Differences in the emergent coding properties of cortical and striatal ensembles

Liya Ma, James M Hyman, Adrian J Lindsay, Anthony G Phillips & Jeremy K Seamans

The function of a given brain region is often defined by the coding properties of its individual neurons, yet how this information is combined at the ensemble level is an equally important consideration. We recorded multiple neurons from the anterior cingulate cortex (ACC) and the dorsal striatum (DS) simultaneously as rats performed different sequences of the same three actions. Sequence and lever decoding was markedly similar on a per-neuron basis in the two regions. At the ensemble level, sequence-specific representations in the DS appeared synchronously, but transiently, along with the representation of lever location, whereas these two streams of information appeared independently and asynchronously in the ACC. As a result, the ACC achieved superior ensemble decoding accuracy overall. Thus, the manner in which information was combined across neurons in an ensemble determined the functional separation of the ACC and DS on this task.

A fundamental goal in neuroscience is to determine the specific function and information processing capabilities of anatomically distinct brain regions. A common approach to this problem is to infer function on the basis of the sensory, motor or cognitive events that evoke neural responses in a given region. Although this approach has been invaluable in advancing our understanding of sensory areas, it can be problematic for regions such as the frontal cortex, which is comprised of neurons that are highly multi-modal. Furthermore, single neuron correlates in frontal regions are often very similar to those found in its efferent targets, such as the dorsal striatum. Perhaps in this case, further insight could be gained by considering how neurons collectively respond rather than what they respond to.

How a neuron responds is determined by many factors, including its biophysical properties, the input it receives and the structure of the local microcircuit. In the neocortex, pyramidal neurons form massive, interconnected networks with interneurons exerting relatively weak inhibition. This arrangement favors recurrent excitation and, with it, persistent activity patterns through time. Overall, cortical ensemble patterns are usually well-balanced, with equal numbers of neurons slightly increasing or decreasing their response to any given event. By contrast, local regions of the striatum receive strong excitatory inputs from the cortex, which generate responses that are temporally restricted by powerful and widespread inhibition produced by local interneurons. This arrangement favors recurrent excitation and, with it, persistent activity patterns through time. Overall, cortical ensemble patterns are usually well-balanced, with equal numbers of neurons slightly increasing or decreasing their response to any given event. By contrast, local regions of the striatum receive strong excitatory inputs from the cortex, which generate responses that are temporally restricted by powerful and widespread inhibition produced by local interneurons. Accordingly, when recorded from behaving animals, striatal neurons tend to be activated transiently and synchronously. This synchronized activity may be an asset in the context of movement generation and learning, but may also be a detriment if it concurrently amplifies variability or ‘noise’ in the neural responses over time. By understanding how these differences might influence the way in which neural ensembles in the frontal cortex versus striatum encode the same information, we hope to gain new insight into their unique functions of these two interrelated brain regions.

To address this issue, we used bundles of tetrodes to simultaneously record from neurons in the ACC and the central portion of the DS while rats performed a sequential action task. Although there is already a rich literature on the responses of individual frontal cortex and striatal neurons on such tasks, our aim was to better understand how coherent ensemble representations emerged from the activity of single neurons. The task required rats to press three levers in different temporal orders (termed sequence blocks) to receive reward. We found that, although single neurons in the ACC and DS represented information about sequences and lever presses with similar overall accuracy, the unique way this information was combined across the neurons through time in the ACC yielded far superior overall ensemble sequence decoding accuracy than we observed in the DS.

RESULTS

Behavior
In our task, different sequences of actions (sequence blocks) consisted of presses on the same physical levers in different temporal orders (Supplementary Fig. 1a,b). Each lever was distinguished by specific cues temporarily affixed to the area immediately surrounding the levers. The sequence of cues was always the same for a given rat, but the cues were moved to different lever locations for each of the two or three sequence blocks. This task design permitted an examination of the manner in which the two regions encoded objective information about discrete lever press actions versus more abstract information about sequences of actions. The rats were required to perform three different sequences of operant responses in a single session, and did so with a high degree of accuracy (percentage correct response: 89.2 ± 5.4%, mean ± s.d.). Overall behavioral performance across different sequences in all sessions did not differ (one-way ANOVA, F2,48 = 0.22, P = 0.80). Across all trials in all sequence blocks, the
Figure 1 Behavior was similar for common segment lever presses in different sequence blocks. (a) Values of the behavioral variables for matched pairs of common-segments lever presses. Common segments involved presses on the same two levers occurring as part of two different sequence blocks (black versus gray). Each column represents a different physical lever, whereas the four rows display the animal’s x coordinate, y coordinate, lever approach angle and lever approach velocity in the six bins surrounding each lever press (the lever press occurred at time 0 in each panel). The two lines in each panel correspond to the value of each of these variables in one of the two sequence blocks. During these intervals, the animals’ location and movement were highly similar across sequence blocks, but were quite different across the three physical levers. Error bars indicate ±SEM. (b) Distributions of the percentage of firing rate variance accounted for when the behavioral variables shown in a were used as factors in a multiple linear regression analysis, performed on all ACC neurons (top) or all DS neurons (bottom). The residual matrices represent the iFR values that remained after the effect of these behavioral variables has been removed. These matrices are used in the analyses shown in subsequent figures.

latencies from the first to the second lever, and those from the second to the third lever, were also equivalent (unpaired t test, t1102 = 0.89, P = 0.38).

We compared neural activity associated with presses on the same physical lever in different sequence blocks. Although we focused on presses on the identical physical lever, the lever may be approached at a different body angle, along a different trajectory or at a different velocity in each sequence block. This could be problematic because differences in movements or movement paths leading up to an action can affect ACC activity. In an attempt to minimize the behavioral variability across sequence blocks, we focused on ‘common segment’ elements. Common segments consist of presses on the same lever when approached from the same preceding lever in two sequence blocks. The average x and y trajectories, the angle of approach to the lever, and the approach velocity were similar for common segment lever presses performed in different sequence blocks (Fig. 1a). None of the behavioral measurements differed significantly across sequence blocks (repeated-measures ANOVA, no effect of sequence for x position: F1,78 = 1.24, P = 0.27; y position: F1,78 = 3.37, P = 0.07; approach angle: F1,78 = 0.0606, P = 0.94; approach velocity: F1,78 = 0.055, P = 0.82; no interaction between time and sequence: x position: F5,390 = 0.061, P = 1; y position: F5,390 = 0.49, P = 0.79; approach angle: F5,390 = 1.91, P = 0.093; approach velocity: F5,390 = 1.18, P = 0.32).

Although similar, they were not completely overlapping, and to determine whether this remaining behavioral variability may have affected the neuronal responses, we used certain behavioral variables (Fig. 1a) as factors in a multilinear regression model performed individually on each neuron (Online Methods). The analysis revealed that the total amount of variability in firing rate that could be accounted for by the five factors was, on average, 3.04% for ACC neurons and 3.39% for DS neurons. In fact, in only 3.99% of the ACC neurons and 5.64% of DS neurons did the behavioral variables collectively account for more than 10% of firing rate variance during the common segment lever press periods (Fig. 1b). Among these neurons, the percentages of variance accounted for by the model did not differ (unpaired t test, t59 = -1.21, P = 0.23). Although this percentage would likely be larger if all task periods had been considered, at least for the common segment lever press periods, any effect of the differences in behavior across sequence blocks was relatively small for most individual neurons.

Single neuron correlates of sequence differentiation

We assessed the selectivity of single ACC and DS neurons for sequences using a signal detection approach by calculating each neuron a selectivity index (SI) for each neuron (Online Methods). The SI for sequence was calculated by comparing the firing rates in the six bins surrounding common segment lever presses performed in different sequence blocks. Even though the effects of behavioral variability on instantaneous firing rates (iFRs) were small for the common segment periods (Fig. 1b), we nevertheless tested whether they could affect SI-based sequence discrimination. We re-calculated the SI values using the residual firing rate matrix generated from the multilinear regression analysis described above, as, in the residual matrix, the effect of these variables had theoretically been regressed off.

Across all neurons recorded (NACC = 637, NDS = 351), 24.0% of individual ACC neurons and 24.0% of individual DS neurons were selective for sequence (that is, had SI values ≥ 0.5), a difference that was not statistically significant (two-way ANOVA, no effect of region, F1,60 = 0.11, P = 0.74; unequal N HSD test, ACC versus DS, P = 1; Fig. 2a). When the calculation of SIs was repeated using the residual matrices, 24.1% of ACC neurons and 25.2% of DS neurons were found to be sequence selective. The differences in the number of sequence-selective neurons detected using the full versus residual matrices were not significant for the ACC (two-way ANOVA, no effect of regression, F1,60 = 0.17, P = 0.69; unequal N HSD test, ACC full versus residual matrices, P = 1) or the DS (unequal N HSD test, P = 0.96). To test the strength of the sequence signals, we compared the SIs of all the neurons possessing SI values ≥ 0.5. The average SIs of these putatively sequence-selective neurons did not differ between the two regions (ACC = 0.79, DS = 0.73, two-way ANOVA, F1,916 = 1.34, P = 0.25; unequal N HSD test, ACC versus DS, P = 0.34; Fig. 2b).

Furthermore, the mean SI values of these neurons were similar if they were calculated using the full or residual matrices (unequal N HSD test, original versus residuals for ACC, P = 1; DS, P = 0.24; Fig. 2b). These results implied that the ACC and DS had notably similar numbers of neurons that were equally selective for sequence differentiation (Fig. 2c–f).

Sequence information by neurons versus ensembles

Next, we asked how single neurons compared with ensembles in terms of sequence differentiation and whether this relationship varied between the two regions. To compare the sequence decoding properties of single neurons versus ensembles on an equal footing, we employed a modified receiver operator characteristic (ROC) analysis approach. The ROC analysis has an advantage in situations in which the distributions are...
unknown because it assesses performance over a range of threshold values rather than being forced to evaluate differences at a single threshold level relative to some theoretical distribution. We investigated the signal detection characteristics of individual ACC and DS neurons by a sequence-specific template creating for each neuron for the 1.4-s period leading up to and including a common segment lever press in half of the trials in one sequence block. As the data were binned at 200 ms, the template was a vector of seven iFR values for each single neuron. The template was then moved through both sequence blocks and bin by bin correlations were calculated. This process was repeated using templates created from different groups of trials and the results were averaged. A ‘true positive’ or ‘hit’ occurred when a correlation score larger than the threshold value was found in the lever press interval of the remaining half of trials of the sequence block from which the original template was constructed. A ‘false alarm’ occurred when a correlation score larger than the threshold value was found between the template and the common segment lever presses performed in the alternate sequence block. If hits and false alarms occur at the same rates at all thresholds, the ROC curve would be a straight line with a 45° slope, with an area under the curve (AUC) of 0.5.

To compare ensembles to single neurons using this approach, we performed a similar analysis, except that, rather than the vectors being seven iFR values, they consisted of N × 7 iFR values, with N being the number of recorded neurons per session. Because the templates were much larger than those of any individual neuron, one would expect a low probability of recurrence of a similar ensemble activity state pattern, thereby reducing false alarm rates. Concurrently, the larger templates would be equally disadvantageous, as it is equally unlikely that a match between the template and the common segment lever presses performed in the alternate sequence block. If hits and false alarms occur at the same rates at all thresholds, the ROC curve would be a straight line with a 45° slope, with an area under the curve (AUC) of 0.5.

For sequence decoding (Fig. 3a), the hit: false alarm ratios for ACC ensembles were higher than 96.77% of individual neurons during the common segment periods of the two sequence blocks. In contrast, only 9.12% of individual DS neurons (Fig. 3b) were superior to the DS ensembles for the same periods (Fig. 3b). Overall, single DS neurons were superior to DS ensembles significantly more often than was the case for single ACC neuron compared with ACC ensembles (Pearson’s Chi-square, \( \chi^2 = 12.1, df = 1, P = 0.0005 \)). It is clear that the reason for this difference is not that DS neurons were superior to ACC neurons on a single neuron basis—in fact they were not (independent-sample t test, \( t_{941} = 1.59, P = 0.11 \))—but rather, when DS neurons were combined into ensembles, their performance was significantly worse than when ACC neurons were combined into ensembles (independent-sample t test, \( t_{13} = 4.44, P = 1.1 \times 10^{-4} \)) (Fig. 3a,b).

To examine the relative superiority of ACC ensembles for sequence decoding in greater detail, we randomly selected neurons from the population of 637 ACC and 351 DS neurons to create ensembles of different sizes and then we performed the ROC analysis on each ensemble (Fig. 3c). This analysis revealed that the ACC ensembles achieved superior signal detection across all ensemble sizes. In addition, the steeper ACC curve, in comparison with the DS curve, indicated that signal detection improved progressively as more ACC neurons were added. Based on the functions fit to the data, we extrapolate that, on average, a randomly drawn ensemble of 112 neurons would be required for the DS to achieve the same level of sequence-signal detection as an ensemble of 19 neurons drawn randomly from the ACC. This is quite notable when one considers that individual neurons in the two areas performed equally on a per neuron basis (Fig. 2a–c). Paradoxically, the DS is at a disadvantage when its neurons are combined into ensembles for sequence decoding.

**Unique activity states represent sequence information**

To understand how ACC ensembles achieved better ROC performance, we probed more deeply into the nature of the ensemble codes themselves. We generated three-dimensional representations of the multiple single unit activity (MSUA) spaces from representative ACC and DS ensembles (Fig. 4a,b). Ensemble activity states during lever presses in one sequence block were tightly clustered, but shifted to another region of the MSUA space when the same lever was pressed as part of a different sequence block. A shift in the MSUA space
sizes were randomly drawn from ACC and DS neuronal populations, and the AUCs were calculated from the sequence signal-detection ROCs. This decoding performance improved with increasing ensemble size in both the ACC and DS, but at a much higher rate in the ACC. Ensembles of different sizes were randomly drawn from ACC and DS neuronal populations, and the AUCs were calculated from the sequence signal-detection ROCs. This process was repeated 100 times at each ensemble size for each region, and the mean and s.e.m. are plotted (ACC: black circles, DS: gray circles). The best fitting trend lines were power functions (ACC: top, DS: bottom), which explained more than 99% of the variance.

To quantify the differences in firing patterns associated with different sequence blocks, we calculated the Mahalanobis distance ($D_{\text{Mah}}$) between population vectors in the MSUA spaces. The $D_{\text{Mah}}$ between the activity states associated with common-segment lever presses in different sequence blocks was significantly larger than the $D_{\text{Mah}}$ between shuffled control blocks (Kruskal-Wallis test, $\chi^2 = 96.34, P = 9.5 \times 10^{-20}$) for both the ACC (Tukey's test, $P = 0$) and the DS (Tukey's test, $P = 1.0 \times 10^{-6}$; Fig. 4c). Once again, the ACC was superior, as the $D_{\text{Mah}}$ between the neural patterns associated with presses on the same lever occurring in different sequence blocks was significantly larger for ACC ensembles than for DS ensembles (Tukey's test, $P = 0.023$).

To show that these activity state patterns are functionally important for sequence decoding, we applied a leave-one-out variant of the $D_{\text{Mah}}$ analysis to the common-segment lever press periods. Using this form of Mahalanobis discriminant analysis (MDA), the correct sequence-block could be accurately predicted in 66.6% of cases for ACC ensembles and in 61.9% of cases for DS ensembles, both of which were significantly better than when the procedure was repeated using shuffled sequence-block assignments (two-way ANOVA, main effect of sequence, $F_{1,140} = 204.89, P = 1.8 \times 10^{-22}$; main effect of region, $F_{1,140} = 20.53, P = 1.3 \times 10^{-5}$; Fig. 4d). Notably, ACC ensembles classified each specific sequence with higher accuracy than DS ensembles (Tukey's test, $P = 0.00022$).

Regional differences in ensemble variance and covariance

The sequence classification and decoding measures used above weigh differences between the patterns (that is, the signal) relative to the variance and covariance in the ensembles across time. Thus, ACC ensembles could be superior to DS ensembles either in terms of the strength of their signals or because they exhibited less variance and covariance through time. To disambiguate these two possibilities, we examined each separately.

To assess potential difference in signal strength, we calculated the Euclidean distance ($D_{\text{Euc}}$) rather than the $D_{\text{Mah}}$ between points in the MSUA spaces associated with each sequence block. $D_{\text{Euc}}$ and $D_{\text{Mah}}$ both measure the distance between the sets of points in the MSUA space, but differ in that $D_{\text{Euc}}$ is a simple measure of the geometric distances between the centers of clusters, whereas $D_{\text{Mah}}$ weighs these distances relative to the individual variances and pooled covariance of the two sets of points. Thus, $D_{\text{Euc}}$ represents a pure measure of the ensemble signal independent of variance and covariance.

Unlike $D_{\text{Mah}}$, $D_{\text{Euc}}$ between common segments in different sequence blocks did not differ between the ACC and the DS ensembles (independent-sample t-test, $t_{70} = -1.09, P = 0.28$), suggesting that the difference between the ACC and DS was not a result of differences in the sequence ‘signal’. This was not unexpected given the similarities in sequence differentiation by single ACC and DS neurons described above.

The higher degree of covariance in the responses of DS neurons to the behavioral variables contained more coordinated variation throughout a given action than did ACC neurons.

The higher degree of covariance in the responses of DS neurons to the behavioral variables may have contributed to the poorer performance of DS ensembles. To test this possibility, we repeated the sequence discrimination analysis described above using the residual matrices. This indeed improved the performance of DS ensembles, as it led to a significant increase in the $D_{\text{Mah}}$ between common segment lever presses (repeated-measures ANOVA and post hoc Tukey's test: residuals>full matrix: $P = 0.00039$; Fig. 5c) and improved sequence decoding accuracy using MDA (repeated-measures ANOVA and post hoc Tukey's test: residuals>full matrix: $P = 0.015$; Fig. 5d). In contrast, repeating the same analysis using the residual ACC matrices had no effect on the ability of ACC ensembles to separate common segments lever presses in different sequence blocks (repeated-measures ANOVA and post hoc Tukey's test: residuals = full matrix: $P = 0.82$; Fig. 5e) or on MDA-based decoding accuracy (repeated-measures ANOVA and post hoc Tukey's test: residuals = full matrix: $P = 0.99$; Fig. 5f).
Figure 4  Sequence information is represented as differences in ensemble activity state patterns in both the ACC and DS. (a) Left, an example MSUA space constructed from the iFRs of all 34 ACC neurons recorded during a single session, visualized in three dimensions using multidimensional scaling. Each dot represents a population vector containing the activities of the entire ensemble during right-lever presses. Dots are colored black if the activities are associated with one sequence block, gray if associated with the alternate sequence block. Right, the average or prototypical activity of ACC neurons recorded during right lever presses in the two sequence blocks. Each bar represents the average normalized iFR for each neuron across all right lever presses in sequence block A (top, black bars) or sequence block B (bottom, gray bars). Data are presented as mean ± s.e.m. (b) Left, an example MSUA space constructed from the iFRs of all 20 DS neurons recorded simultaneously during the same session as in a. Right, the prototypical activity of DS neurons recorded during right lever presses in the two sequence blocks. (c) When common-segment lever press periods were examined, the average separation (DMah in MSUA space) between lever presses performed during different sequence blocks (black bars) were significantly larger than shuffled control blocks (white bars). (d) A leave-one-out prediction procedure based on DMah was used to classify trials as belonging to one of two sequence blocks. Classification accuracy for the actual data (black bars) was significantly better than for trials shuffled data (white bars) for both the ACC and DS, although ACC ensembles performed significantly better than DS ensembles. Error bars in b–d indicate s.e.m. *P < 0.05, **P < 0.001,***P < 0.00001. Figure 5 Consistency of sequence encoding in the ACC and DS within behavioral epochs. (a) Ensemble variance calculated on the basis of activities during the time bins in each pair of common-segment lever presses was higher in the DS than in the ACC. Ensemble variance calculated on the residual iFR matrices (right, white) was smaller than the variance calculated from full iFR matrices (right, black) in the DS, but the two were equivalent in the ACC (left bars). (b) Ensemble covariance was also greater in the DS (right) than in the ACC (left). Ensemble covariance calculated on the residual iFR matrices (white) was smaller than the covariance calculated from full iFR matrices (black) in the DS, but the two were equivalent in the ACC. (c) In DS ensembles, DMah calculated using the residual iFR matrices from the initial behavioral regression (right) was greater than that calculated using the full iFR matrices (left), although both the full iFRs and the residuals still contained robust sequence information. (d) In DS ensembles, sequence classification using the residual iFR matrices (right) was more accurate than when calculated using the full iFR matrices (left). (e) In ACC ensembles, between-sequence DMah calculated using the residuals from the initial behavioral regression was equivalent to that calculated using the full iFRs, and both were greater than shuffled controls. (f) In ACC ensembles, between-sequence leave-one-out MDA using the residual iFR matrices (right) was equivalent to that calculated using the full iFR matrices (left). Error bars indicate s.e.m. *P < 0.05, **P < 0.001, ***P < 0.00001.

Thus, the firing rate variance of DS neurons associated with the measures of behavioral variability (Fig. 1) contributed to its inferior sequence decoding at the ensemble level. This makes intuitive sense given that the levers being compared in the two sequence blocks were identical; thus, the stronger the neurons encoded this commonality, the less likely they would be able to differentiate between the two sequence blocks.

Regional differences in lever press action encoding

The previous analysis suggested that, relative to the ACC, DS ensembles more strongly tracked the commonalities associated with pressing the same physical levers in different sequence blocks. To test whether the spatial lever decoding ability, we again employed the same type of ROC-based analysis described above. In this case, a template was created by randomly selecting half of all presses on a given physical lever. A hit occurred when a correlation score larger than a threshold was found in the remaining half responses on the same lever, whereas a false alarm occurred when a correlation score larger than the threshold value was found between the template and the presses performed on a different lever. This procedure was performed both on single units and on ensembles as outlined above for the case of sequence decoding.

In support of the prediction, the decoding of spatial lever identity by DS neurons was indeed superior to that of ACC neurons on a single-neuron basis (two-way ANOVA, main effect of region, F1,194 = 29.78, P = 5.3 × 10−6). In spite of this, ACC ensembles nevertheless matched DS ensembles in lever-decoding ROC performance (two-way ANOVA, no effect of region, F1,93 = 0.70, P = 0.40). Likewise, the DMah between population vectors associated with responses on the three different levers were similar between regions (two-way ANOVA, no effect of region, F1,194 = 1.08, P = 0.30; both regions discriminated among the three levers significantly better than shuffled controls, two-way ANOVA, F1,194 = 497.3, P = 2.0 × 10−55). Similar results were also obtained using MDA (two-way ANOVA, F1,194 = 770.4, P = 1.8 × 10−69, post hoc Tukey’s test, DS = ACC, P = 0.63). Thus, even though DS neurons were indeed better than ACC neurons in terms of their ability to differentiate unique lever press actions, this advantage did not translate to the ensemble level. It would appear that on this task, the signals carried by individual neurons tended to be less synergistic when combined in the DS relative to the ACC.
Figure 6 Comparison of the timing of maximal sequence and lever differentiation in ACC and DS ensembles. (a) The strength of sequence discrimination (based on rolling \( F \) statistic) fluctuated across bins in the DS population (right bars), where the largest proportion of neurons exhibited maximal sequence specific firing differentiation around 0.6 s before lever press, whereas the smallest proportion exhibited sequence specific firing differentiation at around the actual lever press response itself. Such was not the case in the ACC (left bars), where an equal percentage of neurons exhibited maximal sequence-specific firing differentiation in each of the six bins. (b) The strength of lever discrimination (based on rolling \( F \) statistic) fluctuated across bins in the DS (right bars), where the largest proportion of neurons exhibited maximal lever differentiation 0.6 s before lever press. Such was not the case in the ACC (left bars), where an equal percentage of neurons exhibited maximal lever-specific firing differentiation in each of the six bins. (c) Average level of activity across time bins leading up to all lever presses in all sequence blocks and sessions in all ACC neurons (left) and all DS neurons (right). (d) An example of a DS neuron displaying maximal differentiation between sequences at approximately 0.6 s before presses on the same lever. (e) An example of a DS neuron reaching maximal differentiation between levers at approximately 0.6 s before lever presses in the same sequence block. Shading in (c–e) indicates s.e.m. * \( P < 0.01 \), ** \( P < 0.001 \), *** \( P < 0.0001 \).

ACC versus DS in the timing of sequence and lever signals

What is it about the DS that results in this type of counteractive ensemble effect? One factor may relate to the timing of different signals. To explore this issue, we used a rolling \( F \) statistic to find the proportions of neurons that exhibited their maximal firing differentiation of sequence versus lever identity in each trial. The proportions of ACC neurons differentiating sequences (Pearson’s Chi-square, \( \chi^2 = 9.28 \), df = 5, \( P = 0.098 \); Fig. 6a) or lever presses (Pearson’s Chi-square, \( \chi^2 = 4.79 \), df = 5, \( P = 0.44 \); Fig. 6b) was equal for all time bins in a trial. Furthermore, there was no relationship between the bins in which the maximal differentiation among levers and the bins in which the maximal discrimination between sequences occurred in a trial (Spearman’s rho, \( r = 0.024 \), \( P = 0.36 \)). In other words, the sequence and lever press signals evolved independently in the ACC in each trial.

In marked contrast, the proportions of DS neurons maximally differentiating between sequences (Pearson’s Chi-square, \( \chi^2 = 21.62 \), df = 5, \( P = 0.00062 \); Fig. 6a) or between levers (Pearson’s Chi-square, \( \chi^2 = 17.98 \), df = 5, \( P = 0.0030 \); Fig. 6b) were highly non-uniform across the bins in each trial. Notably, there was a small, but significant, correlation between the bins in which the greatest proportions of DS neurons maximally differentiated sequences and lever press actions (Spearman’s rho, \( r = 0.097 \), \( P = 0.013 \)). Specifically, the maximal point of differentiation for both sequences and lever press actions occurred concurrently at ~600 ms before the actions in the DS (Fig. 6a,b). This was also close to the point at which generalized activity was also highest (Fig. 6c). This effect can be seen in two examples of DS neuron responses (Fig. 6d,e).

Collectively, these results can be summarized as follows. To make a correct choice on this task, a rat must maintain a representation of the overall sequence block as well as the actual physical lever to be pressed. DS neurons collectively exhibited large firing rate fluctuations in response to the specific lever press actions, and this coherent fluctuation appeared at the same point in each trial in which the ensembles maximally differentiated sequence blocks. Because the two streams of information were at odds during common segment lever presses of this task, the decoding accuracy for each stream suffered. In contrast, ACC ensembles excelled, not because they had more neurons selectively coding the correct action or sequence block, but because the two streams of information remained independent across neurons through time.

DISCUSSION

Previous studies have observed robust responses of single neurons in the frontal cortex and striatum during actions\(^\text{16–22,24,27,34,35} \). We focused on how information about sequential actions carried by single neurons is combined into ensemble codes and whether this differs in the ACC versus the DS. We found that the key factor that separated the two regions was the manner in which information was combined across the population through time rather than specific responses of single neurons.

At the single neuron level, ACC and DS neurons consistently performed similarly across various measures including single neuron sequence selectivity and single neuron ROC performance. However, DS neurons performed more poorly on almost all ensemble-based measures of sequence differentiation, including the separation of the sequence activity state patterns based on \( D_{\text{Mah}} \), MDA and ROC performance. The one exception was the separation of the activity state clusters using \( D_{\text{Euc}} \). This was the only measure of pure sequence differentiation (that is, signal) that was completely independent of variance and covariance across time. DS neurons tended to exhibit larger variance and covariance and respond more as a collective. The point at which the largest proportion of DS neurons tended to maximally differentiate sequence identity was also the point at which most neurons differentiated lever identity, and this coincided with the point of maximal overall activity. Given that, in our task, the sequences were different, yet the actual physical levers pressed were identical across the two sequence blocks, if sequence and lever signals emerged simultaneously, but transiently, across the ensemble, the decoding of either would suffer. By contrast, in ACC ensembles, information about the physical lever was functionally independent from the information about the overall sequence (Fig. 5e,f) and, as a result, sequence and lever differentiation in the ACC remained high across all six bins rather than being concurrently maximal in just one or two. It is also important to emphasize that this uniformity through time was not a result of persistent firing in individual ACC neurons that has been well characterized in the frontal cortex. Rather, it resulted from the fact the neurons had mutually independent time courses and were thereby able to tile all six time bins as a population. The asynchrony in the ACC population allowed multiple conflicting sources of information to co-exist, whereas the synchronous nature of DS activity pitted different representations against each other in a small time window. Of course, these differences between the two areas were not absolute, but were nevertheless large enough to substantially affect the accuracy of the representations that emerged at the ensemble level.

Although generally disadvantageous in terms of sequence and lever decoding on this particular task, the inherent synchrony in the DS is vitally important for functions mediated by the DS, such as movement...
The differences in persistence and synchrony between the ACC and DS are therefore functionally important and likely reflect the general properties of the two regions. The larger synchrony of DS over ACC neurons may be related to the local circuitry in the striatum itself that is able to transform tonic excitatory cortical drive into alternating and synchronized activity patterns. In addition, because a small pool of interneurons are able to exert powerful control over many medium spiny neurons (MSNs), it is possible that changes in a single interneuron can effectively turn off a large group of MSNs. A second important difference that we observed between the ACC and DS is the long-tailed firing rate distributions exhibited by populations of DS neurons (Supplementary Fig. 3). Long-tailed distributions are consistent with firing at low rates interspersed with brief periods of very high activity. This activity profile is likely a consequence of the biophysical properties of MSNs that require large synchronous inputs from the cortex to fire, as they possess hyperpolarized resting potentials and strong K⁺ currents that suppress all, but the strongest synchronous inputs. The firing characteristics that we observed in the DS population are therefore consistent with the known physiology and anatomy of the region.

The increased variance and covariance across DS neurons may help tune DS neurons to promote a single coherent, but intermixed, signal. Characteristics that we observed in the DS population are therefore functionally important, as excitatory corticostriatal inputs are appropriate for both the striatum and the neocortex. The authors declare no competing financial interests.

ACKNOWLEDGMENTS

This research was supported by Canadian Institutes of Health Research grants (MOP-93784 and MOP-84319).

AUTHOR CONTRIBUTIONS

J.K.S and L.M. designed the study, L.M. conducted the experiments, L.M., J.M.H., A.I.L. and J.K.S performed data analysis and created the figures, A.G.P. helped interpret the results and write the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Duncan, J. An adaptive coding model of neural function in prefrontal cortex. Nat. Rev. Neurosci. 2, 820–829 (2001).
2. Jung, M.W., Qin, Y., McNaughton, B.L. & Barnes, C.A. Firing characteristics of deep layer neurons in prefrontal cortex in rats performing spatial working memory tasks. Cereb. Cortex 8, 437–450 (1998).
3. Fino, E. & Yuste, R. Dense inhibitory connectivity in neocortex. Neuron 69, 1188–1203 (2011).
4. Packar, A.M. et al. Two-photon optogenetics of dendritic spines and neural circuits. Nat. Methods 9, 1202–1205 (2012).
5. Durstewitz, D., Vittoz, N.M., Floresco, S.B. & Seamans, J.K. Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. Neuron 66, 438–448 (2010).
6. Rich, E.L. & Shapiro, M. Rat prefrontal cortical neurons selectively code strategy switches. J. Neurosci. 29, 7208–7219 (2009).
7. Parthasarathy, H.B. & Graybiel, A.M. Cortically driven immediate-early gene expression reflects modular influence of sensorimotor cortex on identified striatal neurons in the squirrel monkey. J. Neurosci. 17, 2477–2491 (1997).
8. Koós, T. & Tepper, J.M. Inhibitory control of neostriatal projection neurons by GABAergic interneurons. Nat. Neurosci. 2, 467–472 (1999).
9. Gage, G.J., Stoezner, C.R., Mitshok, A.B. & Berke, J.D. Selective activation of striatal fast-spiking interneurons during choice execution. Neuron 67, 466–479 (2010).
10. Berke, J.D., Okatan, M., Skurski, J. & Eichenbaum, H.B. Oscillatory entrainment of striatal neurons in freely moving rats. Neuron 43, 883–896 (2004).
11. Graybiel, A.M., Aosaki, T., Flaherty, A.W. & Kimura, M. The basal ganglia and adaptive motor control. Science 265, 1825–1831 (1994).
12. Graybiel, A.M. Building action repertoires: memory and learning functions of the basal ganglia. Curr. Opin. Neurobiol. 5, 733–741 (1995).
13. Averbeck, B.B. & Lee, D. Effects of noise correlations on information encoding and decoding. J. Neurophysiol. 95, 3633–3644 (2006).
14. Sesack, S.R., Deuch, A.Y., Roth, R.H. & Bunney, B.S. Topographical organization of the efferent projections of the medial prefrontal cortex in the rat: an anterograde tract-tracing study with Phaseolus vulgaris leucoagglutinin. J. Comp. Neurol. 290, 213–242 (1989).
15. Zheng, T. & Wilson, C.J. Corticostriatal combinatorics: the implications of corticostriatal axonal arborizations. J. Neurophysiol. 87, 1007–1017 (2002).
16. Mushihaï, H., Saito, N., Sakamoto, K., Itoyama, Y. & Tanji, J. Activity in the lateral prefrontal cortex reflects multiple steps of future events in action plans. Neuron 50, 631–641 (2006).
17. Averbeck, B.B. & Lee, D. Prefrontal neural correlates of memory for sequences. J. Neurosci. 27, 2204–2211 (2007).
18. Shima, K., Isoda, M., Mushiake, H. & Tanji, J. Categorization of behavioral sequences in the prefrontal cortex. Nature 445, 315–318 (2007).
19. Procyk, E., Tanaka, Y. & Joseph, J.P. Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. Nat. Neurosci. 3, 502–508 (2000).
20. Barone, P. & Joseph, J.P. Prefrontal cortex and spatial sequencing in macaque monkey. Exp. Brain Res. 78, 447–464 (1989).
21. Nakamura, K., Sakai, K. & Hikosaka, O. Neuronal activity in medial frontal cortex during learning of sequential procedures. J. Neurophysiol. 80, 2671–2687 (1998).
22. Neura, Y., Mushihaï, H. & Tanji, J. Integration of temporal order and object information in the monkey lateral prefrontal cortex. J. Neurophysiol. 91, 555–560 (2004).
23. Shidara, M. & Richmond, B.J. Anterior cingulate: single neuronal signals related to degree of reward expectancy. Science 296, 1709–1711 (2002).
24. Ryu, J.W. & Wilson, F.A. Making your next move: dorsolateral prefrontal cortical and planning a sequence of actions in freely moving monkeys. Cogn. Affect. Behav. Neurosci. 4, 430–443 (2004).
25. Li, X. & Ashe, J. Anticipatory activity in primary motor cortex codes memorized movement sequences. Neuron 45, 967–973 (2005).
26. Schmitzer-Torbert, N. & Redish, A.D. Neuronal activity in the rodent dorsal striatum in sequential navigation: separation of spatial and reward responses on the multiple T task. J. Neurophysiol. 91, 2259–2272 (2004).
27. Fuji, N. & Graybiel, A.M. Time-varying covariance of neural activities recorded in striatum and frontoral cortex as monkeys perform sequential-saccade tasks. Proc. Natl. Acad. Sci. USA 102, 9302–9307 (2005).
28. Fuji, N. & Graybiel, A.M. Representation of action sequence boundaries by macaque prefrontal cortical neurons. Science 301, 1246–1249 (2003).
29. Aragón, J.W. & Berdidge, K.C. Coding of serial order by neostriatal neurons: a “natural action” approach to movement sequence. J. Neurosci. 18, 2777–2787 (1998).
30. See, M., Lee, E. & Averbeck, B.B. Action selection and action value in fronto-striatal circuits. Neuron 74, 947–960 (2012).
31. Drow, S.L. & McNaughton, B.L. Selective delay activity in the medial prefrontal cortex of the rat: contribution of sensorimotor information and contingency. J. Neurophysiol. 98, 303–316 (2007).
32. Euston, D.R. & McNaughton, B.L. Apparent encoding of sequential context in rat medial prefrontal cortex is accounted for by behavioral variability. J. Neurosci. 26, 13143–13155 (2006).
33. Krzanowski, W.J. Principles of Multivariate Analysis: a User’s Perspective (Oxford University Press, 2000).
34. Berdyyeva, T.K. & Olton, C.R. Rank signals in four areas of macaque frontal cortex during selection of actions and objects in serial order. J. Neurophysiol. 104, 141–159 (2010).
35. Clover, W.T. & Alexander, G.E. Movement sequence-related activity reflecting numerical order of components in supplementary and presupplementary motor areas. J. Neurophysiol. 80, 1562–1566 (1998).
36. Grillner, S., Hellgren, J., Menard, A., Saltoh, K. & Wikstrom, M.A. Mechanisms for selection of basic motor programs—roles for the striatum and pallidum. Trends Neurosci. 28, 364–370 (2005).
37. Carrillo-Reid, L. et al. Encoding network states by striatal cell assemblies. J. Neurophysiol. 99, 1435–1450 (2008).
38. Calabresi, P., Misgeld, U. & Dott, H.U. Intrinsinc membrane properties of neostriatal neurons can account for their low level of spontaneous activity. Neuroscience 20, 293–303 (1987).
39. Wilson, C.J. & Kawaguchi, Y. The origins of two-state spontaneous membrane potential fluctuations of striatal spiny neurons. J. Neurosci. 16, 2397–2410 (1996).
ONLINE METHODS

Animals. Four experimentally naive male Long-Evans rats (450–550 g) were housed in a facility with 12-h light-dark cycle, with all training and recording occurring during the light cycle. For the duration of the behavioral experiments, the rats were food-restricted to just below 90% of their free-feeding weights. Feeding took place in the home cage after their daily training and recording sessions, and water was available ad libitum in the cages at all times. The animals were single-housed to accommodate the food restriction procedures, recovery from surgery and maintenance of the implant. All procedures were carried out in accordance with the Canadian Council of Animal Care and the Animal Care Committee at the University of British Columbia.

Apparatus. In a large Plexiglas box (25 × 18 inches), a main panel was installed with three equally spaced levers from right to left (Supplementary Fig. 1a). On any given day of sequence training, a unique tactile object (velcro, cardboard or soft foam) was stuck to the lever panel (but not on the lever itself) and the area on the floor immediately in front of the lever panel to symbolize the order in which the 3 levers should be pressed. An area of 25 × 13 inches was left for the rat to move freely. On the opposing-side wall, a food cup was located at the center, with each delivery of reward accompanied by a pure tone. Retractable levers and pellet dispenser were controlled and recorded with a PC via a Med Associate interface system (St. Albans).

Sequence task. All training and recording occurred during the light cycle. The naive subjects were first trained on an FR1 schedule to press each of the three levers. A minimum of 60 presses in 0.5 h, with no less than 15 presses on each lever was required before the rat moving on to the next stage of training. After 3–5 d of FR1 training, the rats were trained on sequence A (right lever→middle lever→left lever). Thereafter, they were trained on sequence B (middle lever→left lever→right lever) and finally sequence C (left lever→right lever→middle lever). In each case, food reward was given after the third correct lever press in the sequence. A lever retracted only when it was pressed in the correct order, and remained extended in the event of an error. For training on all sequences, the percentage of correct response on the third item of the sequence had to reach 75% before moving on to the next trained sequence. The order of lever presses in each sequence was given by tactile objects placed on the panel and the floor in front of the levers. For a given animal, each object consistently designated a single serial position. At any one of the three stages of single-sequence training, if after 3 d of training, the animal still hadn’t reached criterion and if day-to-day improvement ceased, a delay-punishment protocol was introduced to extinguish errors made on the third lever of the given sequence. Specifically, if the third lever was pressed before the first lever, all levers retracted and a 10-s time-out period ensued. This training continued until the animal reached criterion performance. When the criterion had reached criteria on all three sequences, the rat was surgically implanted and allowed 10 d to recover. After recovery, 2–3 refresher sessions on each sequence were given before the first multi-sequence test day.

On the multi-sequence test days, the animals had to perform a minimum of ten trials on each sequence at or above criterion, and switch from sequence to sequence in one of three possible orders pseudo-randomly (Sequence C→Sequence B→Sequence A). In between sequences, the animals were taken out of the box to allow for rearrangement of the tactile objects. Because the task was self-paced, and because it always took the animals more than 1 s between lever presses, we defined each lever-press epoch with reference to the time stamp of the lever press. In all but the ROC analysis, we used six 200-ms time bins, including four bins before the lever-press bin, the lever-press bin itself, plus one bin after the lever press. In the ROC analysis, we included one more bin after the six-bin epoch to have seven numbers for the purpose of having meaningful and reliable correlation coefficients. In addition, after the last lever-press epoch in each trial, a 1-s period was defined as the reward approach period, during which the animals ran from the last lever toward the food cup, before consuming the reward. Neural activities during reward approach and consumption are shown in single neuron examples (Fig. 2d–f), but were not used in any analyses.

Surgery. Stereotaxic surgeries were performed with sterilized-tip procedures under anesthesia by isoflurane. NSAIDs analgesic, antibiotic and a local anesthetic were given before incision. One elliptical-shaped craniotomy was made centered at: AP: +3.2 mm, ML: +1.0 mm and another craniotomy was made centered at AP: +1.2 mm and ML: +3 mm. Once the dura mater was retracted, the bottoms of the two bundles of eight 30-gauge tubes, containing a total of 16 tetrodos, were placed on the cortical surface. The bundles were of cylindrical shape with a bottom radius of ~0.4 mm, and were angled medially by ~15 degrees. The implants were fixed with bone screws and dental acrylic. At the end of the surgery, tetrodes in the anterior bundle were extended by ~1.4 mm into the brain to enter the ACC, and tetrodes in the posterior bundle were extended by ~3 mm to enter the DS. Animals were given 10 d to recover. Prior to each recording session, small adjustments were made with the hyperdricers to maximize the number of neurons recorded.

Acquisition of electrophysiological data. For data acquisition, EIB-36TT with pre-amplifier (Neuralynx), connected to the extracellular electrodes, were plugged into HS-36 headstages and tether cables (Neuralynx). Signals were converted by a Digital Lynx 64 channel system (Neuralynx) and sent to a PC work-station, where electrophysiological and behavioral data were read into Cheetah 5.0 software (Neuralynx). Files were then read into Offline Sorter (Plexon) for spike sorting, based on visually dissociable clusters in three-dimensional projections along multiple axes for each electrode of a tetrode (peak and valley amplitudes, peak-to-valley ratio, principal components and area). Sorting was confirmed by examining auto- and cross-correlations, and ANOVAs were conducted from the two- and three-dimensional projections. Spike timestamps were then read into Matlab (Mathworks) for all further analysis.

Histology. At the end of the studies, the animals were deeply anesthetized using urethane intraperitoneal injection, and a 100-µA current was passed through the electrodes for 30 s. Animals were then perfused with a solution containing 250 ml 10% buffered formalin (vol/vol), 10 ml glacial acetic acid and 10 g of potassium ferrocyanide. This solution causes a Prussian blue reaction, which marks with blue the location of the iron particles deposited by passing current through the electrodes. The brains were then removed and stored in a 10% buffered formalin, 20% sucrose solution (wt/vol) for at least 1 week before being sliced and mounted to determine precise electrode locations. Given that multiple sessions were recorded from individual animals the precise recording locations could not be derived from electrode lesions, but all electrode tracks were inferred between the entrance point and the dyed spot. Representative recording sites are shown in Supplementary Figure 2a,c, and the ranges of recording are shown in Supplementary Figure 2b.d.

iFR. A total of 33 large ensembles were collected from four rats that acquired all three sequences and successfully switched among them within a given session. To obtain an estimate of the neural firing rate for each isolated cell i as a function of time bin t, ri(t), for each spike train in each 200-ms bin, the iFRs were calculated as the reciprocals of the inter-spike intervals, convolved with 20-ms Gaussian kernels and then averaged.54–58 Neurons firing less than 0.14 Hz were excluded from further analysis, as the sample of spikes was too small (250 or less) to be reliably representative of the cell’s activity in relation to behavior. Each lever-press epoch included the 1-s period centered at the moment of lever press, whereas the reward-approach period was the 1-s period immediately after the third lever-press epoch.

Multiple linear regression (MLR). A model with five factors that characterized the animals’ spatial location and movements were used to isolate the effect of behavioral variability on neural activities during the common-segment pairs. A linear model was employed simply because we could not specify any type of consistent nonlinear relationship between the firing rate of the neurons and changes in the behavioral variables. Nevertheless linear regression was appropriate since for each neuron the residuals and the predicted values produced by the model were uncorrelated, indicating that the iFRs indeed had a linear relationship with the behavioral predictors. All behavioral analyses were performed using the common segment lever presses. Common segments included instances where the same two levers were pressed in the same order, but as part of two different sequences. 18 out of the 19 sessions had satisfactory tracking of behaviors and thus were included in MLR and other analyses involving the residual matrices (see below). In 12 out of the 18 sessions, the rats performed three sequences, and we used all three possible pairs of common segments. In 6 of 18 sessions, the rats
performed two sequences, so only one pair of common segments existed in each of these sessions, and was used in the analysis.

Behavioral variables were constructed based on video tracking data. As the video rate was 1 frame per 33.3 ms, we averaged six frames to synchronize the video with the 200-ms iFR bins. The first two factors extracted from the video tracking data were the x and y positions of the animals in the chamber. For visualization purposes (Fig. 1a), the x positions were given with reference to the center of each lever. However, for the model, the x positions were given with reference to a single fixed point (that is, the right wall) across the three levers. To generate the next two factors in the model, we calculated the vector created between the animal’s current position and its position 200 ms prior. The third factor A was the approach angle, or the angle between this vector and the lever. The fourth factor V was the bin-by-bin velocity of the animal during lever approach and was simply the length of this vector at each time step (converted into cm s⁻¹). The fifth factor T was the time (in seconds) since the previous lever press.

\[ p(t) = b_0 + b_1 X(t) + b_2 Y(t) + b_3 A(t) + b_4 V(t) + b_5 T(t) + \epsilon(t) \]  

For a given time bin \( t \), \( p(t) \) represented the neuron’s normalized iFR in that bin. We then examined the distribution of the percentage of variance accounted for by the model (the \( R^2 \) statistic) for each neuron (Fig. 1b). The residual matrix resulting from the regression in each session was used to perform a series of control analyses, including \( D_{\text{Mah}} \), \( D_{\text{Mah}} \)-based leave-one-out error, and ensemble variance and covariance (see below). By comparing the results from these residual-based analyses with those from the full iFR matrices, the effect of behavioral variability on sequence decoding was elucidated.

SI. To examine whether individual units were responsive to sequence, the selectivity for each neuron with respect to each pair of common segments—both of which were associated with the same lever but belonged to different sequence blocks—was obtained by grouping the firing rates into two classes: the iFRs of a given neuron during one common segment were assigned to class A, whereas the iFRs of the same neuron during the other common segment in the pair were assigned to class B. The index was then computed as

\[ d' = \frac{\langle r_f(t) | r \in A \rangle - \langle r_f(t) | r \in B \rangle}{\sqrt{\sigma^2_{I,A} + \sigma^2_{I,B}}} \]  

where \( \langle \cdot \rangle \) denotes the mean.

ROC curve and statistics. The ROC method was used to test classification performance of single neurons and ensembles. ROC analysis has an advantage in situations in which the distributions are unknown because it assesses performance over a range of threshold values rather than being forced to evaluate differences at a single threshold level relative to some theoretical distribution. In a simple example, if a detector’s sensitivity level is set to 0.5, signals (or observations) stronger than this level are reported as positives (or 1) and below as negatives (or 0). Out of all positives, the proportion of incidents when the target was truly present would be the hit rate and the remaining proportion would be the false alarm rate. Each threshold level yields a single dot in the ROC curve. We ran the ROC analysis on correlation scores as described for the classification of both sequences and lever locations. Because correlation coefficients are on a continuous scale, a large number of thresholds were used to produce the detailed ROC curves. We used the AUC from each individual curve for statistical analysis. Because a 45-degree straight-line with an AUC = 0.5 indicates a lack of signal-noise differentiation, the AUCs can be tested for significance by a simple one-sample t-test against a normal distribution with mean = 0.5 and unknown variance. In the case where AUCs from ensembles and from single units were compared, an independent-sample t-test was used. In all of the cases involving multiple groups, a two-way ANOVA was also used with two control groups—random normal distributions with the same sample size and the same variance and mean = 0.5.

To analyze the effect of ensemble size on the between-sequence separation of activity states, ensemble AUCs were calculated between pairs of common segments (that is, lever responses that differed only in their sequence identity) based on randomly selected ensembles with varying sizes \((n = 4, 7, 10, 13, 16, 19)\). For each ensemble size, 100 random draws of \( n \) neurons were performed for each session and the resulting AUC averaged, for a complete representation of the whole data set. For both the ACC and the DS, the average AUCs at each ensemble size were plotted (Fig. 3c) and a power function was fitted for each region. The power function was selected as it accounted for more than 99% of the variance.

MSUA analysis. For population analysis, population vectors \( v(t) = [v_1(t) \ldots v_N(t)] \) were constructed, with \( N \) equal to the number of single units isolated from a given recording session. The term MSUA space refers to the \( N \)-dimensional space spanned by all recorded units and populated by these vectors \( v(t) \). Each dot in the MSUA space represents the state of the entire recorded ensemble in one 200-ms bin. All points corresponding to different 200-ms bins in the epochs of the same behavior are shown in the same color. All statistical analyses were performed in the full space of all recorded units. For the purpose of visualization, multi-dimensional scaling was applied to reduce dimensionality.

To quantify the effects of sequence and lever location on network activity, the Mahalanobis distances (\( D_{\text{Mah}} \)) were computed between the sets of \( N \)-dimensional vectors associated with task epochs of interest. To control for differences in MSUA space dimensionality (that is, ensemble size) in \( D_{\text{Mah}} \) comparisons, a normalization procedure was employed: \( N_{\text{min}} \) was the minimum number of units recorded in any of the data sets to be compared, and \( K_{\text{min}} \) was the minimum number of time bins. For data sets with \( N \) and \( K \) greater than \( N_{\text{min}} \) and \( K_{\text{min}} \), \( N_{\text{min}} \) units and \( K_{\text{min}} \) data points were selected at random and \( D_{\text{Mah}} \) was computed. This procedure was repeated 1,000 times and the results were averaged to make full use of all units and data points recorded. The resultant normalized \( D_{\text{Mah}} \) averages were used in various statistical analyses. To determine the significance level of a given \( D_{\text{Mah}} \) value, between-sequence separation was compared to within-sequence separation (Fig. 4c), and between-lever separation was compared to within-lever separation. To calculate average \( D_{\text{Mah}} \) in a sequence block, bootstrap surrogate blocks were created by randomly shuffling 1-s blocks of the iFR matrices.

The distance between the two shuffled blocks therefore represents the separation between activities during random behavioral events. The process was repeated 100 times and the \( D_{\text{Mah}} \) values averaged. We also used the residual matrices from the linear regression (see above) to calculate the \( D_{\text{Mah}} \) and compared the results to the \( D_{\text{Mah}} \) calculated from the full iFR matrix in order to reveal the influence of the behavioral variables on sequence decoding (Fig. 4c).

The calculation of \( D_{\text{Mah}} \) incorporates three aspects of ensemble activity: the difference in mean firing rates (that is, Euclidean distance or \( D_{\text{Euc}} \)), the variance in each neuron’s activity and covariance between any two cells in the ensemble. We focused on the bin-by-bin variance and covariance, calculated among neurons after averaging across all trials for each time bin within the lever-press epochs. To better understand the difference between ACC and DS in sequence encoding on the ensemble level observed in \( D_{\text{Mah}} \) we also analyzed \( D_{\text{Euc}} \) variance and covariance separately. From each session, 19 cells were randomly drawn and the total covariance was calculated and summed, and the process was repeated for 100 times before the results were averaged for each session. Thus the ensemble bin-to-bin variance was the average of summed variance in a typical 19-cell (that is, \( N_{\text{val}} \)) ensemble (Fig. 5a), and the ensemble covariance shown in Figure 5b is the summed absolute covariance between each cell pair in an ensemble, as we were concerned with the magnitude of the covariance rather than its direction. We also used the residual matrices from the linear regression (see above) to calculate the ensemble variance and covariance to examine the effect of behavioral variation (Fig. 5a, b).

Leave-one-out MDA. In the analysis above, \( D_{\text{Mah}} \) was calculated between clusters. The leave-one-out prediction uses another variant of \( D_{\text{Mah}} \): dot-to-cluster distance. In each distance calculation, a ‘dot’ is the ensemble iFR vector of a single time-bin recorded during a common-segment action. When all the dots for each sequence block are plotted together, there are two clusters in the MSUA space. If the \( D_{\text{Mah}} \) from the dot to its home sequence cluster is shorter than that to the alternative sequence cluster, then a correct classification is counted. The final performance is shown in percentage of correct classifications out of all time-bins tested (Fig. 4d). To have a control for the classification performance, bootstrap surrogate blocks were created by randomly shuffling 1-s blocks of the iFR matrices. The distance between the two shuffled blocks therefore represents the separation between activities during random behavioral events. The process was repeated 100 times and the leave-one-out errors averaged. Two-way ANOVA was then used to test the performance of ensembles from both regions (Fig. 4d). In addition, the residual matrices from the linear regression (see above) were used...
to calculate the leave-one-out errors and their results were compared with the results to those calculated from the full iFR matrix (Fig. 5d,f).

**F statistic for sequence discrimination and for lever discrimination.** The F statistic was calculated separately for sequence classification and lever discrimination. To characterize the cells’ temporal profile of sequence discrimination (Fig. 6a), for each neuron, the F statistic was calculated between the iFRs during each pair of common segments (from two different sequence-blocks) in each of the six time bins, covering the interval from 900 ms before a given action to 300 ms after the action. In other words, the F statistic for sequence was the between-sequence variance divided by the within-sequence variance in a given time bin. We then plotted the frequency distribution of the time bin in which a cell achieved its maximum F statistic for each region (Fig. 6a).

To characterize the level of lever discrimination throughout the lever-press epoch, F statistics were also calculated among each cell’s activities associated with responses on the three levers within each sequence block. The F statistic for lever presses was the between-lever variance divided by the within-lever variance in a given time bin, in a given sequence. The frequency distribution of the time bin in which a cell achieved its maximum F statistic is plotted in Figure 6b.

A Supplementary Methods Checklist is available.

40. Hyman, J.M., Ma, L., Balaguer-Ballester, E., Durstewitz, D. & Seamans, J.K. Contextual encoding by ensembles of medial prefrontal cortex neurons. *Proc. Natl. Acad. Sci. USA* **109**, 5086–5091 (2012).