A geometric representation unveils learning dynamics in primate neurons

Yarden Cohen, Elad Schneidman*, Rony Paz*

Dept. of Neurobiology, Weizmann Institute of Science, Rehovot, ISRAEL 76100

* These authors contributed equally

Abstract

Primates can quickly and advantageously adopt complex rule-based behaviors. We studied acquisition of a classification task while recording single neurons in the dorsal-anterior-cingulate-cortex (dACC) and the Striatum. Monkeys performed trial-by-trial classification on a rich set of multi-cue patterns, allowing de-novo learning with varying conceptual complexities every few days. To examine neural dynamics during the learning itself, we represent each rule with a spanning set of the space formed by the stimulus features. Because neural preference can be expressed by feature combinations, we can track neural dynamics in geometrical terms in this space, allowing a compact universal description of neural trajectories by observing changes in either vector-magnitude and/or angle-to-rule. We find that a large fraction of cells in both regions follow the behavior during learning. Neurons in the dACC mainly rotate towards the policy, suggesting an increase in selectivity that approximates it; whereas in the Putamen we additionally find a prominent magnitude increase, suggesting strengthening of confidence. Moreover, magnitude increases in the striatum followed rotation in the dACC, and finally, the neural policy at the end of the session predicted next-day behavior. Using this novel framework enables tracking of neural dynamics during learning and suggests differential roles of confidence and policy for the different brain regions.

Acknowledgments: We thank Yossi Shohat for animal training, welfare and experimental procedures; Dr. Yoav Kfir for scientific consult, Dr. Eilat Kahana for help with medical and surgical procedures; Dr. Edna Furman-Haran and Fanny Attar for MRI procedures.
E.S. was supported by a European Research Council Grant 311238, an Israel Science Foundation Grant 1629/12, research support from Martin Kushner Schnur and Mr. and Mrs. Lawrence Feis, and a CRCNS grant. R.P. was supported by ISF #2352/19 and ERC-2016-CoG #724910 grant.

Introduction

Learning to classify multi-cue stimuli in order to produce the correct action is an adaptive flexible behavior required from animals on a daily basis. Accordingly, such tasks are commonly used to explore learning strategies in humans [1-5], as well as clinical implications [6-9]. Recent studies have shown that performance depends on complexity [10] and can be predicted using models that rely on high-order features of the stimulus with individual priors [11]. Studies of learning rule-based classification in monkeys have ascribed a major and complementary role to the striatum and regions of the prefrontal-cortex (PFC) [12, 13]. In paradigms that impose category boundaries on multiple stimuli, individual neurons in the PFC exhibit category preference to the different classes [14-21]. Within the PFC, the anterior-cingulate-cortex (dACC) widely projects to striatal regions [22-24], and is involved in several cognitive functions that contribute to the learning process itself beyond the final representation of the rule. Neurons in the dACC represent attention, reflect actions that lead to reward, signal outcome of previous trials, and form and integrate representations of task structure [25-37]. The striatum in turn, receives wide projections from the dACC and plays a role in choosing actions and supplies reinforcement signals which can help to establish a strategy during learning [38-46].

Yet, less is known about how single neurons form representations as learning progresses and gradually becomes relevant to the final classification being imposed. This is mainly because most studies follow extensive training and the neural correlates relate more to the final representation, perception, and recognition, and less to the gradual learning process. In principle, rule-learning can take several forms, and previous studies used few governing rules where the animals learn to assign different outcome probability or value [47, 48], acquisition of arbitrary stimulus-motor associations [49-51], or switching contingencies between the rules [20, 52].

Here, we examine neural dynamics during the learning process of several different rules. To do so, we trained two monkeys (macaca fascicularis) to perform trial-by-trial classification learning based on visual patterns composed of ‘bits’ of black and white squares. Each session required to learn a classification rule on patterns of N=3 bits. Patterns were presented in a pseudo-random non-biasing order (Fig.1a, Methods). Out of the 256 (2^2^3) possible rules, we chose seven rules in which the correct label was determined according to either single, pairwise, or triple-wise dependencies between the bits in the pattern (Fig.1b). These rules are unbiased, namely equally partition the set of 8 patterns; and are independent of each other, namely learning a rule results in chance performance for all other rules. Additionally, we included the majority rule (also unbiased, Fig.1b). We repeated this set of eight rules after >4 weeks to obtain enough neurons recorded per rule. The monkeys did not experience any of these rules prior to recordings. In these conditions of varying the rules every few days, we expected the animals’ performance to vary across sessions and rule types.

Results

Learning classification rules

Both monkeys exhibited within-session learning (Fig.1c), continuous performance improvement in longer sessions (Fig.1d), as well as next-day retention (Fig.1e). As expected from learning eight different rules that change every few days but use similar cues, neither of the monkeys showed retention benefit over the month between rule repetitions (Fig.1f). Despite the hard task, both monkeys achieved learning of all rules in some sessions yet with different levels of accuracy that range across sessions and rules (Significant improvement in 60% and 32% of sessions with 'easy' and 'hard' rules. Binomial tests, see methods, Fig.1g), and with differences in overall performance as reflected also in learning of 2-bit and 4-bit rules (Fig.1h). These results are highly similar to the behavior of human subjects that exhibited similar behavioral variability across individuals, rules, and sessions[11].

We tested and found that error-rates following correct classification show strong dependence on the specific rule (Fig.S1, Kruskal-Wallis test, $p < 1e-15$) and that the stimulus-response association learning of the more salient patterns showed variability across rules and strong dependence on the specific rule (Fig.S2, Kruskal-Wallis test, $p < 1e-4$ for any segment of 10 pattern presentations). This was further supported by rule consistency (Fig.S3) and the ability to learn 4-bit rules (Fig.1h, 16 patterns, 65536 possible rules). This evidence suggests that the monkeys did not use simple memorization or different saliency of the pattern cues,

We conclude that both monkeys learned in a significant number of sessions with some rules being easier to learn than others. The complexity and richness lead to reasonable error rates when compared to over-training on a specific set of rules, and therefore enable examination of neural patterns during successful learning.

Rule-based classification emerges in single neurons

To examine and compare neural responses, we recorded single-units during learning of all 3-bit rules in the dACC (Brodmann area 24), Caudate and Putamen (Fig.2a,b), and divided the sessions into 'easy' and 'hard' rules based on the average success rate (Fig.1g). We found neurons that changed their post-stimulus spiking pattern to differentiate the two category labels ('category-specific' neurons) (Fig.2c, left and middle), whereas other stimulus-responsive neurons maintained no category preference ('stimulus-specific' neurons) (Fig.2c, right). Overall, both the dACC and the Putamen had significantly more category-specific neurons during the late phase of sessions of easy rules compared to hard rules (Fig.2d, dACC: 21% vs. 12.5%, Pu: 22% vs 8% for easy vs. hard,

binomial comparison z-test, $p<0.05$ for both). The same result was obtained when dividing sessions into high and low performance independent of the specific rule, i.e. session-based performance rather than a rule-based performance (Fig. S4), and similarly in each monkey separately (Fig.S4). Moreover, there was a monotonic relation between instantaneous performance and the proportion of neurons showing correct classification (Fig.2e).

Since the category label and the actual choice are correlated during correct performance, we used the occurrence of error trials to distinguish them and compared the correlation between neurons' activity and the correct category to the correlation with the actual choice over trials (Fig.2f, William's test). In sessions of easy rules, more neurons developed significantly stronger category correlations in both dACC and Putamen, whereas significantly more neurons developed larger actual-choice correlations in the Caudate (binomial comparison z-test, $p<0.05$). Finally, in accordance with the aforementioned behavioral findings (Fig.S1-3), even under stringent conditions only very few neurons (3%, methods) had activity that can be attributed to memorizing of specific stimulus patterns; and moreover, most of these neurons were selective and sensory-specific to the all-black or all-white patterns (36/40, Fig. S5).

Together, these findings suggest that the modulation of single neuron activity in the recorded regions reflect different roles during learning: whereas the dACC and Putamen reflect more the learned rule, the Caudate reflected more the actual choice. We therefore focus in the following analyses on the dACC and the Putamen.

A geometric representation to track neural dynamics during learning

We next examine how the modulation of neural activity occurs dynamically and continuously during the learning process itself. To track representations during acquisition, we note that each pattern can be represented by its visual features (Fig.3a), and if we choose a complete set of statistically independent binary functions as features (a minimal spanning basis), we can describe any rule as a weighted combination of these features, i.e. each rule is a vector in this feature space. This choice of the basis also guarantees that a neuron's selectivity is defined by the correlation between its activity to the basis features (*neural-vector*), and we can therefore use rolling regression to identify the neuron's dynamic preference by the trajectory of its neural-vector in this space (Fig.3b).

Importantly, because the neuronal dynamics (neural-vector) is described in terms of movement in feature space and its relation to the rule-vector that varies across sessions, this geometric framework allows comparing neural dynamics across sessions of different rules. Specifically, two processes can contribute to a higher agreement (projection) between the neural-vector and the rule-vector: by increasing vector magnitude (Fig.3b, left) - reflecting a confidence strengthening (being 'louder'), or by rotating the vector towards the direction of the rule-vector (Fig.3b, right) - reflecting a policy change.

Finally, we note that choosing any other representation, i.e. another feature space, can be expressed as a weighted combination of the feature basis, and so our findings below do not depend on the specific set of features we used.

Reducing the high-dimensional (7-d) trajectories in feature-space to changes of angle-to-rule and vector-magnitude reveals several types of single-neuron dynamics. Whereas some neurons remained selective to a feature regardless of the performance improvement (Fig.3c), other neurons rotated towards the correct rule during learning (Fig.3d), and yet other neurons increased their vector magnitude and developed feature selectivity but with no relationship to the correct rule (Fig.3e). Importantly, many neurons performed a more complex path and their trajectories involved changes in several features in parallel (rotation and magnitude, Fig.3f, Fig. S6). The geometric representation of multi-feature trajectories allows interpretation for these as well – such as rotation in the orthogonal subspace ('Category ⊥', e.g. Fig.3f) and magnitude increases in building preference to a combination of several features, that eventually leads to significant agreement with the rule-vector (see Fig.S6 for further examples).

Neural trajectories closely match behavior

We next asked if this neural representation agrees with the behavioral changes in individual sessions. Because the neurons' trajectories in the feature space reflect the change in their category selectivity through the geometric measures of vector angle and magnitude, we estimated the correlation of these measures separately with the behavioral performance curve. We found a large and significant proportion of neurons in both regions for which changes in angle-to-rule were correlated with the behavior (Fig.4a, 32% of the dACC neurons and 37% of the Putamen neurons, Pearson, $p<0.01$). Similarly, a significant proportion of neurons changed their vector-magnitude in correlation with the performance (Fig.4b, 19% of the dACC neurons and 18% of the Putamen neurons, Pearson, $p<0.01$). These populations intersect such that many neurons with vector-magnitude changes also exhibit angle-to-rule changes (Fig. S7). Finally, we tested for the possibility of a temporal lag between the neural dynamics and the behavior. In all cases (for both regions and for both magnitude and angle changes), neurons were equally distributed between preceding change in behavior and slightly following it (Fig.4a,b, right-insets, Means not different than zero, $p>0.1$ for all, t-tests).

We further tested for differences in neural dynamics between regions, and quantified if there is a lag in either angle or magnitude changes between pairs of neurons recorded simultaneously from the dACC and Putamen. This revealed a wide distribution (Fig.4c), but without a specific directional lag ($p>0.1$ for all, t-tests), suggesting that information is being shared in both directions with both short lags (within the same trial) but also longer ones (over few trials), in agreement with the notion of information transfer in the underlying corticostriatal loops.

Representations of confidence and policy across regions

The similar temporal dynamics between neural trajectories and behavior allows relating the neural representations to the different components of learning: increasing vector magnitude reflecting a confidence increase, and rotating the vector towards the rule reflecting a policy change. We find that in easy rules, more neurons rotated towards the rule, i.e. diminished their angle-to-rule, in both the dACC and the Putamen (Fig.4d,e, binomial test comparing easy vs hard rules, dACC: 40% vs 25.5%, Putamen: 49% vs 34%). This rotation gradually progressed along the session as performance improved (Fig.4d,e– right panels).

However, a qualitative difference between the dACC and Putamen is revealed when examining vector-magnitude changes. In the dACC, there was no difference in number of neurons with vector-magnitude changes between easy and hard rules (Fig.4f, binomial test. dACC: 38% vs 34%). In contrast, in the Putamen, significantly more neurons increased their vector-magnitude in easy rules (Fig.4g, binomial test. Putamen: 47% vs 31%). Compared to rotation, vector-magnitude changes in the Putamen became much more prominent towards the end of the session, suggesting strengthening (confidence) of the policy that developed earlier (compare Fig.4d,e, right-panels to Fig.4g, right-panels).

Together, these findings suggest that both the Putamen and dACC reflect a process of policy search for the correct rule, but mainly the Putamen reflects policy strengthening once it is found and the gain in confidence that accompanies successful learning.

Differential dynamics across regions

To examine more closely if the changes in neural properties relate to behavioral improvement, we examined the correlation between two factors: the learning curve and the change in angle-to-rule, namely the change between subsequent windows (rather than the overall angle-to-rule as in Fig.4a). We then compared the proportion of neurons that showed significant correlations in easy vs. hard rules and found that significantly more neurons showed a change in angle-to-rule that followed behavioral changes in easy than in hard rules. This was so in both the striatum and the dACC when we computed the correlation with a lag of 4 trials (Fig.4h), and diminished when the lag was higher (8 trials, Fig.4h top-right) or with zero lag (Fig.4h top-left). There was no difference when the lag was negative i.e. when neural change precedes behavioral change. This suggests that when a search for answers succeeds or fails beyond average, it is followed by a neural vector rotation towards the rule. Although we cannot distinguish if this is due to success/failure itself or to the amount of provided reward, the fact that the rotation was towards the rule suggests it is not reward alone.

In addition, we exploited the simultaneous recordings to test if the two properties, magnitude and rotation, have differential dynamics between regions. We took the vector-magnitude and the rotation-to-rule of each neuron and computed the optimal lag (as in Fig.4c) between the two properties for all simultaneously recorded neurons. According to our hypothesis and in line with previous suggestions, only one distribution of lags was different than zero, that of vector-magnitude in Putamen neurons and angle-to-rule in dACC neurons, with the Putamen magnitude following dACC rotation (Fig.4i, $p<0.004$, t-test; Mean lag -2.8 ± 0.96 trials; Fig.S8, all other comparisons were not different than zero, $p>0.1$ for all, t-tests). Moreover, this was driven mainly by neural pairs recorded in easy rules (Fig.4j, $p<0.01$, t-test, Mean lag -3.56 ± 1.36 trials).

Together, the results suggest that once a strip of successful responses occur, the neural vector in both regions rotate towards the rule, and that rotation in the dACC is followed by magnitude extension in the striatum.

Neural vectors predict next-day (overnight) behavior

If the neural trajectories indeed reflect learning and changes in behavioral policy during a session, then the neural-vector at the end of a session represents the acquired behavior policy and might predict early performance in the next daily session (i.e. an overnight retention process).

To examine this, we projected each neuron's activity vector at the end of a session on the rule of the following day, and computed the correlation between these projections and the early next-day's performance. In Putamen neurons only, this revealed a significant correlation between similarity of neural vectors to the next day's rule and next-day early success rates (Fig.4k, Pearson's $r = 0.24$, $p<0.02$). Namely, the closer the neural vector at the end of the day to the next-day rule, the better will be the initial performance. There was no correlation when we reversed the days as control.

Because the above included the retention benefit of the learned rule when the next-day's rule was identical, we further tested if such priors could also bias learning of new rules in sessions when the next-day rule change. We again projected the neural vector onto the actual next-day's rule and estimated its correlation with the early performance in the next day, but this time only for pairs of days when the rule changed in the second day. There was a significant correlation here as well for Putamen neurons only (Fig.4l, Pearson's $r = 0.35$, $p<0.03$), between the projection of neural activity from the preceding day onto the subsequent rule and performance on that rule early in learning. There was no correlation when we reversed the days as control.

Together, these findings link between the geometric representation and the behavior and demonstrate that the neural projection is indeed a valid representation for the learned rule, and further suggest that it can bias future behavior.

Discussion

We recorded and analyzed dynamic changes in neural representation while monkeys learned to classify multi-cue patterns based on a rich set of potential rules. In these natural conditions of varying eight different rules, the animals' performance varied across sessions and rule types as previously shown in humans [11]. Nevertheless, we found that neurons in both the striatum and the dACC change their firing pattern to eventually represent the learned rule. To address the main challenge of characterizing single-unit dynamics during de-novo learning of rules, we developed an approach that represents each rule in the space formed by a set of features. Because this set of features is a minimal spanning basis, other cartesian coordinates that allow a vector representation are equivalent to it, making the findings generalizable and independent of the specific chosen features. We could therefore summarize each neuron activity by two universal traits – its angle to the rule direction and its magnitude. These two measures have a natural interpretation, the rotating angle reflects learning-related policy or strategy change, and the magnitude increase reflect strengthening or confidence (neural confidence, we do not quantify here cognitive confidence). Further, this has the benefit of denoting all rules that can be formalized by any high-dimensional stimuli space with the same two traits, and hence compare the dynamics of learning across tasks and rules.

Using this framework, we found that in both regions there is a large number of neurons that were dynamically synchronized with the behavior, either in their angle or in their magnitude. This suggests that this neural representation indeed captures the learning process [50] and reorganizes to adapt to the new conditions [53]. Importantly, this representation revealed a dissociation in functionality: neurons in the dACC rotated to decrease their angle-to-rule, namely changed their strategy; whereas neurons in the Putamen changed their activity to reflect both strategy and confidence, by magnitude-increase of the neural-vector that likely reflects strengthening and reinforcement of the correct strategy once identified during learning [39]. In line with this interpretation and the role of the striatum in reinforcement, we found that rotation changes in the dACC were followed by magnitude extensions in the striatum, and more so in easy rules.

In addition, apart from representations of the learned rule, we could also identify neurons with stable or changing correlations in the subspace of the features [54-56], and these were observed during both successful and unsuccessful learning sessions. The exact role of these neurons is yet to be determined, but it seems likely that they reflect a trial-and-error search process during learning or retention of other previously learned rules.

In contrast to the Putamen and the dACC, Caudate neurons showed stronger reflection of the monkeys' instantaneous choice than of the categorization rule. Because in visuomotor associations the dorsal striatum likely represent actions' value [38] and follows the choice representation in lateral PFC [42], yet the role of lateral PFC in category learning and generalization is still under debate [57], our

observations can support the notion that in rule learning Caudate activity reflects value-based action selection [40, 46, 58, 59]. Future studies are required to clarify the caudate specific role in rule-learning. Finally, neurons in the Putamen were related to overnight retention, and their representation at the end of a day predicted next-day behavior. Because we imposed a hard task and learning usually continued over few sessions, this demonstrates that the striatum represents the policy that is used for retention, either within the striatum or by transfer to other regions (e.g. as in consolidation processes). This suggestion is in line with studies showing that the striatum maintains intermediate representations, potentially via sustained activity [60], to allow learning that combines reinforcement and memory under spaced conditions[61, 62].

Our results can further suggest how abnormalities in the cingulate-striatal network can result in maladaptive learning processes that lead to applying incorrect rules, and in extreme cases lead to psychopathologies[63-66]. Overall, we present a new computational framework to examine dynamics of neural changes, and suggest complementing roles for the dACC and Striatum in learning and retention of classification rules.

Author contributions: Y.C, E.S. and R.P. conceived and designed the study; Y.C. performed the experiments and analyses; Y.C., E.S. and R.P. wrote the manuscript.

Figure 1

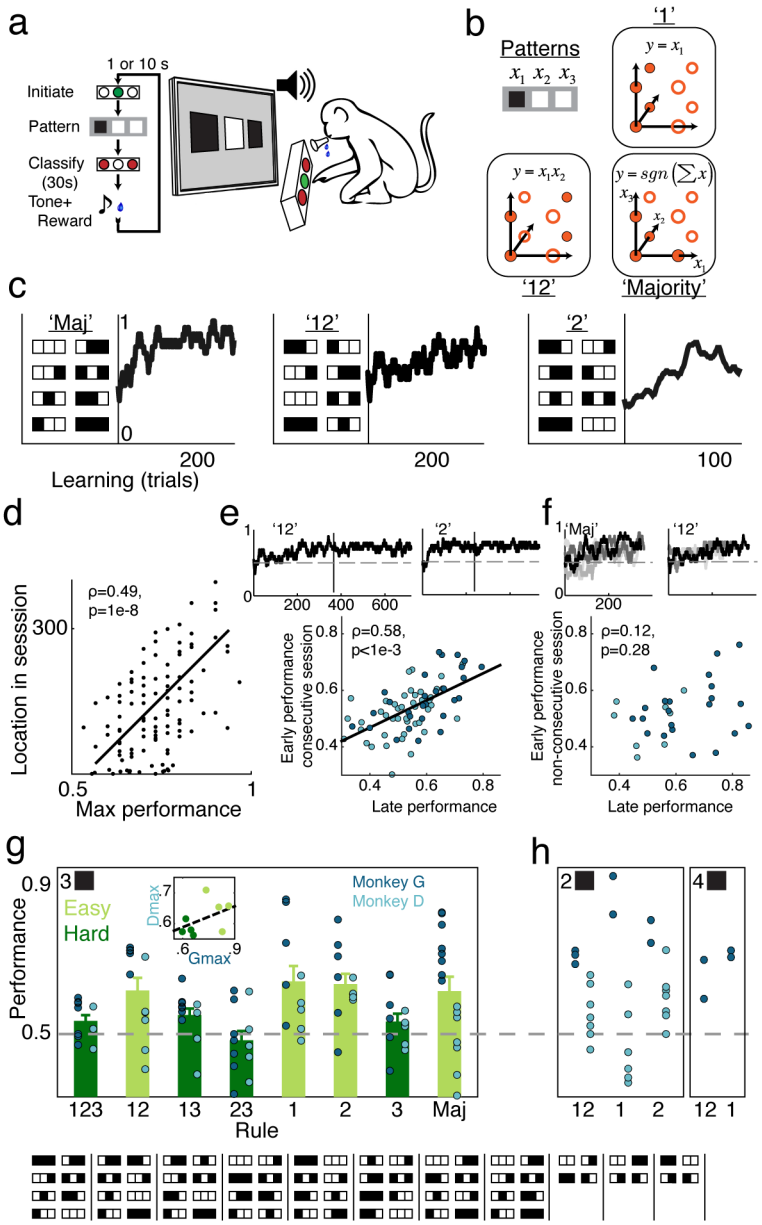

Figure 1. Monkeys learn classifications with varying complexity.

a. Behavioral paradigm: pressing and holding the middle button initiates a new trial. After the pattern appears on the screen the monkey has 30s to classify it with the left or the right button. A fluid reward follows a correct choice and a short timeout after an incorrect choice.

b. A scheme of the rule-based classification. Shown are: one-bit-rule '1' (i.e. decision is based only on the identity of the first bit), two-bit-rule '12' (decision is based on a XOR of bits '1' and '2'), and the Majority-rule (decision is based on summing bits). In each panel the two categories are represented by full and empty circles in the space defined by the individual bits.

c. Learning curves. Shown are individual learning curves from 3 example days each with a different rule (monkey G – two left examples, monkey D – right example). The curves are shown next to the underlying truth table of the rule where patterns are stacked by category.

d. Continuous longer learning within-session is beneficial. Shown is the location within the session of the maximum performance plotted against the maximum performance (best 30 consecutive trials**),** for all sessions.

e. Overnight retention. Shown are two examples of learning curves in consecutive days with vertical lines separating the sessions. Over all sessions, there was a significant correlation between performance at the end of a session and performance at the beginning of the subsequent session with the same rule (monkey D: light blue, $r=0.48$, $p<0.001$; monkey G: dark blue, $r=0.61$, $p<0.001$; Both: $r=0.48$, $p<0.001$).

f. No retention between rule repetitions with >4 weeks separation. The use of eight different rules that use the same cues/stimuli over 4-5 weeks means it is virtually impossible to memorize the different rules. Shown are two examples of three sessions (scales of gray) that repeated the same rule but with >4 weeks between repetitions. Over all sessions, there was no relationship between performance at the end of learning a rule and performance at the beginning of next repetition of the same rule (monkey D: light blue, $r=0.23$, $p>0.5$; monkey G: dark blue, $r=0.21$, $p>0.3$; Both: $r=0.12$, $p>0.2$).

g. Performance in all 3-bit rules, average performance in the last quarter of each session for all rules and all sessions, for both monkeys (D: light blue; G: dark blue). In addition, for each rule, shown is the mean and SE averaged over animals and sessions, and classified into ‘easy’ and ‘hard’ rules (in ‘easy’ rules performance is significantly above chance-level on average). Inset show maximal performance in each rule comparing the monkeys. Truth tables for all rules are shown below.

h. Performance in 2-bit and in 4-bit rules (neural activity was recorded only during 3-bit rules).

Figure 2

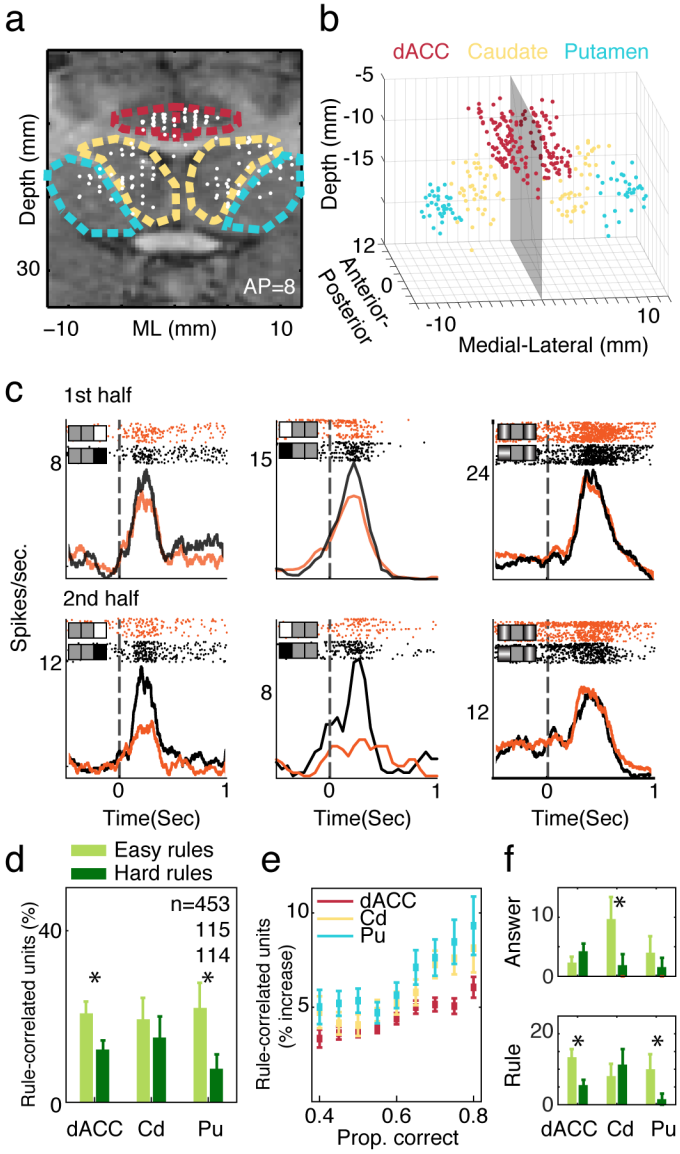

Figure 2. Neurons represent learning of classification.

a. Recording locations projected on a coronal MRI section. Recording locations span ~12mm in the anterior-posterior axis (hence some locations seem to be outside the regions of interest). Red marks borderline of the dACC, Yellow for Caudate and cyan for Putamen.

b. 3-D reconstruction of all recording locations. Gray plane is the midline, anterior-posterior zero is at the anterior-commissure and depth is measured from the dura surface.

c. Rule-learning in neurons. Spiking patterns are divided by the category label (orange and black rasters and PSTHs). Spike times are aligned to the stimulus onset (dashed line). The order of trials is top to bottom. Trials are divided into the first half (upper row) and second half (bottom row) of the

sessions. The two left neurons differentiated between categories in the second half significantly more than in the first half, whereas the right neuron did not learn to differentiate categories.

d. Fraction of single-units whose firing correlated with the rule at the end of learning (Mean and SE). Sessions are divided according to easy vs. hard rules, with more neurons signaling correct classification in easy rules in the dACC and the Putamen ($p<0.05$ for both, binomial tests).

e. Instantaneous performance parallels neural-based categorization. The fraction of neural segments with significant category correlation increases with the mean performance (calculated by rolling regression windows of 40 trials in steps of 4 trials).

f. Proportion of neurons with stronger correlation to the rule (bottom) and with stronger correlation to the actual choice (top), over both successful and error trials (William's test), and comparing easy vs. hard rules. The dACC and the Putamen are significantly correlated more with the rule and not with the actual choice, whereas caudate neurons are significantly correlated with the choice only and not with the rule (significant comparisons are $p<0.05$, binomial tests).

Figure 3

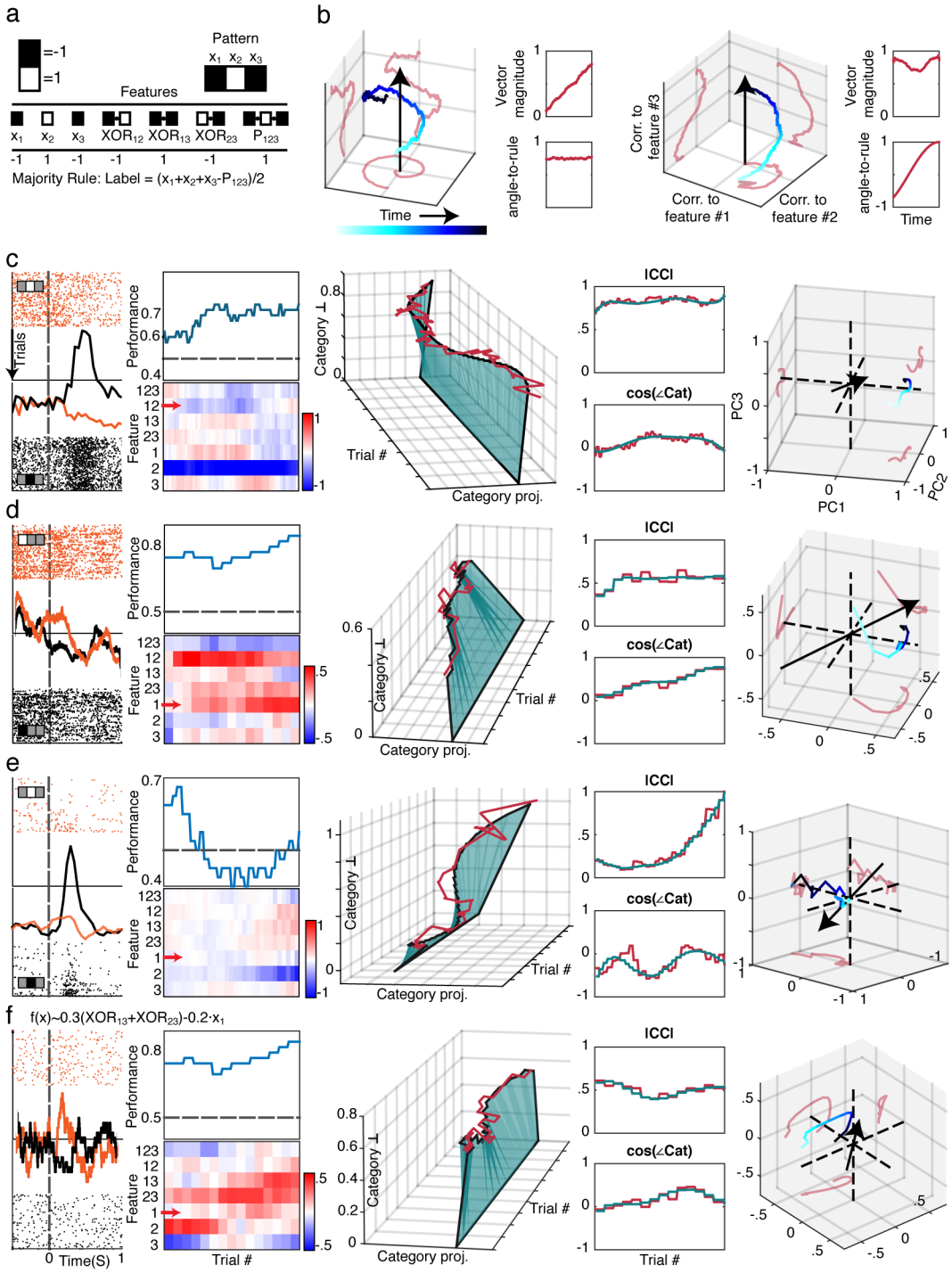

Figure 3. A geometrical representation reveals single-neuron dynamics during learning.

a. The 7 features that form a linear basis (minimal and spanning) for the space in which all 3-bit rules reside. Shown is an example for the representation of the pattern $\vec{x}$ =■□■= $(-\mathbf{1}, \mathbf{1}, -\mathbf{1})$ by its

parity features of: first order, $\{\boldsymbol{x_i}\}_{i=1}^{3}$, second order, $\{\boldsymbol{XOR_{ij}} = \boldsymbol{x_i} \cdot \boldsymbol{x_j}\}_{i<j=1}^{3}$, and third order, $\boldsymbol{P_{123}} = \boldsymbol{x_1} \cdot \boldsymbol{x_2} \cdot \boldsymbol{x_3}$ and the representation of the majority rule as a combination of these features.

b. In this feature space, the change in a neuron's response modulate its projection on the vector that represents the rule (the black arrow). An increase in the projection can result from two processes: a trajectory change that increases the neural-vector magnitude (left) or neural-vector rotation towards the direction of the rule-vector (right), or of course both. The trajectory is color coded for time and red curves depict its projections on the axes. Insets separate the dynamics of vector-magnitude (top) and angle-to-rule (bottom). For presentation purposes only, the neural trajectory is plotted in the space of the three main principle axes (PCA), yet for all analyses the neural-vectors were computed in the full 7D space.

c-f. Examples of single-neuron neural-vector dynamics. Each row is a neuron and shows, from left to right:

1. Rasters aligned to stimulus onset with trials advancing from top to bottom and divided by the sign of the preferred feature at the second half of the session. PSTHs for the top raster (orange) and for the bottom one (black).
2. Top: the behavioral learning curve (20 trials running average). Bottom: the regression correlation coefficients (blue-red color-bar) for each feature in the basis (y-axis) along all trials (x-axis). Red arrow marks the feature that defined the correct label in the specific session.
3. The norm projections of correlation vectors (red) on the category (x-axis) and its orthogonal subspace (Category ⊥, z-axis) over time (y-axis). The blue surface connects (0,0) to the smoothed dynamics to visualize rotation and extension.
4. The correlations' vector magnitude (top) and angle to category (bottom) plotted over time. Red: raw data. Blue: smoothed data.
5. The high dimensional (7D) neural trajectory projected on the three main principle axes (PCA) and curve fitted, color-coded for time progress in the session (light → dark blue). The black arrow is the rule vector in that session and the red curves are the projections of the trajectory on the axes (principal components).

c. A neuron that does not change its selectivity during learning

d. A neuron that rotates towards the correct category by learning the right feature

e. A neuron that extends its vector towards a 'wrong' feature.

f. A neuron that exhibits a complex relationship of vector extension and rotation.

Figure 4

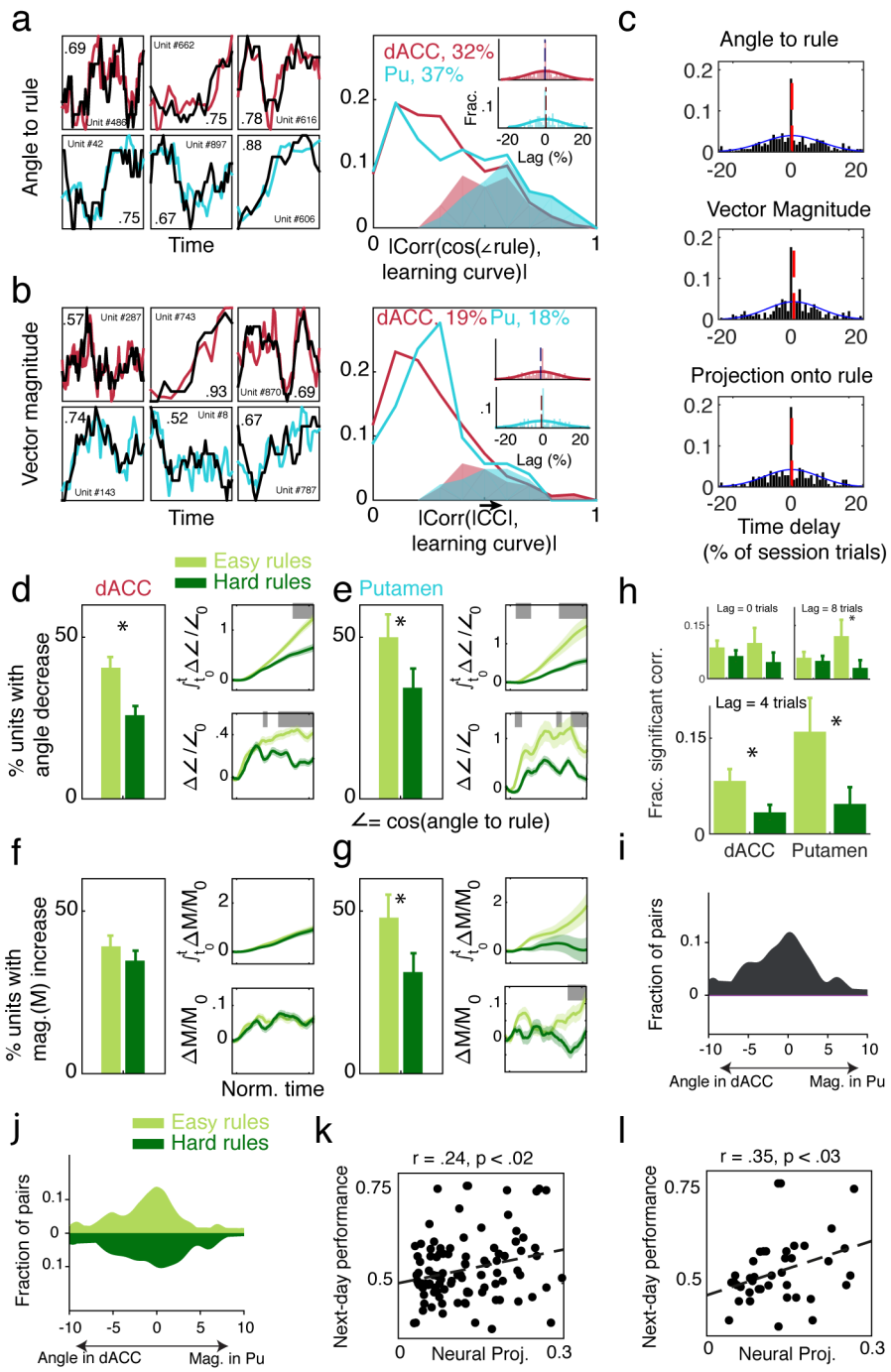

Figure 4. Neural dynamics match learning behavior and varies across regions.

a. Change in angle-to-rule ($\cos(\angle \boldsymbol{Rule})$) during a session overlaid with the performance behavior (black), showing a highly similar temporal pattern. Three dACC neurons (upper row, red) and three Putamen neurons (lower row, cyan), all with significant correlation between neural dynamics and

performance ($p<0.01$ for all, Pearson).
Right panels show histograms of correlation coefficients for all neurons, with shaded area mark neurons exhibiting significant correlations at $p<0.01$. Both regions contained a highly significant number of neurons with dynamics similar to behavior ($p<0.01$ for both, χ^2). Insets show the distribution of trial lags between neural dynamics and behavior, with the mean lag not different than zero ($p>0.1$, t-tests).

b. Same as in (a) for dynamics in vector magnitude ($|\boldsymbol{CC}|$).

c. The distribution of trial lags between neural dynamics of dACC neurons and of Putamen neurons, for all simultaneously recorded pairs in both regions ($p>0.1$ for all, t-tests). Shown are the neural lags when taking into account only angle (upper), magnitude (middle), and combined projection on the rule-vector (bottom).

d,e. Proportion of neurons that significantly decreased their angle-to-rule during the session was different in easy rules vs. hard rules in both the dACC (**d**) and the Putamen (**e**) ($p<0.05$ in both, binomial test). Right panels show the average angle change (bottom, SE in shaded color) and the normalized cumulative change (top). Sessions were time-warped for averaging. Gray bars indicate significant difference between easy and hard rules ($p<0.05$, bootstrap).

f,g. Same as in (d,e) for change in vector magnitude. However, in contrast to rotation in the both regions, only the Putamen showed significant increase in the vector-magnitude in easy vs. hard rules.

h. Local shifts in angle-to-rule follow the behavior in more neurons when comparing easy to hard rules in both regions, but only in a lag of 4 trials (main panel) and not in zero or 8 trials lag (top insets).

i. Distribution of optimal lags for all simultaneously recorded pairs of neurons, between change in vector magnitude of the putamen neuron and angle-to-rule of the dACC neuron. The mean is significantly below zero, indicating that changes in angle in dACC neurons preceded changes in magnitude in Putamen neurons. All other combinations were not significant (Fig.S8).

j. Same as in (i) but separately for easy and hard rules.

k. Neural projection at the end of a session predicts next day performance. The projection of each Putamen neuron's activity at the end of the session (neural-vector) onto the next-day rule (rule-vector), against the mean performance in the beginning of the next-day session (the neural-vector is averaged over the last 25% of the session and performance is averaged over the first 25% of the next day session), showing a positive correlation ($r=0.24$, $p<0.02$, black regression line).

l. Similar to (h) but when the rule changed overnight, and the neural-vector at the end of the session is projected onto the next-day rule, showing a positive correlation ($r=0.35$, $p<0.03$)

Materials and Methods

All surgical and experimental procedures were approved and conducted in accordance with the regulations of the Weizmann Institute Animal Care and Use Committee, following National Institutes of Health regulations and with accreditation from the Association for Assessment and Accreditation of Laboratory Animal Care International.

Two male monkeys (Monkeys G and D, macaca fascicularis, 4-6kg) participated in the experiment. Before data collection, each monkey went through a training phase that acquainted it with all the task components and their sequence in a learning session. Therefore, in the recording sessions the monkeys were familiar with the concept of patterns and classification (Fig.1a). However, the monkeys did not experience any of the rules reported in this manuscript before the electrophysiological recordings began. Each classification rule was replaced every few days and repeated after about a month and after a full cycle of the 8 rules was presented.

The monkeys learned to classify binary patterns of N=3 squares. In each session, the entire set of 2^N possible patterns was presented. The order of patterns was generated by concatenating full sets of randomly ordered 2^N patterns. This process ensured that all patterns appear with the same temporal frequency and that no choice of behavioral rule, apart from the correct one, is beneficial in large portions of the session. For compactness we refer to the rules by their constituent squares. So, for example, in rule '3' the label is determined by the color of the 3rd square and in rule '12' the label is determined by the XOR of squares 1 and 2. See Fig.1 and main text for the list of rules used in this study and during recordings.

Neural recordings

A craniotomy was performed under deep anesthesia and aseptic conditions and a recording chamber (27x27mm) was implanted above the midline and anterior commissure to allow daily electrodes insertion. The chamber's positioning was done according to MRI calculated coordinates with respect to the identified bone structure around the ear canals and eye sockets. Still images were taken during the surgery to record the location of the chamber, the head holder and the screws on the skull for easier extraction process.

After surgery the monkeys were treated with analgesics (Buprenorphine) and antibiotics (Rocephin, Baytril). The monkeys were allowed to recover for 1-2 weeks before the first head restraining in the setup. The fluid consumption regime was gradually reinstated starting two weeks after surgery.

MRI-Based Electrode Positioning: Anatomical MRI scans were acquired before, during, and after the recording period. Images were acquired on a 3-Tesla MRI scanner: (MAGNETOM Trio, Siemens) with a CP knee coil (Siemens). A T1-weighted, three-dimensional gradient-echo (MPRAGE) pulse

sequence was acquired with a repetition time of 2,500 ms, an inversion time of 1,100 ms, an echo time of 3.36 ms, an 8 flip angle, and two averages. Images were acquired in the sagittal plane, 192 x 192 matrix, and 0.63 mm resolution. The first scan was performed before surgery and used to align and refine anatomical maps for each individual animal (relative location of the dACC and the Striatum, and anatomical markers such as the interaural line and the anterior commissure; confirmed using atlas). We used this scan to guide the positioning of the chamber on the skull at the surgery. After surgery, we performed another scan with 2-4 electrodes directed toward the dACC, Putamen and caudate. The regions' depth was calculated from the dura surface and the plane of the top of the chamber. We assessed estimation of electrode tip locations and comparison to the MRI image with <1mm accuracy (mean=0.5mm).

Mapping recording regions: During the first week of electrode insertions we performed a mapping procedure to identify the depth of cell bodies in prominent recording regions. During that week no behavior recordings were made and the fluid restriction was gradually reinstated.

Additionally, with every electrode insertion during the experiment we recorded the depths of cell bodies and were able to reconstruct the boundaries of our regions of interest.

Electrophysiology:

The monkeys were seated in a dark room and each day, up to six microelectrodes (0.6–1.2 MΩ glass coated tungsten, Alpha Omega) were lowered inside a metal guide (Gauge 25xxtw, outer diameter: 0.51 mm, inner diameter: 0.41 mm, Cadence) into the brain using a head-tower and electrode-positioning-system (Alpha-Omega). The guide was lowered to penetrate and cross the dura and stopped once in the superficial layer of the cortex. The electrodes were then moved independently further into either the dACC, Caudate or Putamen. Electrode signals were pre-amplified, 0.3 Hz–6 kHz band-pass filtered, and sampled at 44 kHz; and online spike sorting was performed using a template-based algorithm (Alpha Lab SNR, Alpha Omega). We allowed 15-30 minutes for the tissue and signal to stabilize before starting acquisition and behavioral protocol. At the end of the recording period, offline spike sorting was further performed for all sessions to improve unit isolation (offline sorter, Plexon).

Data Analysis – Behavior

Performance:

Each learning session results in a series of correct and incorrect answers, $\{y_t\}_{t=1:T} \in \{0,1\}^T$, T being the number of trials. To measure learning behavior and account for erratic tendencies we took the following steps:

1) To avoid the behavioral decline that may bias performance at the end of the sessions we disregarded up to the last 10% of the session if it contained only wrong answers. On average we ended up ignoring ~1% or 2-3 trials in each session.

2) We define performance at the end of the session by averaging correct and incorrect answers in the last quarter of the session, $P_{end} = \langle y_t \rangle_{t \in \left[\frac{3}{4}T \to T\right]}$. The confidence level for rejecting the null hypothesis of chance performance follows the regularized incomplete beta function, $I_{1-p}\left(\frac{T}{4} - k, 1 + k\right)$, where $\frac{T}{4}$ is the number of trials in the last quarter of the session and k is the number of correct answers during that segment.

3) Identically to P_{end}, we define P_{start} as the mean performance during the first quarter of the session.

4) In Fig. 1d, we define maximal performance as the best mean performance in 30 consecutive trials, $P_{max} = \max_{\tau \leq T-29} \langle y_t \rangle_{t \in [\tau \to \tau+29]}$.

<u>Easy and hard rules:</u>

We label rules according to the monkeys' ability or inability to recurrently achieve high performance in learning those rules. Thus, given that a subset of rules is labeled 'easy' and another subset is labeled 'hard', we computed the amount of variance, within the set of P_{end}'s that the labeling explains. The R^2 value is: $R^2 = 1 - \frac{\sum_i \left(P_{end}^i - \mu^i\right)^2}{\sum_i \left(P_{end}^i - \mu\right)^2}$ where, $\mu = \langle P_{end} \rangle$ is the mean end-performance of all the sessions and μ^i are the mean end performances of sessions of either 'easy' or 'hard' rules (Fig.1g). We also labeled individual sessions as 'high' or 'low' performance independent of the rules to make sure our results are robust (Fig.S4).

<u>Rule repetition effects</u>

Each classification rule was used for 1-5 consecutive days and repeated after about a month. We compare the monkeys' performance at the end of a session ($P_{end}(n)$ with 'n' standing for the n'th session of the rule) to the mean performance in the first quarter of the following session of the same rule, $P_{start}(n+1)$. The comparison is made by calculating the Pearson correlation between the P_{end}'s and P_{start}'s.

We examined two distinct cases:

1. Taking sessions only from consecutive days, we calculate the correlation, ρ_{across}, of the across-days learning (Fig. 1e).
2. Taking only sessions from the end of a consecutive sequence and the beginning of the following sequence, we calculate the correlation, ρ_{recall}, of the monkeys' ability to recall rules they encountered a month before (Fig. 1f).

<u>Testing for pattern-specific memorization:</u>

We first consider a memorization scheme in which subjects perfectly learn a list of correct pattern-label pairs but do not generalize. The acquisition of such memorized patterns can be an all-or-none event, which means that after a certain pattern-label pair was memorized it will dictate choice behavior. Alternatively, we consider a gradual probabilistic association strengthening process for the observed patterns, which leaves room for errors. Our data rules out the case of all-or-none memorization: Fig. S1 and Fig. S2 demonstrate that our subjects frequently made mistakes on specific patterns even after they were labeled correctly in previous presentations.

We can also rule out memorization of the types mentioned above as the sole mechanism. Relying on memorization alone would mean that all rules on patterns of three bits would be learned in the same rate. Fig S1 clearly shows that this is not the case, and that rule identity plays a key role. Even finer memorization aspects, such as pattern-specific acquisition rates, can also be ruled out from our data. Fig.S2 shows that subjects learned the labeling of the same pattern under different rules at different rates.

Behavior stability tests:

Feature based behavior stability

We want to summarize how consistent were the monkeys in a single number for each session (Fig.S3a). This is done with answers from the last 1/4 of each session. We define as a consistency measure the mean (across patterns) distance of the logistic classifier (fitted to answers in the last 1/4 session) from the chance (0.5) answer.

Namely, if the monkeys adopt a feature based consistent policy at the last quarter of each session, then we can fit their sequence of answers with:

$$P(y=1|x;\vec{\alpha},\gamma)=\frac{1}{1+exp\left(-\gamma-\sum_{\mu}\alpha_{\mu}f_{\mu}(x)\right)} \qquad \textbf{[Eq 1]}$$

where x are the presented patterns, $f_{\mu}(x)$ are the features, and $\vec{\alpha},\gamma$ are fitted to maximize the likelihood of the answers.

A key reason to fit this classifier and not to make the consistency estimation per pattern is that there are much fewer pattern presentation per pattern and in any way calculating per pattern imposes the assumption that the monkeys can tell all patterns apart from each other. The classifier's way doesn't make any assumption beyond a features based behavior policy.

The consistency measure is thus the mean distance from 0.5. or,

$$Consistency=\frac{1}{8}\sum_{x}|P(y=1|x;\vec{\alpha},\gamma)-0.5| \qquad \textbf{[Eq 2]}$$

The chance level for a completely unbiased classifier is $\frac{\sum_{n=0}^{8}\left|\frac{n}{8}-0.5\right|\cdot\binom{n}{8}}{\sum_{n=0}^{8}\binom{n}{8}}=0.1367$

Across sessions stability

Next, we want to check if the monkeys were stable across sessions. Namely, regardless of performance, how similar is the features-based behavior at the last quarter of different sessions. Or, how similar is the behavior when learning the same rule in different sessions.

To avoid cross interference between the within-trial variability (the inconsistency that drives $P(y|x)$ in Eq. 1 close to 0.5) and the across-trial variability (The inconsistency that separates the logistic classifiers that are fitted to different sessions) we threshold the classifiers, $y(x) = [P(y|x) > 0.5]$, and for each rule compare all pairs of sessions. Fig.S3b shows the mean (and SE in errorbars) of the across session similarity score:

$$S_{rule} = \frac{1}{N_{pairs}} \sum_{i<j} \frac{1}{8} \sum_{x} \left[y_i(x) \cdot y_j(x) + \left(1 - y_i(x)\right) \cdot \left(1 - y_j(x)\right) \right] \qquad \textbf{[Eq 3]}$$

where i, j are different sessions of the same rule and N_{pairs} is the number of pairs of sessions with the same rule.

Behavior similarity across rules

To estimate how similar is the monkeys' behavior to the rules they learn we repeat the calculation in Eq. 3 but replace one of the classifiers (y_j) and subtract 0.5 to shift the mean expected overlap to 0. The, above chance level, results are presented in Fig.S3c.

Data Analysis – Neural activity

Single-neuron responses

We expect the learning-relevant cognitive mechanisms to be influenced by both trial-by-trial variations, such as changing behavior and stimulus identity, and by slower processes, namely learning. Studying the learning related dynamics, we are interested in the single unit neural activity that correlates to such inherently variable computational primitives. Namely, we seek a measure of the spiking activity that communicates the variations across trials. Accordingly, for every neuron we examine the spikes in the 500ms following the stimulus onset and bin them into 5x100ms segments to obtain sensitivity to temporal effects in addition to the spike count. The result is a 5-vector of spike counts from each trial, $\vec{V}(t = 1 \ldots T) \in \mathbb{R}^5 \times \mathbb{N}$. In this representation, the component of largest across-trials variance is $\overrightarrow{v^*} = arg \max_{\|\vec{v}\|=1, \vec{v} \in \mathbb{R}^5} Var\left(\vec{V} \cdot \vec{v}\right)$. We then project each 5-dimensional vector on this principle component and get a single number from each trial, $r(t) = \left(\vec{V}(t) - \langle \vec{V} \rangle\right) \cdot \overrightarrow{v^*}$. This number scores the spiking patterns of the neuron with respect to its most prominent fluctuations or change. Importantly, several unrelated processes may contribute to the across-trials variability and in choosing the projection, $r(t)$, as the representation of stimulus neural response, we tune to the largest source of variability, regardless of its nature.

Pattern-specific neurons

We devised a criterion for exemplar preference based on a neuron's firing rate in the 500mSec after the pattern presentation. We build a table of all responses of the neuron to each of the patterns. Next, we compare the sets of responses to each pair of patterns using a rank-sum test for equal medians and treat the distribution of responses as different using a threshold at p<0.05. A neuron is pattern-specific if the distribution of responses to only one pattern is different from all the others (Fig.S5)

Feature-based representation of the rule and neural responses

For a pattern, $\vec{x}$, we chose a basis of features that are polynomials of the variables x_1, x_2, x_3 that take the values ±1. Features, $f_{ijk}(x) = x_1^i \cdot x_2^j \cdot x_3^k$ differ in the polynomial degrees $(i, j, k = 0,1)$ and satisfy, when averaging across all patterns, $\langle f \rangle = 0, \langle f^2 \rangle = 1$.

Lemma 1: Different features in this base are statistically independent.

Proof 1: Let f_1 and f_2 be features in this set and without losing generality assume that they differ in the polynomial degree of x_1 s.t. f_2 doesn't contain x_1. Since $\forall_i, P(x_i = 1) = 0.5$ we get that $P(f_1 = 1|f_2) = P(x_1 = 1) = 0.5 = P(f_1 = 1)$.

Lemma 2: Correlation projections factor in this basis

Proof 2: Let $\boldsymbol{f}$ be a vector in features space, $\boldsymbol{f} = \sum_{I \in \{i,j,k\}} a_I f_I$ such that $\sum_I a_I^2 = 1$. This means that $\langle \boldsymbol{f} \rangle = 0$ and $\langle \boldsymbol{f}^2 \rangle = 1$ (because $\langle f_I \cdot f_J \rangle = 0, \forall_{I \neq J}$). Which leads to $Var(\boldsymbol{f}) = 1$. So, if n is some random variable (say, the neural projection) the correlation $C(\boldsymbol{f}, n) = \frac{\langle (\boldsymbol{f} - \bar{\boldsymbol{f}}) \cdot (n - \bar{n}) \rangle}{\sqrt{Var(\boldsymbol{f}) \cdot Var(n)}} = \sum_I a_I \frac{\langle (f_I - \overline{f_I}) \cdot (n - \bar{n}) \rangle}{\sqrt{Var(f_I) \cdot Var(n)}} = \sum_I a_I \cdot C(f_I, n)$.

This means that if we measure the correlations to the features separately, then the unit vector, $\boldsymbol{a}$, that maximizes $\sum_I a_I C(f_I, n)$ will give us the preferred feature. Also, if $\boldsymbol{a}$ is a rule that we chose in advance, e.g. the one being learned, then the projection $\sum_I a_I C(f_I, n)$ is indeed the rule-correlation (Fig.3b).

Dynamics of representations

To study the dynamics of task related neural correlates we divided each session to partially overlapping windows (40 trials segments with 4 trials jumps). For each neuron, calculating the correlation between its spiking patterns, $r(t)$, following stimulus onset, and the stimulus features, $\vec{f}(\vec{x})$, (as well as to the correct category and the monkey's future answer) yields a set of correlation coefficients, $CC_i(t) = corr\left(r(\tau \in w_t), f_i(\vec{x}(\tau \in w_t))\right)$, for each regression window w_t. These rolling regression coefficients were used to calculate the following measures:

Comparing representation between conditions

To judge whether neurons show rule selectivity during a certain segment of the session (Fig.2d) we test the fraction of regression windows within that segment, that exhibit significant rule correlation (Pearson, p<0.05). This test is done comparatively between sessions of different conditions, and we set a criterion of 10% to declare a neuron as showing rule selectivity during the segment. If there are more than 5 neurons meeting each condition (2 conditions, e.g. easy and hard rules) we use the binomial comparison z statistic, $z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\cdot\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$ with $\hat{p}_1, \hat{p}_2$ the measured success rate in two populations of sizes n_1, n_2 and $\hat{p} = \frac{n_1\hat{p}_1 - n_2\hat{p}_2}{n_1+n_2}$.

Comparing rule vs. answer representations

Since these correlations have a mutual component (the spiking pattern) and interrelate via the performance level, we compare with William's t-statistic for correlated correlations, $t = \frac{(C_{12}-C_{13})\cdot[(n-3)\cdot(1+C_{23})/2D_3]^{\frac{1}{2}}}{\sqrt{1+\frac{(n-3)\cdot(C_{12}+C_{13})^2\cdot(1-C_{23})^2}{8(n-1)D_3}}}$, where C_{12} is the category correlation, C_{13} is the answer correlation, and C_{23} is the correlation between answers and categories. n is the number of trials and D_3 is the determinant of the sample correlations matrix. The statistic is compared to the t – distribution with $n-3$ degrees of freedom (Fig.2f).

Relating neural representation to behavioral performance

To relate any regression measure and performance within a group of neurons we take the following steps:

1. For every 40-trials-long regression window we calculate the mean performance.
2. Given a performance level, we collect all the regression windows with performance within 0.15 of that level and calculate the mean and standard error of the measures of interest.

Angle-to-rule and vector-magnitude

Given a basis of visual features there is a unique spanning of the classification rule in each session. For each regression window we define the angle-to-rule as the angle between the vector of correlation coefficients to visual features and the vector that represents the rule. Similarly, we define the features' correlation magnitude as the norm (L_2) of the correlation coefficients vector.

When presenting the learning related dependence of these geometrical variables over time, we smooth them with 10 percent of the running windows in a session (Fig.3, Fig.4d-g insets)

Session-length standardization

Several calculations require the comparison or grouping of segments from relative session fractions and/or location. To enable this, we standardized the regression measures from each session

to a fixed length of 100 bins. This means that all rolling regressions were stretched to the same length, because a hundred regression windows would only come from 436 trials in a session.

Cells that reduce angle-to-rule or increase vector-magnitude

To quantify neurons that decreased angle-to-rule or increased the vector-magnitude (Fig.4d-g), we compare regression windows in the first 15%-segment of the sessions to regression windows in the last 15%-segment of each session with a 1-tailed t-test. The fractions of cells that passed the test are compared with a 1-tailed binomial test.

Fractional change in neural-vector, angle and magnitude

We calculate the fractional difference from their average baseline values in the initial fraction of regression windows (Fig.4). The resulting traces are smoothed with a 10% window and significant difference between sessions of easy and hard rules is determined with bootstrapping – shuffling the easy/hard label 10,000 times and checking if the correct labeling surpasses the required confidence level (95%).

Optimal lags between time series

Given two time series, e.g. the angle-to-rule of a dACC neuron and the simultaneously-recorded vector-magnitude of a Putamen neuron, we find the shift that maximizes their Pearson correlation. Only pairs with significant correlation in the optimal lag contribute (as in Fig.4i,j).

Relating the neural-vector to next-day behavior

To examine if the learning-related change in the neural-vectors indicate a real shift in the monkeys' preferred policy (Fig.4h,i), we tested if the neurons' preferred feature combination (i.e. their neural-vector) predicts the monkeys' behavior early in the following day. For each neuron we averaged the neural-vector in the late fraction of the rolling regression windows. Then, as a measure of similarity, we calculated the projection of the neural-vector on the subsequent day's rule. In Fig.4h we calculate the Pearson correlations between these neural projections and the mean performance in the early fraction of the next day's session across the neural population. In Fig.4i we repeat the same calculation but only take cases in which the rule was changed between the current and next day.

References

1. Nosofsky, R.M., Kruschke, J.K. & McKinley, S.C. Combining exemplar-based category representations and connectionist learning rules. *Journal of experimental psychology. Learning, memory, and cognition* **18**, 211-233 (1992).

2. Goodman, N.D., Tenenbaum, J.B., Feldman, J. & Griffiths, T.L. A rational analysis of rule-based concept learning. *Cognitive Science* **32**, 108-154 (2008).

3. Shepard, R.N., Hovland, C.I. & Jenkins, H.M. Learning and memorization of classifications. *Psychological Monographs: General and Applied* **75**, 1–42 (1961).

4. Gluck, M.A., Shohamy, D. & Myers, C. How do People Solve the "Weather Prediction" Task?: Individual Variability in Strategies for Probabilistic Category Learning. *Learning & Memory* **9**, 408–418 (2002).

5. Lagnado, D.A., Newell, B.R., Kahan, S. & Shanks, D.R. Insight and strategy in multiple-cue learning. *Journal of Experimental Psychology. General* **135**, 162-183 (2006).

6. Meeter, M., Radics, G., Myers, C.E., Gluck, M.A. & Hopkins, R.O. Probabilistic categorization: how do normal participants and amnesic patients do it? *Neuroscience and Biobehavioral Reviews* **32**, 237-248 (2008).

7. Stuss, D.T.*, et al.* 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 (2000).

8. Speekenbrink, M., Lagnado, D.A., Wilkinson, L., Jahanshahi, M. & Shanks, D.R. Models of probabilistic category learning in Parkinson's disease: Strategy use and the effects of L-dopa. *Journal of Mathematical Psychology* **54**, 123-136 (2010).

9. Shohamy, D., Myers, C.E., Kalanithi, J. & Gluck, M.A. Basal ganglia and dopamine contributions to probabilistic category learning. *Neurosci Biobehav Rev* **32**, 219-236 (2008).

10. Feldman, J. Minimization of Boolean complexity in human concept learning. *Nature* **407**, 630-633 (2000).

11. Cohen, Y. & Schneidman, E. High-order feature-based mixture models of classification learning predict individual learning curves and enable personalized teaching. *Proc Natl Acad Sci U S A* **110**, 684-689 (2013).

12. Seger, C.A. & Miller, E.K. Category learning in the brain. *Annu Rev Neurosci* **33**, 203-219 (2010).

13. Balleine, B.W., Delgado, M.R. & Hikosaka, O. The role of the dorsal striatum in reward and decision-making. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* **27**, 8161-8165 (2007).

14. Cromer, J.A., Roy, J.E. & Miller, E.K. Representation of multiple, independent categories in the primate prefrontal cortex. *Neuron* **66**, 796-807 (2010).

15. Gold, J.I. & Shadlen, M.N. The neural basis of decision making. *Annu Rev Neurosci* **30**, 535-574 (2007).

16. Freedman, D.J. & Assad, J.A. Neuronal Mechanisms of Visual Categorization: An Abstract View on Decision Making. *Annu Rev Neurosci* **39**, 129-147 (2016).

17. Histed, M.H., Pasupathy, A. & Miller, E.K. Learning substrates in the primate prefrontal cortex and striatum: sustained activity related to successful actions. *Neuron* **63**, 244-253 (2009).

18. Freedman, D.J., Riesenhuber, M., Poggio, T. & Miller, E.K. A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* **23**, 5235-5246 (2003).

19. Muhammad, R., Wallis, J.D. & Miller, E.K. A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum. *J Cognitive Neurosci* **18**, 974-989 (2006).

20. Wallis, J.D., Anderson, K.C. & Miller, E.K. Single neurons in prefrontal cortex encode abstract rules. *Nature* **411**, 953-956 (2001).

21. Kim, J.N. & Shadlen, M.N. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat Neurosci* **2**, 176-185 (1999).

22. Heilbronner, S.R., Rodriguez-Romaguera, J., Quirk, G.J., Groenewegen, H.J. & Haber, S.N. Circuit-Based Corticostriatal Homologies Between Rat and Primate. *Biol Psychiatry* **80**, 509-521 (2016).

23. Averbeck, B.B., Lehman, J., Jacobson, M. & Haber, S.N. Estimates of projection overlap and zones of convergence within frontal-striatal circuits. *J Neurosci* **34**, 9497-9505 (2014).

24. Ongur, D. & Price, J.L. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb Cortex* **10**, 206-219 (2000).

25. Heilbronner, S.R. & Hayden, B.Y. Dorsal Anterior Cingulate Cortex: A Bottom-Up View. *Annual Review of Neuroscience, Vol 39* **39**, 149-170 (2016).

26. Saez, A., Rigotti, M., Ostojic, S., Fusi, S. & Salzman, C.D. Abstract Context Representations in Primate Amygdala and Prefrontal Cortex. *Neuron* **87**, 869-881 (2015).

27. Chudasama, Y.*, et al.* The Role of the Anterior Cingulate Cortex in Choices based on Reward Value and Reward Contingency. *Cereb Cortex* (2012).

28. Haroush, K. & Williams, Z.M. Neuronal prediction of opponent's behavior during cooperative social interchange in primates. *Cell* **160**, 1233-1245 (2015).

29. Seo, H. & Lee, D. Behavioral and neural changes after gains and losses of conditioned reinforcers. *J Neurosci* **29**, 3627-3641 (2009).

30. Rushworth, M.F. & Behrens, T.E. Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat Neurosci* **11**, 389-397 (2008).

31. Kolling, N., *et al.* Value, search, persistence and model updating in anterior cingulate cortex. *Nat Neurosci* **19**, 1280-1285 (2016).

32. Rudebeck, P.H., *et al.* Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J Neurosci* **28**, 13775-13785 (2008).

33. Wallis, J.D. & Kennerley, S.W. Contrasting reward signals in the orbitofrontal cortex and anterior cingulate cortex. *Ann N Y Acad Sci* **1239**, 33-42 (2011).

34. Hayden, B.Y. & Platt, M.L. Neurons in anterior cingulate cortex multiplex information about reward and action. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* **30**, 3339-3346 (2010).

35. Mansouri, F.A., Tanaka, K. & Buckley, M.J. Conflict-induced behavioural adjustment: a clue to the executive functions of the prefrontal cortex. *Nat Rev Neurosci* **10**, 141-152 (2009).

36. Seo, H. & Lee, D. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* **27**, 8366-8377 (2007).

37. Lee, D., Rushworth, M.F., Walton, M.E., Watanabe, M. & Sakagami, M. Functional specialization of the primate frontal cortex during decision making. *J Neurosci* **27**, 8170-8173 (2007).

38. Lau, B. & Glimcher, P.W. Value representations in the primate striatum during matching behavior. *Neuron* **58**, 451-463 (2008).

39. Graybiel, A.M. & Grafton, S.T. The striatum: where skills and habits meet. *Cold Spring Harb Perspect Biol* **7**, a021691 (2015).

40. Kim, H.F. & Hikosaka, O. Distinct basal ganglia circuits controlling behaviors guided by flexible and stable values. *Neuron* **79**, 1001-1010 (2013).

41. Jin, X. & Costa, R.M. Shaping action sequences in basal ganglia circuits. *Curr Opin Neurobiol* **33**, 188-196 (2015).

42. Seo, M., Lee, E. & Averbeck, B.B. Action selection and action value in frontal-striatal circuits. *Neuron* **74**, 947-960 (2012).

43. Averbeck, B.B. & Costa, V.D. Motivational neural circuits underlying reinforcement learning. *Nat Neurosci* **20**, 505-512 (2017).

44. Seger, C.A. How do the basal ganglia contribute to categorization? Their roles in generalization, response selection, and learning via feedback. *Neurosci Biobehav Rev* **32**, 265-278 (2008).

45. Merchant, H., Zainos, A., Hernandez, A., Salinas, E. & Romo, R. Functional properties of primate putamen neurons during the categorization of tactile stimuli. *J Neurophysiol* **77**, 1132-1154 (1997).

46. Williams, Z.M. & Eskandar, E.N. Selective enhancement of associative learning by microstimulation of the anterior caudate. *Nat Neurosci* **9**, 562-568 (2006).

47. Yang, T. & Shadlen, M.N. Probabilistic reasoning by neurons. *Nature* **447**, 1075-1080 (2007).

48. Padoa-Schioppa, C. & Assad, J.A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223-226 (2006).

49. Buch, E.R., Brasted, P.J. & Wise, S.P. Comparison of population activity in the dorsal premotor cortex and putamen during the learning of arbitrary visuomotor mappings. *Experimental Brain Research* **169**, 69-84 (2006).

50. Brasted, P.J. & Wise, S.P. Comparison of learning-related neuronal activity in the dorsal premotor cortex and striatum. *The European Journal of Neuroscience* **19**, 721-740 (2004).

51. Mitz, A.R., Godschalk, M. & Wise, S.P. Learning-dependent neuronal activity in the premotor cortex: activity during the acquisition of conditional motor associations. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* **11**, 1855-1872 (1991).

52. Buckley, M.J.*, et al.* Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions. *Science* **325**, 52-58 (2009).

53. Golub, M.D.*, et al.* Learning by neural reassociation. *Nat Neurosci* **21**, 607-616 (2018).

54. Chen, N.H., White, I.M. & Wise, S.P. Neuronal activity in dorsomedial frontal cortex and prefrontal cortex reflecting irrelevant stimulus dimensions. *Experimental Brain Research* **139**, 116-119 (2001).

55. Sadtler, P.T.*, et al.* Neural constraints on learning. *Nature* **512**, 423-426 (2014).

56. Genovesio, A., Brasted, P.J., Mitz, A.R. & Wise, S.P. Prefrontal cortex activity related to abstract response strategies. *Neuron* **47**, 307-320 (2005).

57. Minamimoto, T., Saunders, R.C. & Richmond, B.J. Monkeys Quickly Learn and Generalize Visual Categories without Lateral Prefrontal Cortex. *Neuron* **66**, 501-507 (2010).

58. Yanike, M. & Ferrera, V.P. Representation of outcome risk and action in the anterior caudate nucleus. *J Neurosci* **34**, 3279-3290 (2014).

59. Desrochers, T.M., Amemori, K. & Graybiel, A.M. Habit Learning by Naive Macaques Is Marked by Response Sharpening of Striatal Neurons Representing the Cost and Outcome of Acquired Action Sequences. *Neuron* **87**, 853-868 (2015).

60. Deffains, M.*, et al.* Subthalamic, not striatal, activity correlates with basal ganglia downstream activity in normal and parkinsonian monkeys. *Elife* **5** (2016).

61. Wimmer, G.E., Li, J.K., Gorgolewski, K.J. & Poldrack, R.A. Reward Learning over Weeks Versus Minutes Increases the Neural Representation of Value in the Human Brain. *J Neurosci* **38**, 7649-7666 (2018).

62. Doll, B.B., Shohamy, D. & Daw, N.D. Multiple memory systems as substrates for multiple decision systems. *Neurobiol Learn Mem* **117**, 4-13 (2015).

63. Averbeck, B.B. & Chafee, M.V. Using model systems to understand errant plasticity mechanisms in psychiatric disorders. *Nat Neurosci* **19**, 1418-1425 (2016).

64. Hyman, S.E., Malenka, R.C. & Nestler, E.J. Neural mechanisms of addiction: the role of reward-related learning and memory. *Annu Rev Neurosci* **29**, 565-598 (2006).

65. Salzman, C.D. & Fusi, S. Emotion, cognition, and mental state representation in amygdala and prefrontal cortex. *Annu Rev Neurosci* **33**, 173-202 (2010).

66. Lee, D. Decision making: from neuroscience to psychiatry. *Neuron* **78**, 233-248 (2013).

A geometric representation unveils learning dynamics in primate neurons

Yarden Cohen, Elad Schneidman*, Rony Paz*

Dept. of Neurobiology, Weizmann Institute of Science, Rehovot ISRAEL 76100

* These authors contributed equally

Supplementary Figures

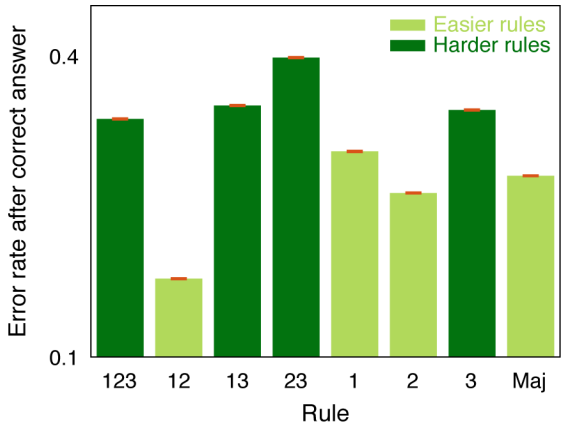

Fig.S1. Error probability after making a correct classification. For each rule (x-axis) we plot the probability of making an error (y-axis) after choosing correctly in the previous presentation of the same pattern. Error bars depict SE (red). The relatively low probability and the strong dependence on the specific rule suggests the monkeys did not memorize specific patterns after a successful classification. Kruskal-Wallis test across all repetitions and patterns validates the effect of the rule on the occurrence of errors after correct classifications ($\chi^2(7)=88.97$, $p < 1e\text{-}15$). Pairwise binomial z-test shows that all the error rates in easier rules are smaller than error rates in harder rules.

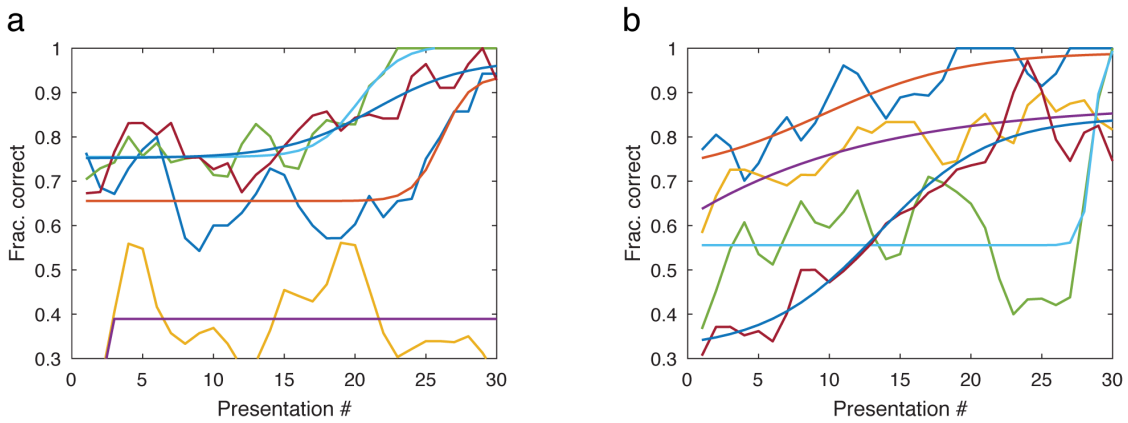

Fig.S2. Acquisition rates for the salient pattern. The most salient pattern, '000', was labeled 'right' and 'left' equally in our set of rules (4+4 rules). Shown are the learning curves (colored lines + sigmoidal fits), the proportion of correct labeling (y-axis, smoothed across days with a running window of 4 presentations) along the pattern presentations (x-axis). The two panels are for the 4 rules in which 000 is labeled 'left' (**a**) and for the 4 it was labeled 'right' (**b**). There was variability across rules in how this most visually salient pattern was learned, showing that patterns that draw more attention are not learned faster or similarly across different rules, and therefore suggesting against simple memorization process. Kruskal-Wallis test across all segments of 10 pattern presentations validates the effect of the rule on the performance in panels a,b. ($p < 1e-4$ in all segments).

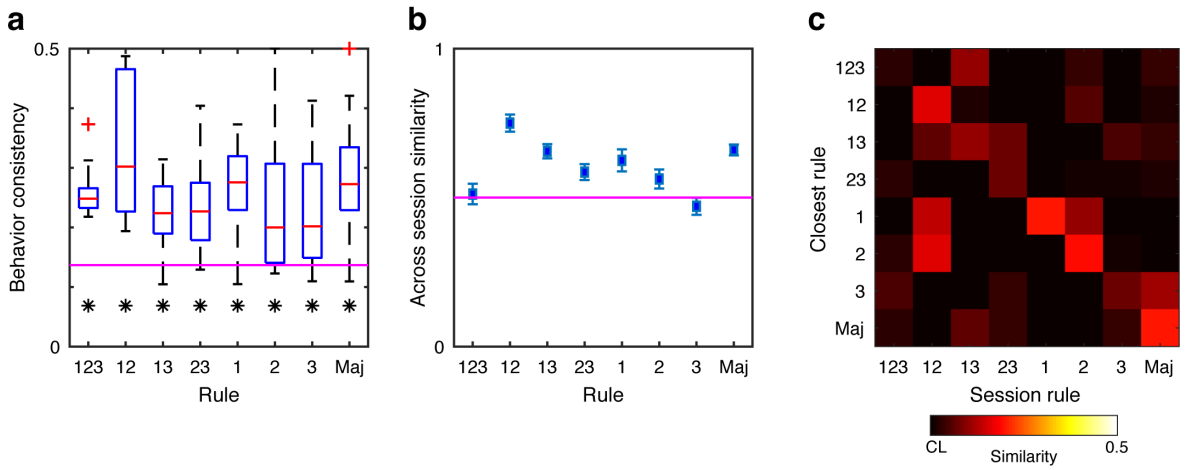

Fig.S3. Behavior consistency. a. The within-session consistency scores (y-axis, methods) for all sessions with a single rule (x-axis). The chance level is marked by the magenta line. Statistical significance (t-test, $p<0.05$) of above chance mean is marked by asterisks. **b.** The between-session pairwise consistency score (y-axis, SE in error bars) across sessions with a single rule (x-axis). The chance level is marked by the magenta line. **c.** The classifier, fitted to the last ¼ of each session in which a certain rule was taught (x-axis), is compared to the set of rules (y-axis, methods). The color scale is from the chance level (CL) to the maximal possible value (0.5). Please see methods for full description of the measures

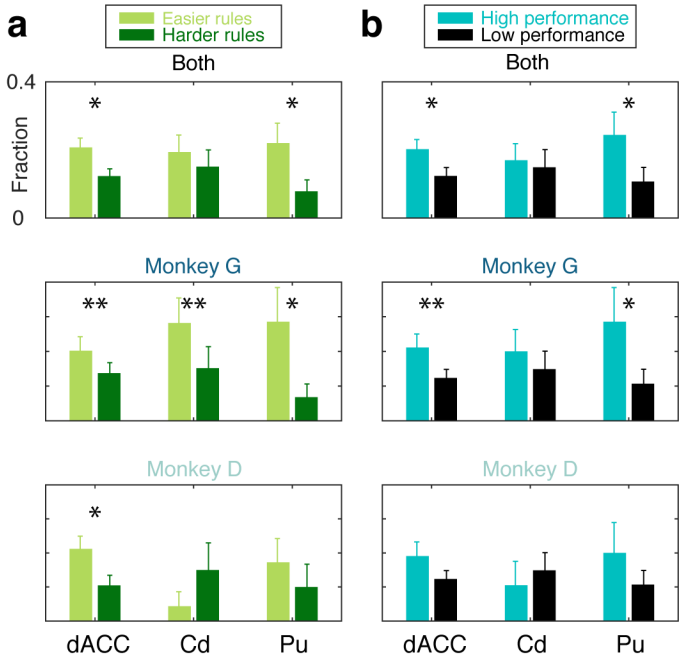

Fig.S4. Robustness of rule-representation in single-units. Similar presentation to Fig.2d: The fractions of neurons that exhibited significant rule correlation during the last third of the session (bars, y-axis. SE in error-bars) are plotted for the 3 regions (x-axis). Apart from grouping both monkeys (top) we also separate to neurons from monkey G (middle) and D (bottom). **a.** Separating sessions by the rule-group. Easy rules in light green and hard rules in dark green. **b.** Separating sessions by performance in each session separately (i.e. independent of the rule). Significant differences by binomial test (*=p<0.05, **=p<0.1).

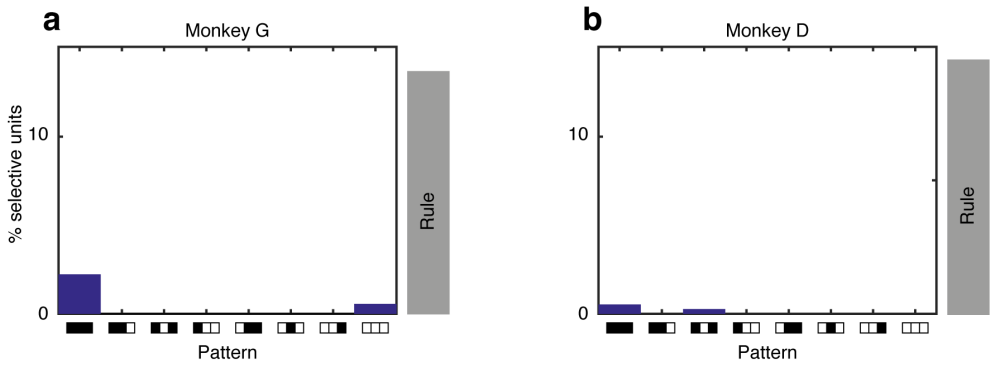

Fig.S5. Pattern selective neurons. The percent of neurons that were selective for a single pattern (bars, y-axis) is plotted for all patterns (x-axis). **a.** monkey G. **b.** monkey D. Gray bars indicate the percent of rule correlated neurons.

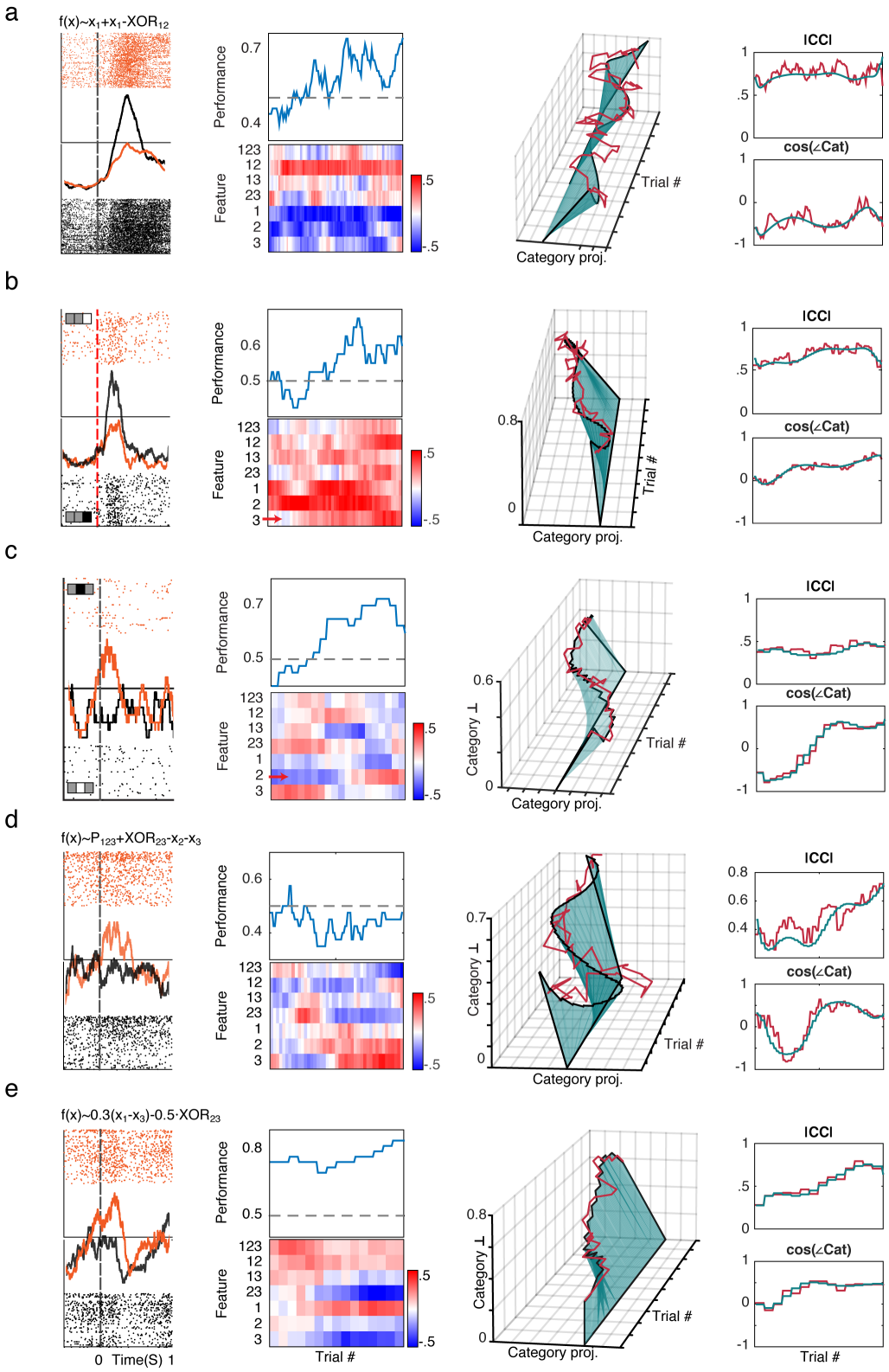

Fig.S6. Examples of learning related neural dynamics. Same format as in Fig.3, showing neurons with stable feature selectivity, high dimensional rotation, magnitude stretching, and complex trajectories.

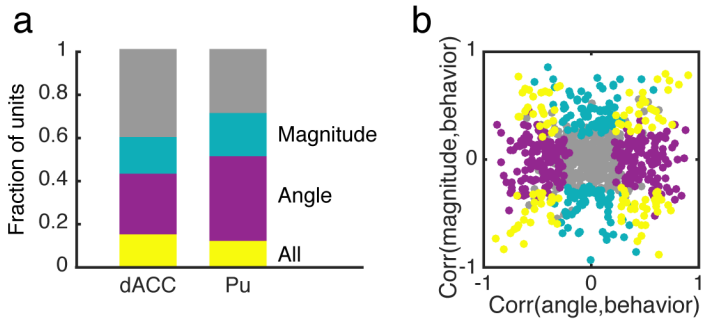

Fig.S7. Significant correlations between geometric neural correlates and learning curves partially overlap. a. Enumeration of the significant (Pearson, $p<0.05$) correlations to all variable combinations in the different regions (x-axis). The colored bars show the fraction of units with significant correlations between the learning curve and both angle-to-rule and correlations vector-magnitudes (yellow), only angle-to-rule (purple), or only vector-magnitudes (turquoise). **b.** For each neuron we compare the magnitude correlation (y-axis) to the angle-to-rule correlation (x-axis). Color-coding as in panel (a).

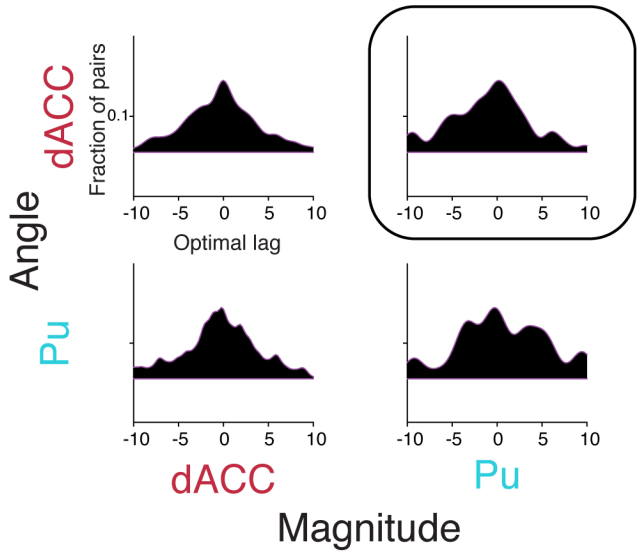

Fig.S8. Optimal lags between neural-vector properties (magnitude and angle) between dACC and Striatal pairs.

For all simultaneously recorded pairs, we computed the optimal lag between the vector-magnitude and the angle-to-rule, for all four possible combinations. Only lags between vector-magnitude in Putamen neurons and angle-to-rule in dACC neurons were significantly different than zero (top-right, also shown in main Fig.4i,j) with the Putamen magnitude following dACC rotation ($p<0.004$, t-test; Mean lag -0.7 ± 0.24 regression steps; all other comparisons were not different than zero, $p>0.1$ for all, t-tests).