Reviewed Preprint
v1 • January 30, 2026
Not revised

Reviewed Preprint
v2 • August 12, 2026
Revised by authors

Ventromedial striatal dopamine dynamically integrates motivated action and reward proximity

Eugenia Z Poh[1,2], **Nicky L Buitelaar**[1], **Gino Hulshof**[1], **Lucie Mazé**[1,4], **Pieter N de Greef**[1], **Georgios Zaverdinos**[1,2,3], **Ingo Willuhn**[1,2]

[1]Netherlands Institute for Neuroscience, Royal Netherlands Academy of Arts and Sciences, Amsterdam, Netherlands, • [2]Amsterdam UMC, location University of Amsterdam, Department of Psychiatry, Amsterdam, Netherlands • [3]Max Planck Institute for Biological Cybernetics, Tübingen, Germany • [4]Institut des Neurosciences Cellulaires et Intégratives, Strasbourg, France

For correspondence: i.willuhn@nin.knaw.nl

Competing interests: No competing interests declared

Funding: See page 29

Reviewing editor: Shelly B Flagel, University of Michigan, United States

eLife Assessment

This manuscript describes a series of studies using four different Go/No Go task variants in combination with fast-scan cyclic voltammetry to determine the role of dopamine release in the ventromedial striatum in action selection, controllability of reward pursuit, effort, and reward approach. The authors conclude that dopamine signals in the ventromedial striatum integrate the invigoration of action initiation with continuous estimation of spatial, but not temporal, proximity to rewards. There is **solid** support for the conclusions, and the findings are **valuable**, with theoretical implications for the role of dopamine in reward-driven behavior.

https://doi.org/10.7554/eLife.109868.2.sa3

Abstract

Dopamine release in the ventromedial striatum (VMS) both invigorates actions and encodes reward-related information, yet how these functions are integrated remains under active debate. To investigate this further, we designed four different versions of a rat Go/No-go task, where we systematically manipulated response requirements, temporal task demands, and controllability of reward pursuit. Dopamine release increased reliably during action initiation (Go) but was delayed during action suppression (No-go), and this difference was insensitive to augmented response demands or controllability. Following response completion, dopamine rose gradually until animals arrived at the reward location, irrespective of reward-delivery timing, prior action demands, or controllability. This proximity dopamine-signal was exaggerated after animals exhibited Pavlovian consummatory behavior during No-go trials, revealing a motivational signal component. Together, these findings indicate that in reward contexts, VMS-dopamine signals successively integrate the invigoration of action initiation with the continuous tracking of spatial - but not temporal - proximity to rewards.

Introduction

Appropriate, situation-dependent behavior in everyday life requires both motor control and information about upcoming rewards. A key neuromodulator involved in these processes is dopamine. Midbrain dopaminergic neurons have been shown to encode a reward-prediction error (RPE) which signals the discrepancy between expected and obtained outcomes in a given situation (Bayer and Glimcher, 2005; Montague et al., 1996; Schultz et al., 1997; Starkweather and Uchida, 2021; Steinberg et al., 2013). This reward-learning signal can also be observed in dopamine release from terminals in the ventromedial striatum (VMS) (Day et al., 2007; Kim et al., 2020; Mikhael et al., 2022; van Elzelingen et al., 2022), one of the main projection targets

of midbrain dopamine neurons. Beyond reward learning, VMS dopamine is also implicated in invigorating actions toward appetitive rewards (Coddington and Dudman, 2018; Collins et al., 2016; Hamid et al., 2016; Howe et al., 2013; Howe and Dombeck, 2016; Mohebi et al., 2019; Nicola, 2007; Phillips et al., 2003; Roitman et al., 2004; Syed et al., 2016), supporting the view that the VMS acts as a motivation-action interface (Mogenson et al., 1980; Salamone and Correa, 2012, 2002). For example, VMS dopamine release increases upon movement towards a known reward location, or when initiating an instrumental response for rewards (Hamid et al., 2016; London et al., 2018; Mohebi et al., 2019; Phillips et al., 2003; Roitman et al., 2004; Syed et al., 2016). Although reward learning and motivation have been widely studied independently, how these two functions are integrated in the VMS remains unclear.

An additional process that influences striatal dopamine is the execution of actions required for achieving a reward. For example, under some conditions, the action-related *effort* required for a reward alters VMS dopamine (Cousins et al., 1996; Gan et al., 2010; Salamone and Correa, 2024). Furthermore, VMS dopamine responds differentially to reward contingencies across instrumental- and Pavlovian-conditioning tasks (Goedhoop et al., 2023; Hamid et al., 2021), and midbrain single-neuron activity differs reportedly for actions that are triggered externally (cue-initiated) versus ones that are under exclusive self control (self-initiated) (Romo and Schultz, 1990). Thus, these data suggest that reward-related VMS dopamine signals may be modulated by the effort or the controllability of reward seeking.

Further evidence for a role of VMS dopamine in motivation comes from the gradual increase in its release as animals approach a known reward location (Hamid et al., 2016; Howe et al., 2013; Kim et al., 2020; Mohebi et al., 2019). This signal scales with spatial proximity to rewards (Howe et al., 2013; Kim et al., 2020), and this prolonged “ramp”-like signal has been suggested to sustain motivational drive (Howe et al., 2013), reinforcing the preceding actions that led to reward (Hamid et al., 2016; Niv and Schoenbaum, 2008). However, such evidence is still limited, and it is unknown whether this gradual increase as animals approach the reward location is modulated by effort and the controllability of reward pursuit, and whether the signal is affected by the preceding action requirement for earning rewards.

In both human and rodent studies, Go/No-go tasks are commonly employed to study the effects of action initiation and suppression. However, action requirements for ‘Go’ and ‘No-go’ trials vary substantially between studies. Here, we developed multiple Go/No-go task variants to elucidate how VMS dopamine integrates reward-related information, action selection, effort and controllability of reward pursuit, and reward approach. More specifically, rats were trained to initiate (Go) or suppress (No-go) actions for rewards in response to an action-cue. In another task variant, we added ‘Free’ trials that required no overt action, which is akin to human ‘No-go’ trials where participants must withhold action for a reward (e.g., not press a button). To determine whether the dopamine differential between action and action suppression is influenced by reward-related contingencies, we manipulated: 1) controllability of reward seeking, defined as the ability to determine when to initiate reward pursuit (Self-vs Cue-initiated trials), 2) trial demands, and 3) timing of reward delivery. We hypothesized that controllability and effort would modulate the dopamine differential between action and action suppression, while the gradual increase in VMS dopamine during reward approach would generalize across contingencies.

By combining real-time measurements of dopamine release with detailed behavioral analysis, we replicate previous findings demonstrating that movement is critical for reward-related dopamine release (Syed et al., 2016). In addition, we demonstrate that VMS dopamine is governed specifically by the initiation of the approach to the reward location and proximity to this location, irrespective of preceding action requirements. Finally, different behavioral strategies of action suppression modulate dopamine release amplitude related to spatial reward proximity. Consummatory Pavlovian behaviors during action suppression are associated with the largest dopamine responses during subsequent reward approach, suggesting that individual differences in Pavlovian drive modulate dopamine encoding of spatial proximity to reward.

Results

To investigate the role of VMS dopamine in reward processing and motor control, we implanted rats with fast-scan cyclic voltammetry (FSCV) microelectrodes (n = 21; Figure 2a) and trained them in four variants of the task: 1) a "short-task variant", in which rats *Self-initiated* trials, 2) a "short-task variant" in which trials were *Cue-initiated* (altered controllability of reward pursuit), 3) a "long-task variant" of the *Self-initiated* task, in which the required duration of rat response action was increased (altered difficulty), and 4) a "long-task variant" of the *Cue-initiated* task (Figure 1a, see *Methods: Behavioral procedures*). We defined controllability as the rats' ability to choose the time point of reward pursuit. In *Cue-initiated* trials, the time point at which trials could be started was dictated by a cue light, whereas in *Self-initiated* trials rats were able to choose intrinsically (control) when to attempt a trial start. There were two trial types in the short-task variants of the task: depending on the (auditory) action cue, rats were required to press a lever (Go) or nose poke (No-go) to obtain reward. In the short-task variants, action-cues switched off after trial completion and a reward was delivered immediately. In the long-task variants, we introduced a third trial type: in addition to Go and No-go action-cues and corresponding required responses, a third (auditory) cue signaled that reward would be delivered without requiring an overt action (Free trials). In the long-task variants, action-cues switched off after trial completion, or in the case of Free trials after 3s, and reward was delivered 2s later (Figure 1a).

Behavioral performance in Go/No-go ("short task") was unaffected by controllability of reward pursuit

In the *Self-initiated* Go/No-go task, similarly used previously by Syed et al. 2016, trials were initiated by an (uncued) nose poke (Figure 1a). Our task design ensured that the time of reward delivery following action-cue onset was matched across trial types. After animals achieved proficiency in the task (see Methods), we measured behavioral performance and real-time dopamine release. Animals could initiate a trial by entering the nose-poke port after the intertrial interval (ITI); importantly, rats spent <20% of the ITI time inside the nose-poke port (Supplementary Figure S1), thus in between trials they mostly left the port. Reward success rates were significantly higher when animals had to initiate actions (Go trials) as compared to trials where the animals had to suppress actions (No-go trials) (Figure 1b; Wilcoxon matched-pairs test, W = 35, p = 0.039).

The activity of midbrain dopaminergic neuronal activity differs depending on the controllability of reward pursuit (Romo and Schultz, 1990). To evaluate how VMS dopamine dynamics during action initiation and suppression are affected by diminished control over trial initiation, we additionally trained our rats on a *Cue-initiated* task variant in which trials could only be initiated (by nose poke) after a cue turned on (Fig 1a;): rats learned to promptly respond to the cue (nose-poke light illumination) to initiate trials (1.68 ± 0.28 s; Supplementary Figure S2), indicating that they learned the contingency. Similar to the *Self-initiated* variant, rats trained in the *Cue-initiated* task showed significantly higher success rates on Go trials as compared to No-go trials (Figure 1c, *Cue-initiated* Go/No-go; W = 54, p = 0.014).

Behavioral performance in Go/No-go/Free ("long task") was unaffected by controllability of reward pursuit

Previous research has shown that spatial proximity of subjects relative to the reward location is reflected in VMS dopamine concentration (Howe et al., 2013), and dopamine responses may differ depending on whether rewards are contingent on a subject's behavior (i.e., Pavlovian vs. instrumental responding for rewards) (Goedhoop et al., 2023; Hamid et al., 2021). Furthermore, VMS dopamine may be influenced by effort (Cousins et al., 1996; Gan et al., 2010; for review, see Salamone and Correa, 2024). To investigate the influence of these factors, we trained rats on the long-task variants that included 'Free' trials and increased Go and No-go trial requirements (Go: "long" repeated presses, 6.52 ± 0.32 presses, vs. "short" presses, 1.80 ± 0.03s; No-go: "long" vs. "short" nose-poke hold; see *Methods*, Figure 1a). Reward-delivery

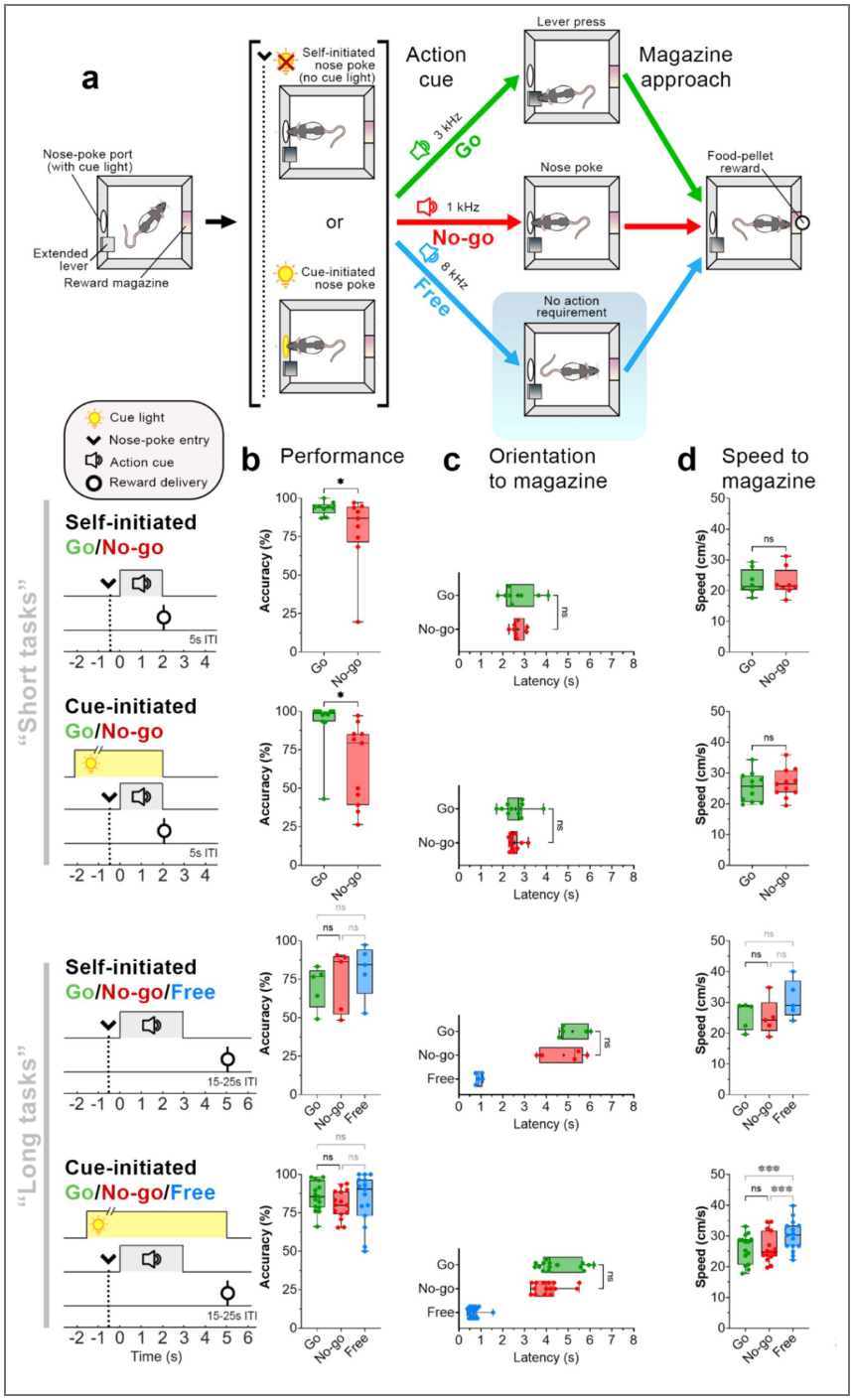


Figure 1. Go and No-go performance is similar across task variants.

a) Top: Schematic of the Go/No-go (short) behavioral task. Rats initiated trials by entering a nose-poke port (arrow, dotted line), either voluntarily (*Self-initiated)* or following nose-poke light illumination (*Cue-initiated*). After staying in the nose-poke port for 0.5s, an auditory stimulus (action cue) instructed rats to either initiate action (lever presses: '*Go*', green) or remain in the nose-poke port ('*No-go*', red). In the *Go/No-go/Free* (long) task, an additional action cue was presented that did not require an action for rewards ('*Free*', blue). Following Free cue onset, rats could approach the reward magazine immediately. *Bottom:* Schematic of task timeline. For all trials, reward-delivery latency (relative to action-cue onset) was similar. Shaded gray area depicts approximate duration of action-cue onset for "short" and "long" task variants ***b-e)*** Success rates across Go/No-go task variants. ***f-i)*** Latency to orient toward the reward magazine following action-cue onset. ***j-m)*** Speed of movement toward the reward magazine after orientation. Significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, Wilcoxon matched-pairs signed rank or *post hoc* Dunn's tests.

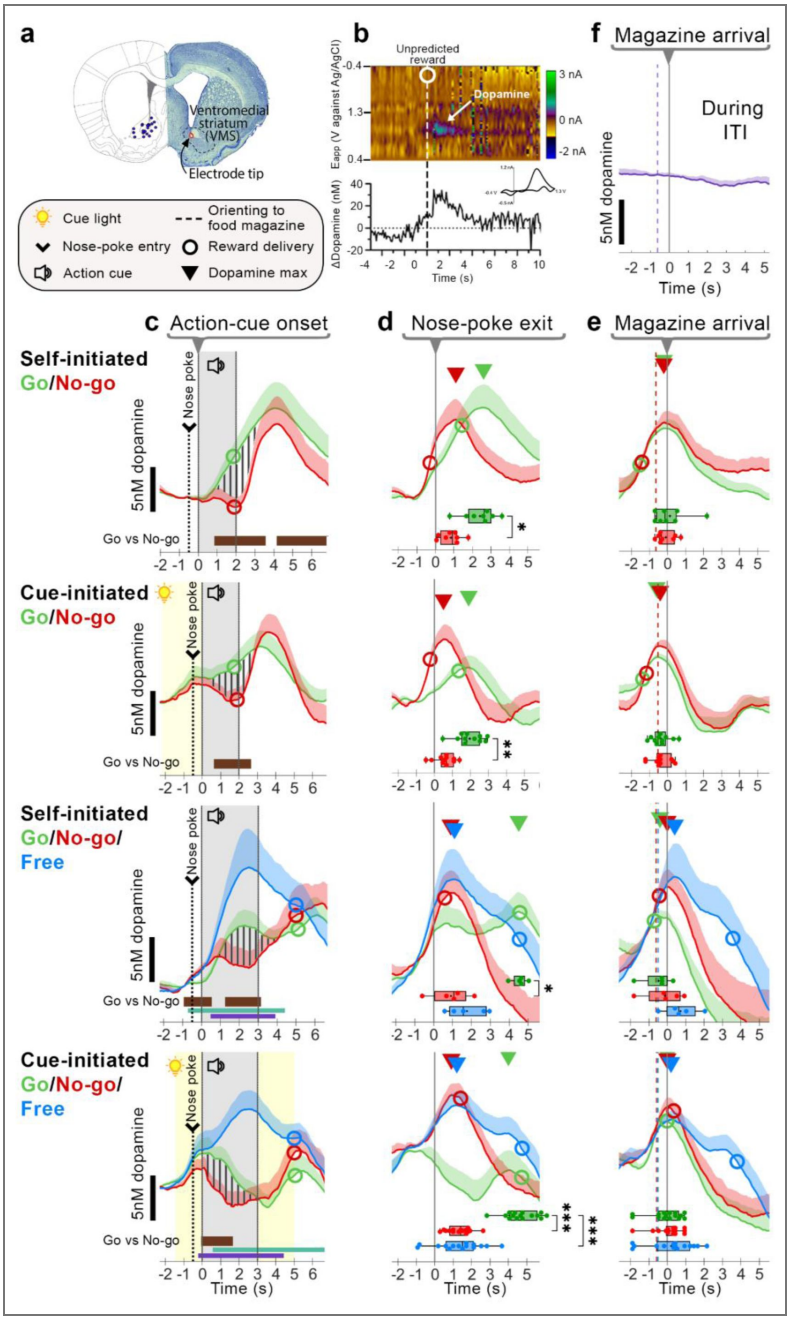


Figure 2. VMS dopamine encodes both action initiation and reward-magazine arrival irrespective of action contingency.

a) Histological verification summary of VMS electrode placements for all animals used in this study (n = 21). **b)** Representative color plot and corresponding dopamine trace in response to an unpredicted pellet. The circle indicates pellet delivery. Inset: Accompanying cyclic voltammogram confirms dopamine detection. ***c-e)*** Average dopamine concentration (nM; mean + SEM) for Go (green), No-go (red), and Free (blue) trials. Triangles denote the average latency-to-maximum dopamine relative to different behavioral epochs. Subject-wise comparisons of dopamine data were made for all alignments. ***d)*** Traces aligned to action-cue onset. Shaded gray area depicts approximate duration of action-cue onset for "short" and "long" task variants. Vertical stripes highlight differences in Go vs. No-go dopamine. Horizontal bars indicate timepoints with significant within-subjects differences between trial types. ***d)*** Traces aligned to nose-poke exit and ***e)*** magazine arrival were analyzed by comparing latency-to-max dopamine release. Box plots depict individual latencies to maximum dopamine release (± 2s) around magazine arrival for each animal. ***f)*** Dopamine release during the inter-trial interval (ITI) as rats approached the reward magazine from the opposite wall, aligned to magazine arrival. Boxplot statistical significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, Wilcoxon matched-pairs signed rank or *post hoc* Dunn's tests.

timing in Free trials matched the timing of long-task Go and No-go trials. In this *Cue-initiated* variant, rats promptly responded to the cue (0.95 ± 0.04 s, Supplementary Figure S1), indicating that they understood the contingency. Success rates did not differ significantly across Go/No-go/Free trial types for both *Self-initiated* (Figure 1d; Kruskal-Wallis test H(2) = 0.400, p = 0.954) and *Cue-initiated* (Figure 1e, H(2) = 1.458, p = 0.482) task variants.

Motivation to approach the reward magazine was similar between Go and No-go trials

To determine whether action initiation (Go) and suppression (No-go) affected motivational drive to approach the reward location, we tracked behavioral performance using DeepLabCut (Mathis et al., 2018). In the short *Self-* and *Cue-initiated* Go/No-go task variants, the latency to orient to the reward magazine (Figure 1f; *Self-initiated*: W = 0, p > 0.999; Figure 1g; *Cue-initiated*: W = 6, p = 0.831), approach speed to the magazine (Figure 1j; *Self-initiated*: W = −5, p = 0.773, Figure 1k; *Cue-initiated*: W = −28, p = 0.240), and latency to arrive at the magazine after action-cue onset (Supplementary Figure S3; *Self-initiated*: W = −4, p = 0.844; *Cue-initiated*: W = 2, p = 0.966) did not differ between Go/No-go trial types.

To assess whether 1) altered controllability of reward pursuit, 2) increased task effort, and 3) the addition of Free trials affected the *core* comparison between action initiation (Go) and suppression (No-go), we examined behavioral performance in the Go/No-go/Free task variants. Trial type significantly influenced latency to orient to the reward magazine in both Go/No-go/Free task variants (Figure 1h; *Self-initiated*: H(2) = 7.600, p = 0.024; Figure 1i; *Cue-initiated*: H(2) = 24.13, p < 0.001). *Post hoc* Dunn's test revealed orienting latencies did not significantly differ between Go and No-go trials (all p > 0.05, Figure 1h, k), but that Free trials significantly differed from Go and No-go trials (*Self-initiated* Go vs. Free: p = 0.034; *Cue-initiated*: p < 0.001, No-go vs. Free: p = 0.002), which was consistent with Free trials requiring no action compared to Go (action) and No-go trials (action suppression). Speed to approach the reward magazine was not significantly different for the *Self-initiated* task (Figure 1l; H(2) = 2.800, p = 0.367) but differed for the *Cue-initiated* Go/No-go/Free task (Figure 1m; H(2) = 19.73, p < 0.001). *Post hoc* Dunn's test revealed that rats moved at similar speeds between Go and No-go trials (*Cue-initiated*), but were significantly faster on Free trials compared to Go and No-go trials (see Figure 1m; p < 0.01), suggestive of increased vigor in Free trials despite matching reward delivery times. In addition, although the latency to arrive at the reward magazine was significantly different between trial types for *Self-* (H(2) = 7.600, p = 0.024) and *Cue-initiated* Go/No-go/Free *tasks* (H(2) = 22.80, p < 0.001), *post hoc* Dunn's test revealed no difference in arrival latencies between Go and No-go trials, but did earlier arrival for Free trials (Supplementary Figure S3). Together, these results indicated that Free trials were associated with earlier approach to the reward location, but the motivational drive to approach was comparable between Go and No-go trials, and was not affected by reward controllability (self- and cue-initiated) nor increasing task effort (short and long-task variant).

VMS dopamine release encodes reward-related action initiation

VMS dopamine recordings (aligned to action-cue onset; Figure 2a,b) showed a rapid increase following presentation of the Go-action cue but no equivalent increase in response to the No-go-action cue (Figure 2c): Within-subject time-series comparisons (Jean-Richard-dit-Bressel et al., 2022) of Go and No-go trials (Figure 2c, '*Self-initiated* Go/No-go', vertically striped area) confirmed significant differences ($p < 0.05$; Figure 2c, thick horizontal bar insets). Despite differences in task requirements such as controllability to attain reward (i.e., *Self-* vs. *Cue-initiated*) and increased effort for Go and No-go trials in the Go/No-go/Free task (i.e., long-task variant, see *Methods*, Figure 1a), within-subject comparisons for all task variants showed a consistent, significant increase in dopamine release on Go trials compared to No-go trials ($p < 0.05$; Figure 2c *Self-* and *Cue-initiated* Go/No-go/Free task, thick horizontal bar insets). In addition, dopamine release was highest for Free trials as compared to Go and No-go trials (p < 0.05; Figure 2c thin horizontal bar insets). Our data suggested that the increase in VMS dopamine release not only

reflected reward-predictive cues but also action initiation. In addition, the difference in VMS dopamine during action initiation (Go) and suppression (No-go) was unaffected by controllability of reward pursuit and increased trial demands.

VMS dopamine release does not only reflect reward-related action initiation

Since the rise in dopamine release for Go trials occurs before rises in No-go trials when traces were aligned to action-cue onset (Figure 2c, vertical line at t=0s), we sought to clarify whether this difference may be due to an earlier moment of action initiation (nose-poke exit) for Go trials as compared to No-go trials. Therefore, we aligned traces to the time when animals departed the nose-poke port (Figure 2d). For Go trials, this was the moment when animals departed the nose-poke port to move to the lever, and for No-go trials, this was the moment animals successfully completed the trial. Analysis of maximum dopamine release (Figure 2d, triangles) differed between trial types: in contrast to No-go and Free trials, maximum dopamine release during Go did not align with nose-poke port exit (Figure 2d). In particular, the latency to maximum dopamine was significantly different between trials, where maximum dopamine release was consistently delayed for Go trials (Figure 2d, boxplot insets; *Self-initiated* Go/No-go: $W = 34$, $p = 0.016$; *Cue-initiated* Go/No-go: $W = 64$, $p = 0.002$; *Self-initiated* Go/No-go/Free: $H(2) = 8.40$, $p = 0.008$; *Cue-initiated* Go/No-go/Free: $H(2) = 23.33$, $p < 0.001$). Overall, our data suggested that differences in Go vs. No-go dopamine could not be attributed to the moment of action initiation (nose-poke departure) alone.

Maximum VMS dopamine release encodes spatial but not temporal proximity to rewards

Previous research has shown that VMS dopamine can not only encode RPE and movement, but can also reflect a subject's proximity to the expected reward location (Hamid et al., 2016; Howe et al., 2013; Mohebi et al., 2019). We therefore realigned dopamine traces to the moment animals arrived at the reward magazine (Figure 2e, vertical line at t=0s). Across all trial types and sessions, dopamine release consistently peaked around the time of magazine arrival (Figure 2e, boxplot insets; *Self-initiated* Go/No-go: $W = 7$, $p = 0.672$; *Cue-initiated* Go/No-go: $W = 3$, $p = 0.916$; *Self-initiated* Go/No-go/Free: $H(2) = 4.526$, $p = 0.111$; *Cue-initiated* Go/No-go/Free: $H(2) = 4.204$, $p < 0.122$). When plotting maximum dopamine as a function of distance to the magazine, the distribution revealed a prominent peak around the magazine-panel (Supplementary Figure S10). Importantly, the timing of maximum dopamine (Figure 2e, *Self-* and *Cue-initiated* Go/No-go/Free, triangles) did not correlate with the timing of reward delivery: in Free trials, average reward-delivery latency (Figure 2e, hollow circles) was delayed following magazine arrival by up to 3s. Realignment of maximum-dopamine concentrations around the time of reward delivery showed that there was no peak distribution around this epoch (Supplementary Figure S11, in particular, see Free trials). In addition, rats oriented to the magazine at similar times, across trial-types and task variants, relative to magazine arrival (Figure 2e; on average 0.6s from orientation to magazine arrival; vertical dashed line). To determine whether maximum dopamine reflected the approach to the reward magazine, or required information about an upcoming reward (e.g., action cues), we analyzed changes in dopamine release during *unprompted* magazine entries in the *Self-initiated* Go/No-go/Free task (Figure 2f, *During ITI*). Our findings demonstrated that unprompted magazine entries during the ITI did not elicit changes in VMS dopamine release. Finally, to determine whether maximum dopamine primarily reflects the earliest predictor of reward, we realigned traces to the moment of action-cue onset in *Self-initiated* trials, and to nose-poke light illumination for *Cue-initiated* trials (Supplementary Figure S12). Our results showed that there was no consistent peak in the distribution of maximum dopamine around cue onset. Together, our results suggested that increases in VMS dopamine required task-relevant features such as knowledge of an upcoming reward and movement towards the reward magazine, where VMS dopamine was maximum upon arrival at the expected reward location.

Motivational state reflected by No-go behavioral strategy correlates with dopamine signal size during reward approach

During the action-cue period of No-go trials, rats had to suppress actions by remaining in the nose-poke port. However, we observed individual differences in behavioral strategies in performing 'action suppression'. We therefore classified No-go trials into three behavioral strategies (*Biting* (purple), *Digging* (teal), *Calm* (maroon); Figure 3a) using a decision-tree classifier based on No-go behavior (Supplementary Figure S4; Video 1-3). Dopamine concentrations during the action-cue period revealed no differences between classifications (Figure 3b, Supplementary Figure S5). This result indicates that behavioral strategy during No-go trials was unrelated to dopamine release *during* the action-cue period, as they demonstrated similar dopamine release for 'action suppression' despite varying No-go behavioral strategies.

However, maximum dopamine concentration measured *after* action-cue offset (Figure 3c) significantly differed between strategies for No-go trials in all task variants (*Self-initiated* Go/No-go: $H(2) = 34.15$, $p < 0.001$; *Cue-initiated* Go/No-go: $H(2) = 16.24$, $p < 0.001$; *Self-initiated* Go/No-go/Free: $H(2) = 22.54$, $p < 0.001$; *Cue-initiated* Go/No-go/Free: $H(2) = 37.48$, $p < 0.001$). *Post hoc* comparisons showed that dopamine levels were consistently higher for *Biting* as compared to *Calm* trials (Figure 3c, *'After action-cue offset: No-go trials'*, all $p < 0.001$). Dopamine was also higher for *Biting* as compared to *Digging* in two task types (Figure 3c). Behaviorally, we observed that Biting trials were characterized by earlier nose-poke hold exits (Supplementary Figure S6). These findings suggest that while dopamine release during the action-cue period is unaffected by action-suppression strategy, differences emerged during reward approach and may reflect differences in motivational drive.

Regrouping animals based on their predominant No-go behavioral strategy revealed a trend in maximum dopamine that resembled individual trial data (Figure 3c, *'No-go (rats)'* and *'No-go (trials)'*, respectively). Therefore, we grouped animals based on their predominant action-suppression strategy to determine whether *Biting* animals also showed elevated dopamine in response to other reward-related information. However, maximum dopamine release in response to an unpredicted pellet did not follow this pattern and instead, they were similar across all groups (Figure 3c, *'Unpredicted pellets (rats)'*). Comparably, grouped rats did not show differences in maximum dopamine for Go and Free trials following action-cue onset (Supplementary Figure S5). These results suggested that the elevated dopamine response during reward approach in *Biting* animals was specific to No-go trials.

To further explore whether VMS dopamine reflects motivational drive within another trial type, we examined Go trials split by the timing of the last lever-press (Supplementary Figure S7). Trials with later last lever-presses (slowest 25% of trials) showed lower maximum dopamine, an effect most pronounced in the long-task variant where lever presses could extend beyond the action-cue period. Together with our classified No-go trial analysis, these results suggested that behavioral indicators of motivational engagement (No-go: *Biting* and nose-poke hold latency; Go: last lever-press latency) modulated maximum dopamine during reward approach.

Discussion

We examined how VMS dopamine integrates motivation and reward-related information during action initiation and suppression in four Go/No-go task variants. In our task, *action initiation* occurred at two distinct points after trial start: 1) when rats began lever pressing (Go), and 2) when rats walked to the reward magazine, either without action requirement (Free) or after successful trial completion (Go and No-go). By contrast, *action suppression* was unique to No-go trials, where animals were required to stay in the nose-poke port. Dopamine release during action (Go) was consistently higher than during action suppression (No-go); this difference was unaffected by controllability of reward pursuit (*Self-* vs. *Cue-initiated*) and differences in effort (i.e., short vs. long). Additionally, our data suggest that VMS dopamine signaled the subjective spatial proximity to reward location. This conclusion is supported by the fact that this gradual increase in dopamine persisted despite differences in the timing of reward delivery (i.e., with or without a 2s-

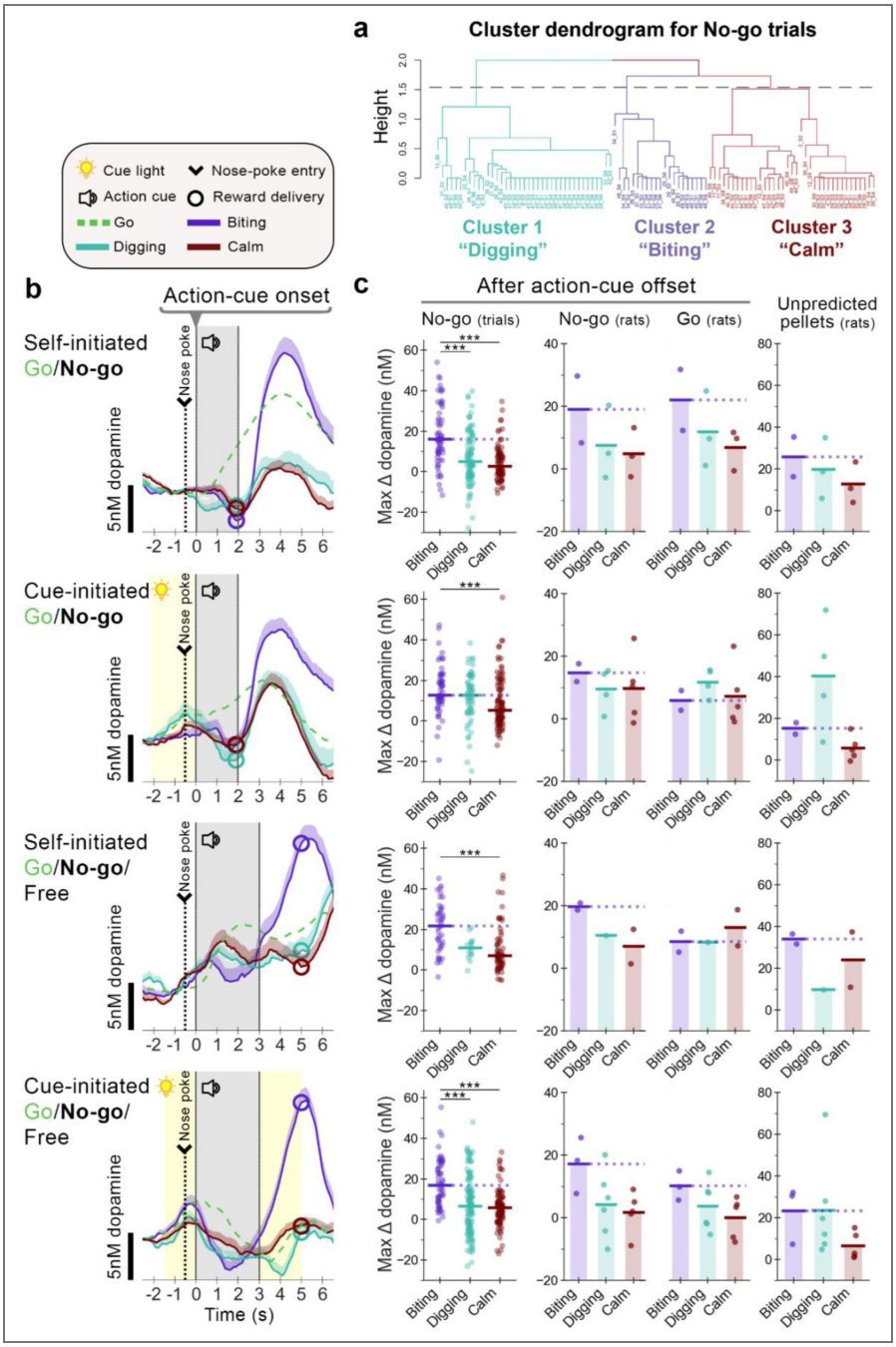


Figure 3. Classified No-go behavior reveals VMS dopamine release following action-cue offset in *'Biting'* trials.

a) Left: Agglomerative hierarchical clustering of behavior during the No-go action-cue period using Ward's linkage and Euclidean distance. A three-cluster solution (dotted grey line) was selected resulting in a balanced distribution of trials across clusters and sessions. These clusters corresponded to three distinct behavioral strategies. *'Biting'* - defined as biting of the nose-poke-port wall; *'Digging'* - digging movements in the nose-poke port; *'Calm'* - all other no-go trials. ***b)*** A supervised decision-tree classifier labeled trials to one of three behavioral groups: *Digging*, *Biting,* and *Calm*. Average dopamine concentration of each classified group is depicted (nM; mean + SEM). ***c)*** Maximum dopamine concentration (nM) following action-cue offset (+2s) was highest in *Biting* trials as compared to other No-go trials. *'After action-cue offset: No-go (trials)'*: Individual No-go trials classified by No-go behavior, and trial-wise statistical Kruskal-Wallis tests were performed for each task-variant. Solid lines depict the median. *'After action-cue offset: No-go (rats)', 'After action-cue offset: Go (rats)',* and '*Unpredicted pellets (rats)*': Rats classified based on their predominant No-go strategy, with no statistical tests performed. Solid lines depict the mean. Significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, *post hoc* Dunn's tests.

delay after action), changes in controllability, and preceding action requirement. Our results indicate that in a reward context, VMS dopamine is not only associated with action execution but continuously tracks spatial, but not temporal, proximity to rewards. Unexpectedly, consummatory Pavlovian behavior in the form of (wall) biting during action suppression led to markedly larger dopamine release during subsequent reward-magazine approach than other behaviors. Together, our findings indicate that VMS dopamine sequentially integrates reward-driven action initiation with the gradual increase in spatial reward proximity.

Reward-related dopamine depends on action initiation irrespective of controllability and effort

VMS dopamine has been associated with the invigoration and initiation of reward-seeking behavior (Ikemoto and Panksepp, 1999; Nicola, 2007; Robinson and Berridge, 1993) as demonstrated by both bulk measurements of VMS dopamine release (Collins et al., 2016; du Hoffmann and Nicola, 2014; Hamid et al., 2016; Mohebi et al., 2019; Nicola, 2010; Phillips et al., 2003; Roitman et al., 2004; Syed et al., 2016; Wassum et al., 2012) and the activity patterns of individual midbrain dopamine neurons (Coddington and Dudman, 2018; Jin and Costa, 2010). Across all four of our task variants, VMS dopamine release was consistently higher during action (Go) than during action suppression (No-go). Thus, by extending the pioneering work by Syed and colleagues (Syed et al., 2016), our findings provide compelling evidence that reward-related dopamine release in the VMS is a function of associated reward-driven action.

In addition to action initiation, recent work suggests that striatal dopamine dynamics are shaped by the controllability of reward contingencies (Bernklau et al., 2024; Goedhoop et al., 2023; Hamid et al., 2021). For instance, the anticipation of operant action initiation during reward-cue presentation is accompanied by more sustained VMS dopamine release compared to cues that signal reward delivery without action requirement (Goedhoop et al., 2023). Similarly, midbrain dopamine recordings showed increased responses to externally triggered actions (i.e., *Cue-initiated*) than to *Self-initiated* actions for rewards (Romo and Schultz, 1990). In contrast, our data indicate that manipulations of controllability in reward pursuit (i.e., *Self-* vs. *Cue-initiated* trials) did not affect the contrast between Go and No-go dopamine dynamics during action execution, but instead only elicited dopamine release at an earlier time point in the trial, presumably triggering an earlier positive RPE in *Cue-initiated* trials (Figure 2a). This discrepancy to previous studies may be rooted in differences in experimental design. Goedhoop and colleagues (2023) compared VMS dopamine during operant versus Pavlovian conditioning, similar to the comparison of Go and Free trials in our task, respectively, which also display differences in VMS dopamine (Figure 2c). On the other hand, the discrepancy may also be rooted in different definitions of "controllability". Here, we defined controllability as the moment animals can initiate reward pursuit (i.e., self-vs cue-initiated), which considers a different aspect of controllability. Thus, we cannot exclude the possibility that dopamine would respond to a more extreme manipulation of controllability Indeed, Bernklau and colleagues (2024) showed that striatal dopamine reflects the animal's perceived locus of control, which shifts from external cues toward the animal as the causal agent with experience. In our paradigm, both self- and cue-initiated task variants may have been equally perceived as agent-controlled (albeit curtailing timing and source of action initiation), but potentially leaving the inferred locus of control unchanged. Together, our results indicate a prominent role for VMS dopamine in reward-related action (Coddington and Dudman, 2018; Ikemoto and Panksepp, 1999; Nicola, 2007) that is not sensitive to moderate changes in task controllability.

Previous studies have shown that dopamine signals may be influenced by the effort required to obtain reward but only for particular task conditions (Cousins et al., 1996; Gan et al., 2010; see for reviews, Salamone and Correa, 2024; Walton and Bouret, 2019). We tested this by varying response demands: short and long Go trials requiring either two lever presses or repeated pressing for three seconds, respectively. No-go trials required matching two or three seconds of action suppression, respectively (Fig 1b). However, such effort manipulations had no detectable impact on action-related dopamine: the relative contrast in VMS dopamine in Go versus No-go

trials did not differ between short and long task variants, suggesting that dopamine does not scale with response duration or physical effort. This is consistent with some previous work (Gan et al., 2010; Walton and Bouret, 2019).

The addition of Free trials in the current study provides critical insight into the role of dopamine in motivated action directed at reward. In Go/No-go tasks designed for human and non-human primates, Go trials often require an overt action (e.g., pressing a button) and No-go trials require passive inaction (e.g., not pressing the button; e.g., Raud et al. 2020). In the present study, Go and Free trials were akin to human Go and No-go trials, respectively. Importantly, our findings demonstrate that VMS dopamine dynamics are highly distinct between action initiation (Go), active inaction (No-go), and passive inaction (Free), for rewards. Therefore, precise definitions of "No-go" trials in both human and animal models are necessary for cross-species translatability of the neural basis of motivated behavior.

VMS dopamine release reflects spatial proximity during approach of reward location

Beyond physical effort, the long-task variants imposed a two-second reward delay after action-cue offset, permitting clearer separation of dopamine signaling during action-cue and reward-approach epochs. Critically, maximum dopamine responses across all four task variants, could not be explained solely by action initiation, as the latency to maximum dopamine differed between trial types when aligned to nose-poke exit (Figure 2d, triangles), diverging from previous reports (Syed et al., 2016). Instead, dopamine release during the long-task variant revealed a gradual increase during magazine (reward location) approach that peaked upon arrival at the magazine, irrespective of whether rewards were received immediately or were delayed (Figure 2d). Importantly, we did not observe an increase in VMS dopamine when animals approached and checked the magazine between trials, indicating the necessity of reward expectation linked to trial performance (Figure 2e). In addition, the signal cannot be attributed to orientation to the reward-magazine - if orientation were the cause, dopamine would not be expected to increase while animals maintain a constant orientation when moving toward the reward magazine. Overall, this finding is consistent with several studies reporting a gradual rise in VMS dopamine, frequently referred to as ramps, as animals actively (Farrell et al., 2022; Hamid et al., 2016; Howe et al., 2013; Krausz et al., 2023; Mohebi et al., 2019) or passively (Farrell et al., 2022; Guru et al., 2020; Kim et al., 2020; Mikhael et al., 2022) move toward reward. This phenomenon was particularly pronounced during Free trials, where dopamine peaked when rats arrived at the reward magazine up to five seconds before reward delivery (Figure 2c,d Go/No-go/Free, circles). Together, our data suggest that while VMS dopamine initially facilitates the initiation of reward-driven action, it subsequently tracks progression toward a reward-associated location, which may serve to sustain motivational drive during reward approach.

The gradual increase in VMS dopamine during reward approach, peaking upon arrival, has been suggested to reflect ongoing state estimation under uncertainty (Farrell et al., 2022; Gershman and Uchida, 2019; Kim et al., 2020; Mikhael et al., 2022). Alternatively, this gradual increase, or ramp, has been suggested to reflect an 'internal model' that continuously updates reward-proximity estimates as animals progress toward the goal (Farrell et al., 2022; Guru et al., 2020). Our findings support the proximity account. In our Go/No-go task, rats traversed the same familiar, relatively short path to the magazine in thousands of trials, minimizing state uncertainty and reducing the need for ongoing state estimation updates. Instead, our findings suggest that VMS dopamine maintained a continuous estimate of spatial reward proximity: arrival at the reward magazine following trial completion served as the most reliable predictor of reward and coincided with maximum dopamine release, consistent with previous findings (Hamid et al., 2016; Howe et al., 2013; Mohebi et al., 2019). While our results clearly demonstrate dopamine signaling related to spatial proximity to rewards, our experimental design cannot definitively distinguish whether VMS dopamine encodes RPE computations or spatial proximity to sustain motivated approach behavior. More modeling and experimental work is needed to test these predictions (but see, Gershman et al., 2024; Lerner et al., 2021).

Dopamine dynamics are linked to motivated action

Our findings reveal that VMS dopamine follows a consistent temporal pattern across trials: dopamine increases when animals initiate reward-directed actions, then continuously tracks spatial reward proximity. Critically, the magnitude of this proximity signal was modulated by motivational engagement, measured by trial-type-specific behavioral indicators (i.e., Go lever-press latencies and No-go behavioral strategies) and cannot be solely explained by changes in vigor. Go trials, in which lever-press requirements were completed faster, were associated with an earlier magazine approach and a larger and earlier increase in dopamine during the subsequent magazine approach than Go trials exhibiting slower action completion (Supplementary Figure 7). In No-go trials, we identified a distinct behavioral phenotype: consummatory Biting responses directed at the wall around the nose-poke device during action suppression were associated with shorter nose-poke hold latencies (Supplementary Figure S6) and an approximately two-fold increase in dopamine during subsequent reward approach, as compared to non-biting strategies across all task variants (Figure 3b). Notably, the increased dopamine signal (associated with earlier lever-press completion (Go) and Biting (No-go)) was specific to reward approach, and was not related to the action-cue period itself (Supplementary Figure S5-S7). Two lines of evidence argue against (Go and No-go) action vigor explaining these dopamine changes. First, the modulation of reward-approach dopamine occurred after the action-cue period, dissociating it temporally from any action vigor expressed during lever-pressing or nose-poke holding. Second, magazine approach-speed, which is a direct measure of post-action-cue vigor (i.e., behavior occurring at the same time as the observed differences in dopamine), showed no consistent relationship with dopamine magnitude across trial classifications (Supplementary Figures S5–S7).

Although Biting responses, which resembled consummatory behaviors typically directed toward rewards, shared similarities with so-called 'sign-tracking' behavior (Flagel et al., 2009; Hearst and Jenkins, 1974), several features distinguish our observations from sign tracking. First, sign tracking is promoted when reward-predictive cues and reward location are in spatial proximity (Christie, 1996; Flagel et al., 2009). However, in our experimental setup, reward-related manipulanda (i.e., levers, nose-poke port) and reward magazine, were positioned on opposing walls of an operant chamber exceeding standard size (Figure 1a). Second, it has been reported that rats do not readily sign-track to nose-poke ports or auditory cues (Beckmann and Chow, 2015; Meyer et al., 2014), both critical for No-go trials in our paradigm. Third, prior work demonstrated enhanced VMS dopamine during Pavlovian behavior towards the cue (Flagel et al., 2011, 2009), whereas we observed amplified dopamine *after* action suppression (and after the cue presentation), during reward approach.

These distinctions suggest that Biting may reflect elevated Pavlovian drive and/or difficulty suppressing reward-directed action, likely due to enhanced motivation (Robinson et al., 2014), rather than sign tracking. Accordingly, our findings point at inter-individual differences in behavioral strategy, analogous to sign and goal-tracking behaviors (Flagel et al., 2011; Robinson et al., 2014). Animals exhibiting heightened task engagement or Palovian drive showed amplified dopamine responses to identical changes in spatial reward proximity, suggesting that encoding of spatial-reward proximity is dynamically modulated by motivational state.

A potential explanation for differences in VMS dopamine may be attributed to Pavlovian biases, wherein reward-associated cues promote approach over suppression, typically resulting in higher Go than No-go performance accuracy. We indeed observed this behavioral asymmetry in our standard Go/No-go task; however, it disappeared when Free trials were included (Figure 1c). Importantly, VMS dopamine continued to differentiate between Go and No-go trials despite the absence of this phenomenon (i.e., more VMS dopamine during the action-cue period for Go than No-go; Figure 2c). Therefore, our data suggest that Pavlovian bias cannot account for the observed differences in dopamine dynamics.

Conclusions

The findings of the present study are consistent with a general role of VMS dopamine in both RPEs and action. However, we further demonstrate that the differential between action and action suppression is not influenced by behavioral controllability or effort. Rather, our results suggest that VMS dopamine primarily encodes the initiation of reward-seeking actions and spatial proximity to rewards, thereby bridging two traditionally separate fields investigating the roles of dopamine in movement and in reward processing, respectively. Building on previous reports that individual dopaminergic neurons multiplex sensory, motor and cognitive information (Engelhard et al., 2019; Kremer et al., 2020), our bulk dopamine-terminal measurements suggest this multiplexing is extended to the broader population level. These findings suggest that population-level VMS dopamine signals serve as a unifying mechanism, conveying information about both motivated action and reward to postsynaptic VMS neurons. We propose that within this framework, VMS dopamine links motor and motivational processes to ensure that actions are maintained until reward receipt.

Methods

1. Animals

At the start of the experiment, adult male Long-Evans rats (250-350g, Janvier Labs) were individually housed on a reversed 12 h light/dark cycle (lights off from 08:00 to 20:00) with controlled temperature and humidity. Water was available *ad libitum*. Following surgical procedures and one week of recovery, rats were food-restricted to 85% of their free-feeding body weight. All animal procedures were performed during the dark phase (between 11:00-19:00) in accordance with Dutch and European law, and approved by the Animal Experimentation Committee of the Royal Netherlands Academy of Arts and Sciences. A total of 21 rats were used in the experiments described in the following sections, where each rat had at least one functional and histologically verified recording electrode.

2. Stereotaxic surgery

Stereotaxic surgery was performed as previously described (Willuhn et al., 2012). Rats were given an analgesic (Metacam, 1-2 mg/kg, s.c.), anesthetized with isoflurane (2-3% maintenance), and placed in a stereotactic frame. Body temperature was maintained at 37°C during surgery using a heating pad and monitored with a probe. The scalp was shaved, disinfected with 70% alcohol, and a midline incision was made. The incision site was treated with lidocaine (100 mg/ml) and the periosteum gently removed to expose the cranium. Holes were drilled for 3-4 anchor surgical screws, two custom-made carbon-fiber microelectrodes(Clark et al., 2010) unilaterally (right hemisphere) targeting the VMS (A-P: +1.2, M-L: +1.5, D-V: −7.1 mm, relative to Bregma (Paxinos and Watson, 2006), and an Ag/AgCl reference electrode that was positioned separately in the forebrain. Electrodes were secured and anchored to the surgical screws with dental acrylic cement. Following surgery, rats were subcutaneously injected with 2 mL saline and placed in a temperature-controlled cabinet until they exhibited mobility. Following implantation of the electrodes, rats were individually housed and were left to recover for at least one week. The location of VMS electrodes can be found in Supplementary Figure S8.

3. Behavioral procedures

Three days before starting behavioral training in the operant boxes, 10-15 sugar pellets/day were introduced into the home cage. Behavioral experiments were performed in modified operant boxes (32 x 30 x 29 cm, MedAssociates Inc.), equipped with a nose-poke hole with an integrated cue-light, house light, auditory-tone generator, food-reward magazine, metal-check grid floor, and retractable levers (Figure 1a). Operant boxes were interfaced with a fast-scan cyclic voltammetry (FSCV) setup for dopamine recordings of awake and freely-moving animals and a

video camera to survey behavior. The operant boxes were stored in individual Faraday cages insulated with sound-absorbing polyurethane foam and ventilated by a fan. Before commencing training or recordings in the operant boxes, rats were habituated to the room for at least 10 min.

Self-initiated task variant

Self-initiated Go/No-go task ("short"; n = 9): Using a behavioral paradigm inspired by Syed and colleagues (Syed et al., 2016), rats initiated a trial by entering the nose-poke port for at least 0.5 s. One of two action (auditory) cues were played, instructing rats to either move towards and press on a unilaterally extended lever at least twice within 5 s ('Go'; 3 kHz) or to remain in the nose-poke hole for another 1.7-1.9 s ('No-go'; 1 kHz). After five correct Go trials, the contralateral lever extended and the opposing lever retracted. Successful trial completion was marked by the offset of the action cue and food pellet dispensed into the reward-magazine immediately. During training and recording sessions, trials were separated with a fixed intertrial interval (ITI) of 5 s.

Self-initiated Go/No-go/Free task ("long"; n = 5)

Go and No-go trial requirements were increased, together with the inclusion of a third trial type (long-task variant). During training, Go trials required an initial lever press within 1.7 s of action-cue onset and a subsequent lever press within the next 1.5 s. During recording days, a correct Go trial occurred when rats made the first lever press between 2.7 to 3.2 s after cue onset. For No-go trials, rats had to stay in the nose-poke hole for at least 2.7 to 3 s after cue onset. Following successful training of the extended Go and No-go trials, a "Free" trial was introduced (8 kHz tone). Rats were permitted to do anything but nose poke or lever press during the Free cue; most rats walked directly to the food-magazine following cue onset. Rewards were delivered with a delay of 2 s following successful trial completion. Trials were separated with a variable ITI between 15-25 s.

Failure to perform the trial's required action resulted in house-light illumination for 5 s (i.e., incorrect trial). During training (i.e., non-recording days), the trial type was repeated to a maximum of three times or until the trial was correctly completed, whichever came first. During dopamine-recording days, trial type repetition did not occur. Success rates for each trial type was calculated by determining the percentage of correct/(correct+incorrect) trials, with incorrect trials only included if the error was made after action-cue onset.

Failure to hold the snout in the nose-poke port during the first 0.5 s before action-cue onset resulted in an immediate house-light illumination for 5 s and presentation of another ITI.

Cue-initiated task variant

The structure of *Cue-initiated* trials was similar to that of *Self-initiated* trials, except that they could only be initiated following the illumination of the nose-poke light. The nose-poke light stayed illuminated for up to 15 s until initiation, and remained illuminated until the end of the trial. All other trial requirements were identical, except that the house light illuminated for 5 s and a new ITI began when rats made a nose poke or lever press during the ITI, or if they did not perform a nose-poke within 15 s of nose-poke light illumination at the start of the trial. A total of n = 11 and n = 15 were included in the *Cue-initiated* Go/No-go and *Cue-initiated* Go/No-go/Free tasks, respectively.

4. Real-time dopamine recordings and analysis

Simultaneous dopamine and video recordings were carried out when rats were able to perform with >60% success rate for each trial type on at least two consecutive training sessions. On recording days, the trial types were counterbalanced. Within a session, Go left, Go right, and No-go trials were presented with 33% probability each, without replacement. This design results in overrepresentation of Go trials, which establishes the prepotent Go response. For sessions with Free trials, trials were presented with a 25% chance without replacement.

FSCV measurement and analysis

Fast-scan cyclic voltammetry (FSCV) at chronically implanted carbon-fiber microelectrodes (Clark et al., 2010) was used to record rapid changes in extracellular dopamine concentration. Prior to recording, microelectrodes were connected to a head-mounted voltammetric amplifier interfaced with a PC-driven data acquisition and analysis system (National Instruments) through an electrical commutator (Clark et al., 2010). Voltammetric scans were repeated every 100 ms (i.e., 10 Hz sampling rate). An alternating potential at the carbon-fiber electrode tip was applied from –0.4V versus Ag/AgCl to +1.3V (anodic sweep) and back to –0.4V (cathodic sweep) at 400V/s (total scan time of 8.5ms), and held at –0.4V between scans. Dopamine traces were isolated from the voltammetric signal by chemometric analysis using a standard training set, based on electrically stimulated dopamine release detected with chronically implanted electrodes (Clark et al., 2010). All data were smoothed with a 10-point median filter and baseline (set at 1s before event of interest, with the exception of 2.5 s before action-cue onset for *Cue-initiated* Go/No-go and Go/No-go/Free tasks) subtraction was performed on a trial-by-trial basis prior to analysis. Dopamine data were aligned to events of interest: action-cue onset, nose-poke exit, and magazine arrival.

Histological verification of recording sites

Following completion of behavioral and recording experiments, animals were terminally anesthetized using sodium pentobarbitone. An electrolytic lesion was performed by running a current on the microelectrodes to mark the FSCV recording sites. Rats were transcardially perfused with saline (0.9% NaCl, w/v) and paraformaldehyde (4% in PBS, w/v). Whole brains were collected and postfixed in Parafresh for 24 h, cryoprotected in sucrose solution (30% in PBS w/v), and cryosectioned coronally (40 µm). Slices were stained with cresyl violet. All animals included in this study had electrodes positioned in the ventromedial striatum (VMS; Supplementary Figure S8).

5. Video recordings and analysis

DeepLabCut pose estimation

For detailed quantification of behavior across each dataset, we employed DeepLabCut (DLC) to estimate body part coordinates of each animal (Mathis et al., 2018; see Supplementary Figure S9). A total of 1500 manually labelled frames from various videos were used to train a ResNet-50 based neural-network.

Labelled body parts included the snout, neck, four back-points, and tail-base. A reference frame from each video was labelled manually to determine the coordinates of the arena floor. Body coordinates were then transformed from pixels into cm using the 'fitgeotform2d' function in MATLAB. Custom analyses were developed in Python to filter DLC coordinates and to analyze behavioral features. Specifically, likelihood cut-off scores for each body part were set at 0.9, large movements in x and y coordinates were removed between each consecutive frame (> 1.5 cm between frames), body-part speeds exceeding a specific speed (mean speed of 9 consecutive frames + 2*SD), and linear interpolation was applied between the nearest valid x and y coordinates to fill in gaps in the data. Data then underwent Gaussian smoothing to reduce noise. This smoothing was performed using a Gaussian window with a SD calculated to capture approximately 99% of the Gaussian probability density function. The smoothing process was applied within a sliding window of 9 frames (29.97 fps) corresponding to approximately 300 milliseconds.

DeepLabCut behavior analysis

Orientation towards the magazine was determined by calculating the angle between the vector from the neck to the snout, and extending this vector to the wall containing the magazine-panel. Frames of animals were labelled as oriented when this vector crossed the length of the magazine-panel. Magazine arrival was determined as the moment the neck was <2 cm from the magazine-panel.

To analyze magazine approach during the ITI, we used DLC-derived coordinates to identify instances where animals were at least 25 cm away from the reward magazine and subsequently moved towards the magazine within 2s (similar to reward approach following a successful trial). This was performed in the *Self-initiated* Go/No-go/Free task, as animals were permitted to interact with the nose-poke hole and lever during the ITI before orienting and approaching the magazine. Only the second approach and subsequent attempts were included in the analysis. Dopamine traces were then analyzed for these specific moments.

Characterization of No-go behaviors

Behavioral strategies employed by rats during No-go trials were manually scored by an experimenter to systematically assess four main features: vigor (scored 0 to 2), and binary indicators for turning, digging, and biting. To characterize vigor, this was based on a scale of 0 to 2: *0* indicating that the animals were not moving, *1* indicating some energy in their movement, and *2* indicating vigorous movement. Turning was defined as the rat rotating the axis of its body while maintaining their snouts in the nose-poke port. Digging was defined as the repeated placement of their paw(s) into the nose-poke hole. Biting was defined by biting the wall of the nose-poke port.

6. Statistical Analysis

Data were processed using custom scripts written in MATLAB and Python, and only correct trials were included. Behavioral data were analyzed using GraphPad Prism (v 10.3.1). Due to violations of normality assumptions, a non-parametric Wilcoxon matched-pairs signed rank test was used for pairwise comparisons (trial type) within each task. Kruskal-Wallis tests were used to examine differences across three trial types, and statistical significance was followed-up with *post hoc* pairwise comparisons using Dunn's test. All tests were two-tailed and statistical significance was set at $p < 0.05$.

Dopamine time-series analysis

Dopamine traces within each recording session were directly compared against each other by bootstrapping the within-subject difference trace (e.g., mean Go trace - mean No-go trace within-subject; Jean-Richard-dit-Bressel et al., 2022). Bootstrapped means were obtained by randomly resampling from subject mean dopamine traces with replacement (1000 iterations). CI limits were derived from 2.5 and 97.5 percentiles of bootstrap distribution, expanded by a factor of $\sqrt{\frac{n}{n-1}}$ (Jean-Richard-dit-Bressel et al., 2020). Significant transients were identified as a period that CI limits did not contain 0 for at least 1 sec (Jean-Richard-dit-Bressel et al., 2020). To determine the timing of maximum dopamine release for realigned traces, we searched for the maximum dopamine concentration −2 and +2 s around magazine arrival.

Clustering No-go behaviors and regrouping animals

Prior to clustering, manually scored trial features were normalized. Averages of characterized No-go scores from each animal across all sessions were then clustered into three groups using agglomerative hierarchical clustering with Ward's linkage method and Euclidean distance as the metric. A decision-tree classifier was implemented in MATLAB to classify labelled trials into their respective groups, and digging and biting were key features identified for group classification (Supplementary Figure S4). For Go/No-go tasks, average dopamine was calculated within a 2s window following action-cue onset; and for Go/No-go/Free tasks, within 3s after action-cue onset. Similarly, to determine maximum dopamine release following action-cue offset, values were determined within 2s for *Go/No-go* tasks and 3s for Go/No-go/Free tasks. Individual trial data were statistically compared using Kruskal-Wallis tests and statistical significance followed-up with *post hoc* Dunn's test.

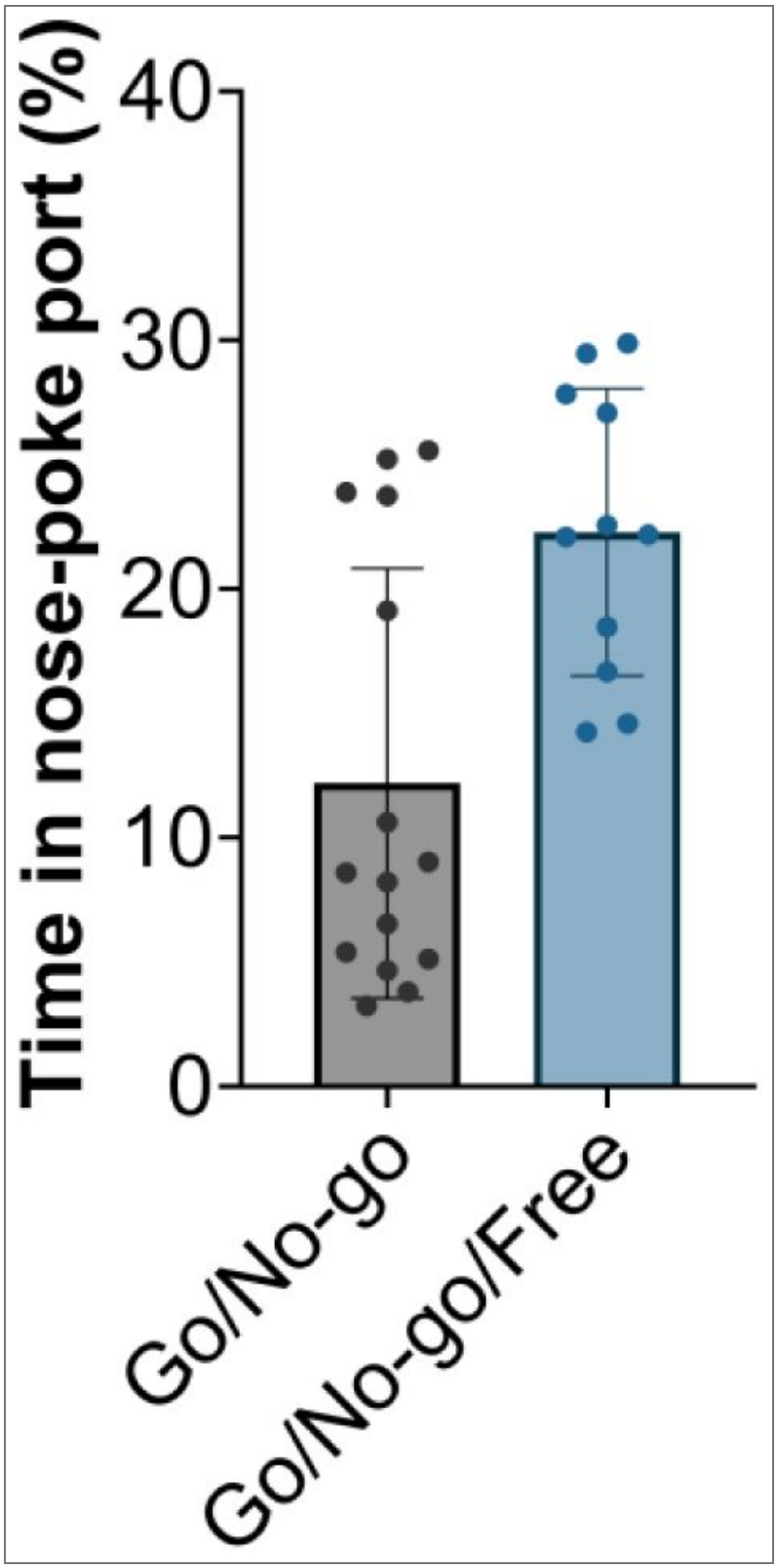
An additional analysis was performed in which animals were regrouped based on their predominant action suppression style within the task (i.e., *Biting*, *Digging* or *Calm*). Regrouped animals were used to compare average and maximum dopamine responses across trial types and in response to unpredicted pellets. Due to limited sample sizes (≤ 2) within certain subgroups, the resulting data were qualified.

Task variant	Animals with FSCV	Total (Biting, Digging, Calm)	
		Trials	Animals with video
Self-initiated Go/No-go	**9**	**229** (61, 89, 79)	8 (2, 3, 3)
Cue-initiated Go/No-go	**11**	**225** (57, 64, 104)	11 (2, 4, 5)
Self-initiated Go/No-go/Free	**5**	**117** (41, 15, 61)	5 (2, 1, 2)
Cue-initiated Go/No-go/Free	**15**	**284** (57, 127, 100)	14 (3, 6, 5)

Table 1. Number of subjects per task variant with FSCV recordings and number of trials in each No-go behavioral classification.

Statistical analyses were performed between classified No-go *trials* due to limited sample sizes per group.

Supplementary figures

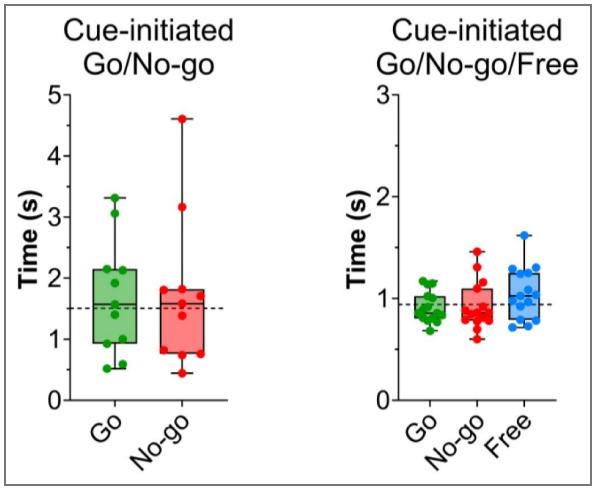


Supplementary Figure S1. Rats did not continuously occupy the nose-poke port during self-initiated intertrial intervals (ITIs). Percentage of time spent in the nose-poke port between trials of *self-initiated* Go/No-go and Go/No-go/Free task variants. Animals were permitted to enter the nose-poke port during the ITI but entry

could not trigger action-cue presentation. The ITI for the *self-initiated* Go/No-go task was fixed at 5s, whereas for Go/No-go/Free task, it was varying between 15-25s.

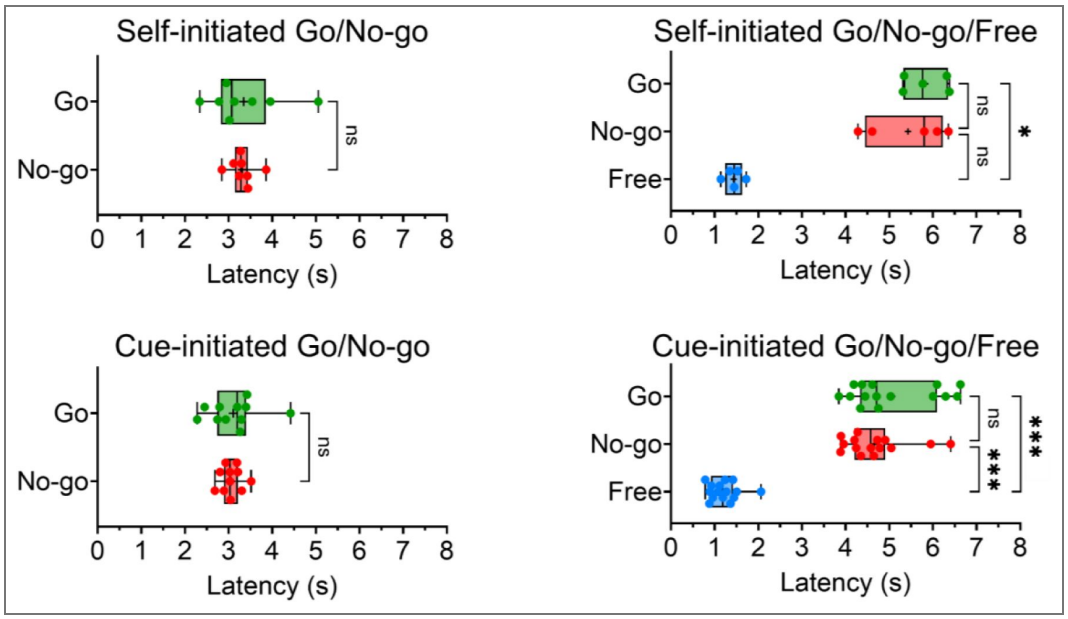
Supplementary Figure S2. Latency to initiate trials in *Cue-initiated* task versions.

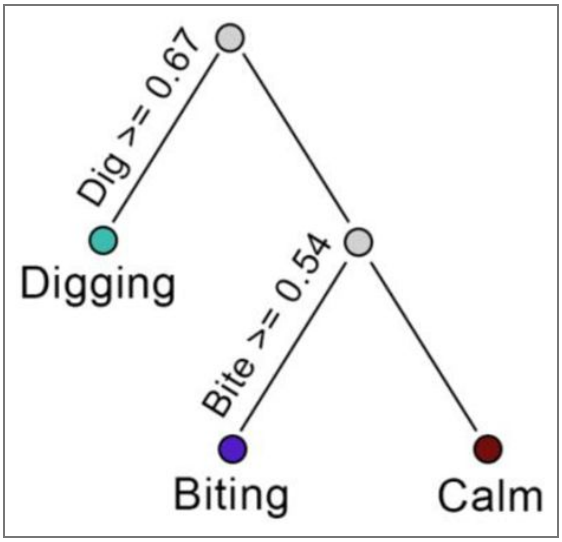
The latency to initiate a trial **following nose-poke light illumination** was similar across trial types within task variants. *Cue-initiated* Go/No-go: W = 18, p = 0.47; *Cue-initiated* Go/No-go/Free: H(2) = 0.93, p =v 0.63. Horizontal dotted lines depict the median response latency to start trials after the cue turned on.



Supplementary Figure S3. Latency to arrive at the reward magazine was not significantly different between Go and No-go trials.

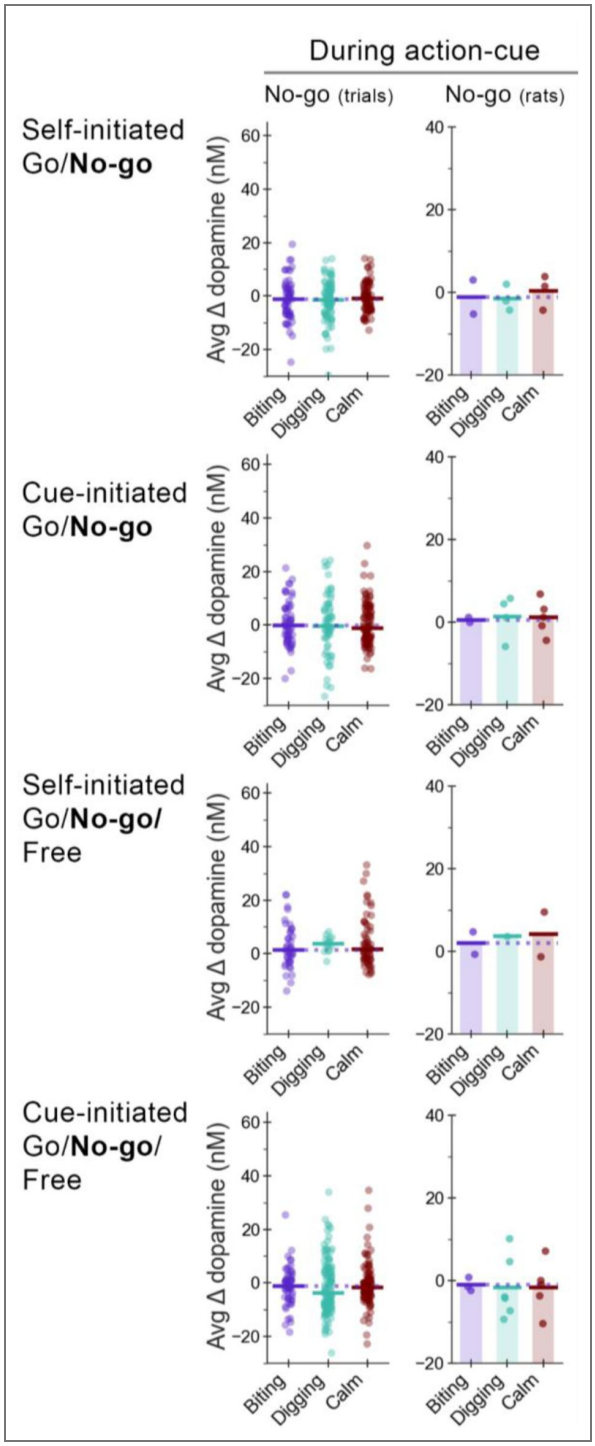
After satisfying response requirements in Go and No-go trials, rats arrived at the reward magazine at similar time points. During Free trials, animals arrived at the magazine within 2s of hearing the action-cue tone. Action-cue onset is at 0s. Significance for *post hoc* Dunn's tests: * p < 0.05, ** p < 0.01, *** p < 0.001.





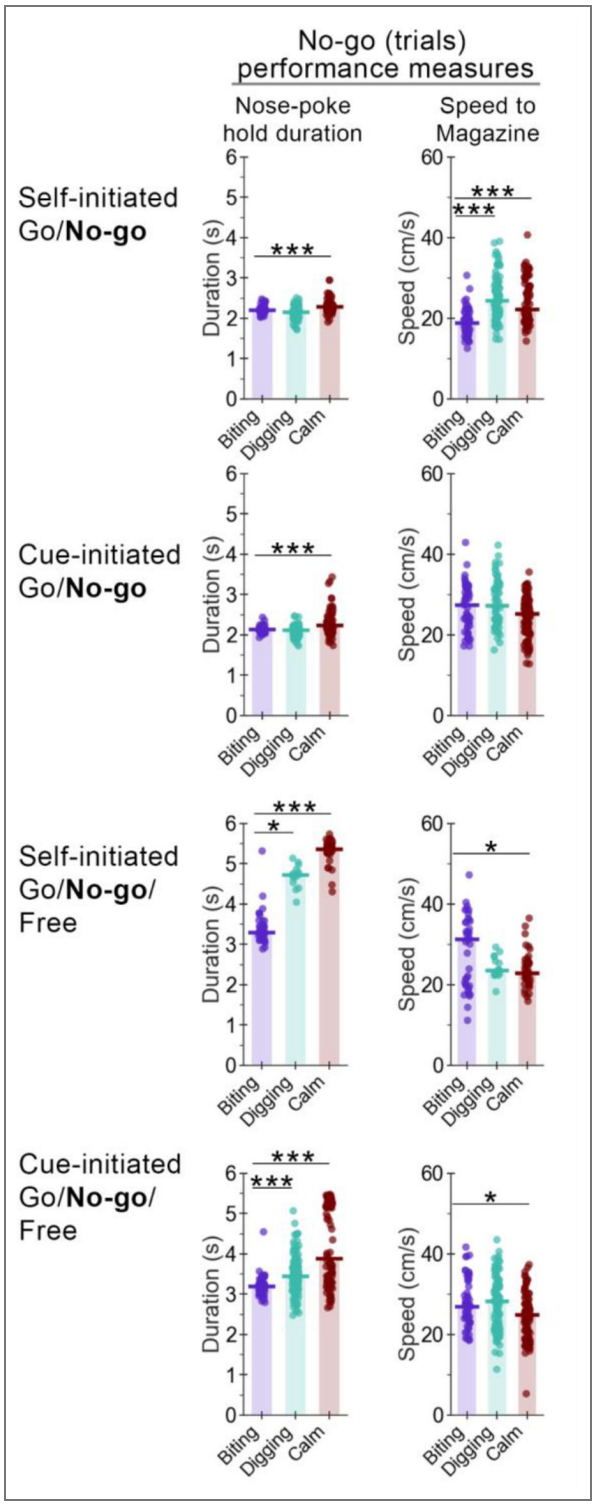
Supplementary Figure S4. Decision-tree classification of No-go trials.

Individual trials (*n* = 855) were categorised into one of three groups.



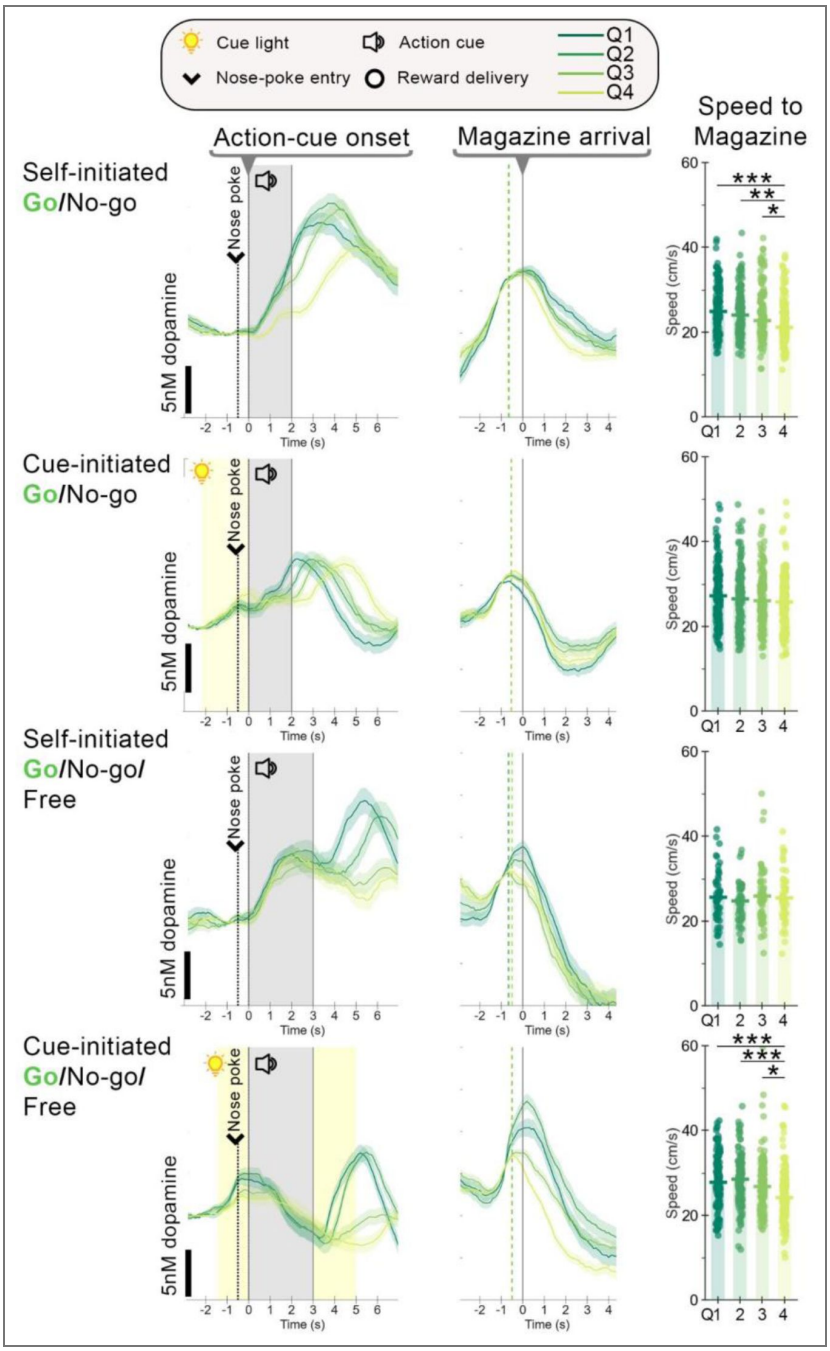
Supplementary Figure S5. Classified No-go behaviors reveal no changes in dopamine during the action-cue epoch.

Column 1) Dopamine concentration during action-cue presentation did not differ significantly between no-go trial classifications (all $p > 0.05$; Kruskal-Wallis test statistics in Statistics table). Solid horizontal lines depict the median. *Columns 2-4)* Rats were grouped according to their predominant No-go strategy (see Methods; horizontal lines depict the mean). *Column 2)* Differences in average dopamine in No-go animals did not qualitatively differ between groups.



Supplementary Figure S6. ***Biting*** **trials were characterized by earlier nose-poke exits but inconsistent magazine approach speeds, ruling out a generalized vigor explanation for reward-approach dopamine changes.**

Column 1) Latency to exit the nose-poke port from action-cue onset. Across Go/No-go/Free task variants, *Biting* trials consistently exhibited the shortest nose-poke hold times. *Column 2)* Speed to approach the reward magazine after nose-poke exit. While Kruskal-Wallis tests were statistically significant for all comparisons ($p < 0.05$; see supplementary statistics table), magazine approach speed showed no consistent relationship with *Biting* trials relative to other classifications. Together, the results rule out 'vigor' during magazine approach as the sole driver of reward-approach dopamine changes. *Post hoc* Dunn's test significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

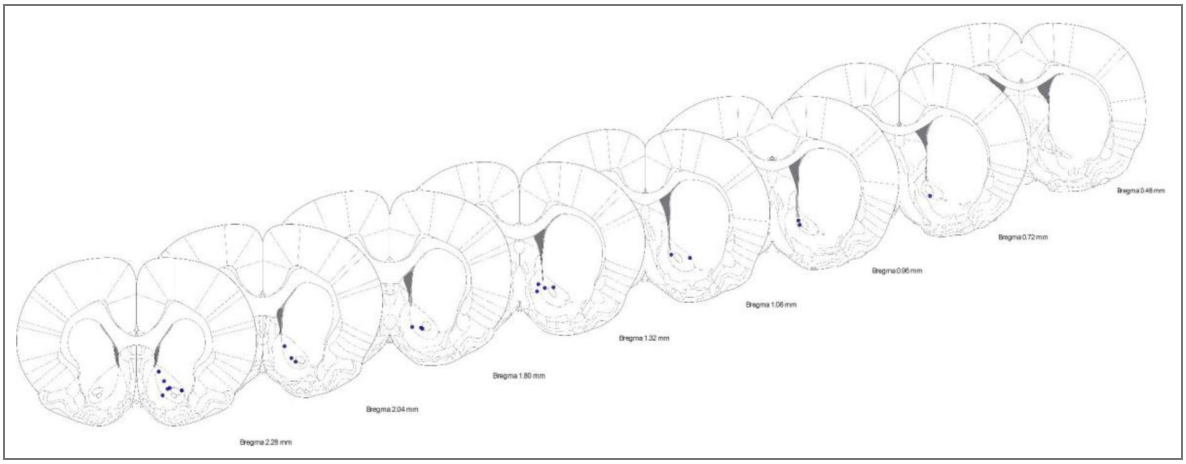


Supplementary Figure S7. VMS dopamine during Go trials tracks magazine arrival after the final lever-press and is independent of vigor.

For each task variant, trials were divided into quartiles based on last lever-press latencies (Q1 = fastest 25%; Q4 = slowest 25%). *Column 1)* When aligned to action-cue onset, the timing of maximum dopamine during reward approach differed between quartiles. *Column 2)* When realigned to magazine arrival, maximum dopamine was aligned to magazine arrival rather than lever press, as evidenced by consistent timing of maximum dopamine across quartiles. *Column 3)* Speed to approach the reward magazine after the last lever press. Kruskal-Wallis tests were statistically significant for Self-initiated Go/No-go and Cue-initiated Go/No-go/Free ($p < 0.001$; see supplementary statistics table). Comparisons of magazine approach speed within each task variant showed no consistent relationship between quartiles, ruling out post-action-cue 'vigor' as a main driver of reward-approach dopamine changes. *Post hoc* Dunn's test significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

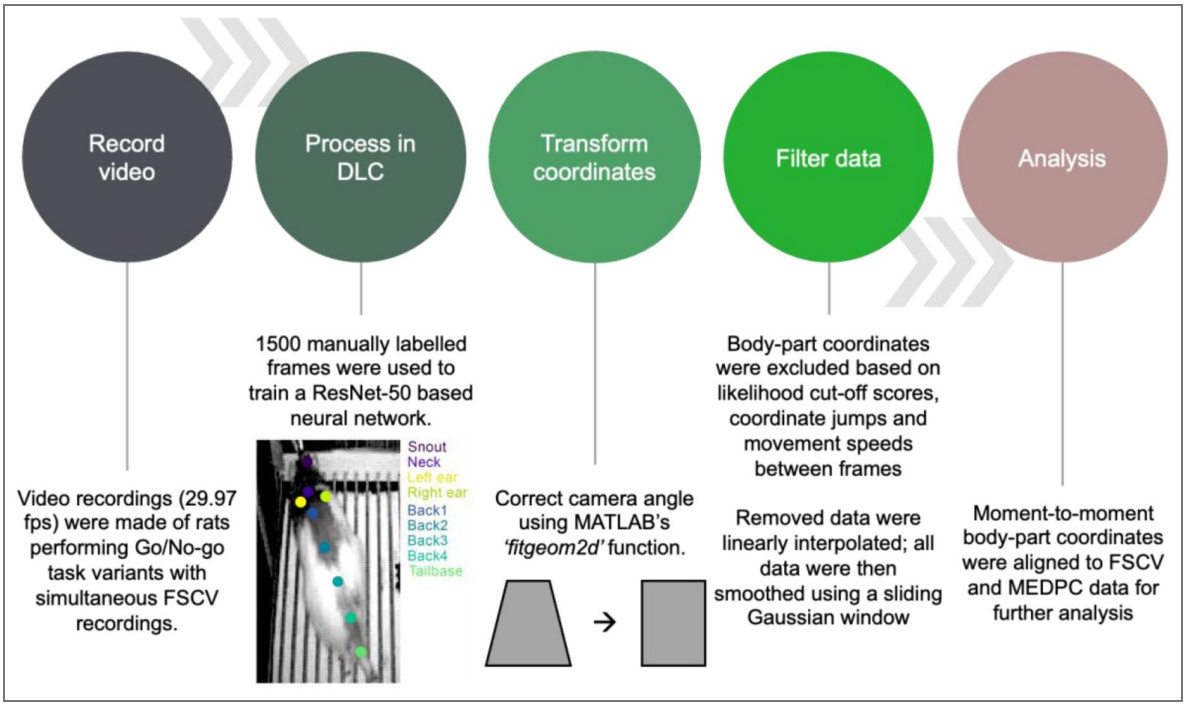
Supplementary Figure S8. **Schematic of all VMS dopamine recording sites.**

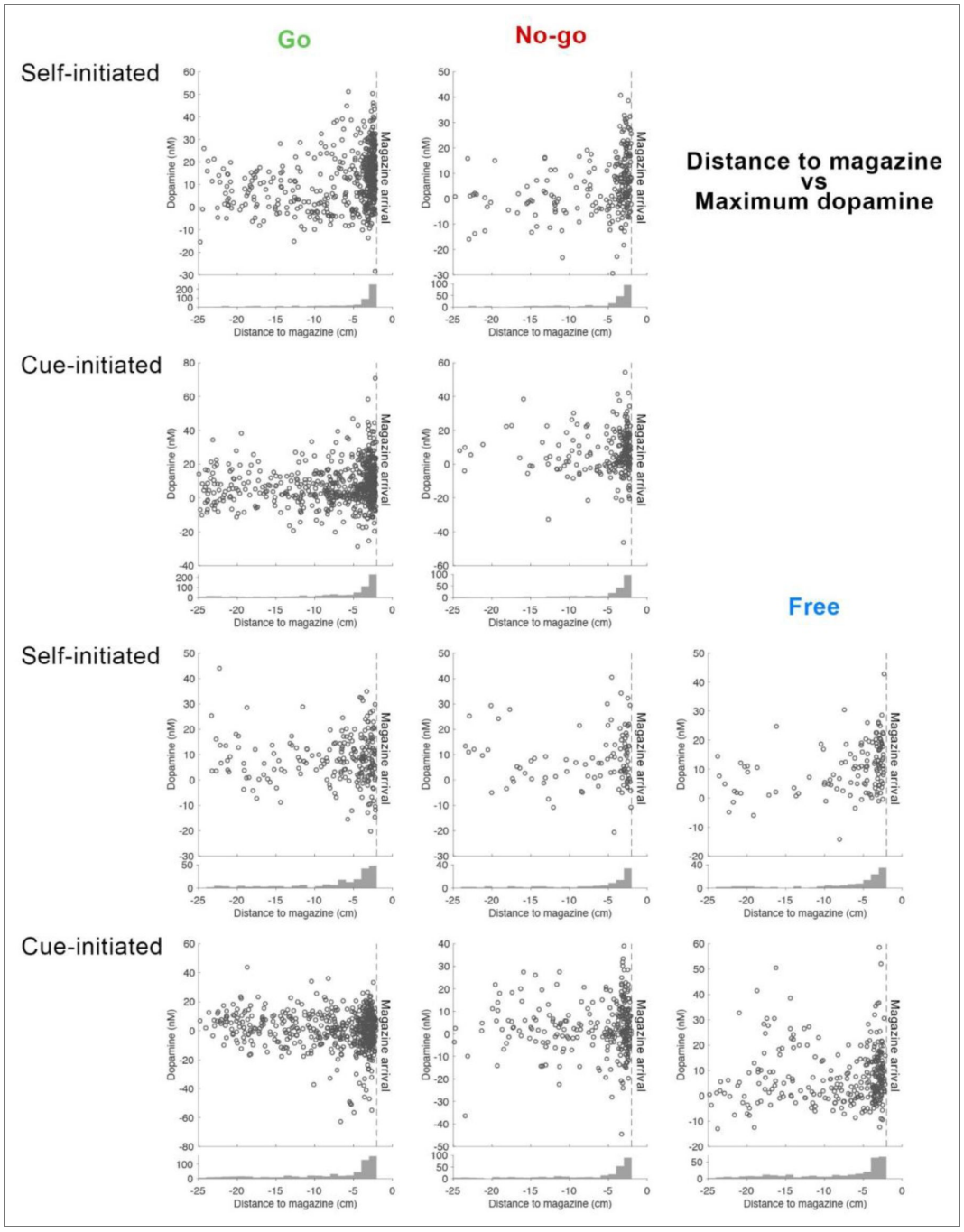
Coronal sections range from 0.48 to 2.28 mm anterior to Bregma. (Paxinos & Watson 1998).



Supplementary Figure S9. **Video analysis workflow.**

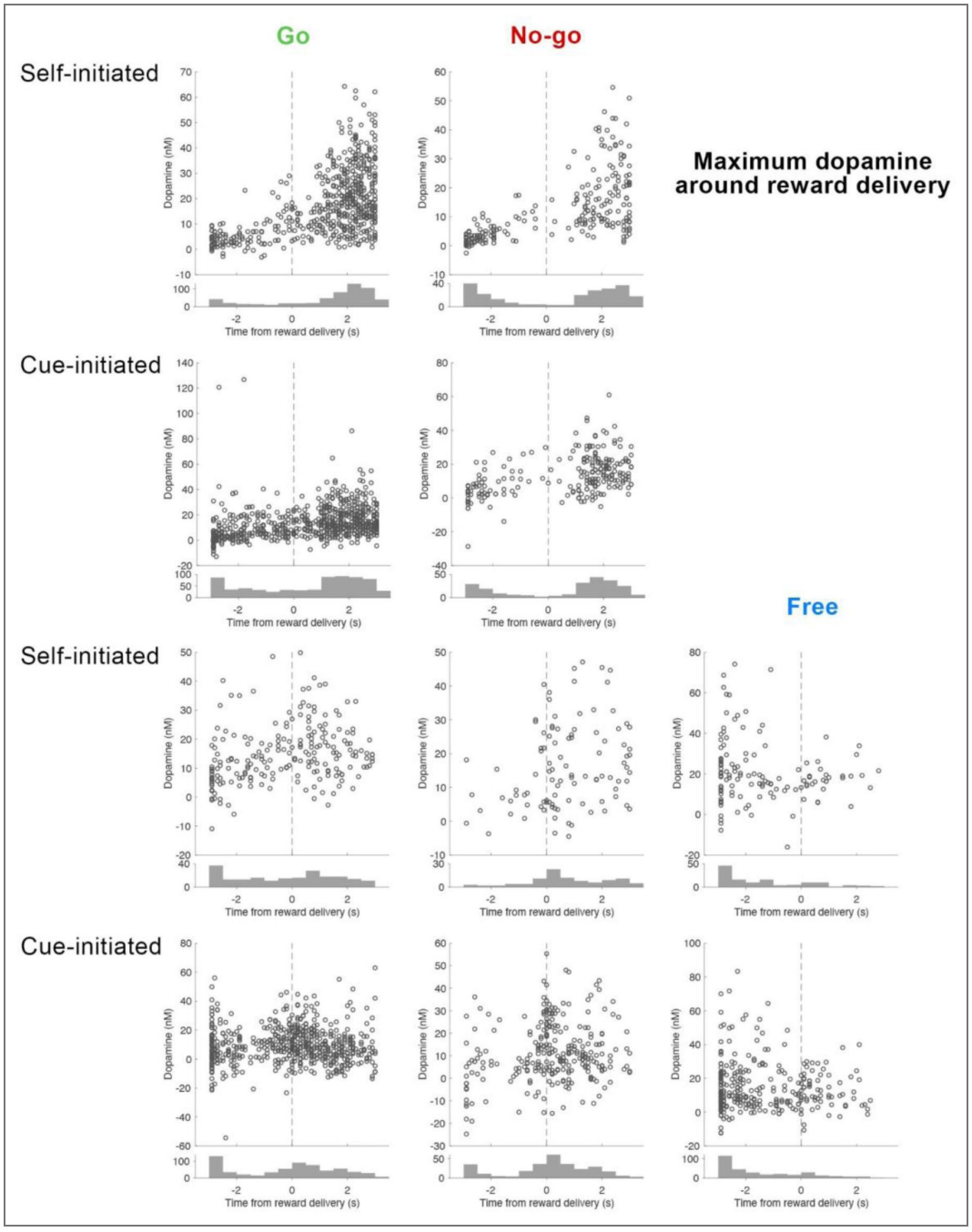
Details for each step can be found in *Methods Section 5: Video recordings and analysis - DeepLabCut pose estimation*. Analysis details can be found under *DeepLabCut behavior analysis*.





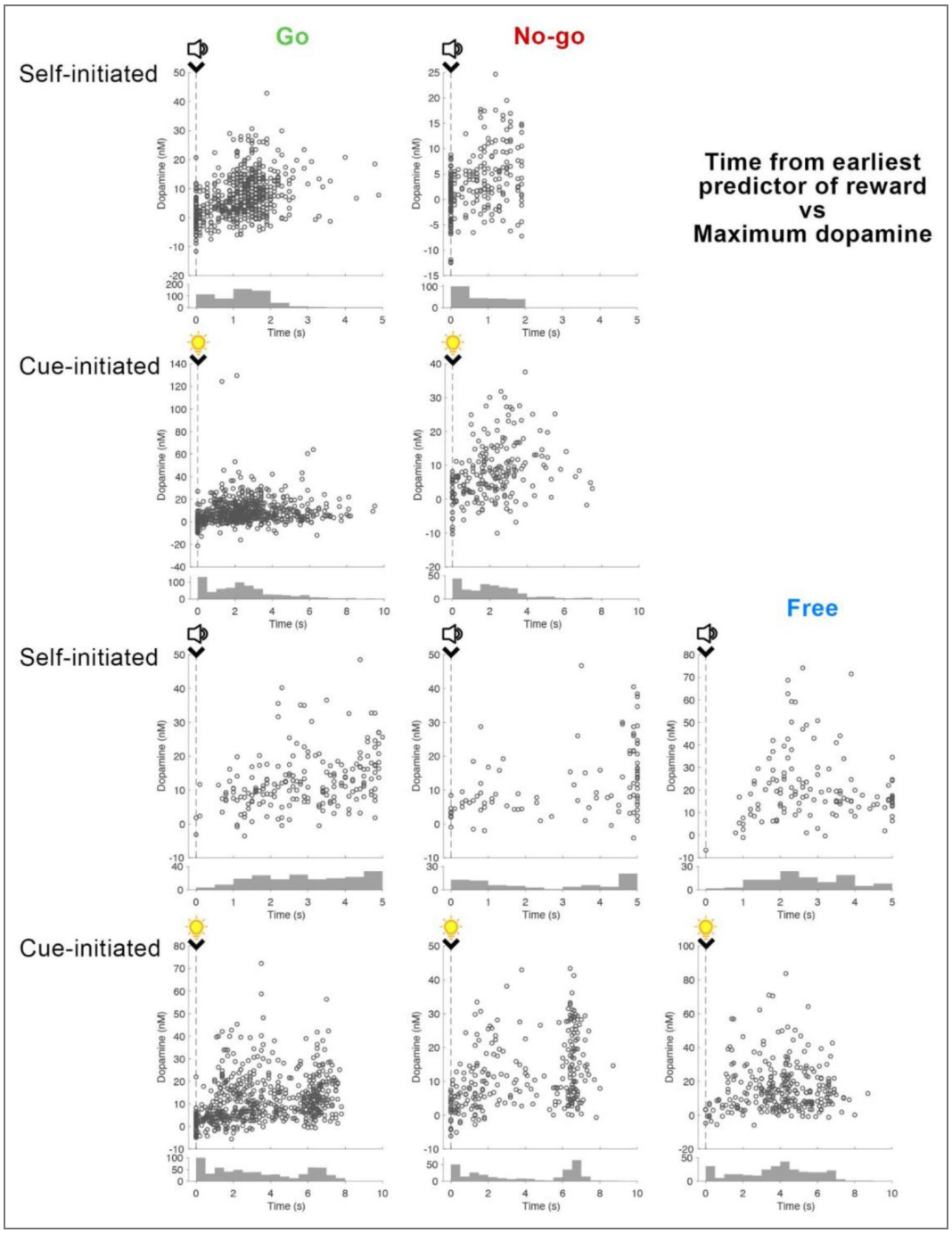
Supplementary Figure S10. Maximum dopamine concentration per trial as a function of distance to the magazine.

The x-axis represents neck distance from the magazine, progressing from furthest to closest (−2cm). Magazine arrival was defined as the moment neck distance reached <2 cm from the magazine-panel (dotted vertical line). The y-axis shows maximum dopamine concentration recorded within each trial. Individual dots represent the maximum dopamine concentration per trial. Bottom insert (histogram) shows the distribution of trial counts across distance bins. Note the accumulation of max dopamine values as animals arrive at the magazine (dotted line).



Supplementary Figure S11. Maximum dopamine concentration around reward delivery (±3s).

The x-axis represents time (s), centered on reward delivery (time = 0 s). The y-axis shows maximum dopamine concentration recorded within each trial. Individual dots represent the maximum dopamine concentration per trial. The bottom insert (histogram) shows the distribution of trial counts across time bins. Note the accumulation of peak dopamine values prior to reward delivery in Free trials but not the other trial types.



Supplementary Figure S12. Maximum dopamine concentration from the earliest predictor of reward to reward delivery.

The x-axis represents time (s), with time = 0 s depicting the earliest predictor of reward in each task variant. The y-axis shows maximum dopamine concentration recorded within each trial. Individual dots represent the maximum dopamine value per trial. The bottom insert (histogram) shows the distribution of trial counts across time bins. For *Self-initiated* task variants, time = 0 s corresponds to the onset of the action cue (speaker symbol). For *Cue-initiated* task variants, time = 0 s corresponds to illumination of the nose-poke light (lightbulb symbol). Note that for *Cue-initiated* Go/No-go/Free, 2 Go trials and 3 Free trials fell outside the x-axis range (0–10 s); for Self-initiated Go/No-go/Free, 7 Go trials and 10 Free trials fell outside this range.

Data availability

Behavioral and voltammetry data have been deposited at *https://osf.io/wcf2d/* and are publicly available. The code for the analyses presented in this paper is openly accessible at *https://osf.io/wcf2d/*.

Acknowledgements

We would like to thank members of the Willuhn laboratory for productive discussions. We thank R. Hamelink and A.S.C. França, for providing technical assistance; M.R. Kandroodi for advice on the hierarchical analysis; L. Fellinger, F. Veen, and T. Arbab for input on the manuscript; H.E.M. den Ouden for constructive feedback. This research was supported by a VICI grant from the Netherlands Organization for Scientific Research VI.C.222.052 awarded to I.W. Funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101066959 was awarded to E.Z.P.

Additional information

Author contributions

EZP and IW designed the experiments; EZP, NB, GH, LM, and PG performed experiments; NB, and GZ, processed video data, EZP preprocessed and analyzed all data; IW, LM, GH, and PG assisted with data analysis; EZP and IW wrote and edited the manuscript.

Funding

Funder	Grant reference number	Author
Netherlands Organisation for Scientific Research	VI.C.222.052	Ingo Willuhn
EC | Horizon Europe | Excellent Science | HORIZON EUROPE Marie Sklodowska-Curie Actions (MSCA)	https://doi.org/10.3030/101066959	Eugenia Z Poh

Author ORCID iDs

Eugenia Z Poh: https://orcid.org/0000-0001-5338-4982
Ingo Willuhn: https://orcid.org/0000-0001-6540-6894

References

Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**:129-141 https://doi.org/10.1016/j.neuron.2005.05.020 | PubMed

Beckmann JS, Chow JJ (2015) Isolating the incentive salience of reward-associated stimuli: value, choice, and persistence. *Learning & memory* **22**:116-127 https://doi.org/10.1101/lm.037382.114 | PubMed

Bernklau TW, Righetti B, Mehrke LS, Jacob SN (2024) Striatal dopamine signals reflect perceived cue–action–outcome associations in mice. *Nature Neuroscience* **27**:747-757 https://doi.org/10.1038/s41593-023-01567-2 | PubMed

Christie J (1996) Spatial contiguity facilitates Pavlovian conditioning. *Psychonomic Bulletin & Review* **3**:357-359 https://doi.org/10.3758/bf03210760 | PubMed

Clark JJ, Sandberg SG, Wanat MJ, Gan JO, Horne EA, Hart AS, Akers CA, Parker JG, Willuhn I, Martinez V, *et al.* (2010) Chronic microsensors for longitudinal, subsecond dopamine detection in behaving animals. *Nature Methods* **7**:126-129 https://doi.org/10.1038/nmeth.1412 | PubMed

Coddington LT, Dudman JT (2018) The timing of action determines reward prediction signals in identified midbrain dopamine neurons. *Nature Neuroscience* **21**:1563-1573 https://doi.org/10.1038/s41593-018-0245-7 | PubMed

Collins AL, Greenfield VY, Bye JK, Linker KE, Wang AS, Wassum KM (2016) Dynamic mesolimbic dopamine signaling during action sequence learning and expectation violation. *Scientific Reports* **6**:20231 https://doi.org/10.1038/srep20231 | PubMed

Cousins MS, Atherton A, Turner L, Salamone JD (1996) Nucleus accumbens dopamine depletions alter relative response allocation in a T-maze cost/benefit task. *Behavioural Brain Research* **74**:189-197 https://doi.org/10.1016/0166-4328(95)00151-4 | PubMed

Day JJ, Roitman MF, Wightman RM, Carelli RM (2007) Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nature neuroscience* **10**:1020-1028 https://doi.org/10.1038/nn1923 | PubMed

du Hoffmann J, Nicola SM (2014) Dopamine Invigorates Reward Seeking by Promoting Cue-Evoked Excitation in the Nucleus Accumbens. *The Journal of Neuroscience* **34**:14349 https://doi.org/10.1523/JNEUROSCI.3492-14.2014 | PubMed

Engelhard B, Finkelstein J, Cox J, Fleming W, Jang HJ, Ornelas S, Koay SA, Thiberge SY, Daw ND, Tank DW, *et al.* (2019) Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature* **570**:509-513 https://doi.org/10.1038/s41586-019-1261-9 | PubMed

Farrell K, Lak A, Saleem AB (2022) Midbrain dopamine neurons signal phasic and ramping reward prediction error during goal-directed navigation. *Cell Reports* **41** https://doi.org/10.1016/j.celrep.2022.111470 | PubMed

Flagel SB, Akil H, Robinson TE (2009) Individual differences in the attribution of incentive salience to reward-related cues: Implications for addiction. *Neuropharmacology* **56**:139-148 https://doi.org/10.1016/j.neuropharm.2008.06.027 | PubMed

Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, Akers CA, Clinton SM, Phillips PE, Akil H (2011) A selective role for dopamine in stimulus–reward learning. *Nature* **469**:53-57 https://doi.org/10.1038/nature09588 | PubMed

Gan JO, Walton ME, Phillips PEM (2010) Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. *Nature Neuroscience* **13**:25-27 https://doi.org/10.1038/nn.2460 | PubMed

Gershman SJ, Assad JA, Datta SR, Linderman SW, Sabatini BL, Uchida N, Wilbrecht L (2024) Explaining dopamine through prediction errors and beyond. *Nature neuroscience* **27**:1645-1655 https://doi.org/10.1038/s41593-024-01705-4 | PubMed

Gershman SJ, Uchida N (2019) Believing in dopamine. *Nature Reviews Neuroscience* **20**:703-714 https://doi.org/10.1038/s41583-019-0220-7 | PubMed

Goedhoop J, Arbab T, Willuhn I (2023) Anticipation of Appetitive Operant Action Induces Sustained Dopamine Release in the Nucleus Accumbens. *The Journal of Neuroscience* **43**:3922 https://doi.org/10.1523/JNEUROSCI.1527-22.2023 | PubMed

Guru A, Seo C, Post RJ, Kullakanda DS, Schaffer JA, Warden MR (2020) Ramping activity in midbrain dopamine neurons signifies the use of a cognitive map. *bioRxiv* https://doi.org/10.1101/2020.05.21.108886

Hamid AA, Frank MJ, Moore CI (2021) Wave-like dopamine dynamics as a mechanism for spatiotemporal credit assignment. *Cell* **184**:2733-2749.e16. https://doi.org/10.1016/j.cell.2021.03.046 | PubMed

Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, Kennedy RT, Aragona BJ, Berke JD (2016) Mesolimbic dopamine signals the value of work. *Nature Neuroscience* **19**:117-126 https://doi.org/10.1038/nn.4173 | PubMed

Hearst E, Jenkins HM (1974) *Sign-tracking: The stimulus-reinforcer relation and directed action* Psychonomic Society.

Howe MW, Dombeck DA (2016) Rapid signalling in distinct dopaminergic axons during locomotion and reward. *Nature* **535**:505-510 https://doi.org/10.1038/nature18942 | PubMed

Howe MW, Tierney PL, Sandberg SG, Phillips PEM, Graybiel AM (2013) Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature* **500**:575-579 https://doi.org/10.1038/nature12475 | PubMed

Ikemoto S, Panksepp J (1999) The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Research Reviews* **31**:6-41 https://doi.org/10.1016/s0165-0173(99)00023-5 | PubMed

Jean-Richard-dit-Bressel P, Clifford CW, McNally GP (2020) Analyzing event-related transients: confidence intervals, permutation tests, and consecutive thresholds. *Frontiers in molecular neuroscience* **13**:14 https://doi.org/10.3389/fnmol.2020.00014 | PubMed

Jean-Richard-dit-Bressel P, Tran J, Didachos A, McNally GP (2022) Instrumental aversion coding in the basolateral amygdala and its reversion by a benzodiazepine. *Neuropsychopharmacology* **47**:1199-1209 https://doi.org/10.1038/s41386-021-01176-2 | PubMed

Jin X, Costa RM (2010) Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature* **466**:457-462 https://doi.org/10.1038/nature09263 | PubMed

Kim HR, Malik AN, Mikhael JG, Bech P, Tsutsui-Kimura I, Sun F, Zhang Y, Li Y, Watabe-Uchida M, Gershman SJ, *et al.* (2020) A Unified Framework for Dopamine Signals across Timescales. *Cell* **183**:1600-1616.e25. https://doi.org/10.1016/j.cell.2020.11.013 | PubMed

Krausz TA, Comrie AE, Kahn AE, Frank LM, Daw ND, Berke JD (2023) Dual credit assignment processes underlie dopamine signals in a complex spatial environment. *Neuron* **111**:3465-3478.e7. https://doi.org/10.1016/j.neuron.2023.07.017 | PubMed

Kremer Y, Flakowski J, Rohner C, Lüscher C (2020) Context-dependent multiplexing by individual VTA dopamine neurons. *Journal of Neuroscience* **40**:7489-7509 https://doi.org/10.1523/jneurosci.0502-20.2020 | PubMed

Lerner TN, Holloway AL, Seiler JL (2021) Dopamine, Updated: Reward Prediction Error and Beyond. *Current Opinion in Neurobiology* **67**:123-130 https://doi.org/10.1016/j.conb.2020.10.012

London TD, Licholai JA, Szczot I, Ali MA, LeBlanc KH, Fobbs WC, Kravitz AV (2018) Coordinated Ramping of Dorsal Striatal Pathways preceding Food Approach and Consumption. *The Journal of Neuroscience* **38**:3547 https://doi.org/10.1523/JNEUROSCI.2693-17.2018 | PubMed

Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, Bethge M (2018) DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience* **21**:1281-1289 https://doi.org/10.1038/s41593-018-0209-y | PubMed

Meyer PJ, Cogan ES, Robinson TE (2014) The form of a conditioned stimulus can influence the degree to which it acquires incentive motivational properties. *PloS one* **9**:e98163 https://doi.org/10.1371/journal.pone.0098163 | PubMed

Mikhael JG, Kim HR, Uchida N, Gershman SJ (2022) The role of state uncertainty in the dynamics of dopamine. *Current Biology* **32**:1077-1087.e9. https://doi.org/10.1016/j.cub.2022.01.025 | PubMed

Mogenson GJ, Jones DL, Yim CY (1980) From motivation to action: Functional interface between the limbic system and the motor system. *Progress in Neurobiology* **14**:69-97 https://doi.org/10.1016/0301-0082(80)90018-0 | PubMed

Mohebi A, Pettibone JR, Hamid AA, Wong J-MT, Vinson LT, Patriarchi T, Tian L, Kennedy RT, Berke JD (2019) Dissociable dopamine dynamics for learning and motivation. *Nature* **570**:65-70 https://doi.org/10.1038/s41586-019-1235-y | PubMed

Montague P, Dayan P, Sejnowski T (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Neuroscience* **16**:1936 https://doi.org/10.1523/JNEUROSCI.16-05-01936.1996 | PubMed

Nicola SM (2010) The flexible approach hypothesis: unification of effort and cue-responding hypotheses for the role of nucleus accumbens dopamine in the activation of reward-seeking behavior. *Journal of Neuroscience* **30**:16585-16600 https://doi.org/10.1523/jneurosci.3958-10.2010 | PubMed

Nicola SM (2007) The nucleus accumbens as part of a basal ganglia action selection circuit. *Psychopharmacology* **191**:521-550 https://doi.org/10.1007/s00213-006-0510-4 | PubMed

Niv Y, Schoenbaum G (2008) Dialogues on prediction errors. *Trends in cognitive sciences* **12**:265-272 https://doi.org/10.1016/j.tics.2008.03.006 | PubMed

Paxinos G, Watson C (2006) *The Rat Brain in Stereotaxic Coordinates: Hard Cover Edition* Elsevier.

Phillips PE, Stuber GD, Heien ML, Wightman RM, Carelli RM (2003) Subsecond dopamine release promotes cocaine seeking. *Nature* **422**:614-618 https://doi.org/10.1038/nature01476 | PubMed

Robinson TE, Berridge KC (1993) The neural basis of drug craving: An incentive-sensitization theory of addiction. *Brain Research Reviews* **18**:247-291 https://doi.org/10.1016/0165-0173(93)90013-P | PubMed

Robinson TE, Yager LM, Cogan ES, Saunders BT (2014) On the motivational properties of reward cues: Individual differences. *Nida* **76**:450-459 https://doi.org/10.1016/j.neuropharm.2013.05.040 | PubMed

Roitman MF, Stuber GD, Phillips PE, Wightman RM, Carelli RM (2004) Dopamine operates as a subsecond modulator of food seeking. *Journal of Neuroscience* **24**:1265-1271 https://doi.org/10.1523/jneurosci.3823-03.2004 | PubMed

Romo R, Schultz W (1990) Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. *Journal of neurophysiology* **63**:592-606 https://doi.org/10.1152/jn.1990.63.3.592 | PubMed

Salamone JD, Correa M (2024) The Neurobiology of Activational Aspects of Motivation: Exertion of Effort, Effort-Based Decision Making, and the Role of Dopamine. *Annual Review of Psychology* https://doi.org/10.1146/annurev-psych-020223-012208 | PubMed

Salamone JD, Correa M (2012) The Mysterious Motivational Functions of Mesolimbic Dopamine. *Neuron* **76**:470-485 https://doi.org/10.1016/j.neuron.2012.10.021 | PubMed

Salamone JD, Correa M (2002) Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behavioural brain research* **137**:3-25 https://doi.org/10.1016/s0166-4328(02)00282-6 | PubMed

Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* **275**:1593-1599 https://doi.org/10.1126/science.275.5306.1593 | PubMed

Starkweather CK, Uchida N (2021) Dopamine signals as temporal difference errors: recent advances. *Current Opinion in Neurobiology* **67**:95-105 https://doi.org/10.1016/j.conb.2020.08.014 | PubMed

Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, Janak PH (2013) A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience* **16**:966-973 https://doi.org/10.1038/nn.3413 | PubMed

Syed ECJ, Grima LL, Magill PJ, Bogacz R, Brown P, Walton ME (2016) Action initiation shapes mesolimbic dopamine encoding of future rewards. *Nature Neuroscience* **19**:34-36 https://doi.org/10.1038/nn.4187 | PubMed

van Elzelingen W, Goedhoop J, Warnaar P, Denys D, Arbab T, Willuhn I (2022) A unidirectional but not uniform striatal landscape of dopamine signaling for motivational stimuli. *Proceedings of the National Academy of Sciences* **119**:e2117270119 https://doi.org/10.1073/pnas.2117270119 | PubMed

Walton ME, Bouret S (2019) What is the relationship between dopamine and effort?. *Trends in neurosciences* **42**:79-91 https://doi.org/10.1016/j.tins.2018.10.001 | PubMed

Wassum KM, Ostlund SB, Maidment NT (2012) Phasic Mesolimbic Dopamine Signaling Precedes and Predicts Performance of a Self-Initiated Action Sequence Task. *Biological Psychiatry* **71**:846-854 https://doi.org/10.1016/j.biopsych.2011.12.019 | PubMed

Willuhn I, Burgeno LM, Everitt BJ, Phillips PEM (2012) Hierarchical recruitment of phasic dopamine signaling in the striatum during the progression of cocaine use. *Proceedings of the National Academy of Sciences* **109**:20703-20708 https://doi.org/10.1073/pnas.1213460109 | PubMed

Peer reviews

Reviewer #1 (Public review):

Summary:

Poh and colleagues investigate dopamine signaling in the nucleus accumbens (ventromedial striatum) in rats engaged in several forms of go/no-go tasks, that differed in reward controllability (self-initiated reward seeking or cue-evoked/quasi-pavlovian), and in the specific timing of the action-reward contingencies. They analysis dopamine recordings made with fast scan cyclic voltammetry and find that dopamine signals vary most consistently to cues that signal a required action (go cues) vs cue signaling action withholding (no go cues). Through various analysis they report that dopamine signals align most clearly with action initiation and with the approach to the reward-delivery location. Collectively these data support aspects of a variety of frameworks related to accumbens dopamine signaling in movement, action vigor, approach, etc.

Strengths:

These studies use several task variants that consolidate a few different components of dopamine signal functions and allow for a broad comparison of many psychological and behavioral aspects. The behavioral analysis is detailed. These results touch on many previous findings, larger showing consistent results with past studies.

Weaknesses:

The paper is dense and could benefit from some revision to increase clarity of the figures, the methods and analysis. The inclusion of many tasks is a strength but also somewhat overshadows specific points in the data, which could be improved with some revision to focus. There is a lack of strong connection between some of the findings, which if revised would help to emphasize the impact of the work.

https://doi.org/10.7554/eLife.109868.2.sa2

Reviewer #2 (Public review):

Here, the authors record dopamine release using fast-scan cyclic voltammetry in the nucleus accumbens/ ventromedial striatum (VMS) while rats perform variants of a go/no-go task. Two versions are self-paced, in that the rat can initiate a trial by nosepoking at the odor port at any time once the ITI had elapsed, whereas the other two require the rat to wait for a cue-light before responding. Two "long" variants also require either more lever-presses on go trials, or a longer nosepoke time for no-go trials, and also incorporate "free" trials in which the rat is rewarded for just heading straight to the food tray. The authors find that dopamine levels increase more during the response requirement for go than no-go trials, indicating a role for invigorating to-be-rewarded actions. Dopamine levels also steadily increased as rats approached the site of reward delivery, and the authors demonstrate quite elegantly that this was not due to orientation to the food tray, or time-to-reward, or action initiation, but instead reflects spatial proximity to the rewarded location. Contrary to previous reports, the authors did not discern any differences in dopamine dynamics depending on whether the trials were cue- or self-paced, and dopamine release did not scale with effort requirements.

The manuscript is well-written and the authors use figures to great effect to explain what could otherwise be a hard-to-parse set of data. The authors make good use of the richness of their behavioral data to justify or negate potential conclusions.

https://doi.org/10.7554/eLife.109868.2.sa1

Author response:

The following is the authors' response to the original reviews.

> ***Public Reviews:***
>
> ***Reviewer #1 (Public review):***
>
> *Summary:*
>
> *Poh and colleagues investigate dopamine signaling in the nucleus accumbens (ventromedial striatum) in rats engaged in several forms of Go/No Go tasks, which differed in reward controllability (self-initiated reward seeking or cue-evoked/quasi-pavlovian), and in the specific timing of the action-reward contingencies. They analyze dopamine recordings made with fast scan cyclic voltammetry, and find that dopamine signals vary most consistently to cues that signal a required action (Go cues) vs cues signaling action withholding (No Go cues). Through various analyses, they report that dopamine signals align most clearly with action initiation and with the approach to the reward-delivery location. Collectively, these data support aspects of a variety of frameworks related to accumbens dopamine signaling in movement, action vigor, approach, etc.*
>
> *Strengths:*
>
> *These studies use several task variants that consolidate a few different components of dopamine signal functions and allow for a broad comparison of many psychological and behavioral aspects. The behavioral analysis is detailed. These results touch on many previous findings, largely showing consistent results with past studies.*
>
> *Weaknesses:*
>
> *The paper could heavily benefit from some revision to 1) increase clarity of the figures, the methods, and the analysis. 2) The inclusion of many tasks is a strength, but also somewhat overshadows specific points in the data, which could be improved with some revision/reworking. 3) Some conclusions are not fully justified. As shown, support for the conclusion "dopamine reflects action initiation but not controllability or effort" is lacking without more analyses and additional context. 4) Further, the notion that the dopamine signals reported here reflect spatial information could be justified more strongly.*

We thank the reviewer for their detailed evaluation and constructive feedback. We have made substantial revisions to address each concern raised:

(1) Clarity of figures, methods and analyses

We have revised the organization of the panels in Figure 1 for clarity.

We have revised Figure 2 and its caption: we have labelled all comparisons depicted in the figure, and now included a line, "Subject-wise comparisons of dopamine data were made for all alignments", to clearly show that all statistical tests shown in Figure 2c-e were performed between subjects.

In the caption of Figure 3, we have now added, "... and trial-wise statistical Kruskal-Wallis tests were performed for each task-variant." to clearly show that statistical tests were performed on the trial-wise level for Figure 3c.

We have now added a table to the Methods section (Table 1), detailing the sample size in each task variant and the number of trials within each No-go classification.

We have added more information in the Methods section: we now include Videos to show the classified No-go behaviors and other trial types (Go and Free); and provide a schematic of the DLC workflow in the supplementary materials (Supplementary Figure S9).

(2) Strengthening specific points in the data

To improve clarity of our main findings, we have revised the layout of the Results section such as including more descriptive headers:

"Behavioral performance in Go/No-go ("short task") was unaffected by controllability of reward pursuit"

"Behavioral performance in Go/No-go/Free ("long task") was unaffected by controllability of reward pursuit"

"Motivation to approach the reward magazine was similar between Go and No-go trials"

"VMS dopamine release encodes reward-related action initiation"

"VMS dopamine release does not only reflect reward-related action initiation"

"Maximum VMS dopamine release encodes spatial but not temporal proximity to rewards"

"Motivational state reflected by No-go behavioral strategy correlates with dopamine signal size during reward approach"

(3) (4) Conclusions drawn from data

We have carefully revised our conclusions to accurately reflect our experimental design and analyses, emphasising our core finding that VMS dopamine was consistently increased in Go versus No-go trials, throughout manipulation of the type of trial start (self- and cue-initiated) and effort manipulation (short and long task variants).

We thank the reviewer for the feedback regarding our VMS dopamine signals reflecting spatial proximity to reward. We performed additional analyses, which we present in Supplementary Figure S10, S11 and S12, to support our interpretation that VMS dopamine encodes spatial proximity to reward.

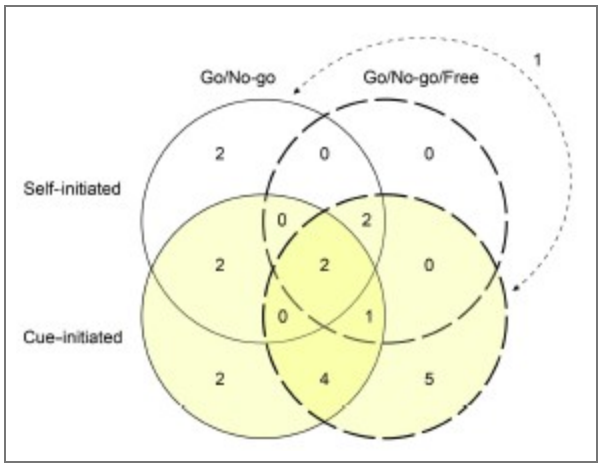
We appreciate the reviewer comment relating to the statement, "Dopamine reflects action initiation but not controllability or effort". We have revised the wording of our conclusion to better reflect our intent, which is to compare the action-selective encoding of dopamine (i.e., action initiation vs action suppression). This subheading is now changed in the Discussion to "Reward-related dopamine depends on action initiation irrespective of controllability and effort". The additional analyses that we have performed are shown in Author response image 1entitled: "Average Go minus No-go dopamine reveals no effect of controllability (self vs. cue-initiated) or effort (short vs. long).

> *Additional details on subjects used in each study, analysis details on trialwise vs subjects-wise data, and other context would be helpful for improving the paper.*

The number of subjects for each task variant was reported in the Methods Section 3: Behavioral procedures in the original submission of this manuscript.

To improve the paper, we now include a table in Methods Section 6: Statistical Analysis (Table 1), detailing the sample size in each task variant (with FSCV recordings) and the number of trials within each No-go classification.

To give more context, we made Author response image 1 to illustrate the number of subjects in each task variant (and their overlap):



Author response image 1. Number of subjects included in each Go/No-go task variant (total *n* = 21). Values (*n*) depict overlap of each subject between task variants. One animal was recorded in self-initiated Go/No-go and Cue-initiated Go/No-go/Free (dotted line with arrowheads).

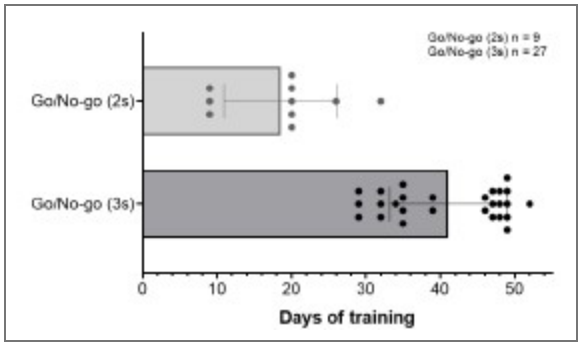
> ***Reviewer #2 (Public review):***
>
> *Here, the authors record dopamine release using fast-scan cyclic voltammetry in the nucleus accumbens/ ventromedial striatum (VMS) while rats perform variants of a Go/No Go task. Two versions are self-paced, in that the rat can initiate a trial by nosepoking at the odor port at any time once the ITI has elapsed, whereas the other two require the rat to wait for a cue-light before responding. Two "long" variants also require either more lever-presses on Go trials, or a longer nosepoke time for No Go trials, and also incorporate "free" trials in which the rat is rewarded for just heading straight to the food tray. The authors find that dopamine levels increase more during the response requirement for Go than No Go trials, indicating a role for invigorating to-be-rewarded actions. Dopamine levels also steadily increased as rats approached the site of reward delivery, and the authors demonstrate quite elegantly that this was not due to orientation to the food tray, or time-to-reward, or action initiation, but instead reflects spatial proximity to the rewarded location. Contrary to previous reports, the authors did not discern any differences in dopamine dynamics depending on whether the trials were cue- or self-paced, and dopamine release did not scale with effort requirements.*
>
> *The manuscript is well-written, and the authors use figures to great effect to explain what could otherwise be a hard-to-parse set of data. The authors make good use of the richness of their behavioral data to justify or negate potential conclusions. I have the following comments.*
>
> *Re: The lack of relationship between effort to acquire reward in the current study and the magnitude of dopamine release, 1) can the authors unpack this a bit more? 2) Why the difference between the Walton and Bouret studies? Were the shifts in effort requirements*

> *comparable across the behavioral tasks? 3) What else could be different between the methodologies?*

We thank the reviewer for the feedback and have responded to each of the three questions below (see points 1-3).

Firstly, we tried to improve the clarity of our research aims. Our primary comparison throughout the manuscript is between Go versus No-go within each task variant. We ask whether the Go/No-go difference in dopamine signaling persists across different response demands. Thus, testing effort was not central to this main question, but rather a feature of the task that did not affect the Go-No-go dopamine difference.

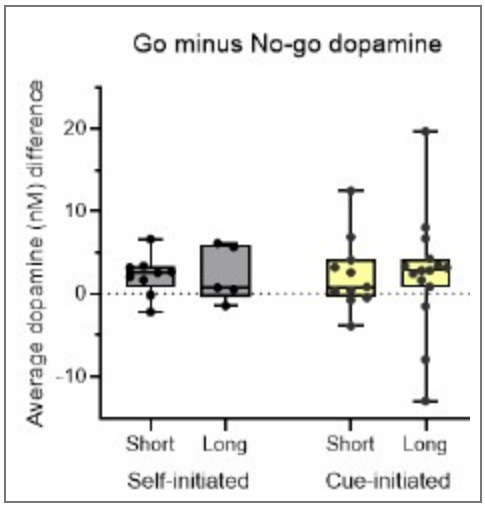
(1) Consistent with this aim, we show that VMS dopamine differs between Go and No-go actions persistently across all task variants despite differences in response requirements (action was always accompanied by greater dopamine release compared to action suppression). Our behavioral-training data suggest that the ability to perform short and long tasks differed: rats were first trained to criterion on either a short (~2 s) or long (~3 s) Go/No-go variant, with the longer variant requiring substantially more training sessions (short: 18.5 ± 7.6 sessions vs long: 41.1 ± 7.9 sessions; see Author response image 2), indicating behavioral demands were higher for the long-task.



Author response image 2.

(2) Regarding the apparent discrepancy with Walton and Bouret (2019), we acknowledge that our original description was imprecise (We wrote: “Previous studies have shown that dopamine signals are influenced by the effort required to obtain rewards”). Our intent was not to suggest a direct contradiction, but rather to emphasize that our findings are consistent with the paper’s broader conclusion that effort encoding by dopamine is limited and highly context-dependent. We have now adjusted the manuscript to better reflect our intent by changing the sentence to, “Previous studies have shown that dopamine signals may be influenced by the effort required to obtain reward but only for particular task conditions (Cousins et al., 1996; Gan et al., 2010; see for reviews, Salamone and Correa, 2024; Walton and Bouret, 2019).”

For added clarity, these were the main results highlighted in the Walton and Bouret review: Gan et al. (2010) demonstrated that VMS dopamine sensitivity to low-effort costs is prominent early in training (≤ 2 training sessions) and diminishes after extended experience (> 9 sessions). Similarly, Hollon et al. (2014) reported that cue-evoked VMS dopamine primarily tracks reward magnitude with minimal modulation by effort. In line with this literature, our rats were highly trained (≥ 9 sessions until the first recording), and exhibited no difference of average Go minus No-go dopamine between short and long task variants within controllability type (see Author response image 3), supporting the idea that extended training exhibits minimal effort-related modulation of VMS dopamine.



Author response image 3. Average Go minus No-go dopamine reveals no effect of controllability (self vs. cue-initiated) or effort (short vs. long). A 2 × 2 Bayesian ANOVA (Cauchy prior: fixed effects $r = 0.5$; random effects $r = 1$) consistently favoured the null model over all alternatives. The main effect of controllability and effort showed moderate evidence of absence (controllability: $BF_{10} = 0.324$; effort: $BF_{10} = 0.309$). The model including both main effects performed more poorly ($BF_{10} = 0.101$), and the full model including a controllability × effort interaction was the least supported of all models examined ($BF_{10} = 0.043$). These results provide moderate evidence in favour of H_0, suggesting that neither controllability, effort, nor their interaction meaningfully predicted average Go minus No-go dopamine responses.

(3) With respect to task comparability and methodological differences, our behavioral paradigm differs in important ways from those highlighted by Walton and Bouret, where effort was often manipulated by training animals to associate cues with different numbers of lever presses within the same session, and typically involved only action initiation. In contrast, our task required both action initiation and action suppression, and changes in response contingencies occurred across separate recording sessions rather than within-session cue-based manipulations. Although these paradigms are not directly comparable, and only had the same dopamine recording technique in common (FSCV), a key takeaway of our results is that regardless of effort differences, VMS dopamine during action initiation is consistently higher than during action suppression.

> *I would argue that the cue- vs self-initiated distinction was pretty minor, given that there was a fixed ITI of 5s. How does this task modification compare to those used previously to show that dopamine release corresponds to behavioral controllability? It would help the reader if the authors could spend more time discussing these disparate findings and looking for points of methodological divergence/commonality.*

We agree that clarifying how our manipulation of controllability compares to prior work improves the manuscript, and we have made the necessary adjustments. However, we would first like to correct an incomplete characterization of the task design.

While the short-task variant used a fixed 5 s inter-trial interval (ITI), the long-task variant employed a variable ITI ranging from 15–25 s. In the long-task variant, the timing of trial onset was less predictable, and we believe this manipulation reduced animals' ability to precisely estimate when reward pursuit could begin. Under these conditions, whether trials were Self-initiated or Cue-initiated had a substantial impact on animals' control over the initiation of reward pursuit. That said, we agree that the Self- versus Cue-initiated distinction overall represents a moderate manipulation of controllability compared to those used in studies that focus on controllability.

A key source of divergence across studies lies in the definition of controllability. We defined controllability as the animals' ability to choose the time point of beginning the reward pursuit, rather than whether an action was required, and have now added the following sentence in the:

- Introduction section: "... controllability of reward seeking, defined as the ability to determine when to initiate reward pursuit (Self- vs Cue-initiated trials)...";

- Results section: "We defined controllability as the rats' ability to choose the time point of reward pursuit. In Cue-initiated trials, the time point at which trials could be started was dictated by a cue light, whereas in Self-initiated trials rats were able to choose intrinsically (control) when to attempt a trial start.";

- Discussion section

Importantly, the action requirements for Go, No-go, and Free trials were identical across these trial-start conditions. We found that the degree to which controllability was manipulated in our task was insufficient to modulate the action-specific VMS dopamine signal (Go vs No-go difference), which remained robust across conditions.

In contrast, controllability has been defined by others as the presence versus absence of an operant action requirement for reward. For example, Goedhoop et al. (2023) directly contrasted operant (lever press required) and Pavlovian (no action required) conditions, removing action execution as a prerequisite for reward. In that context, cues signaling operant control elicited sustained VMS dopamine release, which was interpreted as reflecting anticipation or preparation for executing a learned action. Similarly, Hamid et al. (2021) demonstrated that dopamine "wave" directionality across striatal regions depends on controllability defined by operant versus Pavlovian conditioning.

Taken together, these comparisons (results from the present study and in the literature) suggest that dopamine sensitivity to controllability may depend on how it is manipulated. We have clarified these methodological distinctions in the revised Introduction, Results and Discussion, and emphasized that more extreme manipulations (such as removing action requirements entirely or increasing uncertainty over trial timing) may be necessary to reveal controllability-dependent changes in VMS dopamine signaling. Alternatively, the apparent discrepancies across studies may primarily reflect differences in the underlying definitions of controllability rather than conflicting results.

> ***Reviewer #3 (Public review):***
>
> *Summary:*
>
> *The manuscript by Poh et al. investigated whether dopamine release in the ventral medial striatum integrates information about action selection, controllability of reward pursuit, effort, and reward approach. Rats were implanted with FSCV probes and trained in four Go/No Go task variants:*
>
> *(1) trials were self-initiated and had two trial types (Go vs. No Go) that were auditorily cued,*
>
> *(2) trials were cue-initiated and had two trial types (Go vs. No Go) that were auditorily cued,*
>
> *(3) trials were self-initiated and had three trial types (Go vs. No Go vs. free reward) that were auditorily cued, and effort was increased,*

> *(4) trials were cue-initiated and had three trial types (Go vs. No Go vs. free reward) that were auditorily cued.*
>
> *The authors report that dopamine levels rose during Go trials and slowly rose in No Go trials, but this pattern did not differ across task variants that modified effort and whether trials were cued or initiated. They also report that dopamine levels rose as rats approached the reward location and were greater in rats that bit the noseport while holding during the No Go response.*
>
> *Strengths:*
>
> *(1) Interesting task and variants within the task paradigm that would allow the authors to isolate specific behavioral metrics.*
>
> *(2) The goal of determining precisely what VMS dopamine signals do is highly significant and would be of interest to many researchers.*
>
> *Weaknesses:*
>
> *(1) This Go/No-Go procedure is different from the traditional tasks, and this leads to several problems with interpreting the results:*
>
> *(a) Go/No Go tasks typically require subjects to refrain from doing any action. In this task, a response is still required for the No Go trials (e.g., continue holding the nosepoke). The problem with this modified design is that failure to withhold a response on No Go trials could be because i) rats could not continue holding the response, as holding responses are difficult for rodents, or ii) rats could not suppress the prepotent go response. This makes interpreting the behavior and the dopamine signal in No Go trials very difficult.*

We appreciate the reviewer raising this important methodological consideration. We acknowledge that our Go/No-go task differs from traditional paradigms used in humans and primates (e.g. Raud et al. 2020, 10.1016/j.neuroimage.2020.11658; Eagle, Bari & Robbins 2008, 10.1007/s00213-008-1127-6; Roitman & Loriaux 2013, 10.1152/jn.00350.2013).

However, our design addresses the specific constraints of studying dynamics in freely moving rodents while maintaining the core feature of Go/No-go tasks: requiring suppression of a prepotent response. Our task accomplishes the primary aim of our study, which is to compare VMS dopamine dynamics during action initiation and action suppression, and below we list the reasons why. Therefore, we do not believe that this difference compromises the validity and interpretation of our results.

It has been suggested for decades that the two main processes governed by mesolimbic dopamine are reward learning and motivated action, and our study aimed to better understand how VMS dopamine integrates reward-related information and motivated action, rather than studying them in isolation. To do so, we trained rats in a modified Go/No-go task.

More recent work (Syed et al. 2016; Hamid et al. 2016; Mohebi et al. 2019) demonstrates that VMS dopamine signaling incorporates both action and reward-related information, rather than either of the two alone. Importantly, in freely-moving rodents, examining this relationship requires preventing the approach response that occurs when reward delivery is anticipated (Pavlovian bias, go for rewards). Traditional Go/No-go designs that simply require "doing nothing" would not achieve this control in freely-moving rats, as animals immediately approach the reward magazine as soon as reward is inferred (as seen in our Free trials). Thus, we require a No-go condition, as we and others have defined (Syed et al. 2016), whereby animals have to actively suppress the 'initiation' response. Action initiation is defined at the beginning of the Discussion section: "... at two distinct points after trial start: 1) when rats

began lever pressing (Go), and 2) when rats walked to the reward magazine, either without action requirement (Free) or after successful trial completion (Go and No-go)".

To further strengthen our interpretation that we compare action initiation and suppression, and to facilitate cross-species translation of our results (i.e., rodent to human), we also include Free trials, where reward delivery requires no specific action (which are essentially like "doing nothing" trials in traditional tasks). This addition allowed us to directly compare No-go and Free trials, where animals must actively suppress responding while maintaining task engagement, to a condition where no overt action is required for a reward, respectively. The dramatic difference in VMS dopamine between No-go and Free trials demonstrates that VMS dopamine reflects active action suppression during No-go trials, rather than merely the absence of action requirements. This has now been discussed.

Finally, we only report correct Go, No-go, and Free trials, which differs from that of human go/no-go studies that focus on the failure of appetitive no-go trials (i.e., inhibiting the prepotent response). In the present study, the dopamine signals that we interpret are restricted to successful trials only: action initiation (moving the lever press), action suppression (i.e., suppressing the prepotent Go response while maintaining their position in the nose-poke port), or no action (no lever press, not staying in the port). While this design differs from human Go/No-go paradigms, we believe our study of correctly performed Go, No-go, and Free trials are necessary for isolating action-dependent components of dopamine signaling in freely moving rats (action initiation vs action suppression vs action free).

> *(b) Most Go/No Go tasks bias or overrepresent Go trials so that the Go response is prepotent, and consequently, successful suppression of the Go response is challenging. 1) I didn't see any information in the manuscript about how often each trial type was presented or 2) how the authors ensured that No Go responses (or lack thereof) were reflecting a suppression of the Go response.*

We appreciate the reviewer's attention to this important methodological consideration. The originally submitted version of the manuscript already addressed both concerns raised.

Trial type presentation frequencies

The Methods section describes our trial presentation approach: "On recording days, the trial types were counterbalanced. Within a session, Go left, Go right, and No-go trials were presented with 33% probability each, without replacement. For sessions with Free trials, trials were presented with a 25% chance without replacement."

This design results in overrepresentation of Go trials overall (66% in the short-task; 50% in the long-task), which establishes the prepotent Go response as intended in standard Go/No-go paradigms. To improve clarity, this detail has now been included in the Methods section.

Ensuring No-go responses reflect suppression of Go response

Our paradigm incorporates multiple features that ensure successful No-go performance reflects suppression of the prepotent Go response:

First, the overrepresentation of Go trials (addressed above) establishes response prepotency. Second, during No-go trials, rats must maintain their snout in the nose-poke port for the duration of the action cue, which creates the requirement to suppress the natural tendency to approach rewards (i.e., Pavlovian bias; Jones et al. 2017, 10.1016/j.bbr.2017.05.044; Guitart-Masip et al. 2014, 10.1007/s00213-013-3313-4; Dayan et al. 2006; 10.1016/j.neunet.2006.03.002). In Go trials, such natural bias does not require suppression as the lever can be approached and pressed during the action-cue period. This conflict between the instrumental No-go requirement and the Pavlovian-instrumental bias toward action makes action suppression particularly challenging (consistent with computational accounts of similar paradigms; Lloyd

& Dayan 2023, 10.1371/journal.pcbi.1011569; Jones et al. 2017, Guitart-Masip et al. 2014, Dayan et al. 2006).

Figure 3 provides behavioral evidence of this challenge: animals frequently left the nose-poke port and developed spontaneous motor strategies (such as biting and digging) to stay in the port, suggesting Pavlovian bias interfering with response suppression for rewards. Importantly, all reported No-go data include only correct trials (i.e., those without lever presses), ensuring that the dopamine signal reflects successful response suppression rather than failed Go attempts.

> *(2) The authors observe relatively consistent differences in the DA signal between Go and No Go trials after the action-cue onset. However, the response type was not randomized between trial type, so there is a confound between trial type (Go/No Go) and response (lever/nosepoke). The difference in DA signal may have nothing to do with the cue type, but reflects differences in DA signal elicited by levers vs. nosepokes.*

As stated in the Introduction section and discussed in our rebuttal to point 1a, the focus of our investigation is how VMS dopamine signals differ during action initiation versus action suppression for rewards, as this is a central unanswered question in the dopamine field. More recent work demonstrates that dopamine incorporates not only RPE but also action initiation (Syed et al. 2016; Hamid et al. 2016; Mohebi et al. 2019), and our goal is to further our understanding of action-dependent VMS signals during reward pursuit.

The reviewer suggests that dopamine differences may reflect differences in lever vs. nosepoke rather than cue type (Go vs No-go). We respectfully suggest this concern reflects a misunderstanding by the reviewer of our experimental question. The cue-action relationship is the experimental manipulation itself. It is not possible to study how dopamine encodes instructed action initiation versus suppression without linking specific cues to specific actions. The suggestion to 'randomize' action type across cue types would eliminate the very phenomenon we are investigating: how dopamine signals differ when cues instruct different action requirements.

Our experimental design specifically compares reward pursuit with action requirements (Go trials: lever press; No-go trials: sustained hold) to reward pursuit without action requirements (Free trials: direct magazine approach). This design allows us to isolate how action initiation and action suppression influence reward-related dopamine signaling, which can reveal how the timing of action initiation influences RPE-dopamine. And which is the point of the study: to show how actions influence RPE dopamine signaling.

Supporting this interpretation:

Firstly, trial types were randomly interleaved, and each auditory cue explicitly instructed a specific behavioral response. Our design directly follows established methods demonstrating that VMS dopamine encodes whether actions are initiated or suppressed following action cues (Syed et al. 2016). That study, like ours, intentionally linked cue identity to a specific action requirement to assess how dopamine reflects instructed behavioral control. Thus, the fact that Go and No-go cues map onto different actions is inherent to the question being addressed, not an unintended confound.

Second, as discussed in our response to point 1b, the asymmetry between Go and No-go trials is theoretically essential. Go trials align with Pavlovian approach tendencies (action initiation to reward), while No-go trials create conflict with this bias by requiring action suppression despite the cue being associated with a reward. This Pavlovian-instrumental conflict makes suppression particularly challenging (Lloyd & Dayan 2023, PLoS Comput Biol 10.1371/journal.pcbi.1011569) and allows us to examine the role dopamine in overriding prepotent responses.

Third, the inclusion of Free trials (discussed in point 1a) demonstrates that our findings reflect instructed action control rather than simply motor execution. Free trials require neither lever pressing nor nose poke maintenance, yet show dopamine dynamics distinct from both Go and No-go trials, confirming that dopamine signals encode action requirements beyond motor output per se.

Finally, we demonstrate that VMS dopamine differs in the same trial type (No-go) and can be classified based on different movement patterns (Biting, Digging, Calm). Importantly, the difference in VMS dopamine only appeared after the action was completed, particularly during reward approach (Figure 3). This data argues against the idea that VMS dopamine is particularly tied to the specific operant manipulanda as suggested by the reviewer, but rather, may reflect an internal motivational state for reward.

Together, the aim of the present study is not to redefine Go/No-go paradigms for rodents, but to utilize this task structure to investigate action-dependent dopamine signalling for rewards, which cannot be answered without the cue-action mapping that we have used.

> *(3) Both Go and No Go trials start with the rat having their nose in the noseport. One cue (Go cue) signals the rat to remove their nose from the noseport and make two lever responses in 5 seconds, whereas the other cue (No Go cue) signals the rat to keep their nose in the noseport for an additional 1.7-1.9 s. The authors state that the time between cue onset and reward delivery was kept the same for all trial types, and Figure 1 suggests this is 2 s, so was reward delivered before rats completed the two lever presses? I would imagine reward was only delivered if rats completed the FR requirement, but again, the descriptions in the text and figures are incongruent.*

The reviewer asks whether reward was delivered before rats completed the two lever presses and notes incongruence between text and figures. We respectfully note that these details were stated in the originally submitted version of the manuscript (see below).

Reward delivery timing

The reviewer asks whether reward was delivered before rats completed the two lever presses, which refers to the short-task variant. No - reward was always delivered immediately after the second lever press for all Go trials. This is described in Methods Section 3: Behavioral procedures - Self-initiated task variant. For added clarity, we have now added the term “immediately”: “... food pellet dispensed into the reward-magazine immediately.”

In the Results section, we report that the average latency to complete two lever presses was 1.8s, which closely matches the 1.7-1.9s nose-poke hold maintenance required for No-go trials in the “short” variant. Thus, the time point of reward delivery was matched between Go and No-go trial types.

Representation of reward delivery timing in figure and text

The reviewer's confusion appears to stem from the schematic representation in Figure 1 and the task variant structure. There were two overarching task variants with different trial requirements:

“Short-task” variants: Go trials required two lever presses (completed on average in 1.8s);

No-go trials required 1.7-1.9s nosepoke maintenance

“Long-task” variants: Go trials required a ‘rewarded’ press to occur 2.7-3.2s after cue onset (completed within ~3s); No-go trials required 2.7-3s nosepoke maintenance.

For simplicity in depicting action-cue onset in Figure 2c, we used grey shading with a speaker icon at approximately 0-2s and 0-3s to represent these two variants. This schematic representation was not intended to indicate the precise reward delivery time, which (as stated in the Methods) occurred only upon successful completion of trial requirements in the short-task variant, or 2s after successful completion of trial requirements in the long-task variant. To improve clarity, we have adjusted the legend of Fig. 2 for more clarity, adding "Shaded gray area depicts approximate duration of action-cue onset for "short" and "long" task variants.", and included more information under Results: "In the short-task variants, action-cues switched off after trial completion and a reward was delivered immediately" and "n the long-task variants, action-cues switched off after trial completion, or in the case of Free trials after 3s, and reward was delivered 2s later (Figure 1a)."

> *(4) The manuscript is difficult to understand because key details are not in the main text or are not mentioned at all. I've outlined several points below:*
>
> *(a) The author's description in the manuscript makes it appear as a discrimination task versus a Go/No Go task. I suggest including more details in the main text that clarify what is required at each step in the task. Additionally, providing clarity regarding what task events the voltammetry traces are aligned to would be very useful.*

We respectfully note that the requested details were already present in the originally submitted version of the manuscript (see below). However, we acknowledge that the task design is complex and may benefit from additional clarity in the main text to aid reader comprehension.

Behavioral task

The reviewer suggests our task appears more like a discrimination task than a Go/No-go task. We acknowledge that our paradigm differs from traditional Go/No-go tasks used in humans and primates, as discussed in our responses to points 1a and 2. However, we classify this as a Go/No-go task because it shares the defining feature: requiring action initiation (Go) and suppression (No-go). This classification is consistent with established rodent literature examining action initiation versus suppression (Syed et al. 2016).

Moreover, as discussed in our response to point 1a, we included Free trials specifically to demonstrate that No-go trials require active suppression rather than discrimination alone. The distinct dopamine dynamics across Go, No-go, and Free trials confirm that our task captures action initiation, action suppression, and action-free states, which is an important contrast that we needed to address our research question about action-dependent dopamine signaling.

The key requirements for each trial type and task variant are described at the beginning of the Results section. A full description of each step required in the task was provided in the Methods Section 3: Behavioral procedures, to avoid repetition in the Results. Specifically:

Self-initiated Go/No-go task variant ("short")

Self-initiated Go/No-go/Free task variant ("long")

Cue-initiated Go/No-go and Go/No-go/Free task variant

To improve clarity, we have added a reference in the Results section to the Methods: Behavioral procedures for additional procedural details.

Voltammetry trace alignment

The events to which voltammetry traces are aligned were stated in the legend:

"... when traces were aligned to action-cue onset..."

"... aligned to the time when animals departed the nose-poke port..."

"... we realigned traces to the moment animals arrived at the reward magazine... "

Figure 2c legend: "c) Traces aligned to action-cue onset"

Figure 2d-e legend: "d) Traces aligned to nose-poke exit and e) magazine arrival...."

However, to improve clarity, we have now added "aligned to action-cue onset.. " to make the trace alignment more immediately apparent when results are first presented, and added: "Dopamine data were aligned to events of interest: action-cue onset, nose-poke exit, and magazine arrival."

> *(b) How many subjects were included in each task variant? The text makes it seem like all rats complete each task variant, but the behavioral data suggest otherwise. Moreover, it appears that some rats did more than one version. Was the order counterbalanced? If not, might this influence the DA signal?*

The number of subjects for each task variant was reported in the Methods Section 3: Behavioral procedures, where each task variant description includes the corresponding sample size ("Self-initiated Go/No-go task ('short'; n =9)", "Self-initiated Go/No-go/Free task ("long"; n = 5)", "A total of n = 11 and n = 15 were included in the Cue-initiated Go/No-go and Cue-initiated Go/No-go/Free tasks, respectively."

For added clarity, we have also added a table for the separation of No-go trials and the number of subjects that it has come from in the Methods.

Task variant completion

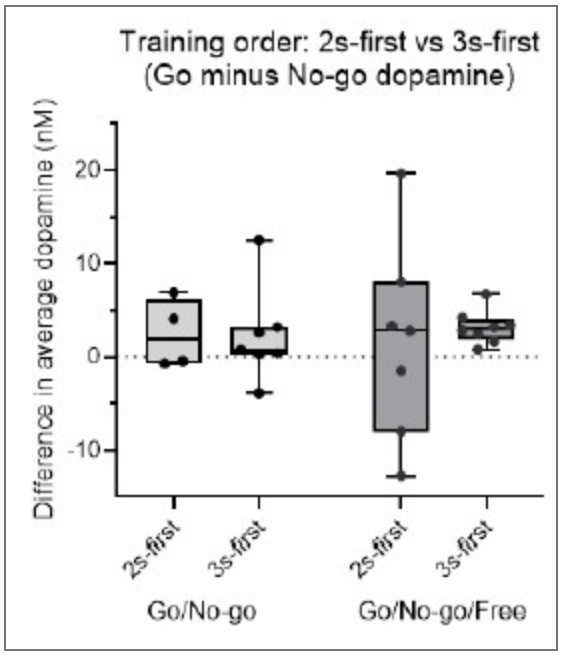
Not all animals completed all task variants. As stated in Methods Section 4: Real-time dopamine recordings and analysis, animals had to achieve >60% success rate for each trial type on at least two consecutive training sessions to proceed to recording. Other reasons include electrode degradation before all recordings could be completed (See Author response image 1).

Training order

We trained four cohorts of animals. One cohort was trained first in the short-task variant

(Cue-initiated Go/No-go), and the remaining three cohorts were trained first in Cue-initiated Go/No-go (3s) long-task variant (i.e., without Free trials, behavioral and FSCV data not presented in the manuscript). Training order was not fully counterbalanced due to constraints described above.

Following additional analyses, our data suggest that training order did not influence our core comparison of Go minus No-go dopamine. To directly address whether training order influenced dopamine signals, we separated animals based on whether they were first trained in the Cue-initiated Go/No-go (2s) or the Cue-initiated Go/No-go/Free (3s). We calculated the average dopamine of each rat during the action-cue period and then calculated the difference between them (Author response image 4). We observed absence of evidence of a difference between the groups.



Author response image 4. Average Go minus No-go dopamine reveals no effect of the initial training variant (2s-first vs 3s-first) or test task version (Go/No-go vs. Go/No-go/Free). A 2 × 2 Bayesian ANOVA (Cauchy prior: fixed effects r = 0.5; random effects r = 1) consistently favoured the null model over all alternatives. The main effect of the initial training variant and test task version both showed moderate evidence of absence (initial training variant: BF_{10} = 0.378; test task version: BF_{10} = 0.367). The model including both main effects performed more poorly (BF_{10} = 0.134), and the full model including an initial training variant × test task version interaction was the least supported of all models examined (BF_{10} = 0.069). These results provide moderate evidence in favour of H_0, suggesting that neither the variant animals were first trained on, the task version administered at test, nor their interaction meaningfully predicted average Go minus No-go dopamine responses.

> *(5) There is a major challenge in their design and interpretation of the dopamine signal. Both trial types (Go and No Go) start with the rat having their nose in the noseport. An auditory cue is presented for 2-3 s signaling to the rat to either leave the noseport and make a lever response (Go trial) or to stay in the noseport (No Go trial). The timing of these actions and/or decisions is entirely independent, so it is not clear to me how the authors would ever align these traces to the exact decision point for each trial type. They attempt to do this with the nose-port exit analysis, but exiting the noseport for a Go trial (a rat needs to make 2 lever presses and then get a reward) versus a No Go trial (a rat needs to go retrieve the reward) is very different and not comparable.*

We respectfully disagree with the reviewer's assertion that our data alignment approach is problematic. Aligning neural activity to specific behavioral epochs that occur at different times and across conditions is a widely used method to investigate the relationship between neural activity and behavior. Just to mention some examples: data collected with fiber photometry (e.g., Tan et al. 2026, doi: 10.1038/s41386-026-02368-4; Hart et al. 2024, doi: 10.1016/j.celrep.2024.113828) and voltammetry (Hamid et al. 2016, doi: 10.1038/nn.4173; Syed et al. 2016, doi:10.1038/nn.4187).

Alignment method

We intentionally designed the task so that overall action timing is matched between trial types (as described in our response to point 3), while specific behavioral epochs occur at different times. This allows us to compare dopamine dynamics during comparable behavioral events across Go and No-go trials (e.g., nose-poke exit).

We align data to three critical behavioral epochs, stated in the Methods Section 4: Real-time dopamine recordings and analysis - FSCV measurement and analysis: action-cue onset, nose-poke exit, and magazine arrival. Each alignment addresses a specific aspect of our research question:

Action-cue onset alignment captures VMS dopamine dynamics when animals have explicit knowledge of trial type and the required action. This allows us to characterize how dopamine evolves following correct action selection, which is central to our research question about how dopamine differs during successful Go versus No-go action execution, as well as no overt action (Free) in the long-task variant.

Nose-poke exit alignment captures dopamine dynamics at the moment animals initiate movement. The reviewer suggests that exiting for Go versus No-go trials is "very different and not comparable" because subsequent actions differ (two lever presses vs. direct reward retrieval). However, this is precisely our experimental manipulation: we compare dopamine signals when animals exit the nose-poke port to perform different actions. This comparison is both valid and necessary to address our research question (does dopamine encode action initiation?).

Magazine arrival alignment captures dopamine dynamics at reward approach. This allows us to differentiate between spatial proximity to reward from other concepts including temporal proximity (how soon is reward) and action requirements.

We acknowledge that we cannot identify the precise moment of decision formation. In fact, the precise moment of decision formation is irrelevant for our question. However, our aim is to characterize VMS dopamine dynamics during successful action execution for rewards, following cues associated with specific actions.

> *(6) The voltammetry analysis did not appear to test the hypotheses the authors outlined in the intro. All comparisons were done within task variants (DA dynamics in Go vs. No Go trials, aligned to different task events), but there were no comparisons across task variants to determine if the DA signal differed in cued vs self-initiated trials.*

Our aim was to investigate whether VMS dopamine signals consistently differed between action initiation and action suppression during reward pursuit. To test this within-variant contrast (Go > No-go), we manipulated how reward pursuit is initiated (self- vs cue-initiated) and the “effort” requirements (short vs long task variants), and our results show that they did not affect the differential between Go and No-go.

The consistent Go > No-go dopamine that we observed across all task variants, together with the consistent increase during magazine approach, supports our conclusion that VMS dopamine integrates motivated action and reward.

The reviewer suggests that we should have compared dopamine signals across self- vs cue-initiated task variants. We acknowledge this is an interesting, but entirely different, question and have addressed it in the Discussion section. We note that differences in controllability altered the time course of increased VMS dopamine, presumably by triggering earlier positive RPEs in Cue-initiated tasks as compared to Self-initiated tasks (illumination of the nose-poke light being the earliest predictor of reward). However, since our primary research question relates to the difference in VMS dopamine between action initiation and suppression, our results show that this difference was unaffected in two variants (short and long), strengthening our conclusions about the relationship between action and reward-related dopamine signaling.

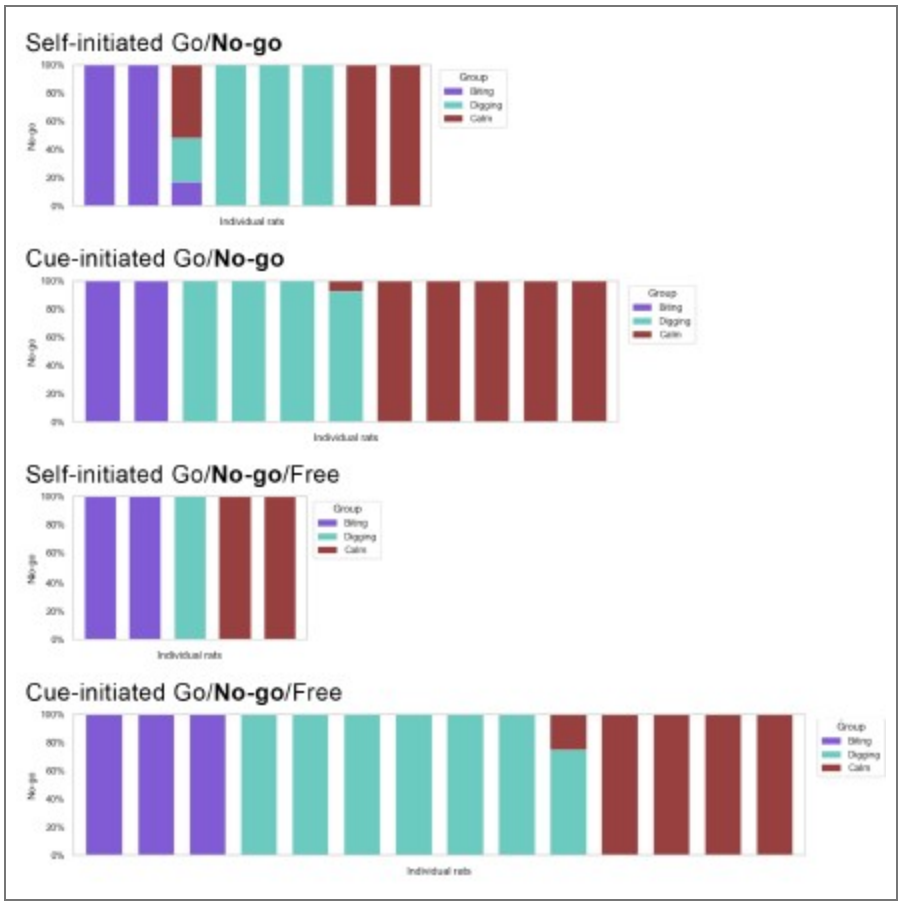
> *(7) Classification of No Go behaviors was interesting, but was not well integrated with the rest of the paper and was underdeveloped. It also raised more questions for me than*

> *answers. For example:*
>
> *(a) Was the behavior classification consistent across rats for all No Go trials? If not, did the DA signal change within subjects between biting vs digging vs calm?*
>
> *(b) If "biting rats" were not always biting rats on every No Go trial, then is it fair to collapse animals into a single measure (Figure 3C).*
>
> *(c) Some of the classification groups only had 2 or fewer rats in them, making any statistical comparison and inference difficult.*

Behavioral classification for each rat was consistent across trials (i.e., 100%, see Author response image 5). Upon reviewing the consistency of classifications within individual animals, we found that “Biting” animals exhibited biting behavior across the majority of their No-go trials. Only one animal (in the Self-initiated Go/No-go "short" variant) showed mixed classifications across trials, occasionally exhibiting digging or calm behavior. For all other animals, the predominant behavioral classification was highly consistent within subjects across sessions. The occasional trials where “biting” animals did not bite were too infrequent to permit meaningful within-animal comparisons. Therefore, we believe collapsing animals by their predominant behavioral phenotype in Figure 3C is appropriate and accurately represents stable individual differences in No-go response strategies.

As stated in the Methods Section 6: Statistical Analysis - Clustering No-go behaviors and regrouping animals (last-line), we specifically avoided between-subjects statistical comparisons for groups with $n \leq 2$, as this would be inappropriate (see Author response image 5), and reported qualitative observations only. These exploratory findings at the individual-trial level suggest behavioral heterogeneity during No-go trials, that others may use for future investigation, but do not form primary conclusions.

The behavioral classification is integrated with our central findings on VMS dopamine encoding spatial proximity. Our results demonstrate that individual variation in action suppression strategy, in particular Biting behaviors, consistently manipulates the timing of max dopamine release during subsequent reward approach, but not during the action itself (Figure 3b-c). This links our observations of action-dependent dopamine (Figure 2c) with spatial reward approach (Figure 2e). We believe that our findings shed new light into the understanding of how action modulates reward-related dopamine dynamics at the individual level. This has been discussed in Discussion section: Dopamine dynamics are linked to motivated action.



Author response image 5. Rats predominantly stick to a particular strategy to perform No-go trials. Each bar represents an individual animal, and colours represent the % of each classification type.

> ***Recommendations for the authors:***
>
> ***Reviewer #1 (Recommendations for the authors):***
>
> *(1) Figures: It would be helpful to have more panel labels on the figures - 1C, for example, labels 4 different dataset panels, similar to a few other cases. This would really help to improve readability, as there are a lot of task types, trial types, behavioral measures, and labels to sift through. The figures overall are very busy, and it is a bit of a challenge to process the tasks and trial comparisons, as well as interpret what the quantification insets mean.*

We thank the reviewer for their suggestion. We have added more panel labels to Figure 1 to improve the ability to follow with the main text. We hope that the reviewer finds it acceptable.

> *(2) Figures: A bit more specifically, in Figure 2, the boxplot insets are pretty hard to see, and it's not clear what scale they are on or what data they reflect. Similarly, it's unclear what the horizontal bars reflect in terms of which conditions are being compared. Why are box plots used for some comparisons, and why are some comparisons based on time series bootstrapping, but others are not clear? I would consider broadly reworking this figure and its description for clarity. Figure 3, by comparison, is easier to understand - the quantifications are clearer and labeled.*

We have now labeled the comparisons being depicted by the horizontal bars in Fig 2c. We have also clarified the boxplot analysis in Fig 2d and Fig 2e by adding 'Max dopamine' labels to the figure, and have made these analysis methods more explicit in the figure legend.

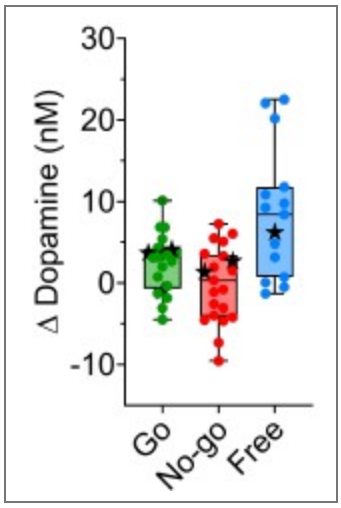
We used time-series bootstrap analysis to identify when dopamine signals diverged between trial types, which depicts dopamine differences during distinct action requirements. The latency-to-max quantification provides a summary measure to test specific encoding hypotheses (temporal vs. spatial proximity to reward).

> *(3) Broadly, more clarity on the FSCV analysis is warranted.*
>
> *(a) Targeting: It looks like the dataset contains a mix of medial shell and mostly core accumbens placements. The paper treats VMS as a uniform dopamine region, but it is more standard to separate core and shell (and also other parts of the shell) into subregions. Many of the reported encoding profiles here are known to differ across the accumbens. So, some consideration of this seems appropriate - a minimal signal-behavior comparison for the shell vs core subgroups, for example.*

We thank the reviewer for raising this point. Upon careful re-examination of our histological analysis, we identified an error that occurred when we made the overlay of electrode placements across rostrocaudal planes (to project placements onto a single plane for the sake of simplicity; Fig 2a): We incorrectly assigned some recordings to the nucleus accumbens shell. We corrected this error, which shows that the vast majority of recordings were in the core, with only 2 animals in the shell (black stars). We have adjusted Fig 2a to reflect this, and added an anatomically more complete illustration of electrode placements across rostrocaudal planes as Supplementary Figure S8.

While we acknowledge reported differences between core and shell dopamine in some contexts, the small number of shell placements precludes meaningful statistical comparison. However, to determine whether average dopamine concentrations differed between core and shell during the action-cue period, we have plotted the values in Author response image 6. The fact that shell data mostly falls centrally into the overall core-data distribution suggests no consistent difference in dopamine release.

Furthermore, in our experience (and that of colleagues (personal communication)) with appetitive operant tasks, core and shell FSCV dopamine signals do not substantially differ for action-selective encoding and reward approach. Given the sample distribution and our focus on general VMS function in Go/No-go behavior, we believe pooling these regions is appropriate.



Author response image 6. Average dopamine release during the action cue in nucleus accumbens core (circles) and shell (stars) animals showed no distinct separation between regions. Each symbol represents an animal.

> *(b) Design: In my understanding of the design, the main distinction between the short and long task variants is a 2-second versus a 3-second required nose poke hold. 3 seconds here is "long" and more "difficult". I'm not sure I agree that a 1-sec distinction really reflects a difference in task difficulty or effort. Can the authors point to a past paper that demonstrates this variation is sufficient to engage a behavioral difference and/or a neural encoding difference? Broadly, some justification of the validity of this manipulation is needed, I think.*

While we lack direct citations for this specific manipulation, we believe that the 2s and 3s hold requirements represent meaningful differences in difficulty based on our extensive rat-behavior experience and behavioral evidence in Author response image 2.

Importantly, the difficulty of No-go trials does not stem merely from the required time to hold their snouts in the nose-poke port, but from suppressing the motivational/Pavlovian bias to approach reward-associated cues. In No-go trials, subjects must suppress the prepotent tendency to immediately approach reward-related stimuli and instead maintain active suppression of this approach behavior. Even the 2s hold is challenging as animals tend to perform better on Go trials compared to No-go trials (Fig 1b and 1c), demonstrating the inherent difficulty of response suppression even at the shorter duration. The additional 1-second substantially increases this demand, as it represents a 50% increase in hold duration.

Our training data clearly demonstrate the difficulty in reaching task criterion when increasing the required action (for Go and No-go trials) from 2s to 3s. Across four cohorts of animals trained in Go/No-go task variants, one cohort that was trained first in the 2s task variant, and the remaining three cohorts were trained in the 3s task variant of Go vs No-go. Animals required 41.1 ± 7.9 (n = 27) sessions to learn the 3s hold (approximately 8 weeks), versus

18.5 ± 7.6 sessions for the 2s hold (approximately 4 weeks; mean ± SEM). This indicates that despite only a 1-second difference in required action performance, animals needed more than double the number of training days to reach criterion, clearly indicating differential effort demands.

> *(c) Figure 2 results: the authors state that because there is a greater DA signal to Go vs No Go cues in all the task variants, this means that controllability of reward pursuit and increased task effort do not affect VMS dopamine. But the magnitude of the signals looks*

> *different across the task variants - it looks clearly stronger overall in the self-initiated tasks, for example. Given that dopamine signals are not compared across task variants (I think the tasks are all between-subjects?), I don't think the above conclusion is justified.*

We respectfully clarify that our conclusion does not claim controllability and effort have no effect on dopamine magnitude, but rather that these manipulations do not affect the action-selective difference in dopamine (Go > No-go). Our central finding is that the relative difference between Go and No-go remains consistent across all task variants (within-subjects comparison).

We did not perform across-variant comparisons of absolute dopamine magnitudes because that was not our primary research question. Our focus was to understand whether dopamine differs between action initiation and suppression, and whether this difference can be modulated by controllability or effort.

We acknowledge the reviewer's observation that absolute magnitudes appear larger in self-initiated vs cue-initiated task variants. We believe that this likely reflects differences in RPE timing rather than controllability per se: in cue-initiated tasks, the nose-poke light provides an early trial-start signal, distributing RPE temporally across the trial. In self-initiated tasks, trial-initiation and action requirements are temporally integrated. Though understanding how controllability affects absolute dopamine magnitude is an interesting question for future work (e.g., using sophisticated regression-based encoding models), it was beyond the scope of our current investigation, which focuses on action-selective encoding.

> *(4) Broadly, I don't think these data, as shown, support the conclusion "dopamine reflects action initiation but not controllability or effort" without more analysis and additional context.*

We have revised the wording of our conclusions throughout the manuscript to better reflect our intent, which is to compare the action-selective encoding of dopamine (i.e., action initiation vs action suppression). It is now "Reward-related dopamine depends on action initiation irrespective of controllability and effort".

> *(a) Figure 3 - more description of the classified behaviors would be helpful for interpreting this part of the data. When are the behaviors occurring - during the hold cue? Or is the classification related to what they do immediately after holding? Or something in between> I guess I'm not sure what digging and biting are in the context of a nose poke hold. As described, it's not clear what the signal differences relate to - movement differences? Generally, it's not clear what to make of the behaviors. They seem to emerge spontaneously, but it's not clear whether the specific actions mean anything, so it's a bit difficult to know what to glean from the dopamine is greater during "biting". It's a very different movement pattern, so perhaps this result relates to that, rather than task engagement or motivational drive per se?*

We thank the reviewer for the comment. We have added relevant information in the figure caption and in the Results section to clarify that classified behaviors occurred during the action-cue period (for No-go trials, the hold cue; Figure 3a caption). In addition, we have included Videos to better depict the classified No-go behaviors during the action-cue period.

We agree with the comment that these classified behaviors, such as biting, seem to emerge spontaneously. Our interpretation of these behaviors is that they may represent the motivational state of each subject. Most importantly, whereas the behavioral differences occurred during the action-cue period (while animals had to suppress actions and stay within the nose-poke port), the difference in VMS dopamine was only observable after this behavior was completed. Thus, the movement pattern per se is likely not relevant to the dopamine release occurring after its completion. This has been discussed in the Discussion section: Dopamine dynamics are linked to motivation action.

Based on our videos, it appears as though Digging could be perceived as more vigorous (i.e., more general movement in the nose-poke port). However, we did not observe more dopamine during the action-cue period of Digging trials as compared to Biting trials. Furthermore, more vigor during the action-cue period (e.g. Digging trials) did not result in more dopamine during the reward approach period. Together, the data suggest that another process may underlie the large increase in VMS dopamine in Biting trials during reward approach, such as varying attribution of incentive salience.

> *(b) In some cases, but not all, dopamine measurement comparisons are done on a total trial basis, and in others, it seems to be subject averages. It's not clear why different approaches are used for different parts of the data. But also, for the trialwise analysis, what statistical steps were taken to incorporate the subject as a random factor in the analysis? If that is not done, then a trial-wise analysis artificially increases the power for the stat (n=trial#).*

We used different analytical approaches depending on sample size and data structure. To compute differences in Go vs No-go dopamine within each animal, as intended by our experimental design, we performed subject-level comparisons (Figure 2).

For the behavioral classification analysis (No-go, Figure 3), we performed trial-level analyses to increase the statistical power and better characterize this unexpected and interesting phenomenon. We explicitly chose not to perform subject-level group comparisons because

(1) some groups had only n=2-3 animals, making subject-level statistics underpowered, and (2) behavioral classifications were highly stable within individual animals (see Author response image 5). We acknowledge that formal between-group comparisons (across subjects) are underpowered due to small n, but the stability of within-subject No-go behavioral strategy and qualitatively distinct VMS dopamine profile suggest that these differences may be biologically meaningful and worthy of future investigation in larger samples. We have made these limitations more explicit in the Results.

> *This relates to Figure 3, where all trial data are shown next to individual subjects - the subject-wise group comparisons are between 2-5 or so rats, which is quite low. In Figure 2, a subject n of 27 is listed, so it's not clear why this analysis is on such a small set of rats. Generally, it's not clear how many rats/subjects are in each data bit. The methods say only 5 rats are in the long self-initiated task, but 15 in the cue-initiated task. Clarity in all this is needed, including in the figures/captions.*

We have now added detail about the statistical test performed in Figure 3's caption:

"After action-cue offset: No-go (trials)': Individual No-go trials classified by No-go behavior, and trial-wise statistical Kruskal-Wallis tests were performed for each task-variant." We have also added in a table in Methods (Table 1), showing the number of trials in each No-go classification, and animals regrouped based on their predominant No-go strategy (see Methods section: Statistical analysis - Clustering No-go behaviors and regrouping animals).

For the small sample sizes based on the regrouping of animals based on their predominant No-go strategy, we have now added in the caption, "... Rats classified based on their predominant No-go strategy, with no statistical tests performed."

> *(5) I'm also a little confused about the paper's narrative that the dopamine data reflect spatial (but not temporal) proximity to reward - it seems that this conclusion is based on the dopamine signal peaking at magazine entry, but that is different, I think, than a spatial signal per se (space is not manipulated in this study). I think more analysis of the signals during the magazine approach behaviors would be helpful, and possibly comparing rewarded vs unrewarded approaches. The emphasis, including in the title, that a major take-home of the data is that dopamine encodes reward proximity, is not really borne out by the current analyses. Reward expectation is not manipulated independently of the approach action, so it's hard to pin this on "space" vs "reward is soon". This is admittedly a general complexity in characterizing dopamine ramps.*

As the reviewer notes, we acknowledge that 'spatial proximity' and 'reward is soon' are challenging to fully dissociate in appetitive approach paradigms. However, we believe that our data and new additional analyses, which is now included in Results: Maximum VMS dopamine release encodes spatial but not temporal proximity to rewards and Supplementary Figures S10-12, provide compelling evidence that VMS dopamine primarily reflects spatial proximity to the expected reward location, rather than temporal proximity to reward delivery.

Evidence against full temporal encoding:

(1) Max VMS dopamine occurred up to 3s before reward delivery in Free trials (Fig 2e, open circles vs triangles). Furthermore, when we calculated the max values of individual trials of realigned traces, maximum dopamine does not consistently coincide with reward delivery across trial types (new Supplementary Figure S11).

(2) If dopamine encoded temporal proximity from the earliest reward-predictive cue, we would expect consistent accumulation from cue onset (action cue for self-initiated, nose-poke light for cue-initiated). However, realigned trials also did not show a consistent accumulation around these events (new Supplementary Figure S12).

Evidence for spatial encoding:

(1) Max dopamine consistently occurred when animals arrived at the magazine across all trial types and task variants, regardless of when reward was actually delivered (Fig. 2e).

(2) Across individual trials, max dopamine values concentrated around the magazine-panel, with a striking accumulation when animals were in close proximity to it (new Supplementary Figure S10) showing a distance-dependent distribution.

(3) Assessment of unrewarded magazine approaches during the intertrial interval (ITI) revealed no increase in dopamine release (Fig 2f), indicating that VMS dopamine requires task-relevant reward expectation.

> *(6) Examples of the DLC workflow in a supplement would be appropriate. Also, video examples of the 3 kinds of behaviors from the clustering analysis could be useful for understanding what they are/what they mean.*

We have added supplementary figures showing the DeepLabCut workflow (Supplementary Figure S9) and Videos 1-3 demonstrating the three behavioral classifications (biting, digging, calm) during No-go trials.

> *(7) Referencing/scholarship: I would suggest broadening the citation pool for the paper to include more older work that has established the notion that dopamine signaling and the accumbens act as a motivation-action interface, as this has been a longstanding notion since at least the 1980s. There is also a sizable literature on dopamine signaling of effort, some of which would be appropriate to cite here.*

We thank the reviewer for the recommendation and have now expanded our citations to include foundational literature to work from the 1980s-90s. These can be found in the Introduction, Results, and Discussion.

> ***Reviewer #2 (Recommendations for the authors):***
>
> *(1) Lines 353-357- This came as a surprise, as the relevant results are only featured in supplementary information. These should be moved to the main manuscript. As both the biting behavior and faster lever-press completion lead to larger peak dopamine, does this represent response vigor?*

We appreciate the reviewer's interest in these data. However, we believe that the data that we report in "Supplementary Fig 7: Quartile analysis of Go trials shows coordinated changes in last lever-press timing and VMS dopamine, does not warrant movement to the main manuscript."

The purpose of the lever-press timing analysis was to demonstrate that maximum dopamine release coincides with the moment animals arrived at the magazine, rather than with the action period itself. By sorting Go trials based on last lever-press latency, we show temporal coordination between action completion and dopamine timing but critically, dopamine peaks after action completion, not during it.

This temporal dissociation argues against a 'response vigour' interpretation. If dopamine encoded motor vigour, we would expect the signal to coincide with or precede the vigorous action. Instead, both the lever-press data (Supplementary Figure S7) and the classified No-go trial data (Figure 3, biting behavior) show that changes in max dopamine occur after the actions themselves, during the subsequent approach to reward.

Together, these findings demonstrate that VMS dopamine reflects spatial approach to the reward location following action completion, not the vigour of the required actions per se. The lever-press analysis serves as supporting evidence for this temporal relationship but does not introduce a novel finding that warrants main figure emphasis. We mentioned this temporal coordination in Results to ensure readers are aware of the converging evidence while maintaining focus on our central findings regarding action initiation versus suppression.

> *(2) Line 13- the experimental work cited refers to midbrain dopamine neurons, rather than dopamine release within the VMS. Please correct.*

We thank the reviewer for pointing out this error. We have now corrected the citation to refer to studies measuring striatal dopamine rather than midbrain dopamine neurons.

> *(3) Line 166 - Shouldn't this say "consistently delayed for No Go trials"? Looks like peak dopamine occurs later for these trial types.*

This statement refers to traces aligned to nose-poke exit (not action-cue onset). The peak Go dopamine occurred later than other trial types following exit (Green arrows in Figure 2d).

> *(4) Lines 391-392 - however however*

Thank you for the comment, we have adjusted the text.

https://doi.org/10.7554/eLife.109868.2.sa0