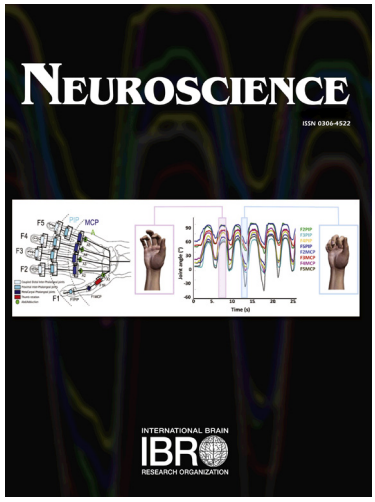
Accepted Manuscript



Outcome modulation across tasks in the primate dorsolateral prefrontal cortex

Encarni Marcos, Simon Nougaret, Satoshi Tsujimoto, Aldo Genovesio

PII: S0306-4522(17)30813-8
DOI: https://doi.org/10.1016/j.neuroscience.2017.11.019
Reference: NSC 18134

To appear in: *Neuroscience*

Received Date: 30 June 2017
Revised Date: 11 October 2017
Accepted Date: 11 November 2017

Please cite this article as: E. Marcos, S. Nougaret, S. Tsujimoto, A. Genovesio, Outcome modulation across tasks in the primate dorsolateral prefrontal cortex, *Neuroscience* (2017), doi: https://doi.org/10.1016/j.neuroscience.2017.11.019

Outcome modulation across tasks in the primate dorsolateral prefrontal cortex

Encarni Marcos[1], Simon Nougaret[1], Satoshi Tsujimoto[2,3], Aldo Genovesio[1]

[1]Department of Physiology and Pharmacology, Sapienza University of Rome, Italy

[2]Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University, Kyoto, Japan

[3]The Nielsen Company Singapore Pte Ltd

Animals need to learn and to adapt to new and changing environments so that appropriate actions that lead to desirable outcomes are acquired within each context. The prefrontal cortex (PF) is known to underlie such function that directly implies that the outcome of each response must be represented in the brain for behavioral policies update. However, whether such PF signal is context dependent or it is a general representation beyond the specificity of a context is still unclear. Here, we analyzed the activity of neurons in the dorsolateral PF (PFdl) recorded while two monkeys performed two perceptual magnitude discrimination tasks. Both tasks were well known by the monkeys and unexpected changes did not occur but the difficulty of the task varied from trial to trial and thus the monkeys made mistakes in a proportion of trials. We show a context independent coding of the response outcome with neurons maintaining similar selectivity in both task contexts. Using a classification method of the neural activity, we also show that the trial outcome could be well predicted from the activity of the same neurons in the two contexts. Altogether, our results provide evidence of high degree of outcome generality in PFdl.

Keywords: neurophysiology; prefrontal cortex; outcome; reward; task context;

INTRODUCTION

During decision making, different outcomes associated with alternative options are evaluated and, after a competitive process, one option is selected among the others (Smith and Ratcliff, 2004, Gold and Shadlen, 2007, Marcos et al., 2013). Such selected option is normally achieved with an action that, ideally, leads to a desirable outcome. However, real-world situations are usually not ideal and the selection of an action over the others is generally influenced by different sources of noise and/or by suboptimal inference that can lead to erroneous choices (Beck et al., 2012). Thus, it is fundamental to monitor each outcome so that behavioral policies can be updated or enhanced accordingly. Previous studies have shown that neurons in the basal ganglia (Hikosaka et al., 1989, Apicella et al., 1991, Bowman et al., 1996, Hollerman et al., 1998), the orbitofrontal cortex (Tremblay and Schultz, 2000), the amygdala (Nakamura et al., 1992), the supplementary eye field (Amador et al., 2000), the primary motor cortex (Ramakrishnan et al., 2017) and the prefrontal cortex (PF) (Watanabe, 1989, Kobayashi et al., 2006) represent the outcome of a response. The outcome encoding has been explored in a broad field, including the manipulation of reward probability within a task context (Apicella et al., 2009), the temporal predictability of stimuli within different contexts (Sardo et al., 2000), the encoding of outcome during instrumental or free reward conditions (Ravel et al., 2001) or during different rules within a task (Mansouri et al., 2006). However, most of these studies have focused on the neural representation of the outcome within a task, and whether such representation is specific or not to the task context has been mainly disregarded. Here, we address this question by analyzing the activity of neurons recorded from the dorsolateral PF (PFdl) of two monkeys while they performed two perceptual discrimination tasks: a distance and a duration discrimination task. Within each task, trials had different difficulty leading to errors. The two tasks had a similar structure but they differed in the rule that had to be applied in order to succeed. The accuracy of the animals was similar, making the subjective value for correct and incorrect cases comparable in both task contexts and providing an ideal experimental framework to compare outcome coding among tasks.

Neurophysiological studies showed that neurons in PFdl represent task rules and strategies (White and Wise, 1999, Wallis et al., 2001, Genovesio et al., 2005, Genovesio et al., 2008, Tsujimoto et al., 2008). Moreover, patients with PF damage show inability to change behavior after a task rule has changed suggesting that PF is fundamentally involved in the flexible adaptation to task switches (Anderson et al., 1991, Stuss et al., 2000, Goldstein et

al., 2004). To be able to achieve such flexibility, our brain needs to monitor the outcome so that rules can be updated or integrated. Our experimental tasks offered a stable environment for the monkeys in which the two rules were kept constant with no surprises. Monkeys had to report which of two stimuli, sequentially presented on a screen, was farther from a reference point or lasted longer in a distance and in a duration discrimination task, respectively. In our previous work (Genovesio et al., 2014, Marcos et al., 2016), we have shown that even in such cases of constant-rule context, PF neurons keep track of the previous trial outcome potentially helping to preserve or even improve the task performance and also possibly preparing to face plausible unexpected changes. We have also shown that the outcome representation in PF can vanish either after the presentation of an external event or with the passage of time depending on the specific neuronal dynamics (Marcos et al., 2016). The present analysis in the PFdl use part of the same dataset from previous studies (Genovesio et al., 2014, Marcos et al., 2016) and aim at extending our findings on outcome coding by examining for the first time the outcome modulation using two experimental task contexts to test the level of specificity/generality of outcome monitoring.

EXPERIMENTAL PROCEDURES

Experimental task

Two adult male rhesus monkeys (*Macaca mulatta*, 8.5 and 8.0 kg) performed a distance discrimination task and a duration discrimination task. Monkeys had to discriminate which of two stimuli sequentially presented on a screen was either farther from a reference point or longer in duration, respectively. They sat in a monkey chair with their head fixed and 29 cm from a video monitor. Three infrared switches were placed in front of the monkeys, within reach, as an interface between them and the tasks. All procedures conformed to the *Guide for the Care and Use of Laboratory Animals* (1996) and were approved by the National Institute of Mental Health Animal Care and Use Committee.

The temporal sequence of tasks events is shown in Figure 1A and 1B. The trials in both tasks had the same temporal structure but with different durations of the two main stimuli. Each stimuli could be either a blue circle of 3° diameter or red square of 3×3° dimensions. If the first stimulus (S1) was the blue circle then the second stimulus (S2) was the red square and viceversa. Each trial started when the monkeys pressed the central switch. A central stimulus was shown for 400 or 800 ms followed by the onset of S1. In the distance

task, the duration of S1 was always 1000 ms and it was always located at a distance of 8-48 mms (in steps of 8 mm) above or below the reference point. In the duration task, the presentation of S1 lasted from 200 ms to 1200 ms (in steps of 200 ms) and it was always presented at the center of the screen. After the disappearance of S1, a first delay (D1) of 400 or 800 ms preceded the subsequent presentation of S2. In the distance task, the duration of S1 was always 1000 ms and it appeared below the reference point if S1 had appeared above and above otherwise. Its exact distance was chosen from the same interval as S1 distances but it never equaled the distance of S1. In the duration task, the length of S2 was chosen from the same interval as S1's duration but without considering the one used for S1. In this case, S2 was always presented at the center of the screen. A second delay (D2) of 0, 400 or 800 ms followed the disappearance of S2 and preceded the reappearance of the two stimuli, which served as a "go" signal. Each stimulus was presented either 40 mm to the right or 40 mm to the left of the central stimulus, pseudorandomly chosen. The monkeys had to select, within 6 s, one of the two stimuli based on which one was presented either farther from the reference point in the distance task or had lasted longer in the duration task. Before the go signal the monkeys could not plan any motor response. Correct choices were rewarded with 0.1 ml fluid whereas incorrect choices were followed by an acoustic feedback. All variables of the task, such as duration of D1 and D2 or color and shape of the two stimuli, were pseudorandomly determined. For more details about the two experimental tasks, see Genovesio et al. (2009) and Genovesio et al. (2011).

Surgery

Recording chambers were implanted over the exposed dura mater of the left frontal lobe, along with head restraint devices. Aseptic techniques were used together with isofluorane anesthesia (1–3%, to effect). Monkey 1 had two, 18 mm-diameter chambers, and Monkey 2 had a single, 27×36-mm chamber.

Histological Analysis

Electrolytic lesions (15 mA for 10 s, anodal current) were made at selected locations. After 10 days, the animal was deeply anesthetized and afterward, perfused through the heart with formaldehyde-containing fixative. We plotted recording sites on Nissl-stained coronal sections by reference to the recovered electrolytic lesions and the marking pins inserted

during perfusion. Recordings were predominantly taken from area 8, area 46 and a small population of area 12. Figure 1C shows the recording sites.

Data Collection

We monitored eye position with an infrared oculometer (Arrington Recording, Scottsdale, AZ) and recorded single cells using quartz insulated, platinum-iridium electrodes (0.5–1.5 MΩ at 1 kHz), positioned by a 16-electrode drive assembly (Thomas Recording, Giessen, Germany). The electrodes were arranged in a concentric array with 518 mm spacing. Spikes were discriminated online using the Multichannel Acquisition Processor (Plexon, Dallas, TX) and confirmed with the Offline Sorter (Plexon).

Neural analyses

Neural stability. To assess the stability of the neurons recorded in both tasks, we calculated an index that combined the similarity of the mean waveform (W) and the interspike interval (ISIH) in the two tasks (Dickey et al., 2009). In brief, the similarity of W was calculated as the Pearson's correlation coefficient between the mean waveforms obtained in each task. The similarity of ISIH was assessed by first fitting the ISIH distributions from each task with a mixture of three log-normal distributions using an expectation algorithm and then computing a score (I) as:

$$I(A,B) = \sqrt{\sum_{i=1}^{i=8} \frac{(A_i - B_i)^2}{\sigma_i^2}}$$

where A and B represent the eight parameters used to fit the ISIH distributions in the distance and duration task, respectively, and σ is the variance of the parameters obtained from a sample set. Both scores W and I are then normalized and combined to obtain one unique score that will serve as the criterion to classify neurons as stable:

$$W' = tanh^{-1}(W)$$

$$I' = log(I)$$

$$S = (x - \mu_{pos})^T \Sigma_{pos}^{-1} (x - \mu_{pos}) - (x - \mu_{neg})^T \Sigma_{neg}^{-1} (x - \mu_{neg})$$

where $= \binom{W'}{I'}$, μ_{pos} and μ_{neg} are the mean score values of true-positives and true-negatives obtained from a sample set and Σ_{pos} and Σ_{neg} are their covariance. If the score S is lower than a threshold (T) the neuron is considered to be stable across tasks. The values of μ_{pos},μ_{neg}, Σ_{pos}, Σ_{neg} and T are those obtained from Dickey et al. (2009). From the total number of recorded neurons (N=443), we classified 97% of them as being stable across tasks (N=428).

Neural response modulation. For visual purposes, we computed the mean firing rate of the populations using windows of 50 ms in steps of 5 ms. The statistical difference between firing rate of the population in correct and incorrect trials was calculated using a two-tailed permutation test (10,000 iterations) calculated with non-overlapping windows of 50 ms. We only reported time bins as significant when more than two consecutive time bins showed significance.

To identify the neurons that modulated their activity in response to the feedback obtained, we sorted the trials by correct or incorrect and performed a one-way ANOVA. We used the neural activity calculated in the interval between 200 ms to 500 ms after the feedback was provided. Neural activity was used as the variable and outcome as the factor of the one-way ANOVA. The neurons that exhibited a significant modulation of their activity were classified as outcome selective in the duration task (n_{dur}) or outcome selective in the distance task (n_{dist}). A subset of these neurons was further classified as outcome selective in both (n_c). The significance of the number of neurons in this subset was assessed with a hypergeometric distribution test (Casella and Berger, 1990). The test estimates the probability that k samples or more, drawn from a group of N samples, belong to a specific subset of m samples. In our case, we calculated the probability of getting k or more neurons ($k=n_c$) outcome selective in the duration task when selecting m neurons ($m=n_{dist}$) from the N total number of neurons (see Marcos et al. (2017)).

We used the same interval for the receiver operating characteristic (ROC) analysis (Green and Swets, 1966). The ROC analysis provides an estimate of the performance of a binary classifier. In other words, it gives an estimate of overlap between preferred and nonpreferred firing rate distributions. The reported ROC values are defined as the area under the ROC curve (auROC) and they correspond to the probability of correctly classifying an individual data point. auROC values are between the interval 0 and 1 where 0 and 1 are maximum selectivity with opposed preference and 0.5 corresponds to non-selectivity. Normalized values of auROC are calculated within the interval 0.5 and 1, so that it considers only the absolute selectivity.

Neural response classification. We assessed how well individual neurons represented the goal features (blue and red choice) or the goal location (left and right) by using a classification method based on the peristimulus time histogram (Foffani and Moxon, 2004). We sorted the trials by goal features or goal location and we used the individual-trial neural activity in the period between 200 and 500 ms after feedback onset, calculated with a bin size of 100 ms, as the predictor variable. We calculated the proportion of correctly classified trials for each neuron individually. We selected one trial (test trial) and calculated a template of activity for each condition as the mean firing rate of all remaining trials. Then, to assess how well a trial could be classified, we calculated the Euclidean distance between the test trial and the template. The trial was classified as belonging to the condition with the lowest distance. We repeated this procedure for all trials recorded from a neuron.

To assess how well the population represented correct and incorrect outcomes, we used the same classification method using a neuron-dropping analysis (Foffani and Moxon, 2004, Lebedev et al., 2004, Lebedev, 2014). We sorted the trials by correct or incorrect outcome and we calculated the activity in the interval between 200 and 500 ms, with a bin size of 100 ms, after feedback onset. In each interaction, we randomly selected one trial from each neuron (test trials set) belonging to the same condition and calculated a template of activity for each neuron and condition without considering the test trials, as above. The test trials were classified as those that belong to the condition with the lowest sum of computed Euclidean distances calculated between each trial in the test trials set and its corresponding neuron's templates. We repeated this procedure 1,000 times for each condition. The neuron-dropping analysis consisted in removing one neuron from the group and repeating the same procedure until the group was composed by only one neuron.

RESULTS

Two monkeys performed the distance and duration discrimination tasks (Figure 1A). Their performance was high and similar in both tasks (79% for distance task and 81% for duration task for Monkey 1 and 80% for both distance and duration tasks for Monkey 2). Moreover, their responses were faster and more accurate for easier discriminations. For a detailed description of their behavior refer to Genovesio et al. (2009) and Genovesio et al. (2011).

To investigate the outcome coding at the population level, we sorted the trials by correct or incorrect and calculated the mean firing rate of all recorded neurons aligned to

feedback onset. We observed that after feedback onset the mean firing rate of the population significantly differed between correct and incorrect cases (Figure 2). In particular, the mean firing rate was higher for incorrect than correct trials in both tasks, but for a slightly longer time interval for the distance (left panel of Figure 2A) than for the duration task (left panel of Figure 2B). In both tasks, the population activity was significantly higher for incorrect than correct trials in the interval between 200 ms and 500 ms after feedback onset. Next, we calculated individual firing rates in this interval for incorrect versus correct choices. Right panels of Figure 2 show the individual modulation of activity during distance and duration tasks. Black dots correspond to neurons that significantly modulated their activity related to the trial outcome. Consistent with the mean firing rate of the population, the majority of the neurons responded with higher activity for incorrect than for correct outcomes.

Next, we identified the neurons that significantly coded the outcome in one or both tasks. With this method, from the initial database of 428 neurons, 77 neurons were classified as outcome selective in both tasks whereas 61 neurons were classified as outcome selective only in the distance task and 53 neurons as only outcome selective in the duration task (Figure 3A). The number of overlapping neurons was significantly greater than that the one expected by chance (n_{chance}=49; $p < 10^{-15}$, hypergeometric distribution test) indicating a relationship between the occurrence of an outcome neural response in one task and the one in the other task. Consistent with this result, the number of overlapping neurons was also significantly higher than expected by chance (n_{chance}=38; $p < 10^{-7}$, hypergeometric distribution test; 60 outcome-selective neurons only in the distance task, 61 only in the duration task and 55 in both tasks) for an equal number of correct and incorrect trials (mean ± SEM, 30.82 ± 0.55 in the distance task and 28.46 ± 0.56 in the duration task), i.e. when only the first correct trials that matched the number of incorrect trials were considered for analysis. Interestingly, a significant proportion of neurons exhibited higher firing rate for incorrect than for correct outcomes (70% and 65% of neurons for distance and duration tasks, respectively; $p < 0.001$ in both tasks, binomial test). Moreover, most of the neurons maintained the same preference across tasks even if they did not show a significant modulation of their activity (bottom panel of Figure 3A). Indeed, 75% (46/61) of the outcome-selective neurons only in the distance task showed the same preference in the duration task and, in the same way, 83% (44/53) of the outcome-selective neurons only in the duration task showed the same preference in the distance task. To assess their degree of selectivity, we used a ROC analysis and plotted the auROC values of one task against the

other for all categories of neurons. The neurons that were significantly modulated in both tasks exhibited a higher degree of selectivity compared to the context-specific ones ($p < 0.001$ when comparing normalized values of auROC during the distance and duration task separately). Interestingly, even if the neurons significantly modulated their outcome related activity only in one task context, their preference for correct or incorrect outcomes was often shared also in the other task as demonstrated by the high and significant value of correlation between auROC values in both tasks ($p < 10^{-7}$, $r = 0.57$ and $r = 0.66$ for outcome-selective neurons in the distance and the duration tasks, respectively). Thus, even if differences of outcome response magnitude exist, neural outcome selectivity might be, in most neurons, represented in a generalist manner in PF.

Figure 4 shows the population activity of the outcome-selective neurons. We considered the neurons that modulated their activity in response to the outcome in at least one task and we further divided the neurons by their preferred outcome, i.e. correct or incorrect. As expected, the difference in neural activity between preferred and nonpreferred outcomes began to be significant soon after the feedback onset. The time onset of this difference (mean ± SEM, 169.29 ms ± 26.38 ms and 173.72 ms ± 22.19 ms for correct- and incorrect-selective neurons, respectively, in the distance task and 126.90 ms ± 22.41 ms and 190.70 ms ± 24.07 ms for correct and incorrect-selective neurons, respectively, in the duration task) was not significant between correct and incorrect trials, either within each task or for the same outcome across tasks ($p > 0.05$, Mann-Whitney U-test).

We further investigated whether the outcome-selective neurons were modulated by the goal features and goal location during correct and incorrect trials. To do so, we performed a classification (see Experimental Procedures) using the mean activity of neurons in the interval between 200 and 500 ms after feedback with trials sorted by blue or red goal or by left or right location. This method allowed us to quantify the amount of information that can be decoded from the activity of neurons. The decoding values for goal features and location in correct and incorrect trials were all above chance in the two tasks (Figure 5). Importantly, although the decoding of goal features and location for the two possible outcomes did not significantly differ in any case ($p > 0.05$, paired t-test), we observed that the proportion of correctly decoded goal features was greater and the proportion of correctly decoded goal location was lower in the incorrect trials compared to the correct trials. Although not significant, these differences in goal coding were common to both tasks showing a further similarity in the response of outcome-selective neurons across tasks.

Figure 6 shows two examples of outcome-selective neurons. Figure 6A shows a neuron that showed significantly greater activity after incorrect than correct responses in both tasks. In this case, the activity of the neuron was modulated by the outcome after the feedback onset and during our interval of analyses (200 ms to 500 ms after feedback) in the distance (left panel of Figure 6A) and in the duration tasks (right panel of Figure 6A). The other neuron presented in Figure 6B was significantly modulated by the outcome in the distance (left panel of Figure 6B) but not in the duration task (right panel of Figure 6B). However, although this neuron was not outcome selective in the duration task, it exhibited a non-significant higher activity after errors, indicating that this neuron showed a similar trend in firing rate following errors during both tasks. The auROC values of these two examples are pointed by arrows in Figure 3B.

To assess whether the neurons classified by the one-way ANOVA as outcome selective only in one task might, similarly to the neuron presented in Figure 3B, also exhibit an outcome modulation in the other task although with less strength, we performed a classification analysis. To this purpose, we computed the percentage of correct classification when looking at the activity, sorted by correct/incorrect trials, of one or more neurons (see Experimental Procedures). The percentage of correctly classified trials, as correct or incorrect outcome, was notably higher for neurons which represent the outcome in both tasks (Figure 6C and 6D). Interestingly, the classification for the neurons that significantly encoded the outcome in one task, when tested in the task in which they did not significantly modulate their activity was notably above chance (Figure 6C and 6D). This indicates that, although with less power, most neurons represented the outcome independently of the context even when their outcome-related modulation did not reach statistical significance.

DISCUSSION

In the present study, we examined the context generality or specificity of the outcome representation in PFdl. Our analyses focused on two different perceptual discrimination tasks that required applying different rules: choose the farther or the longer stimulus, in the distance or the duration discrimination task, respectively. We have tested the independence of the outcome modulation in these two tasks and we have found that neurons in PFdl were similarly modulated in both cases, rather than in a context independent way. In other words, most of the neurons that represented the outcome in one context were, with similar

selectivity, also doing so in the other context as shown by the high correlation in the outcome auROC values. Nevertheless, we observed that the task could have a modulatory influence on the outcome signal but mainly in the decrease of neural response selectivity.

Our experimental setup provides an ideal framework to investigate the specificity or the generality of outcome coding in PFdl. The two tasks were alike but only on the surface. On the one hand, the sequence of events was equivalent in the two experimental tasks with only changes in the duration and position of the two stimuli. On the other hand, the tasks differed in the rule – duration or distance discrimination - that needed to be applied in other to successfully perform a trial. Thus, task structures were as similar as possible for allowing comparison while the difference in the behavioral rule provided two clearly distinguishable contexts. Moreover, correct and incorrect trials were signaled equally in both cases allowing for a controlled comparison between outcome coding in two different task contexts. Our study complements that from Kennerley and Wallis (2009) in which the coding of the outcome under different manipulations of its value was investigated. Although less prevalently than neurons in the anterior cingulate cortex, they found that many cells in PFdl coded the outcome of a choice across multiple decision variables, such as presence of reward, payoff or cost. Our results show that neurons in PFdl also code the same choice outcome across two different task contexts.

In order to optimally exploit an environment, animals need to learn which actions lead to desired outcomes. When faced with an unknown context, proper responses are learned by exploring which of them maximizes reward. Such learning that includes trial and error and delayed reward is called reinforcement learning (Sutton and Barto, 1998) and it takes place when animals learn a new task or when faced with a dynamic context. In these situations, they learn and update the association between context and actions to maximize their reward rate. Some neurophysiological studies have examined the encoding of choice and outcome during reinforcement learning tasks (Barraclough et al., 2004, Tsujimoto and Sawaguchi, 2005, Seo et al., 2007). It has been suggested that PFdl is involved in updating the decision strategy based on previous choice and outcome during an oculomotor free-choice task in which the probability of reward depends on an opponent's choice (Barraclough et al., 2004, Seo et al., 2007).

In our case, the monkeys were well trained before the beginning of the recordings and the task rules constant and well known by them. Nevertheless, outcome evaluation and past

goal or response monitoring (Tsujimoto et al., 2012, Genovesio et al., 2014) can be advantageous also beyond reinforcement learning in the context of the exploitation/exploration trade off to explore new possibilities for which PF cortex is thought to play an important role (Koechlin, 2014). For example, outcome monitoring could help to discover strategies that can speed-up learning, as it has been found in the conditional learning paradigm in which monkeys could adopt the repeat-stay or the change-shift strategies to increase success rate (Bussey et al., 2001, Genovesio et al., 2005). The outcome monitoring in our task might be essential not to learn to discriminate magnitudes but rather to maintain the monkey's behavioral performance by continuously reinforcing the high order rule to choose the stimulus with the greatest magnitude. Alternative simpler rules such as selecting the red square, the blue circle, the first or the second presented object, the right or the left goal should always be discarded and it is possible that goal and outcome monitoring contributed to suppress potential alternative rules. Indeed, Buckley et al. (2009) reported that lesions to periprincipal sulcus region, but not to other parts of PF, impair maintenance of abstract rules. Moreover, the monitoring of the outcome is also critical for flexible adaptation to unexpected changes in the environment (Sakai, 2008, Stoet and Snyder, 2009) but our study could not address the role of the outcome neurons when rules change. It is also possible that the outcome information could contribute to maintain the different thresholds of the two tasks, although the feedback delivered from the outcome did not provide information on the actual magnitudes of the two stimuli.

We have also tested whether the coding of goal features and goal location is greater in correct compared to error trials, as reported in Tsujimoto et al. (2011), but we did not observe any significant difference. Based on the contrast between orbital PF (PFo) and PFdl, Tsujimoto et al. (2011) discussed their findings from a model that PFo provides a strategy signal to PFdl, which selects a response based on the strategy and the previous response. Thus, the discrepancy between this previous study and our current finding may stem from the fact that in Tsujimoto et al. (2011) the coding of goal location in the error trials was soon replaced by the representation of the future goal, which could be determined after negative feedback.

In an additional study, using memory- and sensory-guided saccade tasks, Tsujimoto and Sawaguchi (2005) showed that a subset of neurons in PF coded the response and outcome differently for memory or sensory decisions in probabilistic tasks in which the reward was immediately delivered or with delay. In this case, after the choice, the immediate

or delayed delivery of the reward occurred with 50% of probability and thus the outcome was not critical for the animals in order to adapt their behavioral policies. In our study, however, the outcome represented a feedback on the correctness of their perceptual discriminations. This difference might account for the higher context dependent specificity described in the previous study compared to ours. One possible interpretation is that when the reward/no reward event has no relevance as feedback on the success or failure it becomes represented within a specific local task-related circuit as for other variables like task epochs durations (Genovesio et al., 2016) or values of magnitude (Marcos et al., 2017).

Our findings resonates instead with the high level of rule-independent outcome modulation found by Mansouri et al. (2006) when monkeys switched between matching to sample rules in a Wisconsin Card Sorting. Most of the neurons showed error-related modulation relative to the pre-feedback period in a rule-independent manner. An important difference between our results and their earlier results (Mansouri et al., 2006) is that in their task errors occurred mainly in the transition from one rule to the other and therefore the error signal represented also a loose-shift. Hence, the same neurons could increase activity after errors in trials with different rules simply because they signaled the loose-shift rule in both tasks rather than error per se. In our study, instead, each task rule could not be applied in the other task and the outcome signal could not represent a loose-shift signal but just inform that an error occurred when applying the task rule.

The outcome monitoring function should be also considered in the context of the PF-basal ganglia circuit. The PF-basal ganglia pathways constitute an integrative system where the two partners seem to play distinct roles in the response learning process. According to this view, PF plays its role when new rules need to be learned to replace the previous ones whereas basal ganglia could make previously learned rules more effective based on environmental context (Wise et al., 1996) via their cholinergic and dopaminergic neuromodulator systems, which differently shape the projection neurons responses to reward-related events (Morris et al., 2004). The dopaminergic innervation of PF neurons is not as dense as in the basal ganglia and PF dopamine seems to play its main role in the learning of novel associations and not in maintaining the familiar ones (Puig and Miller, 2012, Puig et al., 2014, Puig and Miller, 2015). In our study, there were no new rules or new associations to learn and therefore, the context-independent signal observed could then just reflect the monitoring of the behavior leading to the reward (Ichihara-Takeda and Funahashi, 2006). Indeed, PFdl neurons are not much sensitive to free-reward delivery in contrast to

orbitofrontal neurons and their activity during the reward period reflects the achievement or not of a correct performance (Ichihara-Takeda and Funahashi, 2006). The outcome modulation observed in our task might therefore have a monitoring function. This unspecific or context-independent signal about the correctness of the trial can be in turn conveyed to the input structures of the basal ganglia where signals from several prefrontal cortical areas converge (Haber and Knutson, 2010) and in which both timing signals (Chiba et al., 2008) and context-dependent reward coding (Apicella, 2007) have been described.

The coding of different task events in PF has been found to be both general and specific. On the same dataset on duration and distance magnitude discrimination we have previously assessed different types of task context dependency among the same population of cells. We found that similarly to the outcome coding of the present study, PF neurons represented the future goal independently of the task context (Genovesio et al., 2012). On the other hand, we have shown that neurons in PF represented the duration of a task interval in a task context dependent manner (Genovesio et al., 2016). Likewise, both the absolute and the relative coding of a magnitude – distance/duration - were highly context dependent (Genovesio et al., 2012, Marcos et al., 2017). For example, neurons with higher activity in working memory for long durations of a stimulus did not exhibit the same response modulation for long distances of a stimulus (Genovesio et al., 2012, Marcos et al., 2017). In summary, outcome and goal seem to be the most general representations in contrast to neurons coding absolute and relative values of magnitude in a context dependent manner. Our new results represent a step further in the understanding of the level of overlap between the neural computations that contribute to space and time coding. Thus, they help to enrich the debate on whether there is or not a common magnitude representation, going beyond the simplest questions focused on perception or estimation levels.

Past studies focusing on reward-related modulation in the orbitofrontal cortex have shown that the outcome can be represented independently on how the reward is signaled (Thorpe et al., 1983, Morrison et al., 2011). The function of individual neurons in economic decisions has been shown to be stable in different contexts although neurons remap to encode goods differently across experimental blocks (Xie and Padoa-Schioppa, 2016). Future studies should address whether such stability in the orbitofrontal cortex extends also to the outcome modulation in terms of success and failure.

ACKNOLEDGMENTS

We thank Ms. Ping-yu Chen for the histological material, Dr. Andrew Mitz for engineering, and Dr. Steve Wise for his support during all phases of this project. This work was partially supported by the Division of Intramural Research of the National Institute of Mental Health (Z01MH-01092) and by the Italian Ministry of Research, "Fondo per gli investimenti della ricerca di base" FIRB 2010 (Protocol number: RBFR10G5W9_001).

AUTHOR CONTRIBUTIONS

A.G. and S.T. collected the data. E.M., S.N. and A.G. designed the study. E.M. analyzed the data and prepared the figures. All authors interpreted the results and wrote the paper.

ADDITIONAL INFORMATION

Competing financial interests: The authors declare no competing financial interests.

REFERENCES

Amador N, Schlag-Rey M, Schlag J (2000) Reward-predicting and reward-detecting neuronal activity in the primate supplementary eye field. J Neurophysiol 84:2166-2170.

Anderson SW, Damasio H, Jones RD, Tranel D (1991) Wisconsin Card Sorting Test performance as a measure of frontal lobe damage. J Clin Exp Neuropsychol 13:909-922.

Apicella P (2007) Leading tonically active neurons of the striatum from reward detection to context recognition. Trends Neurosci 30:299-306.

Apicella P, Deffains M, Ravel S, Legallet E (2009) Tonically active neurons in the striatum differentiate between delivery and omission of expected reward in a probabilistic task context. Eur J Neurosci 30.

Apicella P, Ljungberg T, Scarnati E, Schultz W (1991) Responses to reward in monkey dorsal and ventral striatum. Exp Brain Res 85:491-500.

Barraclough D, Conroy ML, Lee D (2004) Prefrontal cortex and decision making in a mixed-strategy game. Nat Neurosci 7:404-410.

Beck JM, Ma WJ, Pitkow X, Latham PE, Pouget A (2012) Not noisy, just wrong: the role of suboptimal inference in behavioral variability. Neuron 74:30-39.

Bowman EM, Aigner TG, Richmond BJ (1996) Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. J Neurophysiol 75:1061-1673.

Bussey TJ, Wise SP, Murray EA (2001) The role of ventral and orbital prefrontal cortex in conditional visuomotor learning and strategy use in rhesus monkeys. Behav Neurosci 115:971-982.

Casella G, Berger RL (1990) Statistical Inference. Belmont, CA: Duxbury.

Chiba A, Oshio K, Inase M (2008) Striatal neurons encoded temporal information in duration discrimination task. Exp Brain Res 186:671-676.

Dickey AS, Suminski A, Amit Y, Hatsopoulos NG (2009) Single-unit stability using chronically implanted multielectrode arrays. J Neurophysiol 102:1331-1339.

Foffani G, Moxon KA (2004) PSTH-based classification of sensory stimuli using ensembles of single neurons. J Neurosci Methods 135:107-120.

Genovesio A, Brasted PJ, Mitz AR, Wise SP (2005) Prefrontal cortex activity related to abstract response strategies. Neuron 47:307-320.

Genovesio A, Seitz LK, Tsujimoto S, Wise SP (2016) Context-dependent duration signals in the primate prefrontal cortex. Cereb Cortex 26:3345-3356.

Genovesio A, Tsujimoto S, Navarra G, Falcone R, Wise SP (2014) Autonomous encoding of irrelevant goals and outcomes by prefrontal cortex neurons. J Neurosci 34:1970-1978.

Genovesio A, Tsujimoto S, Wise SP (2008) Encoding problem-solving strategies in prefrontal cortex: activity during strategic errors. Eur J Neurosci 27:984-990.

Genovesio A, Tsujimoto S, Wise SP (2009) Feature- and Order-Based Timing Representations in the Frontal Cortex. Neuron 63:254-266.

Genovesio A, Tsujimoto S, Wise SP (2011) Prefrontal cortex activity during the discrimination of relative distance. J Neurosci 31:3968-3980.

Genovesio A, Tsujimoto S, Wise SP (2012) Encoding goals but not abstract magnitude in the primate prefrontal cortex. Neuron 74:656-662.

Gold JI, Shadlen MN (2007) The neural basis of decision making. Annu Rev Neurosci 30:535-574.

Goldstein B, Obrzut JE, John C, Ledakis G, Armstrong CL (2004) The impact of frontal and non-frontal brain tumor lesions on Wisconsin Card Sorting Test performance. Brain Cogn 54:110-116.

Green DM, Swets JA (1966) Signal Detection Theory and Psychophysics. New York: Wiley.

Haber SN, Knutson B (2010) The reward circuit: linking primate anatomy and human imaging. Neuropsychopharmacology 35:4-26.

Hikosaka O, Sakamoto M, Usui S (1989) Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. J Neurophysiol 61:814-832.

Hollerman JR, Tremblay L, Schultz W (1998) Influence of reward expectation on behavior-related neuronal activity in primate striatum. J Neurophysiol 80:947-963.

Ichihara-Takeda S, Funahashi S (2006) Reward-period activity in primate dorsolateral prefrontal and orbitofrontal neurons is affected by reward schedules. J Cogn Neurosci 18:212-226.

Kennerley SW, Wallis JD (2009) Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. Eur J Neurosci 29:2061-2073.

Kobayashi S, Nomoto K, Watanabe M, Hikosaka O, Schultz W, Sakagami M (2006) Influences of rewarding and aversive outcomes on activity in macaque lateral prefrontal cortex. Neuron 51:861-870.

Koechlin E (2014) An evolutionary computational theory of prefrontal executive function in decision-making. Philos Trans R Soc Lond B Biol Sci 369.

Lebedev MA (2014) How to read neuron-dropping curves? Frontiers in Systems Neuroscience 8:102.

Lebedev MA, Messinger A, Kralik JD, Wise SP (2004) Representation of attended versus remembered locations in prefrontal cortex. PLoS Biol 2:e365.

Mansouri FA, Matsumoto K, Tanaka K (2006) Prefrontal cell activities related to monkeys' success and failure in adapting to rule changes in a wisconsin card sorting test analog. J Neurosci 26:2745-2756.

Marcos E, Pani P, Brunamonti E, Deco G, Ferraina S, Verschure P (2013) Neural variability in premotor cortex is modulated by trial history and predicts behavioral performance. Neuron 78:249-255.

Marcos E, Tsujimoto S, Genovesio A (2016) Event- and time-dependent decline of outcome information in the primate prefrontal cortex. Sci Rep 6:25622.

Marcos E, Tsujimoto S, Genovesio A (2017) Independent coding of absolute duration and distance magnitudes in the prefrontal cortex. J Neurophysiol 117:195-203.

Morris G, Arkadir D, Nevet A, Vaadia E, Bergman H (2004) Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. Neuron 43:133-143.

Morrison SE, Saez A, Lau B, Salzman CD (2011) Different time courses for learning-related changes in amygdala and orbitofrontal cortex. Neuron 71.

Nakamura K, Mikami A, Kubota K (1992) Activity of single neurons in the monkey amygdala during performance of a visual discrimination task. J Neurophysiol 67:1447-1463.

Puig MV, Antzoulatos EG, Miller EK (2014) Prefrontal dopamine in associative learning and memory. 282:217-229.

Puig MV, Miller EK (2012) The role of prefrontal dopamine D1 receptors in the neural mechanisms of associative learning. Neuron 74:874-886.

Puig MV, Miller EK (2015) Neural substrates of dopamine D2 receptor modulated executive functions in the monkey prefrontal cortex. Cereb Cortex 25:2980-2987.

Ramakrishnan A, Byun YW, Rand K, Pedersen CE, Lebedev MA, Nicolelis MAL (2017) Cortical neurons multiplex reward-related signals along with sensory and motor information. Proceedings of the National Academy of Sciences of the United States of America 114:E4841-E4850.

Ravel S, Sardo P, Legallet E, Apicella P (2001) Reward unpredictability inside and outside of a task context as a determinant of the responses of tonically active neurons in the monkey striatum. J Neurosci 21:5730-5739.

Sakai K (2008) Task set and prefrontal cortex. Annu Rev Neurosci 31:219-245.

Sardo P, Ravel S, Legallet E, Apicella P (2000) Influence of the predicted time of stimuli eliciting movements on responses of tonically active neurons in the monkey striatum. Eur J Neurosci 12:1801-1816.

Seo H, Barraclough DJ, Lee D (2007) Dynamic signals related to choices and outcomes in the dorsolateral prefrontal cortex. Cereb Cortex 17:i110-i117.

Smith PL, Ratcliff R (2004) Psychology and neurobiology of simple decisions. Trends Neurosci 27:161-168.

Stoet G, Snyder LH (2009) Neural correlates of executive control functions in the monkey. Trends Cogn Sci 13:228-234.

Stuss DT, Levine B, Alexander MP, Hong J, Palumbo C, Hamer L, Murphy KJ, Izukawa D (2000) Wisconsin Card Sorting Test performance in patients with focal frontal and posterior brain damage: effects of lesion location and test structure on separable cognitive processes. Neuropsychologia 38:388-402.

Sutton RS, Barto AG (1998) Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA.

Thorpe SJ, Rolls ET, Maddison S (1983) The orbitofrontal cortex: neuronal activity in the behaving monkey. Exp Brain Res 49:93-115.

Tremblay L, Schultz W (2000) Reward-related neuronal activity during go-nogo task performance in primate orbitofrontal cortex. J Neurophysiol 83:1864-1876.

Tsujimoto S, Genovesio A, Wise SP (2008) Transient Neuronal Correlations Underlying Goal Selection and Maintenance in Prefrontal Cortex. Cereb Cortex 18:2748-2761.

Tsujimoto S, Genovesio A, Wise SP (2011) Comparison of strategy signals in the dorsolateral and orbital prefrontal cortex. J Neurosci 31:4583-4592.

Tsujimoto S, Genovesio A, Wise SP (2012) Neuronal activity during a cued strategy task: comparison of dorsolateral, orbital, and polar prefrontal cortex. J Neurosci 32:11017-11031.

Tsujimoto S, Sawaguchi T (2005) Context-dependent representation of response-outcome in monkey prefrontal neurons. Cereb Cortex 15:888-898.

Wallis JD, Anderson KC, Miller EK (2001) Single neurons in prefrontal cortex encode abstract rules. Nature 411:953-956.

Watanabe M (1989) The appropriateness of behavioral responses coded in post-trial activity of primate prefrontal units. Neurosci Lett 101:113-117.

White IM, Wise SP (1999) Rule-dependent neuronal activity in the prefrontal cortex. Exp Brain Res 126:315-335.

Wise SP, Murray EA, Gerfen CR (1996) The frontal cortex - basal ganglia system in primates. Crit Rev Neurobiol 10:317-356.

Xie J, Padoa-Schioppa C (2016) Neuronal remapping and circuit persistence in economic decisions. Nat Neurosci 19:855-861.

Figure captions

Figure 1. Experimental tasks and recording locations. Temporal presentation of events during distance **(A)** and duration **(B)** discrimination tasks. In both tasks, two stimuli were sequentially presented on a screen, separated by a delay and followed by the reappearance of the two stimuli that served as targets. Monkeys had to select which of the two stimuli was presented farther, in the distance task, or lasted longer, in the duration task. After their choice feedback (Fb) was provided. **(C)** Location sites of the recordings obtained from the two monkeys.

Figure 2. Mean firing (FR) of the neural population and individual cells. Trials are sorted by correct and incorrect choices. Left plots, temporal dynamics of the population aligned to the time of feedback (Fb) for both the distance **(A)** and the duration **(B)** tasks. In both cases, green lines are trials that resulted in incorrect responses whereas orange are those that were correct. The difference between firing rates is significantly different between conditions in the interval between 100 ms and 550 ms after feedback in the distance task (A) and between 200 ms and 500 ms after feedback in the duration task (B) ($p < 0.025$, two-tailed permutation test). Shaded areas are SEM. Right plots, mean activity calculated in the period between 200 ms and 500 ms after feedback delivery for correct against incorrect trials in the distance **(A)** and duration **(B)** tasks. Black dots are the neurons that significantly modulated their activity in this period.

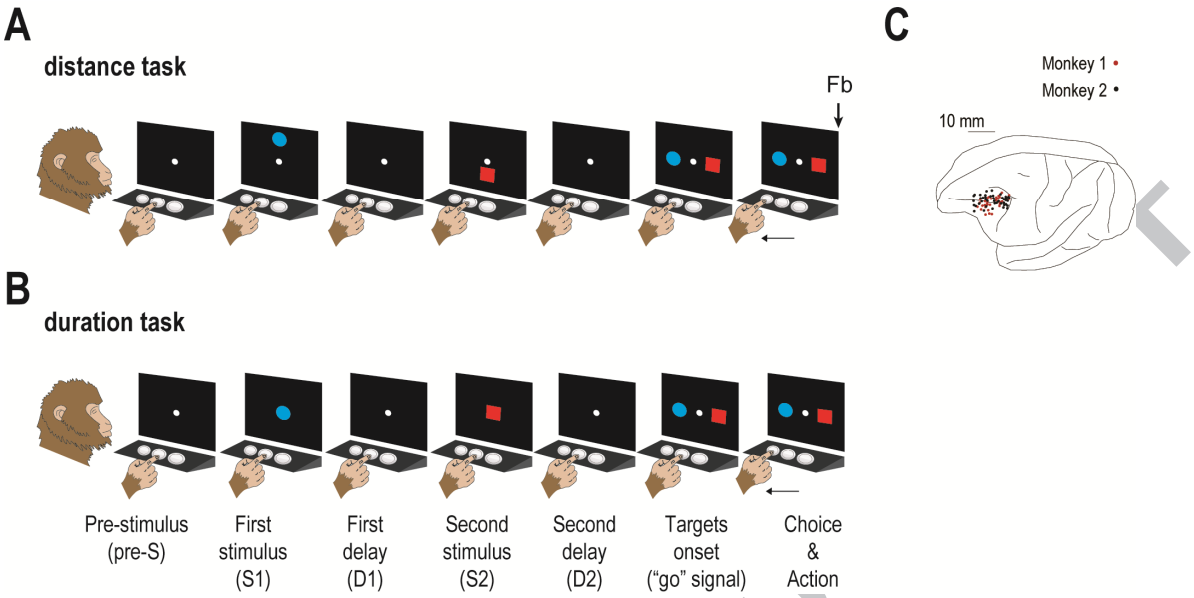
Figure 3. Selectivity and classification of outcome neurons. **(A)** Number of neurons belonging to each group: outcome selective only in the distance task, common neurons and outcome selective only in the duration task. The bottom inset indicates the proportion of neurons in each group that maintained preference in both tasks. **(B)** auROC values of neurons modulated by outcome. Gray dots are the values for neurons with their activity significantly modulated only in one of the two tasks (cyan, distance task and red, duration task). Black dots auROC values calculated for neurons that significantly coded the outcome in both tasks (common neurons). The auROC values are significantly correlated for all cases ($p<10^{-7}$; $r=0.77$ for all cases, $r=0.89$ for common neurons, $r=0.57$ for outcome selective in the distance task, $r=0.66$ for outcome selective in the duration task). Green arrows indicate the two examples of Figure 4.

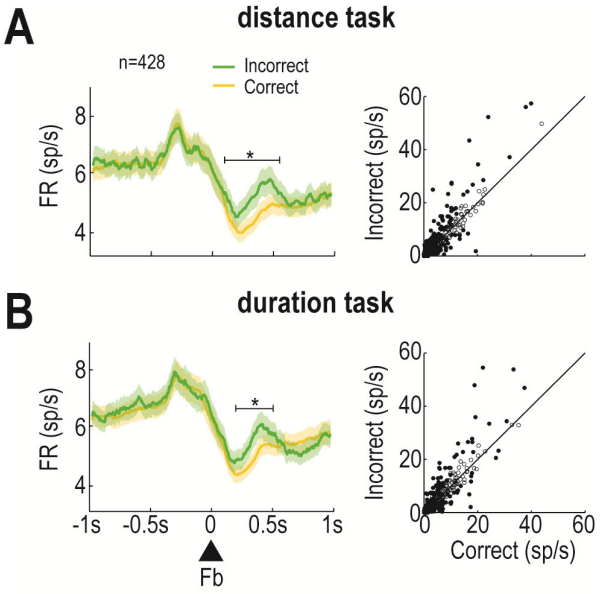
Figure 4. Population activity of outcome-selective neurons. Mean activity of outcome-selective neurons aligned to the time of feedback (Fb) during the distance task **(A)** and the

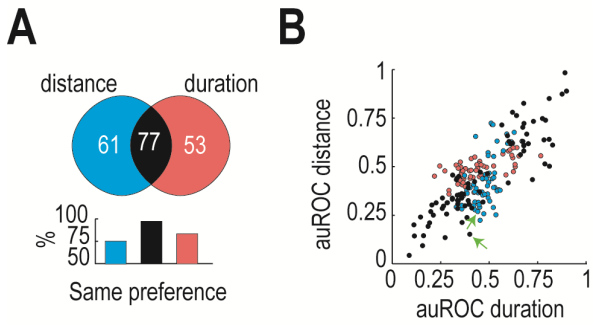
duration task **(B)** for incorrect (green) and correct (orange) trials for neurons sorted by correct (*Left panels*) or incorrect (*Right panels*) preferred outcome (* = $p < 0.025$, two-tailed permutation test). Shaded areas are SEM.

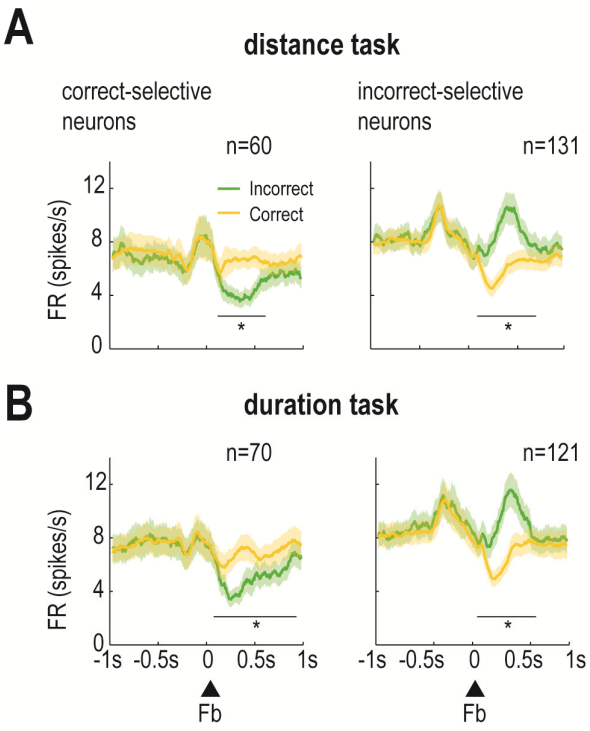
Figure 5. Percentage of correct classified trials for goal features and location. Proportion of trials correctly classified as blue or red goal (*Left panels*) and as left or right location (*Right panels*) sorted by correct (dark gray) and incorrect (light gray) trials for the distance **(A)** and the duration **(B)** tasks when all outcome-selective neurons (n=191) are used. Error bars are SEM.

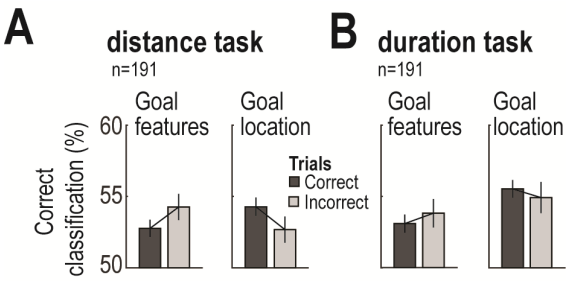
Figure 6. Example of outcome neurons and percentage of correct trials classification with neuron-dropping analysis. **(A)** Raster plot and mean firing rate of a neuron with higher activity for incorrect (green) than for correct (orange) trials in both tasks. **(B)** Example of a neuron significantly selective for incorrect trials in the distance but not the duration task. In both examples, neural activity is aligned to the feedback. **(C-D)** Proportion of correctly classified trials for distance **(C)** and duration **(D)** tasks. The activity of neurons divided by groups based on their previous classification (cyan for outcome selective only in the distance task, black for common neurons and red for outcome selective only in the duration task) is used for decoding. Dashed line indicates the chance level.

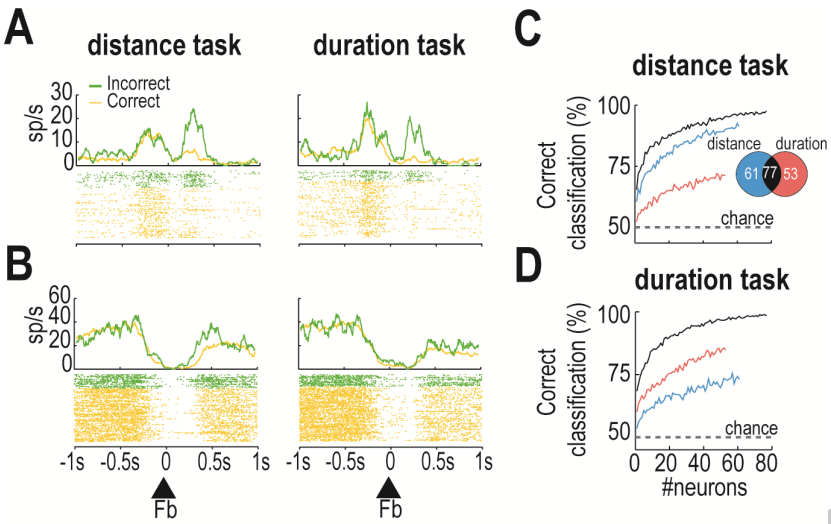












Highlights

- The outcome of a response is coded at the population level in PFdl.
- Outcome-selective neurons code the outcome independently of the task context.
- PFdl neurons have similar outcome selectivity across tasks.