Distinct basal ganglia circuits controlling behaviors guided by flexible and stable values

Hyoung F. Kim[1,*] and **Okihide Hikosaka**[1]

[1]Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20892

SUMMARY

Choosing valuable objects is critical for survival, but their values may change flexibly or remain stable. Therefore, animals should be able to update the object values flexibly by recent experiences and retain them stably by long-term experiences. However, it is unclear how the brain encodes the two conflicting forms of values and controls behavior accordingly. We found that distinct circuits of the primate caudate nucleus control behavior selectively in the flexible and stable value conditions. Single caudate neurons encoded the values of visual objects in a regionally distinct manner: flexible value coding in the caudate head, and stable value coding in the caudate tail. Monkeys adapted in both conditions by looking at objects with higher values. Importantly, inactivation of each caudate subregion disrupted the high-low value discrimination selectively in the flexible or stable context. This parallel complementary mechanism enables animals to choose valuable objects in both flexible and stable conditions.

INTRODUCTION

We choose between objects based on their values, which we learn from past experience with rewarding consequences (Awh et al., 2012; Chelazzi et al., 2013). The values of some objects change flexibly, and we have to search valuable objects based on their consequent outcome (Barto, 1994; Dayan and Balleine, 2002; Padoa-Schioppa, 2011; Rolls, 2000). On the other hand, the values of some other objects remain unchanged, and we have to choose the valuable objects based on the long-term memory. Since the stable value formed by repetitive experiences is reliable, we may consistently choose the object regardless of the outcome (Ashby et al., 2010; Balleine and Dickinson, 1998; Graybiel, 2008; Mishikin et al., 1984; Wood and Neal, 2007). Both flexible and stable value-guided behaviors are critical to choose the valuable objects efficiently. If we rely only on flexible values, we would always have to make an effort to find valuable objects by trial and error. If we rely only on stable values, we would fail to choose valuable object if their values have changed recently. Therefore, our brain must acquire both flexible and stable values of objects to guide each behavior.

However, the flexible and stable values are often mutually conflicting (stability-flexibility dilemma) (Abraham and Robins, 2005; Anderson, 2007; Daw et al., 2006; Liljenström, 2003). For the flexible value any short-term change in object value matters, and the memory must be updated quickly. For the stable value only a long-term change matters, and the

*Correspondence hyoung.f.kim@gmail.com.

memory must be updated only slowly so that small changes can be ignored. It is still unclear how the brain encodes both flexible and stable values to guide choice behavior accordingly. It would be difficult for a single neural circuit to process the potentially conflicting values. One alternative hypothesis would be that the brain has two independent mechanisms, one encoding flexible memories and the other encoding stable memories to guide choice behavior differently in each situation. Notably, the parallel process has been suggested to be a fundamental feature of the brain anatomically and functionally (Alexander et al., 1986). Especially, the basal ganglia have well-known parallel anatomical circuits connected from cortical regions to output structures (Alexander et al., 1986; Kemp and Powell, 1970; Szabo, 1970, 1972). In particular, the caudate nucleus receives inputs from a large portion of the cerebral cortex including the prefrontal and temporal cortex (Saint-Cyr et al., 1990; Selemon and Goldman-Rakic, 1985; Yeterian and Van Hoesen, 1978) through which visual object information is processed (Kim et al., 2012; Yamamoto et al., 2012). We thus hypothesized that the caudate nucleus contains parallel functional units which process object value information independently.

To test this hypothesis, we performed two experiments, first aiming at neuronal information processing and then behavioral causality. These experiments together suggested that the head and tail of the primate caudate nucleus have distinct functions, the head guiding controlled behavior based on flexible values and the tail guiding automatic behavior based on stable values.

RESULTS

To examine the value representation and the behavior control by the caudate nucleus, we used flexible and stable value procedures (Figure 1). Figure S1 shows the underlying concept. In each case the monkey experienced fractal objects with high values and low values. In the flexible value procedure (Figure S1A), objects changed their values frequently and the monkey must adapt to the changes flexibly. This is a short-term learning process. In the stable value procedure (Figure S1B), objects retained their values (i.e., high or low) stably across repeated learning. This is a long-term learning process. The testing of the long-term memory was done in a separate experimental context in which objects were no longer associated with the previously assigned values.

We used saccades to the fractal objects as the behavioral measure (Figure 1). In the flexible value procedure (Figure 1A), the saccade to one object was followed by a reward and the other object was associated with no reward, and this contingency was reversed frequently. To examine the short-term behavioral learning, we measured the target acquisition time after a go cue (the disappearance of the fixation point). As the value of each object changed block-wise, the target acquisition time changed accordingly: the monkeys made saccades more quickly to the high-valued object than the low-valued object (Figure 1B) (difference of target acquisition time: 57.7 ms, $p \ll 0.001$, two-tailed t-test). On choice trials (see Experimental Procedures) the monkeys mostly chose the high-valued object (average: 83.9 $\pm 0.8\%$). These saccades can be called "controlled saccades," because they were controlled by reinforcing feedbacks delivered just after the saccades.

During learning of stable values (Figure 1C), the saccades to a set of objects were always followed by a reward (high-valued) and the saccades to a different set of objects were always followed by no reward (low-valued), and this was repeated across days (see Figure S2 for detail). To examine the long-term behavioral memory, we used a free-looking task (Figure 1D) and a free-viewing procedure (Figure S2D). These tests were done at least one day after the learning session, and the saccades were followed by no reward. Yet, the monkey made saccades to the objects automatically, and did so more likely to high-valued

objects than low-valued objects. The preference to the high-valued objects emerged slowly across several daily learning sessions and then remained stably after 4 daily sessions of learning (Figure S2D), as reported previously (Yasuda et al., 2012). Therefore, to analyze the neuronal and behavioral coding of stable object values, we used fractal objects that the monkey had learned for more than 4 daily sessions. When such well learned objects were used in the free-looking task, the likelihood of saccades to high-valued objects was significantly higher than to low-valued ones (Figure 1D-right) (difference of automatic looking: 18.9%, p < 0.01, two-tailed t-test). These saccades can be called "automatic saccades," because they were not followed by any reinforcing feedbacks delivered just after the saccades.

To test whether the caudate nucleus controls the saccade behavior to choose high-valued objects, we recorded spike activity of single neurons in the caudate nucleus using the flexible and stable value procedures. We first found that many neurons in the caudate nucleus responded to visual objects, confirming previous studies (Brown et al., 1995; Caan et al., 1984; Rolls et al., 1983; Yamamoto et al., 2012). The ratios of neurons that responded to fractal objects relative to the encountered neurons in the three caudate regions (Figure 3A) were: head 163/845 (19.3 %), body 109/381 (28.6 %), tail 107/205 (52.2 %). Their responses were often modulated by the values associated with the objects. Importantly, neurons in different regions of the caudate nucleus were influenced by flexible and stable values differently. Figure 2 shows the activity of three example neurons which were recorded from three caudate regions.

In the flexible value procedure (Figure 1A and 2A), the three example neurons in the caudate responded to the fractal objects with a phasic activation, but in different ways (Figure 2C). The caudate head neuron was activated by the objects when their values were high (Figure 2C-left, red), but not when their values were low (Figure 2C-left, blue) (p << 0.001, two-tailed Wilcoxon rank-sum test). The caudate body neuron weakly encoded the flexible reward values of the objects (Figure 2C-center): its responses to the objects were more prolonged when their values were low. In contrast, the responses of the caudate tail neuron were not influenced by the flexibly changing values of the objects (Figure 2C-right).

In the stable value procedure, the same three caudate neurons behaved quite differently compared to the activity in the flexible value procedure. To test the neuronal activity, we serially presented fractals without any object-reward contingency while the monkey was fixing on the center dot (Figure 2B). The caudate head neuron, which responded selectively to objects with high flexible values (Figure 2C-left), became nearly silent in the stable value procedure (Figure 2D-left). The caudate body neuron, which weakly encoded negative flexible values (Figure 2C-center), showed little bias based on stable values (Figure 2D-center). In contrast, the caudate tail neuron, which was not influenced by objects' flexible values (Figure 2C-right), now showed a clear bias toward objects with high stable values (Figure 2D-right) (p << 0.001, two-tailed Wilcoxon rank-sum test).

The regional difference in flexible/stable value encoding, exemplified in Figure 2, was commonly present among caudate neurons. This is shown in Figure 3B and C as the averaged responses of all neurons responding to fractal objects in the three caudate subregions. Since different caudate neurons responded more strongly to high-valued objects or to low-valued objects (positive and negative neurons in Figure 3D,E), we averaged the neurons' responses (using cross-validation) separately for the neurons' preferred value (magenta) and the non-preferred value (black) (Figure 3B,C). The bias in activity based on flexible values appears strongest in the caudate head and weakest in the caudate tail (Figure 3B, yellow). In contrast, the bias in activity based on stable values appears strongest in the caudate tail and weakest in the caudate head (Figure 3C, yellow). Similar trends were

observed for both positive and negative neurons (Figure S3). These conclusions were confirmed by the subregional difference in the proportion of neurons that showed a statistically significant bias based on flexible values (Figure 3D) or stable values (Figure 3E). The flexible and stable value biases showed two opposing gradients across the caudate head, body and tail. An analysis of individual neurons supports these conclusions (Figure 4).

We considered factors that might confound our interpretation. First, the regional difference might be caused by the monkey's long-term experience of the experimental procedure. This is unlikely, however, because we recorded from the three caudate subregions in a temporally counter-balanced manner along the whole experimental project. Second, the regional difference might depend on the difference in the testing procedure (i.e., saccades to objects in the flexible value procedure, not in the stable value procedure). However, this possibility was not supported by a supplementary experiment using the flexible value-fixation task (Figure S4).

We so far have shown that 1) the flexible and stable values are represented in the caudate subregions differentially (particularly caudate head and tail) and 2) the flexible and stable values induce controlled and automatic saccades respectively. These results suggest that the caudate nucleus contains parallel mechanisms: the caudate head guides controlled saccades based on flexible values and the caudate tail guides automatic saccades based on stable values. Our data support this hypothesis, as shown below. Since caudate body neurons showed an intermediate coding pattern, we will focus on the comparison between the caudate head and tail.

First, neurons in the caudate head, not tail, showed value-differential activity before controlled saccades. In the flexible value procedure which induced controlled saccades (Figure 1A,B), the value-differential response of caudate head neurons (Figure 5A-left, yellow) emerged in parallel with the change in the monkey's target acquisition time (Figure 5B-left, yellow). Such a correlation was absent in caudate tail neurons (Figure 5, right).

Second, the flexible-stable dichotomy was observed using the object-value assocation learning task (Figure S2) in different contexts (Figure 6). When the monkey learned the values of novel objects (Figure 6A-left), neurons in the caudate head, not tail, acquired the value-differential response (Figure 6B,C-left), as in the flexible value procedure (Figure 5A). In contrast, when well-learned objects were introduced after more than one-day retention, the monkey showed a clear bias in the target acquisition time from the beginning throughout the session (Figure 6A-right). This was paralleled by the stably maintained value bias in caudate tail neurons (Figure 6C-right) which was absent during the new learning (Figure 6C-left). Notably, caudate head neurons showed no value bias initially, although they quickly acquired it (Figure 6B-right). These results suggest that neurons in the caudate tail, not head, can support the value-differential saccades when previously well-learned objects are unexpectedly presented.

Third, in the flexible value procedure, the overall pre-saccadic activity of caudate head neurons was significantly stronger to preferred value objects than non-preferred value objects (Figure 7A-top), but such a bias was not detected in the caudate tail (Figure 7A-bottom). In the free-looking task (as part of the stable value procedure) which induced automatic saccades (Figure 1D), caudate tail neurons showed pre-saccadic activity which was significantly stronger to preferred value objects than to non-preferred value objects (Figure 7B, bottom). Such presaccadic activity was absent in caudate head neurons (Figure 7B-top). The caudate tail-specific activity preceding automatic saccades was confirmed using a free-viewing procedure (Figure S5) in which four objects, chosen randomly on each

trial, were presented simultaneously and the monkey looked at them with no reward consequence.

To further test the flexible-stable dichotomy hypothesis, we selectively inactivated the caudate head or the caudate tail by injecting a $GABA_A$ receptor agonist, muscimol (Figure 8A). The inactivation of the caudate head disrupted the initiation of saccades in the flexible value task (which we call controlled saccades) (Figure 8B-top). Before the inactivation, the target acquisition time on single object trials was significantly shorter for high-valued objects than for low-valued objects (Figure 1B, Figure S6B-left). This bias of controlled saccades decreased significantly during the caudate head inactivation (Figure 8B-top), but only for contralateral saccades (from 69.7 ms to 20.4 ms; $p < 0.01$, paired t-test). The bias decrease was largely due to earlier saccades to low-valued objects (S6B-top). The caudate head inactivation also disrupted the choice of the high-valued objects in the flexible value task (Figure S7C-top), again only for contralateral saccades ($p < 0.05$, paired t-test), when four, not two, objects were used. However, the caudate head inactivation did not affect saccades in the stable value procedure using either the free-looking task (Figure 8C-top) or the free-viewing procedure (Figure S8B).

In contrast, the inactivation of the caudate tail specifically disrupted the initiation of saccades in the stable value task (free-looking task) (Figure 8C-bottom). Before the inactivation, the likelihood of saccades to the presented object (which we call automatic saccades) was higher for high-valued objects than for low-valued objects (Figure 1D, Figure S6D-left). This bias of automatic saccades disappeared during the caudate tail inactivation (Figure 8C-bottom), but only for contralateral saccades (from 19.9% to −1.2%; $p < 0.01$, paired t-test). The bias decrease was largely due to more frequent saccades to low-valued objects (S6D-bottom). Among the saccades made to the presented object, there was no change in latency. The caudate tail inactivation also disrupted the choice of the high-valued objects in the stable value task, again only for contralateral saccades (free-viewing procedure, Figure S8C; see Figure S5 for neuronal activity). However, the caudate tail inactivation did not affect saccades in the flexible value procedure in either the single object trials (Figure 8B-bottom) or the choice trials (Figure S7C-bottom).

DISCUSSION

Our results demonstrate that two subregions of the caudate nucleus, head and tail, distinctly encode the flexible and stable values of visual objects, and these value memories guide behavior in controlled and automatic manners, selectively and respectively. This provides an answer to a long-standing question about the function of the parallel neural circuits in the basal ganglia. The parallel circuits are thought to serve different functions, such as oculomotor, motor, cognitive, and emotional functions (Alexander et al., 1986). However, it is unclear how these circuits coordinate with each other during adaptive behavior. Our data suggest that the caudate subregions work integratively but independently, aiming at a unitary goal – choosing valuable objects.

How can parallel and independent mechanisms work for a unitary goal? We propose that caudate head and tail work in a mutually complementary manner. Their complementary features are twofold: information and behavior, as discussed below.

Flexible value coding is useful to find valuable objects if their values change frequently. This is the function that the caudate head contributes to. Single neurons of the caudate head change their responses flexibly to inform which objects are recently more (or less) valuable. Their responses rely on short-term memory or working memory. Such flexibility is an essential feature of cognitive functions (Kehagia et al., 2010). Indeed, many neurons in

"cognitive" brain areas encode flexible object values (Kim et al., 2008; Padoa-Schioppa, 2011; Rolls, 2000; Thorpe et al., 1983; Tremblay and Schultz, 1999).

However, the caudate head does not retain the value information, once the reinforcing feedback is not delivered immediately. This is problematic because the information would not allow us (and animals) to choose valuable objects until we experience an actual reward. The caudate tail, as part of the stable value system, would compensate for this limitation. Single neurons in the caudate tail respond to objects differentially based on the previous, long-term experience of the objects (see Yamamoto et al., 2013 for details). This information would enable us to choose valuable objects without updated feedback. Such stable value information would underlie visual skills (Gottlieb, 2012; Shiffrin and Schneider, 1977; Wood and Neal, 2007). However, the caudate tail may work inadequately in a flexible condition, since it is insensitive to recent changes in object values.

Clearly, the caudate head and tail, together but in parallel, provide a robust capacity for choosing valuable objects efficiently. If a set of objects changes their values frequently, neurons in the caudate head adapt to the changes by quickly altering their responses to the objects based on recent outcomes. If another set of objects retains their values for a long time, neurons in the caudate tail retain the sensitivity to the objects and, when any of the objects appears, react to it automatically regardless of the outcome; this occurs quickly to many objects.

Behaviorally, our inactivation experiments indicate that the caudate head and tail guide saccades aiming at valuable objects in different manners. The caudate head guides controlled saccades based on the flexible values (with immediate feedbacks), whereas the caudate tail guides automatic saccades based on the stable values (with no feedback). Consistent with these results, neurons in these caudate subregions showed value-differential pre-saccadic activity, but in different contexts: caudate head neurons during controlled saccades vs. caudate tail neurons during automatic saccades. Notably, the inactivation of the caudate head as well as tail appeared to decrease the suppression of saccades to low-valued objects, rather than decrease the facilitation of saccades to high-valued objects. This may be determined by the balance between the direct and indirect pathways (Hikosaka et al., 2000). How the balance might be controlled remains to be studied.

The controlled saccades guided by caudate head and the automatic saccades guided by the caudate tail may be equivalent to a well-documented dichotomy of behavior, such as goal-directed behavior vs. skill (or habit) (Balleine and Dickinson, 1998), controlled vs. automatic behavior (Schneider and Shiffrin, 1977), and System 2 vs. System 1 (Evans, 2008). Several lines of evidence in human neuroimaging, human clinical, animal lesion, and physiological studies suggest that different regions of the basal ganglia are involved in controlled vs. automatic behavior (Ashby and Maddox, 2005; Balleine and O'Doherty, 2010; Hikosaka et al., 1999; Redgrave et al., 2010; Yin and Knowlton, 2006). Human neuroimaging data suggest that subregions of the basal ganglia become active differentially depending on planning, skill acquisition, reward prediction, and feedbacks (Balleine and O'Doherty, 2010; Seger, 2008; Wunderlich et al., 2012). Human patients with Parkinson's disease are impaired in cognitive tasks that require flexible adaptations to environmental changes, such as set-shifting and value reversal (Brown and Marsden, 1990; Cools et al., 1984; Kehagia et al., 2010; Lees and Smith, 1983). On the other hand, Parkinson's disease patients are also impaired in probabilistic category learning tasks which require visual skills (Ashby and Maddox, 2005; Knowlton et al., 1996; Shohamy et al., 2004). Patients with Huntington's disease may show profound impairments in visual recognition (Lawrence et al., 1998), even early in the disease when neurodegeneration is detected mainly in the caudate tail (Vonsattel et al., 1985). Notably, monkeys with lesions in the caudate tail are

deficient in visual skills (Fernandez-Ruiz et al., 2001). The stable value information in the caudate tail may be transmitted to the superior colliculus through the substantia nigra pars reticulata so that monkeys make saccades preferentially to high-valued objects (Yasuda et al., 2012).

Although these studies individually provide important data, it has been difficult to reach a unified view on basal ganglia functions. Our recording and inactivation experiments on the primate caudate head and tail provide new insights in understanding how the basal ganglia normally control behavior in multiple but integrative ways and how behavior can be disrupted in multiple ways in basal ganglia disorders.

EXPERIMENTAL PROCEDURES

General procedures

Two adult male rhesus monkeys (*Macaca mulatta*, 8-9 kg) were used for the experiments. All animal care and experimental procedures were approved by the National Eye Institute Animal Care and Use Committee and complied with the Public Health Service Policy on the humane care and use of laboratory animals. We implanted a plastic head holder and a recording chamber to the skull under general anesthesia and sterile surgical conditions. The chamber was tilted laterally by 25° and was aimed at the caudate head, body and tail. Two search coils were surgically implanted under the conjunctiva of the eyes to record eye movements. After the monkeys fully recovered from surgery, we started training them with flexible and stable value procedures.

Neural recording

While the monkey was performing a task, we recorded the activity of single neurons in different subregions in the caudate nucleus using conventional method. The recording sites were determined with 1 mm-spacing grid system, with the aid of MR images (4.7T, Bruker) obtained along the direction of the recording chamber. Single-unit recording was performed using glass-coated electrodes (Alpha-Omega). The electrode was inserted into the brain through a stainless-steel guide tube and advanced by an oil-driven micromanipulator (MO-97A, Narishige). The electric signal from the electrode was amplified with a bandpass filter (2 Hz-10kHz; BAK, Mount Airy, MD) and collected at 1 kHz. Neural spikes were isolated on-line using a custom voltage-time window discrimination software (MEX, LSR/NEI/NIH).

Behavioral tasks

Behavioral tasks were controlled by a QNX-based real-time experimentation data acquisition system (REX, Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health (LSR/NEI/NIH), Bethesda, Maryland). The monkey sat in a primate chair, facing a fronto-parallel screen 33cm from the monkey's eyes in a sound attenuated and electrically shielded room. Visual stimuli generated by an active matrix liquid crystal display projector (PJ550, ViewSonic) were rear-projected on the screen. We created the visual stimuli using fractal geometry. Their sizes were approximately $8° \times 8°$.

Flexible value procedure—This procedure allowed us to examine behavioral and neuronal encoding of flexible object values as they were being updated in blocks of trials (Figure 1A and Figure S1A). Therefore, learning (of object values) and testing (of the monkey's behavior and of neuronal activity) were done in one task procedure, as illustrated in Figure 1A. For each monkey a fixed set of two fractal objects (say, A and B) was used as the saccade target (except in some experiments used for the muscimol-induced inactivation, see below). Each trial started with a central white dot presentation, which the monkey was

required to fixate. After 700 ms, while the monkey was fixating on the central spot, one of the two fractal objects was chosen pseudo-randomly and was presented at one of two diagonally symmetric positions (one of them at the neuron's preferred location). The preferred position was determined using a saccade task in which another fractal, as the target, was presented at different positions. The fixation spot disappeared 400 ms later, and then the monkey was required to make a saccade to the object within 4 s. The monkey received a liquid reward 300 ms after making a saccade to one object (e.g., A), but received no reward after making a saccade to the other object (e.g., B). During a block of 30 to 40 trials, the object-reward contingency was fixed, but it was reversed in a following block (e.g., B-high/A-low) without any external cue. While a neuron was being recorded, these two blocks (A-high/B-low and B-high/A-low) were alternated in blocks (their order counterbalanced across neurons). Most trials (24-32 out of 30-40 trials) were single object trials: one of the two objects was presented and the monkey had to make a saccade to it. The purpose of the single object trials was to examine how quickly the saccade is made to the presented object (target acquisition time, see Data analysis). The rest of trials (6-8 out of 30-40 trials) were choice trials: two objects were presented at the same time, one at the neuron's preferred position and the other at the diagonally symmetric position. The monkey had to choose one of the objects by making a saccade to it to obtain the reward associated with the chosen object. The purpose of the choice trials was to examine how likely the saccade is made to the high-valued object (choice rate, see Data analysis). If the monkey failed to make a saccade correctly on either single object or choice trials, the same trial was repeated. In each recording session, these two types of block were repeated at least twice. This flexible value procedure was modified in a supplementary experiment (Figure S4) in which the monkey had to keep fixating the central spot while an object was presented (400 ms) until a trial ended. In some experiments for the muscimol-induced inactivation of caudate subregions (see Figure S7), four familiar fractal objects were used in a 2-2 format (C&D-high/E&F-low and E&F-high/C&D-low). Half of 32 trials (one block) were single object trials. The other half were choice trials: two objects were simultaneously presented on one side, either contralateral or ipsilateral to the inactivation site. Other procedures were the same as the 1-1 format described above.

Stable value procedure—To examine behavioral and neuronal encoding of stable object values, we conducted the learning procedure and the testing procedure separately on different days (Figure 1C and Figure S1B). In the learning procedure, the monkey experienced visual objects repeatedly in association with consistent reward values and thus learned their stable values (Figure 1C and Figure S2). In the testing procedure, monkey's saccade behavior and neuronal activity were examined using different tasks (see Figure 1D and 2B). To focus on stable object values, the testing procedure was applied to objects that had been learned for more than 4 daily sessions. Below we explain in detail (1) the learning procedure, (2) the procedure for testing neuronal activity, and (3) the procedure for testing saccade behavior.

1. Procedure for learning stable object values (Figure S2). To create a fixed bias among fractal objects in their reward values, we used an object-directed saccade task. In each session, a set of eight fractals was used as the target and was presented at one of five positions (right, up, left, bottom, and center). The monkey made a saccade to the target to obtain a liquid reward. Half of the fractals were always associated with a liquid reward (high-valued objects), whereas the other half were associated with no reward (low-valued objects). One training session consisted of 160 trials (20 trials for each object). Each set was learned in one learning session in one day. The same sets of fractals were used repeatedly for learning across days (or months), throughout which each object remained to be either a high-valued object or a low-valued object. Monkey 1 and 2 learned 608 and 456 fractals, respectively,

among which 312 and 176 fractals were learned extensively (> 4 daily sessions). The long-term learning continued during the whole experimental project. Note that individual object sets were learned with variable intervals (6.4 ± 0.3 days) for two reasons: 1) There were too many object sets to be learned in one day; 2) some object sets were removed from the list of learning to test the effects of memory retention (though this is not the subject of the current study). The test of stable value coding (described below) was done by choosing some sets of objects (usually >2 sets: >16 objects) from the well-learned sets of objects (61 sets: 488 objects).

2. Procedure for testing saccade behavior (Figure 1D and Figure S2D). To examine the monkey's coding of stable object values, we used a free-looking task and a free-viewing procedure. In the free-looking task (Figure 1D), each trial started with a white dot presentation at one of the three positions (up, center and down in centerline), and the monkey was required to fixate it for 500 ms. A fixed time (100 or 200 ms) after the fixation point disappeared, one of a set of eight fractal objects (Figure 1C) was chosen pseudo-randomly and was presented on the right or left side. The monkey was free to look at the object or anywhere else in 2 s, but no reward was given. If the monkey made a saccade to the object within the 2 s, the object disappeared 300 ms after the saccade. On one third of the trials the fixation point was followed by the delivery of a reward without an object presentation. This reward trial was used to maintain the monkey's arousal and motivation level. In the free-viewing procedure (Figure S2D), each trial started with a central white dot presentation, and the monkey was required to fixate it. After 300 ms of fixation, four of a set of eight fractal objects were chosen pseudo-randomly and were presented simultaneously for 2 s. The monkey was free to look at these objects for 2 s, but no reward was given. After a blank period (500 ms), another four objects were presented. On half of the trials, instead of the objects, a white dot was presented at one of four positions. If the monkey made a saccade to it, a reward was delivered. This reward trial was used to maintain the monkey's arousal and motivation level. Each object was presented at least 16 times in one session.

3. Procedure for testing neuronal activity (Figure 2B). To examine the neuronal coding of stable object values, we used a passive viewing task. In each session a set of eight fractal objects was used as the visual stimuli. While the monkey was fixating on a central spot of light, the fractal objects were presented sequentially in the neuron's preferred position in a pseudo-random order (presentation time: 400 ms, inter-object interval time: 500-700 ms). The preferred position was determined using the passive viewing task in which another fractal was presented at various positions. After every 1–4 object presentation, a reward was delivered 300 ms later. The reward was thus not associated with particular objects. Each object was presented at least six times in one session. The neuronal coding of stable object values was tested after long-term learning (> 4 daily sessions) and after a sufficient retention period (> 1 day after the last learning session). For each neuron more than 2 sets of fractals (i.e., > 16 fractals) were used for the testing.

Inactivation of caudate nucleus

To inactivate each region of the caudate nucleus, muscimol (GABA$_A$ agonist) was injected into the head or tail of the caudate nucleus (Figure 8A) (Hikosaka and Wurtz, 1985). The injection was done in either the right or left side of the caudate nucleus of each monkey. To accurately locate the injection site, we recorded single or multiple neuronal activity before the injection and confirmed that the neurons were sensitive to flexible or stable values of fractal objects. For this purpose, we used a custom-made injectrode consisting of an epoxy-coated tungsten microelectrode for neuron recording and a silica tube for muscimol

injection. Before the injection while the injectrode was positioned at the injection site, the monkey performed the flexible value procedures (Flexible value task, Figure 1A; flexible value-choice task, Figure S7) and the stable value procedures (free-looking task, Figure 1D; free-viewing procedure, Figure S2D), and the data were used as a pre-injection control. We injected 1 μl of 5.12 mM muscimol (Sigma) at the speed of 0.2 μl/min. Starting 5 min after the injection, the monkey was required to resume the flexible and stable value tasks. The tests were repeated several times until 2-3 h after the injection. We performed the inactivation experiments after collecting most of the behavioral and neuronal data.

Data analysis

We analyzed the neuronal and behavioral discriminations of high-valued and low-valued objects. To assess the neuronal discrimination, we first measured the magnitude of the neuron's response to each fractal object by counting the numbers of spikes within a test window in individual trials. For stable object-value learning, the test window was 0–400 ms after the onset of the object in the passive-viewing task. For flexible object value learning, the test window was 0-400 ms after the onset of the object in the object-directed saccade task. The neuronal discrimination was defined as the area under the receiver operating characteristic (ROC) based on the response magnitudes of the neurons to high-valued objects versus low-valued objects (Figure 4). The statistical significance of the neuronal discrimination was tested using two-tailed Wilcoxon rank-sum test.

We also assessed the overall neuronal discrimination of object values in the subregions of the caudate nucleus (head, body, and tail) (Figure 3). Since some caudate neurons responded more strongly to high-valued objects (i.e., positive neurons) while others to low-valued objects (i.e., negative neurons), we first determined each neuron's preferred value by comparing the magnitude of the neuron's response to high-valued objects and to low-valued objects. This was done by computing an ROC area based on the numbers of spikes within the test window in individual trials. We then averaged the responses of individual neurons in each subregion separately for the neurons' preferred value and the non-preferred value. This was done by using a cross-validation method. Specifically, trials in one recording session were divided into the odd and even numbered trials. Either odd or even numbered trials were randomly chosen for determining the neuron's preferred value (using the ROC analysis), and the other was used for computing the average response. The cross-validation method precluded any artificial result of neuronal discrimination due to an arbitrary choice of the preferred value.

To assess the behavioral discrimination, we used several measures. For the flexible value procedure, we computed the target acquisition time (single object trial, Figure 1B) and the choice rate (choice trials, Figure S7). For the stable value procedure, we computed the probability of automatic looking (single object-presenting trials – free-looking task, Figure 1D) and the choice rate (choice trials – free-viewing procedure, Figure S2). The target acquisition time was defined as the time from the go-signal (i.e., the disappearance of the fixation point) until the gaze reached the presented object (Figure 1A). To assess the behavioral discrimination across multiple test sessions, we computed an ROC area based on the target acquisition times for high-valued vs. low-valued objects (Figure 6A). The probability of automatic looking was defined as the probability of trials in which a saccade was made to the presented object (Figure 1D). The choice rate was defined as follows: $(nSACh-nSACl)/(nSACh+nSACl)$ where $nSACh$ and $nSACl$ are the numbers of saccades toward high-valued and low-valued objects, respectively.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Abraham WC, Robins A. Memory retention – the synaptic stability versus plasticity dilemma. Trends Neurosci. 2005; 28:73–78. [PubMed: 15667929]

Alexander GE, DeLong MR, Strick PL. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu. Rev. Neurosci. 1986; 9:357–381. [PubMed: 3085570]

Anderson P. Flexibility and stability in the innovating economy. Admin. Sci. Quart. 2007; 52:333–335.

Ashby FG, Maddox WT. Human category learning. Annu. Rev. Neurosci. 2005; 56:149–178.

Ashby FG, Turner BO, Horvitz JC. Cortical and basal ganglia contributions to habit learning and automaticity. Trends Cogn. Sci. 2010; 14:208–215. [PubMed: 20207189]

Awh E, Belopolsky AV, Theeuwes J. Top-down versus bottom-up attentional control: a failed theoretical dichotomy. Trends Cogn. Sci. 2012; 16:437–443. [PubMed: 22795563]

Balleine BW, Dickinson A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology. 1998; 37:407–419. [PubMed: 9704982]

Balleine BW, O'Doherty JP. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology. 2010; 35:48–69. [PubMed: 19776734]

Barto AG. Reinforcement learning control. Curr. Opin. Neurobiol. 1994; 4:888–893. [PubMed: 7888773]

Brown RG, Marsden CD. Cognitive function in Parkinson's disease: from description to theory. Trends Neurosci. 1990; 13:21–29. [PubMed: 1688671]

Brown VJ, Desimone R, Mishkin M. Responses of cells in the tail of the caudate nucleus during visual discrimination learning. J. Neurophysiol. 1995; 74:1083–1094. [PubMed: 7500134]

Caan W, Perrett DI, Rolls ET. Responses of striatal neurons in the behaving monkey. 2. Visual processing in the caudal neostriatum. Brain Res. 1984; 290:53–65. [PubMed: 6692139]

Chelazzi L, Perlato A, Santandrea E, Della Libera C. Rewards teach visual selective attention. Vision Res. 2013; 85:58–72. [PubMed: 23262054]

Cools AR, Van den Bercken JH, Horstink MW, Van Spaendonck KP, Berger HJ. Cognitive and motor shifting aptitude disorder in Parkinson's disease. J. Neurol. Neurosur. Ps. 1984; 47:443–453.

Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. Nature. 2006; 441:876–879. [PubMed: 16778890]

Dayan P, Balleine B. Reward, Motivation, and Reinforcement Learning. Neuron. 2002; 36:285–298. [PubMed: 12383782]

Evans J. Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. Annu. Rev. Psychol. 2008; 59:255. [PubMed: 18154502]

Fernandez-Ruiz J, Wang J, Aigner TG, Mishkin M. Visual habit formation in monkeys with neurotoxic lesions of the ventrocaudal neostriatum. P. Natl. Acad. Sci. USA. 2001; 98:4196–4201.

Gottlieb J. Attention, learning, and the value of information. Neuron. 2012; 76:281–295. [PubMed: 23083732]

Graybiel AM. Habits, rituals, and the evaluative brain. Annu. Rev. Neurosci. 2008; 31:359–387. [PubMed: 18558860]

Hikosaka O, Nakahara H, Rand MK, Sakai K, Lu X, Nakamura K, Miyachi S, Doya K. Parallel neural networks for learning sequential procedures. Trend Neurosci. 1999; 22:464–471. [PubMed: 10481194]

Hikosaka O, Takikawa Y, Kawagoe R. Role of the basal ganglia in the control of purposive saccadic eye movements. Physiol. Rev. 2000; 80:953–978. [PubMed: 10893428]

Hikosaka O, Wurtz RH. Modification of saccadic eye movements by GABA-related substances. I. Effect of muscimol and bicuculline in the monkey superior colliculus. J. Neurophysiol. 1985; 53:266–291. [PubMed: 2983037]

Kehagia AA, Murray GK, Robbins TW. Learning and cognitive flexibility: frontostriatal function and monoaminergic modulation. Curr. Opin. Neurobiol. 2010; 20:199–204. [PubMed: 20167474]

Kemp JM, Powell TP. The cortico-striate projection in the monkey. Brain. 1970; 93:525–546. [PubMed: 4990231]

Kim S, Hwang J, Lee D. Prefrontal coding of temporally discounted values during intertemporal choice. Neuron. 2008; 59:161–172. [PubMed: 18614037]

Kim S, Cai X, Hwang J, Lee D. Prefrontal and Striatal Activity Related to Values of Objects and Locations. Front. Neurosci. 2012; 6:108. [PubMed: 22822390]

Knowlton BJ, Mangels JA, Squire LR. A neostriatal habit learning system in humans. Science. 1996; 273:1399–1402. [PubMed: 8703077]

Lawrence AD, Sahakian BJ, Robbins TW. Cognitive functions and corticostriatal circuits: insights from Huntington's disease. Trends Cogn. Sci. 1998; 2:379–388. [PubMed: 21227253]

Lees A, Smith E. Cognitive deficits in the early stages of Parkinson's disease. Brain. 1983; 106:257–270. [PubMed: 6850270]

Liljenström H. Neural stability and flexibility: a computational approach. Neuropsychopharmacology. 2003; 28(Suppl 1):S64–73. [PubMed: 12827146]

Mishkin, M.; Malamut, B.; Bachevalier, J. Neurobiology of Learning and Memory. The Guilford Press; New York: 1984. Memories and habits: two neural systems.; p. 65-77.

Padoa-Schioppa C. Neurobiology of economic choice: a good-based model. Annu. Rev. Neurosci. 2011; 34:333–359. [PubMed: 21456961]

Redgrave P, Rodriguez M, Smith Y, Rodriguez-Oroz MC, Lehericy S, Bergman H, Agid Y, DeLong MR, Obeso JA. Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. Nat. Rev. Neurosci. 2010; 11:760–772. [PubMed: 20944662]

Rolls ET. The orbitofrontal cortex and reward. Cereb. Cortex. 2000; 10:284–294. [PubMed: 10731223]

Rolls ET, Thorpe SJ, Maddison SP. Responses of striatal neurons in the behaving monkey. 1. Head of the caudate nucleus. Behav. Brain Res. 1983; 7:179–210. [PubMed: 6830651]

Saint-Cyr JA, Ungerleider LG, Desimone R. Organization of visual cortical inputs to the striatum and subsequent outputs to the pallido-nigral complex in the monkey. J. Comp. Neurol. 1990; 298:129–156. [PubMed: 1698830]

Schneider W, Shiffrin RM. Controlled and automatic human information processing: I. Detection, search, and attention. Psychol. Rev. 1977; 84:1–66.

Seger CA. How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. Neurosci. Biobehav. Rev. 2008; 32:265–278. [PubMed: 17919725]

Selemon LD, Goldman-rakic PS. Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkey. J. Neurosci. 1985; 5:776–794. [PubMed: 2983048]

Shiffrin RM, Schneider W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychol. Rev. 1977; 84:127–189.

Shohamy D, Myers CE, Onlaor S, Gluck MA. Role of the basal ganglia in category learning: how do patients with Parkinson's disease learn? Behav. Neurosci. 2004; 118:676–686. [PubMed: 15301595]

Szabo J. Projections Nucleus from the in the Rhesus of the Monkey Caudate. Exp. Neurol. 1970:1–15. [PubMed: 4985789]

Szabo J. The course and distribution of efferents from the tail of the caudate nucleus in the monkey. Exp. Neurol. 1972; 572:562–572. [PubMed: 4631188]

Thorpe SJ, Rolls ET, Maddison S. The orbitofrontal cortex: neuronal activity in the behaving monkey. Exp. Brain Res. 1983; 49:93–115. [PubMed: 6861938]

Tremblay L, Schultz W. Relative reward preference in primate orbitofrontal cortex. Nature. 1999; 398:704–708. [PubMed: 10227292]

Vonsattel JP, Myers RH, Stevens TJ, Ferrante RJ, Bird ED, Richardson EP. Neuropathological classification of Huntington's disease. J. Neuropathol. Exp. Neurol. 1985; 44:559–577. [PubMed: 2932539]

Wood W, Neal DT. A new look at habits and the habit-goal interface. Psychol. Rev. 2007; 114:843–863. [PubMed: 17907866]

Wunderlich K, Dayan P, Dolan RJ. Mapping value based planning and extensively trained choice in the human brain. Nat. Neurosci. 2012; 15:786–791. [PubMed: 22406551]

Yamamoto S, Monosov IE, Yasuda M, Hikosaka O. What and where information in the caudate tail guides saccades to visual objects. The J. Neurosci. 2012; 32:11005–11016.

Yamamoto S, Kim HF, Hikosaka O. Reward value-contingent changes of visual responses in the primate caudate tail associated with a visuomotor skill. J. Neurosci. 2013; 33:11227–11238. [PubMed: 23825426]

Yasuda M, Yamamoto S, Hikosaka O. Robust Representation of Stable Object Values in the Oculomotor Basal Ganglia. J. Neurosci. 2012; 32:16917–16932. [PubMed: 23175843]

Yeterian EH, Van Hoesen GW. Cortico-striate projections in the rhesus monkey: the organization of certain cortico-caudate connections. Brain Res. 1978; 139:43–63. [PubMed: 413609]

Yin HH, Knowlton BJ. The role of the basal ganglia in habit formation. Nat. Rev. Neurosci. 2006; 7:464–476. [PubMed: 16715055]
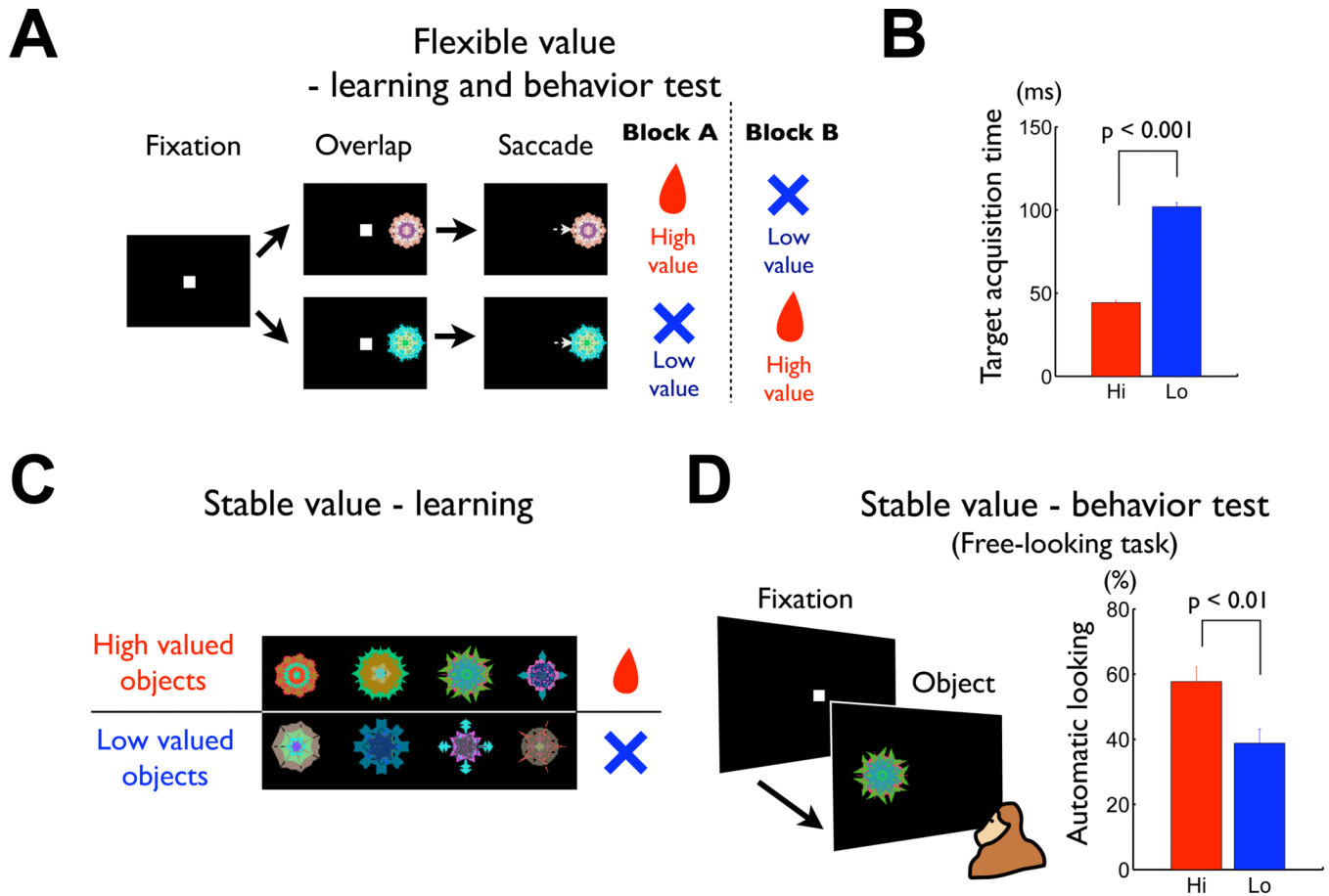
Figure 1. Learning and behavioral testing procedures for flexible and stable values
(A-B) Flexible value procedure. **(A)** Sequence of events in the flexible object-value association task. While the monkey was fixating on a central white dot, one of two fractal objects was presented at the neuron's preferred position. After 400 ms the fixation spot disappeared and the monkey was required to make a saccade to the object. In a block of 30-40 trials a liquid reward was delivered after the saccade to one object, but not the other object. The object-reward contingency was reversed in the next block. **(B)** Average target acquisition time, defined as the time from the offset of the fixation spot to the time when the eye position entered the object area (Hi, high-valued object; Lo, low-valued object). **(C-D)** Stable value procedure. **(C)** An example set of fractal objects that were associated with stable values. During long-term learning, the upper four objects were always associated with a reward (high-valued objects) while the lower four were always associated with no-reward (low-valued objects). **(D)** Free-looking task for testing saccades to objects with stable values. On each trial one of the eight objects in one set (as in **(C)**) was presented 100 or 200 ms after the fixation spot disappeared, and the monkey was free to look at it without any reward outcome (left). After long-term learning the monkey was more likely to look at the high-valued objects (right). See also Figure S1 and S2.
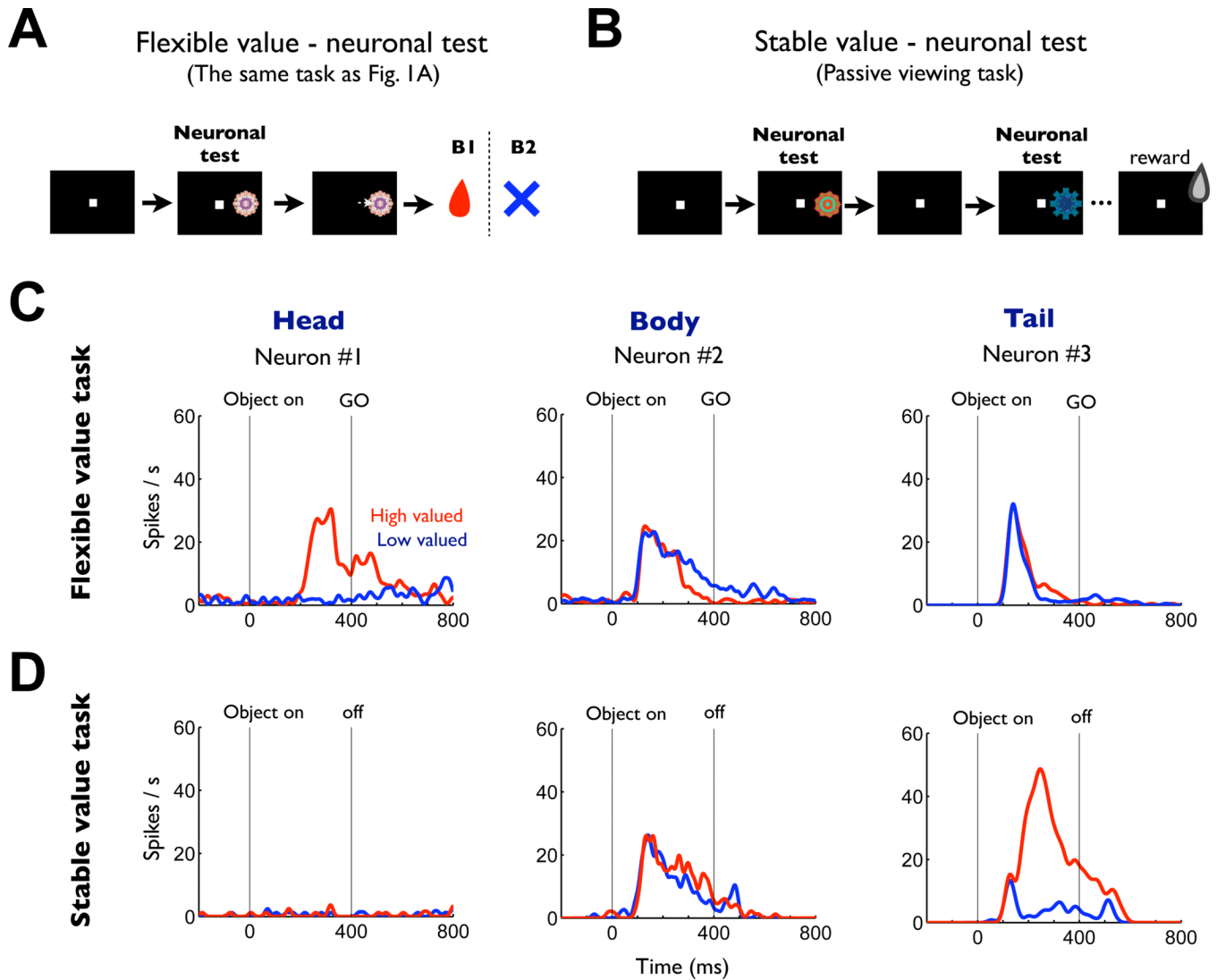
Figure 2. Flexible and stable value coding of the same three neurons in the caudate nucleus
(A) Procedure for testing neurons' responses to objects with flexible values (same as in Figure 1A). **(B)** Procedure for testing neurons' responses to objects with stable values. While the monkey was fixing on a central white dot, 2–6 fractal objects were chosen pseudorandomly from a set of 8 objects and were presented sequentially in the neuron's preferred location (each for 400 ms). The monkey was rewarded 300 ms after the final object disappeared. **(C)** Responses of three representative neurons recorded in the caudate head, body and tail to high-valued objects (red, 24 trials) and low-valued objects (blue, 24 trials) during the neuronal test period of flexible value procedure. They are shown by spike density functions (SDFs), which are aligned at the onset of object presentation. **(D)** Responses of the same three neurons shown separately for high (red, 32 trials) and low valued objects (blue, 32 trials) during the neuronal test period of stable value procedure.
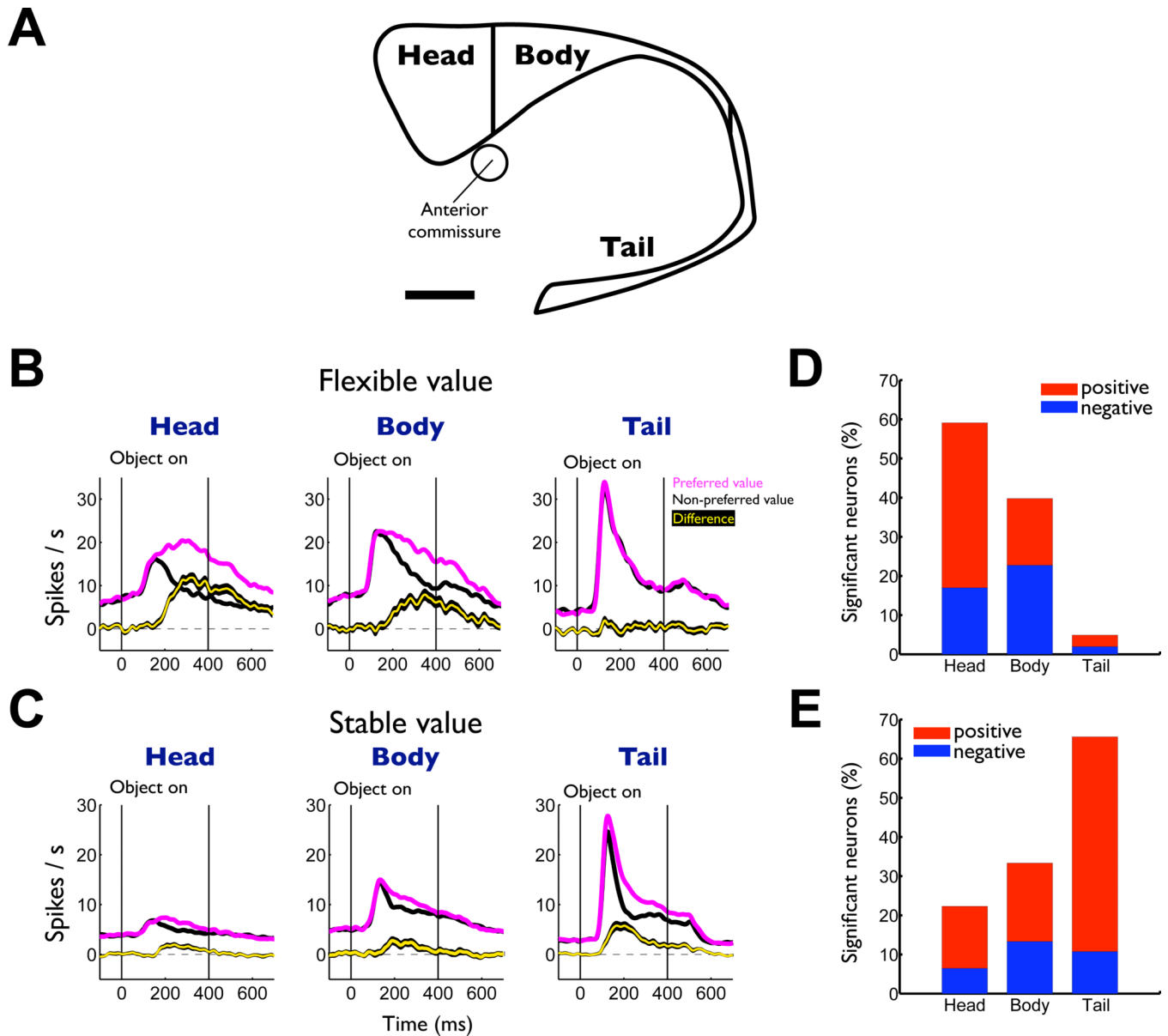
Figure 3. Differential encoding of flexible and stable values in subregions of the caudate nucleus
(A) MRI-reconstructed image of caudate nucleus and its subregions. We delineated the subregions of the caudate nucleus by the anterior commissure (head-body) and the genu (body-tail). Scale bar indicates 5 mm. **(B)** Average responses to the flexibly valued objects of neurons in the head (n = 159), body (n = 88) and tail (n = 102) of the caudate nucleus. Neuronal responses were averaged for the neurons' preferred values (magenta) and non-preferred values (black) using a cross-validation method (see supplementary information). The yellow line indicates the difference between the preferred and non-preferred responses (mean ± SE). **(C)** Average responses to the stably valued objects in the three caudate subregions (head, n = 129; body, n = 86; tail, n = 92). **(D)** Proportions of flexible value coding neurons in the three caudate subregions. Red: positive neurons that responded more strongly to high-valued objects. Blue: negative neurons that responded more strongly to low-valued objects. **(E)** Proportions of stable value coding neurons. See also Figure S3 and S4.
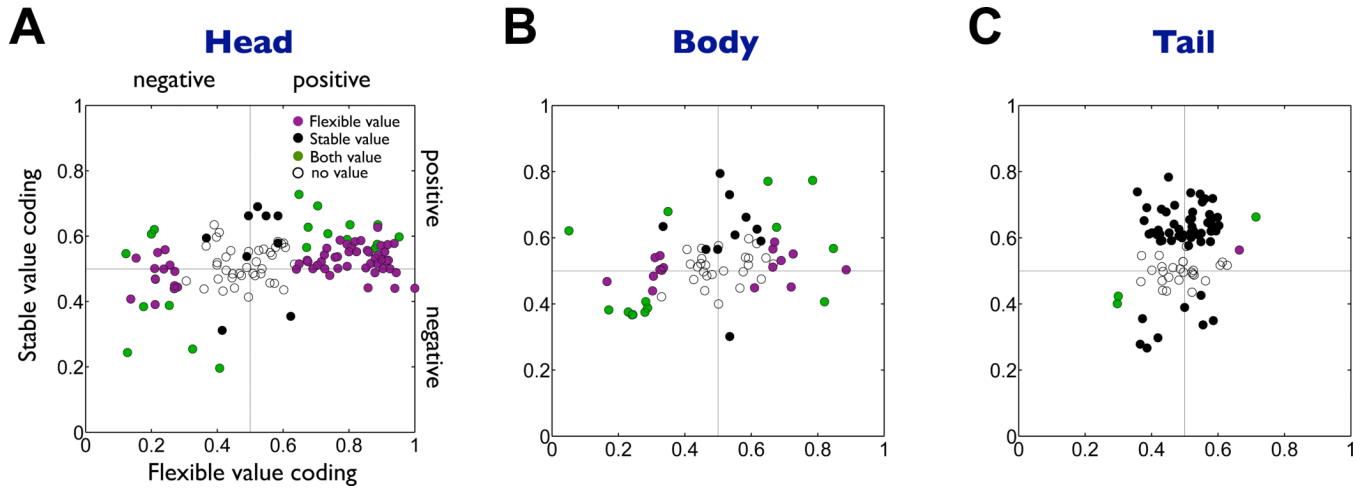
Figure 4. Flexible and stable values by individual neurons in the caudate subregions
(A-C) Comparison between flexible value coding (abscissa) and stable value coding
(ordinate) in the caudate head (n=125), body (n=65) and tail (n=87). Plotted for each neuron
(each dot) are the magnitudes of the flexible and stable value coding measured by ROC
areas. ROC areas higher and lower than 0.5 indicate positive and negative value coding,
respectively. Purple: neurons encoding only flexible values. Black: neurons encoding only
stable values. Green: neurons encoding both flexible and stable values. White: neurons
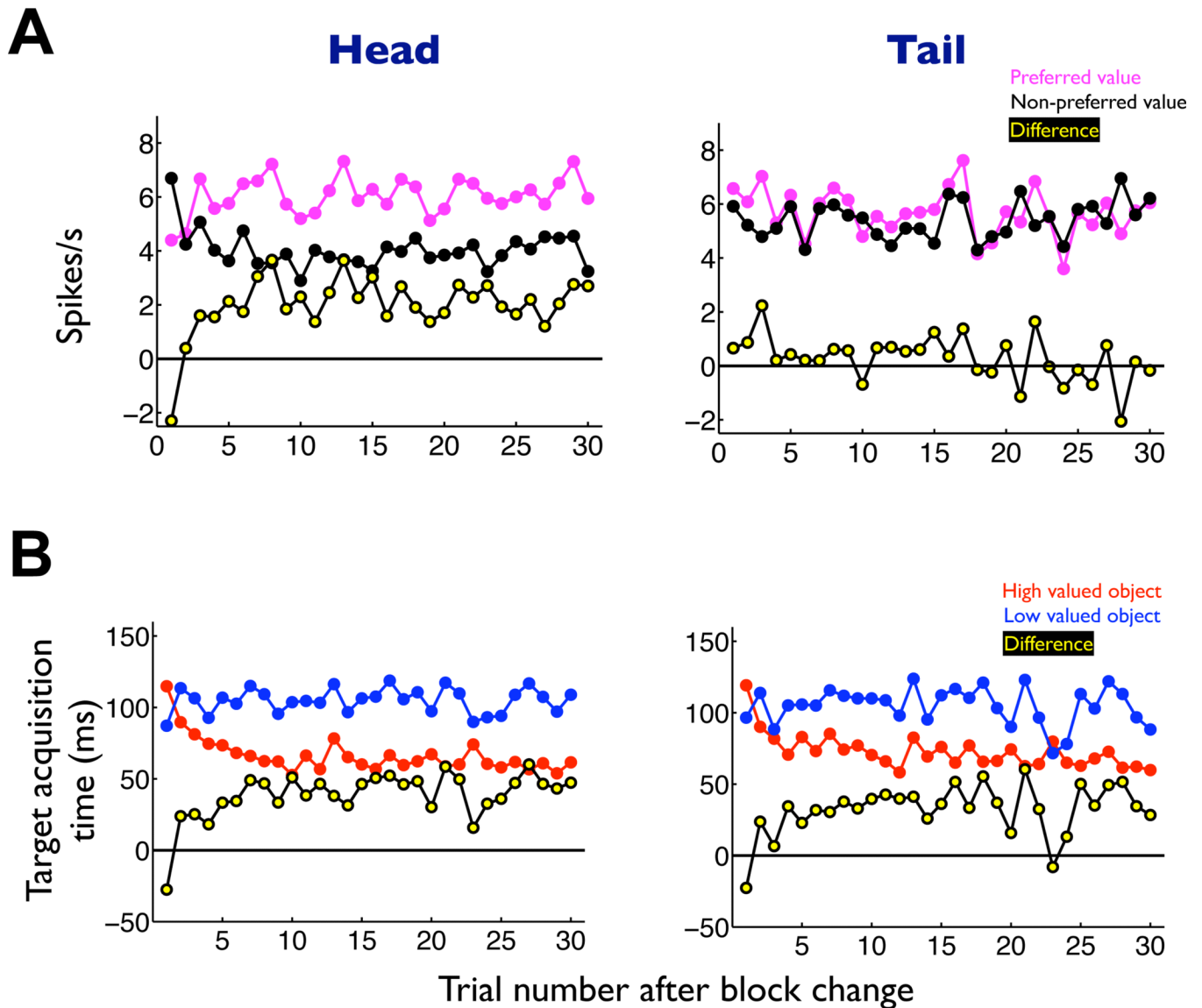encoding neither value.

Figure 5. Time courses of the neuronal and behavioral responses in the flexible value procedure, shown separately for the caudate head and tail

(A) Neuronal responses plotted against the number of trials after the reversal of the object-reward contingency. The responses were averaged for the neurons' preferred values (magenta) and non-preferred values (black). Yellow dotted line indicates the difference between the preferred and non-preferred responses. **(B)** Behavioral responses while the neuronal responses were obtained in the caudate head and tail. The target acquisition times were averaged for the high valued object (red) and low valued object (blue). Yellow dotted line indicates the difference of target acquisition time between the high and low valued objects.
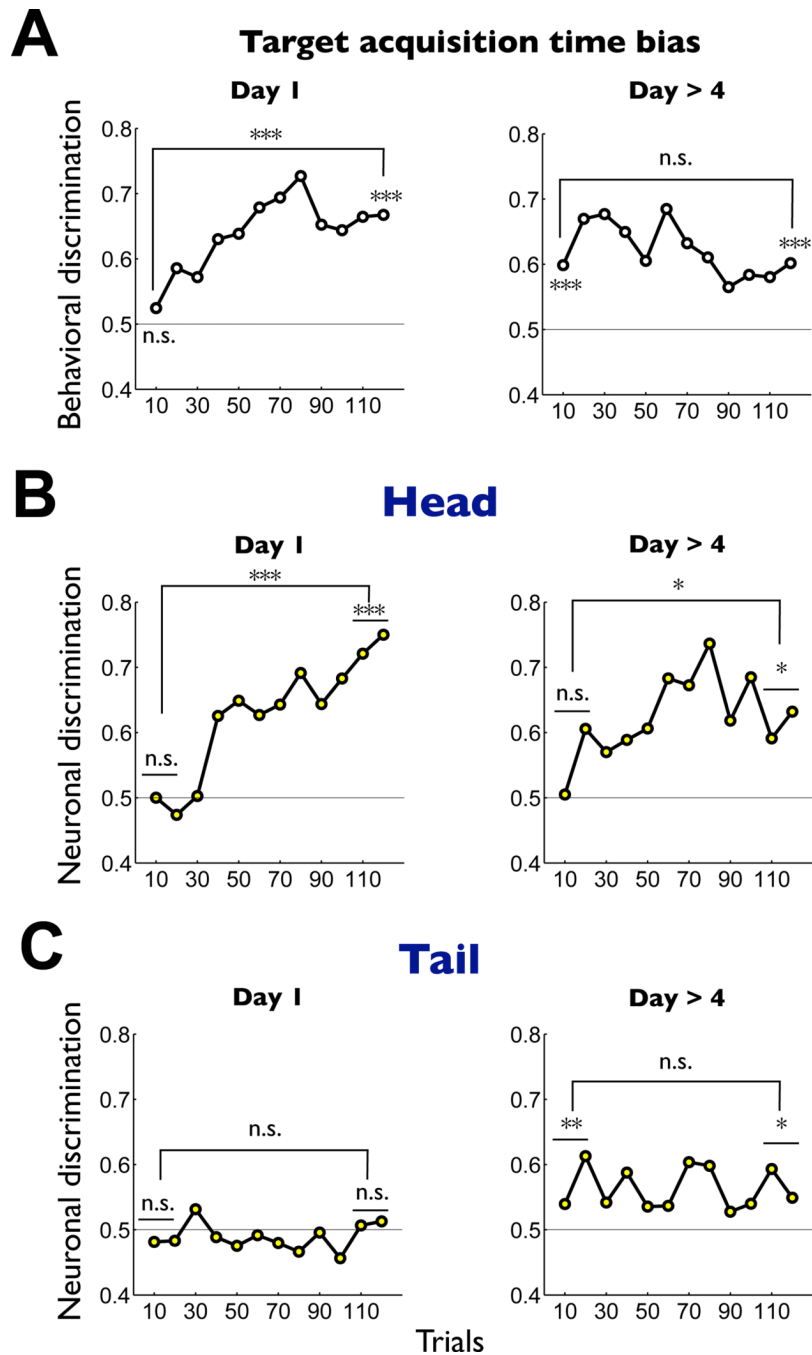
Figure 6. Changes in the neuronal and behavioral responses during the combination of short-term and long-term learning of object values
The monkey performed the object value learning task (Figure S2) in two stages of long-term learning: 1) initial stage where novel objects were used as the saccade target (Day 1), and 2) late stage where objects had been learned more than 4 times (Day > 4). **(A)** Behavioral discrimination based on the target acquisition time. **(B)** Caudate head neuronal discrimination (Day 1, n = 18; Day > 4, n = 23). **(C)** Caudate tail neuronal discrimination (Day 1, n = 29; Day > 4, n = 21). Each graph shows the short-term changes in the discrimination within a session of the object value learning (number of trials: 120). Each data point indicates the discrimination in a sub-block of 10 trials, computed as an ROC area.

For the behavioral data (A) an ROC area larger than 0.5 indicates a stronger preference of high-valued objects. For the neuronal data (B and C) the ROC areas were normalized so that the neuron's preferred value is plotted upward. Specifically, if the original ROC areas showed a negative slope during learning, the values were flipped with respect to 0.5 (i.e., neutral value). The statistical significance of the discrimination was tested for the early and late stages of the short-term learning (1 sub-block = 10 trials for behavioral discrimination, 2 sub-blocks = 20 trials for neuronal discrimination) in two ways: 1) comparison with the neutral value (ROC 0.5), and 2) comparison between the early and late stages. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, n.s. non-significant.
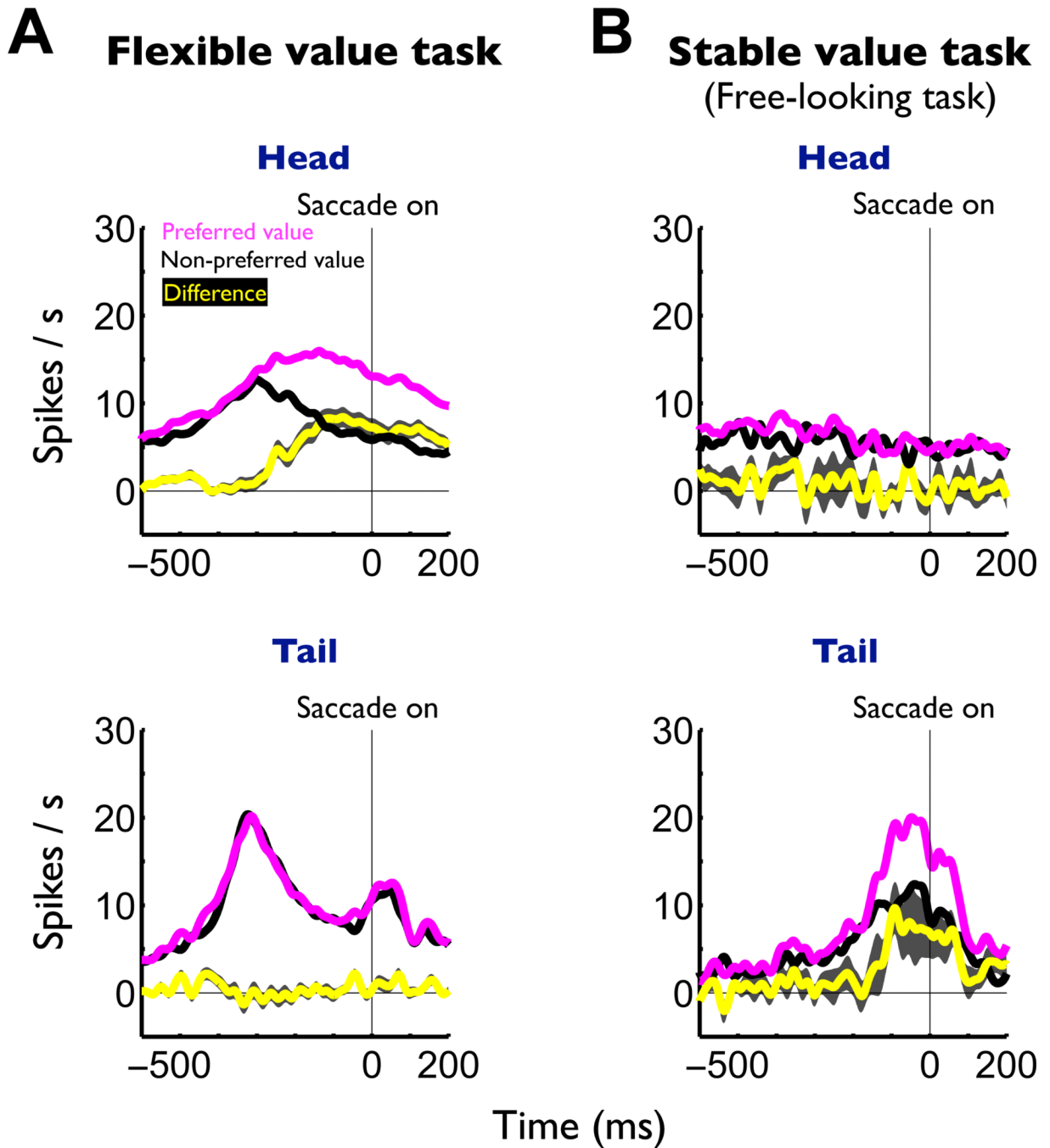
Figure 7. Presaccadic neuronal activity in the flexible and stable value procedure, shown separately for the caudate head and tail

(A) Averaged neuronal activity aligned on the onsets of saccades to preferred values (magenta) and non-preferred values (black) in the flexible value procedure (Figure 1A). The yellow line indicates the difference in neuronal activity between the preferred and non-preferred values (caudate head, n = 159; tail, n = 102; mean ± SE). (B) Averaged neuronal activity aligned on the onsets of saccades to preferred values (magenta) and non-preferred values (black) in the stable value procedure (free-looking task, Figure 1D). The yellow line indicates the difference in neuronal activity (caudate head, n = 20; tail, n = 20; mean ± SE). See also figure S5.

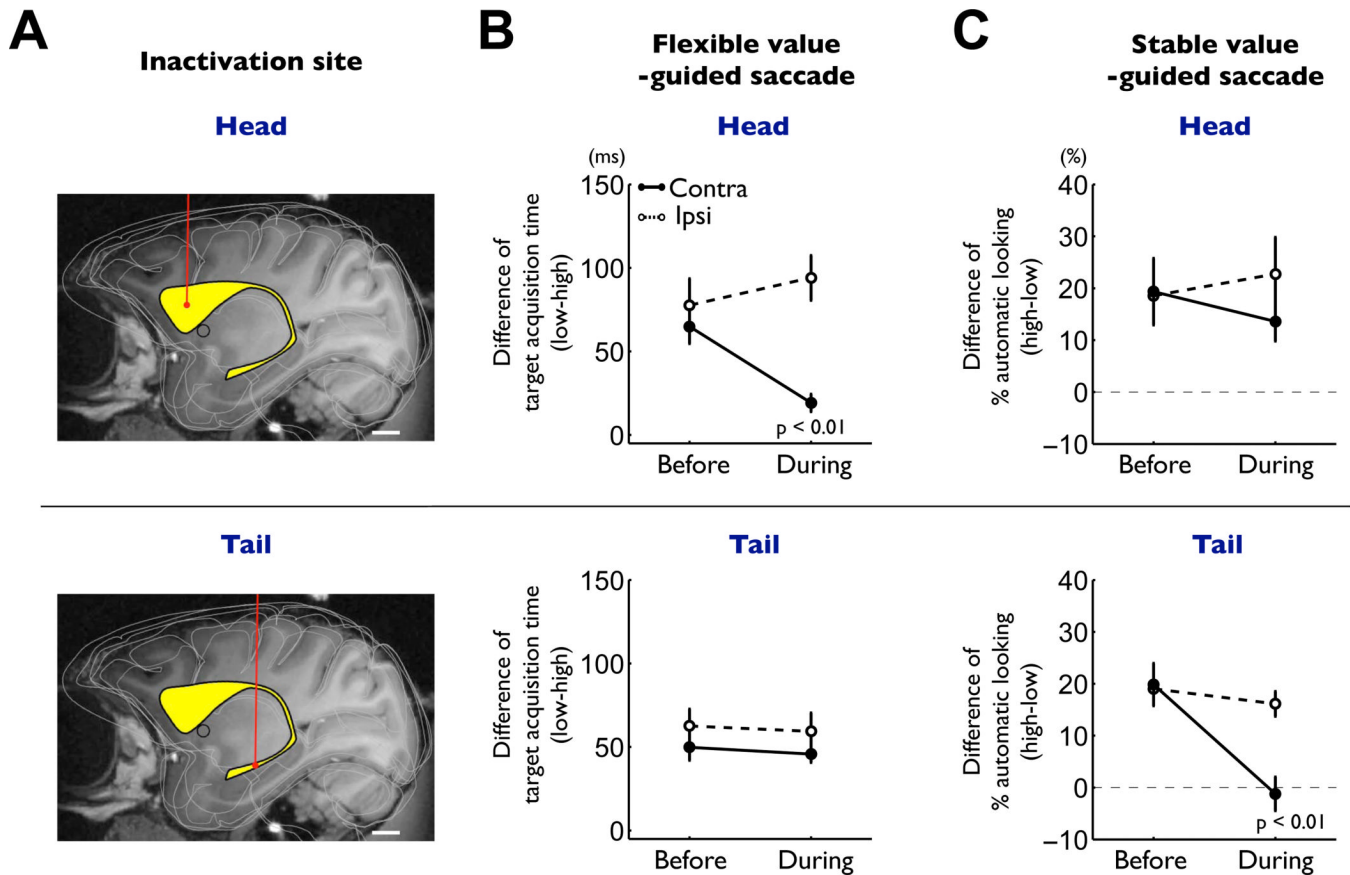**Figure 8. Differential impairments of flexible and stable value-guided saccades by caudate head and tail inactivations**

(A) Example injection sites of muscimol in the caudate nucleus (yellow structure) reconstructed on an MR image: caudate head (top) and tail (bottom). Scale bar indicates 5 mm. **(B)** Effects on the controlled saccades in the flexible value procedure (Figure 1B). The differences in the target acquisition time between high- and low-valued objects are plotted before and during inactivation (mean ± SE). Data are shown separately for caudate head inactivation (top, n = 6) and tail inactivation (bottom, n = 9), and for contralateral saccades (solid line) and ipsilateral saccades (dashed line). **(C)** Effects on the automatic saccades in the stable value procedure (free-looking task, Figure 1D). The differences in the probability of automatic looking between high- and low-valued objects are plotted before and during inactivation (mean ± SE). Same format as in (B). See also figure S6-8.