Differential Encoding of Action Selection by Orbitofrontal and Striatal Population Dynamics

Authors:

Long Yang[1] and Sotiris C. Masmanidis[1,2,*]

Affiliations:

[1]Department of Neurobiology, University of California Los Angeles, Los Angeles, California 90095, USA.

[2]California Nanosystems Institute, University of California Los Angeles, Los Angeles, California 90095, USA.

***Address for correspondence**

Sotiris Masmanidis, 650 Charles E Young Dr. South, Los Angeles, CA 90095, USA.

E-mail: smasmanidis@ucla.edu

ORCID numbers:

Long Yang: 0000-0001-8317-8768

Sotiris C. Masmanidis: 0000-0002-8699-3335

Running head:

Differential frontostriatal encoding of action selection

Abstract:

Survival relies on the ability to flexibly choose between different actions according to varying environmental circumstances. Many lines of evidence indicate that action selection involves signaling in corticostriatal circuits, including the orbitofrontal cortex (OFC) and dorsomedial striatum (DMS). While choice-specific responses have been found in individual neurons from both areas, it is unclear whether populations of OFC or DMS neurons are better at encoding an animal's choice. To address this, we trained head-fixed mice to perform an auditory guided two-alternative choice task, which required moving a joystick forward or backward. We then used silicon microprobes to simultaneously measure the spiking activity of OFC and DMS ensembles, allowing us to directly compare population dynamics between these areas within the same animals. Consistent with previous literature, both areas contained neurons that were selective for specific stimulus-action associations. However, analysis of concurrently recorded ensemble activity revealed that the animal's trial-by-trial behavior could be decoded more accurately from DMS dynamics. These results reveal substantial regional differences in encoding action selection, suggesting that DMS neural dynamics are more specialized than OFC at representing an animal's choice of action.

New and Noteworthy:

While previous literature shows that both OFC and DMS represent information relevant to selecting specific actions, few studies have directly compared neural signals between these areas. Here we compared OFC and DMS dynamics in mice performing a two-alternative choice task. We found that the animal's choice could be decoded more accurately from DMS population activity. This work provides among the first evidence that OFC and DMS differentially represent information about an animal's selected action.

Introduction:

Solving the problem of which action to choose is a critical and incompletely understood aspect of brain function. Previous work has revealed this process relies strongly on the basal ganglia and their interactions with cortical networks (Alexander et al. 1986; Frank 2011; Humphries et al. 2006; Mink 1996). Corticostriatal circuits play a prominent role in enabling animals to choose favorable actions (Hwang et al. 2019; Pennartz et al. 2009; Rothwell et al. 2015; Sharpe et al. 2019; Znamenskiy and Zador 2013). Among these circuits are the orbitofrontal cortex (OFC) and dorsomedial striatum (DMS), which have both been shown to influence action selection (Balleine et al. 2007; Bradfield et al. 2015; Murray and Izquierdo 2007; Ostlund and Balleine 2007; Tai et al. 2012; Yin et al. 2005). Altered function of projections from OFC to DMS, or neighboring corticostriatal pathways, has been implicated in maladaptive action selection processes found in obsessive compulsive disorder (Ahmari et al. 2013; Burguiere et al. 2013; Corbit et al. 2019). A separate line of evidence relying on recordings of neural activity, has shown that both areas encode information that is relevant for action selection (Feierstein et al. 2006; Gremel and Costa 2013; Guo et al. 2019; Ito and Doya 2015; Kimchi and Laubach 2009; Moorman and Aston-Jones 2014; Nonomura et al. 2018; Schultz and Romo 1992; Seo et al. 2012; Shin et al. 2018; Stalnaker et al. 2012). However, information processing in these areas is diverse and not exclusively linked to action selection (Padoa-Schioppa and Assad 2006; Samejima et al. 2005; Wang et al. 2013; Wilson et al. 2014). Therefore, while behavioral studies have established a role for OFC and DMS in action selection, the extent to which encoding of an animal's choice of action differs or agrees between these areas is less clear (Seo et al. 2012; Sharpe et al. 2019). Since the striatum integrates excitatory input from multiple sources (Friedman et al. 2015; Guo et al. 2019; Lee et al. 2019; Reig and Silberberg 2014; Sippy et al. 2015), DMS dynamics may diverge significantly from those in OFC. However, until now there has not been a systematic effort to compare concurrently measured neural activity in these areas.

Here we compared simultaneously recorded dynamics in OFC and DMS using a high throughput electrode measurement technique, which gave access to the spiking activity of dozens of neurons in each area. Recordings were carried out in head-fixed mice performing an auditory guided two-alternative choice task. Using population decoding methods, we found that DMS activity performed better than OFC at representing an animal's choice of action. These regional differences were absent in data from another group of animals which were trained on a one-alternative choice task, confirming that the differential effects reflect action- rather than auditory tone-specific activity. Moreover, the results of population decoding analysis were more consistent than those from analyzing the selectivity of individual neurons, suggesting that monitoring dynamics of neural ensembles provides an effective way to compare computational properties across brain areas (Bakhurin et al. 2017). This work represents among the first efforts to directly compare information processing in frontal cortical and striatal areas during an action selection task. Taken together, the results suggest that although an animal's choice of action is encoded in both OFC and DMS dynamics, this neural representation is more refined in DMS ensembles.

Materials and Methods:

Animals. All animal procedures were approved by the University of California, Los Angeles Chancellor's Animal Research Committee. Experiments involved male wild type (C57BL/6J) mice, 8 – 12 weeks old at the time of the first surgery (stock no. 000664, Jackson Laboratory).

Surgery. Surgical procedures were performed under aseptic conditions and isoflurane anesthesia on a stereotaxic apparatus (Model 1900, Kopf Instruments). In the first surgery, a pair of stainless steel head fixation bars was attached to each side of the skull with dental cement (Metabond, Parkell). In the second surgery, performed on the day prior to electrophysiological recording, we drilled a small rectangular craniotomy over OFC, DMS, and cerebellum for the electrical reference wire.

Two-Alternative Choice Task. After a 7 day recovery period from the first surgery, mice (n = 6) were food restricted to maintain their weight at around 90% of their baseline level, and given water ad libitum. Behavior was monitored and controlled online with custom LabVIEW programs (National Instruments). All behavior-related signals were simultaneously recorded by another data acquisition system (C3100, Intan Technologies) for offline data analysis (sampling rate: 25 kHz). During training (one session per day), mice rested their hindlimbs and right forelimb on an acrylic body tube (44 mm inner diameter, part number 8585K26, McMaster-Carr). First, over a 3 – 4 day period, animals were habituated to the head fixation apparatus and to consume unconditional rewards (7.7 µl, 10% sweetened condensed milk). Second, they were given access to a joystick (part number 679-2501-ND, Digikey), whose centering spring was modified to reduce the operating force (~ 0.1 N). The joystick lever was made of a 2 mm diameter stick (part number 23-400-112, Fisher Scientific) covered with surgical tape, and placed below the left forelimb. The lever was constrained to move forward and backward. Animals were initially rewarded for making arbitrary joystick movements (1 – 3 days). In the next session, animals had first to release or hold the joystick in the central position for at least 0.1 s (defined as ±0.05 arbitrary units in Figure 1), then push or pull the lever past the threshold position to obtain reward (threshold for forward: 0.25; backward: -0.25). The session was ended after obtaining over 200 rewards from either forward or backward movements. Third, on subsequent sessions we trained mice on the two-alternative choice task with auditory cues. Animals initiated each trial by releasing or holding the joystick in the central position for a pre-defined time (release period). Moving the joystick out of the central position during the release period resulted in resetting the timer. The release period was initially set to 0.1 s, but increased by 0.01 s after each rewarded trial. Conversely, if animals did not receive any reward for 40 s the release period was reduced by 0.01 s. The final release period was 1 s. The release period was followed by the cue period, in which a pure auditory tone was presented for 100 ms (cue 1: 3 kHz; cue 2: 16 kHz, pseudorandom order, the same cue was

never presented more than twice in a row). In the ensuing movement period, moving the joystick forward in response to cue 1 beyond a pre-set threshold position (0.16), or backward in response to cue 2 (-0.14), resulted in immediate reward delivery. Animals were not prohibited from initiating their joystick response during the cue period, but correct choices were only rewarded after cue offset. Failure to respond or make a correct choice in the movement period transitioned the task to the inter-trial interval (ITI). The movement period was initially set to 10 s and reduced to a final value of 3 s in steps of 0.07 s after each rewarded trial. The ITI was initially set to 1.5 s and increased to 8 s in steps of 0.065 s after each rewarded trial. On the final stage of training (13 – 44 days), the duration of the release and movement period was fixed to its final value, and the ITI was drawn from a uniform distribution between 5.5 – 10.5 s. Electrophysiological recordings were carried out at the end of the final stage. Including all stages of training, animals belonging to the two-alternative choice task group were trained for 18 – 49 days including the recording session.

One-Alternative Choice Task. A separate group of mice (n = 6) was trained on a task identical to the two-alternative choice task, except the rewarded response to both auditory cues was movement of the joystick lever in the same direction (forward). The final stage of training had a duration of 3 – 13 days. Including all stages of training, animals belonging to the one-alternative choice task group were trained for 9 – 17 days including the recording session.

Offline Behavioral Data Analysis. The joystick voltage signal was downsampled to 1 kHz for offline analysis. The baseline position was calculated from the mean joystick position in the 1 s release period. There were three types of behavioral responses to each cue: hit (moving the joystick past the threshold position in the correct direction, resulting in reward), non-responsive (joystick movement failing to cross the threshold position), and error (moving the joystick in the incorrect direction). The hit and error rates were defined as:

$$Hit\ rate = N_{hit}/(N_{hit} + N_{NR} + N_{error})$$

$$Error\ rate = N_{error}/(N_{hit} + N_{NR} + N_{error})$$

where N_{hit}, N_{NR}, and N_{error} represent the number of hit, non-responsive, and error trials, respectively.

Electrophysiology. Silicon microprobes containing a total of 256 electrodes (Yang et al. 2020) were used to simultaneously record from ventral and lateral orbital subregions of OFC, as well as DMS (128 electrodes per area; probe model 128K in OFC and 128DN in DMS). To confirm the probe location, the silicon shafts were coated with a fluorescent dye (DiD, Thermo Fisher Scientific) before insertion. The target insertion coordinates relative to bregma were: 2.6 mm anterior, 0.8 – 1.4 mm lateral), 2.7 mm ventral in OFC, and 1.0 mm anterior, 1.1 – 1.55 mm lateral, 3.5 mm ventral in DMS. Recordings were in the right hemisphere, contralateral to the joystick manipulation arm. Data were acquired at a sampling rate of 25 kHz, bandpass filtered from 300 – 7000 Hz, and spike sorted with Kilosort (Pachitariu et al. 2016). Analysis of neural activity only included putative single-unit clusters, and did not further classify units into different subpopulations such as medium spiny projection neurons or fast spiking interneurons. There was only one recording session per animal.

Single-Neuron Response Selectivity Analysis. Analysis of single-neuron responses only used hit trials, separated into two trial types corresponding to cue 1 and 2. Spike trains were aligned to either the time of cue onset or movement threshold, binned in 1 ms steps, and smoothed with a Gaussian kernel (15 ms standard deviation). Each neuron's selectivity for one of the two trial types was assessed with a receiver operating characteristic (ROC) analysis (Feierstein et al. 2006), which was applied on activity 0 – 100 ms of cue onset (for cue-aligned data), or 0 – 100 ms before movement threshold (for movement-aligned data). The selectivity index, with a range of ±1, was calculated from the area under the ROC curve (auROC) as follows:

$$Selectivity\ index = 2(auROC - 0.5)$$

where positive and negative values corresponded to higher responsiveness for cue 2 and cue 1 trials, respectively. The significance of neural selectivity was assessed with bootstrapping 100 times, and using a threshold probability value of 0.05.

Population Decoding Analysis. All decoding analysis was performed on concurrently recorded ensembles. The goal of the decoding analysis was to train a classifier that could distinguish between cue 1 and cue 2 hit trials at different times relative to cue onset or movement threshold. Decoding was based on a linear support vector machine (SVM) learning algorithm (Chang and Lin 2011), with 80% of trials used for training and the remaining 20% for testing the accuracy (half of the test trials were from cue 1, the other half from cue 2). Within each animal, the number of cue 1 and cue 2 trials used in the decoder was matched (range: 98 – 344 for decoding on hit trials, and 30 – 211 for decoding on non-responsive trials). To enable a pairwise comparison between simultaneously recorded data in OFC and DMS, within each animal we also matched the number of OFC and DMS neurons (range: 15 – 83). This process was repeated 1000 times with different random sampling of trials and neurons. The mean decoding accuracy per animal was calculated from the average performance over the 1000 iterations. We separately trained a decoder on data in which the test trial labels were randomly shuffled. Decoding analysis was performed on 100 ms segments of unsmoothed spiking data, and repeated in 10 ms time steps. The decoding latency was defined as the time step in which the mean accuracy crossed the 95% confidence interval (CI) calculated from 1000 iterations of the shuffled data. The results note that in some animals, the accuracy in OFC or DMS never crossed the 95% CI, and thus could not be included in the analysis of decoding latency. To examine the effect of population size on decoding accuracy, we separately trained and tested an SVM decoder with a population of 5, 10, 15, 20, 30, or 40 randomly chosen neurons, and calculated the average accuracy across 1000 random drawings.

Experimental Design and Statistical Analyses. Statistical analysis was carried out with standard functions in MATLAB (MathWorks) and Prism (GraphPad Software). Data collection and analysis were not performed with blinding to the conditions of the experiments. No statistical methods were used to predetermine sample size, but our sample sizes are similar to those reported in previous publications. The sample size, type of test and probability values are indicated in the figure legends. Data distribution was assumed to be normal, but this was not formally tested. T-tests were always two sided. In all figures, the convention is $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, and not significant (ns) $p > 0.05$.

Data and Code Accessibility. Data sets and code used in this study are available from the corresponding author upon request.

Results:

Action selection task for head-fixed mice

We developed an auditory guided two-alternative choice task involving forelimb manipulation of a joystick. Head-fixed mice were trained to move a joystick lever forward in response to a low frequency tone, and backward in response to a high frequency tone (n = 6 mice, **Figure 1A,B,C**). A sweetened milk reward was delivered immediately after making the correct choice, with both cues associated with the same volume of liquid. After training, mice made a statistically similar proportion of correct (hit trial) and incorrect (error trial) choices after each type of cue (**Figure 1D,E**). The hit rate on cue 1 trials was uncorrelated with the hit rate on cue 2 trials (Pearson correlation, $r = 0.34$, $p = 0.51$). Additionally, there was no significant difference in lever press time relative to cue onset (**Figure 1F,G**). This appears to show that there was no bias in the animal's perceived value of performing forward or backward joystick movements. Nevertheless, a potential concern with this task is that if OFC or DMS activity discriminates between different auditory tone

frequencies (Guo et al. 2018), this would interfere with the interpretation of action selectivity. To account for potential auditory coding differences in the absence of action selection, we introduced a one-alternative choice task in a separate group of mice. Here, both cues were associated with joystick movement in the same direction (forward), and equal reward volume (n = 6 mice, **Figure 1H,I,J**). Again, the response rate was similar between the two cue types (**Figure 1K,L**), and there was no significant difference in lever press time (**Figure 1M,N**). We therefore reasoned that any differences found between OFC and DMS neural activity in the two-alternative, but not the one-alternative choice task, would provide strong evidence for differential encoding of action selection in these brain areas.

Higher proportion of cue-selective neurons in DMS

In order to examine activity in OFC and DMS during task performance in well-trained mice, we used silicon microprobes to record from dozens of neurons in these areas in parallel (**Figure 2**). We first assessed the cue-aligned response of individual neurons recorded from the two-alternative choice task group (**Figure 3A**). We used ROC analysis to calculate their selectivity during the cue period. We found neurons in both areas that showed selectivity for one of the trial types (**Figure 3B**), but the proportion of cue-selective neurons was significantly higher in DMS (**Figure 3C**). A sizable fraction of neurons recorded from the one-alternative choice task group were also selective for one of the auditory tones (**Figure 3D,E**), and again DMS contained a greater proportion of cue-selective neurons than OFC (**Figure 3F**). The proportion of total selective cells pooled from all animals was also found to be similar between the two task groups (OFC: 124/415 of cells in the two-alternative versus 109/380 in the one-alternative choice task, chi square test, p = 0.71; DMS: 256/379 in the two-alternative versus 325/457 in the one-alternative choice task, chi square test, p = 0.26). Therefore, the regional differences in cue period selectivity found in the two-alternative choice task group appear to be at least partly driven by stronger auditory tone discrimination in DMS, in the absence of overt action selection processes.

Next, to examine neural activity during the movement period, which is less likely to be biased by auditory tone discrimination, we performed ROC analysis on data aligned to the time of movement threshold (**Figure 4A**). For recordings made from the two-alternative choice task group, a subset of neurons showed selectivity for one of the trial types (**Figure 4B**), and the proportion of selective neurons was statistically similar in the two brain areas (**Figure 4C**). Data from the one-alternative choice task group also did not show a significant difference in the fraction of selective neurons between OFC and DMS during the movement period (**Figure 4D,E,F**). On the other hand, the proportion of total selective cells pooled from all animals was significantly higher for the two-alternative choice task group (OFC: 146/415 of cells in the two-alternative versus 38/380 in the one-alternative choice task, chi square test, $p < 0.0001$; DMS: 164/379 in the two-alternative versus 81/457 in the one-alternative choice task, chi square test, $p < 0.0001$). Thus, while recordings from both types of tasks contained neurons with selective responses in the movement period, the number of these selective cells was enriched in the task with two instead of one possible actions. These results suggest that neural selectivity in the movement period was largely due to encoding of the animal's choice of action, rather than persistent effects of auditory tone discrimination. However, significant regional differences in the selectivity index were only observed for cue- but not movement-aligned data, which raises potential ambiguities in interpreting the findings in the context of action selection.

More accurate decoding of selected action from DMS population dynamics

Similar to previous electrophysiological studies using two-alternative choice tasks, our data showed substantial heterogeneity in the response properties of individual neurons (Feierstein et al. 2006; Guo et al. 2019). Thus, determining the proportion of selective neurons may not be a reliable way to compare information processing across brain areas. Since our recording approach provided simultaneous access to multiple neurons from OFC and DMS, we were able to explore this issue by applying decoding methods to compare trial-by-trial population dynamics. SVM

decoders were trained to distinguish between cue 1 and cue 2 hit trials based on the activity of simultaneously measured ensembles.

Decoder performance in each area was quantified in terms of two parameters—accuracy in assigning test trials to the correct cue type, and latency to reach a statistically significant accuracy level (95% CI). In the two-alternative choice task group, decoding accuracy in both areas increased rapidly after cue onset (**Figure 5A**). However, on average, DMS population activity had both a significantly higher accuracy and lower latency of decoding than OFC (**Figure 5B,C**). In contrast, in the one-alternative task group, neither the accuracy nor latency was statistically distinguishable between the two brain areas (**Figure 5D,E,F**). These results were qualitatively different from the single-neuron selectivity index analysis in the same dataset (**Figure 3F**).

To determine how decoding accuracy depended on the population size, we compared the accuracy of SVM decoders that were tested on different numbers of neurons. The accuracy improved with greater population size, but only the two-alternative choice task group showed significant differences between OFC and DMS (**Figure 6A,B**), consistent with the finding that DMS contains a better population code for action selection.

To confirm that the higher accuracy in DMS in the two-alternative choice task reflects more accurate encoding of action, rather than auditory tone selectivity, we also applied the SVM decoder on a different subset of trials (non-responsive), in which animals failed to move the lever to the threshold position (**Figure 7A**). There was no significant bias in decoding accuracy between OFC and DMS on non-responsive trials, although the difference was close to being significant ($p = 0.056$, **Figure 7B**). On average a small amount of lever movement was observed even on non-responsive trials, thus accurate decoding on these trials is likely to reflect action selectivity to a smaller extent than on hit trials.

In the two-alternative choice task group, decoding accuracy, but not latency, remained significantly better in DMS for movement-aligned data on hit trials (**Figure 8A,B,C**), again in marked contrast to the results of analysis of single-neuron selectivity (**Figure 4C**). However,

decoding accuracy in the movement period was not statistically distinguishable between the two areas in data from the one-alternative choice task group (**Figure 8D,E,F**).

Discussion:

Taken together, the decoding results suggest that DMS neural populations are significantly better at representing an animal's selected action than an equivalent number of concurrently recorded cells in OFC. These regional differences were apparent both for the cue and movement period. Importantly, significant differences in decoding accuracy were only observed in data from the two-alternative, but not the one-alternative choice task group, suggesting that these effects reflect action rather than auditory tone selectivity. Moreover, more accurate decoding in DMS was only observed on hit trials, but not on non-responsive trials (**Figure 7**), providing further evidence that these results are unlikely to reflect regional variations in auditory tone selectivity. Finally, the results obtained with population decoding methods were qualitatively different from those obtained with analysis of single-neuron selectivity, which yielded less consistent findings between the cue and movement period.

This work used joystick-based action selection tasks specifically developed for experiments in head-fixed mice (Nonomura et al. 2018). In combination with silicon microprobe recording tools targeting OFC and DMS in parallel, this approach enabled a regional comparison of neural dynamics within the same animal and behavioral session. Selectivity of dynamics to specific trial types was assessed both at the level of individual cells using ROC analysis, and at the level of neural populations using SVM decoding methods. There were some important qualitative differences between these modes of analysis. Overall, individual neuron ROC analysis gave internally inconsistent results, showing regional differences in the proportion of selective cells for cue- but not movement-aligned data (**Figure 3C** and **4C**). Furthermore, since the trends were similar for data from the two-alternative and one-alternative choice tasks, it was unclear whether selectivity reflected action or auditory tone discrimination. By comparison, population

decoding methods were internally consistent, showing regional differences in decoding accuracy for both cue- and movement-aligned data (**Figure 5C** and **8C**). The most parsimonious explanation for the discrepancy between the two methods used to analyze neural selectivity, is that population decoding methods revealed aspects of computation in OFC or DMS which were less apparent or accessible from ROC analysis of individual neuron firing.

Our findings appear to disagree with a study in non-human primates, which concluded that lateral prefrontal cortex contains a better representation of action selection than dorsal striatum (Seo et al. 2012). This discrepancy may arise from differences in which two functionally distinct subregions of frontal cortex and striatum were compared (Ito and Doya 2015; Izquierdo 2017), the methods used to assess neural selectivity, or the details of the behavioral task. On the other hand, other studies have challenged the view that OFC strongly influences an animal's current choice of action (Gardner et al. 2017), but instead suggest that this area participates in learning or updating choices on future trials (Constantinople et al. 2019; Miller et al. 2018). More generally, numerous lines of evidence suggest that OFC represents information that is not exclusively related to, and may even be distinct from, action selection, including spatial goals (Feierstein et al. 2006), value (Padoa-Schioppa and Assad 2006), action value (Simon et al. 2015), and complex cognitive maps representing the relationship between multiple facets of an animal's environmental and behavioral state (Wilson et al. 2014). Furthermore, since the approaches in this work are correlative, we cannot conclude that DMS serves a more critical behavioral role in action selection than OFC. Rather, populations of DMS neurons appear to more accurately encode information about an animal's choice of action compared to OFC, at least under the conditions used here.

Some studies have suggested that behavior shifts from DMS to DLS-dependent during the transition from goal-directed to habitual responses (Gremel and Costa 2013; Thorn et al. 2010; Yin et al. 2009). This could potentially influence the interpretation of our findings; however, based on the robust encoding of action selection in DMS we speculate there was insufficient training to

362 make responding habitual. Alternatively, the types of tasks employed here may not support habit
363 formation.

364 A potential caveat of this work is that OFC and DMS may use distinct strategies to encode
365 information in the tasks with one and two alternative choices. In particular, the one-alternative
366 choice task requires animals to respond to a cue without needing to select between different
367 possible actions. Additionally, despite more protracted training, behavioral performance was
368 worse in animals belonging to the two-alternative choice task group. For these reasons, there
369 may be limitations in comparing results across the two task conditions. Nevertheless, the results
370 suggest differences in how populations of OFC and DMS neurons represent information in the
371 two-alternative choice task. Since that task requires correct action selection, the observed
372 encoding differences are unlikely to be exclusively due to differences in auditory cue encoding.

373 There are a number of potential mechanisms for the observed difference in OFC and DMS
374 population dynamics. One possibility is that other sources of input may convey complementary
375 signals to DMS (Ponvert and Jaramillo 2019). Another study found that medial prefrontal cortex
376 (mPFC), which also projects to DMS, contains a greater proportion of action-selective units than
377 OFC (Simon et al. 2015), suggesting that mPFC exhibits more accurate encoding of action
378 selection than OFC. There is also some indirect support from our observation that decoding was
379 not only more accurate, but reached a significant level of performance on average 87 ms earlier
380 in DMS than OFC (**Figure 5C**). This temporal lag even raises the possibility that action selection
381 signals propagate from DMS to OFC through basal ganglia-cortical feedback loops (Oldenburg
382 and Sabatini 2015). Since corticostriatal projections are a major source of excitatory drive in the
383 striatum (Emmons et al. 2017; Shepherd 2013), another possibility is that striatal microcircuits
384 amplify or refine incoming cortical signals to improve information processing (Bakhurin et al.
385 2017). This could occur via gain modulation effects at corticostriatal synapses (Lee et al. 2019),
386 or contributions from local interneurons which may modify striatal output (Gittis et al. 2010).

Grants:

SCM was supported by NIH grants NS100050, NS096994, DA042739, DA005010, and NSF NeuroNex Award 1707408.

Disclosures

The authors declare the absence of any conflict of interest.

References:

Ahmari SE, Spellman T, Douglass NL, Kheirbek MA, Simpson HB, Deisseroth K, Gordon JA, and Hen R. Repeated cortico-striatal stimulation generates persistent OCD-like behavior. *Science* 340: 1234-1239, 2013.

Alexander GE, DeLong MR, and Strick PL. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu Rev Neurosci* 9: 357-381, 1986.

Bakhurin KI, Goudar V, Shobe JL, Claar LD, Buonomano DV, and Masmanidis SC. Differential Encoding of Time by Prefrontal and Striatal Network Dynamics. *J Neurosci* 37: 854-870, 2017.

Balleine BW, Delgado MR, and Hikosaka O. The role of the dorsal striatum in reward and decision-making. *J Neurosci* 27: 8161-8165, 2007.

Bradfield LA, Dezfouli A, van Holstein M, Chieng B, and Balleine BW. Medial Orbitofrontal Cortex Mediates Outcome Retrieval in Partially Observable Task Situations. *Neuron* 88: 1268-1280, 2015.

Burguiere E, Monteiro P, Feng G, and Graybiel AM. Optogenetic stimulation of lateral orbitofronto-striatal pathway suppresses compulsive behaviors. *Science* 340: 1243-1246, 2013.

Chang CC, and Lin CJ. LIBSVM: A Library for Support Vector Machines. *Acm T Intel Syst Tec* 2: 2011.

Constantinople CM, Piet AT, Bibawi P, Akrami A, Kopec C, and Brody CD. Lateral orbitofrontal cortex promotes trial-by-trial learning of risky, but not spatial, biases. *eLife* 8: 2019.

Corbit VL, Manning EE, Gittis AH, and Ahmari SE. Strengthened Inputs from Secondary Motor Cortex to Striatum in a Mouse Model of Compulsive Behavior. *J Neurosci* 39: 2965-2975, 2019.

Emmons EB, De Corte BJ, Kim Y, Parker KL, Matell MS, and Narayanan NS. Rodent Medial Frontal Control of Temporal Processing in the Dorsomedial Striatum. *J Neurosci* 37: 8718-8733, 2017.

Feierstein CE, Quirk MC, Uchida N, Sosulski DL, and Mainen ZF. Representation of spatial goals in rat orbitofrontal cortex. *Neuron* 51: 495-507, 2006.

Frank MJ. Computational models of motivated action selection in corticostriatal circuits. *Curr Opin Neurobiol* 21: 381-386, 2011.

Friedman A, Homma D, Gibb LG, Amemori K, Rubin SJ, Hood AS, Riad MH, and Graybiel AM. A Corticostriatal Path Targeting Striosomes Controls Decision-Making under Conflict. *Cell* 161: 1320-1333, 2015.

Gardner MPH, Conroy JS, Shaham MH, Styer CV, and Schoenbaum G. Lateral Orbitofrontal Inactivation Dissociates Devaluation-Sensitive Behavior and Economic Choice. *Neuron* 96: 1192-1203 e1194, 2017.

Gittis AH, Nelson AB, Thwin MT, Palop JJ, and Kreitzer AC. Distinct roles of GABAergic interneurons in the regulation of striatal output pathways. *J Neurosci* 30: 2223-2234, 2010.
Gremel CM, and Costa RM. Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nature communications* 4: 2264, 2013.
Guo L, Walker WI, Ponvert ND, Penix PL, and Jaramillo S. Stable representation of sounds in the posterior striatum during flexible auditory decisions. *Nature communications* 9: 1534, 2018.
Guo L, Weems JT, Walker WI, Levichev A, and Jaramillo S. Choice-Selective Neurons in the Auditory Cortex and in Its Striatal Target Encode Reward Expectation. *J Neurosci* 39: 3687-3697, 2019.
Humphries MD, Stewart RD, and Gurney KN. A physiologically plausible model of action selection and oscillatory activity in the basal ganglia. *J Neurosci* 26: 12921-12942, 2006.
Hwang EJ, Link TD, Hu YY, Lu S, Wang EH, Lilascharoen V, Aronson S, O'Neil K, Lim BK, and Komiyama T. Corticostriatal Flow of Action Selection Bias. *Neuron* 104: 1126-1140 e1126, 2019.
Ito M, and Doya K. Distinct neural representation in the dorsolateral, dorsomedial, and ventral parts of the striatum during fixed- and free-choice tasks. *J Neurosci* 35: 3499-3514, 2015.
Izquierdo A. Functional Heterogeneity within Rat Orbitofrontal Cortex in Reward Learning and Decision Making. *J Neurosci* 37: 10529-10540, 2017.
Kimchi EY, and Laubach M. Dynamic encoding of action selection by the medial striatum. *J Neurosci* 29: 3148-3159, 2009.
Lee K, Bakhurin KI, Claar LD, Holley SM, Chong NC, Cepeda C, Levine MS, and Masmanidis SC. Gain Modulation by Corticostriatal and Thalamostriatal Input Signals during Reward-Conditioned Behavior. *Cell reports* 29: 2438-2449 e2434, 2019.
Miller KJ, Botvinick MM, and Brody CD. Value Representations in the Orbitofrontal Cortex Drive Learning, not Choice. *bioRxiv* 2018.
Mink JW. The basal ganglia: focused selection and inhibition of competing motor programs. *Progress in neurobiology* 50: 381-425, 1996.
Moorman DE, and Aston-Jones G. Orbitofrontal cortical neurons encode expectation-driven initiation of reward-seeking. *J Neurosci* 34: 10234-10246, 2014.
Murray EA, and Izquierdo A. Orbitofrontal cortex and amygdala contributions to affect and action in primates. *Annals of the New York Academy of Sciences* 1121: 273-296, 2007.
Nonomura S, Nishizawa K, Sakai Y, Kawaguchi Y, Kato S, Uchigashima M, Watanabe M, Yamanaka K, Enomoto K, Chiken S, Sano H, Soma S, Yoshida J, Samejima K, Ogawa M, Kobayashi K, Nambu A, Isomura Y, and Kimura M. Monitoring and Updating of Action Selection for Goal-Directed Behavior through the Striatal Direct and Indirect Pathways. *Neuron* 2018.
Oldenburg IA, and Sabatini BL. Antagonistic but Not Symmetric Regulation of Primary Motor Cortex by Basal Ganglia Direct and Indirect Pathways. *Neuron* 86: 1174-1181, 2015.
Ostlund SB, and Balleine BW. The contribution of orbitofrontal cortex to action selection. *Annals of the New York Academy of Sciences* 1121: 174-192, 2007.
Pachitariu M, Steinmetz N, Kadir S, Carandini M, and Harris KD. Kilosort: realtime spike-sorting for extracellular electrophysiology with hundreds of channels. *bioRxiv* 2016.
Padoa-Schioppa C, and Assad JA. Neurons in the orbitofrontal cortex encode economic value. *Nature* 441: 223-226, 2006.
Pennartz CM, Berke JD, Graybiel AM, Ito R, Lansink CS, van der Meer M, Redish AD, Smith KS, and Voorn P. Corticostriatal Interactions during Learning, Memory Processing, and Decision Making. *J Neurosci* 29: 12831-12838, 2009.
Ponvert ND, and Jaramillo S. Auditory Thalamostriatal and Corticostriatal Pathways Convey Complementary Information about Sound Features. *J Neurosci* 39: 271-280, 2019.
Reig R, and Silberberg G. Multisensory integration in the mouse striatum. *Neuron* 83: 1200-1212, 2014.

Rothwell PE, Hayton SJ, Sun GL, Fuccillo MV, Lim BK, and Malenka RC. Input- and Output-Specific Regulation of Serial Order Performance by Corticostriatal Circuits. *Neuron* 88: 345-356, 2015.
Samejima K, Ueda Y, Doya K, and Kimura M. Representation of action-specific reward values in the striatum. *Science* 310: 1337-1340, 2005.
Schultz W, and Romo R. Role of primate basal ganglia and frontal cortex in the internal generation of movements. I. Preparatory activity in the anterior striatum. *Exp Brain Res* 91: 363-384, 1992.
Seo M, Lee E, and Averbeck BB. Action selection and action value in frontal-striatal circuits. *Neuron* 74: 947-960, 2012.
Sharpe MJ, Stalnaker T, Schuck NW, Killcross S, Schoenbaum G, and Niv Y. An Integrated Model of Action Selection: Distinct Modes of Cortical Control of Striatal Decision Making. *Annual review of psychology* 70: 53-76, 2019.
Shepherd GM. Corticostriatal connectivity and its role in disease. *Nat Rev Neurosci* 14: 278-291, 2013.
Shin JH, Kim D, and Jung MW. Differential coding of reward and movement information in the dorsomedial striatal direct and indirect pathways. *Nature communications* 9: 404, 2018.
Simon NW, Wood J, and Moghaddam B. Action-outcome relationships are represented differently by medial prefrontal and orbitofrontal cortex neurons during action execution. *J Neurophysiol* 114: 3374-3385, 2015.
Sippy T, Lapray D, Crochet S, and Petersen CC. Cell-Type-Specific Sensorimotor Processing in Striatal Projection Neurons during Goal-Directed Behavior. *Neuron* 88: 298-305, 2015.
Stalnaker TA, Calhoon GG, Ogawa M, Roesch MR, and Schoenbaum G. Reward prediction error signaling in posterior dorsomedial striatum is action specific. *J Neurosci* 32: 10296-10305, 2012.
Tai LH, Lee AM, Benavidez N, Bonci A, and Wilbrecht L. Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nat Neurosci* 15: 1281-1289, 2012.
Thorn CA, Atallah H, Howe M, and Graybiel AM. Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron* 66: 781-795, 2010.
Wang AY, Miura K, and Uchida N. The dorsomedial striatum encodes net expected return, critical for energizing performance vigor. *Nat Neurosci* 16: 639-647, 2013.
Wilson RC, Takahashi YK, Schoenbaum G, and Niv Y. Orbitofrontal cortex as a cognitive map of task space. *Neuron* 81: 267-279, 2014.
Yang L, Lee K, Villagracia J, and Masmanidis SC. Open source silicon microprobes for high throughput neural recording. *J Neural Eng* 17: 016036, 2020.
Yin HH, Knowlton BJ, and Balleine BW. Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *Eur J Neurosci* 22: 505-512, 2005.
Yin HH, Mulcare SP, Hilario MR, Clouse E, Holloway T, Davis MI, Hansson AC, Lovinger DM, and Costa RM. Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nat Neurosci* 12: 333-341, 2009.
Znamenskiy P, and Zador AM. Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. *Nature* 497: 482-485, 2013.

Figure Legends

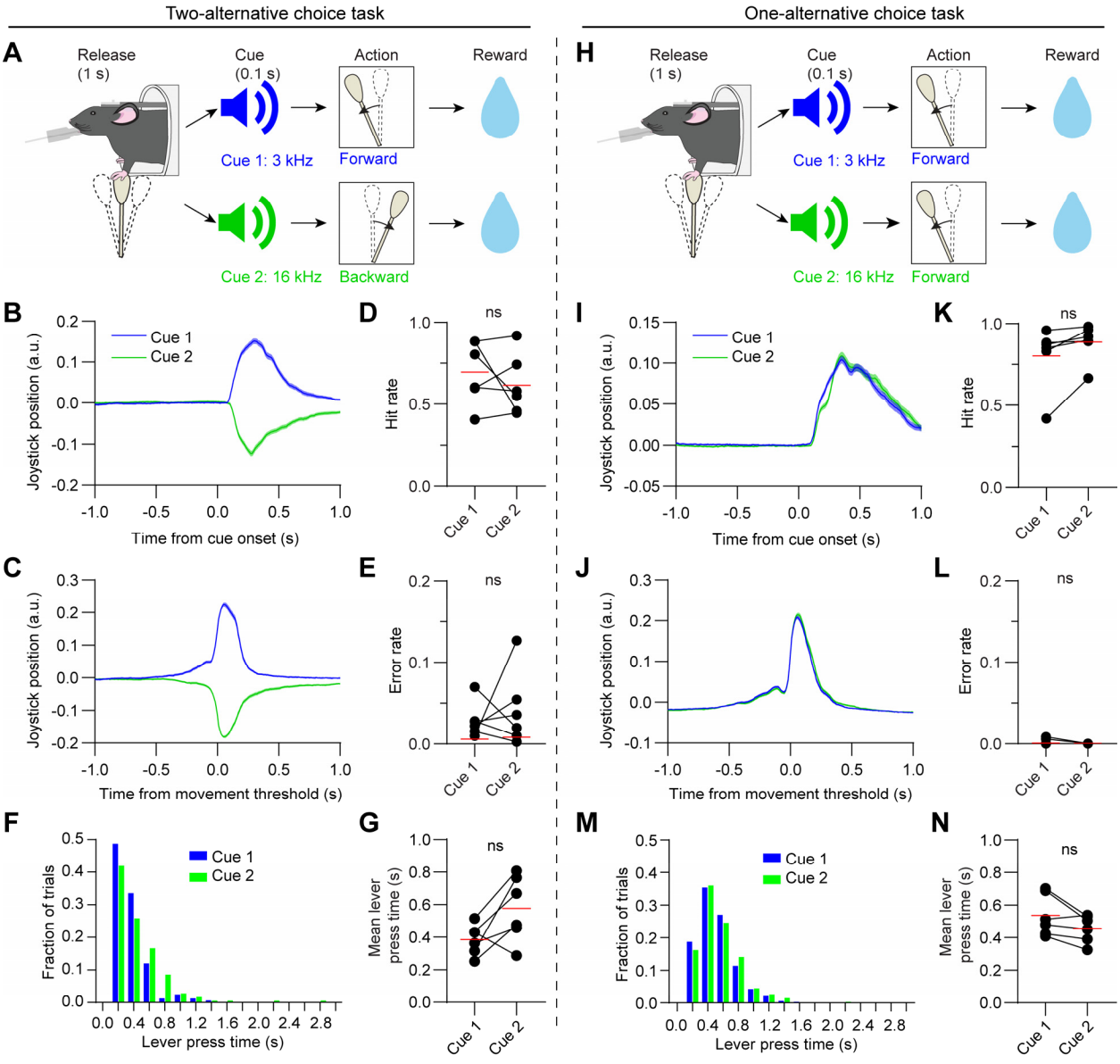


Figure 1. A two-alternative and one-alternative choice task for head-fixed mice.

A) Illustration of the two-alternative choice task paradigm. Head-fixed mice had to first release or hold in the joystick in the central position for 1 s for the trial to begin. They received reward for moving a joystick forward in response to a low frequency tone (cue 1, 3 kHz, 100 ms), or backward in response to a high frequency tone (cue 2, 16 kHz, 100 ms).

B) Joystick position signal in response to the two cues on hit trials, aligned to cue onset. Data represent the mean of one well-trained animal from the two-alternative choice task group on the recording day. Shaded areas represent standard error of the mean (SEM).

C) Same as B but aligned to the time of joystick movement threshold.

D) There was no significant difference in hit rate across all animals tested on the two-alternative choice task (n = 6 mice, paired t-test, $t_5 = 0.92$, $p = 0.4$). Red lines represent the mean values.

E) There was no significant difference in error rate across all animals tested on the two-alternative choice task (n = 6 mice, paired t-test, $t_5 = 0.54$, $p = 0.61$).

F) Distribution of the lever press times on cue 1 (blue) and cue 2 (green) hit trials, for the same animal as B and C.

G) There was no significant difference in mean lever press time across all animals tested on the two-alternative choice task (n = 6 mice, paired t-test, $t_5 = 2.34$, $p = 0.066$).

H) Illustration of the one-alternative choice task paradigm, used to characterize cue-selective neural responses in the absence of action selectivity. This task is similar to the two-alternative choice task, except both cues are associated with forward joystick movement.

I) Joystick position signal in response to the two cues on hit trials, aligned to cue onset. Data represent the mean of one well-trained animal from the one-alternative choice task group on the recording day.

J) Same as I but aligned to the time of joystick movement threshold.

K) There was no significant difference in hit rate across all animals tested on the one-alternative choice task (n = 6 mice, paired t-test, $t_5 = 2.29$, $p = 0.07$).

L) There was no significant difference in error rate across all animals tested on the one-alternative choice task (n = 6 mice, paired t-test, $t_5 = 1.55$, $p = 0.18$).

M) Distribution of the lever press times on cue 1 (blue) and cue 2 (green) hit trials, for the same animal as I and J.

N) There was no significant difference in mean lever press time across all animals tested on the one-alternative choice task (n = 6 mice, paired t-test, $t_5 = 2.32$, $p = 0.068$).

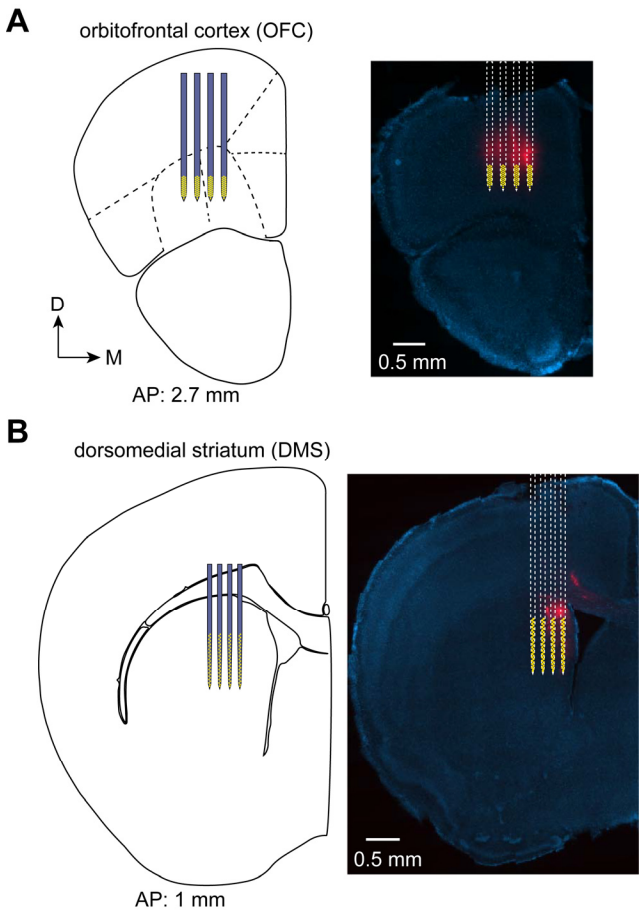
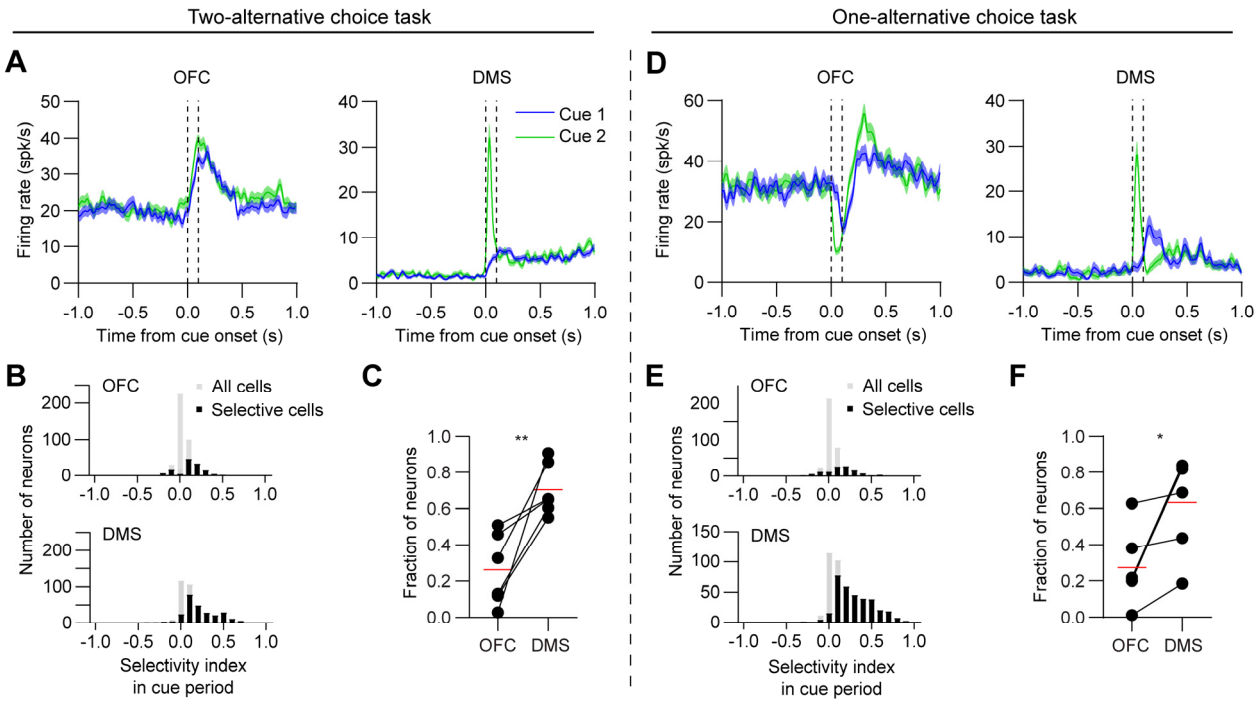


Figure 2. Simultaneous electrophysiological recordings in OFC and DMS.

A) Left, a 128 electrode silicon microbrobe with four shafts targeted the ventral and lateral orbital subregions of OFC. Yellow markings indicate the location of the electrode recording sites. The anterior-posterior (AP) position from bregma is indicated at the bottom. Right, fluorescence image of a brain slice showing the track made by the DiD dye coated on the probe (red), with a superimposed drawing of the silicon microprobe. Slices were stained with DAPI (blue).

B) Same as A, but for a separate 128 electrode silicon microprobe targeting DMS in the same animal.



564 **Figure 3**. Higher proportion of selective neurons in DMS in the cue period.

565 **A)** Mean firing rate versus time of one OFC (left) and DMS (right) neuron, with data aligned to
566 cue onset. Data in panels A-C are from animals in the two-alternative choice task group. Blue
567 and green lines represent cue 1 (low frequency) and cue 2 (high frequency) hit trials,
568 respectively. The dashed vertical lines demarcate the time interval used to calculate the
569 selectivity index in the cue period. Shaded areas represent SEM.

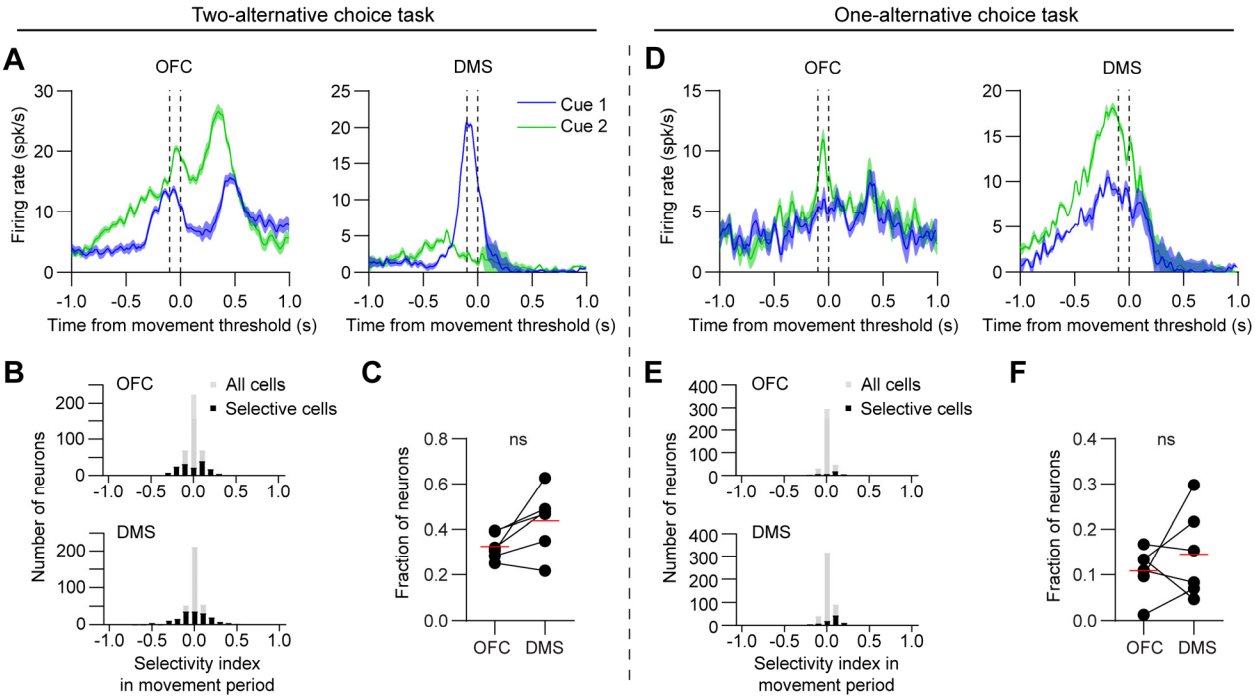
570 **B)** Selectivity index distribution of OFC (top) and DMS (bottom) neurons in the cue period.
571 Positive (negative) values reflect stronger responses to cue 2 (cue 1). Grey bars represent all
572 cells, and black bars represent significantly selective cells. Neurons are pooled across all 6
573 animals in the two-alternative choice task group (n = 124 out of 415 cells in OFC, and 256 out
574 of 379 cells in DMS were selective).

575 **C)** The fraction of neurons per animal that were selective in the cue period was significantly
576 higher in DMS compared to OFC (n = 6 mice, paired t-test, $t_5 = 4.1$, $p = 0.0096$). Red lines
577 represent the mean values.

D) Mean firing rate versus time of one OFC (left) and DMS (right) neuron, with data aligned to cue onset. Data in panels D-F are from animals in the one-alternative choice task group.

E) Selectivity index distribution of OFC (top) and DMS (bottom) neurons in the cue period. Neurons are pooled across all 6 animals in the one-alternative choice task group (n = 109 out of 380 cells in OFC, and 325 out of 457 cells in DMS were selective).

F) The fraction of neurons per animal that were selective in the cue period was significantly higher in DMS compared to OFC (n = 6 mice, paired t-test, $t_5 = 3$, $p = 0.03$).



585 **Figure 4**. Similar proportion of selective neurons between OFC and DMS in the movement period.

586 **A)** Mean firing rate versus time of one OFC (left) and DMS (right) neuron, with data aligned to
587 movement threshold. Data in panels A-C are from animals in the two-alternative choice task
588 group. Blue and green lines represent cue 1 (low frequency) and cue 2 (high frequency) hit
589 trials, respectively. The dashed vertical lines demarcate the time interval used to calculate the
590 selectivity index in the movement period. Shaded areas represent SEM.

591 **B)** Selectivity index distribution of OFC (top) and DMS (bottom) neurons in the movement period.
592 Positive (negative) values reflect stronger responses to cue 2 (cue 1). Grey bars represent all
593 cells, and black bars represent significantly selective cells. Neurons are pooled across all 6
594 animals in the two-alternative choice task group (n = 146 out of 415 cells in OFC, and 164 out
595 of 379 cells in DMS were selective).

C) The fraction of neurons per animal that were selective in the movement period was not significantly different between OFC and DMS (n = 6 mice, paired t-test, $t_5 = 2.3$, $p = 0.07$). Red lines represent the mean values.

D) Mean firing rate versus time of one OFC (left) and DMS (right) neuron, with data aligned to movement threshold. Data in panels D-F are from animals in the one-alternative choice task group.

E) Selectivity index distribution of OFC (top) and DMS (bottom) neurons in the movement period. Neurons are pooled across all 6 animals in the one-alternative choice task group (n = 38 out of 380 cells in OFC, and 81 out of 457 cells in DMS were selective).

F) The fraction of neurons per animal that were selective in the movement period was not significantly different between OFC and DMS (n = 6 mice, paired t-test, $t_5 = 0.9$, $p = 0.43$).

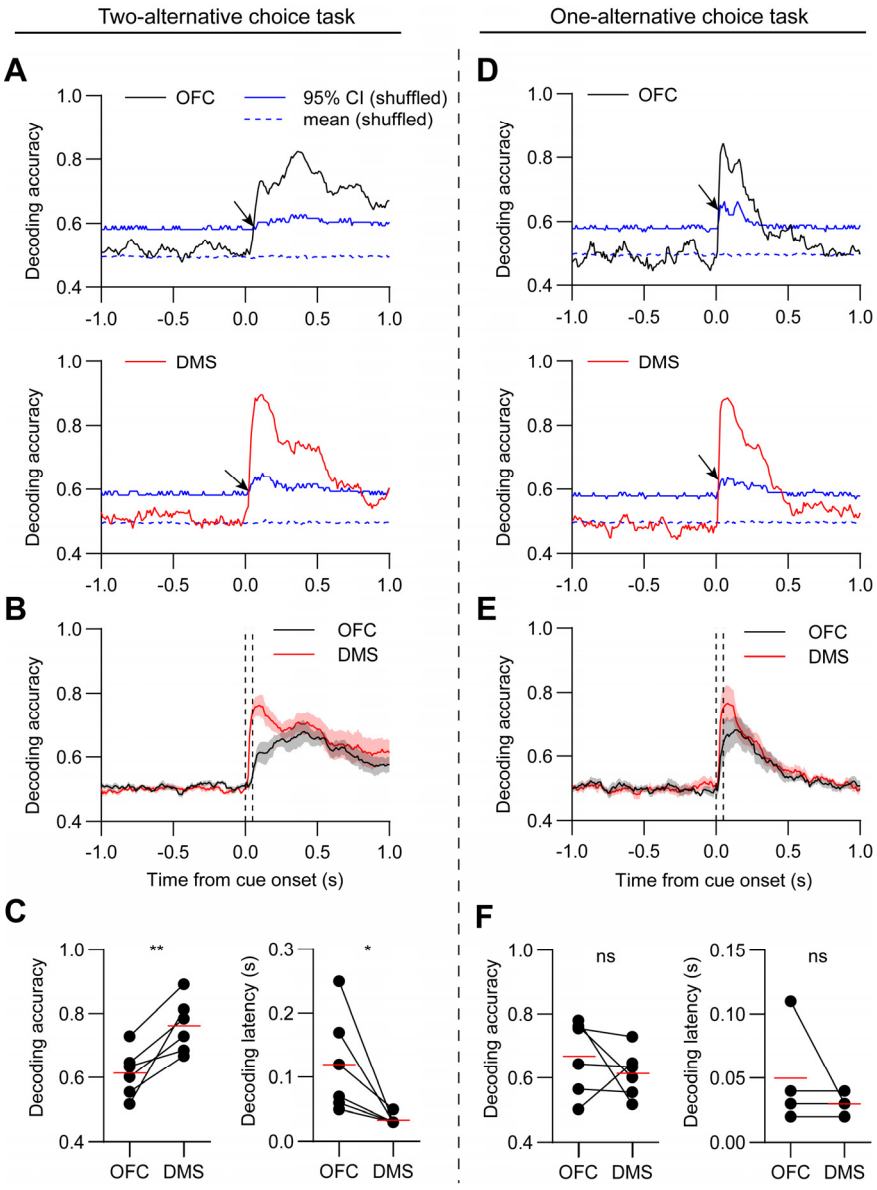


Figure 5. More accurate and rapid decoding from DMS population activity in the cue period for the two- but not one-alternative choice task.

A) Mean decoding accuracy versus time in OFC (top, black) and DMS (bottom, red) of one animal with data aligned to cue onset. Data in panels A-C are from animals in the two-alternative choice task group. The decoder was trained to distinguish cue 1 from cue 2 hit trials. The dashed and solid blue lines represent the mean and 95% CI of the decoder performance tested on shuffled data. Decoding was performed on 100 ms segments of data and repeated

in 10 ms steps. The arrows indicate the decoding latency, corresponding to the time step in which the mean decoding accuracy crossed the 95% CI.

B) Mean decoding accuracy versus time of all animals (n = 6 mice). The dashed vertical lines demarcate the 100 ms cue period used to calculate the decoding accuracy in C. Shaded areas represent SEM.

C) Left, the mean decoding accuracy in the cue period was significantly higher in DMS relative to OFC (n = 6 mice, paired t-test, $t_5 = 4.4$, $p = 0.007$). Right, the mean decoding latency was significantly smaller in DMS relative to OFC (n = 6 mice, paired t-test, $t_5 = 2.7$, $p = 0.04$). Red lines represent the mean values. The average decoding latency was 120 ± 33 ms in OFC, 33 ± 3 ms in DMS, (mean ± SEM relative to cue onset). On average the decoding latency was 87 ms lower in DMS relative to OFC (representing the difference between the values at the red lines).

D) Mean decoding accuracy versus time in OFC (top, black) and DMS (bottom, red) of one animal with data aligned to cue onset. Data in panels D-F are from animals in the one-alternative choice task group.

E) Mean decoding accuracy versus time of all animals (n = 6 mice). The dashed vertical lines demarcate the 100 ms cue period used to calculate the decoding accuracy in F.

F) Left, the mean decoding accuracy in the cue period was not significantly different between OFC and DMS (n = 6 mice, paired t-test, $t_5 = 0.9$, $p = 0.39$). Right, the mean decoding latency was not significantly different between OFC and DMS (n = 4 mice, paired t-test, $t_3 = 1$, $p = 0.39$). Note that decoders trained on data from two animals failed to cross the 95% CI level, thus those animals could not be included in the decoding latency analysis. The average decoding latency was 50 ± 20 ms in OFC, 30 ± 4 ms in DMS, (mean ± SEM relative to cue onset).

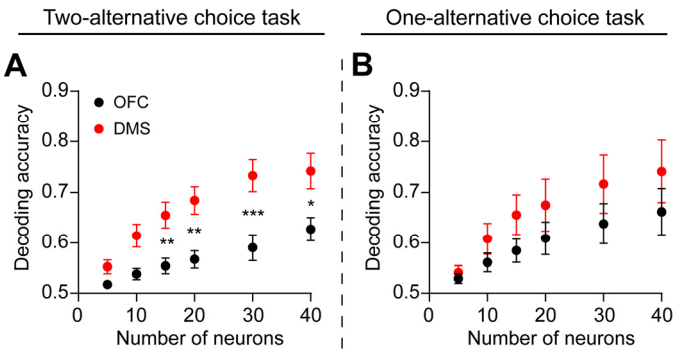


Figure 6. Decoding accuracy improves with population size.

A) Decoding accuracy in the cue period as a function of number of neurons in OFC and DMS for hit trials on the two-alternative choice task group. Due to differences in the number of simultaneously recorded neurons across animals, the number of animals is not equal for all data points (n = 6 mice for a size of 5, 10, 15, 20 neurons, n = 5 mice for a size of 30 neurons, and n = 4 mice for a size of 40 neurons). A two-way ANOVA revealed a significant effect of the number of neurons ($F_{5,54}$ = 12.1, p < 0.0001). Post hoc Sidak's test between OFC and DMS: 5 neurons, p = 0.804; 10 neurons, p = 0.068; 15 neurons, p = 0.007; 20 neurons, p = 0.001, 30 neurons, p = 0.0003; 40 neurons, p = 0.014.

B) Decoding accuracy in the cue period as a function of number of neurons in OFC and DMS for hit trials on the one-alternative choice task group. Due to differences in the number of simultaneously recorded neurons across animals, the number of animals is not equal for all data points (n = 6 mice for a size of 5, 10, 15 neurons, and n = 5 mice for a size of 20, 30, 40 neurons). A two-way ANOVA revealed a significant effect of the number of neurons ($F_{5,54}$ = 5.6, p = 0.0003). Post hoc Sidak's test between OFC and DMS: 5 neurons, p > 0.99 10 neurons, p = 0.924; 15 neurons, p = 0.637; 20 neurons, p = 0.793, 30 neurons, p = 0.638; 40 neurons, p = 0.622.

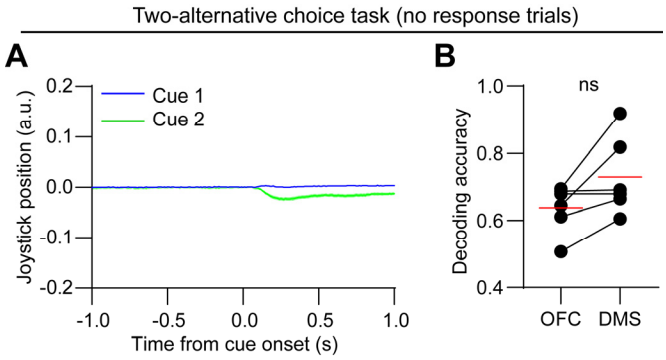
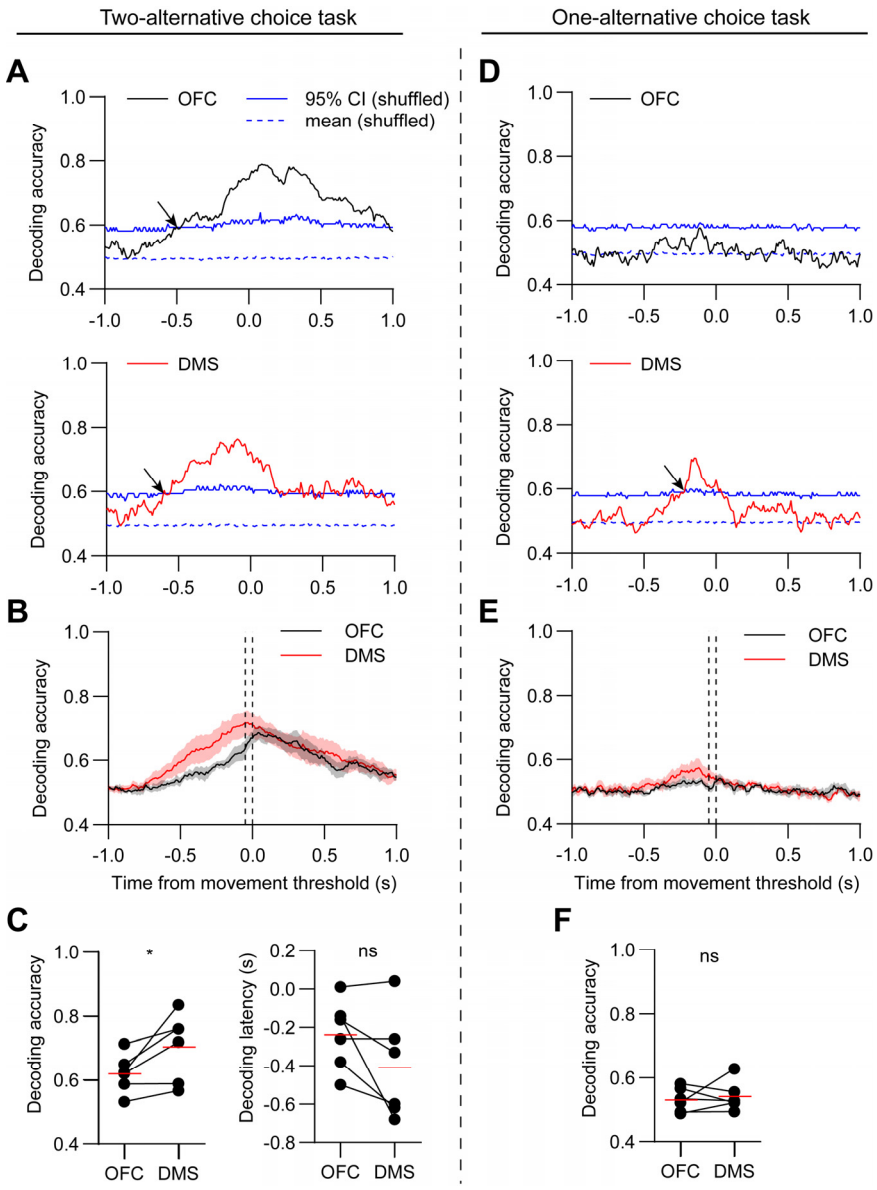


Figure 7. Decoding accuracy is not significantly different on non-responsive trials.

A) Joystick position signal in response to the two cues on non-responsive trials, aligned to cue onset. Data represent the mean of one well-trained animal from the two-alternative choice task group on the recording day (same animal as that shown in Figure 1B). Shaded areas represent standard error of the mean (SEM).

B) The mean decoding accuracy in the cue period was not significantly different between OFC and DMS on non-responsive trials (n = 6 mice, paired t-test, $t_5 = 2.5$, $p = 0.056$).



662 **Figure 8.** More accurate decoding from DMS population activity in the movement period for the
663 two- but not one-alternative choice task.

664 **A)** Mean decoding accuracy versus time in OFC (top, black) and DMS (bottom, red) of one animal
665 with data aligned to movement threshold. Data in panels A-C are from animals in the two-
666 alternative choice task group. The dashed and solid blue lines represent the mean and 95%
667 CI of the decoder performance tested on shuffled data. The arrows indicate the decoding
668 latency.

B) Mean decoding accuracy versus time of all animals (n = 6 mice). The dashed vertical lines demarcate the 100 ms movement period used to calculate the decoding accuracy in C. Shaded areas represent SEM.

C) Left, the mean decoding accuracy in the movement period was significantly higher in DMS relative to OFC (n = 6 mice, paired t-test, $t_5 = 2.8$, $p = 0.04$). Right, the mean decoding latency was not significantly different between OFC and DMS (n = 6 mice, paired t-test, $t_5 = 2$, $p = 0.1$). The average decoding latency was -238 ± 37 ms in OFC, -408 ± 113 ms in DMS, (mean ± SEM relative to movement threshold).

D) Mean decoding accuracy versus time in OFC (top, black) and DMS (bottom, red) of one animal with data aligned to movement threshold. Data in panels D-F are from animals in the one-alternative choice task group.

E) Mean decoding accuracy versus time of all animals (n = 6 mice). The dashed vertical lines demarcate the 100 ms movement period used to calculate the decoding accuracy in F.

F) The mean decoding accuracy in the movement period was not significantly different between OFC and DMS (n = 6 mice, paired t-test, $t_5 = 0.5$, $p = 0.64$). Note that for all animals, decoders trained on data from OFC or DMS failed to cross the 95% CI level, thus decoding latency analysis could not be performed.