



Published in final edited form as:

*Curr Biol.* 2018 July 23; 28(14): 2206–2217.e5. doi:10.1016/j.cub.2018.05.017.

## Prefrontal cortex represents long-term memory of object values for months

Ali Ghazizadeh<sup>1,2,3</sup>, Simon Hong<sup>1,4</sup>, Okihide Hikosaka<sup>1,5</sup>

<sup>1</sup>Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health Bethesda, MD 20892, USA

<sup>2</sup>Electrical Engineering Department, Sharif University of Technology, Tehran, 11365-11155 Iran

<sup>3</sup>Brain Engineering Center and School of Cognitive Sciences, Institute for Research in Fundamental Sciences, Tehran, 19395-5746, Iran

<sup>4</sup>Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>5</sup>National Institute on Drug Abuse, National Institutes of Health, Baltimore, MD 21224, USA

### Summary

As a central hub for cognitive control, prefrontal cortex (PFC) is thought to utilize memories. However, unlike working or short-term memory, the neuronal representation of long-term memory in PFC has not been systematically investigated. Using single-unit recordings in macaques, we show that PFC neurons rapidly update and maintain responses to objects based on short-term reward history. Interestingly, after repeated object-reward association, PFC neurons continue to show value-biased responses to objects even in the absence of reward. This value-biased response is retained for several months after training and is resistant to extinction and to interference from new object reward learning for many complex objects (>90). Accordingly, the monkeys remember the values of the learned objects for several months in separate testing. These findings reveal that in addition to flexible short-term and low-capacity memories, primate PFC represents stable long-term and high-capacity memories, which could prioritize valuable objects far into the future.

### In Brief:

In this article, Ghazizadeh et al, present a systematic investigation of value memory in ventrolateral prefrontal cortex neurons in macaque monkeys lasting from a few minutes to several months. This value memory is shown to be robust against forgetting and interference and allows animals to find many valuable objects long after reward learning.

---

Corresponding author and lead contact: Ali Ghazizadeh, Electrical Engineering Department, Sharif University of Technology, Azadi Avenue, Tehran, 11365-11155, Iran, aliaghazizadeh@gmail.com.

Author contributions:

AG and OH designed the experiment. AG, SH and OH contributed to data collection. AG analyzed the data. AG and OH wrote the first draft of the manuscript. All authors contributed to editing and revising the manuscript.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Declaration of interest:

The authors declare no competing financial interests.

## Keywords

Prefrontal cortex; Object values; Long-term memory; High-capacity memory; Single-unit recording; Macaque monkey

---

## Introduction

Many animals including invertebrates [1] and vertebrates [2] are capable of rapidly learning values associated with arbitrary stimuli. Such short-term adaptability is thought to be critical when encountering stimuli in novel environments. In primates, PFC is thought to underlie various forms of visual reward association learning such as conditional visuomotor learning [3,4], concurrent visual discrimination[5,6] and object-reward association[7]. A popular concept is that PFC works flexibly so that the animal can think and behave intelligently[8,9]. Such flexibility is based on the comparison of recently acquired information and behavioral goals and therefore is based on short-term memories. In this scheme, PFC can rapidly learn object reward associations to meet current task demands, whereas long-term retention and consolidation of reward association could depend on other brain areas. Indeed, such a dichotomy between new learning and long-term storage of value memories is recently reported in basal ganglia [10,11] and resembles standard consolidation theory of episodic memories in that different structures are involved in initial memory formations and long-term storage[12]. These data may suggest that PFC is specifically involved in one side of the dichotomy: short-term working memory.

On the other hand, there is also evidence that implicates PFC in encoding and retrieval of certain memories in longer time scales[13–15]. With regards to reward, long-term memory is critical ecologically for activities, such as foraging, when cues associated with food are not encountered for protracted periods due to seasonal availability. On the maladaptive side, reward-based long-term memory may underlie behaviors such as drug addiction. However direct evidence regarding the neural representation of reward-based long-term memories across a wide time range in primate PFC is currently lacking.

We have recently shown that PFC connects with both short and long-term value memory networks in basal ganglia[16]. Furthermore, using fMRI, we recently found signatures of reward-based long-term memory in PFC[17]. We thus hypothesize that not only PFC neurons should be sensitive to reward learning with new objects but should also reflect previously learned associations at longer memory periods. Interestingly, our results show that this is indeed the case. We found that PFC responses encoded both reward-based short-term and long-term memories for several months with a high capacity for many complex objects.

## Results

To examine the retention of learned values, we recorded the activity of single neurons in and ventral to the principal sulcus (ventrolateral PFC areas 8Av, 46v and 45 [vlPFC], Figure S1) in two monkeys (monkeys B and R with 159 and 191 neurons in left and right hemispheres, respectively, Table 1). Each monkey viewed fractal objects that were repeatedly paired with

low or high rewards in a biased reward training task (Figure 1A left), dividing the objects into “bad” and “good” categories, respectively. The fractals trained with reward were later tested in a passive viewing task to probe reward-based memory signals while recording neurons in vIPFC (Figure 1A, right).

To examine the retention of learned values in long-term memory, we tested the neurons’ responses to good and bad objects, hours, days, weeks or months after they were last trained with reward (Figure 1B, left). During the memory test, good and bad objects were intermixed and shown one at a time and without reward in the neuron’s receptive field (mapped prior to memory test, Figure S2, see STAR Methods). Animals were rewarded for maintaining fixation after random intervals independent of objects. We utilized many random fractals (~100) for each monkey in value memory test (Figure 1C). This was to ensure that neural responses could not be attributed to idiosyncratic features in each stimulus and also to gauge the capacity of the long-term memories.

During the memory retention period for a given set of objects, reward association was not repeated for that set, but reward training with other objects was performed (Figure 1B–C, >230 novel fractals not used in memory test as well as previously trained fractals). The training with novel objects was done in part to examine the stability of value memories in vIPFC in the presence of new object-reward learning (retroactive interference). But in addition, this reward training, allowed us to examine reward learning signals in individual vIPFC neurons for novel objects in their first day of training and for previously trained objects after multiple days of training (Figure 1B, right). The same strategy was used to test memory at multiple scales in individual neurons despite the fact that recordings were acute. A neuron could be tested with fractals that were trained months ago in one block and with fractals that were trained hours or days ago in the next block in the same session (Figure 1B, right).

### **Object value encoding emerges and persists across reward training trials in vIPFC**

First, we examined the learning of object values in vIPFC neurons and their short-term value retention (memory across trials) during learning (Figure 2). Figure 2A shows two example neurons that had similar firing to good and bad objects in the beginning of training, but developed differential firing to good and bad objects by the end of the first training session (one with higher firing to good objects referred to as good-preferring neuron and one with higher firing to bad objects referred to as bad-preferring neuron). Value preference of neurons was determined at the end of the training session (with cross validation, see STAR Methods). Across the population, similar learning was observed when comparing the first five and last five reward pairings with an object in the first day of training (Day 1) and after >10 days of training (using previously trained objects) (Figure 2B). The difference in response to good and bad objects emerged rapidly within the first few trials and then grew with a slower rate (Figure 2C, Day 1). After multiple days of value training, the response difference became larger (Figure 2C, Day >10), indicating that the slower learning continued to enhance response differences across days. Interestingly, this differential firing to objects emerged mainly from a reduction of firing to the non-preferred object values

while the response to the preferred value remained unchanged across the population on average (Figure 2D).

Average discriminability of good and bad objects (value AUC, see STAR Methods) among vIPFC neurons was already significant within the first 5 trials and continued to grow across trials and days (Figure 2E). These results confirm previous findings about rapid value learning in vIPFC, but also reveal a second slower time course that can take many days to fully develop. The differential response to good and bad objects did not change within a training session after 10 days of training (Figure 2C, main effect of trials after 10 days  $F_{4,165}=0.3$ ,  $P > 0.8$ , 1<sup>st</sup> compared to 2–5<sup>th</sup> trials  $t_{33}=1.7$   $P=0.09$ ). Animal behavior measured by sporadically added choice trials also showed rapid learning within a few trials in the first day followed by a slower phase of improvement afterwards (Figure S3A–B).

To examine the retention of learned values during training, the discriminability of good and bad objects for objects seen 1–4 trials ago was compared to objects seen 5–8 trials ago. During the intervening trials for a given object, other objects in a set of 8 were trained with reward. Population average showed retention of learned values despite the intervening trials (Figure 2F–I). This was true even on the first day of training despite overall smaller AUC values (Figure 2F–G) compared to the AUC values after multiple days (Figure 2H–I). Accordingly, animal's choice of good objects was not affected by the number of intervening trials for objects seen 1–4 trials ago compared to 5–8 trials ago in both the first and after >10 days of training (Figure S3C–D).

### Encoding of object values persists in vIPFC from minutes to months

Next, we tested the reward-based long-term memory after >10 days of training when the differential response to good and bad objects seemed to be fully developed (Figure 2C). Value memory was tested by examining neuronal activity during passive viewing of objects that were trained in the same day (minutes to hours ago) or days, weeks and months before (Table 1). Some neurons were tested in more than one memory period by exposing the neuron in the same recording session to objects with different time lapses after that last training. Figure 3A shows two such example neurons each tested with objects trained days before and months before. Both neurons were excited more strongly by good compared to bad objects (good-preferring neurons), showing reward-based long-term memory days after training (Figure 3A days-later). Importantly, this reward-based long-term memory was also observed for other objects that were seen more than 3 months ago (Figure 3A months-later). Similarly, neurons with higher firing to bad compared to good objects (bad-preferring neurons) retained value memories for objects trained days and months before (Figure S4). The population average using the neurons' value preference (cross validated, see STAR Methods) showed a remarkably similar pattern of activity and significant value discrimination across the four memory periods tested (Figure 3B). The firing difference to good and bad objects, its rise time and duration was virtually unchanged by passage of time (Figure 3C).

To quantify the effect of value memory on object responses in vIPFC, we calculated AUCs for discrimination of good and bad objects (value AUCs) for each neuron during passive viewing. The overall distribution of value AUCs across all neurons showed a

significant shift toward preference for good objects (higher firing to good compared to bad) (Figure 4A). A large percentage of neurons showed significantly positive AUC values (46% good-preferring), but some neurons showed significantly negative AUC values (13% bad-preferring). Notably, except for differences in average firing pattern to good and bad objects (Figure 4B), good- and bad-preferring neurons were not distinguishable based on various physiological criteria including spike shape, basal firing rate, visual and value differentiation onsets and response variability across objects (Figure S5).

Across the vIPFC population, the value memory significantly increased the pairwise discriminability between good and bad objects compared to discriminability of objects within good and bad categories (Figure 4C). However, no relationship between within and between value category discrimination among neurons was observed beyond what is expected from chance (see STAR Methods). This suggests that the degree of object selectivity in a neuron was not predictive of its sensitivity to value memory.

Finally, and consistent with data shown in Figure 3B–C, the average value AUC was stable across memory periods for many months across the vIPFC population (Figure 4D, left). The stability of value memory was also observed separately in good- and bad-preferring neurons (i.e. similar AUCs across memory periods). The percentages of good- and bad-preferring neurons in the population were also stable across the memory periods (Figure 4D, right). Similar results were observed separately for each monkey (Figure S6A–D). Across the population, object selectivity was relatively low and overall did not change across memory periods, regardless of neuron value preference (good- or bad-preferring) or object categories (good or bad). This suggests that value discrimination across memory periods did not become sparsely focused on a few objects (i.e. no significant increase in sparsity).

Importantly, retention of value discrimination was also observed within neurons tested with multiple memory periods (Figure S7A–B). For this population, the value AUCs in earlier and later memory periods were not significantly different and had the same sign (i.e. good- or bad-preferring in both periods) for the majority of neurons (>89%; 69 out of 79 neurons, Figure S7B). Such sustained value memory was also observed in neurons that were additionally tested in value learning task (Figure S7C–D). For this group of neurons, the average value AUC during passive viewing task tended to be somewhat lower (not significant) than AUC during the training task (Figure 2E vs Figure S6C: training AUC=0.77 vs passive viewing AUC=0.71 n=34, difference  $t_{33}=1.8$  P=0.07). Nevertheless, the value discriminability during passive viewing was not sensitive to passage of time for several months after training (Figure S7C). The majority of the neurons tested in value learning (67%; 23 out of 34 neurons) showed significant value learning in training task and significant value memory during passive viewing. This suggests that the same group of neurons in vIPFC encode reward-based short-term (across training trials) memory as well as long-term memory. Furthermore, the fact that the population value AUC was well-sustained across the memory periods, despite monkeys being engaged in reward training with many other objects (>230 objects), reveals the formation of stable long-term memories that were resistant to retroactive interference.

### **Value encoding of objects in vIPFC is resistant to extinction**

Repeated exposure of reward-predicting cues without reward can result in the extinction of value signals[18]. Since the passive viewing task consisted of repeated presentation of a given object without contingent reward, the stability of value signals in vIPFC neurons against extinction could be tested as well. Consistent with reports on repetition suppression [19], there was a reduction in overall firing as a function of number of exposures for a given object. However, the differential firing to good and bad objects remained stable across trials and there was no decrease in average value AUC (Figure 4F). This was the case separately for bad- and good-preferring neurons averaged across monkeys (Figure 4G) and also separately for each monkey (Figure S6E–F). Such resistance to extinction is previously seen in over-trained habitual behaviors [20] and is consistent with the development of object skills created by long-term training with reward[21,22].

### **vIPFC can encode and retain values of a large number of objects**

In order to ensure that the value signals present in vIPFC population were not driven by only a few objects that were well-remembered, we examined the average response to all the objects in each monkey (52/52 and 48/48 good/bad fractals in monkeys B and R respectively). In both monkeys, the vIPFC population response (combined across memory periods) to good objects was larger than bad objects consistent with a high capacity mechanism (Figure 5A). Across all the objects used, the population response in vIPFC could be used with high fidelity to discriminate good and bad objects (Figure 5A inset, 0.99 and 0.95 AUC in monkeys B and R, respectively  $p < 1e-10$ ). The number of objects that were correctly classified at the optimal threshold by the ideal observer was 100 out of 104 in monkey B and 85 out of 96 in monkey R (see STAR Methods). This suggests that representation of value memory in vIPFC has high capacity for a large number of objects (high-capacity memory). This memory capacity can also be verified by looking at the average AUC of a given good object vs other bad objects and vice versa across the population separately for neurons with positive and negative AUCs (determined from Figure 4A  $G > B$  or  $B > G$ ). This analysis further confirmed that value memory in both good- and bad-preferring neurons is not driven by a few objects rather a clear separation of AUCs between good and bad objects is observed for a large number of objects (Figure 5B).

### **Intact behavioral memory of object values from days to months**

Given the retention of learned object values in vIPFC for months, we predicted that monkeys should remember learned values long after last reward learning. To test this behaviorally, free viewing gaze bias toward good objects [23,24] was used as an index of memory strength[25]. The persistence of gaze bias was tested across memory periods while reward-learning with other objects was underway (Figure 6A–B). Consistent with neural retention of learned value, behavioral results showed a significant gaze bias toward good objects that was unchanged from days to months in both monkeys (Figure 6C).

## **Discussion**

We asked whether and how long-term memory of object values is represented in the ventrolateral prefrontal neurons. Prefrontal cortex is often associated with working memory

that has a short-term and low-capacity storage[26,27]. Despite evidence for a role in encoding and retrieval of certain long-term memories[14,15,28], the neuronal representation of value memory in the vIPFC across a wide time range from a few trials to several months was not systematically investigated. Our results showed that vIPFC neurons not only rapidly learn and maintain values in short-term memory across the training trials (Figure 2) but also retain them for long periods of up to a few months and for a surprisingly large number for visual objects (Figure 1, 3–5). Value memories were found to be resistant to interference from reward learning with other objects and to extinction from repeated exposures of learned objects without reward (Figure 4). Such resistance to extinction following long-term reward training suggests that the memory signals in vIPFC are related to habits and skills rather than mediating a goal-directed function [20–22]

Notably while short-term and long-term memories of object values are encoded by distinct regions in basal ganglia [29–31], they are found in the same region and within the same neurons in the vIPFC (Figure 2, Figure S7). It is possible that the value learning and memory signals in vIPFC are mediated by basal ganglia circuitry. Specifically, recent studies in our lab have shown that posterior basal ganglia (pBG), which includes caudate tail (CDt) and caudal-lateral substantia nigra reticulata and compacta (clSNr and clSNc), encode long-term value memories of objects days after training [10,32,33]. Interestingly, the vIPFC region investigated in this study is shown to be targeted by clSNr (the pBG output) via thalamus[34]. This puts vIPFC within the network that processes and stores reward-based long-term memories. On the other hand, vIPFC also projects heavily to caudate head (CDh) [16] which is known to be sensitive to recent reward histories[10] and receives feedback from CDh via basal ganglia thalamocortical loops[34]. This puts vIPFC in a prime position to be informed of both recent and old reward histories for objects. Encoding both reward-based short and long-term memories, enables vIPFC to play a key executive role in real life when both recent and old memories should be taken into account for optimal decision making.

vIPFC is also reciprocally connected with the inferior temporal cortex (IT) [35,36] and these connections contribute to object discrimination learning and memory [13,37,38]. Since IT cortex is the main input source to CDt [39], it is possible that reward-based long-term memories may also be found in responses of IT neurons. However, it is likely that due to higher object selectivity in IT compared to PFC [40–42](and see Figure S5), the responses in the latter is better suited for discrimination of objects based on learned values. Indeed, our results showed that the value discriminability in vIPFC was independent of the modest object selectivity present in this region (Figure 4C). Nevertheless, our results are consistent with our recent findings using fMRI that implicated both vIPFC and IT cortex and their connectivity in discrimination of valuable objects in long-term memory[17].

The fact that single vIPFC neurons were responding similarly within bad and good object categories after long memory periods, may be related to previous findings that implicated PFC in learning object categories[43,44]. Indeed, it is possible that the act of associating high reward with some objects and not others resulted in implicit categorization of object into two categories[45]. If so the sustained responses in vIPFC may be better interpreted as category memory rather than value memory. However, the fact that good-preferring neurons

largely outnumbered bad-preferring neurons in the current experiment argues against such interpretation (Figure 4A–B). This is because in a pure category learning there is no a-priori reason for one category to be over-represented as a preferred response among the neuronal population as often observed in perceptual categorization[43]. A higher proportion of reward preferring neurons in PFC is also reported previously[45]. Thus, we believe that a simple categorization framework does not capture the biased preference for valuable objects in vIPFC.

We also note that while behavioral and neural value learning both showed a rapid rise within the first five trials in the first day of learning (Figure 2 and Figure S3), behavioral learning seemed to almost plateau in the first day while vIPFC value discrimination continued to improve across days. One possibility is that the mapping of value discrimination in vIPFC to choice behavior is non-linear. This is likely to be true since neural discrimination of good and bad objects (value AUC) can be combined across many neurons for highly accurate choice of good objects even when value AUC per neuron is modest. In this case increasing the AUC per neuron beyond a certain point will have diminishing effects in choice rate that is already close to 100%. Alternatively, the improvement in value discriminability in vIPFC across days may be used for processes that go beyond simply knowing which object to choose. For instance, we recently showed that biased object-reward association can create a value pop-out for efficient visual search but only after training is repeated for many days and long after choice plateaued at 100%[22]. Given the prominent vIPFC projections to superior colliculus (SC)[46], the gradual increase in AUC after multiple days could lead to development of value pop-out and rapid gaze orientation toward good among bad objects. Notably, discrimination of objects by values in long-term memory happened rapidly (~100ms) in vIPFC (Figure S5A). Such rapid detection of valuable objects not only helps the animal to obtain rewards but to obtain them more frequently (i.e. maximizing rewards per unit time), an “object skill” that is important for competitive fitness[21]. The strong value memory developed across days in vIPFC can also interact with the short-term feature based attention[42] for rapid detection and prioritization of good objects[22]. One mechanism for such interaction could be via a top-down modulation to bias sensory processing in the early parts of ventral stream[47].

While there is accumulating evidence on the role of vIPFC [48–50] and more generally frontal cortexes[51–53] in value-based learning and decision making across monkeys and humans[54], such involvement may generalize to other motivationally salient but non-rewarding situations. Indeed, previous studies have shown that many vIPFC neurons encode short-term memory for motivationally salient outcomes[52,55]. Notably, Kobayashi et al[55] found that short-term memories for appetitive and aversive outcomes are encoded by partially non-overlapping populations in lateral PFC. This suggests that short-term memory of value and salience could independently co-exist in vIPFC. We have recently shown that days-long memory arising from distinct experiences such as reward, aversiveness, uncertainty and novelty, can similarly modify object salience and division of attention between multiple objects[24]. These results predict the existence of a long-term memory mechanism for non-rewarding dimensions of learned salience. This may turn-out to be encoded by single vIPFC neurons or by different group of neurons or regions that also influence downstream areas such as SC to guide gaze and attention toward important



objects. Indeed, the activations seen in vIPFC could mediate or be mediated by attention or emotional responses to good objects. Examination of the long-term memory signals across such different dimensions in vIPFC can also help determine the degree to which the activations seen here are related to attentional processes rather than value per se [56]. Nevertheless, we note that even in that case the attentional processes will be contingent on the long-term retention of learned stimulus value, which is a form of associative memory.

Finally, while recent findings point to PFC as a likely candidate for storage of long-term memories[57], our results do not rule out the possibility that the observed reward-based long-term memory signals in vIPFC may be partially or fully routed to it by other structures. Previous findings suggest that such object-outcome memories are likely to be independent of areas mediating episodic memories such as hippocampus[58,59] pointing to non-declarative origins for value memory. However, apart from the discussed connections between vIPFC, basal ganglia and IT cortex, vIPFC also connects with other regions such as orbitofrontal cortex (OFC), anterior cingulate cortex (ACC) and amygdala all of which are known to mediate reward-based decision making [60–62] and could be involved in formation of stable object values. While there is some evidence about the precedence of basal ganglia relative to PFC in associative learning and formation of reward-based short-term memory[3], whether reward-based long-term memory formation emerges in a single area or separately across various components of this interconnected network is an open question.

In summary, our results revealed a long-term high capacity memory mechanism in the primate prefrontal cortex for discrimination of objects based on their old values. Repeated reward association created differential object selectivity in vIPFC (Figure 4C). The stability of object discrimination for many months in vIPFC is consistent with the stability of behavioral memory of object values reported here and in our previous studies. Maintaining memory of an object's old value is important in real life where many objects are experienced and must be efficiently detected in future encounters. This system allows animals and humans to robustly adapt to their environments to find previously rewarding objects, accurately and quickly. On the other hand, this long-term high capacity memory could also be relevant for mal-adaptive behaviors where rewarding drug cues persistently activate cortical areas to drive drug seeking behavior. This may explain why disruption of such prefrontal activations can be an effective therapeutic approach[63].

## STAR\*Methods

### CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources, data and code should be directed to and will be fulfilled by the Lead Contact, Dr. Ali Ghazizadeh (alieghazizadeh@gmail.com).

### EXPERIMENTAL MODEL AND SUBJECT DETAILS

**Subjects and surgery**—Two adult male rhesus monkeys (*Macaca mulatta*) were used in all tasks (monkeys B and R ages 7 and 10, respectively). 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. Both animals underwent surgery under general anesthesia during which a head holder and a recording chamber were implanted on the head and scleral search coils for eye tracking were inserted in the eyes. The chamber was tilted laterally and was placed over the left and right prefrontal cortex (PFC) for monkeys B and R, respectively (25° tilt for B and 35° tilt for R, Figure S1A). After confirming the position of recording chamber using MRI, a craniotomy over PFC was performed during a second surgery.

## METHOD DETAILS

**Sample size**—To calculate the sample size needed, we wanted to be able to detect changes of at least 0.05 (i.e.  $AUC > 0.55$  or  $< 0.45$ ) in mean population AUC (Figure 4A) with power 80% and significance level of 0.05. Our initial recording in each monkey showed a standard deviation of about 0.2 in AUC histogram across the population. Given these values and assuming a normal distribution for AUC, the total number of neurons to have significant results separately in each monkey, was found to be 126 neurons per animal (<http://powerandsamplesize.com/Calculators/Test-1-Mean/1-Sample-Equality>). Given the fact that we had further division of neurons among memory periods as well as a reward learning task, we collected more than this minimum sample size in each monkey (159 and 191 neurons in monkeys B and R, respectively).

**Stimuli**—Visual stimuli with fractal geometry were used as objects [64]. One fractal was composed of four point-symmetrical polygons that were overlaid around a common center such that smaller polygons were positioned more toward the front. The parameters that determined each polygon (size, edges, color, etc.) were chosen randomly. Fractal diameters were on average  $\sim 7^\circ$  (ranging from  $5^\circ$ - $10^\circ$ ). Monkeys saw many ( $> 300$ ) fractals (B: 104, R: 96 fractals used to test value memory and B: 234, R: 308 additional fractals for continuous reward learning during the memory periods) half of which were randomly selected to be associated with large reward (good objects) and the other half with small reward (bad objects).

**Task control and Neural recording**—All behavioral tasks and recordings were controlled by custom written visual C++ based software (Blip; [www.robilis.com/blip](http://www.robilis.com/blip)). Data acquisition and output control was performed using National Instruments NI-PCIE 6353. During the experiment, head-fixed monkeys sat in a primate chair and viewed stimuli rear-projected on a screen in front of them ( $\sim 30$ cm) by an active-matrix liquid crystal display projector (PJ550, ViewSonic). Eye position was sampled at 1 kHz using a scleral search coils. Diluted apple juice (33% and 66% for monkey B and R respectively) was used as reward. Reward amounts could be either small (0.08ml and 0.1ml for monkey B and R, respectively) or large (0.21ml and 0.35ml for monkey B and R, respectively).

The recording sites were determined with the aid of gadolinium-filled chamber and grid using MR images (4.7T, Bruker, Figure S1). Activity of single isolated neurons were recorded with acute penetrations of glass coated tungsten electrodes (AlphaOmega, 250 $\mu$  total thickness). The dura was punctured with a sharpened stainless-steel guide tube and the electrode was inserted into the brain through the guide tube by an oil-driven micromanipulator (MO-972, Narishige) until neural background or multiunit was

encountered. The electrode would then be retracted and adjusted slowly until the surface of the brain was determined. Recording depth of encountered neurons were measured from this surface depth (Figure S1). The electric signal from the electrode was amplified and filtered (2 Hz-10 kHz; BAK amplifier and pre-amps) and was digitized at 1kHz. Neural spikes were isolated online using voltage-time discrimination windows. Spike shapes were digitized at 40kHz and recorded for 4.5ms for at least 300 spikes per neuron. An attempt was made to record all well-isolated and visually responsive neurons (visually response to neutral familiar fractals using receptive field mapper or passive viewing tasks or to flashing white dots in various locations). This was done to ensure that results were unbiased and representative of the population activity of visually responsive neurons in PFC. This resulted in a total of 350 PFC neurons in both monkeys (Table 1). In a few sessions, recording was done to confirm the location of frontal eye field (FEF) in both monkeys. Neurons with pre-saccadic discharge were found in FEF locations shown in both monkeys as reported previously [65](Figure S1B). In monkey B, the FEF location was confirmed by low-intensity electrical stimulation (~35 $\mu$ A) which evoked downward saccades reliably.

**Receptive field mapping task**—In this task, the animal had to keep fixating a central white dot (2°) while fractal objects were shown in one of 33 locations spanning eight radial directions and eccentricities from 0 to 20 degrees in 5 degree steps. Fractals covered fixation when shown on center. To measure visual responses unaffected by value and to reduce effects of object selectivity, multiple (at least 8) neutral familiar objects not used in value training were used in this task. Objects were shown sequentially with 400ms on and 200ms off schedule. Central fixation would remain on between object presentations. Animals were rewarded for fixating after each object with probability 0.125 after which an ITI of 1–1.5s with black screen would ensue. Locations were visited once orderly along radial directions, then orderly along the eccentricity circles and finally once randomly, resulting in 99 object presentations in one block of mapping. For some neurons, more than one block of mapping was performed. While mapping was done for all neurons in this study to determine the RF at experiment time, the data from mapping task itself was saved and analyzed for 221 neurons (Figure S2B).

**Value Training: saccade task**—Each session of training was performed with a set of eight fractals (4 good / 4 bad fractals). A trial started after central fixation on a white dot (2°), after which one object appeared on the screen at one of the four peripheral locations (10–15° eccentricity) or center (force trials, Figure 1A). In some sessions fractals were shown on 8 radial directions (45° divisions). After an overlap period of 400ms, the fixation dot disappeared and the animal was required to make a saccade to the fractal. After 500 $\pm$ 100ms of fixating the fractal, a large or small reward was delivered. Diluted apple juice (33–66%) was used as reward. The displayed fractal was then turned off followed by an inter-trial intervals (ITI) of 1–1.5s with a black screen. Breaking fixation or a premature saccade to fractal during overlap period resulted in an error tone (<7% of trials). A correct tone was played after a correct trial. Normally a training session consisted of 80 trials with each object pseudo-randomly presented 10 times. This created different intervening trials for reward association with a given object which are used in Figure 2F–I to examine retention of learned values across trials. Each object was trained in at least 10 sessions prior to test

of long-term value memory. When saccade task was done with neural recording (Figure 2, monkeys B and R with 23 and 11 neurons, respectively), first day of training consisted of 160 trials (20 trials/ object) and after >10 days training consisted of at least 40 trials (5 trials/ object). To check the behavioral learning of object values, choice trials with one good and one bad object were included randomly in one out of five trials (20% choice trials, Figure S3). The location and identity of fractals were randomized across choice trials. During the choice the two fractals were shown in diametrically opposite locations and monkey was required to choose one by looking and holding gaze for  $500 \pm 100$ ms one fractal after which both fractals were turned off and corresponding reward (large or small) would be delivered. Only a single saccade was allowed in choice trials.

**Value memory: passive viewing task**—Neural discrimination of good and bad objects was measured in the same day (minutes to hours after) or 1–6 days, 1–4 weeks (7–29 days) or 1–4 months (30–142 days) after last value training session using a passive viewing task. A trial started after central fixation on a white dot ( $2^\circ$ ). Animal was required to hold a central fixation while good and bad objects were displayed randomly with 400ms on and 400ms off schedule. Animal was rewarded for continued fixation after a random number of 2–4 objects were shown. Objects were shown close to the location with maximal visual response for each neuron as determined by receptive field mapping task (Figure S2). When this maximal location was close to center ( $<5^\circ$ ) passive viewing was sometimes done by showing objects at the center. During a block of passive viewing a set of eight fractals (4 good / 4 bad fractals, a column in Figure 1C) were used with 5–6 presentation per object. In many cases, more than one block was acquired for a given set (often 20 trials/obj) as shown in Figure 4F.

**Behavioral memory: free-viewing task**—Each free-viewing session consisted of 15 trials. In any given trial, four fractals would be randomly chosen from 4 good and 4 bad objects. Location and identity of fractals shown in a trial would be chosen at random. Thus, a given trial could have anywhere between 0–4 good objects shown in any of the four corners of an imaginary diamond or square around center ( $15^\circ$  from display center, Figure 6B). Fractals were displayed for 3 seconds during which the subjects could look at (or ignore) the displayed fractals. There was no behavioral outcome for free viewing behavior. After 3 seconds of viewing, the fractals disappeared. After a delay of 0.5–0.7s, a white fixation dot appeared in one of nine random locations in the screen (center or eight radial directions). Monkeys were rewarded for fixating the fixation dot. This reward was not contingent on free viewing behavior. Next display onset with 4 fractals was preceded by an ITI of 1–1.5s with a black screen. Monkey B did 11 and 11 and monkey R did 10 and 9 sessions of free viewing for days- and months-old values, respectively.

## QUANTIFICATION AND STATISTICAL ANALYSIS

**Neural data analysis**—Responses were time-locked to object-onset for analysis in all tasks (receptor field mapping, saccade task and passive viewing tasks). The analysis epoch was from 50–350ms in receptive field mapping task (Figure S2) and 100–400ms after object onset in saccade task and passive viewing tasks. Average firing to good and bad objects and their difference were calculated during the analysis epoch (e.g. Figure 2C–D, 3B–C). The discriminability based on learned values was measured from mean firing during analysis

epoch across trials using area under receiver operating characteristic curve (AUC). Wilcoxon rank-sum test was used for AUC significance for individual neurons (Figure 4A).

**Value preference cross validation:** The value preference of each neurons was determined using a cross validation method [66]. The sign of firing difference between good and bad objects in odd trials was used to determine value preference in even trials and vice-versa (i.e. if a neuron fired higher for good compared to bad in odd trials, this firing preference label was applied to even trials and similarly value preference label in odd trials came from even trials). The odd and even trials were then combined using their cross validated value preference labels. Average firing to preferred and non-preferred values (e.g. Figure 2B, Figure 3B) and AUC of preferred vs non preferred values (e.g. Figure 2E, Figure 4D) were constructed using these cross validated value preferences for each neuron. All neurons were used in AUC averages whether value preferences were significant or not (i.e. we did not select only significant neurons for population analysis).

Absolute pairwise AUC between objects (reflected around 0.5) were calculated in the analysis epoch (Figure 4C). Significant shift above or below diagonal was determined by comparing the distribution of difference between pairwise AUCs denoted by y-axis and x-axis for each neuron (actual data) with the same distribution made from shuffling good and bad labels (e.g. first 4 objects good and second 4 objects bad labels as shown in Figure 3A, shuffled to be odd objects good and even objects bad) and using a t-test between the average of distribution in real data compared to averages of shuffled data distributions for 10 such shufflings. To determine whether between value discrimination was related to the within value discrimination across the population a regression analysis was used. Since for any arbitrary grouping of objects the between and within object discrimination tend to be positively correlated (higher within object discrimination resulting in higher between object discrimination) and to account for this effect, the regression was done first for the shuffled data  $y^s = a^s x^s + b^s$  where  $y^s$  and  $x^s$  were between and within value AUCs. Then for the real data we had  $y^r - a^s x^r = a^r x^r + b^r$  where  $y^r$  and  $x^r$  were between and within value AUCs and  $a^s$  was the slope from the shuffled data regression. The  $a^r$  shows if there is any relationship for between and within value AUCs beyond what can be expected from arbitrary groupings of data into good and bad objects (beyond chance). Using this method  $a^s = 0.8$  in shuffled data and  $a^r = -0.01$ ,  $P=0.81$ . Thus, the degree of between value category discriminability was not related to within value discriminability across the population.

**Non-parametric test for neuron response:** To determine whether and at which time points the firing difference between preferred and non-preferred value was different from zero (Figure 3B), we used a permutation test to obtain a null distribution of differences by a Monte-Carlo method with 1000 iterations in which value preference of each neuron was randomly assigned and an F-statistic for the difference was calculated for each time point across the PSTH (-200 to 800ms). Since the data in adjacent time points cannot be considered independent and to correct for multiple comparisons, the statistics extracted from each iteration was the maximum F-statistic across all time points following the method proposed by Blair and Karniski 1993 [67]. The max F-statistic from the real data was then compared with the null distribution of max F-statistics to determine the significance of the

whole differential firing. The significant time points in each trace (black dots in Figure 3B) was then determined as points with F-statistic with  $p < 0.05$  compared to this null distribution (resulting in familywise error  $< 0.05$ ). For comparing the differential firing across memory periods (Figure 3C), we extended this permutation method to do non-parametric equivalent to one-way ANOVA by permuting the memory labels across differential firings and doing 1000 iterations to obtain the null distribution. An F-statistic with four levels across time was calculated and its max was used to make the null distributions against which the max F-statistic from the real data was compared and was found not to be significant ( $P > 0.8$ ).

**Onset detection procedure:** Custom written MATLAB functions were used to detect response onset (visual and value onsets in each neuron, Figure S5A). Briefly, for each neuron, raw firing for visual onset or firing difference for value onset was transformed to z-scores using -200ms to 30ms after object onset as the baseline. First response peak after object onset was detected using MATLAB *findpeaks* with a minimum peak height of 1.64 corresponding to 95% confidence interval. The onset was determined as the first valley before this peak (valleys within baseline) using *findpeaks* on the inverted response.

**Object selectivity in PFC:** A common metric for measuring object selectivity in a given neuron is the response sparsity [40]. Sparsity of a neuron that responds equally to all objects is zero and for a neuron that responds only to one object and does not fire to any other objects is one. Using this metric, we found the average sparsity of the PFC neurons to be  $0.095 \pm 0.11$  (sd) which is lower than values reported for IT cortex (e.g. often  $> 0.3$ ) consistent with previous literature [41,42]. Sparsity of good- and bad- preferring neurons were like each other at 0.085 and 0.088, respectively, and were slightly lower than sparsity of non-significant value neurons at 0.109 (Figure S5C left). Sparsity of objects calculated separately among good and bad objects for the three neuron types (good- and bad-preferring and non-significant value), showed that the reduction of sparsity happens for bad objects in bad-preferring and good objects in good-preferring neurons, Figure S5C right). This may be taken as reduction of object selectivity in the preferred value for value selective neurons. However, the sparsity metric reduces if a positive shift is added to all object firings. Thus, the reduction of sparsity for objects in the preferred value category may have been simply caused by an upward shift in firing. On the other hand, if preferred value changes the object firing in a multiplicative fashion, sparsity metric will not change. These two possibilities can be examined using, coefficient of variation (CV, see STAR Methods), which behaves similar to sparsity in this regard. While we found some increase in the sd for firing to the preferred value objects (Figure S5D), the increase in the sd was less than predicted by a multiplicative effect such that CV showed a trend to be smaller for preferred values (Figure S5E) thus explaining the overall reduction of sparsity.

Another metric that is informative for object selectivity and could be used by downstream areas is the overall discriminability of objects. Instead of asking whether the firing of a given neuron is sparse to a few objects, one can ask how well any two objects can be discriminated from each other given the neurons response to multiple objects. Furthermore, a discriminability metric such as AUC is robust to addition and multiplication (i.e. does not change by the same constant added or multiplied for both groups). Thus, in the case of

value coding neurons comparison of pairwise AUC within bad and good object maybe fairer than sparsity which is sensitive to additive shifts. Average absolute pairwise AUC was not different across neuron types and no difference in pairwise discriminability of good and bad objects was observed within good- and bad-preferring neurons (Figure S5F). Thus, while value training had large effect on discriminability of good vs bad objects (Figure 4A–C) within good and bad object discriminabilities were not different across the population.

**Quantifying object selectivity:** Sparsity of responses to objects (Figure S5C) was determined for each neuron (at least 8 objects per neuron, half good) using the following formula separately for good and bad objects as:  $S = (1 - A) / \left(1 - \frac{1}{n}\right)$  where  $A = (\sum_i r_i/n)^2 / \sum_i (r_i^2/n)$ . The standard deviation of response to objects (Figure S5D) was calculated in each value category as:  $SD = \left(\frac{1}{n-1} \sum_i (r_i - \bar{r})^2\right)^{0.5}$ . The coefficient of variation (Figure S5E) was calculated for each value category as:  $CV = \frac{SD}{\bar{r}}$ . In all formulæ 'n' is the number of stimuli in a category and 'r<sub>i</sub>' is the average firing in the analysis epoch for a given stimuli and  $\bar{r}$  is the average of 'r<sub>i</sub>'. To calculate sparsity, SD and CV, for a given neuron, the value of these metrics were first calculated within value category and then were averaged across the two.

For these analyses or others that were agnostic to the effect of memory (e.g. Figure 4A,C,F), data from multiple memory periods were collapsed (e.g. for a neuron tested with 4 good objects in days and 4 good objects in months, a total of 8 good objects were considered regardless of memory period).

**Memory capacity:** To examine the memory capacity in PFC (Figure 5), responses to individual objects were averaged across all the neurons recorded with that object during passive viewing. The average population responses to the object were then normalized by the mean firing in 100–400ms window after object onset to all objects used in a given passive viewing session (4 good/4 bad objects) and then plotted together for all objects (Figure 5A). This normalization allowed us to compare responses of objects across the population where neurons recorded with different sets were different. To quantify the size of memory in PFC, we used the sum of true positive and true negatives at optimal criterion (percurve OPTROCPT). For Figure 5B, the AUC of a given object vs all the other objects from the other category was calculated based on the average firing in 100–400ms window after object onset. For example if a neuron was recorded with 4 good and 4 bad objects, AUC of each good object vs all 4 bad objects was calculated. This was repeated for all neurons recorded with that set and the AUCs for the same object was averaged across neurons with the same value preference and plotted along with the s.e.m. along the object number. The same process was repeated for the bad objects. Good- and bad-preferring neurons were determined by the results shown in Figure 4A based on whether AUC >0.5 or <0.5. (All neurons whether significant value coding or not are included in Figure 5)

**Reconstruction of recorded locations on standard atlas:** To create recording location on a standard brain (D99 brain[68], Figure S1B), the most anterior, posterior, lateral and medial recorded locations (4 recorded corners) for each monkey were determined and the

corresponding coronal sections in the native space and the standard atlas were matched by visual inspection. The 4 recorded corners were scaled and sheared (affine transformation) to match the corresponding corners in the lateral surface of the standard brain and the corresponding transformation was applied to all the recorded location in the grid (black dots, Figure S1B) and overlaid on the standard brain along with the anatomical demarcations color-coded according to standard atlas [69].

**Free-viewing analysis**—Gaze locations were analyzed using custom written MATALB functions in an automated fashion and saccades (displacements  $>2.5^\circ$ ) vs stationary periods were separated in each trial [22,24]. Behavioral memory of good vs bad objects was based on their discriminability (AUC) using first saccade in each session after display onset. The overall AUC was then averaged across sessions (Figure 6C).

**Statistical tests and significance levels**—Chi-squared test was used to test main effect of memory period on proportions (e.g. Figure 4D right). One-way and two-way ANOVAs were used to test main effect of memory period (e.g. Figure 4D left) or trials on neural responses (e.g. Figure 4F–G) or behavior (e.g. Figure S3A). Error-bars in all plots show standard error of the mean (s.e.m). Significance threshold for all tests in this study was  $p < 0.05$ . *ns*: not significant, \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$  (two-sided).

## DATA AND SOFTWARE AVAILABILITY

Requests for the reported recording data and for the custom-written MATLAB scripts that were used for data analysis should be directed to the lead contact.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements:

This research was supported by the Intramural Research Program at the National Institutes of Health, National Eye Institute. We thank Hikosaka lab and Bruce Cumming for valuable discussions.

## References:

1. Carew TJ, and Sahley CL (1986). INVERTEBRATE LEARNING AND MEMORY - FROM BEHAVIOR TO MOLECULES. *Annu. Rev. Neurosci* 9, 435–487. [PubMed: 2423010]
2. Skinner BF (1938). *The Behavior of Organisms* (Appleton-Century-Crofts, New York). Spaulding, WD, Storms, L., Goodrich, V., Sullivan, M.(1986), *Appl. Exp. Psychopathol. Psychiatr. Rehabil. Schizophr. Bull* 12, 560–577.
3. Pasupathy A, and Miller EK (2005). Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* 433, 873–876. [PubMed: 15729344]
4. Sakagami M, and Tsutsui K (1999). The hierarchical organization of decision making in the primate prefrontal cortex. *Neurosci. Res* 34, 79–89. [PubMed: 10498334]
5. Brush ES, Rosvold HE, and Mishkin M (1961). EFFECTS OF OBJECT PREFERENCES AND AVERSIONS ON DISCRIMINATION LEARNING IN MONKEYS WITH FRONTAL LESIONS. *J. Comp. Physiol. Psychol* 54, 319-.



6. Easton A, and Gaffan D (2001). Crossed unilateral lesions of the medial forebrain bundle and either inferior temporal or frontal cortex impair object-reward association learning in Rhesus monkeys. *Neuropsychologia* 39, 71–82. [PubMed: 11115656]
7. Kobayashi S, Lauwereyns J, Koizumi M, Sakagami M, and Hikosaka O (2002). Influence of reward expectation on visuospatial processing in macaque lateral prefrontal cortex. *J. Neurophysiol* 87, 1488–1498. [PubMed: 11877521]
8. Kane MJ, and Engle RW (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychon. Bull. Rev* 9, 637–671. [PubMed: 12613671]
9. Barraclough DJ, Conroy ML, and Lee D (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci* 7, 404–410. [PubMed: 15004564]
10. Kim HF, and Hikosaka O (2013). Distinct Basal Ganglia circuits controlling behaviors guided by flexible and stable values. *Neuron* 79, 1001–1010. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23954031>. [PubMed: 23954031]
11. Yasuda M, and Hikosaka O (2015). Functional territories in primate substantia nigra pars reticulata separately signaling stable and flexible values. *J. Neurophysiol* 113, 1681–1696. Available at: 10.1152/jn.00674.2014. [PubMed: 25540224]
12. Winocur G, Moscovitch M, and Bontempi B (2010). Memory formation and long-term retention in humans and animals: Convergence towards a transformation account of hippocampal-neocortical interactions. *Neuropsychologia* 48, 2339–2356. [PubMed: 20430044]
13. Tomita H, Ohbayashi M, Nakahara K, Hasegawa I, and Miyashita Y (1999). Top-down signal from prefrontal cortex in executive control of memory retrieval. *Nature* 401, 699–703. [PubMed: 10537108]
14. Blumenfeld RS, and Ranganath C (2007). Prefrontal cortex and long-term memory encoding: An integrative review of findings from neuropsychology and neuroimaging. *Neuroscientist* 13, 280–291. [PubMed: 17519370]
15. Euston DR, Gruber AJ, and McNaughton BL (2012). The role of medial prefrontal cortex in memory and decision making. *Neuron* 76, 1057–1070. [PubMed: 23259943]
16. Griggs WS, Kim HF, Ghazizadeh A, Costello MG, Wall KM, and Hikosaka O (2017). Flexible and stable value coding areas in caudate head and tail receive anatomically distinct cortical and subcortical inputs. *Front. Neuroanat* 11, 106. [PubMed: 29225570]
17. Ghazizadeh A, Griggs W, Leopold DA, and Hikosaka O (2018). Temporal-prefrontal cortical network for discrimination of valuable objects in long-term memory. *Proc. Natl. Acad. Sci. U. S. A* 115.
18. Todd TP, Vurbic D, and Bouton ME (2014). Behavioral and neurobiological mechanisms of extinction in Pavlovian and instrumental learning. *Neurobiol. Learn. Mem* 108, 52–64. [PubMed: 23999219]
19. Grill-Spector K, Henson R, and Martin A (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci* 10, 14–23. [PubMed: 16321563]
20. Dickinson A, Balleine B, Watt A, Gonzalez F, and Boakes RA (1995). Motivational control after extended instrumental training. *Anim. Learn. Behav* 23, 197–206.
21. Hikosaka O, Yamamoto S, Yasuda M, and Kim HF (2013). Why skill matters. *Trends Cogn. Sci* 17, 434–441. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23911579>. [PubMed: 23911579]
22. Ghazizadeh A, Griggs W, and Hikosaka O (2016). Object-finding skill created by repeated reward experience. *J. Vis* 16.
23. Yamamoto S, Kim HF, and Hikosaka O (2013). Reward value-contingent changes of visual responses in the primate caudate tail associated with a visuomotor skill. *J. Neurosci* 33, 11227–11238. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23825426>. [PubMed: 23825426]
24. Ghazizadeh A, Griggs W, and Hikosaka O (2016). Ecological origins of object salience: Reward, uncertainty, aversiveness, and novelty. *Front. Neurosci* 10.
25. Eichenbaum H, Yonelinas AP, and Ranganath C (2007). The medial temporal lobe and recognition memory. *Annu. Rev. Neurosci* 30, 123–152. [PubMed: 17417939]
26. Goldman-Rakic PS (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. *Compr. Physiol*

27. Miller EK, Erickson CA, and Desimone R (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J Neurosci* 16, 5154–5167. Available at: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=8756444](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8756444). [PubMed: 8756444]
28. Osada T, Adachi Y, Kimura HM, and Miyashita Y (2008). Towards understanding of the cortical network underlying associative memory. *Philos. Trans. R. Soc. B-Biological Sci* 363, 2187–2199.
29. Kim HF, Ghazizadeh A, and Hikosaka O (2014). Separate groups of dopamine neurons innervate caudate head and tail encoding flexible and stable value memories. *Front. Neuroanat* 8.
30. Hikosaka O, Kim HF, Yasuda M, and Yamamoto S (2014). Basal Ganglia Circuits for Reward Value-Guided Behavior. *Annu. Rev. Neurosci.* Vol 37 37, 289–+.
31. Hikosaka O, Ghazizadeh A, Griggs W, and Amita H (2017). Parallel basal ganglia circuits for decision making. *J. Neural Transm.* 1–15.
32. Yasuda M, Yamamoto S, and Hikosaka O (2012). Robust representation of stable object values in the oculomotor Basal Ganglia. *J. Neurosci* 32, 16917–16932. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23175843>. [PubMed: 23175843]
33. Kim HF, Ghazizadeh A, and Hikosaka O (2015). Dopamine Neurons Encoding Long-Term Memory of Object Value for Habitual Behavior. *Cell* 163, 1165–1175. [PubMed: 26590420]
34. Middleton FA, and Strick PL (2002). Basal-ganglia “projections” to the prefrontal cortex of the primate. *Cereb. Cortex* 12, 926–935. [PubMed: 12183392]
35. Gerbella M, Borra E, Tonelli S, Rozzi S, and Luppino G (2013). Connectional heterogeneity of the ventral part of the macaque area 46. *Cereb. Cortex* 23, 967–987. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22499799>. [PubMed: 22499799]
36. Webster MJ, Bachevalier J, and Ungerleider LG (1994). CONNECTIONS OF INFERIOR TEMPORAL AREAS TEO AND TE WITH PARIETAL AND FRONTAL-CORTEX IN MACAQUE MONKEYS. *Cereb. Cortex* 4, 470–483. [PubMed: 7530521]
37. Browning PGF, Easton A, and Gaffan D (2007). Frontal-temporal disconnection abolishes object discrimination learning set in macaque monkeys. *Cereb. Cortex* 17, 859–864. [PubMed: 16707734]
38. Simons JS, and Spiers HJ (2003). Prefrontal and medial temporal lobe interactions in long-term memory. *Nat. Rev. Neurosci* 4, 637–648. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12894239>. [PubMed: 12894239]
39. Saint-Cyr JA, Ungerleider LG, and Desimone R (1990). Organization of visual cortical inputs to the striatum and subsequent outputs to the pallido-nigral complex in the monkey. *J. Comp. Neurol* 298, 129–156. [PubMed: 1698830]
40. Rolls ET, and Tovee MJ (1995). SPARSENESS OF THE NEURONAL REPRESENTATION OF STIMULI IN THE PRIMATE TEMPORAL VISUAL-CORTEX. *J. Neurophysiol* 73, 713–726. [PubMed: 7760130]
41. Woloszyn L, and Sheinberg DL (2012). Effects of Long-Term Visual Experience on Responses of Distinct Classes of Single Units in Inferior Temporal Cortex. *Neuron* 74, 193–205. [PubMed: 22500640]
42. Bichot NP, Heard MT, DeGennaro EM, and Desimone R (2015). A Source for Feature-Based Attention in the Prefrontal Cortex. *Neuron* 88, 832–844. [PubMed: 26526392]
43. Freedman DJ, Riesenhuber M, Poggio T, and Miller EK (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* (80-. ). 291, 312–316. Available at: [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=11209083](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11209083).
44. Pan XC, and Sakagami M (2012). Category representation and generalization in the prefrontal cortex. *Eur. J. Neurosci* 35, 1083–1091. [PubMed: 22487038]
45. Pan X, Sawa K, Tsuda I, Tsukada M, and Sakagami M (2008). Reward prediction based on stimulus categorization in primate lateral prefrontal cortex. *Nat. Neurosci* 11, 703. [PubMed: 18500338]
46. Borra E, Gerbella M, Rozzi S, and Luppino G (2015). Projections from Caudal Ventrolateral Prefrontal Areas to Brainstem Preoculomotor Structures and to Basal Ganglia and Cerebellar Oculomotor Loops in the Macaque. *Cereb. Cortex* 25, 748–764. [PubMed: 24068552]

47. Zhou HH, and Desimone R (2011). Feature-Based Attention in the Frontal Eye Field and Area V4 during Visual Search. *Neuron* 70, 1205–1217. [PubMed: 21689605]
48. Watanabe M (1996). Reward expectancy in primate prefrontal neurons. *Nature* 382, 629–632. [PubMed: 8757133]
49. Seo H, Barraclough DJ, and Lee D (2007). Dynamic signals related to choices and outcomes in the dorsolateral prefrontal cortex. *Cereb. Cortex* 17, I110–I117. [PubMed: 17548802]
50. Tsutsui KI, Grabenhorst F, Kobayashi S, and Schultz W (2016). A dynamic code for economic object valuation in prefrontal cortex neurons. *Nat. Commun* 7.
51. Wallis JD, and Miller EK (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur J Neurosci* 18, 2069–2081. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14622240>. [PubMed: 14622240]
52. Roesch MR, and Olson CR (2004). Neuronal activity related to reward value and motivation in primate frontal cortex. *Science (80-.)*. 304, 307–310. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15073380>.
53. Rushworth MFS, Noonan MP, Boorman ED, Walton ME, and Behrens TE (2011). Frontal Cortex and Reward-Guided Learning and Decision-Making. *Neuron* 70, 1054–1069. [PubMed: 21689594]
54. Dixon ML, and Christoff K (2014). The lateral prefrontal cortex and complex value-based learning and decision making. *Neurosci. Biobehav. Rev* 45, 9–18. [PubMed: 24792234]
55. Kobayashi S, Nomoto K, Watanabe M, Hikosaka O, Schultz W, and Sakagami M (2006). Influences of rewarding and aversive outcomes on activity in macaque lateral prefrontal cortex. *Neuron* 51, 861–870. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16982429>. [PubMed: 16982429]
56. Maunsell JHR (2004). Neuronal representations of cognitive state: reward or attention? *Trends Cogn. Sci* 8, 261–265. [PubMed: 15165551]
57. Kitamura T, Ogawa SK, Roy DS, Okuyama T, Morrissey MD, Smith LM, Redondo RL, and Tonegawa S (2017). Engrams and circuits crucial for systems consolidation of a memory. *Science (80-.)*. 356, 73–78.
58. Knowlton BJ, and Squire LR (1993). The learning of categories: parallel brain systems for item memory and category knowledge. *Science (80-.)*. 262, 1747–1749.
59. Malamut BL, Saunders RC, and Mishkin M (1984). Monkeys with combined amygdalo-hippocampal lesions succeed in object discrimination learning despite 24-hour intertrial intervals. *Behav. Neurosci* 98, 759. [PubMed: 6487412]
60. Rolls ET, Critchley HD, Mason R, and Wakeman EA (1996). Orbitofrontal cortex neurons: Role in olfactory and visual association learning. *J. Neurophysiol* 75, 1970–1981. [PubMed: 8734596]
61. Kennerley SW, Walton ME, Behrens TEJ, Buckley MJ, and Rushworth MFS (2006). Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci* 9, 940–947. [PubMed: 16783368]
62. Paton JJ, Belova MA, Morrison SE, and Salzman CD (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* 439, 865–870. [PubMed: 16482160]
63. Terraneo A, Leggio L, Saladini M, Ermani M, Bonci A, and Gallimberti L (2016). Transcranial magnetic stimulation of dorsolateral prefrontal cortex reduces cocaine use: A pilot study. *Eur. Neuropsychopharmacol* 26, 37–44. [PubMed: 26655188]
64. Miyashita Y, Higuchi SI, Sakai K, and Masui N (1991). GENERATION OF FRACTAL PATTERNS FOR PROBING THE VISUAL MEMORY. *Neurosci. Res* 12, 307–311. [PubMed: 1660991]
65. Bruce CJ, and Goldberg ME (1985). PRIMATE FRONTAL EYE FIELDS .1. SINGLE NEURONS DISCHARGING BEFORE SACCADES. *J. Neurophysiol* 53, 603–635. [PubMed: 3981231]
66. Bromberg-Martin ES, Hikosaka O, and Nakamura K (2010). Coding of Task Reward Value in the Dorsal Raphe Nucleus. *J. Neurosci* 30, 6262–6272. [PubMed: 20445052]
67. Blair RC, and Karniski W (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology* 30, 518–524. [PubMed: 8416078]
68. Reveley C, Gruslys A, Frank QY, Glen D, Samaha J, Russ BE, Saad Z, Seth AK, Leopold DA, and Saleem KS (2016). Three-dimensional digital template atlas of the macaque brain. *Cereb. Cortex*

69. Saleem KS, and Logothetis NK (2012). A combined MRI and histology atlas of the rhesus monkey brain in stereotaxic coordinates (Academic Press).

Author Manuscript

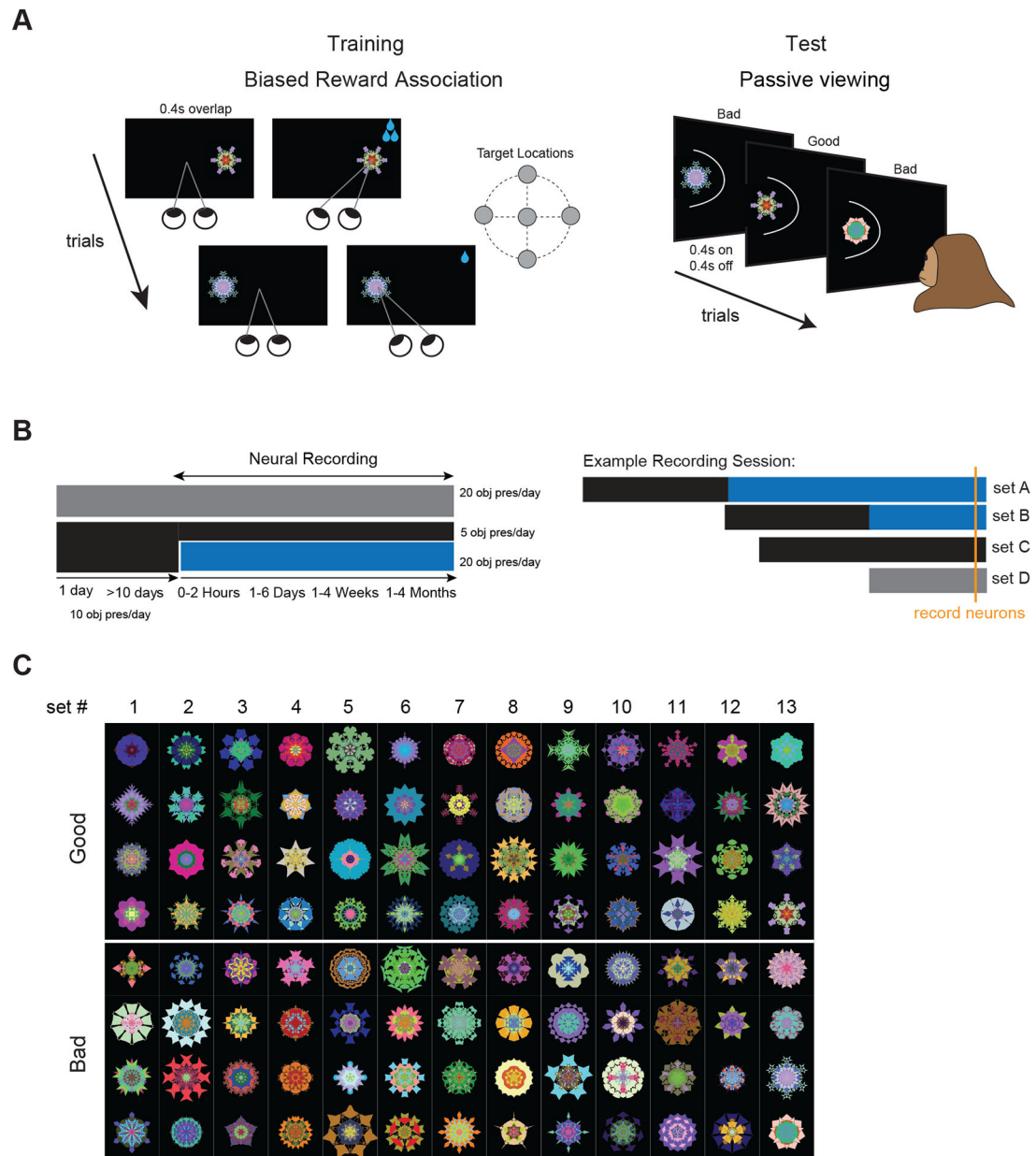
Author Manuscript

Author Manuscript

Author Manuscript

### Highlights

- Ventrolateral prefrontal cortex (vlPFC) encodes learning of object values.
- Learned object values are retained in vlPFC for at least several months.
- Value memory in vlPFC has a high capacity and resists interference or extinction.
- Macaques remember previously high-valued objects several months after learning.



**Figure 1. Object value learning and memory paradigm.**

(A) Reward training: Monkeys first fixated centrally and after the instruction (fixation off) made a saccade to a  $10^{\circ}$ - $15^{\circ}$  peripheral target (or stayed at center if target shown centrally) and held gaze to receive low or high reward (bad and good objects, respectively). Paired reward was always the same for a given object in all trials (left). Passive viewing test: Monkeys kept central fixation while good and bad objects were shown randomly and sequentially in the neuron's receptive field (RF) without reward (right). (B) Acute neural recordings were performed during passive viewing (blue block) to test value memory in objects previously trained with reward (black block) and during reward training of new objects (gray block) and previously trained objects (black block >10 days). Reward-based long-term memory was measured hours, days, weeks or months after last object-reward

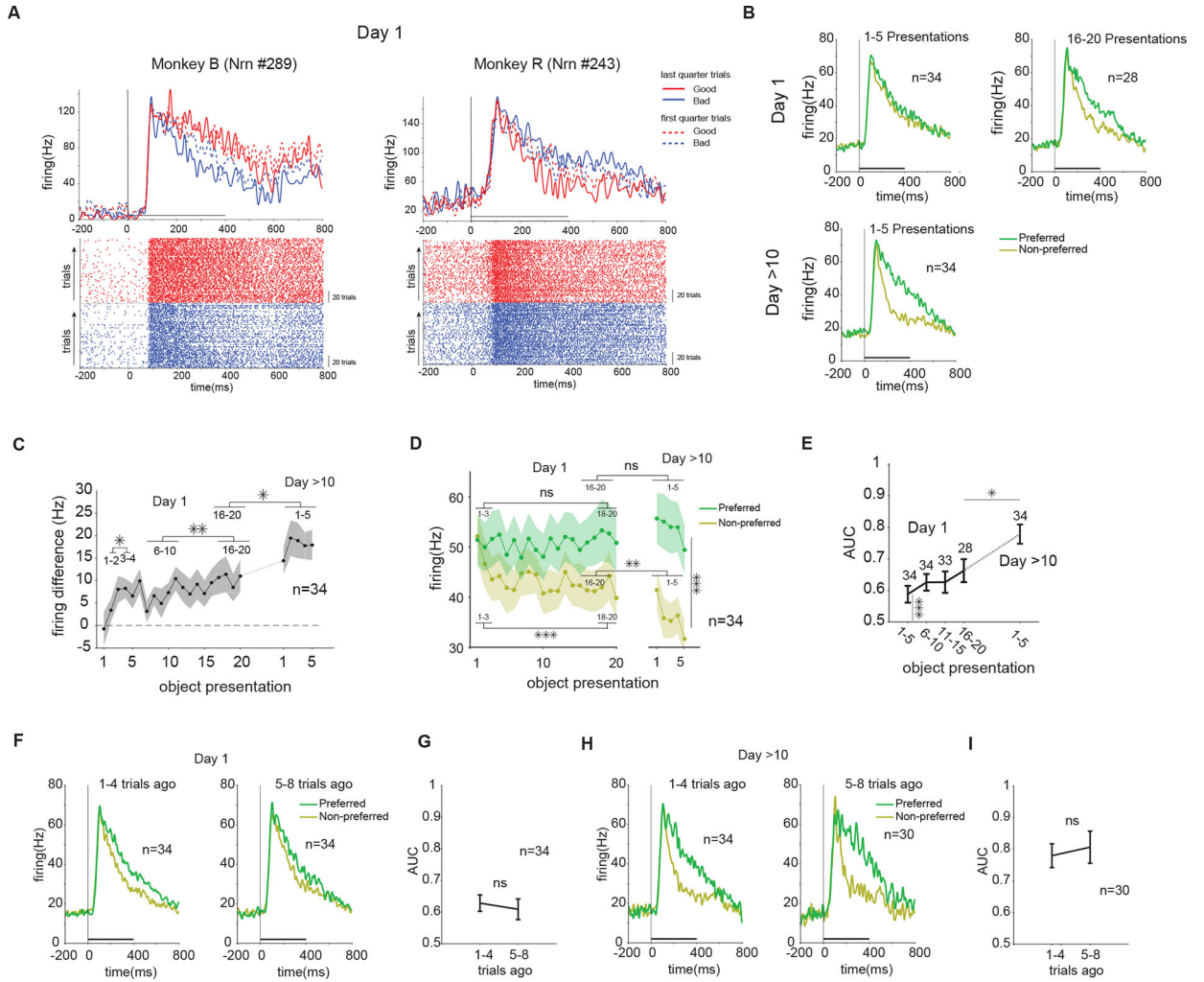
association, during which monkeys were still trained with many other objects (left). Different memory scales within individual neurons could be examined by using objects trained months ago in one block and objects trained days ago in another block. To test the effect of learning, a given neuron could be recorded using novel objects in one block and previously trained objects in another block (example recording session, right). (C) Fractals (>100) used for monkey B in value memory test. Fractals were simultaneously trained and tested in sets of 8 objects. See also Figures S1 and S2

Author Manuscript

Author Manuscript

Author Manuscript

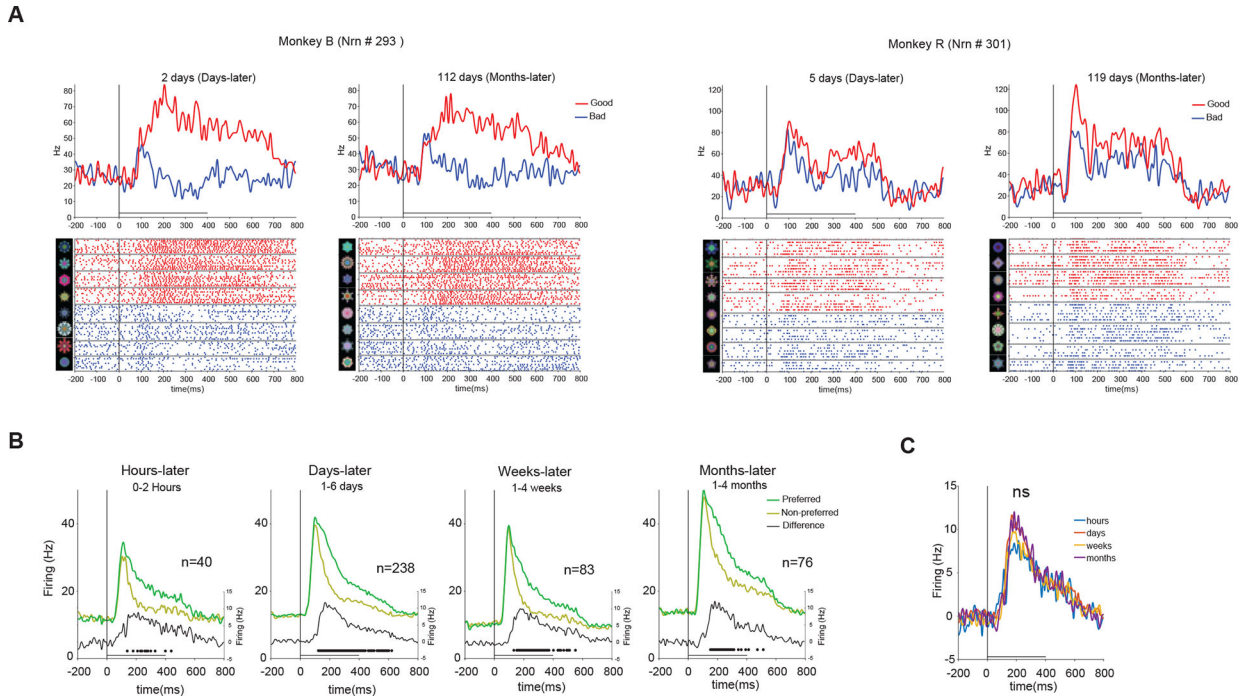
Author Manuscript



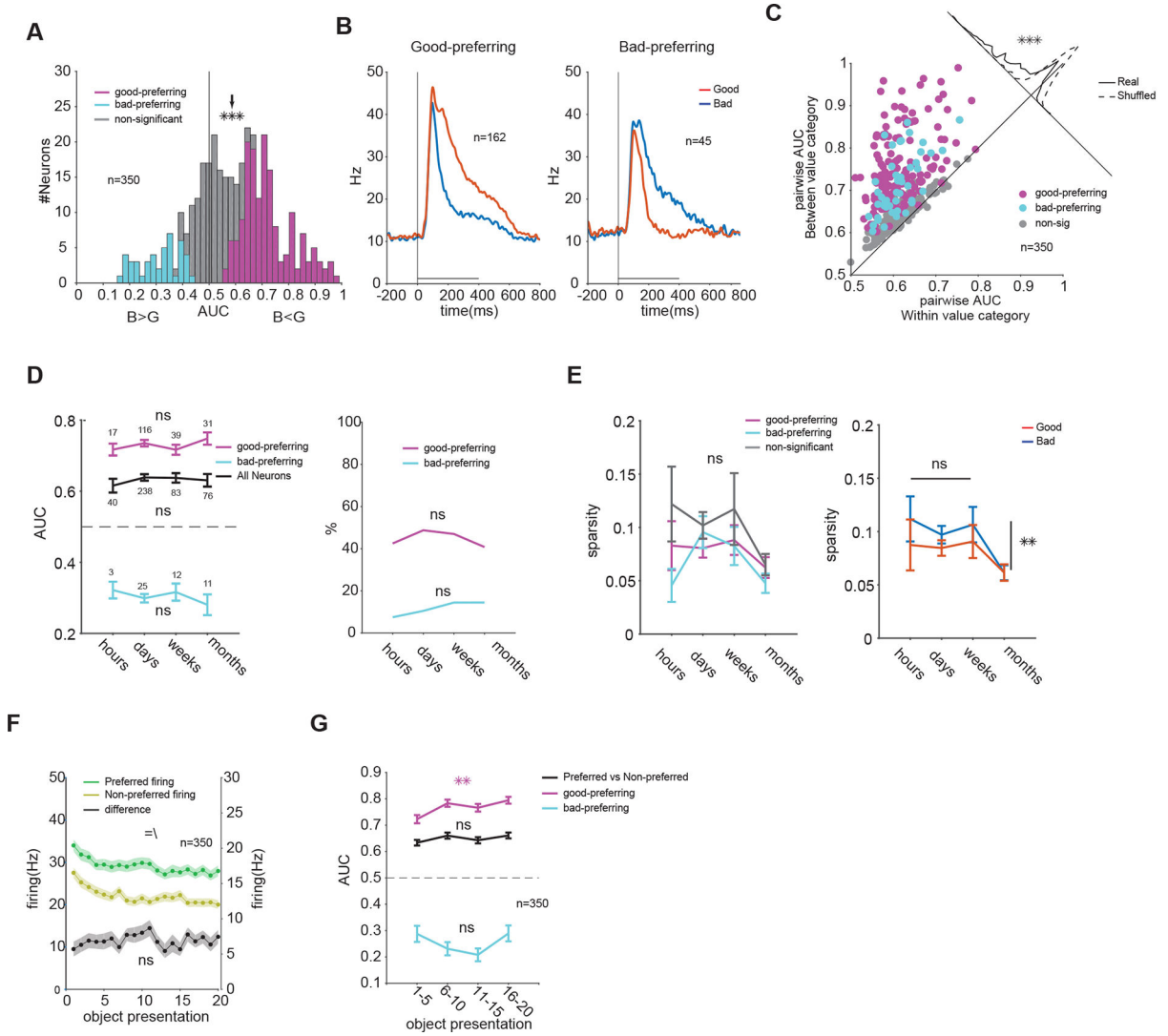
**Figure 2. PFC learns and retains object values across trials during reward training.** (A) Examples of single neurons tested during reward learning. Top row shows average firing to 4 good and 4 bad objects in the beginning and end of training (first and last quarter of trial in a training session). Bottom row shows raster plot of firing to good and bad objects with dots indicating spike times. Y-axis is ordered by trial number from the lowest number at the bottom. Good and bad objects were trained in pseudo-random mixed order, but are displayed separately. (B) Average firing rate to the preferred and non-preferred value in the first and last five object presentations in the first training day and in the first five object presentations after >10 training days. (C) Population average of firing difference to preferred vs non-preferred value across reward training trials for each object in the first day and after >10 days of training ( $F_{24,792}=2.7$ ,  $P<10^{-4}$  main effect of training, paired- $t>2.2$ ,  $P<0.035$  all post-hoc tests). (D) Same as C, but for population average firing rate to preferred and non-preferred value. In Day 1, the firing reduction was significant for non-preferred value ( $t_{84}=4.2$ ,  $P<1e-4$ ), but no significant change was observed for preferred value ( $t_{84}=0.5$ ,  $P=0.6$ ). After >10 days, the difference of preferred and non-preferred value was significant ( $F_{1,330}=38$   $P<1e-8$ ) and the response of non-preferred value showed a further reduction, compared to the end of learning in the first day ( $t_{144}=3.2$ ,  $P<1e-2$  but for preferred value



$t_{144}=1.6$ ,  $P=0.09$ ) (E) Population average of preferred vs non-preferred value AUC across reward training trials in the first day and after >10 days of training ( $F_{4,158}=5.9$ ,  $P<10^{-3}$  main effect of training, paired- $t>2.5$ ,  $P<0.02$  all post-hoc tests). (F) Average firing to preferred vs non-preferred values for objects not seen 1–4 vs 5–8 trials ago and (G) Population average retention of value AUC for objects not seen 1–4 vs 5–8 trials ago during the first day of training ( $t_{33}=0.7$ ,  $P=0.5$ ). (H-I) Same as F-G but after >10 days of training ( $t_{29}=0.6$ ,  $P=0.6$ ). Horizontal bar in A, B, F, H: Time from object onset to fixation offset. See also Figure S3



**Figure 3. PFC retains memory of object values in the absence of reward for several months.** (A) Examples of single neurons tested in the same session with good and bad objects that were trained with reward days and months before (one example from in each monkey). Top row shows average firing to good and bad objects in a set. Bottom row shows raster plot of firing to each object in the set with dots indicating spike times. Actual fractals used are shown to the left and grouped into good and bad fractals. (B). Population average firing (left y-axis) to preferred vs non-preferred values and differential firing (right y-axis) to preferred vs non-preferred value hours, days, weeks and months after reward association. The differential firing was significant in all memory periods ( $p < 1e-2$ ). The dots show significant time points in difference (corrected for multiple comparison, Family-wise error  $< 0.05$ , see STAR Methods) (C) Differential firing shown in B overlaid for 4 memory periods. The differential firing was not significantly different across memory periods ( $P = 0.83$ , 1-way permutation test, see STAR Methods). Horizontal bar: object-on duration in A-C and subsequent similar plots in passive viewing task. See also Figure S4



**Figure 4. Value memory creates good vs bad object discriminability in PFC neurons which is resistant to passage of time and to extinction.**

(A) Distribution of good vs bad object discrimination (AUC) across all neurons (arrow marks population average good vs bad AUC: 0.58,  $t_{349}=9.8$ ,  $P<10^{-19}$ ). (B) Average firing of good- and bad- preferring neurons (C) Average pairwise discriminability of objects between good and bad category (y-axis) compared to within good and bad category (x-axis) (correlation non-sig,  $P>0.8$ , see STAR Methods). Oblique distribution on top right shows deviation from the diagonal for the actual and shuffled data ( $t_9=116$ ,  $P<10^{-15}$ ). (D) Population average of preferred vs non-preferred value AUC (left, for all  $F_{3,433}=0.7$ ,  $P=0.5$ , for good-preferring  $F_{3,199}=0.8$ ,  $P=0.4$  and for bad-preferring  $F_{3,47}=0.5$ ,  $P=0.6$ ) and percentage of good- and bad-preferring neurons across 4 memory periods (right,  $\chi^2_3 < 2.2$ ,  $P>0.5$ ). (E) Object selectivity as measured by sparsity in good- and bad-preferring neurons and in neurons non-significant for value (left,  $F_{3,424}=1.6$ ,  $P=0.16$ ) and for good and bad objects across all neurons ( $F_{3,864}=3.2$ ,  $P=0.02$ , post hoc sparsity in month lower than other periods  $P<0.03$  lsd). (F) Population average firing to preferred and non-preferred values (left axis) and their difference (right axis) as a function of repeated exposure to objects in the

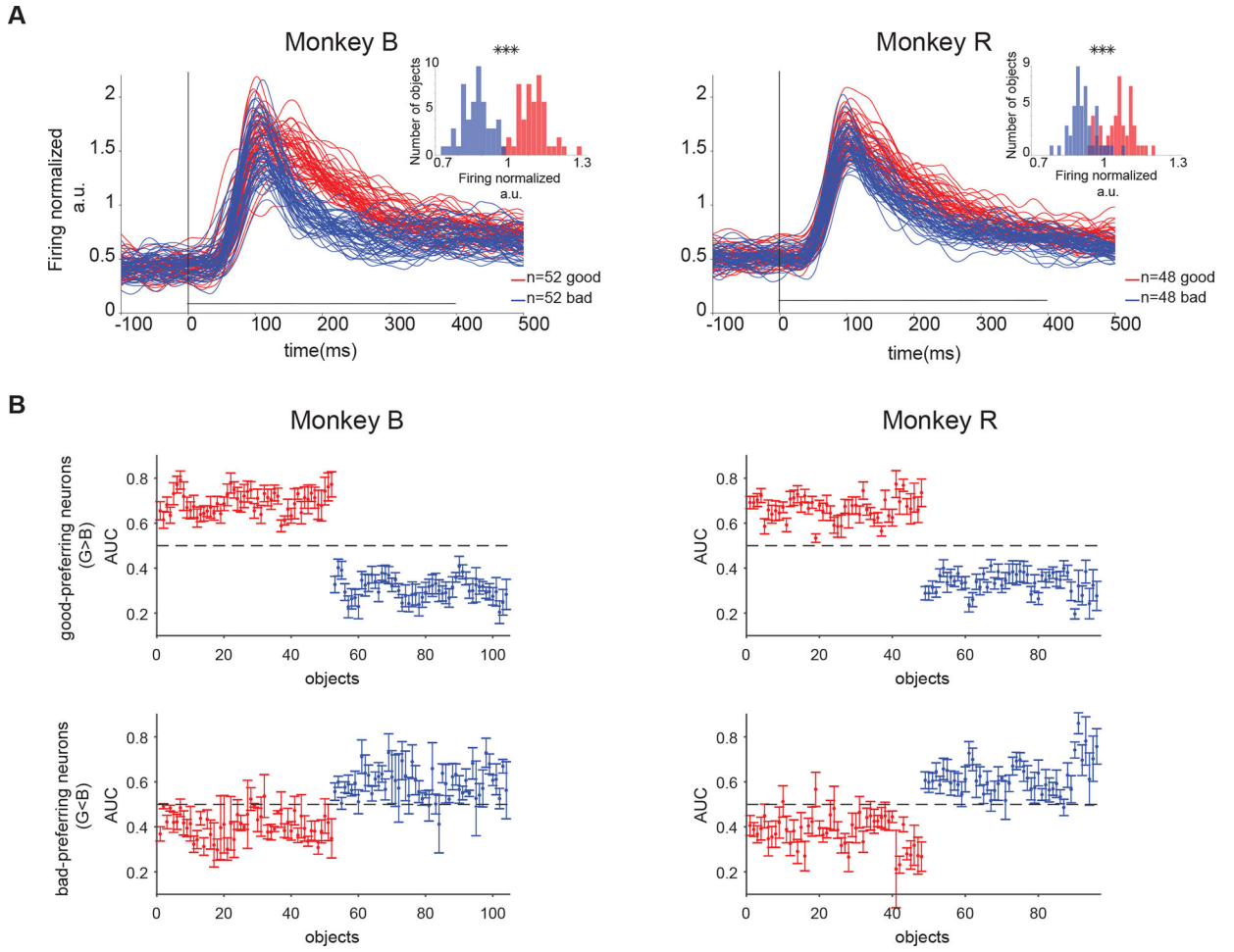
absence of reward during one passive viewing block. The presentations of a given object were not in consecutive trials and were usually intervened by other object presentations during passive viewing, but are plotted as consecutive trials for each object (main effects of value  $F_{1,13960}=270$ ,  $P<10^{-59}$ [=] and of trials  $F_{19,13960}=3.5$ ,  $P<10^{-6}$ [], interaction  $P>0.9$ , main effect of trials on firing difference  $F_{19,6980}=0.68$ ,  $P>0.8$ ). (G) Population average preferred vs non-preferred AUC as a function of repeated exposure to objects ( $F_{3,1396}=1.46$ ,  $P=0.2$ ) and good vs bad AUCs averaged for good- ( $F_{3,644}=4.9$ ,  $P<0.01$ ) and bad-preferring ( $F_{3,176}=2.1$ ,  $P=0.09$ ) neurons. Data in A-C and F-G are collapsed over all memory periods. See also Figures S5 to S7

Author Manuscript

Author Manuscript

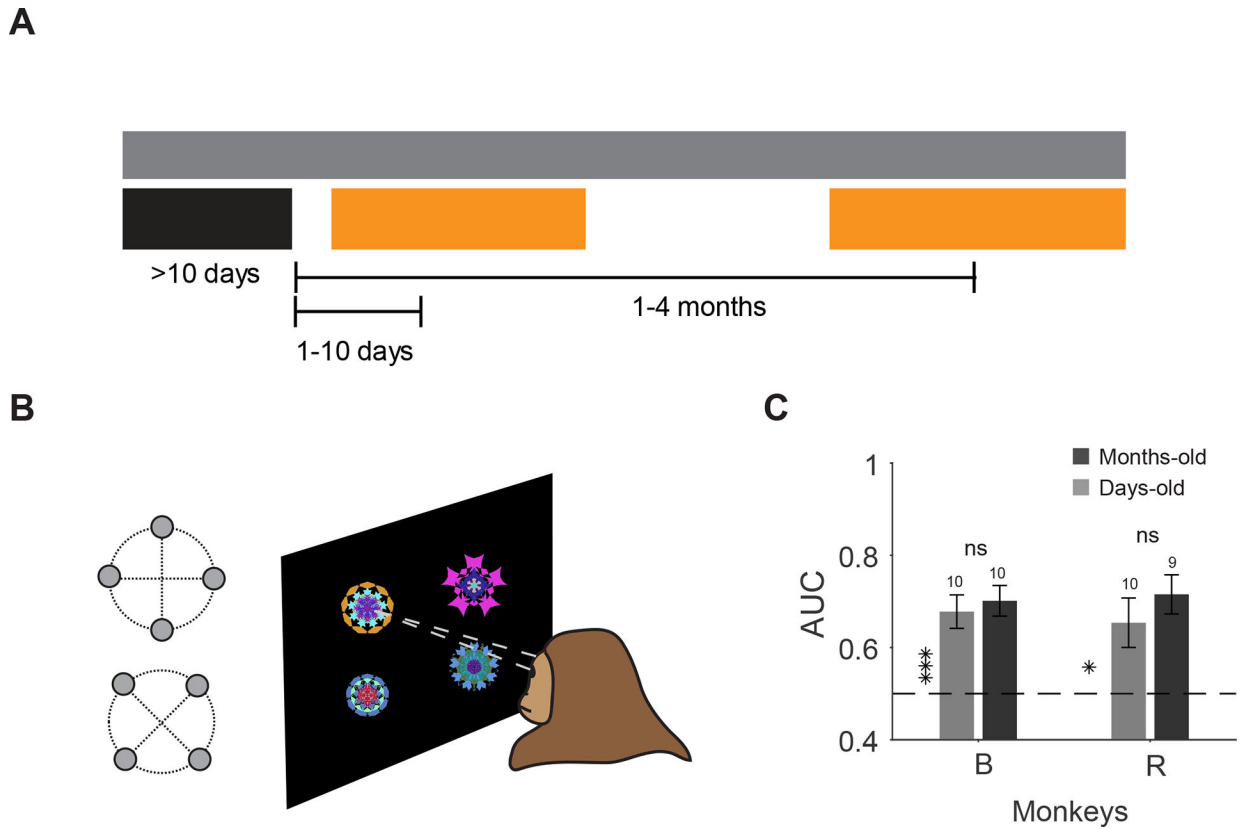
Author Manuscript

Author Manuscript



**Figure 5. PFC value memory has high-capacity for objects**

(A) Average population response to all objects (good and bad) seen by monkey B (left) and R (right). The population response was normalized by mean firing in 100–400ms after object onset to all good and bad objects in a set shown in passive viewing. Small axes show histogram of response to all good and bad object (mean value in 100–400ms after object onset) for monkey B (left) and R (right). (B) Average AUC for each good object vs all other bad objects (red) and for each bad object vs all other good objects (blue) in a set across good-preferring (top row) and bad-preferring neurons (bottom row) in both monkeys.



**Figure 6. Behavioral discrimination of good and bad objects days and months after last reward association.**

(A) Objects were tested in a free viewing task days or months after reward training while reward learning with other objects were in progress. (B) Free viewing task: good and bad objects were randomly selected and shown to the monkey for viewing in the absence of reward. (C) Behavioral discriminability (AUC) of objects based on days-old and months-old values as measured by first saccade after display onset in both monkeys (monkey B:  $t_{10} > 4.9$ ,  $P < 10^{-3}$  monkey days-old and months-old AUC,  $t_{20} = 0.63$ ,  $P = 0.63$  difference in AUC, monkey R:  $t_9 = 2.8$ ,  $P = 0.01$  days-old and  $t_8 = 5.9$ ,  $P < 10^{-3}$  months-old AUC,  $t_{17} = 0.89$ ,  $P = 0.38$ ).

**Table 1.**  
**Number of neurons recorded in each monkey and across various memory periods.**

Neurons separated by the number of memory periods in which they are tested (left side, e.g. if a neuron was recorded with objects from two memory periods, it would be counted under column labeled two) and total number of neurons recorded in each memory period (right side, e.g. if a neuron was recorded with objects in days and months memory periods, it would be counted once under each corresponding column)

<i>Neurons separated by number of memory periods in which they are recorded</i>				<i>Neurons separated by the memory period label</i>					
Monkeys	Number of Memory Periods			Sums	Memory Period Labels				Sums
	One	Two	Three		Hours	Days	Weeks	Months	
<b>B</b>	130	27	2	159	24	95	28	43	190
<b>R</b>	141	44	6	191	16	143	55	33	191
<b>Sums</b>	271	71	8	350	40	238	83	76	350

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript