

Activation of Dorsal Raphe Serotonin Neurons Underlies Waiting for Delayed Rewards

Katsuhiko Miyazaki,¹ Kayoko W. Miyazaki,¹ and Kenji Doya^{1,2}

¹Neural Computation Unit, Okinawa Institute of Science and Technology, Okinawa 904-0412, Japan, and ²Computational Neuroscience Laboratories, Advanced Telecommunications Research Institute International, Kyoto 619-0288, Japan

The serotonergic system plays a key role in the control of impulsive behaviors. Forebrain serotonin depletion leads to premature actions and steepens discounting of delayed rewards. However, there has been no direct evidence for serotonin neuron activity in relation to actions for delayed rewards. Here we show that serotonin neurons increase their tonic firing while rats wait for food and water rewards and conditioned reinforcement tones. The rate of tonic firing during the delay period was significantly higher for rewards than for tones, for which rats could not wait as long. When the delay was extended, tonic firing persisted until reward or tone delivery. When rats gave up waiting because of extended delay or reward omission, serotonin neuron firing dropped preceding the exit from reward sites. Serotonin neurons did not show significant response when an expected reward was omitted, which was predicted by the theory that serotonin signals negative reward prediction errors. These results suggest that increased serotonin neuron firing facilitates a rat's waiting behavior in prospect of forthcoming rewards and that higher serotonin activation enables longer waiting.

Introduction

Serotonin [5-hydroxytryptamine (5-HT)] has long been implicated in a variety of motor, cognitive, and affective functions (Jacobs and Azmitia, 1992; Adell et al., 2002; Hensler, 2006), such as locomotion, sleep–wake cycles, and mood disorders. A wide range of literature shows that a reduced level of 5-HT in the CNS promotes impulsive behaviors (Evenden, 1999; Cardinal, 2006), including motor impulsivity (failure to suppress inappropriate action) and choice impulsivity (choice of small immediate rewards over larger delayed rewards). However, recent studies that assessed the effects of serotonergic manipulation on impulsivity reported mixed results. Administration of serotonin selective reuptake inhibitors, which increase the extracellular serotonin concentration, increased selections of a large delayed reward over small immediate rewards (Bizot et al., 1988, 1999), meaning a decrease in impulsive choice. Although forebrain serotonin depletion caused impulsive choice behavior in some studies (Wogar et al., 1993; Mobini et al., 2000; Denk et al., 2005), other studies demonstrated no effect (Winstanley et al., 2003), only a transient effect (Bizot et al., 1999), or an effect on motor impulsivity but not on choice impulsivity (Harrison et al., 1997; Winstanley et al., 2004). Recent studies showed that 5-HT_{2A} receptor antagonists caused a reduction in motor impulsivity, whereas 5-HT_{2C} receptor antagonists caused an increase (Higgins et al., 2003; Fletcher et al., 2007). A possible reason for such inconsistent results in

these manipulation studies is the complex compensatory mechanisms regulating 5-HT neuron firing, presynaptic release, and postsynaptic receptor expression. Measurement of 5-HT neuron activity related to impulsive behavior is crucial in clarifying the role of 5-HT in impulsivity.

Previous recording studies of the dorsal raphe nucleus (DRN), a center of forebrain 5-HT projection (Jacobs and Azmitia, 1992), revealed activation of putative 5-HT neurons correlated to the level of behavioral arousal (Jacobs and Fornal, 1999), salient sensory stimuli (Heym et al., 1982; Waterhouse et al., 2004; Ranade and Mainen, 2009), and rhythmic motor outputs (Fornal et al., 1996). Recent recording studies in monkey (Nakamura et al., 2008) and rat (Ranade and Mainen, 2009) demonstrated that DRN neurons display firing correlated with diverse behavioral events, including rewards and conditioned cues, but no study has established a link between 5-HT neuron activity in the DRN and impulsive behaviors. In the present study, we sought to provide direct evidence as to whether activation of the forebrain 5-HT system is involved in the regulation of impulsive behaviors. To achieve this, we recorded the firing of 5-HT neurons in the DRN while rats performed a task requiring waiting for rewards and a conditioned reinforcement tone. Using *in vivo* microdialysis measurements, we recently discovered that 5-HT efflux in the rat DRN increases when the animals perform a task that requires waiting for delayed reward (Miyazaki et al., 2010). However, a limitation of microdialysis measurements is that the temporal resolution is in the order of minutes. By recording from single 5-HT neurons, we demonstrate that the firing rates of these neurons increase specifically during periods in which the animals wait for a forthcoming reward. We also tested the hypothesis that serotonin neurons encode an aversive prediction error signal (Daw et al., 2002) by analyzing the 5-HT neuron responses to reward omission but found no significant responses predicted by the hypothesis.

Received July 12, 2010; revised Oct. 16, 2010; accepted Oct. 25, 2010.

We thank M. Ito for help in logistic regression analysis for serotonin neuron classification. We thank K. Arakaki for her very kind and excellent animal care.

This article is freely available online through the *J Neurosci* Open Choice option.

Correspondence should be addressed to Katsuhiko Miyazaki or Kenji Doya, 1919-1 Tancha, Onna, Okinawa 904-0412, Japan. E-mail: miyazaki@oist.jp, doya@oist.jp.

DOI:10.1523/JNEUROSCI.3714-10.2011

Copyright © 2011 the authors 0270-6474/11/310469-11\$15.00/0

Materials and Methods

Subjects. All experimental procedures were performed in accordance with guidelines determined by the Okinawa Institute of Science and Technology Animal Experiment Committee. Three male Long-Evans rats weighing 300–340 g at the start of the behavioral training sessions were each housed one per cage. The housing area was temperature controlled at 24°C and maintained on a 12 h light/dark cycle (lights on from 8:00 A.M. to 8:00 P.M.). All training and test sessions were conducted during the light period for 5–6 d/week. The rats were deprived of food and water in the home cage and received their daily food and water reward only during the experimental session (~10–15 g/d and 15–20 ml/d, respectively).

Behavioral apparatus and training. A free operant task that we designate a sequential food–water navigation task was used. Rats were individually trained and tested in a cylindrical apparatus 1.5 m in diameter with a 45-cm-high wall; three identical-looking cylinders that were used as the tone site, the food site, and the water site were fixed in an isosceles triangle (Fig. 1A). Each cylinder had a rectangular window of 4 cm long, 3 cm across, and 2.3 cm in depth positioned 5 cm above the floor. All of windows had an infrared photo-beam crossing their entrance to detect the rats' nose-poke responses positioned at a depth of 0.5 cm and 1 cm from the bottom. The open-field apparatus was surrounded by a soundproof box (2.5 m square and 2.4 m in height). Four 100 W lamps were set at the four corners of the box and indirectly illuminated the open field. One speaker and one camera were positioned above the open field. The position of the rat was monitored by a video tracking system (CV-2000; Keyence). When the rat poked its nose through the small window of three cylinders, the control infrared photo-beam was interrupted to detect the rats' responses. The tone site nose poke induced an 8 kHz tone (tone 1: 0.4 s, 70 dB) from the speaker. At the food site, a small food pellet (45 mg) was delivered into the window through a food dispenser. At the water site, a spout of water (length, 0.5 cm; duration, 2.5–4 s) was protruded at the front wall of the window positioned 0.5 cm below from the center.

The start of the sequential food–water navigation task was signaled by having the four lamps on, and one lamp on indicated its termination. The behavioral instrumental response in this task is to hold the rat's nose in a fixed posture in either the tone site window waiting for conditioned reinforcement tone or the reward site windows waiting for food or water rewards. This task required rats to make alternate visits and nose pokes to the food site and the water site via the tone site visit and nose poke. Rats initiated a trial by keeping nose poking in a fixed posture to achieve continuous interruption of the photo-beam at the tone site during a delay period until tone 1 was presented, signaling that a reward was available at one of the reward sites. After tone 1 presentation, the rat was required to continue nose poking at one of the reward sites during another delay period until the reward was delivered (Fig. 1B). To continue the task, rats had to alternately visit two reward sites via the tone site. In this task, food is always rewarded first. We named the delay periods that preceded the tone and the rewards (food and water) the tone delay and

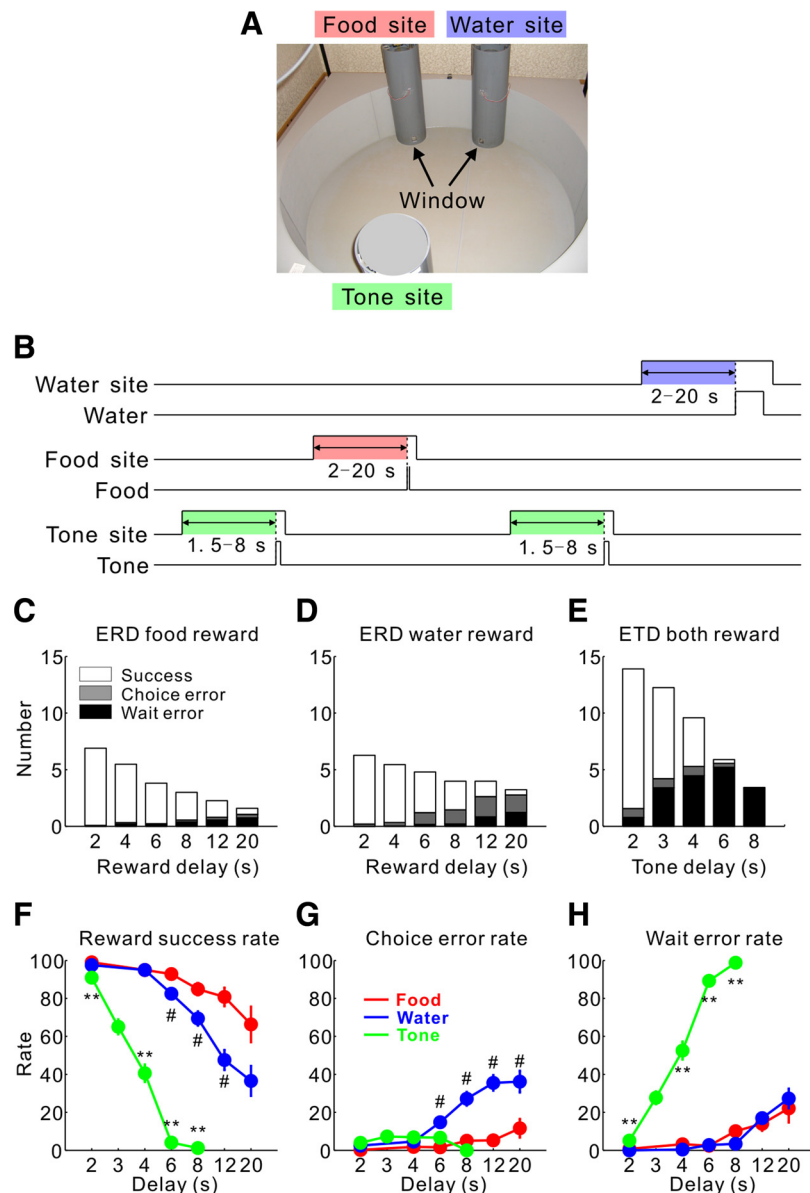


Figure 1. Design of the behavioral task and rats' performance. **A**, Open field, reward cylinders (food site and water site), and a tone cylinder (tone site) for the task. Windows for nose pokes (reward locations) are indicated. The tone cylinder also has a small window at the same position as the reward cylinders. **B**, Schematic of the movements required of rats to receive rewards at the food and water sites. To start the task, rats have to visit, insert, and keep their noses in the tone site until the tone (8 kHz, 0.4 s) is presented (tone delay). Green, red, and blue areas indicate tone delay, food reward delay, and water reward delay, respectively. For details regarding the tone and reward delay periods, see Materials and Methods. **C–E**, Numbers of reward successes, choice errors, and wait errors under the ERD condition (food, **C**; water, **D**) and ETD condition (both rewards, **E**). Under the ERD condition, the number of tone wait errors is negligible, and therefore these are not counted. For similar reasons, the number of reward wait errors under the ETD condition is not counted. The number of successes for food and water are merged. The number of choice errors for food and water and the number of tone wait errors for food and water are also considered together. **F–H**, Reward success rate (**F**), choice error rate (**G**), and wait error rate (**H**) during the ERD condition (red, food; blue, water) and the ETD condition (green, both rewards). ** $p < 0.001$, significant differences compared with the ERD condition (Mann–Whitney U test). # $p < 0.01$, significant differences between food and water under the ERD condition (Mann–Whitney U test). Data from three rats are plotted as mean \pm SEM.

the reward delay, respectively. Depending on whether tone 1 preceded the presentation of food or water, we refer to tone 1 as either the food tone or the water tone, respectively. The food and water delay periods were always set at the same length. During the initial training period, the tone delay was fixed at 0.2 s and the reward delay was set at 0 s. The reward sites did not change until rats could receive a reward at one of the sites.

Two types of error were present in this task: reward choice error and wait error. The reward choice error occurred when the rat nose poked at an incorrect reward site after tone 1. The food choice error entailed the

rat nose poking the water site when the correct nose poke was at the food site, whereas the reverse was the case for the water choice errors. The wait error indicated that the rat failed to wait for the tone or the rewards during delay period by keeping his nose in a fixed posture, which was detected by a brake of continuous interruption of the photo-beam. The reward choice error was signaled by a 500 Hz tone (tone 2: 0.4 s, 72 dB) immediately after a nose poke at the wrong reward site. The occurrence of the four types of wait error (food tone, water tone, food, and water wait error) was not signaled. Rats could start the next trial at any time after reward consumption or after making a reward choice error or a wait error. The rats were trained daily for a period of 2 h. It took 3 weeks or less for the rats to learn the sequential food–water navigation task.

Surgical procedures. After rats had mastered the sequential food–water navigation task, they were anesthetized with Equithesin (3 ml/kg, i.p.) and stereotactically implanted with a stainless steel guide cannula (25 gauge) into the DRN (from bregma: posterior, -7.8 mm; lateral, 0 mm; ventral, -4.5 mm) according to the atlas of Paxinos and Watson (1998). The guide cannula was fixed onto the skull and anchored with dental acrylic and stainless steel screws. The recording electrodes consisted of a bundle of either 8 or 12 Formvar-insulated, 25 μ m, nichrome wires (AM761500; A-M Systems) that were inserted into a 29 gauge stainless steel guide cannula. The cannula/bundle was mounted on a movable microdrive, which enabled adjustment of the electrode position (Miyazaki et al., 1998). The tips of the electrodes were spread out to ~ 0.3 mm, cut with sharp surgical scissors to extend ~ 1 mm beyond the cannula, and electroplated with gold to obtain an impedance of 100–200 k Ω at 1 kHz. This electrode assembly was chronically implanted dorsal to the DRN through the guide cannula at 5.5 mm ventral relative to bregma, flat skull. After surgery, the animals were housed individually and allowed >1 week to recover. After the operation, gentamicin (2 mg/kg, postoperative) was systemically administered once a day for 1 week.

Electrophysiological recordings during four reward condition tests. After recovery from surgery, rats were retrained the sequential food–water navigation task in which tone delay and reward delay were extended up to 2 s. Rats were required to keep their noses in the site window during both delay periods. Once they had been trained to perform this basic task, we introduced four reward condition tests in each recording session: constant delay (CD) condition, extended reward delay (ERD) condition, extended tone delay (ETD) condition, and water omission (WO) condition. Under the CD condition, the delay period at all sites was fixed at 2 s and the task lasted for either 600 or 900 s. Under the ERD condition, the tone delay was fixed at 1.5 s, whereas the reward delay at both sites was increased gradually every 300 s (2, 4, 6, 8, 12, 20 s or 3, 5, 7, 9, 12, 20 s), and the task lasted for a total of 1800 s. Under the ETD condition, the reward delay was fixed at 2 s, but the tone delay was increased gradually every 300 s (2, 3, 4, 6 and 8 s) and the task lasted for 1500 s. Under the WO condition, the delivery of the reward at the water site was omitted, but the reward was still presented at the food site. Under the WO condition, the constant delay condition (450–600 s) was followed by water omission (600–900 s) before returning to the constant delay condition (450–600 s). Under all conditions, there was no presentation of an explicit environmental signal that indicated changes in either the reward conditions or the delay length. We measured the number of acquired rewards (food and water sites), reward choice errors (food and water sites), and wait errors (food, water, and tone sites).

Neural activity was recorded during the four reward condition tests. The wires were screened for activity on a daily basis; if no activity was detected, the rat was removed and the electrode assembly was advanced by 30 or 60 μ m. Otherwise, the active wires were selected for recording, a session was conducted, and the electrode was advanced at the end of each recording day. The placement of electrodes was estimated by depth and confirmed with histology. Neural activity was first passed through a field-effect transistor (unity gain) in a head stage mounted on the recording assembly and then differentially amplified (10,000 \times) and filtered (50 Hz to 3 kHz). During the recording session, the position of the rat was monitored using light-emitting diodes on the head stage at 40 Hz. Spike waveforms and behavioral data (position of rat, tone presentation, and breaks in the photo-beam) were simultaneously acquired by a computer via a Power1401 interface (Cambridge Electronic Design). Neural activ-

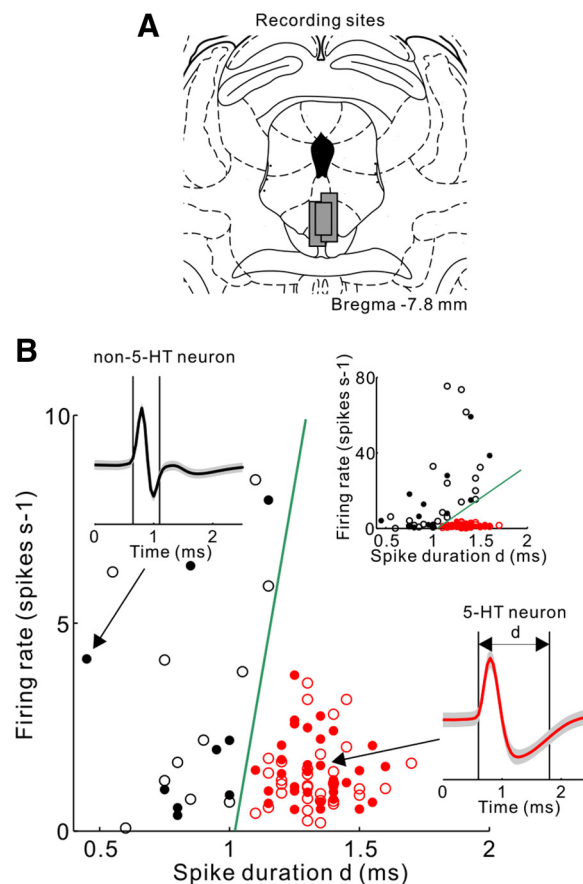


Figure 2. Locations and classification of putative serotonin neurons. **A**, Location of the electrode track in each rat; boxes are reconstructed from the final electrode position. **B**, Results of classification of putative serotonin neurons based on spike duration (d in waveform inset, x-axis, in milliseconds) and baseline firing rate (y -axis, spikes s^{-1}). Filled red circles show neurons with decreased firing rates after systemic injection of 8-OH-DPAT. Filled black circles show putative non-5-HT neurons on the basis of no response to 8-OH-DPAT. The boundary (solid green line) was obtained with a logistic regression model from the results of neural responses to 8-OH-DPAT (see Materials and Methods). Open red and black circles, which were not tested by 8-OH-DPAT, are classified into putative 5-HT and non-5-HT neurons, respectively. Inset classification plot shows the range of firing rates. Inset waveforms show examples of putative serotonin (red) and non-serotonin (black) neurons. Gray shadings indicate SD.

ity data were digitized at 20 kHz. Single units were isolated offline using a template-matching procedure with Spike2 software (Cambridge Electronic Design). The sorted files were then processed and analyzed in Matlab (MathWorks).

Identification of serotonin neurons. In this study, we used the classic criteria for identifying DRN 5-HT neurons, namely, broad spikes, slow regular firing, and suppression by 5-HT_{1A} receptor agonist pattern (Aghajanian et al., 1978; Vandermaelen and Aghajanian, 1983). Baseline spike trains were analyzed to obtain measures of the mean firing rate and spike width. The spike waveforms were averaged over 600 s, and their width was determined as the time between a 5% positive deviation from baseline to the point of half return to baseline after a negative peak (Fig. 2B, inset). The neurons were screened for a wide waveform and baseline firing rate. To identify serotonin neurons, 48 neurons were tested with a 5-HT_{1A} receptor agonist, 8-hydroxy-2-(di- n -propylamino)tetrilin (8-OH-DPAT) (250 μ g/kg, i.p.), after 1-d recording sessions. The 8-OH-DPAT test consisted of ~ 10 min of baseline recording, 8-OH-DPAT injection, and ~ 60 min postinjection recording. The firing rates of putative serotonin neurons decreased, whereas presumed non-serotonin neurons did not respond to 8-OH-DPAT injection. The relationship between spike waveform width and firing rate was plotted. To establish a boundary between serotonin and non-serotonin neurons, a logistic regression model was used. Other recorded neurons that were not tested

with 8-OH-DPAT injection were also plotted by waveform width versus firing rate, and serotonin neurons were identified by the boundary.

Data analysis. To calculate the baseline firing rate of each 5-HT neuron, neural activity was recorded for 60 s before each task condition. Fifteen 2-s periods during which the rats did not nose poke any sites were arbitrarily selected and averaged to obtain a measure of the mean baseline firing rate. The Mann–Whitney *U* test ($p < 0.01$) was used to compare these 15 epochs of baseline firing activity with the activity recorded during the delay period and the periods of reward consumption. We defined the delay period activity under the CD condition as the rate of discharge recorded 0–2 s after the rats inserted their noses into each site. Food and water consumption activity under the CD condition was defined as the firing rate during 1–3 and 0.5–2.5 s after the onset of the reward presentation, respectively. We used the Wilcoxon's signed-rank test to compare differences between the baseline activity and the delay period activity for the population of 5-HT neurons. Perievent time histograms of single neurons and the population activity were generated by averaging activities in a 10 ms bin and was smoothed by a Gaussian filter with $\sigma = 5$ bin (50 ms).

Monitor of rats' mouth movements during the delay periods. We did not monitor rats' mouth movement during tone and reward delay periods in this study. We monitored rats' behavior during neural recording. However, as a result of the low resolution of video, it was difficult to examine whether rats showed preparatory licking or chewing during reward or tone delay periods. To monitor rats' mouth movement, a small camera (CV-022; Keyence) was attached inside the cylinders that monitor rats' mouth movement from an anterior view during tone and reward delay periods. Three rats were newly trained the food–water navigation task, and rats' mouth movement were examined during the CD condition (tone delay, reward delay, and spout duration were 2 s). Numbers of tongue protrusion were counted by video analysis. Spout licking was monitored by a touch sensor connected to the spout (D5C-1DA0; Omron).

Histology. At the end of the last recording session, an overdose of pentobarbital was administered to each rat, and a 10 μ A positive current was passed for 20 s through one or two recording electrodes in the assembly to mark the final recording position. Each rat was perfused with 10% Formalin containing 3% potassium hexacyanoferrate (II), and the brain was removed. Sections (60 μ m) were cut with an electrofreeze microtome and stained with cresyl violet; the position of the electrodes was confirmed as a dot of Prussian blue.

Results

Waiting for reward and conditioned reinforcement tone

We recorded single-unit activity in the DRNs of rats engaged in a sequential food–water navigation task. This task required rats to make alternate visits to two reward sites (a food site and a water site) via a non-rewarding tone site in an open field (1.5 m in diameter) (Fig. 1A). The rats were trained to first nose poke to the tone site, which caused a tone (8 kHz, 0.4 s), and then head to one of the reward sites (Fig. 1B) (see Materials and Methods). We defined two types of delay period: the tone delay period (this period preceded the tone at the tone site) and the reward delay period (this period took place before the receipt of rewards at the food and water sites). The animals were tested under four different conditions: (1) the CD condition, under which the reward and tone were delivered 2 s after the nose poke, (2) the ERD condition, under which the delays before the delivery of food and water were extended up to 20 s, (3) the ETD condition, under which the delay to tone delivery was extended up to 8 s, and (4) the WO condition, under which water was not delivered (Fig. 1B) (see Materials and Methods).

Under the CD condition, rats successfully acquired both food and water rewards in >93% of the trials (supplemental Fig. S1, available at www.jneurosci.org as supplemental material). Under the ERD condition, rats could wait for the delayed rewards for up to 8 or 12 s. As the delay was extended, the number of choice

errors (heading to an incorrect reward site) and wait errors (premature exit from reward sites) increased (Fig. 1C,D). In contrast, under the ETD condition, rats could rarely wait the tone for 6 s or more (Fig. 1E). The success rate for the reward sites dropped more quickly under the ETD condition than under the ERD condition (Fig. 1F).

Under the ERD condition, success rate was significantly higher at the food site than at the water site for 6, 8, and 12 s delays (Mann–Whitney *U* test, $p < 0.01$) (Fig. 1F). Under the ERD condition, the choice error rate was significantly higher at 6, 8, 12, and 20 s delays when the rat was required to go to the water site (Mann–Whitney *U* test, $p < 0.01$) (Fig. 1G). Under the ETD condition, no significant differences in success rates were observed before heading to food and water sites. The tone wait error rate under the ETD condition was significantly higher than the food and water wait error rates under the ERD condition at 2, 4, 6, and 8 s delays (Mann–Whitney *U* test, $p < 0.001$) (Fig. 1H). Under the ERD condition, no significant differences in wait error rates were detected between the food and water sites. Under the ETD condition, no significant differences in tone error rates were observed for food and water. These results show that the rats had more tolerance for delays preceding food and water than that preceding the conditioned tone and that they had motivational bias toward food over water.

5-HT neurons in reward and non-reward delays

We recorded 104 neurons from the DRNs of three rats (Fig. 2A). To identify serotonin neurons, we used three criteria: wide spike waveform, low tonic firing rate, and suppression by 5-HT_{1A} autoreceptor activation (Aghajanian et al., 1978; Vandermaelen and Aghajanian, 1983), although some exceptions have been recently reported, such as slow-firing non-5-HT neurons (Allers and Sharp, 2003), theta rhythmic firing of 5-HT neurons (Kocsis et al., 2006), and expression of 5-HT_{1A} receptors in a subset of non-5-HT neurons (Kirby et al., 2003; Marinelli et al., 2004). We tested the responses of 48 neurons to systemic injection of the 5-HT_{1A} receptor agonist 8-OH-DPAT and observed suppression of the firing rates of 33 neurons (Fig. 2B, filled red dots). We took the spike duration and the firing rate of these neurons as the training data and derived a separation line to distinguish putative 5-HT neurons using logistic regression in the two-dimensional parameter space (Fig. 2B). We considered 63 putative 5-HT neurons with >95% probability for additional analysis.

Figure 3A shows representative activity of a single DRN neuron under the CD condition. The neuron increased its firing rate during both the food delay and the water delay periods. However, this neuron did not show an increased firing rate during the tone delay period. This indicates that the increased firing observed during the reward delay period was not a simple correlate of the motor behavior (namely, the rat keeping its nose in a hole), which is common to both reward and tone delay periods. Instead, this firing depends on the prediction or intention of acquiring the forthcoming reward. As shown in Figure 3B, some other neurons increased their firing rates during both the reward delay and tone delay periods.

Figure 3, C and D, shows the firing rates of 63 putative 5-HT neurons during the reward delay and tone delay periods compared with the baseline firing rates during the rest period. Fifty of the 63 neurons significantly increased their firing rates during at least one of the two reward delay periods (compared with baseline; Mann–Whitney *U* test, $p < 0.01$). Of these 50 neurons, 38 also significantly increased firing during the tone delay period (Mann–Whitney *U* test, $p < 0.01$). No neuron showed a signifi-

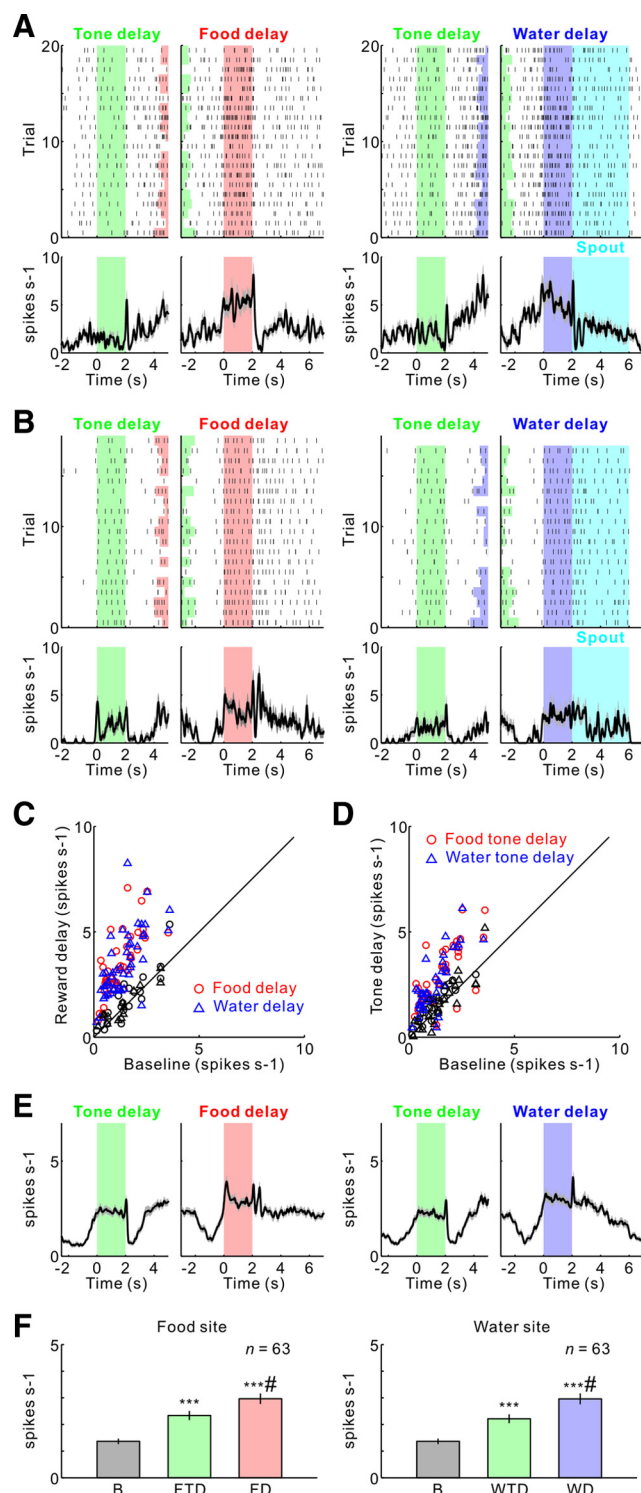


Figure 3. Activity of serotonin neurons during tone delay and reward delay periods. **A, B**, Activity of two example neurons recorded in the dorsal raphe nucleus shown separately for food (left) and water (right) during the CD condition (tone delay and reward delay are 2 s). For each reward, raster plots of neural activity (top) and perievent time histograms smoothed with a Gaussian filter (SD of 50 ms) (bottom) are aligned at the time of tone site entry (left) and at the time of reward site entry (right). Raster plots represent neural activity in the order of occurrence of trials for each reward site from bottom to top. Each dot represents a spike. The tone for food and water site is food tone and water tone, respectively. Reward choice error and reward wait error trials are excluded. Green, red, and blue areas indicate tone delay, food delay, and water delay periods, respectively. Light blue areas indicate water spout presenting period. **C, D**, Comparison between baseline activity and reward delay activity (**C**) and tone delay activity (**D**) in DRN 5-HT neurons ($n = 63$). Open circles indicate food delay activity (**C**) and tone delay

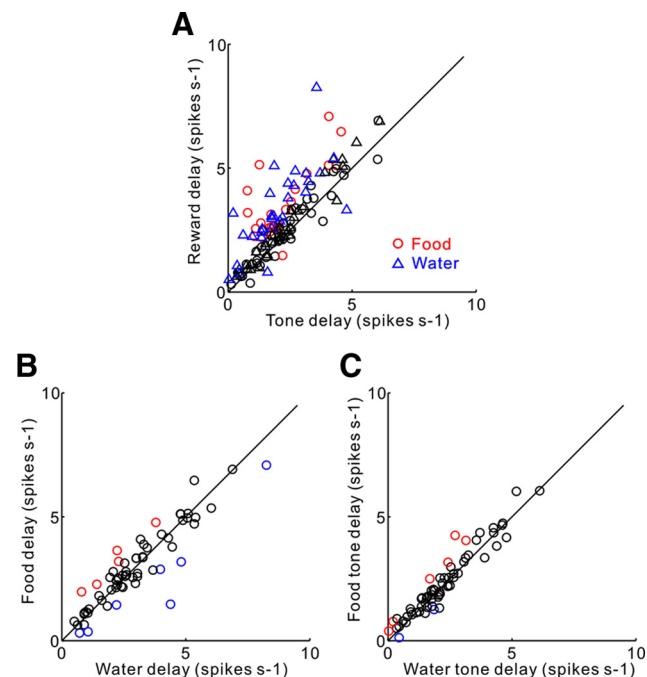


Figure 4. Comparison between tone delay activity and reward delay activity of DRN serotonergic neurons under the constant delay condition. **A**, Comparison between tone delay activity and reward delay activity in DRN 5-HT neurons ($n = 63$). Open circles indicate the comparison between tone delay activity that precedes food and food delay activity. Open triangles indicate the comparison between tone delay activity that precedes water and water delay activity. Red and blue represent activity in which the difference in firing rate was statistically significant (Mann–Whitney U test, $p < 0.01$). **B**, Comparison between food delay activity and water delay activity in DRN 5-HT neurons ($n = 63$). Red and blue indicate neurons with statistically significant higher food and water delay responses, respectively (Mann–Whitney U test, $p < 0.01$). **C**, Comparison between the tone delay activity that precedes food and the tone delay activity that precedes water in DRN 5-HT neurons ($n = 63$). Red and blue indicate neurons with a significantly higher tone delay response preceding food and a higher tone delay response preceding water, respectively (Mann–Whitney U test, $p < 0.01$).

cant increase in firing rate only during the tone delay period. The majority of neurons had higher firing rates during reward delay periods than during tone delay periods (Fig. 4A).

Figure 3E shows the population activity of the 63 putative 5-HT neurons measured under the CD condition. Compared with baseline levels, the activity was significantly higher during both the reward and the tone delay periods. It was also higher during the reward delay than during the tone delay (Wilcoxon's signed-rank test, food delay vs food tone delay, $p < 10^{-6}$; water delay vs water tone delay, $p < 10^{-7}$) (Fig. 3F). We used exactly the same apparatus for the tone and reward sites, so that the sensory input for the rat in the nose-poke hole was the same for the three sites. Place-specific activities despite the same immediate sensory inputs have been reported in hippocampus or related areas (Moser et al., 2008). However, in such place coding activi-

activity preceding food (**D**). Open triangles indicate water delay activity (**C**) and tone delay activity preceding water (**D**). Red and blue indicate statistically significant delay activity (Mann–Whitney U test, $p < 0.01$). Black, No significant difference from baseline. **E**, Averaged activity of the 63 neurons during the CD condition. Compared with baseline levels, the activity was significantly higher during both the reward and the tone delay periods. It was also higher during the reward delay than during the tone delay (Wilcoxon's signed-rank test, food delay vs food tone delay, $p < 10^{-6}$; water delay vs water tone delay, $p < 10^{-7}$) (Fig. 3F). We used exactly the same apparatus for the tone and reward sites, so that the sensory input for the rat in the nose-poke hole was the same for the three sites. Place-specific activities despite the same immediate sensory inputs have been reported in hippocampus or related areas (Moser et al., 2008). However, in such place coding activi-

ties, different neurons have higher activities at different places. The different firing of 5-HT neurons at reward and tone sites should not be such place coding activity because the higher activity was seen predominantly at the reward sites rather than at the tone site (Fig. 4A). No significant differences were observed between the 5-HT neural activities during food and water delay periods or between tone delay periods preceding food and water site visits (Wilcoxon's signed-rank test, food delay and water delay activity, $p = 0.29$; food tone delay and water tone delay activity, $p = 0.20$) (Figs. 3F, 4B,C).

The 5-HT neural activities during both food and water reward consumption periods were significantly higher than the baseline (Wilcoxon's signed-rank test, $p < 10^{-6}$) (supplemental Fig. S2, available at www.jneurosci.org as supplemental material). Although some neurons responded to both during reward delay and reward consumption periods (supplemental Fig. S2B, available at www.jneurosci.org as supplemental material), the population neural activity was significantly higher during reward delay than reward consumption (Wilcoxon's signed-rank test, food delay vs food consumption, $p = 0.0032$; water delay vs water consumption, $p = 0.00054$) (supplemental Fig. S2C, available at www.jneurosci.org as supplemental material). Some reward consumption activity may be related to rhythmic oral–buccal movements, such as licking and chewing (Fornal et al., 1996).

Many 5-HT neurons exhibited a single spike at the onset of the conditioned tone and the sound of food magazine and water spout operation at a latency of 60–90 ms (supplemental Fig. S3, available at www.jneurosci.org as supplemental material), similar to previous reports (Heym et al., 1982; Waterhouse et al., 2004; Ranade and Mainen, 2009).

In contrast to 5-HT neurons, non-5-HT neurons showed diverse response profiles (supplemental Fig. S4A,B, available at www.jneurosci.org as supplemental material). As a population, non-5-HT neurons did not show significant firing rate increase from baseline in response to any behavioral events (food tone delay, water tone delay, food delay, water delay, and food consumption) except during water consumption period activity (supplemental Fig. S4C,D, available at www.jneurosci.org as supplemental material).

Recent research showed that DRN neurons with high firing rates are GABA neurons, and slow-firing DRN neurons include both 5-HT and non-5-HT neurons (Allers and Sharp, 2003). Thus, we examined putative non-5-HT neurons separately based on their baseline firing rates (higher or lower than 5 Hz) (supplemental Fig. S5, available at www.jneurosci.org as supplemental material). Low baseline firing non-5-HT neurons ($n = 15$) showed significantly stronger reward delay and tone delay activities (Wilcoxon's signed-rank test, food delay vs baseline, $p = 0.0026$; water delay vs baseline, $p = 0.0034$; food tone delay vs baseline, $p = 0.002$; water tone delay vs baseline, $p = 0.00031$) (supplemental Fig. S5A, left and middle, available at www.jneurosci.org as supplemental material). However, in contrast to the putative 5-HT neurons, there was no significant difference between reward delay and tone delay activities (Wilcoxon's signed-rank test, food delay vs food tone delay, $p = 0.52$; water delay vs water tone delay, $p = 0.49$) (supplemental Fig. S5A, right, available at www.jneurosci.org as supplemental material). No significant differences were observed between the neural activities during food and water delay periods or between the tone delay periods preceding food and water site visits (Wilcoxon's signed-rank test, food delay vs water delay, $p = 0.85$; food tone delay vs water tone delay, $p = 0.93$) (supplemental Fig. S5C, left, available at www.jneurosci.org as supplemental material).

High baseline firing non-5-HT neurons ($n = 26$) showed no significant firing rate increase during reward delay or tone delay (Wilcoxon's signed-rank test, food delay vs baseline, $p = 0.26$; water delay vs baseline, $p = 0.33$; food tone delay vs baseline, $p = 0.28$; water tone delay vs baseline, $p = 0.79$) (supplemental Fig. S5B, left and middle, available at www.jneurosci.org as supplemental material). There was no significant difference between reward delay and tone delay activities (Wilcoxon's signed-rank test, food delay vs food tone delay, $p = 0.45$; water delay vs water tone delay, $p = 0.42$) (supplemental Fig. S5B, right, available at www.jneurosci.org as supplemental material). In contrast to the putative 5-HT and low firing non-5-HT neural activities, food delay activity of high firing non-5-HT neurons was significantly stronger than water delay activity (Wilcoxon's signed-rank test, food delay vs water delay, $p = 0.002$; food tone delay vs water tone delay, $p = 0.073$) (supplemental Fig. S5C, right, available at www.jneurosci.org as supplemental material).

Effect of extended reward and tone delay

When the delay periods were extended, the higher 5-HT neural activity ($n = 47$) was prolonged until the end of the delay periods. The neuron shown in Figure 5, A and B, extended its higher firing rate during both reward and tone delay periods under the ERD and ETD conditions. The population neural activity during the successful trials was higher than the baseline in all reward delays (Wilcoxon's signed-rank test, 2–20 s, $p < 0.001$) (supplemental Fig. S6, available at www.jneurosci.org as supplemental material). The average firing rates during both food and water delays were not affected by the delay length (one-way ANOVA, food site, $p = 0.63$; water site, $p = 0.44$) (Fig. 5C). However, during the longer delays (6–20 s), the firing rate in the last 2 s showed a slight but significant decrease compared with that in the first 2 s (Wilcoxon's signed-rank, food site, $p = 0.0033$; water site, $p = 0.0023$) (Fig. 5D).

To examine the relationship between the 5-HT activity and tolerance for extended delays, we compared the neural activities during successful and wait error trials under the ERD condition (Fig. 6A,B). Although increased 5-HT neural activity was sustained until the end of both food and water delay periods in successful trials, it decayed close to the baseline in wait error trials (Fig. 6A). The average firing rate during the last 2 s of the waiting period was significantly lower in wait error trials than in successful trials (Wilcoxon's signed-rank test, $p < 0.001$) (Fig. 6B). We also found lower firing rates in choice error trials (Fig. 6C,D). In the choice error trials in which the rat mistakenly went to the food site (i.e., a water choice error), the 5-HT neural activity sharply dropped at the presentation of the error tone after the nose poke (Fig. 6C, left). Interestingly, when the rat mistakenly headed to the water site (i.e., a food choice error), the 5-HT neural activity was lower even before presentation of the error tone; this may reflect lower motivation or confidence of the rat (Fig. 6C, right). During the first 2 s of both the food and water delay periods, the average firing rates were significantly lower in the choice error trials than in successful trials (Wilcoxon's signed-rank test, $p < 0.0001$) (Fig. 6D).

Effect of water reward omission

Under the WO condition, the rat had to nose poke to the water site but was not required to wait there before heading to the tone site. In the session shown in Figure 7A, during the first trial of the WO condition that followed the CD condition, the rat waited for the omitted reward for ~9 s before it moved to the tone site. As represented by the red marks in Figure 7A, in subsequent trials,

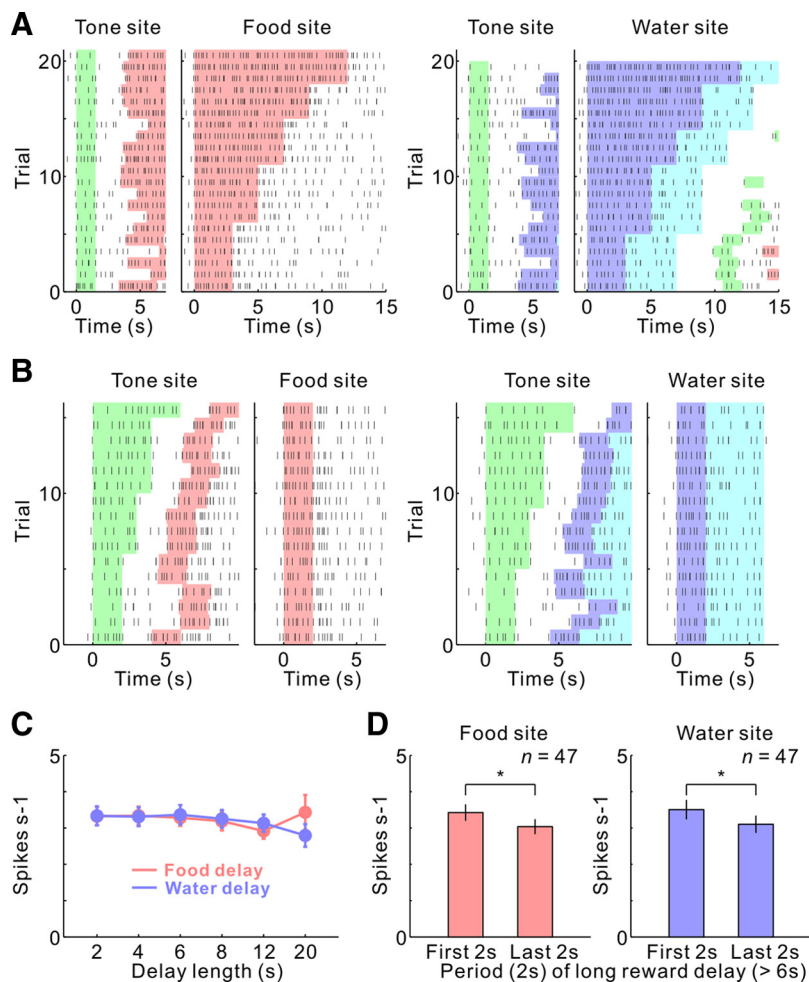


Figure 5. Modulation of serotonin neural activity to extended reward delay and extended tone delay. **A**, Example of a 5-HT neuron during the extended reward delay test (the reward delay was gradually increased every 5 min; tone delay, 1.5 s). **B**, Activity of an example neuron (same neuron as **A**) during the extended tone delay test (the tone delay was gradually increased every 5 min; reward delay, 2 s). **C**, Average firing rate of serotonin neurons during reward delay periods of different lengths (mean ± SEM). The average firing rates of 3, 5, 7, and 9 s reward delay periods are merged with those of 2, 4, 6, and 8 s reward delay periods, respectively. Red indicates the food delay firing rate. Blue indicates the water delay firing rate. Food site: 2 s delay ($n = 46$), 4 s delay ($n = 45$), 6 s delay ($n = 47$), 8 s delay ($n = 43$), 12 s delay ($n = 37$), and 20 s delay ($n = 15$). Water site: 2 s delay ($n = 46$), 4 s delay ($n = 45$), 6 s delay ($n = 47$), 8 s delay ($n = 46$), 12 s delay ($n = 39$), and 20 s delay ($n = 13$). **D**, Average firing rate during the first 2 s and the last 2 s of the long reward delay period (6–20 s) (mean ± SEM). * $p < 0.01$, Wilcoxon's signed-rank test.

the rat spent a shorter period of time waiting for the omitted reward, and the abandonment of waiting was preceded by a cessation of 5-HT neuron firing. This effect can be appreciated more readily by examining the raster plot and the perievent time histogram aligned at the time of exit from the water site during the water omission condition (Fig. 7A, right).

For the first 2 s after entry to the water site, the population activity of 5-HT neurons ($n = 24$) in the WO trials (when rats waited for longer than 3 s) was similar to the activity observed during the CD trials; however, during the performance of WO trials, this activity declined toward baseline levels in the last 2 s before exit from the site (Fig. 7B). The average firing rate during the initial 2 s after the water site entry in the WO condition was not significantly different from that in the CD condition (Wilcoxon's signed-rank test, $p = 0.13$). The average firing rate during the last 2 s before site exit in the WO condition was significantly lower than those during the initial 2 s after the site entry in both CD and WO conditions (Wilcoxon's signed-rank test, $p < 0.0001$) (Fig. 7C). This reduction in the sustained activity of the

5-HT neurons time locked to the rats' exit from the reward site was similar to that observed in the wait error trials under the ERD condition (Fig. 6A,B). These results demonstrate that an increase in 5-HT neural activity is predictive of whether the animal continues to wait for a possible reward.

In the opponent theory of serotonin and dopamine, central serotonin has been considered to be the crucial substrate for the aversive motivational system (Deakin and Graeff, 1991; Daw et al., 2002). The opponent theory predicts that unexpected punishment or loss of reward causes a phasic response of serotonin neurons (Daw et al., 2002). Thus, we investigated the responses of 5-HT neurons to the unexpected omission of the water reward. None of the putative 5-HT neurons recorded in the WO condition showed any significant increase in activity around the time of expected reward when the water reward was omitted. Figure 7D shows the population activity of the 24 5-HT neurons around the expected time of water delivery during the WO condition (left: first three trials after switching from CD condition; right: all trials in which rats waited for >3 s). The average firing rates during 1 s before and after the time of expected water delivery were not significantly different for the first three trials after switching to the WO condition (Wilcoxon's signed-rank test, $p = 0.31$) (Fig. 7E, left). When all trials in the WO condition were averaged, a significant decline in the firing rate was observed before and after the time of the expected reward (Wilcoxon's signed-rank test, $p = 0.0055$) (Fig. 7E, right). The opponent theory also predicts suppression of 5-HT neurons in response to an unexpected reward. Thus, we examined the responses of 5-HT neurons to unexpected water spout presentation when reward condition was changed from the WO condition to the CD condition but did not find any suppression of firing (supplemental Fig. S7A,B, available at www.jneurosci.org as supplemental material). The average firing rates during the second after the time of water delivery did not significantly change in the first trial or first three trials after switching from WO condition to CD condition compared with all other trials with a water reward (Wilcoxon's signed-rank test, $p = 0.26$ for first trial and $p = 0.10$ for first three trials) (supplemental Fig. S7C, available at www.jneurosci.org as supplemental material). Non-5-HT neurons ($n = 7$) also did not respond to unexpected water omission (Wilcoxon's signed-rank test, $p = 0.22$ for the first three trials after switching to the WO condition) or to unexpected water spout presentation when reward condition was changed from the WO condition to the CD condition (Wilcoxon's signed-rank test, $p = 0.47$ for first trial and $p = 0.22$ for first three trials).

In choice error trials, the error tone at the entry to a wrong site signaled no reward, which the rat would have expected. Thus, we

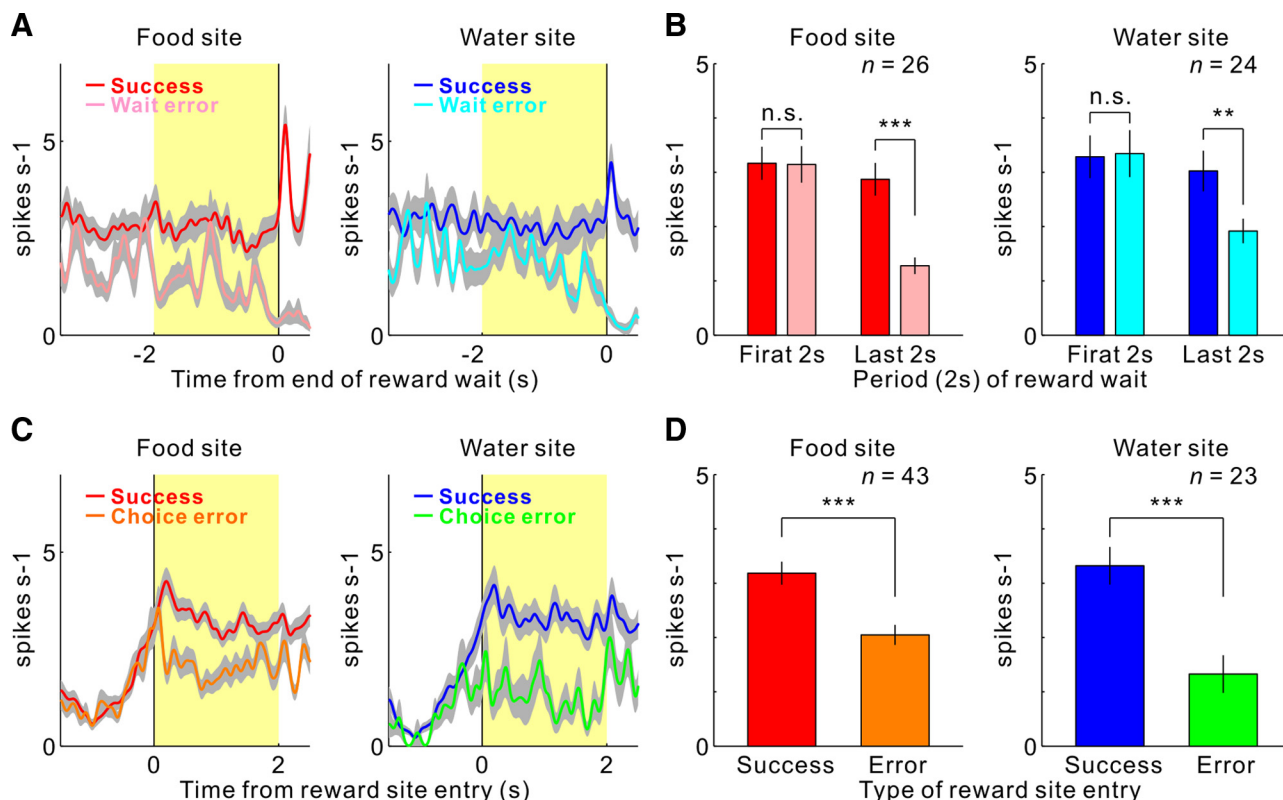


Figure 6. Activity of serotonin neurons during reward wait error and reward choice error in the extended reward delay test. **A**, Population activity aligned to the onset of the reward presentation (red, food; blue, water) and to the reward wait error (pink, food wait error; cyan, water wait error) (left, food site, $n = 26$; right, water site, $n = 24$). Gray shadings represent SEM. Light yellow areas indicate the periods that were used to analyze average firing rate. **B**, Average firing rate during the first and last 2 s of the waiting period after entry into the reward site in a case of successful entry (red, food; blue, water) and in the case of wait error entry (pink, food wait error; cyan, water wait error) (left, food site, $n = 26$; right, water site, $n = 24$; \pm SEM). **C**, Population activity aligned to reward site success entry (red, food; blue, water) and to reward choice error entry (orange, water choice error; green, food choice error) (left, food site, $n = 43$; right, water site, $n = 23$). **D**, Average firing rate during 2 s after reward site entry in a case of successful entry (red, food; blue, water) and in a case of choice error entry (orange, water choice error; green, food choice error) (left, food site, $n = 43$; right, water site, $n = 23$; \pm SEM). ** $p < 0.01$, *** $p < 0.0001$, Wilcoxon's signed-rank test. n.s., Not significant.

compared the DRN neural activity during the first one second from reward site entry between successful and choice error trials in the extended reward delay condition. In the 5-HT neurons, the activity in choice error trials was significantly lower than in successful trials (Wilcoxon's signed-rank test, $p < 0.0001$) (Fig. 6C). In the non-5-HT neurons, the average firing rates were significantly lower in the water choice error trials than in successful trials (Wilcoxon's signed-rank test, $p = 0.0072$; $n = 16$), although they were not significantly different between food choice error and successful trials (Wilcoxon's signed-rank test, $p = 0.83$; $n = 11$). These results also do not support encoding of negative reward prediction error by DRN neurons.

Discussion

We monitored the activity of DRN neurons while rats performed a food–water navigation task, in which we manipulated the length of delays to both food and water rewards and to a conditioned tone stimulus. 5-HT neurons were identified by their spike shape, firing rate, and response to 5-HT_{1A} receptor agonist using a logistic regression model. The principal observation of the present study is that the 5-HT neurons showed increased firing rates while rats waited for the primary rewards and the conditioned reinforcement tone. Furthermore, we found that the sustained 5-HT neural activity during the reward delay period was significantly higher than that during the tone delay period, meaning that the activity was not simply attributable to the nose-poking behavior, which was required at both reward and tone

sites. We also found that increased 5-HT neural activity ceased before rats gave up waiting for possible future rewards, both in wait error trials under extended delay conditions and during adaptively truncated waiting in water omission trials. We found no increase in 5-HT neuron firing at the time of expected reward delivery in water omission trials. These results provide the first pieces of direct evidence that serotonin neural activity in the DRN increases during waiting for delayed rewards. The present results are consistent with our previous finding in microdialysis measurement experiments that 5-HT efflux significantly increased while rats performed a very similar task in the delayed reward condition but not in the reward omission condition compared with the immediate reward condition (Miyazaki et al., 2010). The present study revealed that waiting behavior for a delayed reward is indeed the critical behavioral event during which 5-HT neuronal firing is increased.

Based on a wide range of literature showing that reduced levels of 5-HT in the CNS promote impulsive behavior (Evenden, 1999; Cardinal, 2006) and our experience of autonomous adaptive agents (Doya and Uchibe, 2005; Doya, 2007), we previously proposed that 5-HT controls the timescale of reward evaluation (Doya, 2002), which is an important parameter that affects the character of learned behaviors. The present results support our hypothesis that a higher level of central serotonin activity extends the temporal horizon of future reward evaluation and promotes behaviors aiming for delayed rewards (Doya, 2002). To date, the

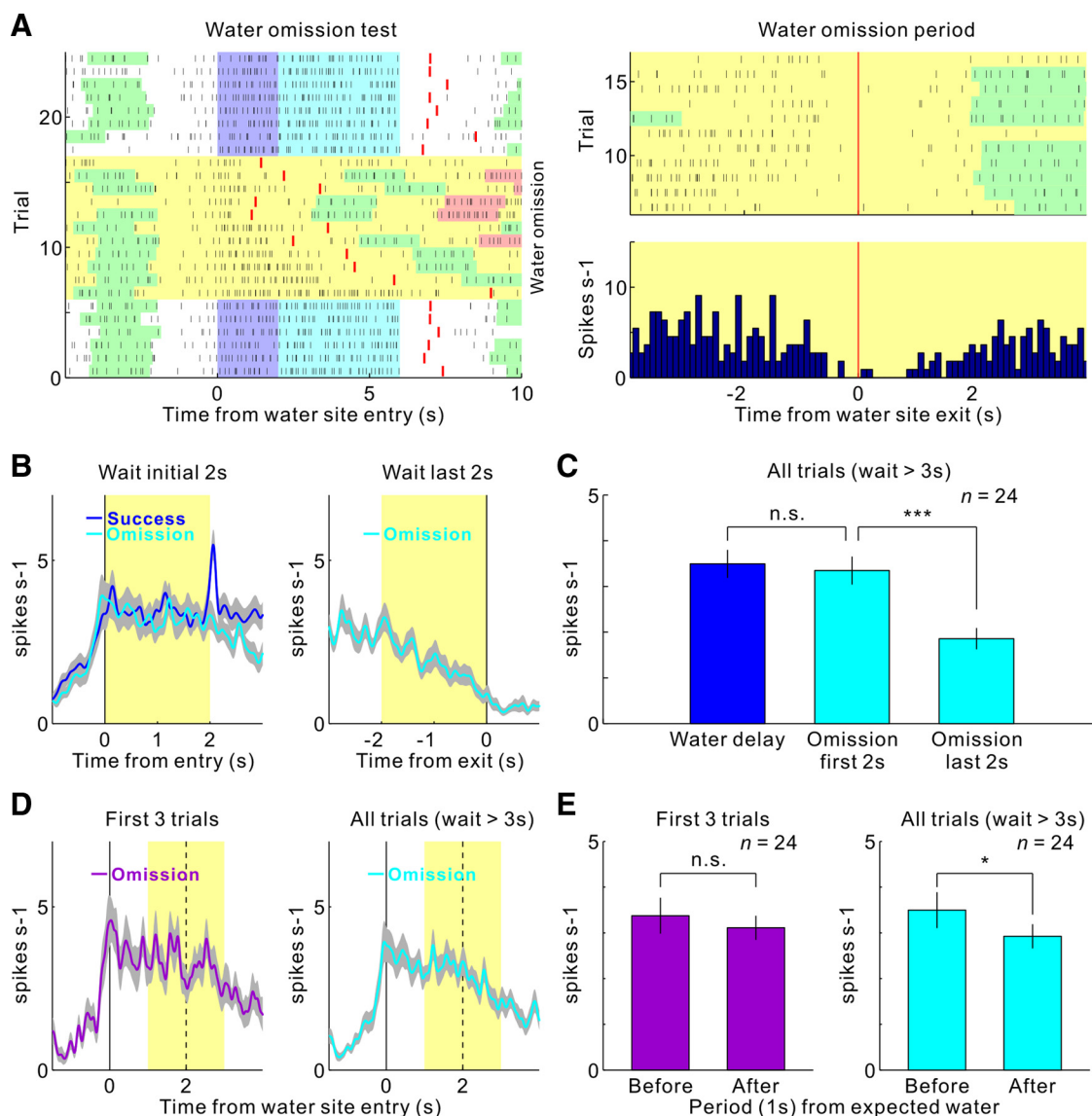


Figure 7. Effects on expected reward omission on serotonin neural activity. **A**, Example of a 5-HT neuron aligned to water site entry during the water omission test. Light yellow areas indicate the water omission period. Red dots indicate water site exit. During the water omission period, food site success entry was rewarded. Right, The raster (top) and the histogram (bottom, 100 ms bins) during water omission period is aligned to the onset of water site exit. **B**, Population activity aligned to water site rewarded entry (blue) and to water omission entry (cyan, left; $n = 24$). Population activity aligned to water site exit after water omission entry (right; $n = 24$). Gray shadings represent SEM. Light yellow areas indicate the periods that were used to analyze average firing rate. **C**, Average firing rates during a 2 s period after water site rewarded entry, after water omission entry, and before water site exit ($n = 24$; \pm SEM). **D**, Population activity of the first 3 trials (left) and all trials (right) after changing to the water omission condition ($n = 24$). **E**, Average firing rate for 1 s before and 1 s after the expected water reward was omitted (left, first 3 trials; right, all trials; $n = 24$; \pm SEM). * $p < 0.01$, Wilcoxon's signed-rank test. n.s., Not significant.

role of 5-HT in impulsive behavior has mainly been examined by lesion and pharmacological manipulation, and there has been little evidence that the serotonin system is activated during behaviors requiring impulse control. Using both microdialysis (Miyazaki et al., 2010) and unit recording techniques, we demonstrated that serotonin neurons in the DRN are prominently activated during periods in which the rat is required to wait for a delayed reward. This increased serotonin activity may facilitate waiting behavior for a delayed reward.

DR neurons are known to fire tonically in conjunction with rhythmic oral–buccal movements (such as licking and chewing) as well as in response to tactile stimuli applied to the face (Fornal et al., 1996). In a supplemental experiment, we used a small camera inside the nose-poke holes to monitor the mouth movements of rats that were trained in the same procedure. The rats did not

show any preparatory chewing or licking movement during the delay period (supplemental Fig. S8, available at www.jneurosci.org as supplemental material). These observations reject the possibility that the increased 5-HT neural activity during delay periods were attributable to rhythmic oral–buccal movements, although more research is necessary to show that 5-HT activity for delayed reward is not related to various motor responses. The possibility that increased 5-HT neural activity is attributable to sensory stimuli to the whisker touching the side walls of the nose-poke holes is also unlikely because the activity of 5-HT neurons started to increase before nose pokes (Fig. 3E) and diminished before exit from the reward sites when rats gave up waiting (Figs. 6B, 7C).

There are reciprocal projections between the DRN and the substantia nigra and the ventral tegmental area (Lee and Geyer,

1984; Herve et al., 1987; Kalenit et al., 1988). Pharmacological studies have shown opposing effects of 5-HT and dopamine in conditioned and unconditioned behaviors (Fletcher et al., 1995; Fletcher and Korth, 1999), and a computational model that assumes an opponent interaction between 5-HT and DA (Daw et al., 2002) has been proposed. It is well known that dopamine reports predicted reward (Schultz, 1998), and the opponent theory posits that serotonin should report predicted punishment or loss of reward (Daw et al., 2002). In the present study, we used a water omission test in which there was no explicit cue to indicate a change from reward delivery to omission; therefore, the first trial in the omission condition presented an unexpected loss of reward, or punishment. We did not find any phasic 5-HT response at the timing of omitted reward. In a recent monkey study, the DRN neurons showed a tonic reward-related response in both small- and large-reward preference manners but not any phasic response to unexpected change from a large reward to a small reward (Nakamura et al., 2008). These results give no support to the opponent hypothesis of 5-HT and dopamine, although a recent recording study in rats reported that a small subset of DRN neurons did respond to reward omission around the expected time of reward (Ranade and Mainen, 2009). In the present study, only the appetitive condition was used to examine 5-HT neural activity. Using aversive conditions (e.g., presentation of quinine solution instead of water omission) would further clarify the role of 5-HT neural activity. Recently, the lateral habenula neurons of monkeys were shown to be excited by a no-reward-predicting target and exhibited a phasic response to unexpected reward omission (Matsumoto and Hikosaka, 2007), suggesting that the habenula system as a more likely opponent of the dopaminergic system.

In our task, there was no choice between immediate and delayed rewards, although the rat could choose whether to wait for a delayed reward or to quit the task. The wait errors associated with lower 5-HT neuron firing suggest that 5-HT can affect choice involving delayed rewards. A recent microdialysis study reported a significant increase in 5-HT efflux in the medial prefrontal cortex of rats performing an inter-temporal choice task compared with “yoked” rats performing no choice (Winstanley et al., 2006). Recording of serotonin neural activity during an inter-temporal choice task would further clarify the role of the serotonergic system in delayed reward discounting and impulsive choice.

We introduced both food and water reinforcement to test whether 5-HT neurons respond differently for different kinds of rewards. Although we did not give separate cues indicating which reward site the rats should head to, the rats could internally discriminate which reward site they came from and which they should head to. The difference in the success rate for the food and water sites in the extended delay condition (Fig. 1F) further suggests that the rats remembered which site gave which reward. We previously found that neurons in the nucleus accumbens and the medial prefrontal cortex exhibited reward-type-dependent activity during the reward delay period (Miyazaki et al., 1998, 2004). In the pigeon *Nidopallium caudolaterale*, a functional analog of the mammalian prefrontal cortex, sustained neural activity during reward delay periods was dependent on the delay and the amount of reward (Kalenscher et al., 2005). Human functional magnetic resonance imaging studies showed brain activity correlated with the values of delayed rewards (McClure et al., 2004; Tanaka et al., 2004, 2006). In this study, only a small number of 5-HT neurons showed reward-type-dependent responses (Fig. 4B), and there was no significant difference in the population

activities during the delay period before food and water rewards (Fig. 3F). The average firing rates during both food and water delays were not affected by the delay length (Fig. 5C). Thus, the sustained 5-HT neuron activity during the delay period is not a simple correlate of anticipated reward value.

Our present study suggests that the increased activity of 5-HT neurons in the DRN is associated with the animal's preparedness to wait for forthcoming rewards or events required to acquire rewards. Through their dense projections to forebrain areas, the activity of the 5-HT neurons may facilitate behaviors that are associated with acquisition of delayed rewards. It remains to be solved how 5-HT efferents modulate cellular and network properties to facilitate waiting behavior for delayed reward and how the afferents to DRN regulate the 5-HT neural activities.

References

- Adell A, Celada P, Abellán MT, Artigas F (2002) Origin and functional role of the extracellular serotonin in the midbrain raphe nuclei. *Brain Res Brain Res Rev* 39:154–180.
- Aghajanian GK, Wang RY, Baraban J (1978) Serotonergic and non-serotonergic neurons of the dorsal raphe: reciprocal changes in firing induced by peripheral nerve stimulation. *Brain Res* 153:169–175.
- Allers KA, Sharp T (2003) Neurochemical and anatomical identification of fast- and slow-firing neurones in the rat dorsal raphe nucleus using juxtacellular labelling methods in vivo. *Neuroscience* 122:193–204.
- Bizot JC, Thiébot MH, Le Bihan C, Soubrié P, Simon P (1988) Effects of imipramine-like drugs and serotonin uptake blockers on delay of reward in rats. Possible implication in the behavioral mechanism of action of antidepressants. *J Pharmacol Exp Ther* 246:1144–1151.
- Bizot J, Le Bihan C, Puech AJ, Hamon M, Thiébot M (1999) Serotonin and tolerance to delay of reward in rats. *Psychopharmacology (Berl)* 146:400–412.
- Cardinal RN (2006) Neural systems implicated in delayed and probabilistic reinforcement. *Neural Netw* 19:1277–1301.
- Daw ND, Kakade S, Dayan P (2002) Opponent interactions between serotonin and dopamine. *Neural Netw* 15:603–616.
- Deakin JFW, Graeff FG (1991) 5-HT and mechanisms of defence. *J Psychopharmacol* 5:305–315.
- Denk F, Walton ME, Jennings KA, Sharp T, Rushworth MF, Bannerman DM (2005) Differential involvement of serotonin and dopamine systems in cost-benefit decisions about delay or effort. *Psychopharmacology (Berl)* 179:587–596.
- Doya K (2002) Metalearning and neuromodulation. *Neural Netw* 15:495–506.
- Doya K (2007) Reinforcement learning: computational theory and biological mechanisms. *HFSP J* 1:30–40.
- Doya K, Uchibe E (2005) The Cyber Rodent Project: exploration of adaptive mechanisms for self-preservation and self-reproduction. *Adapt Behav* 13:149–160.
- Evenden JL (1999) Varieties of impulsivity. *Psychopharmacology (Berl)* 146:348–361.
- Fletcher PJ, Korth KM (1999) Activation of 5-HT_{1B} receptors in the nucleus accumbens reduces amphetamine-induced enhancement of responding for conditioned reward. *Psychopharmacology (Berl)* 142:165–174.
- Fletcher PJ, Tampakeras M, Yeomans JS (1995) Median raphe injections of 8-OH-DPAT lower frequency thresholds for lateral hypothalamic self-stimulation. *Pharmacol Biochem Behav* 52:65–71.
- Fletcher PJ, Tampakeras M, Sinyard J, Higgins GA (2007) Opposing effects of 5-HT(2A) and 5-HT(2C) receptor antagonists in the rat and mouse on premature responding in the five-choice serial reaction time test. *Psychopharmacology (Berl)* 195:223–234.
- Fornal CA, Metzler CW, Marrosu F, Ribiero-do-Valle LE, Jacobs BL (1996) A subgroup of dorsal raphe serotonergic neurons in the cat is strongly activated during oral-buccal movements. *Brain Res* 716:123–133.
- Harrison AA, Everitt BJ, Robbins TW (1997) Central 5-HT depletion enhances impulsive responding without affecting the accuracy of attentional performance: interactions with dopaminergic mechanisms. *Psychopharmacology (Berl)* 133:329–342.
- Hensler JG (2006) Serotonergic modulation of the limbic system. *Neurosci Biobehav Rev* 30:203–214.

- Hervé D, Pickel VM, Joh TH, Beaudet A (1987) Serotonin axon terminals in the ventral tegmental area of the rat: fine structure and synaptic input to dopaminergic neurons. *Brain Res* 435:71–83.
- Heym J, Trulsson ME, Jacobs BL (1982) Raphe unit activity in freely moving cats: effects of phasic auditory and visual stimuli. *Brain Res* 232:29–39.
- Higgins GA, Enderlin M, Haman M, Fletcher PJ (2003) The 5-HT_{2A} receptor antagonist M100,907 attenuates motor and “impulsive-type” behaviours produced by NMDA receptor antagonism. *Psychopharmacology (Berl)* 170:309–319.
- Jacobs BL, Azmitia EC (1992) Structure and function of the brain serotonin system. *Physiol Rev* 72:165–229.
- Jacobs BL, Fornal CA (1999) Activity of serotonergic neurons in behaving animals. *Neuropsychopharmacology* 21:9S–15S.
- Kalén P, Skagerberg G, Lindvall O (1988) Projections from the ventral tegmental area and mesencephalic raphe to the dorsal raphe nucleus in the rat. Evidence for a minor dopaminergic component. *Exp Brain Res* 73:69–77.
- Kalenschers T, Windmann S, Diekamp B, Rose J, Güntürkün O, Colombo M (2005) Single units in the pigeon brain integrate reward amount and time-to-reward in an impulsive choice task. *Curr Biol* 15:594–602.
- Kirby LG, Perner L, Valentino RJ, Beck SG (2003) Distinguishing characteristics of serotonin and non-serotonin-containing cells in the dorsal raphe nucleus: electrophysiological and immunohistochemical studies. *Neuroscience* 116:669–683.
- Kocsis B, Varga V, Dahan L, Sik A (2006) Serotonergic neuron diversity: identification of raphe neurons with discharges time-locked to the hippocampal theta rhythm. *Proc Natl Acad Sci U S A* 103:1059–1064.
- Lee EH, Geyer MA (1984) Dopamine autoreceptor mediation of the effects of apomorphine on serotonin neurons. *Pharmacol Biochem Behav* 21:301–311.
- Marinelli S, Schnell SA, Hack SP, Christie MJ, Wessendorf MW, Vaughan CW (2004) Serotonergic and nonserotonergic dorsal raphe neurons are pharmacologically and electrophysiologically heterogeneous. *J Neurophysiol* 92:3532–3537.
- Matsumoto M, Hikosaka O (2007) Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* 447:1111–1115.
- McClure SM, Laibson DI, Loewenstein G, Cohen JD (2004) Separate neural systems value immediate and delayed monetary rewards. *Science* 306:503–507.
- Miyazaki KW, Miyazaki K, Doya K (2010) Activation of central serotonergic system in response to delayed but not omitted rewards. *Eur J Neurosci*. Advance online publication. Retrieved November 22, 2010. doi:10.1111/j.1460-9568.2010.07480.x.
- Miyazaki K, Mogi E, Araki N, Matsumoto G (1998) Reward-quality dependent anticipation in rat nucleus accumbens. *Neuroreport* 9:3943–3948.
- Miyazaki K, Miyazaki KW, Matsumoto G (2004) Different representation of forthcoming reward in nucleus accumbens and medial prefrontal cortex. *Neuroreport* 15:721–726.
- Mobini S, Chiang TJ, Ho MY, Bradshaw CM, Szabadi E (2000) Effects of central 5-hydroxytryptamine depletion on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology (Berl)* 152:390–397.
- Moser EI, Kropff E, Moser MB (2008) Place cells, grid cells, and the brain's spatial representation system. *Annu Rev Neurosci* 31:69–89.
- Nakamura K, Matsumoto M, Hikosaka O (2008) Reward-dependent modulation of neuronal activity in the primate dorsal raphe nucleus. *J Neurosci* 28:5331–5343.
- Paxinos G, Watson C (1998) The rat brain in stereotaxic coordinates, Ed 4. San Diego: Academic.
- Ranade SP, Mainen ZF (2009) Transient firing of dorsal raphe neurons encodes diverse and specific sensory, motor and reward events. *J Neurophysiol* 102:3026–3037.
- Schultz W (1998) Predictive reward signal of dopamine neurons. *J Neurophysiol* 80:1–127.
- Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S (2004) Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat Neurosci* 7:887–893.
- Tanaka SC, Samejima K, Okada G, Ueda K, Okamoto Y, Yamawaki S, Doya K (2006) Brain mechanism of reward prediction under predictable and unpredictable environmental dynamics. *Neural Netw* 19:1233–1241.
- Vandermaelen CP, Aghajanian GK (1983) Electrophysiological and pharmacological characterization of serotonergic dorsal raphe neurons recorded extracellularly and intracellularly in rat brain slices. *Brain Res* 289:109–119.
- Waterhouse BD, Devilbiss D, Seiple S, Markowitz R (2004) Sensorimotor-related discharge of simultaneously recorded, single neurons in the dorsal raphe nucleus of the awake, unrestrained rat. *Brain Res* 1000:183–191.
- Winstanley CA, Dalley JW, Theobald DE, Robbins TW (2003) Global 5-HT depletion attenuates the ability of amphetamine to decrease impulsive choice on a delay-discounting task in rats. *Psychopharmacology (Berl)* 170:320–331.
- Winstanley CA, Dalley JW, Theobald DE, Robbins TW (2004) Fractionating impulsivity: contrasting effects of central 5-HT depletion on different measures of impulsive behavior. *Neuropsychopharmacology* 29:1331–1343.
- Winstanley CA, Theobald DE, Dalley JW, Cardinal RN, Robbins TW (2006) Double dissociation between serotonergic and dopaminergic modulation of medial prefrontal and orbitofrontal cortex during a test of impulsive choice. *Cereb Cortex* 16:106–114.
- Wogar MA, Bradshaw CM, Szabadi E (1993) Effect of lesions of the ascending 5-hydroxytryptaminergic pathways on choice between delayed reinforcers. *Psychopharmacology (Berl)* 111:239–243.