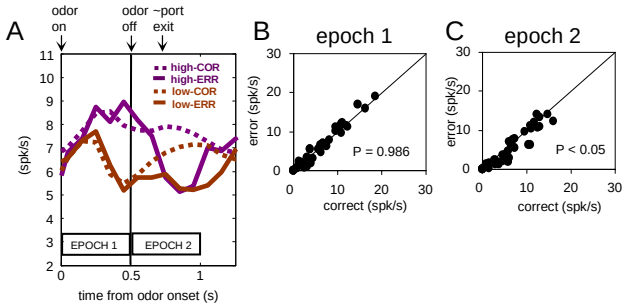
Supplemental Figure. 3

Figure S3. Activity in VTA non-dopaminergic neurons. The average firing rate of the 222 non-dopaminergic neurons broken down by if they fired more (A) or less (B) strongly during the delivery of reward as compared to baseline.



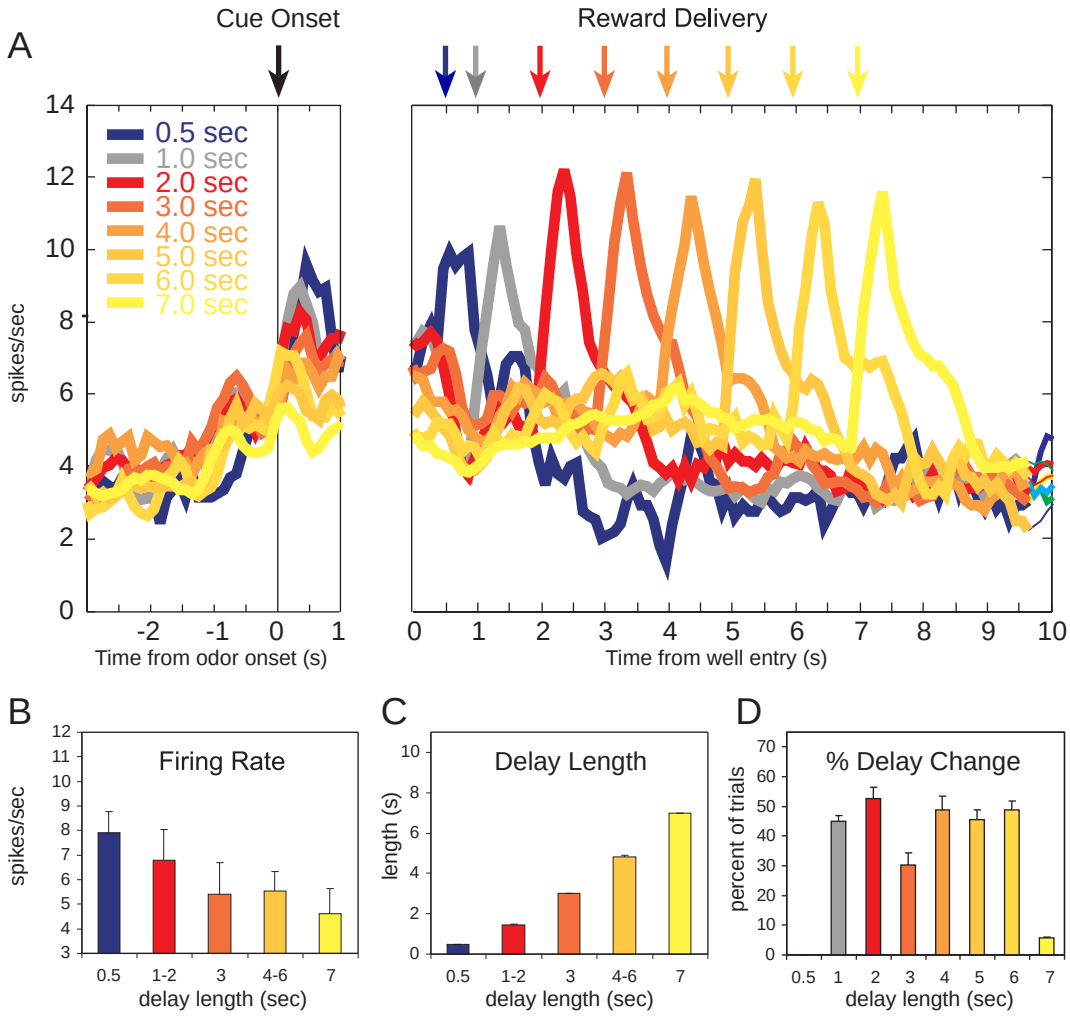
Supplemental Fig. 4

Figure S4. Effect of response direction on cue-evoked firing in VTA dopamine neurons. Plots the value index for each neuron, collapsed across reward and delay measures, for response made to left (x-axis) and right (y-axis). Value index (high – low/ high + low) was computed using average firing rates starting at odor onset and ending at the time of odor port exit.



Supplemental Figure. 5

Figure S5. Cue-evoked activity in putative dopaminergic neurons on correct and incorrect forced-choice trials. A. Population activity averaged over the 19 cue/reward responsive neurons, collapsed across direction and value manipulation for (1) correct high-value, forced-choice trials (purple-dashed), (2) incorrect high-value, forced-choice trials (purple-solid), (3) correct low-value, forced-choice trials (brown-dashed), and (4) incorrect low-value, forced-choice trials (brown-solid). B and C plot the average firing rate taken for each error trial and paired correct trial over the 500 ms after odor onset ('epoch 1') and the 500 ms after odor offset ('epoch 2'), respectively. * $p < 0.05$; chi-square.



Supplemental Fig. 6

Figure S6. Activity in VTA dopamine neurons reflects time to rather than titration of delayed rewards. A. Average firing rate for all 19 cue/reward responsive dopamine neurons over forced- and free-choice trials. Color indicates the length of the delay preceding reward delivery from 0.5 to 7 s. Activity is aligned on odor onset (left) and well entry (right). To plot the average firing rate over all neurons, free-choice trials had to be included because not all recording sessions had a sufficient sample of each delay under forced-choice conditions, however statistical analysis was restricted to forced choice trials by collapsing across 1-2 sec delays and 4-6 second delays. B. Height of each bar indicates the average firing rate during odor sampling during performance of forced-choice trials of different delays. C. The average length of delay associated with each of the bars indicated in B. D. Percentage of trials of given delay on which the length of the delay had changed from its previous value.

SUPPLEMENTAL DATA

Changes in neural activity in response to changes in delay and reward size in different populations of neurons recorded in VTA and SNpc

In the main text, neurons were divided into those with waveforms characteristic of dopaminergic neurons and those with non-dopaminergic waveforms. Further the analysis focused on 19 VTA putative dopamine neurons that were cue/reward responsive and 14 that were cue responsive only. Of the 14 cue-responsive neurons, none encoded value during odor sampling. This is further illustrated in Figure S1A-B, which plots the activity of the 14 cue-responsive neurons, broken down by whether they fired more (A-B) or less strongly (C-D) for cues as compared to baseline. In contrast, the activity of the 19 cue/reward responsive neurons (all excitatory), as illustrated in the main text and reproduced here (Figure 2A), clearly encoded prediction errors and the relative value of reward-predicting cue.

The fact that not all putative dopamine neurons exhibited prediction errors and value encoding is interesting in light of recent work suggesting that electrophysiological properties may not reliably distinguish dopaminergic neurons [1,2]. Non-reward responsive neurons reported here may be non-dopamingeric even though their waveforms clustered with those responsive to reward delivery. This would support a procedural role for functional screening in behavioral recording studies when examining dopaminergic activity.

Also shown are the 2 neurons recorded from substania nigra pars compacta (SNpc). Unfortunately, the neuronal yield associated with recording dopaminergic neurons from SNpc was extremely low making it difficult to make meaningful comparisons across VTA and SNpc. With that said, we did observed predictions errors in these two neurons in SNpc (Figure S2B). However neither reflected the value of the predicted reward during odor sampling.

As expected, most of the neurons recorded in VTA did not exhibit waveforms characteristic of dopaminergic neurons. The population response of these 222 neurons is illustrated in Figure S3, broken down by whether the neurons showed an increase (A) or decrease (B) in firing to reward delivery (compared to baseline). In contrast to the

population response of the putative dopaminergic neurons, shown in Figure S1A, firing in these neurons – even in those that fired to reward - was not modulated by the error in predicted value but rather seemed to generally increase or decrease in response to reward throughout the session. Of the 222 non-dopaminergic neurons, 13 were significantly modulated by the size of the expected reward and 14 were significantly modulated by the length of the expected delay; however this is no more than one would expect from chance alone (chi-square; p's > 0.3). Furthermore, neither population exhibited a significant correlation between size and delay manipulations. We conclude that value encoding was a unique attribute of dopamine neurons that were reward-responsive.

Activity in dopaminergic neurons is independent of cue identity and response direction

In the main text, neural activity was analyzed regardless of cue identity and response direction. Previous reports have shown that dopaminergic neurons encode reward prediction errors independent of the identity of predictive cues or direction of response required. This was also true for the results described here. This is evident in Figure S4, which plots the difference between high (averaged across short and big) and low (averaged across long and small) valued conditions, for those responses made to the left (x-axis) and right well (y-axis). Encoding of the value of responses left and right was highly correlated. This contrasts with the directional encoding of value we have previously found in other areas in this task [3].

Cue-evoked activity on error trials correctly signals the value of the reward that is available

In the main text, we compared cue-evoked activity on free- versus forced-choice trials to determine whether neuronal activity in the dopaminergic neurons encoded the value of the available rewards independent of the reward that was subsequently selected (Figure 6, main text). Cue-evoked activity was higher on the forced-choice trials when the cue predicted the high value outcome, however, under free-choice conditions, this difference did not exist. Instead cue-evoked activity was the same as that on the high

value forced-choice trials, regardless of whether the rat ultimately responded at the high value well or the low value well.

Here we examined whether similar effects might also be evident when comparing cue-evoked activity on correct and incorrect forced-choice trials. This might be the case if, after learning, incorrect trials were not actually errors but rather reflected exploratory behavior. In other words, the rats might have been essentially checking to see if the rewards had changed again. We paired each incorrect forced-choice trial with the immediately preceding and following correct forced-choice trials. The average population response on these trials, collapsed across value manipulation and direction, is shown in Figure S5A. For both correct and incorrect trials, cue-evoked activity was higher when the cue predicted the high value outcome, even when the rat responded to the wrong well. Indeed there was no significant difference in neural activity between correct and incorrect trials during odor presentation. Notably, after odor sampling, activity between correct and incorrect trials did differ, declining on incorrect trials. These effects are quantified in Figure S5B and C, which plot the average firing rate for incorrect (x-axis) and correct (y-axis) trials during epoch 1 (odor sampling) and epoch 2 (500 ms after odor offset). That activity on errors still accurately reflected the reward available suggests that, at least for neural activity in VTA, the incorrect trials were not due to miscoding of the proper response for reward.

Dopamine neurons encode the time to delayed reward rather than its uncertainty due to titration of the delay

In the main text, we demonstrate that some dopamine neurons encode the relative value of an immediate versus delayed reward. Generally changes in firing to delayed rewards might reflect the cost of lost opportunities or the uncertainty of future rewards. Although the delayed reward was always delivered in our task, future rewards are thought to be inherently uncertain compared to immediate rewards. In addition, the timing of the delayed reward was titrated to discourage but not eliminate responding (see methods). This was necessary so that we could compare neural activity at the two wells on free-choice trials in each block. However titrating the delay caused the timing of the delayed reward to be less consistent than that of the immediate reward, particularly at the start of

each delay block, when the time to reward increased rapidly on one side to its new value. However, comparison of cue-evoked activity on trials of different delays showed that activity was inversely related to delay length and was not related to the frequency of titration. This is illustrated in Figure S6 which plots the average firing rate over all 19 cue/reward responsive neurons, independently for each delay length. Cue-evoked activity was maximal at the short delay (0.5 s) and then declined with each increase in the delay (Figure S6A-C). Importantly, this incremental decline with increasing delay did not reflect the rate of which the delay was changing, which was low for 0.5 s, 3 s and 7 s delays and high for 1-2 s and 4-6 s delays (Figure S6D). Thus it seems unlikely that the influence of delayed reward on neural activity observed here is an artifact of titration per se, although changes in activity may still reflect the more general uncertainty inherent in delayed rewards. Instead, the reduced activity of dopamine neurons during long forced-choice trials (as well as small-reward trials) may reflect an error in reward prediction and/or the reduced value of the reward.