Dopamine neurons encode errors in predicting movement trigger occurrence

Benjamin Pasquereau and Robert S. Turner

Department of Neurobiology, Center for Neuroscience and The Center for the Neural Basis of Cognition, University of Pittsburgh, Pittsburgh, Pennsylvania

Submitted 27 May 2014; accepted in final form 18 November 2014

Pasquereau B, Turner RS. Dopamine neurons encode errors in predicting movement trigger occurrence. J Neurophysiol 113: 1110–1123, 2015. First published November 19, 2014; doi:10.1152/jn.00401.2014.—The capacity to anticipate the timing of events in a dynamic environment allows us to optimize the processes necessary for perceiving, attending to, and responding to them. Such anticipation requires neuronal mechanisms that track the passage of time and use this representation, combined with prior experience, to estimate the likelihood that an event will occur (i.e., the event’s “hazard rate”). Although hazard-like ramps in activity have been observed in several cortical areas in preparation for movement, it remains unclear how such time-dependent probabilities are estimated to optimize response performance. We studied the spiking activity of dopamine neurons in the substantia nigra pars compacta of monkeys during an arm-reaching task for which the foreperiod preceding the “go” signal varied randomly along a uniform distribution. After extended training, the monkeys’ reaction times correlated inversely with foreperiod length, highlighting a progressive anticipation of the go signal according to its hazard rate. Many dopamine neurons modulated their firing rates as predicted by a succession of hazard-related prediction errors. First, as time passed during the foreperiod, slowly decreasing anticipatory activity tracked the elapsed time as if encoding negative prediction errors. Then, when the go signal appeared, a phasic response encoded the temporal unpredictability of the event, consistent with a positive prediction error. Neither the anticipatory nor the phasic signals were affected by the anticipated magnitudes of future reward or effort, or by parameters of the subsequent movement. These results are consistent with the notion that dopamine neurons encode hazard-related prediction errors independently of other information.

Although hazard-like activity has been observed in several cortical areas over the course of the foreperiod preceding an action (Genovese et al. 2006; Janssen and Shadlen 2005; Mauritz and Wise 1986; Niki and Watanabe 1979; Premereur et al. 2011; Renoult et al. 2006), it remains unclear how such time-dependent probabilities are estimated and updated in the brain. Converging evidence suggests that this ability is critically dependent on intact dopamine function. Interval timing is distorted by dopaminergic agonists or antagonists (Cheng et al. 2007; Coull et al. 2011; Drew et al. 2003; Matell et al. 2006; Meck et al. 2012; Narayanan et al. 2012; Rammayer and Classen 1997; Rammayer 1993), and difficulties in time perception are often found in disorders associated with dopaminergic dysfunction such as Parkinson’s disease (Harrington et al. 1998; Pastor et al. 1992), schizophrenia (Carroll et al. 2008, 2009; Elvevag et al. 2003; Tysk 1983), and attention-deficit/hyperactivity disorder (Rubia et al. 2009).

Most current research on the activity of midbrain dopamine neurons focuses on their responses to primary rewards or reward-associated sensory cues (Bromberg-Martin et al. 2010b; Schultz 2006, 1998; Wise 2004). Findings from those studies position dopamine neuron activity as a neurophysiological correlate of the reward prediction error signal posited by classical models of reinforcement learning (Bayer and Glimcher 2005; Hollerman and Schultz 1998; Schultz 1998; Schultz et al. 1997; Sutton and Barto 1998). As a facet of that prediction error signal, dopamine responses are sensitive to temporal aspects of reward expectation, both at the time of reward-predicting stimuli and at the time of the reward itself (Fiorillo et al. 2008; Kobayashi and Schultz 2008). For instance, when a Pavlovian-conditioned sensory stimulus precedes reward delivery by a fixed-duration delay (i.e., fixed across successive trials), the dopamine response to the stimulus is smaller for longer delays (reflecting a temporal discounting of reward value), whereas the response to the reward is larger for longer delays (possibly as a result of increasing temporal uncertainty with longer delays). Significantly, when Fiorillo et al. (2008) varied stimulus-to-reward intervals randomly from trial to trial, reward-evoked dopamine responses were instead smaller for longer delays, suggesting that the magnitude of the reward-evoked dopamine response was anticorrelated with the hazard rate of reward delivery. The generality of that interpretation is uncertain, however. For example, Bromberg-Martin et al. (2010a) reported that dopamine responses to a start-of-trial stimulus were larger when the delay to stimulus delivery was longer.

In the present study, we explored this understudied dimension of dopamine neuron activity by investigating whether and how these neurons change their activity in relation to an action-related event that varies temporally (i.e., a go signal that
triggers a reaching movement). Unlike the majority of recent studies which have used Pavlovian tasks (Fiorillo et al. 2003, 2008; Joshua et al. 2008; Kobayashi and Schultz 2008; Matsuno and Hikosaka 2009; Tobler et al. 2005), we studied dopamine activity in monkeys that were conditioned operantly to perform an arm-reaching task. Our results expand on the pioneering observations of Schultz et al. (1993) in which a phasic dopamine signal was observed only when a movement trigger was temporally uncertain. In the present study, we report that dopamine neurons of the substantia nigra pars compacta (SNc) modulate their activity according to an action-related event’s subjective hazard rate, consistent with an encoding of error signals related to the estimation of event probability. This activity was not influenced by trial-to-trial variations in task performance or the anticipated levels of effort or reward, suggesting a role in task monitoring.

MATERIALS AND METHODS

Animals. Two rhesus monkeys (monkey C: 8 kg, male; monkey H: 6 kg, female) were used in this study. Procedures were approved by the Institutional Animal Care and Use Committee of the University of Pittsburgh (protocol no. 12111162) and complied with the Public Health Service Policy on the humane care and use of laboratory animals (amended 2002). The animals were housed in individual primate cages in an air-conditioned room where water was always available. Monkeys were under food restriction to increase their motivation in task performance. Throughout the study, the animals were monitored daily by an animal research technician or veterinary technician for evidence of disease or injury, and body weight was documented weekly. If a body weight <90% of baseline was observed, the food restriction was stopped.

Apparatus. Monkeys were trained to execute reaching movements with the left arm using a two-dimensional (2-D) torqueable exoskeleton (KINARM; BKIN Technologies, Kingston, ON, Canada). Visual targets and cursor feedback of hand position were presented in the horizontal plane of the task by a virtual reality system. Distinct force levels were applied at the shoulder and elbow joints by two computer-controlled torque motors so as to simulate kinetic friction loads opposing movement of the hand. The force was constant for any hand movement above a low threshold velocity (0.5 cm/s) and directly opposed the direction of movement. A detailed description of the apparatus and paradigm can be found in Pasquevuar and Turner (2013).

Effort-benefit reaching task with variable foreperiod. Monkeys were trained to perform a reaching task that manipulated movement incentive and effort independently (Fig. 1). On each behavioral trial, the animal was required to align the cursor with two visual targets (radius: 1.8 cm) displayed in succession (Fig. 1A). A trial began when a target appeared at the start position and the monkey made the appropriate arm movement to align the cursor with this target. The monkey maintained this position for a random-duration hold period (2.4–4.6 s) during which 1) an instruction cue was displayed at the start position (0.5-s duration) and 2) the peripheral target appeared (0.7–1.3 s after offset of the instruction cue, at the same location for all trials, 7.5 cm distal to the start position). At the end of the start position hold period, the start position target disappeared (the go signal), thereby cueing the animal to move the cursor to the peripheral target. The time interval between appearance of the peripheral target (the warning stimulus) and go signal (the imperative stimulus) varied randomly from trial to trial (uniform distribution between 0.5 and 1.5 s at 1-ms resolution). This variable time interval was defined as the preparatory foreperiod. The monkey was required to hold the cursor at the start position during the foreperiod and then initiate an arm movement after the go signal to capture the peripheral target within <0.8 s. At a random delay after successful target capture (0.7–1.3 s), food reward was delivered via a sipper tube attached to a computer-controlled peristaltic pump (1 drop = ~0.5 ml, puree of fresh fruits and protein biscuits). For the instruction cues, symbols indicated the level of upcoming force that the animal would encounter (black, 1.8, or 3.2 N) and cue colors indicated the size of reward (1, 2, or 3 drops of food; Fig. 1B). Nine distinct cue types (3 force levels × 3 reward sizes) were presented in pseudorandom order across trials with equal probability. The trials were separated by 1.5- to 2.5-s intertrial intervals (randomly distributed at 100-ms resolution) during which the screen was black.

Surgery. After reaching asymptotic task performance, animals were prepared surgically for recording using aseptic surgery under isoflurane inhalation anesthesia. An MRI-compatible plastic chamber (custom-machined PEEK, 28 × 20 mm) was implanted with stereotoxic
guidance over a burr hole allowing access to the SNc (Horsley-Clark coordinates; anterior 11, lateral 4, depth 0). The chamber was positioned in the parasagittal plane with an anterior-to-posterior angle of 20°. The chamber was fixed to the skull with titanium screws and dental acrylic. A titanium head holder was embedded in the acrylic to allow fixation of the head during recording sessions. Prophylactic antibiotics and analgesics were administered postsurgically.

**Localization of the recording site.** The anatomic location of the SNc and proper positioning of the recording chamber to access it were estimated from structural MRI scans (Siemens 3T Allegra Scanner, voxel size: 0.6 mm). An interactive 3-D software system (Cicerone) was used to visualize MRI images, define the target location, and predict trajectories for microelectrode penetrations (Micinovic et al. 2007). Electrophysiological mapping was performed with penetrations spaced 1 mm apart. The boundaries of brain structures were identified on the basis of standard criteria including relative location, neuronal spike shape, firing pattern, and responsiveness to behavioral events (e.g., movement, reward). By aligning microelectrode mapping results (electrophysiologically characterized X-Y-Z locations) with structural MRI images and high-resolution 3-D templates of individual nuclei derived from an atlas (Martin and Bowden 1996), we were able to gauge the accuracy of individual microelectrode penetrations and determine chamber coordinates for the SNc.

**Recording and data acquisition.** During recording sessions, a glass-coated tungsten microelectrode (impedance: 0.7–1 MΩ measured at 1,000 Hz) was advanced into the target nucleus using a hydraulic manipulator (MO-95; Narishige). Neuronal signals were amplified with a gain of 10K, bandpass filtered (0.3–10 kHz), and continuously sampled at 25 kHz (RZ2; Tucker-Davis Technologies, Alachua, FL). Individual spikes were sorted using Plexon off-line sorting software (Plexon, Dallas TX). The timing of detected spikes and relevant task events was sampled digitally at 1 kHz. Dopamine neurons were identified according to the following standard criteria: 1) location within the SNc, 2) polyphasic extracellular waveforms, 3) low tonic irregular spontaneous firing rates (0.5–8 spikes/s), and 4) long duration of action potentials (>1.5 ms). The duration of action potential waveforms was calculated as the interval from the beginning of the first negative inflection (>2 SD of voltage measured) to the succeeding positive peak.

**Analysis of behavioral data.** We analyzed the way the animals performed the behavioral task to test 1) whether the behavior varied according to the levels of anticipated reward and effort, and 2) whether behavior showed evidence of hazard-like anticipation during the variable foreperiod. Data were collected after animals had extensive experience with the reaching task (6 mo for monkey H, 12 mo for monkey C). Kinematic data (digitized at 1 kHz) derived from the exoskeleton were numerically filtered and combined to obtain position and tangential velocity (Fig. 1C). The time of movement onset was determined as the point at which tangential velocity crossed an empirically derived threshold (0.2 cm/s). Reaction times (RT; interval between go signal and movement onset), movement durations (MD; interval between movement onset and capture of the target), peak velocity (maximum tangential velocity), and accelerations were computed for each trial. Trials with errors (3.5% of trials) or RTs defined as outliers (4.45× median absolute deviations, equivalent to 3× SD for a normal distribution, 3.7% of trials) were excluded from the analysis to avoid potentially confounding factors such as inattention. Two-way ANOVAs were used to test these kinematic measures for interactive effects of incentive reward and effort.

We used an animal’s RTs at different time points during the foreperiod to infer the animal’s time-varying judgment of the probability of go signal occurrence. The duration of the preparatory foreperiod, defined as the time interval between the appearance of the peripheral target and the go signal, varied randomly between trials at 1-ms resolution and ranged with a flat distribution from 0.5 to 1.5 s. In other words, the probability of go signal appearance at each of 1,000 potential onset times was \( P = 1/1,000 \) (Fig. 2A). Ideally, an observer’s anticipation of an event should be governed by the event’s objectively determined hazard rate, which is defined as the conditional probability that an event will occur given that it has not yet occurred (Fig. 2B). Here, the objective hazard rate of the go signal is the probability that it will occur at time \( t \) divided by the probability that it has not yet occurred:

\[
h(t) = f(t) / [1 - F(t)],
\]

where \( f(t) \) corresponds to the probability distribution of the go times, and \( F(t) \) is the cumulative distribution of probabilities. However, the actual anticipatory behavior of animals on foreperiod tasks differs significantly from what would be expected of an ideal observer (Tsunoda and Kakei 2008, 2011), even after an animal has been trained extensively on a specific foreperiod probability distribution. This implies that the “subjective” representation of a hazard function that an animal builds up with experience is non-ideal. Much of the divergence of subjective hazard functions from the ideal can be explained by an animal’s uncertainty about the passage of time, an uncertainty that grows with the length of the time interval [from Weber’s law for time estimation (Gibbon 1977)]. Thus we modeled each animal’s subjective hazard rate based on the assumption that elapsed time was known with an uncertainty that scaled with elapsed time in the foreperiod (Fig. 2, C and D). More specifically, the probability density function \( f(t) \) was blurred by normal distributions.

---

Fig. 2. Estimation of subjective hazard rates. A and B: the waiting time between target presentation and the go signal varied randomly from trial to trial between 0.5 and 1.5 s according to a uniform (flat) probability distribution. Probability density function (A) and hazard rates (B) were calculated according to this time schedule. C and D: for each animal, the subjective probability density function \( C \) and hazard rates \( D \) were estimated using the monkey’s task performance. E and F: subjective hazard rates were found that yielded the best linear fit between reaction times (RTs) averaged across sessions (means ± SE) and the estimated probability that the go signal was about to occur (i.e., the subjective hazard rate). Goodness of fit is furnished for each monkey (\( R^2 \) and \( P \) value). RTs were averaged across 9 reward/effort combinations; \( n \) denotes the number of sessions.
whose standard deviations are proportional to elapsed time (Janssen and Shadlen 2005; Tsunoda and Kakut 2008, 2011), as follows:

\[ \tilde{f}(t) = \frac{1}{\alpha t} \sqrt{2\pi} e^{-\frac{(t - r)^2}{2\sigma^2}}, \]

(2)

where \( \alpha \) is the coefficient of variation that reflects an individual animal’s uncertainty in time estimation. The subjective hazard rate was then obtained by substituting \( \tilde{f}(t) \) and \( \tilde{f}(t) \) into Eq. 1. We reasoned that an animal’s anticipation of go signal presentation, as measured by RTs at different time points across the foreperiod, would be modulated monotonically according to the animal’s judgment of the go signal’s hazard rate at those time points. On the basis of that logic, we identified a coefficient of variation \( \alpha \) that maximized the linear relationship between RTs (averaged across sessions separately for 20 foreperiod time points; bins of 50 ms each) and subjective hazard rate. To be specific, the MATLAB function “fminsearch” (The MathWorks) was used to find the best coefficient of variation \( \alpha \) that provided the maximum likelihood fit between RT and subjective hazard rate.

To test for any effects of go signal anticipation (i.e., the subjective hazard rate) on task performance and potential interactions with reward/effort values, we used a series of two-way ANOVAs (hazard rate × reward quantity and hazard rate × effort level) with behavioral variables (i.e., RT, MD, velocity, and acceleration). The task design involved 3 reward values combined with 3 effort levels, and we used 20 foreperiod time points in our analysis, making a total of 180 types of trials. Because of this high number of conditions and the relative limited number of trials recorded (mean = 244 trials), it was not possible to perform a straightforward three-way ANOVA. To overcome this obstacle, we collapsed the reward or effort conditions and performed a series of two-way ANOVAs.

We then investigated the possibility that foreperiod history influenced the monkeys’ RT performance. Sequential effects have been described previously in similar variable-foreperiod paradigms (Niemi and Naatanen 1981) such that a longer foreperiod on the preceding trial leads to a slower RT on the current one. Consequently, it has been assumed that temporal preparation does not depend exclusively on the conditional probability of the foreperiod duration in the current trial, but that the hazard rate function also is influenced by the time at which the go signal was presented in the immediately preceding trial (Alejia 1975). To test for such an effect related to recent history, RTs were compared between three categories of preceding trial foreperiod duration (short: <0.8 s, medium: 0.8–1.2 s, long: >1.2 s), and the potential interaction with hazard rate was examined using a two-way ANOVA. In addition, the gaze position of eyes, detected using an infrared camera system (240 Hz; ETI-200, ISCAN, Woburn, MA), was used to measure the distance of the eyes from the start position target during the foreperiod. All of the data analyses were performed using custom scripts in the MATLAB environment (The MathWorks).

Neuronal data analysis. Neuronal recordings were accepted for analysis based on electrode location, recording quality and duration (>200 trials). Adequate single-unit isolation was verified by testing for appropriate refractory periods (>1.5 ms) and waveform isolation (signal-to-noise ratio superior to 3 SD).

For different task events (e.g., instruction cue, go signal, and movement onset), continuous neuronal activation functions [spike density functions (SDFs)] were generated by convolving each discriminated action potential with a Gaussian kernel (20-ms variance). Mean peri-event SDFs (averaged across trials) for each of the effort/ reward conditions were constructed. A neuron’s baseline firing rate was calculated as the mean of the SDFs across the 500-ms epoch immediately preceding cue instruction. A phasic response to a task event was detected by comparing SDF values during a peri-event epoch (500 ms) relative to a cell’s baseline firing rate (\( P < 0.01 \), r-test). A neuron was judged to be task-related if it generated a significant phasic response for at least one of the nine task conditions. Response onset and offset times were defined as the times at which the SDF first and last crossed the \( P = 0.01 \) threshold, respectively. To compare the neural responses (magnitude, latency, and duration) to different task events, we used SDFs averaged for the high reward value only, because phasic dopamine response to the instruction cue event was larger for this task condition (Pasquearu and Turner 2013). Magnitude corresponded to the maximum activity detected in SDF relative to the baseline. Then, for neurons that had a significant instruction-evoked or go-evoked response, we tested for effects of reward and effort using a single-neuron-based analysis. Mean SDFs for each of the nine trial types (3 reward quantities × 3 effort levels) were searched independently, and the point of largest deviation across all trial types was taken as the time of the neuron’s cue-related response. Single-trial spike counts (SC) were extracted from a 100-ms window centered on the time of the maximal response. Two-way ANOVAs were used to test whether these cue-evoked SCS were influenced by 1) reward, 2) effort, or 3) the additive effects of reward and effort.

For neuronal responses that occurred during the premovement period (i.e., during the RT), we tested whether the onset of the response was time locked, from trial to trial, to the occurrence of the go signal or to movement onset. First, for each valid trial, phasic changes in discharge rate were detected in the spike train using the Legendy surprise method (Legendy and Salaman 1985; Pasquearu and Turner 2011). Single-trial changes in discharge were defined as groups of two or more spikes whose interspike intervals (ISIs) were unusually short compared with other ISIs measured in the spike train during the preceding second. We used a surprise threshold of 4. Second, the onset times of single-trial changes in discharge were tested for significant temporal correlation (i.e., time locking) with the times of go signal and movement onset. Such time locking has been considered to be evidence for an underlying functional linkage between the behavioral event and the linked neural discharge (Hanes et al. 1995; Turner and Anderson 1997). If a neuron’s discharge has a close temporal relationship to the go signal, then the time locking between movement onset and the neuron’s initial change in discharge should covary, across trials, with the behavioral RT. If, on the other hand, response onset times show a tighter temporal linkage to the movement onset, then the interval duration between the go signal and change in discharge should vary with the RT. To test for these linkages, Pearson correlations (\( P < 0.05 \)) were calculated for interval durations (e.g., go-response and response-movement) vs. the behavioral RT. A significant positive correlation between the go-response interval and RT in the absence of a correlation between response-movement interval and RT was taken to indicate that changes in discharge were time locked to movement onset. Conversely, a significant positive correlation between response-movement interval and RT, but not go-response interval and RT, implied that the initial change in discharge was time locked to the go signal. If neither or both correlations were significant, then the time locking was considered to be indeterminate.

To test whether neuronal activity encoded an animal’s estimate of go signal probability, we performed a series of regressions comparing SC with the subjective hazard rate. We used a sliding window procedure around the time of go signal onset to determine the timing and strength of such an influence. For each neuron, we counted spikes in a 200-ms test window that was stepped in 25-ms increments from −1,000 to +1,000 ms relative to go signal. For each step, a linear regression was computed with SC as the dependent variable and subjective hazard rate as the independent variable. To test whether individual regressions were significant, we shuffled spike counts across trials repeatedly (1,000 times) and compared actual regression coefficients to the 99% confidence interval of coefficients yielded by shuffling. The threshold for significance (\( P < 0.01 \)) was validated by calculating the likelihood of type I (false positive) errors in a parallel analysis of activity during 200-ms windows preceding the target...
presentation (i.e., during a time period when there should be no neuronal encoding). Second, we investigated whether hazard-related changes in neuronal discharge rate were modulated by trial-to-trial differences in effort/reward values. With SC as the dependent variable, a series of regression models were performed to test for possible interactions with reward or effort values:

\[ SC_i = \beta_0 + \beta_1 \text{Reward}_i + \beta_2 \text{Effort}_i + \beta_3 \text{Reward}_i \times \text{Effort}_i, \]  

where \( SC_i \) is the number of spikes in the \( i \)th time bin in the sliding window procedure, and all variables were normalized (Z-score) to obtain standardized regression coefficients (\( \beta_{0-3} \)). To test whether individual regression coefficients were significant, we used the same shuffling method described above.

The influence of hazard rate on the activity of dopamine neurons before and after appearance of the go signal was then analyzed using two distinct time windows. The pre-go window (which reflected foreperiod information) was defined as the firing rate from \(-200 \) to \(+50 \) ms relative to go signal, and the post-go window measured the firing rate in a 150-ms window centered on the time of the maximal go-evoked response. Single-trial measures of a neuron’s pre-go and post-go activity were sorted according to the trial’s foreperiod duration (20 bins of 50 ms each) and then regressed against subjective hazard rate.

In addition, we performed population-based analyses using mean peri-go SDFs. For neurons that were found to phasically respond to the go signal, we tested whether population-averaged activities (during instruction and go signal) encoded effort and/or reward levels by using a two-way ANOVA combined with the sliding window procedure. To examine the population-averaged activity of hazard-encoding neurons, we then dissociated trials into three distinct categories according to the probability that the movement trigger was about to occur (low rate: \(<1.3 \times 10^{-3} \), medium rate: \(1.3–2.3 \times 10^{-3} \), high rate: \(>2.3 \times 10^{-3} \)). Mean peri-go signal SDFs (averaged across neurons) were constructed for each category of hazard rate and compared using a series of one-way ANOVAs.

Finally, to test for possible encoding of parameters related to movement execution, we used time-resolved multiple linear regressions. We simultaneously tested whether trial-to-trial neuronal activity was modulated by the subjective hazard rate, task factors (reward and effort), and measures of performance (RT, maximum velocity, and initial acceleration). In other words,

\[ SC_i = \beta_0 + \beta_1 \text{Reward}_i + \beta_2 \text{Effort}_i + \beta_3 \text{RT}_i + \beta_4 \text{Velmax}_i + \beta_5 \text{Acc}_i, \]  

where \( SC_i \) is reward and \( Acc \) is acceleration. To test whether individual regression coefficients (\( \beta_{0-5} \)) were significant, we shuffled spike counts across trials repeatedly (1,000 times) and compared actual regression coefficients to the 99% confidence interval of coefficients yielded by shuffling.

**Results**

Before presenting results related to the foreperiod epoch itself, we first establish that the animals performed the behavioral task correctly and show that the neurons studied met standard criteria for midbrain dopamine neurons.

**Behavioral effect of instructions.** The nine conditioned stimuli presented during the instruction period of the reaching task (Fig. 1B) effectively communicated nine different levels of motivational value. This was evidenced by consistent effects on the animals’ task performance. Reaction times, error rates, and kinematic measures were affected by both the expected reward magnitude (RT: \( F > 11.65, P < 0.001 \); error rates: \( F > 14.54, P < 0.001 \); movement durations: \( F > 11.04, P < 0.001 \); accelerations: \( F > 18.22, P < 0.001 \); maximum velocities: \( F > 22.05, P < 0.001 \)) and force (RT: \( F > 10.75, P < 0.001 \); error rates: \( F > 9.48, P < 0.001 \); movement durations: \( F > 40.44, P < 0.001 \); accelerations: \( F > 34.04, P < 0.001 \); maximum velocities: \( F > 83.62, P < 0.001 \)). Overall, movements were faster and more accurate in trials in which large benefits and low efforts were anticipated. Details about effort/reward effects on monkeys’ behavior can be found in our previous study (Pasquereau and Turner 2013).

**Neuronal responses to instructions.** The population of neurons studied here responded to instruction cues as would be expected of classically described dopamine neurons (Fiorillo et al. 2003; Hollerman and Schultz 1998; Matsumoto and Hokosaka 2009; Morris et al. 2006; Tobler et al. 2005). While the monkeys performed the reaching task, we recorded single-unit activity from 141 putative dopamine neurons in the SNc. The instruction-evoked responses of these neurons have been described in detail elsewhere (Pasquereau and Turner 2013), and we summarize those results here. Similar to previous studies (Fiorillo et al. 2003; Hollerman and Schultz 1998; Matsumoto and Hokosaka 2009; Morris et al. 2006; Tobler et al. 2005), we identified dopamine neurons on the basis of location and standard electrophysiological criteria. At the beginning of each trial, most dopamine neurons (74%, 104/141) emitted a phasic increase in discharge above the baseline firing rate following the presentation of informative cues (\( P < 0.01 \), t-test). Consistent with the basic predictions of reinforcement theory (Sutton and Barto 1998) and previous observations (Schultz 2006), the magnitude of the cue-evoked responses scaled monotonically with reward quantity in more than half of the cue-responsive neurons (61% of neurons; \( P < 0.05 \), 2-way ANOVA). Notably, in a subset of these neurons (\( n = 12 \)), the reward-encoding phasic responses were discounted by the predicted physical effort (\( P < 0.05 \)) in a manner that closely matched behavioral performance (for details see Pasquereau and Turner 2013).

The results thus far have shown that the size of reward and the level of effort signaled by early instruction cues affected both the magnitude of the dopamine neuron responses evoked by the cues and the way our animals performed the subsequent reaching movement. These results show that the task’s instruction cues effectively communicated motivationally significant cost and benefit information to the animals.

**Estimation of subjective hazard rate.** We next determined how the variable foreperiod before presentation of the go cue affected a monkey’s behavior. On the basis of previous literature (Niemi and Naatanen 1981; Tsunoda and Kakei 2011, 2008; Vallesi et al. 2007), we predicted that after extensive exposure to the task’s variable foreperiod schedule, monkeys would gradually ramp up their anticipation of the arrival of the go signal according to a hazard function of the foreperiod interval (Fig. 2). Consistent with this prediction, both monkeys had visuo-manual RTs that declined gradually as a function of the probability that the go signal was about to occur (monkey C: \(-10\% \), monkey H: \(-21\% \); \( F > 7.77, P < 0.001 \), repeated-measures 1-way ANOVA). Under the assumption that an animal’s RT performance was modulated monotonically according to its judgment of the probability of go signal occurrence (Gibbon et al. 1997; Jansen and Shadlen 2005; Tsunoda...
and Kakei 2008, 2011), we estimated subjective hazard rates (as expressed by Eqs. 1 and 2) by finding the value for α that was most effective at linearizing the RT-hazard rate relationship. To be more specific, we estimated the coefficient of variation (α) that reflects an individual animal’s uncertainty in the time estimation by finding the value for α that maximized the linear fit between RT and hazard rate (Fig. 2, E and F). This procedure was performed separately for each monkey by using measures of performance averaged across sessions (81 sessions for monkey C and 60 sessions for monkey H). As predicted, the subjective hazard rates differed markedly from the objective hazard rate (compare Fig. 2, B and D) but approximated the task performance well for both animals (P < 0.001 and $R^2 > 0.82$, linear regression). Notably, the resulting subjective hazard rates were indistinguishable between the two monkeys (Fig. 2, C and D), indicating identical perceptions of time in this variable foreperiod schedule and similar estimations of time-dependent probabilities.

Other measures of motor performance, such as movement duration ($F < 0.31, P > 0.05$) and maximum velocity ($F < 0.30, P > 0.05$), were not affected by the hazard rate, suggesting that the monkeys’ vigor to execute the reaching movement was unchanged across foreperiod durations. In addition, the gaze position of the animals’ eyes did not vary significantly across foreperiod durations ($F < 1.64, P > 0.05$), suggesting that the monkeys remained equally attentive and engaged in the task across the waiting time. Thus the changes in RT described above (Fig. 2, E and F) could not be explained by a change in the gaze position of the animals’ eyes.

Subjective hazard rates were uninfluenced by incentives and history. We showed above that the animals attended to the task’s motivation-related factors (the predicted levels of reward and effort) as evidenced by effects on the performance of the subsequent reaching movements (above and Pasquaereau and Turner 2013). We wondered whether these motivation-related factors also influenced an animal’s estimation of the go-signal’s hazard rate. We used two-way ANOVAs to test for interacting effects on RTs of the subjective hazard rate and reward size or effort level (Fig. 3). No significant interactions were found between the hazard rate and reward magnitude ($F < 1.06, P > 0.05$; Fig. 3A) or hazard rate and effort level ($F < 1.3, P > 0.05$; Fig. 3B). Thus the monkeys’ time-evolving estimation of the probability of occurrence of the go signal was not influenced by the incentive value associated with the instruction cue. Next, to test whether the recent history of foreperiods presented influenced the monkeys’ estimation of the hazard rate, we compared RTs according to the preceding trial foreperiod duration (Fig. 3C). Three categories of trials were identified on the basis of the duration of the foreperiod in the immediately preceding trial (short foreperiod: $<0.8$ s, medium: $0.8–1.2$ s, long: $>1.2$ s). Contrary to previous studies in humans (Alegría 1975; Niemi and Naatanen 1981), we found no effects of the preceding trial’s foreperiod on the current trial’s RT and no interactions with the hazard rate ($F < 2.14, P > 0.05$, 2-way ANOVA). Thus, for our data set, collected from overtrained animals, the foreperiod hazard function was not modulated by recent history.

Overall, the behavioral results can be explained by a model in which, after extensive training on the task, monkeys modulated their motor preparation according to a subjective hazard rate function based on the learned foreperiod schedule and the perceived passage of time moment-to-moment during the current foreperiod. This motor preparation process was not influenced by an animal’s motivational state or by the duration of recently experienced foreperiods.

Neuronal responses to the go signal. Next, having laid the groundwork needed to validate the behavioral task and the dopamine cell population, we set out to characterize how dopamine neurons responded to the go signal, a timing cue delivered after a variable foreperiod in the middle of the operant task. Of the 141 neurons studied, 67 (48%; $P < 0.01$, $t$-test) showed a phasic increase in discharge rate during the period following presentation of the go signal and preceding movement onset. To clarify whether these phasic signals were more closely related to delivery of the go signal or initiation of movement, we studied the temporal locking of the evoked responses. In Fig. 4, A and B, respectively, we show the spiking activity of the same example neuron aligned on presentation of the go signal and movement onset. The time of onset of the phasic increase in firing was detected in single trials using a burst detection algorithm (Legendy and Salcman 1985). In this example, a close temporal linkage was observed between the timing of go signals and the onset of the phasic response. Indeed, whereas the intervals between go
signals and neural responses were relatively constant and did not correlate with RTs (ρ = 0.16, P = 0.23; Fig. 4C), the intervals between neural responses and movement onset were strongly correlated with RTs (ρ = 0.88, P < 0.001). For the population of dopamine neurons in which time locking could be determined (43/67; Fig. 4D), phasic responses were temporally linked to the go signal (38/43, 88%) far more often than to movement onset (5/43, 12%; χ² = 25, P < 0.001). Moreover, as a population, go-linked neurons (i.e., neurons in which the response-to-movement interval correlated with RTs) had significantly larger mean correlation coefficients than those of movement-linked neurons (mean ρ = 0.48 ± 0.02 vs. 0.23 ± 0.02, P < 0.001, t-test; Fig. 4D). Thus the phasic responses of dopamine neurons following the go signal were, for the most part, temporally linked to the go signal and, consequently, had a weak relation to movement execution. The subset of dopamine neurons identified as go-responding was anatomically intermixed with the other cells in the substantia nigra (Fig. 5). No difference was observed in the locations of the different subtypes of neurons in either animal (95% of the cell distributions overlapped).

The phasic neural responses triggered by the go signal (n = 67) differed in several ways from those evoked by the earlier instruction cues (n = 104; Fig. 6, A and B). Response magnitudes were 42% smaller after the go signal (median: 6.9 spikes/s) compared with instruction-evoked responses (median: 11.9 spikes/s; P < 0.001, Kolmogorov-Smirnov 2-sample test; Fig. 6C). In addition, response onset latencies were delayed 31% later after the go signal (medians: 152 vs. 116 ms; P < 0.001; Fig. 6D), and responses lasted 28% longer than instruction-evoked responses (medians: 118 vs. 92 ms; P < 0.001; Fig. 6E).

Importantly, none of the go-evoked responses encoded information about the expected reward quantity or the physical effort (F < 0.04, P > 0.05, cell-by-cell 2-way ANOVA). In Fig. 7, we show the population-averaged activities of go-responding neurons only (n = 67), constructed separately around the instruction cues (Fig. 7A) and the go signal (Fig. 7B) for the different types of trials related to the nine distinct reward/effort combinations. Whereas the go-responding subpopulation of neurons did not encode incentive values in their responses to the go signal (P > 0.05, 2-way ANOVA combined with the sliding window procedure), the same subset of cells significantly encoded the expected size of reward during the first 400 ms of the presentation of the instruction cue (P < 0.05; Fig. 7A). As described above, 61% of instruction-evoked responses encoded reward and/or effort information in the cell-by-cell analysis (see Neuronal responses to instructions). Thus the go-responding subset of dopamine neurons did encode the reward value associated with each trial, only not in response to the go signal.

The go-related phasic responses distinguished themselves from the instruction-evoked responses (described above and in Pasquereau and Turner 2013) and from previous descriptions of direct sensory-driven dopamine responses (Redgrave and Gurney 2006) by beginning later, lasting longer, and not encoding motivational information. We wondered whether these long-latency signals carried information related to behavior at this stage of the task. In particular, since the go signal acted primarily as a temporal cue, we wondered if go-related responses encoded information about task timing and time-dependent probabilities of occurrence.

**Dopamine neurons encode hazard-related signals.** To examine potential effects of the hazard rate on dopamine activity around the time of the go signal, we performed time-resolved linear regressions with neural activity as the dependent variable. For each recorded cell (n = 141), we tested whether the firing rate was monotonically modulated by the subjective hazard rate estimated previously on the basis of each animal’s performance (Fig. 8A). Dopamine neurons recorded from both monkeys were pooled for this analysis because no differences were found in location, neuronal spike shape, or responsiveness to task events. The activity of a substantial fraction of dopamine neurons (27%) was affected by the hazard rate during the time of the go-evoked phasic response, as evidenced by the fraction of neurons showing significant effects within the 400 ms immediately following the go signal (38/141 neurons with at least one significant bin within the period post-go; P < 0.01, permutation test across trials with different foreperiods). Notably, the activity of 18% (25/141 neurons with at least one significant bin within the period pre-go; P < 0.01, permutation test) of neurons was also modulated by...
hazard rate during the 500 ms immediately preceding presentation of the go signal (Fig. 8A).

A population analysis of hazard-dependent neurons (n = 38, identified by the linear regression described above; Fig. 8B) showed that the magnitude of go-evoked phasic responses decreased as a function of the hazard rate such that responses were maximal when the go signal was least likely to appear (P < 0.001, time-resolved ANOVA within the post-go period) and that responses were attenuated as appearance of the go signal became more and more likely. In addition, a close examination of neural activity revealed that mean tonic firing rates preceding appearance of the go signal decreased as a function of the hazard rate (P < 0.05; within the pre-go period; Fig. 8B). At the end of the shortest foreperiod (i.e., 0.5 s after the target presentation), when the probability of occurrence of the go signal was very low, the population mean firing rate was unchanged relative to the population’s tonic baseline level (P > 0.05, Mann-Whitney U-test; Fig. 8C). The population firing then decremented slowly across the time interval when the go signal could potentially appear until the foreperiod ended with a go signal. Preceding appearance of the go signal, we observed that the longer the foreperiod, the higher the hazard rate and the lower the population firing rate (Fig. 8C). Notably, this alteration in dopamine activity was unlikely to be a correlate of the elapsed time between events because no change from baseline firing rate was detected before the shortest foreperiod (i.e., before the earliest possible time the go signal could appear; P > 0.05, Mann-Whitney U-test; Fig. 8C). The dopamine activity decremented only across the time period when the go signal could potentially appear, thus suggesting an encoding related to appearance of the go signal.

To investigate whether the subjective hazard rate was specifically encoded by a subset of neurons (n = 38; Fig. 8) or whether these response patterns could be a more general property of dopamine neurons, we studied the population activity of all go-responding neurons (n = 67) according to the subjective hazard rate. Without restricting our analysis to those identified by the linear regression (described above; Fig. 8A), we still found that the go-evoked phasic response decreased with increasing hazard rate such that responses were maximal when the probability of go occurrence was lowest (P < 0.001; time-resolved ANOVA). Activity preceding the go signal, however, was not modulated with hazard rate (P > 0.05) in this larger population of neurons.

The effects of hazard rate on firing rate, described above at the population level, were also evident in the activity of individual neurons. In Fig. 9A, we show the activity of an example neuron (from a population of n = 38) with individual trials in the raster diagram sorted according to the duration of that trial’s foreperiod. The cell’s firing rate declined gradually as a function of the hazard rate prior to presentation of the go signal (pre-go: −42%, P < 0.01 linear regression; Fig. 9B). The phasic response that followed presentation of the go signal was also attenuated as a monotonic function of time-dependent probabilities of occurrence (post-go: −64%, P < 0.01, linear regression; Fig. 9C) such that no phasic signal was evident in trials that had the longest foreperiods (bottom raster lines in Fig. 9A; i.e., trials with highest hazard rates).

**Dual coding of hazard-related signals.** Dopamine neurons appeared to encode information about go signal probability in two separate ways: 1) by slowly decreasing firing rates below the baseline firing rate prior to go signal presentation, and 2) by increasing firing rate phasically above baseline following go signal appearance. To characterize this apparent dual coding of hazard-related signals in dopamine neurons, we fitted each neuron’s activity during pre-go and post-go epochs as linear...
functions of the subjective hazard rate (Fig. 9, B and C). The hazard-dependent change in pre- or post-go activity was then measured as the difference in the linear fitted line between the extremes of the hazard function (high hazard rate minus low hazard rate; Fig. 9, D and E). Nearly all significantly modulated neurons had negative changes in firing rate in both pre-go (mean ± SE: −1.7 ± 0.23 spike/s; P < 0.001, t-test; Fig. 9D) and post-go epochs (−6.2 ± 0.74 spike/s; P < 0.001; Fig. 9E), reflecting a general rule: the more likely the go signal, the lower the firing rate. The hazard-related modulation of post-go activity could not be attributed solely to a carryover of pre-go reductions into the post-go epoch because the hazard-dependent changes in post-go activity were substantially larger than those in the pre-go activity (P < 0.001, paired t-test; Fig. 9F). Furthermore, the mean slope of the regression coefficients was significantly more negative for post-go phasic responses than for pre-go activity (P < 0.001, paired t-test; Fig. 9G). These observations are in general agreement with previous reports of a marked asymmetry in the strength of dopamine neuron encoding between increases and decreases relative to baseline firing rate (Fiorillo et al. 2003; Nakahara et al. 2004). This effect is presumably due to the low basal activity of these cells. A population analysis found that hazard-dependent encoding in pre-go and go-evoked phasic responses differed by a factor of 3.7 (Fig. 10). Dopamine neurons therefore encoded time-dependent probabilities of occurrence with a higher gain after the appearance of the go signal than during the variable foreperiod. However, despite this asymmetry, the activity of dopamine neurons accurately encoded the subjective hazard rate both during the foreperiod and in the phasic response to the go signal. Indeed, the fit of the population means from both analysis epochs to the animal’s subjective hazard function was remarkably good (R² = 0.94 and 0.95 for pre-go and post-go epochs, respectively; Fig. 10).

Hazard-related encoding is independent of anticipated effort/reward magnitude. Because considerable evidence implicates dopamine neurons in the motivation-related modulation of motor performance (Bromberg-Martin et al. 2010b; Salamone and Correa 2002; Wise 2004), we investigated whether the hazard-related encoding was modulated by different motivation-related factors.
consistent with the observation noted above (see Neuronal sions with neural activity as the dependent variable (seeMA-
mance. We performed time-resolved multiple linear regres-
presentation encodes information related to motor perfor-
tested whether dopamine activity around the time of go signal
related encoding during pre- and post-go periods was not
influenced by motivational aspects of the task such as the
reward value or the effort (all P > 0.01, shuffle test). This is
consistent with the observation noted above (see Neuronal responses to the go signal) that go-evoked responses did not
encode information about the expected reward quantity or
physical effort.

No direct relationship to task performance. Finally, we
tested whether dopamine activity around the time of go signal
presentation encodes information related to motor performance. We performed time-resolved multiple linear regres-
sions with neural activity as the dependent variable (see MA-
erials and METHODS, Eqs. 3 and 4). This analysis revealed that hazard-related encoding during pre- and post-go periods was not influenced by motivational aspects of the task such as the reward value or the effort (all P > 0.01, shuffle test). This is consistent with the observation noted above (see Neuronal responses to the go signal) that go-evoked responses did not encode information about the expected reward quantity or physical effort.

Notably, we found two distinct features of dopamine neuron activity that correlated with this hazard-dependent behavior. First, as time passed during the foreperiod, the activity of dopamine neurons tracked the subjective hazard rate via a slow decrease in firing rate (Figs. 9 and 10). Second, when the go signal appeared, the magnitude of the evoked phasic response encoded the degree of unpredictability of that event. These patterns of dopamine activity were similar to the “prediction errors” formalized in reinforcement learning theory as the difference between reality and prediction (Bayer and Glimcher 2005; Schultz et al. 1997; Sutton and Barto 1998), although in the present study, the encoding described around the go signal was not modulated in concert with trial-by-trial variations in motivation-related aspects of the task (Fig. 7B) and did not resemble a salience-like signal. These observations are discussed in more detail below. Furthermore, we confirmed previous reports (Schultz et al. 1997) that dopamine neuron activity is not related directly to movement (Fig. 4) or to measures of motor performance (i.e., kinematics).

Dopamine has long been implicated in learning and the control of incentive motivation (Berridge and Robinson 1998; Bromberg-Martin et al. 2010b; Salamone and Correa 2002; Schultz 2006; Wise 2004). More specifically, the spiking activity of dopamine neurons is thought to play a crucial role in the acquisition of cue-reward associations and in regulating motivated, reward-seeking brain states. Dopamine neurons respond with a phasic burst of activity to many types of events including primary rewards, sudden novel stimuli, and reward-associated stimuli (Ljungberg et al. 1992; Romo and Schultz 1990; Schultz and Romo 1990). Many studies have focused on the responses of dopamine neuron to reward-related events. Indeed, there is considerable evidence (Fiorillo et al. 2003, 2008; Joshua et al. 2008; Kobayashi and Schultz 2008; Morris

**DISCUSSION**

The present study revealed that dopamine neurons dynami-
cally modulate their activity with the passage of time in a way
that reflects hazard-related information. This finding suggests a
role for dopamine neurons as a teaching signal for estimating and updating time-dependent probabilities necessary for be-
havioral anticipation and response preparation.

To study time-dependent processing, we used a variable foreperiod schedule that allows monkeys to probabilistically anticipate the arrival of the go signal according to a hazard rate function based on prior experience (Gibbon et al. 1997; Jans-

Fig. 8. Coding of time-dependent probabilities in dopamine activity. A: the fraction of neurons that significantly modulated their activity according to the hazard rate around the time of the go signal (P < 0.01, sliding window linear regression). B and C show the population-averaged activities of dopamine neurons that showed significant hazard-dependent phasic responses. B: spike density functions aligned on go signal presentation were grouped according to the probability that the movement trigger was about to occur (low, yellow; medium, red; high, black). The gray frame indicates time bins where significant hazard encoding was detected (P < 0.05, ANOVA). C: spike density functions were then aligned on the time of the earliest possible go signal (i.e., 500 ms after target presentation). The width of the spike density function line indicates the population SE. The mean tangential velocities of the hand for each set of trials are plotted at bottom. Vertical black lines indicate the mean time of presentation of the go signal for each class of hazard rates.
predicting reward value and 2) predicting reward timing. The temporal precision of encoding has been studied primarily using unexpected changes in reward timing (Fiorillo et al. 2008; Hollerman and Schultz 1998). When a reward is expected but does not occur, then the dopamine activity is suppressed shortly (~100 ms) after the time that reward is usually delivered, even in the absence of any external cue (Hollerman and Schultz 1998). Reward prediction errors are therefore internally timed and make the dopamine signal a putative teaching signal for time-sensitive reward predictions.

Our results show that dopamine neurons can encode errors in predicting the timing of movement trigger events independent of the anticipated magnitude of a future reward. The same subset of neurons did encode reward information when it was first signaled by the instruction cue (Fig. 7). It is likely that the go-evoked dopamine responses were insensitive to reward value because the go cue provided no additional information about reward magnitude than was already indicated a few seconds earlier by the instruction period (Fig. 7A). Our multi-step task allowed us to observe hazard-related encoding that was unmodulated by trial-by-trial variations in an animal’s motivational state (effort/reward trade-off).

Previous studies have shown that dopamine neurons can decrease their tonic spiking activity before the occurrence of variably timed reward-related events (e.g., reward, trial start cue, reward-predictive cue), suggesting the presence of an internal mechanism that computes continuous temporal predictions during the delay period (Bromberg-Martin et al. 2010a; Fiorillo et al. 2008; Nomoto et al. 2010). Our data are consistent with that proposal, demonstrating that dopamine neurons decrease their activity in prediction of the go signal (Fig. 8C). This change of activity accumulated in size with each moment of the foreperiod when the go signal could have occurred but did not, and the sign of the change (a decrease) was the same as would occur if reward itself were omitted or delayed (Schultz 2006). In this sense, the pre-go dopamine signal identified here could be considered an encoding of a succession of negative prediction errors. When the go signal occurred, we then observed a phasic response, the magnitude of which reflected how unexpected the go signal was, consistent with an encoding of positive prediction error. This last observation clarifies why phasic dopamine responses to a go signal were reported by Schultz et al. (1993) to be present only after a variable (i.e., temporally uncertain) foreperiod and absent

Fig. 9. Features and prevalence of hazard-dependent dopamine signals. A: for one example dopamine neuron, spike density function and rasters were constructed around the time of go signal occurrence (time 0). Individual trials in the raster diagram were sorted from top to bottom according to the duration of the foreperiod. Purple tick marks indicate time of target presentation, marking the start of the foreperiod. Average firing rates were measured in 2 time windows to detect hazard-dependent effects on firing rates in the pre-go (B: dark gray) and post-go (C: light gray) time periods. Low and high hazard firing rates obtained by linear fitting are indicated by green circles. D–G show population-based analysis of hazard-dependent signals. Histograms of changes in firing rate for pre-go (D) and post-go (E) periods. ***P < 0.001, t-test. F: fitted low hazard (x-axis) and high hazard (y-axis) activities for each neuron modulated by the probability that the go signal occurred. G: comparison of regression coefficients obtained from linear fittings between pre-go and post-go signals (e.g., reflecting the slopes of the green lines shown in B and C, respectively).

Fig. 10. Dual coding of hazard rate by dopamine signals. Population-averaged spike rates (means ± SE) from pre-go (black) and post-go (gray) analysis windows plotted relative to a trial’s subjective hazard rates. Before averaging, each neuron’s activity was normalized by subtracting the baseline firing rate. Statistics on the goodness of linear fit are furnished for the 2 analysis windows ($R^2$ and P value).
when the foreperiod was of fixed duration (i.e., temporally certain).

Similar hazard-dependent signals were observed in response to primary rewards, when the reward was delivered after a variable delay (Fiorillo et al. 2008). As with our results, phasic dopamine responses in that study were larger after shorter delay intervals. Bromberg-Martin et al. (2010a), however, observed an inverted relationship in dopamine responses to a trial start stimulus after variable intertrial intervals, with stronger phasic signals after longer delays. Distinct from those previous results, the activity we describe in this work was in response to a sensory cue that provided temporal information about the required time of movement execution. Although shorter foreperiods were globally associated with shorter delays to obtain the reward in our paradigm, we believe that the modulation of go-evoked dopamine responses described here is unlikely to be explained by a simple discounting of a reward-related response by the cost of time. First, the only form of temporal discount encoding that has been described for dopamine neurons has been a prospective type of encoding (i.e., encoding the expected future delay until reward delivery) (Fiorillo et al. 2008; Kobayashi and Schultz 2008) rather than a retrospective encoding, which would have to be the case here (i.e., encoding the delay that has already elapsed). Second, past reports of temporal discounting in dopamine neurons used delays that varied substantially in length (Fiorillo et al. 2008; Kobayashi and Schultz 2008), whereas in our paradigm the foreperiod varied on a much shorter subssecond time scale with the explicit intention of minimizing the influences of the cost of time. Third, we found no evidence that go-evoked responses were sensitive to trial-by-trial variations in expected value (anticipated reward quantity and effort level in Fig. 7B), suggesting a relative independence of this response from the future reward.

In addition, our results confirm the classical view that dopamine neurons do not signal information about somatic movement per se (DeLong et al. 1983; Schultz et al. 1983). Even though a large fraction of dopamine neurons changed their firing rate around the time of movement onset, trial-by-trial analysis of the timing of those responses showed that they were linked temporally to the appearance of the sensory cue in a large majority of the cases (Fig. 4). Furthermore, the spike rates of dopamine neurons during the perimovement period did not correlate with any of a wide variety of measures of motor performance (see RESULTS, No direct relationship to task performance). Although consistent with the classical view (DeLong et al. 1983; Schultz et al. 1983), our observations are at variance with more recent studies that have suggested a direct link between tonic (Romo and Schultz 1990) or phasic (Bouret et al. 2012; Jin and Costa 2010; Kiyatkin 1988; Puryear et al. 2010) dopamine signals and motor performance.

The role of phasic dopamine activity in processing nonrewards, events remains a topic of study (Bromberg-Martin et al. 2010b). Some studies have reported that dopamine neurons also encode a form of salience for the purpose of shifting attention to unpredicted events or increasing the general motivation to detect and respond to situations of high importance (Berridge and Robinson 1998; Bromberg-Martin et al. 2010b; Horvitz 2000). These salience-encoding neurons are characterized by bursts of activity evoked at short latency to both rewarding and aversive events (Fiorillo et al. 2013; Joshua et al. 2008; Matsumoto and Hikosaka 2009; Nomoto et al. 2010). Although we did not use aversive events in this study, the go-related activity that we observed appears unlikely to be a salience-like signal as based on four points: First, unlike the short latencies described for salience-like encoding (Fiorillo et al. 2013), responses to the go signal typically occurred at long latencies (>100-ms latency; Fig. 6). Second, dopamine activity was dually modulated as a function of hazard rate in a way that combined decreases in anticipation of the go signal with increases when the go signal occurred (Fig. 9). This suggests that, unlike the encoding of salience, positive and negative valences were encoded by opposing changes in firing rate. Third, the salience-like signal is reported to be a good predictor of the speed of subsequent orienting and approach behaviors (Bromberg-Martin et al. 2010b; Satoh et al. 2003). We found no such relationship between the go-related activity of dopamine neurons and parameters of the subsequent reaching movement (see RESULTS, No direct relationship to task performance). Fourth, whereas salience-like signals were recently described to be more common in the dorsolateral portion of the SNc (Matsumoto and Hikosaka 2009), we found that corresponding neurons were not restricted to a specific subregion of the SNc (Fig. 5). Thus it appears unlikely that the hazard-related go signal activity described here is simply a form of salience-related activity.

Our findings point to a role for dopamine neuron signaling that supplements the traditional perspective of dopaminergic involvement in incentive motivation and salience signaling (Berridge and Robinson 1998; Montague et al. 1996; Schultz et al. 1997; Wise 2004). Dopamine activity that encodes errors in predicting motor-phase timing could be used in recipient brain areas to estimate and update the subjective hazard rate to prepare and optimize adaptive behavioral responses. Learning the relative time-dependent probabilities of impending actions is exploited to improve performance in at least two ways. First, anticipation of an event can be used to enhance attention, representation, and perception at task-appropriate moments as has been shown for visual (Correa et al. 2005; Ghose and Maunsell 2002) and auditory (Jaramillo and Zador 2011) cortical areas. This phenomenon has been recently suggested by one study that reported time-dependent changes in dopamine activity correlated with attentional levels (Totah et al. 2013). Second, anticipatory knowledge of an event’s hazard function enables the nervous system to prepare motor output prior to the event (Bestmann et al. 2008). Sensory cues that predict action allow a gradual buildup of preparatory activity in premotor and motor cortex prior to action (Crammond and Kalaska 2000; Mauritz and Wise 1986; Renoult et al. 2006; Roux et al. 2006; Salinas and Romo 1998; Tanji and Evarts 1976). Critically, a significant role for dopamine in temporal aspects of both perception and motor control has been suggested from observations that patients with Parkinson’s disease, a disorder associated with dopamine depletion, are impaired on a variety of tasks that require time-dependent estimation or movement planning (Benecke et al. 1986; Freeman et al. 1993; Nagasaki et al. 1978; Pastor et al. 1992). In addition, dopaminergic medication improves motor (O’Boyle et al. 1996) and perceptual (Malapani et al. 1998) performance, and pharmacological studies have shown that dopamine drugs (agonists and antagonists) can speed or slow the “internal clock” (i.e., the internal representation of time) in neurologi-
cally normal subjects (Cheng et al. 2007; Drew et al. 2003; Matell et al. 2006; Meck et al. 2012; Rammsayer and Classen 1997; Rammsayer 1993). Overall, the hazard-related activity we describe may be a neurophysiological correlate of the roles hypothesized for dopamine in temporal processing and related roles in regulating sensory anticipation and motor preparation.

REFERENCES


Luce RD. Response times: their role in inferring elementary mental organization. New York: Oxford University Press, 1986.


DOPAMINE ENCODING OF HAZARD-RELATED ERRORS


