Successful choice behavior is associated with distinct and coherent network states in anterior cingulate cortex

Christopher C. Lapish†, Daniel Durstewitz†‡, L. Judson Chandler†, and Jeremy K. Seamans††

1Department of Psychiatry and Brain Research Centre, University of British Columbia, 2211 Wesbrook Mall, Vancouver, BC, V6T 2B5; 2Centre for Theoretical and Computational Neuroscience, University of Plymouth, Portland Square, Plymouth, PL4 8AA, United Kingdom; and 3Department of Neuroscience, Medical University of South Carolina, 67 President Street, Charleston, SC 29425

Edited by Charles F. Stevens, The Salk Institute for Biological Studies, La Jolla, CA, and approved June 22, 2008 (received for review April 29, 2008)

Successful decision making requires an ability to monitor contexts, actions, and outcomes. The anterior cingulate cortex (ACC) is thought to be critical for these functions, monitoring and guiding decisions especially in challenging situations involving conflict and errors. A number of different single-unit correlates have been observed in the ACC that reflect the diverse cognitive components involved. Yet how ACC neurons function as an integrated network is poorly understood. Here we show, using advanced population analysis of multiple single-unit recordings from the rat ACC during performance of an ecologically valid decision-making task, that ensembles of neurons move through different coherent and dissociable states as the cognitive requirements of the task change. This organization into distinct network patterns with respect to both firing-rate changes and correlations among units broke down during trials with numerous behavioral errors, especially at choice points of the task. These results point to an underlying functional organization in cell assemblies in the ACC that may monitor choices, outcomes, and task contexts, thus tracking the animal’s progression through “task space.”

cell assemblies | decision making | multiple single-unit recordings | neural coding | population analysis

There is general agreement that processing of information in the neocortex is done by networks of cells operating in a coordinated fashion rather than working independently, as in a purely feed-forward-type architecture. The most popular concept to describe these processes dates back to Hebb’s (1) proposal on cell assemblies, which has been investigated from a variety of experimental perspectives (2–8) and has formed the foundation of a number of computational approaches (9–12). There is now accumulating evidence that the transient organization of neurons into dynamic ensembles and the sequential transitions among them may form the basis for cortical information processing (2, 3, 5, 7, 13–15).

Although such processes have been investigated in some depth for perceptual and spatial domains, much less is known regarding the network dynamics that govern higher-order cognitive processes, such as action and outcome monitoring. The anterior cingulate cortex (ACC) has been the subject of increased interest as a region that plays a role in action monitoring, a supervisory cognitive function that is especially important for optimal decision making in challenging and novel situations (16–19). A number of diverse single-unit correlates accompanying these processes have been observed in the ACC (20–27), yet it is unknown how single neurons organize into functional networks that could serve these functions.

In the present study, we investigated functional ensemble dynamics within the ACC during the performance of the delayed win-shift radial-arm maze, a task with distinct cognitive phases and high ecological validity (28). The locations of rewards changed between and within trials, such that the animal had to continually monitor its own actions and track the changing environmental contingencies as it progressed through the task. Using sophisticated statistical methods, we found that the recorded population appeared to track each aspect of the task by entering distinct and separable network states that broke down on trials with numerous errors.

Results

Population Activity Organizes into Specific Patterns During Distinct Task Epochs. Multiple single-unit recordings from the rat ACC (dorsal medial agranular prefrontal cortex; see supporting information [SI] Fig. S1A) were obtained from 10 rats over a total of 27 trials of the delayed-win-shift radial arm-maze task. The task consisted of the following behaviorally dissociable epochs illustrated in Fig. 1A: the periods surrounding correct arm choices during the training phase (TrC) and test phase (TsC), respectively; periods surrounding incorrect choices during the test phase (TsI; there were very few incorrect choices during the training phase); the period surrounding the point when the animal reached a food cup (TrR, during the training phase, and TsR, during the test phase); the entire delay phase (D); and all of the remaining periods intervening between arm choices and reward epochs during the training (Tr) and test phases (Ts).

Consistent with previous studies (24, 29–33), we observed a number of single-unit correlates that were associated with arm entries, reward processing, specific movements, and behavioral errors (Fig. 1E). However, the focus of the present study was the functional patterns of network activity associated with the distinct task epochs rather than single-unit correlates.

To analyze population activity, spike trains of all units were first convolved with Gaussian functions and binned at 200 ms, yielding smoothed instantaneous firing rates (iFR) for each unit as a function of time. For each time bin, the iFR of the N simultaneously recorded units were combined into population vectors embedded within an N-dimensional state space, here termed multiple-unit activity (MUA) space. For the purpose of visualization, we used multidimensional scaling to obtain 3D projections of these N-dimensional MUA spaces as shown in Fig. 1B and C. In these graphs, each dot represents the entire state of the recorded network within one 200-ms bin of the task, and all population vectors (dots) belonging to different 200-ms bins
of the same task epoch are shown in the same color. As Fig. 1B illustrates (see also Movies S1 and S2 and Fig. S2), different task epochs segregate in MUA space, i.e., MUA vectors belonging to the same task epoch tend to cluster within similar regions of MUA space, whereas MUA vectors belonging to different task epochs populate different regions, implying that the population as a whole differentiates between individual task epochs. Averaging across all points in MUA space belonging to the same task epoch yields an N-dimensional “prototype vector” representing the mean iFRs of all units within this epoch, illustrating the differential patterns of activation as a function of task epoch (Fig. 1D).

This organization into distinct patterns appeared to be functionally relevant to the successful completion of a trial. Fig. 1B illustrates population activity for an animal that performed perfectly on a trial. However, when the population activity for the same animal was reexamined during a trial in which six errors were committed, the segregation in MUA space broke down (Fig. 1C; see also Movies S1 and S2 and Fig. S2). This collapse of segregation in MUA space was also reflected in the lack of differentiation between task-epoch iFR prototypes (Fig. 1D). On this trial, only one unit showed clear differential activity, and it appeared to be most tightly linked to the signaling of behavioral errors (Fig. 1E). Therefore, successful task performance was characterized by the neural population attaining distinct task-epoch-specific patterns that changed dynamically with the cognitive requirements of the task. A failure to exhibit such distinct population patterns was correlated with errors in choosing the correct arms.

To quantify these phenomena statistically, activity patterns were analyzed across all animals used in this study by computing a separation error for each trial and every pair of task epochs. This was done by fitting a hyperplane (linear discriminant function) that optimally separated the clouds of points associated with each pair of task epochs using discriminant analysis and by determining for all of these pairs the relative number of population vectors (i.e., dots in MUA space) that were assigned to the wrong task epoch according to this linear classifier. Because the separability of task-epoch points is also affected by other factors such as the dimensionality of the MUA space and the total number of points/epoch, conservative surrogate data, obeying the same temporal continuity constraints as the original data, were constructed and used to test the significance of separation between any pair of task epochs (see Methods for details). Fig. 24 shows the separation errors for each of the original pairwise comparisons and for the respective surrogates constructed for each pair of task epochs averaged across all trials and animals. Each of the pairwise comparisons reached significance except for the comparison between test-phase incorrect (TsI) and correct (TsC) choices (paired t tests using corrected 0.050 levels according to the Holm–Bonferroni method, (34), α = 0.05; see SI Text and Fig. S3 for additional separation measures and
Individual Neurons Are Generally not Selective for Single-Task Epochs. As described above, ensemble activity significantly differentiated between almost all pairs of task epochs. To address the question to what extent this differentiation may be rooted in highly task-epoch selective responses of single ACC neurons, a selectivity index for each unit and task epoch was computed that compared the average activity of a unit during a given epoch with the average activity across all other epochs (see Methods for details), and the significance of this index was again tested using surrogates. A significant selectivity index of a unit for a particular epoch implies that its average activity during that epoch significantly deviated (positively or negatively) from the average activity across all remaining epochs. For each of the eight task epochs, between 13% and 27% of all units were found to be task-epoch-selective according to this definition (Fig. 2B). However, cells may significantly modulate their average iFR compared with the grand mean during more than one task epoch (e.g., units 5 and 9 in Fig. 1E). Indeed, ≈80% of task-epoch-selective units exhibited a significant modulation in at least one other epoch (Fig. 2C, gray bars). Conversely, this implies that <4% of all units recorded were highly selective for just one epoch. On the other hand, if any two of the epochs were compared with each other, 72% of the epoch-selective units were active in only one of the two instances and could therefore account for their separation (Fig. 2C, white bars). Thus, although each unit was generally modulated during more than one epoch, all pairs of task epochs differed in their constitution of selective units.

Different Task Epochs Are also Associated with Unique Coalitions Among Neurons. Our results thus far show that different cognitive phases of the task are associated with specific patterns of iFR activity across all recorded units. To examine whether these iFR patterns are also accompanied by epoch-specific correlations among units, the absolute (zero time lag), iFR Pearson correlation coefficients averaged across all pairs of units as a function of task epoch were computed (using 200-ms bins as before; Fig. 2D, dark bars). For each task epoch, except for the incorrect choices, iFR comodulations were significantly above chance (t tests, all $P < 0.005$) based on comparisons with data sets where we shuffled the iFR bins (Fig. 2D, light gray bars). To assess whether these correlated activity changes were indeed task-epoch-specific or simply a general feature of the neuronal pairs recorded, we extracted the pairs with the 20% highest correlations of task epoch with regard to the most highly correlated pairs. A given task epoch shared only ≈20–30% on average of its most highly correlated pairs with any other task period (Fig. 2E; Fig. S4), suggesting that coalitions among neurons formed and dispersed with each task epoch. As a further confirmation of this observation, we constructed surrogates by recombining iFR bins from different task epochs while maintaining the temporal relation among units (i.e., no shuffling of iFR bins). In all cases except for the TsR epochs ($t$ test, $P < 0.1$), these across-task-epoch iFR correlations were significantly lower than the within-task-epoch correlations ($t$ tests, all $P < 0.005$), even after correction by the shuffle predictor (Fig. S5). Hence, different task epochs are not only differentiated by unique patterns of changes in firing rate but also through task epoch specific coalitions among units.

Ensemble Organization Is Diminished on Trials with Numerous Behavioral Errors. To address the functional importance of the distinct network patterns, changes in population activity as a function of behavioral errors were examined. For the entire dataset, trials were divided according to a median split based on the number of incorrect choices. The resulting groups agreed with previously defined criteria of asymptotic performance on this task (0–1 vs. 2 + incorrect choices; ref. 35). For the combined delay and test phases, separability was significantly worse for trials with many incorrect choices versus trials with 0–1 incorrect choices (Fig. 3A; $t$ test, $P < 0.01$). Furthermore, Fig. 3A shows that the breakdown in MUA space affected mostly comparisons involving test-phase choice and reward epochs (TsC and TsR), and less...
so basal test vs. delay phase epochs (Ts vs. Dl). There was not only an overall significant decrease in separation but also a lower number of individual trials that yielded significant separation for comparisons involving TsC and TsR epochs (Fig. 3B). In addition, task-epoch specific iFR correlations among units also tended to decrease as a function of behavioral errors, but only with TsC did this reduction reach statistical significance (Fig. 3C; t test, P < 0.05). In conclusion, on trials where animals made two or more choice errors, network dynamics were severely compromised as evidenced by a failure to organize into distinct task-epoch-specific patterns both in terms of firing rate activity and correlations among units. This was especially true at choice points during the test phase, suggesting that errors are associated with a failure of the ACC to enter into unique and coordinated network states particularly at those times where decisions about arm entries are made.

**Discussion**

The present study explored population coding within the ACC while animals foraged for food in an ecologically valid radial arm maze task involving distinct cognitive phases. Many single-unit correlates (e.g., of choices or rewards) were observed. In most cases, however, single units did not limit their firing rate changes to one specific event type but were active across multiple cognitively defined task epochs. On the other hand, each task epoch was characterized by a unique pattern of firing-rate changes across units and correlations among units. The behavioral importance of this functional organization was supported by the fact that these population patterns tended to break down on trials with numerous errors.

**Cell Assemblies and Dynamic Population Patterns.** The cell assembly framework first introduced by Hebb (1) could provide a mechanistic basis for the observed ensemble patterns. A cell assembly is defined by a functional group of neurons entertaining relatively strong recurrent excitatory connections among each other, while potentially inhibiting pyramidal cells of other assemblies. ACC networks may consist of various partially overlapping cell assemblies encoding various cognitive events within the decision making process as illustrated in Fig. 4. As a result of such functional arrangements, at the single-neuron level, there would be numerous combinations of enhanced and depressed activity changes across different cognitive events, whereas each assembly as a whole would be associated with a unique pattern of firing rate changes and a unique pattern of within-assembly correlations (due to the recurrent excitatory connections within but not across assemblies), as observed here (Figs. 1 and 2).

Many other approaches, both similar to and different from ours, have been used to identify ensemble organization in multiple single-unit data. These include Hidden Markov Models, which also revealed distinct patterns and transitions among them in vivo (14, 36), and principal component or cluster analysis to unravel functional groupings (37), hierarchical organization of task-coding assemblies (7), or simply as a means for visualizing distinct population states and their connecting trajectories (15). At the level of precise spiking times, task-phase-specific alliances among neurons were demonstrated in primate prefrontal areas using spike train cross-correlograms (13) and repeating patterns of spikes aligned among multiple neurons (6, 38) were also taken as evidence for ensemble organization. Hence, a number of
studies employing different analysis methodologies provide converging evidence for the organization of neurons into functional ensembles at different temporal scales.

**ACC and Functional Assemblies in Behavioral Monitoring.** Transient lesions of the rat ACC produced significant impairments on the same task used here (35), causing numerous revisits to previously baited arms indicative of an inability to update responding as reward contingencies changed. Consistent with the putative function of the ACC in monitoring and updating response policies (39), single-unit activity in this area is most often related to specific movements, movement sequences, or response choices, as well as to rewards, specific action-reward pairs, or to the detection of response errors (21–27, 29–33). One important aspect of these single-unit correlates is their highly dynamical nature, with neural firing rates changing as a function of task context and reward magnitude even for the very same movements (24, 32, 40), or, vice versa, with the same neurons coding for very different movements in different tasks (29, 40). This high degree of flexibility and context dependence supports the view of the ACC as a device for monitoring or attending to actions (16, 17).

Within this framework, the functional organization into ensembles and the transitions among them may help to bind different cognitive attributes of a response strategy together, and to connect actions to subsequent outcomes (39). Performing a complex or novel task requires an animal to integrate various sensory, motor, reward, and memory aspects, and the transient formation of dynamic ensembles in the ACC as observed here may reflect this process. In this sense, the assemblies monitor the animal’s progression in a task-dependent frame of reference or “task space.” The sequential transitions among ACC assemblies and the transepoch activity of some of the units recorded may in addition help to connect the temporally separated components of action-outcome chains. Sequential transitions among active populations as well as persistent activity of single frontal neurons have been proposed previously as a means to link events through “task space.” The sequential transitions among ACC assemblies and the transepoch activity of some of the units recorded may in addition help to connect the temporally separated components of action-outcome chains. Sequential transitions among active populations as well as persistent activity of single frontal neurons have been proposed previously as a means to link events through “task space.”

**Methods**

**Data Acquisition.** Male Long-Evans rats (225–250 g, Harlan) were deeply anesthetized with 100 mg/kg ketamine and 10 mg/kg xylazine and placed into a stereotaxic device where multielectrode arrays were inserted through a cranial hole centered at A/P 2.2; M/L 0.8; D/V 2.5 relative to bregma and offset 10° from the vertical. After recovery, rats received 1 trial per day of the delayed win-shift radial arm maze described in detail elsewhere (35). During the training phase, four open arms were chosen randomly and baited, whereas the remaining four arms were blocked by a door. Upon visiting all four baited arms, the animal was locked into the last arm it visited, and the lights were turned out. After a 60- to 90-sec delay, the light was turned on, the door opened, and the animal began the testing phase. During the testing phase, all eight arms were open, and the rat had to visit the four arms that were blocked during the training phase. An error was scored as an entry into an arm that has been visited previously during either the training or test phase. All behavior was recorded with an online frame-capture COHU camera synchronized with timesteps created by the Neuralynx recording system (Custom Software). Off-line analysis of the video was used to collect the timesteps of behaviorally relevant events, from which event files were created with Event Session Splitter (Neuralynx). At the end of the experiments, each animal was deeply anesthetized with pentobarbital and transcardially perfused with 4% paraformaldehyde (P6148, Sigma). The brains were then collected, sectioned at 50 µm, and placements were observed with a dissecting microscope at 20× (Cambridge Instruments). See SI Text for in-depth description of arrays and unit recording and isolation parameters.

**Data Analysis.** All spike trains were first convolved with Gaussian functions to yield smoothed firing-rate functions and minimize the impact of random spike-time jitter at the borders between bins. Spikes were then binned at 200 ms (approximately the inverse of the average firing rate of ~4.8 Hz), and all simultaneously recorded neurons were combined into N-dimensional vectors iFR(t). To confirm the visually apparent separation among task epochs statistically, a linear classifier was constructed for each trial and pair of task epochs by determining an optimally separating (N – 1)-dimensional hyperplane via discriminant analysis (e.g., ref. 44). The relative number of misclassified points was taken as an index of separability and was compared to surrogates constructed in the following way: Pairs of task epochs were first combined, and from these unions of points k contingent segments were randomly drawn, where k is the number of contingent segments for the original task epoch. To determine significance for each dataset individually, a nonparametric test was applied: significance at the 5% level was assigned by establishing whether the original classification error was among the 5% lowest within the set of 1 original and 99 surrogate classification errors.

To examine percentages of epoch-selective cells, for each cell α and task epoch p, a selectivity index was computed as

\[ s_\alpha(p) = \frac{\| iFR_\alpha(t)_{p} - \langle iFR_\alpha(t)_{q} \rangle \|}{\| iFR_\alpha(t)_{p} + \langle iFR_\alpha(t)_{q} \rangle \|} \]

where \( \| \cdot \| \) denotes the average across segments of epoch p, or across all of the combined other epochs q, respectively. Significance of each index \( s_\alpha(p) \) at the 5% level was again established by a nonparametric comparison with 99 surrogates, where in this case k contingent segments, with k being the number of segments making up epoch p, were drawn at random from the combination of all epochs in the task.

For the correlation analysis, absolute standard zero-lag Pearson correlation coefficients \( r_{p, n}(m) \) were computed among scalar times series \( iFR_{\alpha}(t) \) and \( iFR_{\beta}(t) \) for all pairs of simultaneously recorded neurons \( n,m \in \{1, N\}, n \neq m \), and separately for all task epochs p. To test significance nonparametrically, 99 surrogates were constructed for each \( r_{p, n}(m) \) by randomly shuffling the bins within series \( iFR_{\alpha}(t) \) and \( iFR_{\beta}(t) \). For comparison of within-task-epoch to across-task-epoch correlations, equal time slices were first drawn from all task epochs to ensure that the surrogates drew from every task epoch with equal likelihood. Surrogates were then constructed by drawing for each comparison at random 20 bins from the union of time-equalized task epochs, original correlations were also recomputed for the time-equalized task epochs, and both original and across-epoch surrogate correlations were corrected by the mean from 99 shuffles.

For the comparison of the two behavioral error groups, all separation errors for each dataset were first normalized to the average separation error within the respective set of surrogates. However, the two behavioral error groups based on the median split did not differ with regard to either the average number of recorded neurons (ME = 13.8 for the low and ME = 14.2 for the high error group, two-sided t test: \( P > 0.86 ) \) or the standard deviation of this number (SD = 5.2 for the low and SD = 5.8 for the high error group), i.e., the dimensionality of the MUA spaces for those two groups was the same.

Further details of methods and methodological considerations can be found in SI Text.

**Acknowledgments.** We thank Drs. A. G. Phillips, J. Cohen, J. Hyman, N. Vitoz, T. Wennekers, and D. Euston for helpful comments during the preparation of the manuscript. This work was funded by the National Institutes of Health (Grants MH0655924, MH064569, and AA10983) and the Canadian Institute of Health Research (Grant MOP-84319).

Supporting Information

Lapish et al. 10.1073/pnas.0804045105

SI Text

Data Acquisition. Custom multielectrode arrays were built using 25-µm-diameter tungsten wires (California Fine Wires) in a 36-gauge silica tube yielding a spacing of 150 µm, arranged in a 2 by 12 matrix with the length spanning the anterior-posterior axis of the ACC (Fig. SI.4). They were then attached via gold pins to an EIB-27 board and a HS-27 headstage (Neuralynx). A commutator connected the HS-27 to 4 Lynx-8 programmable amplifiers and an EPP-27 patch panel running Cheetah acquisition software (Neuralynx). Signals were sampled at 30,305 Hz filtered between 600–6,000 Hz. Online spike detection signals were amplified 5–10,000 times and thresholds set at 75–125 µV or >3 times the noise amplitude. Clustering was performed through KlustaKwik and Neuralynx’s SpikeSort 3D software. They were then manually assigned into clusters by Neuralynx SpikeSort 3D, where multiple parameters were used to effectively visualize clusters with the most often used combination of spike height, trough, and energy. Each cluster was then run through an ISI filter to remove any ISIs <10 ms and any duplicate timestamps from the data set. Data analysis of the clusters was performed by custom written routines in Matlab (MathWorks).

Data Analysis. The following provides additional details and discussion regarding our analysis methods as well as on additional checks we have performed.

MUA space separation. For the smoothing of spike trains, Gaussian functions with 50 ms standard deviation and an integral of one were used.

Solely for the purpose of visualization (Fig. 1), metric multidimensional scaling (MDS) was used to obtain 3D projections of the N-dimensional spaces (N = number of simultaneously recorded cells). Metric MDS finds a lower-dimensional projection of the data points while attempting to preserve the original distances among all points within this projection.

Prototypes of task epochs were computed by averaging across all vectors of that task epoch. In Fig. 1D, the difference of each task epoch prototype from the grand average across all task epochs is presented to highlight the differential activities.

To confirm the visually apparent separation among task epochs statistically, a linear classifier was constructed for each trial and pair of task epochs by determining an optimally separating (N-1)-dimensional hyperplane via linear discriminant analysis (e.g., ref. 1). This linear classifier is optimal in the sense of a maximum-likelihood criterion under the assumption that the data are multivariate normally distributed. We chose it for its computational simplicity, which is important given the thousands of surrogates we had to evaluate, its straightforward statistical interpretation, and the absence of any user-defined parameters that might make the process more subjective (as required in many other machine-learning approaches). In Fig. S3, we furthermore show results from two other related statistics, namely the Mahalanobis distance between the group means (which can be viewed as a multivariate extension of d2 in signal detection theory), and the \( R^2 \) statistics (explained variance) obtained by multiple regression where class (task epoch) membership is correlated with IFR(t). Although classification based on multiple regression is equivalent to that obtained by linear discriminant analysis for the two-category case (see ref. 2, Ex. 4.2, for details), \( R^2 \) provides a finer-grained (real-valued) statistics than the classification error.

Because the separation performance will depend on the dimensionality of the space and the numbers of points to be separated (e.g., \( N \) points can always be perfectly linearly separated in an (N-1)-dimensional space), surrogates were constructed and used as baseline. A simple way to obtain such surrogates would be to randomly reassign for each comparison the original population vectors to the two task epochs that are to be compared, and then recalculate the classification error. However, as described in Methods, we used a more strict criterion to account for potential temporal contingencies within the clouds of points: Pairs of task epochs were first combined, and from these unions of points \( k \) contingent segments were randomly drawn, where \( k \) is the number of contingent segments for the original task epoch. For instance, there are four correct choice epochs consisting of \( 5 \times 200-\text{ms} \) bins each embedded within the test phase, and hence for a comparison TsC-Ts four segments of five consecutive points are drawn at random from the combined Ts and TsC points. To determine significance across all data sets, the average across the 99 surrogates for each set was obtained and compared to the original classification errors across all data sets by paired t tests with alpha-levels corrected according to the Holm–Bonferroni method (3) [another modified Bonferroni method suggested by Cross and Chaffin (4) yielded identical results].

As noted in the main text, a significant separation among task epochs could on average (across all task epoch pairs) be achieved in ~40% of all trials, with contributions from all trials and animals. More specifically, for each trial from each animal there was a minimum of 2 significantly differentiated pairs of task epochs and an average of 10.6 (SD = 4.2). Hence, the results are neither specific to one or a few data sets nor to just a few animals, but are a general phenomenon across trials and animals.

As a further check of the results of the separation error analysis, we constructed “metasurrogates” by randomly shuffling the event times within each task phase and subjecting this new data set with randomly defined task epochs to the full analysis. In this case, with only one exception none of the comparisons between shuffled event epochs from the same task phase reached significance.

Low vs. High Behavioral Error Group Comparisons. With regard to the comparison of the two behavioral error groups, we note that the differences among these groups in overall separation (averaged across all epochs) did not reach significance if only training phase-associated comparisons were included (\( t \) test, \( P > 0.1 \)), i.e., when delay- and test-phase comparisons were excluded. This was expected, because almost all errors were committed during the test phase, i.e., the two groups were comparable in their performance during the training phase.

We also emphasize that Fig. 3A indicates a general breakdown of separation in the high-error group, not one that is confined to those particular choices that were incorrect. As a further test, we compared separation from the basal test phase during correct and incorrect choices for the high-error group only, and found no significant differences between them (\( \mu_{TsC-Ts} = 0.83, \mu_{TsI-Ts} = 0.87, P > 0.6 \) paired t test), whereas both values were significantly larger than the respective mean for correct choices in the low error group (\( \mu_{TsC-Ts} = 0.61, P < 0.05 \) for both). Hence, the lack of MUA space separation is related to whether in general many or few behavioral errors were committed but not to particular choices made (e.g., specific arm locations) or whether they are correct or incorrect.
**Correlation Analysis.** As noted in *Methods*, for comparison of within-task-epoch to across-task-epoch correlations (Fig. S5), equal-time slices were first drawn from all task epochs. More precisely, given that TrC, TrR, TsC, and TsR all consisted of 4 × 5 consecutive bins, 4 × 5 consecutive bins were also drawn from the Tr, Di, and Ts epochs at random, and Ts1 was incorporated only if it had at least 4 × 5 bins, which was also the maximum allowed. This procedure ensured that the surrogates drew from every task epoch with equal likelihood. They were now constructed by drawing for each comparison at random 20 bins from the union of time-equalized task epochs, and correlation coefficients were recalculated between time series $iFR_m(t)$ and $iFR_n(t)$ for all pairs of neurons $n,m$, and for each surrogate set $s$. Original correlations were also recomputed for the time-equalized task epochs, and both original and across-epoch surrogate correlations were corrected by the mean from 99 shuffles (within original or surrogate epochs $p$ or $s$, respectively), to rule out any potential dependence on absolute firing rates. This process was then repeated 99 times for each dataset, yielding statistically reliable estimates. Note that this is a very conservative test as it retains all correlations among units that may exist across time-equalized task epochs.

All analysis routines were custom-written in Matlab or C++.

**Further Analysis Results.** To yield some information about the temporal precision of the task epoch-dependent organization into iFR and correlational patterns, the iFR($t$) vectors of each neuron were shifted by a random number of bins drawn from a Gaussian distribution with zero mean and standard deviations of $\sigma_{\text{shift}} = 1, 3, 6, 8, \text{or } 10$ bins (100 of such shift surrogates were created for each data set and value of $\sigma_{\text{shift}}$). Both MUA space separation (averaged across all pairs of epochs) and the average iFR correlations monotonically decayed as the standard deviation $\sigma_{\text{shift}}$ was increased (Fig. S6). For the MUA space separation (Fig. S6A), this progressive deterioration became significant (paired t test, $P < 0.05$) from at least $\sigma_{\text{shift}} = 3$ bins onwards, and continued to significantly deteriorate as $\sigma_{\text{shift}}$ was further increased. This suggests that pattern formation in accordance with task epochs may be a more coarsely grained process, which is not so surprising given that task epochs themselves stretch out across significant periods of time ($\geq 1$ s). Note in particular that even the largest shift ($\sigma_{\text{shift}} = 10$) should have only little effect on all comparisons across the major task phases (Tr, Di, Ts), i.e., 19/28 of all epoch-pair comparisons which involve samples from different task phases may still yield significant separation, as the major task phases extend over tens of seconds (explaining the rather small, although significant, effect sizes between consecutive steps in Fig. S6A). In contrast, iFR correlations started to significantly decay for $\sigma_{\text{shift}} \geq 1$ bins (Fig. S6B), suggesting that the correlational patterns among units are highly sensitive to a shift of the timing of the instantaneous firing rates. This is even more remarkable as these correlations were averaged across all conditions and neural pairs, including probably many uncorrelated pairs which dilute the overall effect, and as the shifting function had a mean of zero, i.e., both positive and negative (integer number) shifts could occur with equal likelihood, thus potentially preserving many of the original temporal relations in the shift surrogates (at least for small $\sigma_{\text{shift}}$).

Fig. S1. (A) A schematic of the spatial layout of microarrays and a sagittal section of the rat ACC showing the location of recording wires (gray box). A coronal section of rat PFC showing a representative placement with the tip of the arrow identifying the recording site in the ACC (Upper Right). (B) Three-dimensional projection derived from the waveform properties of two cells isolated from a single wire. (C) The path of a rat during the test phase of a representative trial (blue line), and the orientation of the animals' body (yellow oval) and head position (brown square). (D) The radial arm maze is shown in green with the path of the animal superimposed in black. In this trial the animal made four correct choices in the test phase and the choices are arranged left to right in a row. The path/position of the animal is shown only for the period when the iFR of one of the five cells deviated significantly from the overall mean iFR. These paths correspond to the test phase iFR shown in Fig. 1E.
Fig. S2. Two-dimensional plots of each pair of axes from the MUA space plots in Fig. 1B and C.
Fig. S3.  (A) Percentage of individual trials on which a significant separation between task epochs could be achieved, as a function of task epoch-pair comparison. (B) Mahalanobis distance between the mean iFR vectors from each pair of task epochs for original and surrogate data (this measure is basically a multivariate equivalent of $d'$ in signal detection theory). These Mahalanobis distances are significantly larger for the original than for the surrogate data, indicating significantly better separation, for all epoch-pair comparisons except for TsC vs. Tsl. This is in perfect agreement with the results of the separation error analysis. (C) Likewise, multiple regression $R^2$ is significantly larger for all original compared to surrogate task-epoch pairs except for the Tsc-Tsl comparison.
Fig. S4. Four examples of task-epoch-selective iFR correlations. (A) Example of a cell pair with a positive iFR correlation during TrC only. (B) Example of a cell pair with a positive iFR correlation during TsR only. (C) Example of a cell pair with a negative iFR correlation during TsC only, yet a positive correlation during TrR. (D) Example of a cell pair with a negative iFR correlation during both correct choice epochs (TrC and TsC, as well as TsR).
Fig. S5. Comparison of within- to across-task epoch iFR correlations, corrected by mean absolute iFR correlations within shuffled surrogates. Black bars: Average absolute corrected iFR correlation within each of the 8 task epochs. Gray bars: Average absolute corrected iFR correlation within surrogates compiled by recombining segments from all 8 task epochs at random (see Methods for details). Error bars = SEM. Within task epoch correlations were significantly higher than across task epoch correlations for all epochs ($P < 0.005$) except for TsR ($P < 0.1$).
Fig. S6.  
(A) Deterioration of MUA space separability as a function of the degree of temporal shifting of the neural iFR(t) vectors (in terms of the standard deviation, $\sigma_{\text{shift}}$, of the average shift in units of 200 ms iFR bins). The differences to the baseline condition ("0", the original data) became significant for $\sigma_{\text{shift}} \geq 3$, and each difference among consecutive steps became significant as well. 
(B) Decay of absolute iFR correlations as a function of $\sigma_{\text{shift}}$. Correlations were averaged across all pairs of each data set (likely to include many uncorrelated pairs) and all task epochs. All differences to the original data were significant for $\sigma_{\text{shift}} \geq 1$, and the first couple of consecutive differences were significant as well as indicated by the stars. Error bars – SEM.
Movie S1. 3D rotational animation of Fig. 18.
Movie S2. 3D rotational animation of Fig. 1C.