Behavioral/Cognitive

Amphetamine Exerts Dose-Dependent Changes in Prefrontal Cortex Attractor Dynamics during Working Memory

Christopher C. Lapish,[1]* Emili Balaguer-Ballester,[2,4]* Jeremy K. Seamans,[3] Anthony G. Phillips,[3] and Daniel Durstewitz[4,5]

[1]Department of Psychology, Stark Neuroscience Institute, Institute for Mathematical Modeling and Computational Sciences, Indiana University—Purdue University Indianapolis, Indianapolis, Indiana 46202, [2]Faculty of Science and Technology, Bournemouth University, Poole, BH12 5BB, United Kingdom, [3]Department of Psychiatry, Djavad Mowafaghian Centre for Brain Health, University of British Columbia, Vancouver, British Columbia, Canada, V6T 1Z3, [4]Department of Theoretical Neuroscience, Bernstein Center for Computational Neuroscience, Central Institute of Mental Health, Medical Faculty Mannheim/Heidelberg University, D-68159 Mannheim, Germany, and [5]School of Computing and Mathematics, Faculty of Science and Environment, Plymouth University, Plymouth, PL4 8AA, United Kingdom

Modulation of neural activity by monoamine neurotransmitters is thought to play an essential role in shaping computational neurodynamics in the neocortex, especially in prefrontal regions. Computational theories propose that monoamines may exert bidirectional (concentration-dependent) effects on cognition by altering prefrontal cortical attractor dynamics according to an inverted U-shaped function. To date, this hypothesis has not been addressed directly, in part because of the absence of appropriate statistical methods required to assess attractor-like behavior *in vivo*. The present study used a combination of advanced multivariate statistical, time series analysis, and machine learning methods to assess dynamic changes in network activity from multiple single-unit recordings from the medial prefrontal cortex (mPFC) of rats while the animals performed a foraging task guided by working memory after pretreatment with different doses of D-amphetamine (AMPH), which increases monoamine efflux in the mPFC. A dose-dependent, bidirectional effect of AMPH on neural dynamics in the mPFC was observed. Specifically, a 1.0 mg/kg dose of AMPH accentuated separation between task-epoch-specific population states and convergence toward these states. In contrast, a 3.3 mg/kg dose diminished separation and convergence toward task-epoch-specific population states, which was paralleled by deficits in cognitive performance. These results support the computationally derived hypothesis that moderate increases in monoamine efflux would enhance attractor stability, whereas high frontal monoamine levels would severely diminish it. Furthermore, they are consistent with the proposed inverted U-shaped and concentration-dependent modulation of cortical efficiency by monoamines.

Key words: multiple single-unit recordings; dopamine; neural computation; neural dynamics; neuromodulation; prefrontal cortex

Introduction

Several psychiatric disorders, including schizophrenia, attention deficit hyperactivity disorder, and bipolar disorder, are characterized by deficits in cognitive function. These deficits are linked to altered neuromodulatory drive (Howes and Kapur, 2009; Seeman, 2011) primarily within the prefrontal cortex (PFC; Goldman-Rakic et al., 2004). Monoamines, such as dopamine (DA), norepinephrine (NE), and serotonin (5-HT), exert power-ful modulatory effects on neural circuits that support higher cognitive functions (Williams et al., 2002; Aston-Jones and Cohen, 2005a,b; Durstewitz and Seamans, 2008). Each of these neuromodulators increases delay- and response-related neural firing, while suppressing background activity of PFC neurons in primates performing a working memory task (Sawaguchi and Goldman-Rakic, 1994; Williams et al., 2002), sometimes interpreted as an increase in the signal-to-noise.

A central feature of drugs that alter monoamine function in the cortex is their inverted U-shaped dose–response profile. With respect to working memory, performance is improved by low doses of D_1 agonists, whereas high doses impair performance (Sawaguchi et al., 1990a,b; Arnsten et al., 1994; Sawaguchi and Goldman-Rakic, 1994; Seamans et al., 1998; Aujla and Beninger, 2001). NE also appears to affect physiology and behavior in an inverted U manner. At lower concentrations, NE acts on $\alpha2_A$ adrenoceptors to increase "signals" in working memory tasks, whereas high levels, acting on $\alpha1$ receptors, generally suppress cell firing (Aston-Jones and Cohen, 2005a,b). Treatment with the indirect monoamine agonist D-amphetamine (AMPH) also reveals an inverted U-shaped dose–response profile, because low

doses (<1 mg/kg) improve but higher doses (>2 mg/kg) impair performance on delay tasks (Aultman and Moghaddam, 2001; Shoblock et al., 2003).

Progress has been made in understanding the biophysical mechanisms that may underlie differential monoamine modulation of "signal" versus "noise" and inverted U-shaped concentration–response functions. Computational models suggest that DA, and to a certain extent 5-HT, exert their functional effects by altering attractor dynamics within the PFC via modulation of various voltage-gated and synaptic currents (Durstewitz et al., 1999, 2000; Compte et al., 2000; Brunel and Wang 2001; Cano-Colino et al., 2013). Specifically, DA appears to stabilize attractor states and impede transitions between them at optimal/intermediate extracellular concentrations (thus enhancing information maintenance). In contrast, at high concentrations, DA weakens attractor states, thereby facilitating transitions and enhancing cognitive shifting and flexibility (Durstewitz and Seamans, 2008). These predictions are derived directly from *in vitro* electrophysiological findings and are supported indirectly by the differential modulation of cortical neuron activity *in vivo* by monoamines. However, this hypothesized regulation of PFC attractor dynamics through monoamines has never been demonstrated directly.

Previous work by our group characterized task-related neuronal ensemble dynamics in the medial PFC (mPFC) during working memory and decision-making using a delayed spatial win–shift (DSWSh) task (White and McDonald, 2002). Performance on this task was impaired after transient inactivation of the mPFC or blockade of mPFC dopaminergic receptors (Seamans et al., 1995, 1998). Multiple single-unit recordings revealed that mPFC networks move through discrete activity state patterns that exhibit attractor-like dynamics during each cognitively defined epoch of the task (Lapish et al., 2008; Balaguer-Ballester et al., 2011). The current study uses these techniques to test directly the computationally derived hypothesis that different AMPH levels differentially regulate PFC attractor dynamics. AMPH was chosen because it enhances monoamine efflux in the PFC (Berridge and Stalnaker 2002; Pum et al., 2007) and has also been used to alter cognitive function and brain activity in humans (Mattay et al., 2003; Tipper et al., 2005).

Materials and Methods

Pharmacology. AMPH was obtained from Sigma and dissolved in 0.9% sterile saline (SAL) before injection. All injections were intraperitoneal and given in a volume of 1.0 ml based on the weight of the animal.

Electrophysiology and behavior. All animals in this study were male Long–Evans rats treated in accordance with the ethical guidelines endorsed by the University of British Columbia and Canadian Council for Animal Care. For a detailed description of the surgical and probe-making procedures, see Lapish et al. (2008). Static single-wire probes were fabricated from 25 μm tungsten wire (California Fine Wires) and arranged in a 2 × 12 array with 150 μm spacing between each recording site. The length of the array was implanted along the rostral–caudal axis of the anterior cingulate cortex, with electrode arrays centered on the following coordinates: anteroposterior, 2.2 mm; mediolateral, 0.8 mm; dorsoventral, −2.5 mm relative to bregma. Animals were placed on a reverse light/dark cycle on arrival in the colony and given *ad libitum* access to food for 1 week. Surgery required for electrode implantation was then performed, and the animals were allowed 2 weeks to recover before training. All electrophysiological data were acquired via a Neuralynx recording system and analyzed offline. Behavioral data were captured via a video camera (Cohu) and recorded in Noldus Ethovision (Noldus) and was synchronized with the electrophysiological recording system and then scored offline. The running speed and behavioral path of each animal were extracted from the movie, but, because of distortion in the

movie, accurate measurements could only be determined for 6 of 10 animals in the experiment.

All animals were trained on the DSWSh task using an eight-arm radial arm maze (Fig. 1A). Each trial consisted of a training and a test phase separated by a 1 min delay. The training phase commenced by opening four of eight arms to allow acquisition of a single sugar pellet (Noyes) from each arm. After retrieval of the final sugar pellet in the training phase, the animal was restricted to that arm with the lights extinguished for the duration of the delay. After the delay, the test phase began with all eight arms open for exploration. Errors were scored as re-entries into previously visited arms. Animals received one trial per day until they reached a criterion of one error or less for 2 consecutive days. On the next day, animals received injections of SAL or AMPH. Behavioral training commenced again on the following day until the animals met a criterion of one error or less for 2 consecutive days, whereupon the second drug injection was delivered in a counterbalanced manner.

Behavioral epochs were defined according to Lapish et al. (2008) and Balaguer-Ballester et al. (2011) and were informed by previous literature describing macro and micro choices (Brown, 1992). The entire 60 s delay was considered as a single epoch in all analyses. For the choice and reward epochs, spike trains were collected for 2 s periods surrounding each behavioral event. A choice was recorded when the approach toward a specific arm was initiated, and a reward was noted when the animal's nose reached the food cup. The width of these time windows were chosen to strike an optimal balance: they are sufficiently wide to allow for statistically robust characterization of neural trajectories as they move into and out of epoch-specific states yet not too wide to cause too much overlap with preceding or subsequent states or other confounding events. Approach behavior was characterized by an orienting response toward a specific arm and subsequent entry into it. Choice and reward epochs were distinguished by whether they occurred during the training or test phases of the task, because these phases presumably entail different cognitive and memory requirements (for more details, see Lapish et al., 2008).

An initial experiment was performed in rats not implanted with tetrode arrays to characterize the dose–response relationship between AMPH dose and task performance. After criterion performance was attained, groups of rats received injections of SAL or AMPH 15 min before the task. After these experiments, two additional groups of rats with electrode implants were trained on the task and, after reaching criterion levels of performance, received counterbalanced injections of SAL and 1 mg/kg AMPH or SAL and 3.3 mg/kg AMPH.

Construction of neural state spaces and statistical analyses. Analyses were restricted to datasets that contained at least eight units resulting in a total of $n = 10$ animals that either received counterbalanced SAL/1.0 mg/kg AMPH ($n = 5$) or counterbalanced SAL/3.3 mg/kg AMPH ($n = 5$). A total of 380 units were recorded in these animals. Ensemble sizes (low-dose SAL: 18, 16, 28, 27, 18; low-dose AMPH: 17, 9, 40, 28, 15; high-dose SAL: 11, 12, 18, 12, 20; high-dose AMPH: 14, 22, 20, 10, 25) were not significantly different across treatments conditions (Kruskal–Wallis test, $\chi^2_{(3,16)} = 2.49, p > 0.47$). Although a subset of animals in this study was included in the studies by Lapish et al. (2008) or Balaguer-Ballester et al. (2011), each specific dataset used here is reported for the first time, i.e., comes from recording sessions different from those reported previously. Units acquired during each recording session were treated as independent samples, because they could not be confirmed unequivocally as being from the same neurons across treatments. Spike trains from the m simultaneously recorded units were convolved with Gaussian kernels with bandwidths (variances) optimized by multivariate kernel density estimation as described by Duong and Hazelton (2005). This method yields spike density estimates that optimize a statistical error criterion based on unbiased cross-validation (for similar approaches, see Yu et al., 2009, Omi and Shinomoto, 2011). Spike density estimates were then binned such that 90% of all bins contained at most one spike, resulting in a bin width of 100 ms used for all animals. All single-unit spike densities were combined into m-dimensional population vectors, $\mathbf{v} = \{v_1(t), v_2(t), \ldots, v_m(t)\}$, with components $v_i(t_j)$ for each unit i as a function of time bin t_j (Lapish et al., 2008; Balaguer-Ballester et al., 2011). Units for which $\langle v_i \rangle$ < 2% of the most responsive unit (the one with maximum average firing
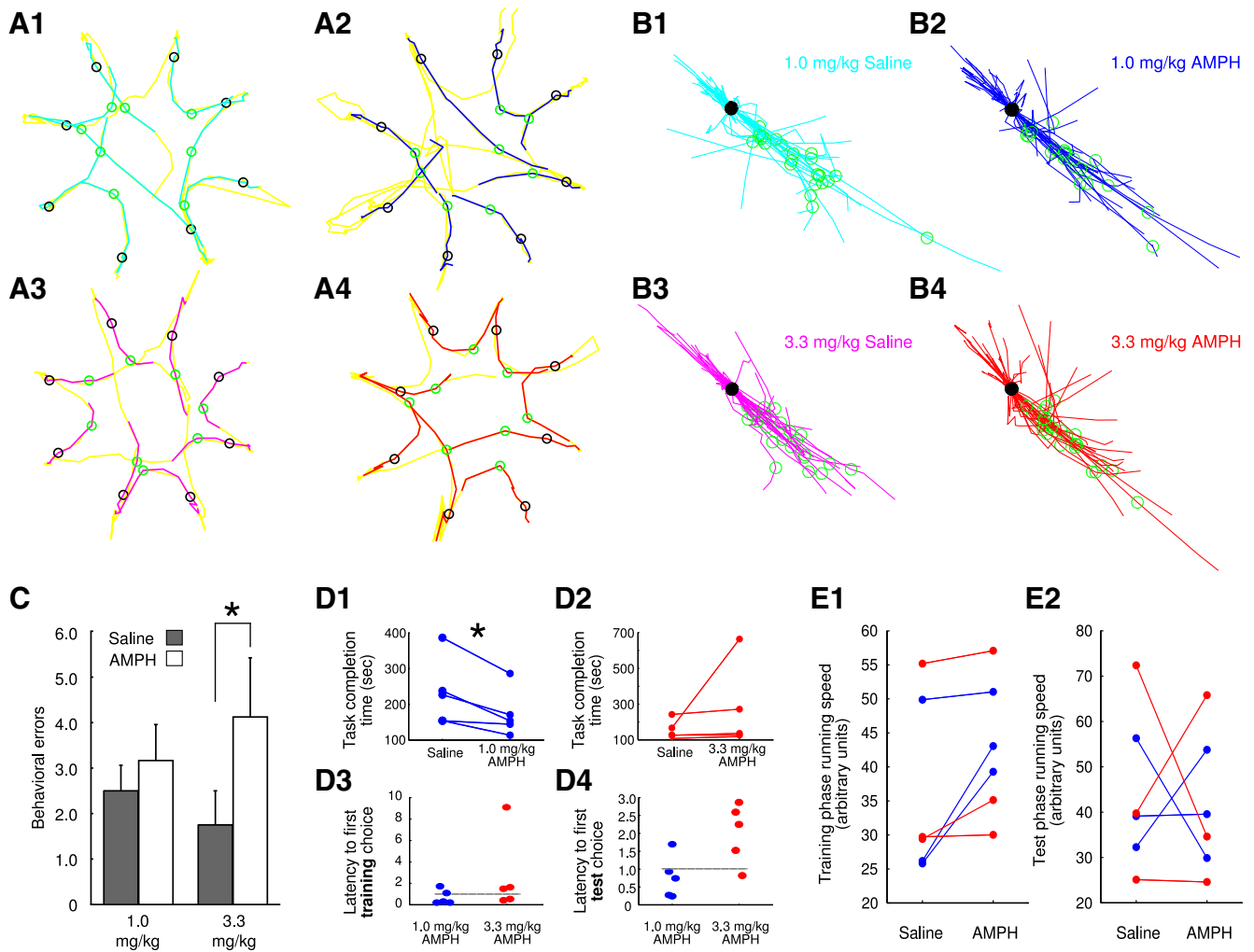
Figure 1. Effects of AMPH on behavior. ***A***, The path of a SAL-treated animal (***A1***) and the same animal treated with 1.0 mg/kg AMPH (***A2***) is shown. Similarly, the path is shown from an animal given SAL (***A3***) or 3.3 mg/kg AMPH (***A4***). In these panels, the path of the animal during the choice and reward epochs is highlighted in cyan, blue, magenta, or red, corresponding to treatment group. Behavior that does not belong to the choice or reward epoch is shown in yellow. The point in the trajectory that was assigned as the "choice point" is denoted with the green circle, whereas the reward point is denoted by the black circle. The delay and incorrect choice epochs are omitted from these figures for clarity. In ***B***, every test and training phase correct choice is shown for six animals. As previously, the color of the trajectory denotes treatment group, and the circles denote the choice (green) and reward (black) points. To visualize the approach of the animal to the reward, the location was rotated as if each reward was in the same location, and, as such, no appreciable changes in approach behavior were detectable. ***C***, The 1.0 mg/kg dose of AMPH had no significant effect on behavioral errors during the DSWSh. In contrast, the 3.3 mg/kg dose increased errors relative to SAL injection. In ***D***, the time to complete the task in the 1.0 mg/kg group (***D1***) and the 3.3 mg/kg group (***D2***) was assessed. Each line represents an animal. Decreased task completion time was observed in the 1.0 mg/kg group (*$p < 0.05$, paired t test). No detectable differences were observed in the normalized (AMPH/Saline) latency to first training phase choice (***D3***) or the first test phase choice (***D4***). In ***E***, running speed was assessed across each treatment group during the training (***E1***) and test (***E2***) phases. A trend toward an increase in running speed was observed in the training phase ($p = 0.07$, paired t test) but not in the test phase.

rate) were excluded for statistical robustness. Estimated spike densities from AMPH conditions were transformed such that the average rate $\langle v_i \rangle$ and the range of v_i were exactly the same for each unit under AMPH and the matched SAL condition i.e., $\langle v_i \rangle^{AMPH} = \langle v_i \rangle^{SAL}$; $\max(v_i^{AMPH}) - \min(v_i^{AMPH}) = \max(v_i^{SAL}) - \min(v_i^{SAL})$. Specifically, AMPH neuron data were first scaled to the same range as the corresponding SAL neuron data and then shifted by a constant to remove the mean difference as follows:

$$v_i^{AMPH} = \left\{ \frac{(v_i^{AMPH} - \min(v_i^{AMPH}))}{\max(v_i^{AMPH}) - \min(v_i^{AMPH})} \right\} \times (\max(v_i^{SAL}) - \min(v_i^{SAL})) + \min(v_i^{SAL}), \quad (1)$$

$v_i^{AMPH} = v_i^{AMPH} - \langle v_i^{AMPH} \rangle + \langle v_i^{SAL} \rangle$. This transformation served merely to ensure that velocity vectors (see below) were comparable directly across different treatment conditions and to rule out any trivial dependence on rate differences among treatment conditions. However, to ensure that results did not depend on the kind of transformation

performed, two additional tests were run: (1) rather than upscaling AMPH rates to match those of the corresponding SAL rates, the latter were downscaled by the same procedure to match those of the AMPH units (i.e., AMPH and SAL were swapped in Eq. 1); (2) rather than just rescaling rates, spikes were added explicitly to the AMPH units drawn at random from the corresponding spike density functions of the units (or, alternatively, from the spike density functions of the matched SAL units) to artificially enhance the spike rate in the AMPH treatment conditions. None of these procedures changed notably the results. In fact, running the separation error (SE) statistics for proportions of added spikes of 0 (no change), 10, 20, or 40% did not reveal any differences between these conditions across different expansion orders (Kruskal–Wallis test, $\chi^2_{(3,36)} = 3.77$, $p = 0.288$), nor were results for "reversely" (from AMPH to SAL) rescaled data different from those with rescaling as given in Equation 1 across expansion orders (Kruskal–Wallis test, $\chi^2_{(1,18)} = 0.46$, $p = 0.49$).

In the following, the original space of the m single-unit recordings is referred to as the multiple single-unit activity (MSUA) space. As done

previously (Balaguer-Ballester et al., 2011), this space will then be expanded by adding the following new dimensions to it: (1) time-lagged versions $v_i(t - \tau_1), \ldots v_i(t - \tau_p)$ with lags τ (fixed at 100 ms) of each original variable $v_i(t)$ (referred to as delay-embedding space of order $O = 1$); and (2) products of the original variables, for instance, $v_1(t) \times v_2(t) \times v_3(t)$ or $v_2(t) \times v_4(t)^2$, up to some specified order $O > 1$ (referred to as O^{th} order delay-interaction space). The underlying rationale is the following: the time-lagged terms $v_i(t - \tau)$, also called a delay embedding in nonlinear dynamics (Takens, 1981; Sauer and Yorke, 1993; Kantz and Schreiber, 2004), expand the original space of firing rate variables by additional axes that contain information about the system dynamics not captured by the current state of the observed variables alone. Intuitively, it may be clear that the underlying neural system may have arrived at the current state of observed firing rates through many different routes, whereas considering more and more of the history of firing rate vectors that led to the current state will more and more constrain this space of possibilities. This allowed us to disentangle more detailed aspects of the precise system dynamics (like the convergence to attractor states) that were not accessible from the current state alone [Takens, 1981; Sauer and Yorke, 1993; these ideas have been formalized by a number of theorems in nonlinear dynamics and evaluated on multiple single-unit recordings by our group previously (Balaguer-Ballester et al., 2011)]. A preliminary analysis using autoregressive (AR) models on the firing rates of the units further revealed that one-time lag was optimal for predicting the future rate values of the units, and that either just using the synchronous rates of the other units as predictors or including higher-order lags in the AR models deteriorated prediction performance (a sign of overfitting). This result suggests that inclusion of just one lag term per unit in the delay-embedding space is sufficient, in line with our previous observations that one-time lag is sufficient for disambiguating neural trajectories (Balaguer-Ballester et al., 2011).

Whereas the first step (delay embedding) helped with "disentangling" neural trajectories through additional dynamical information in the neural state space, the second step of adding instantaneous firing rate (iFR) cross-product (interaction) terms to the space, also called a "multinomial basis expansion" in the statistics literature (Schölkopf et al., 1998; Hastie et al., 2009), addressed another issue: attractors (attracting manifolds) may be highly nonlinear and convoluted objects in the original neural state space, and, for these reasons, it may be difficult to establish their existence or a directed (convergent) flow toward these objects. Basis expansions (adding functions of the recorded variables to the model) have been used amply in statistics to achieve better separability of nonlinear object classes in the expanded space (for instance, N points from two classes nonlinearly entangled in a lower-dimensional space can always be perfectly linearly separated in a $N - 1$ dimensional space, provided they do not lie on a lower-dimensional plane). We have demonstrated previously that, in the present neurophysiological context, basis expansions help to distinguish different attractors of the neural dynamic (Balaguer-Ballester et al., 2011). In principle, many different types of functions of the original firing rate variables may be used in a basis expansion, but here we chose simple cross-product terms because these additional axes have a straightforward biological interpretation in terms of comodulation of neural iFRs. More precisely, an Oth-order unit-interaction term is defined by (omitting lags for notational convenience) the following:

$$\varphi(t) \equiv \varphi(t, m, o): = \prod_{i=1}^{m} v_i(t)^{k_i} \text{ such that } o = \sum_{i=1}^{m} k_i, \quad (2)$$

where the k_i are the exponents of the iFR variables. The Oth-order delay-interaction space consists of all combinations of Oth-order firing-rate products φ with orders $o = 1 \ldots O$. Vectors in this high-dimensional space will be denoted by Φ. For instance, a vector in the space corresponding to $O = 2$ is defined by the binomial expansion:

$$\Phi(t) = \{ \sqrt{2} \times v_1(t), \ldots, \sqrt{2} \times v_m(t), \sqrt{2} \times v_1(t)v_2(t), \ldots, \sqrt{2}$$
$$\times v_1(t)v_m(t), \ldots, v_1^2(t), \ldots, v_m^2(t)\}. \quad (3)$$

The constant factors in this expansion permit the product of any two vectors in the expanded space to be represented by a simple function

of the original variables (i.e., without explicitly expanding the space), termed a kernel function. The so-called "kernel representation" (Schölkopf et al., 1998) thus tremendously facilitates and eases numerical issues associated with computations in these potentially very high-dimensional vector spaces (e.g., $\sim 10^5$ dimensions for $m = 30$ neurons and $O = 3$; for additional details on this approach, see (Balaguer-Ballester et al., 2011). In summary, by the first step of adding lagged variables, neural trajectories would be disambiguated, whereas the second step of augmenting the state space by axes defined through the multinomial basis expansion (O) above will enable the linear separation of classes with nonlinear boundaries.

As noted above, dimensions in the expanded embedding spaces can be interpreted in terms of cross-products of unit firing rates. To make this relation more explicit and hence lend better biological interpretability to some of the results obtained with the basis expansions, we also studied differences between the treatment conditions with respect to second-order (Pearson's) and higher-order cross-correlations. The classical Pearson correlation coefficient between firing rates v_1, v_2 of any two units is defined as follows:

$$C(v_1, v_2) = \frac{\sum_t (v_1(t) - \bar{v}_1)(v_2(t) - \bar{v}_2)}{\sqrt{\sum_t (v_1(t) - \bar{v}_1)^2} \sqrt{\sum_t (v_2(t) - \bar{v}_2)^2}}. \quad (4)$$

Likewise, a third-order correlation coefficient among three units is given by

$$C(v_1, v_2, v_3)$$
$$= \frac{\sum_t (v_1(t) - \bar{v}_1)(v_2(t) - \bar{v}_2)(v_3(t) - \bar{v}_3)}{\sqrt[3]{\left(\sum_t (v_1(t) - \bar{v}_1)^3 \sum_t (v_2(t) - \bar{v}_2)^3 \sum_t (v_3(t) - \bar{v}_3)^3 \right)}}, \quad (5)$$

and so on for higher-order terms. Product terms, such as $v_1 v_2$ or $v_1 v_2 v_3$ that appear in the definition of these qth-order Pearson-type correlation coefficients also constitute dimensions in the expanded state space. That is, many of the dimensions in the expanded spaces are similar to multi-unit correlations, only that they are not mean corrected (centered) and not standardized (by the variance terms in the denominator).

Analyses in the expanded embedding spaces were performed by means of a regularized kernel-Fisher's discriminant analysis (kernel-FDA; Mika et al., 2000; Yang et al., 2004; Saadi et al., 2007) and regularized kernel-principal component analysis (kernel-PCA; Schölkopf et al., 1998). Regularization techniques (Schölkopf et al., 1998; Hastie et al., 2009; Witten and Tibshirani, 2011) are used commonly in statistics to minimize out-of-sample prediction errors (avoiding overfitting to the data at hand) by penalizing the dimensionality of the space (the degrees of freedom of the covariance matrix). The degree of regularization is defined by a single parameter, λ, that was obtained through cross-validation (Hastie et al., 2009) from a previous study with multiple trials from the same task as used here (Balaguer-Ballester et al., 2011) and held constant for all analyses presented here (because only single trials were available in the current study, across-trial cross-validation was not possible based on the current data alone). kernel-FDA (equivalent to ordinary FDA for $O = 1$) extracts a lower-dimensional subspace from the expanded high-dimensional state space with axes that provide optimal separation between predefined sets (classes) of data points (for c classes, this space will be $c - 1$ dimensional). FDA is a linear transform similar to PCA, only that dimensions are not the most variance-explaining ones as in PCA but rather the ones along which the overlap between c class-specific distributions of data points is minimized. Here, kernel-FDA was used to visualize neural dynamics in a three-dimensional (3D) space (by retaining the three most discriminating directions from the expanded space), as well as for obtaining a (linear) classifier based on the projection of the data points onto the one most-discriminating axis extracted. In general, discriminant analysis classifiers are Bayesian optimal if the data come from multivariate normal distributions. Because projection of the data from a high-dimensional space onto a single axis leads to sums of many random variables (although not independent ones), normal assumptions hold

approximately by virtue of the central limit theorem (as shown by Balaguer-Ballester et al., 2011). Given normal distribution assumptions, for each population pattern, $v(t)$ (i.e., the location of the trajectory in state space at time t), one can compute the posterior probability $P(C|v(t))$ that the animal is currently in task epoch C, given that delay-interaction pattern. Based on these posterior probabilities, population patterns $v(t)$ were assigned to task epochs, and a measure of separation among neural trajectories from different task epochs was defined as the proportion of misclassified data points, the SE herein (Lapish et al., 2008; Balaguer-Ballester et al., 2011).

For statistical significance testing of the classification performance captured by the SE and for testing attractor-like properties described below, nonparametric bootstrap approaches were used, with bootstrap samples constructed by randomly shuffling non-overlapping stretches of the time series that retained entire trajectories from a given task epoch (i.e., block permutation; performed independently for each animal), such that each sample preserved all temporal autocorrelations up to the length of the relevant task epoch (1000 replications for one-sided comparisons at $p = 0.001$; Efron and Tibshirani, 1993; Lapish et al., 2008; Balaguer-Ballester et al., 2011). Importantly, bootstraps constructed via block permutations may also preserve some statistical properties of the original series other than linear autocorrelations but should destroy most of the relation to the original task-phase labels (because there are a total of six classes here, chance level for correct classification is at ~17%). Attractor-like properties of neural states related to cognitive task epochs were assessed by checking for consistently correct classification of entire task-epoch trajectories: a trajectory that converges to, returns to, or cycles within a supposedly attracting set A in neural state space should terminate in a series of consistently correctly classified points (i.e., that all lie in A). If that was the case for a particular trajectory considered, it was labeled "convergent," otherwise "divergent." As another bootstrap test for this convergence, the temporal sequence of binned firing rates was randomized for all neurons within task epochs, i.e., without destroying assignment of data points to task epochs (in contrast to the bootstraps defined above), but just their temporal order.

The methods used here have been shown previously to be robust with regards to dataset size (number of recorded units), for a range of bin sizes, and for different SDs of the Gaussian smoothing function used to compute spike densities (Balaguer-Ballester et al., 2011). We stress that the multinomial basis expansions are used here primarily as a statistical tool for analyzing (disentangling) trajectory dynamics and not for assessing high-order neural correlations, although, as noted above, here we also attempted to more directly connect these concepts empirically by also studying higher-order correlations directly (additional details and mathematical justifications can be found in the study by Balaguer-Ballester et al., 2011).

Results

Behavior and firing rate measures

A dose–response experiment determined the optimal doses of AMPH for behavioral effects on performance of the DSWSh task in a group of rats not implanted with multielectrode arrays. Each animal received a single injection of SAL ($n = 6$), 1.0 ($n = 8$), 2.0 ($n = 6$), 3.3 ($n = 8$), or 4.0 ($n = 6$) mg/kg AMPH. We observed a dose-dependent effect on errors after AMPH injections (Kruskal–Wallis test, main effect of dose, $\chi^2_{(5,34)} = 27.66$, $p = 0.0001$; normality rejected according to Lilliefor's test, $p = 0.001$). The higher 3.3 mg/kg dose of AMPH produced a significant increase in errors compared with SAL-treated animals (Wilcoxon's signed-rank test, $n = 14$, $W = 22$, $p = 0.002$). Although 3.3 mg/kg may elicit stereotypy (Grilly and Loveland, 2001), in the present study, no detectable stereotypies were observed and animals were still capable of foraging for food, albeit with an increased behavioral error rate. In this dose–response experiment, only a subset of animals (two of six) treated with the highest dose of 4.0 mg/kg were unable to complete the task because of excessive stereotypic behavior. Given these behavioral data, two

doses of AMPH were chosen for subsequent experiments: (1) a lower dose (1.0 mg/kg) that produced no overt behavioral deficits; and (2) a higher dose (3.3 mg/kg), the lowest dose to cause deficits in cognitive performance without inducing stereotypy. Two additional groups of animals were implanted with multiple electrode arrays and given a single injection of either the low or high dose of AMPH, as well as a counterbalanced injection of SAL. Clear movie tracking data were available from $n = 3$, in each of the 1.0 and 3.3 mg/kg conditions, and the path traveled by two representative animals is shown in Figure 1A. Visual inspection of these paths reveals no detectable change in approach behavior to the reward across treatments (Fig. 1B). As was the case for non-implanted animals in the initial dose–response study, implanted animals given 3.3 mg/kg AMPH exhibited a deficit in task performance (Fig. 1C; Wilcoxon's signed-rank test, $n = 5$, $W = 0$, $p < 0.05$, normality rejected according to Lilliefor's test, $p < 0.05$). Performance after the 1.0 mg/kg dose was unaffected (Wilcoxon's signed-rank test, $n = 5$, $W = 3$, $p > 0.29$; normality rejected according to Lilliefor's test, $p < 0.05$) relative to the SAL condition. A decrease in the time to complete the task, measured from the first to last correct choice, was observed for the 1.0 mg/kg group ($t_{(4)} = 3.64$, $p < 0.03$; Fig. 1D1). However, no AMPH-induced changes were observed in the normalized latencies (AMPH latency/SAL latency) to make the first choice of the training or test phase in either treatment group. A trend toward an increased running speed was observed in the training phase between SAL- and AMPH-treated animals (repeated-measures t test, $t_{(5)} = 2.3$, $p = 0.07$; Fig. 1E1,E2).

An analysis of neural firing after AMPH and SAL injections revealed a significant suppression of the mean iFR in the 3.3 mg/kg AMPH condition. In contrast, a small but significant increase in the overall average iFR was observed with the 1.0 mg/kg dose (dose × treatment interaction, $F_{(1,370)} = 8.73$, $p < 0.004$; Fig. 2B). The latter effect is consistent with previous reports of AMPH-induced changes in the firing rate of cortical neurons (Stone, 1976; Homayoun and Moghaddam, 2006; Gulley and Stanis, 2010).

Optimal neural state space under control conditions

We next determined, from the set of m simultaneously recorded units with iFRs $v_i(t)$ in time bin t, the neural state space in which neural trajectories were optimally unfolded and separated using the statistical methods described by Balaguer-Ballester et al. (2011). Here, the term "trajectory" refers to the temporal evolution of neural population activity through a space constructed from the spike rates of the recorded units, i.e., to a series of temporally consecutive vector points in this space. Thus, these trajectories capture how neural dynamics evolve in time with respect to task stages and behavioral events. As explained in detail in Materials and Methods, the neural state spaces considered here may consist of the m single-unit iFRs (termed the MSUA space in the following), the m single-unit iFRs plus additional variables consisting of time-lagged versions of the original iFR variables (referred to as the delay-embedding space of order $O = 1$), or the delay-embedding space including additional variables that describe interactions among units (iFR cross-product terms) up to some specified (expansion) order O (referred to as the Oth-order expansion delay-interaction space). Thus, the delay-embedding space ($O = 1$) will have dimensions corresponding to variables such as $v_i(t - \tau_1), \ldots v_i(t - \tau_p)$ with lags τ, in addition to the original iFR variables $v_i(t)$, whereas expansion spaces with $O > 1$ will contain variables such as $v_1(t) \times v_2(t) \times v_3(t)$ or $v_2(t) \times v_4(t)^2$, consisting of up to O iFR
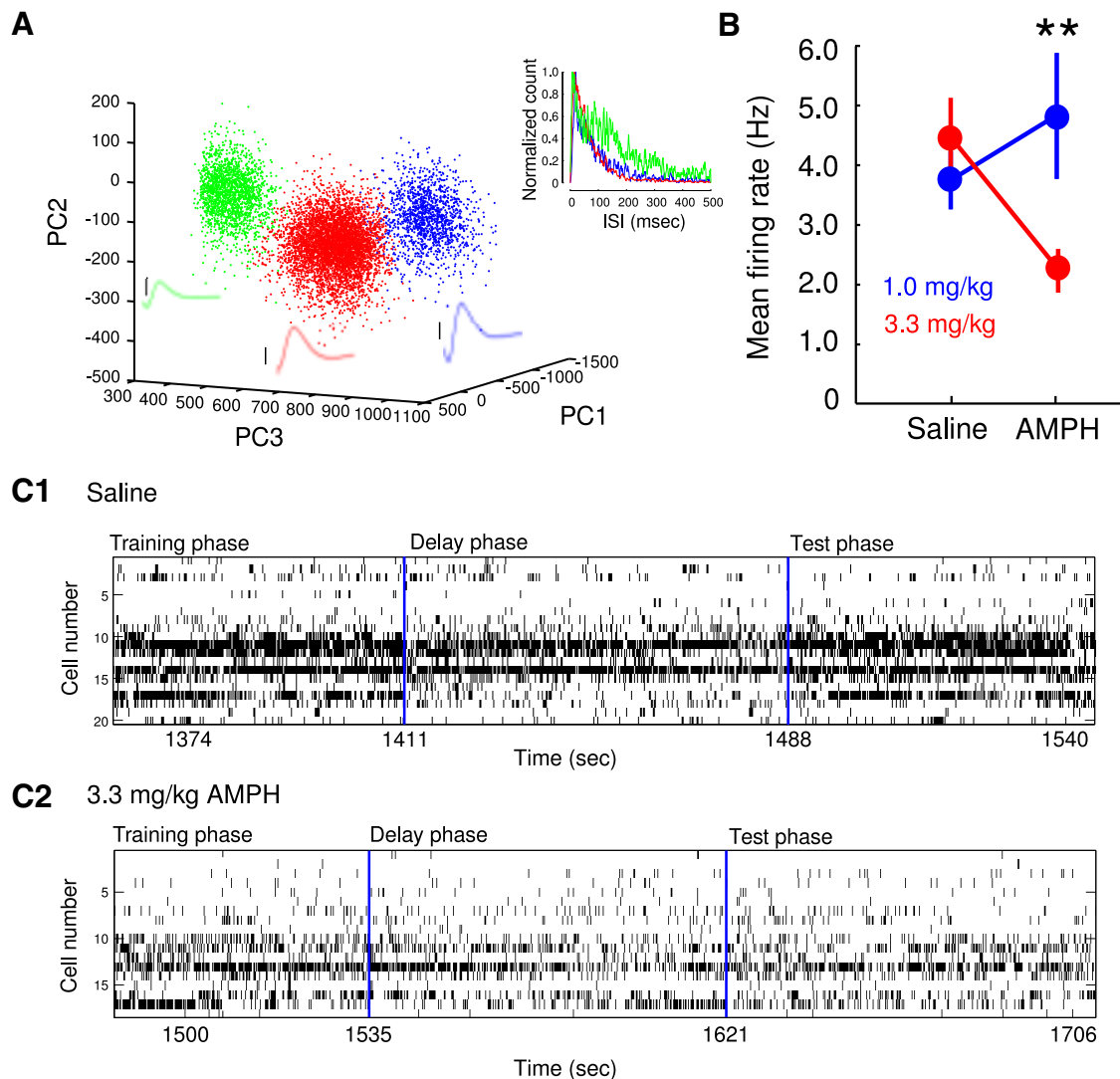
Figure 2. Effects of AMPH on neural firing. **A**, A representative scatter plot from a single electrode, in which three different single units were isolated via spike sorting. The mean waveform corresponding to each color-coded cluster are shown in the scatter plot. Black scale bars next to each waveform, 50 μV. The inset shows the color-coded interspike interval (ISI) distribution corresponding to each unit. **B**, Changes in overall average iFR $\pm$ SEM in the DSWSh task as a function of AMPH concentration. Representative examples of iFRs for all units throughout performance of the DSWSh task for a SAL-treated (**C1**) and a 3.3 mg/kg AMPH-treated (**C2**) animal. Note the general decrease in firing rate across the task after 3.3 mg/kg AMPH. ******$p < 0.01$, ANOVA, treatment $\times$ dose interaction.

product terms, capturing interactions among the units (for exact mathematical definition, see Eq. 2). The rationale behind this is explained in detail in Materials and Methods and by Balaguer-Ballester et al. (2011) and will also become evident from the results presented below.

Figure 3A shows an example of a 3D FDA projection of the original MSUA space consisting only of the current spiking rates of the units (i.e., no cross-product and no delay terms), whereas Figure 3B shows the same data within a 3D projection obtained from the optimally expanded high-dimensional space, by which we mean the space to which unit iFR cross-product terms were added as dimensions up to an order O (see above), such that task-epoch-specific trajectories could be separated optimally (see Materials and Methods, Movies 1, 2). In addition to the actual data points (neural population states), the graphs give the difference vectors between temporally consecutive pairs of data points (arrows), thereby indicating the speed and direction of movement (the "flow field") of the neural population state at each point. (Note that the set of Euclidean distances in the original spaces cannot be fully reproduced in the lower-dimensional pro-

jection, and hence some properties of the flow will inevitably be distorted. In addition, we have omitted time bins that do not correspond to choices and rewards for clarity, introducing some gaps into the representation of the flow.) Although in Figure 3A the flow field is relatively disordered, as indicated by no apparent pattern in the orientation and length of the velocity vectors as a function of task epoch (indicated by the color coding), the picture changes dramatically when flow fields are displayed in the optimally expanded space (Fig. 3B), with the appearance of a clear neural orbit corresponding to the entire task-epoch sequence. Although the neural population state seems to move relatively fast between different task phases, it slows down as it approaches the center of each task phase, indicating that task phases act as semi-attracting (or metastable) entities (Rabinovich et al., 2008; Balaguer-Ballester et al., 2011; Rabinovich and Varona, 2011). This observation is quantified in the inset of Figure 3B, which shows the normalized velocity of the trajectory in this 3D representation versus the distance of the trajectory from the centroid of each task epoch. As the distance from the centroid increases, so does the velocity of the trajectory.
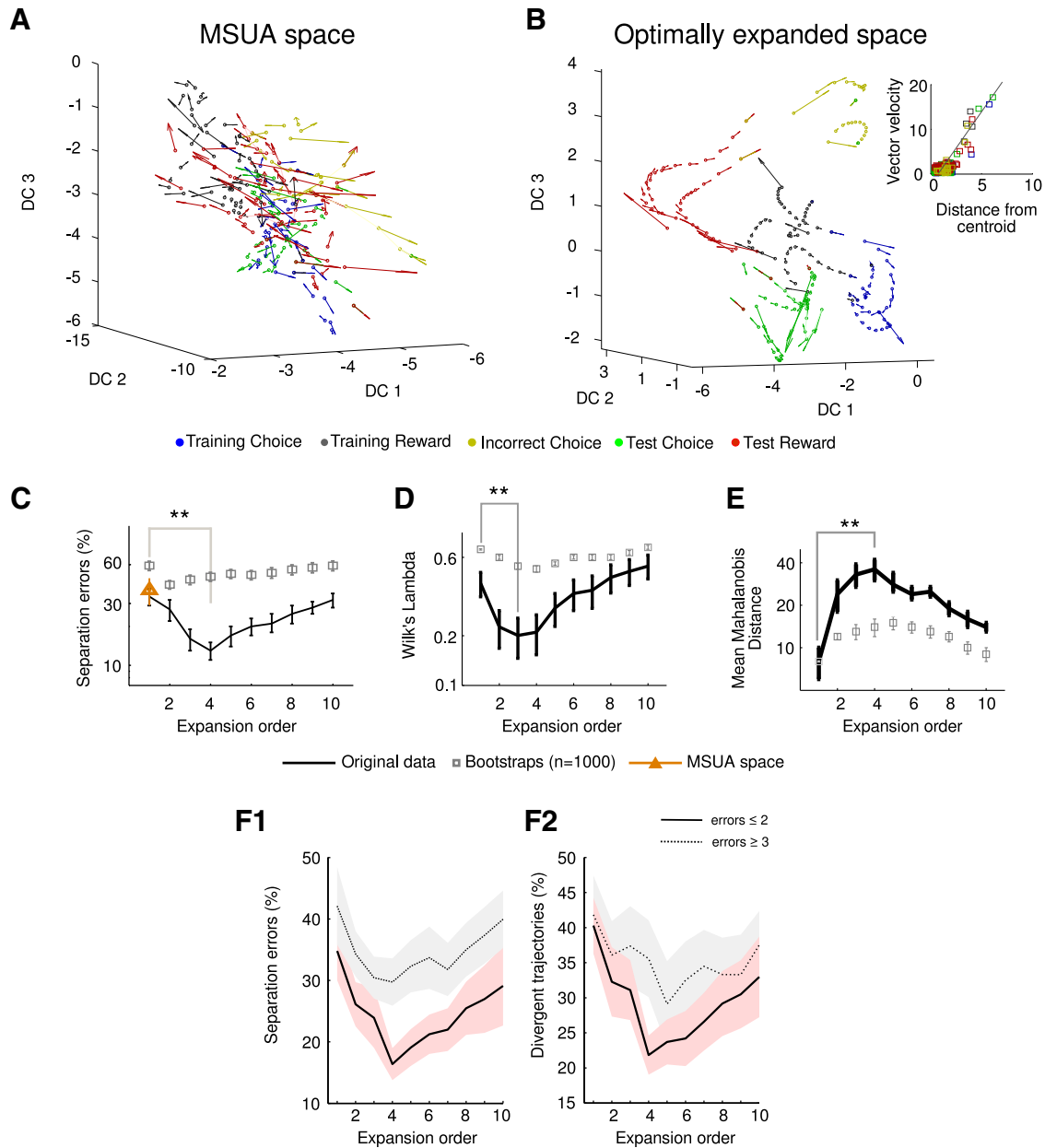
Figure 3. Task-epoch states and trajectories are optimally resolved in expanded state space. ***A***, Neural flow field from the MSUA state space projected onto the three maximal discriminant coordinates (DC) obtained by kernel-FDA (data from 1 SAL-injected animal; dimensions were orthonormalized). In this projection, the flow field is very disordered, occluding the flow of neural trajectories. ***B***, Same as in ***A*** for the optimally expanded space, in which task epochs are optimally resolved and the flow field much more ordered. Note the fast movement of the flow (large velocity vectors) when population activity transits from one task epoch to another, whereas movement slows down (small velocity vectors) toward the center of the task-epoch states. The inset shows the velocity of the population vector as a function of its distance to the corresponding task-epoch centroid, in which the velocity increases the farther away from the centroid. ***C***, SEs (mean ± SEM) as a function of multinomial expansion order *O* initially decrease up to *O* = 4 and then increase again as *O* > 4 (**statistical comparison of orders 1 and 4, Wilcoxon's signed-rank test, *W* = 692, *p* = 0.0004, normality rejected according to Lilliefor's test, *p* < 0.03). Block-permutation bootstraps (gray square markers) differ significantly from the original data. Note that the optimal expansion order is approximately *O* = 3–5. The same effect of expansion order is observed for Wilk's lambda (***D***) and mean Mahalonobis distance (***E***). ***F***, Both SEs and the percentage of divergent trajectories increase in animals that make at least three behavioral errors.

Consistent with Figure 3, *A* and *B*, the SE (as defined in Materials and Methods) is significantly lower for the optimally expanded space than for the original MSUA space (Fig. 3*C*, triangle) and for the *O* = 1 space (i.e., the MSUA space with the addition of lagged iFR values but not iFR cross-product terms). Multiple comparisons across the 10 expansion orders indicate significant differences (Kruskal–Wallis test, $\chi^2_{(9,290)}$ = 32.3, *p* = 0.002; normality rejected according to Lilliefor's tests, *p* < 0.033). Pairwise comparisons indicate that *O* = 1 is significantly different from orders 2 to 6 (Wilcoxon's signed-rank test, *W* < 1150, *p* < 0.03,

Bonferroni-corrected.). For higher-order expanded spaces (*O* > 6), SEs significantly increase again and are similar to *O* = 1, which replicates findings by Balaguer-Ballester et al. (2011).

Statistical measures of task phase separation commonly used in multivariate ANOVA computed on the full discriminant subspace further confirmed optimal embeddings at expansion orders *O* = 2–5 (Fig. 3*D*,*E*). In addition, analyses were conducted with nonparametric bootstraps that retained all features of the original time series that are not related to the structure imposed by the task itself (Fig. 3*C*–*E*, gray squares). For all statistics of
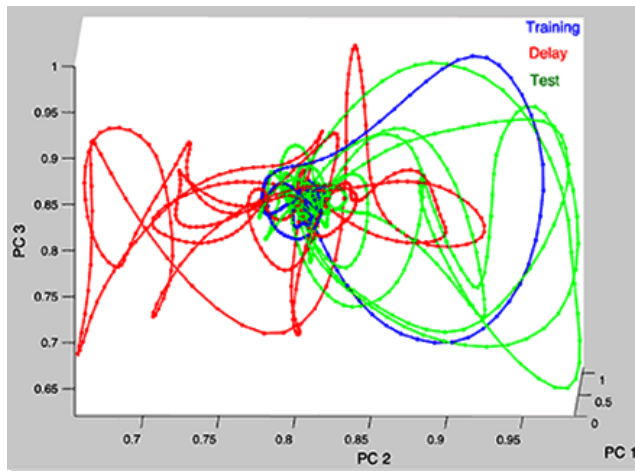
Movie 1. 3D state-space visualizations obtained by kernel-PCA for the optimally expanded space ($O = 4$) following a saline injection.
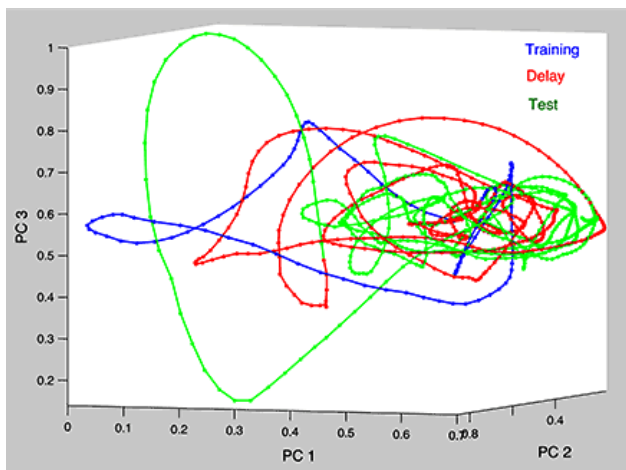
Movie 2. 3D state-space visualizations obtained by kernel-PCA for the optimally expanded space ($O = 4$) following a saline injection in the 3.3 mg/kg AMPH group.

task-epoch separation that we have explored, bootstrap data did not show any notable dependency on the expansion order, which is in contrast to the original data. Also in agreement with our previous work (Lapish et al., 2008; Balaguer-Ballester et al., 2011), an increase in SEs and in the number of divergent trajectories (Fig. 3F) was observed in animals performing poorly (at least three behavioral errors) compared with those performing well (no more than two errors), which indicates that neural state spaces become more disordered when behavioral performance deteriorates.

AMPH alters the structure of neural state spaces in a dose-dependent and bidirectional manner
In this and all following analyses, the firing rates of all single units were normalized to have the same mean and range under SAL and treatment conditions, to yield neural state space representations directly comparable among conditions (and rule out any possibly confounding effects through mere mean rate differences; see Materials and Methods). Indeed, after this transformation, when SE was examined across SAL and AMPH conditions in the original (non-expanded) MSUA space, no significant differences were observed across treatment conditions ($F_{(3,116)} = 1.53$, $p > 0.2$). In contrast, when the optimally expanded spaces were ex-

amined, animals treated with 1.0 mg/kg AMPH exhibited a clear separation of task epochs and ordered neural flow fields (Fig. 4A; see also Movie 3), whereas this organization completely broke down in animals treated with the 3.3 mg/kg dose of AMPH (Fig. 4B; see also Movie 4). With this higher dose, orbits corresponding to the different task phases could not be discerned, even in the optimally expanded space. These differences in state space organization were confirmed statistically across the whole sample (Fig. 4C), yielding a highly significant dose × treatment interaction for SE ($F_{(1,116)} = 14.79$, $p = 0.0002$) within the optimal space. In agreement with predictions of the inverted U-shaped dose–response theory, the mean SE decreased for the 1.0 mg/kg dose of AMPH compared with SAL animals (Wilcoxon's signed-rank test, $W = 283$, $p < 0.02$; normality rejected according to Lilliefor's test, $p < 0.002$), whereas SE was significantly increased in animals treated with the 3.3 mg/kg dose (Wilcoxon's signed-rank test, $W = 599$, $p < 0.03$; normality rejected according to Lilliefor's test, $p < 0.002$), indicating a dose-dependent and bidirectional effect of AMPH on neural state space separation.

Furthermore, although SE reached a significant minimum at expansion orders 3–5 for the 1.0 mg/kg treated animals (Fig. 5A1; comparison of order 1 to order 4, Wilcoxon's signed-rank test, $W = 692$, $p = 0.0004$, normality rejected according to Lilliefor's test, $p < 0.03$), this was not the case for the 3.3 mg/kg condition (Fig. 5A2). Although a minimum still occurred at these orders for 3.3 mg/kg AMPH, this did not differ significantly from the SE at $O = 1$, and, in general, there was no significant improvement in task-epoch separation for $O > 1$ in these animals ($U_{(30)} > 1043$, $p > 0.05$ for all pairwise comparisons). Thus, unlike the clear benefits obtained with the SAL-treated and 1.0 mg/kg AMPH-treated animals, there is no apparent advantage in expanding the space to include higher-order iFR interactions for animals treated with 3.3 mg/kg AMPH. These observations were generally consistent across all separate pairwise comparisons of behavioral epochs in terms of SE (Fig. 6). In summary, although the 1.0 mg/kg dose of AMPH improved separation among all task epochs in the optimally expanded spaces relative to SAL injection, in contrast, the 3.3 mg/kg dose strongly reduced separation for most task-epoch pairs.

To foster the conceptual understanding of the higher-order terms in the basis expansion, we examined higher-order rate correlations more directly in terms of the standard statistical definition of the bivariate Pearson's cross-correlation and its direct higher-order generalizations (see Materials and Methods). Examining the second- and higher-order cross-correlations among units provides additional insights into what underlies the improved state space separation in the 1.0 mg AMPH group, in contrast to the diminished separation for the 3.3 mg AMPH group. Indeed, across the range of orders examined (orders 2–4), these rate correlations were always highest for the 1.0 mg AMPH group and always lowest for the 3.3 mg AMPH group (rank-sum test, $p < 0.0002$; Fig. 5D1). The fact that unit correlations at all orders examined were weak for the 3.3 mg AMPH group explains why expanding the neural state spaces through rate product terms does not help much in improving separation for this condition. Correlations and thus product terms in the basis expansion simply did not add any additional information that would help for separating trajectories for high AMPH or, even worse, may have added noise in this case. Thus, although after a low dose of AMPH network dynamics seem to become more coherent and coordinated compared with SAL, network dynamics appear to become mostly disorganized and incoherent when rats received the higher 3.3 mg/kg dose of AMPH. Additionally, Figure 5D2
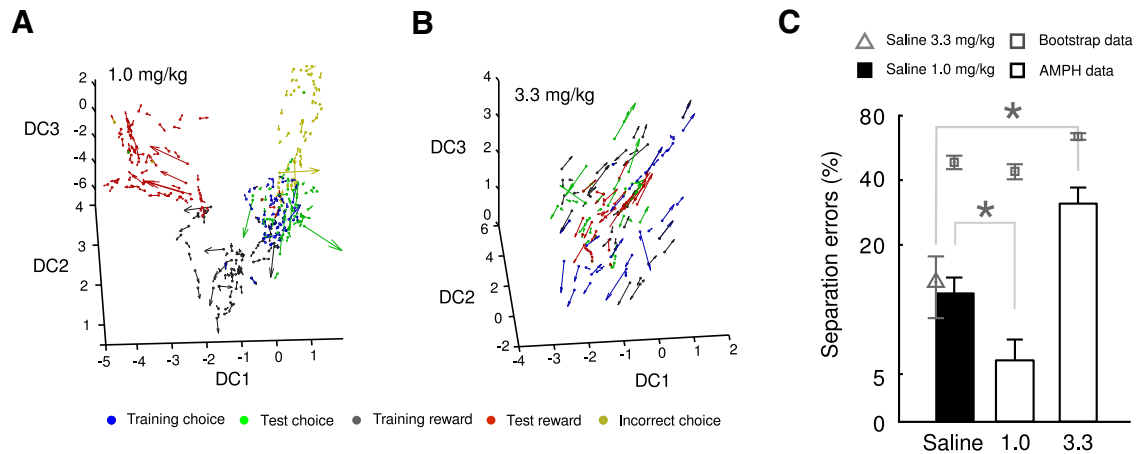
Figure 4. AMPH exerts a bidirectional and dose-dependent effect on task-phase separation in neural state space. **A**, State-space projections onto the three maximal discriminant coordinates (DC) derived from the optimally expanded space in an animal treated with 1.0 mg/kg. **B**, Same as **A** for an animal treated with 3.3 mg/kg AMPH. **C**, SEs by task condition: relative to SAL injection, SEs decrease significantly after the 1.0 mg/kg AMPH dose but increase after the 3.3 mg/kg AMPH dose in the optimally expanded spaces. *$p < 0.03$, Wilcoxon's signed-rank test.
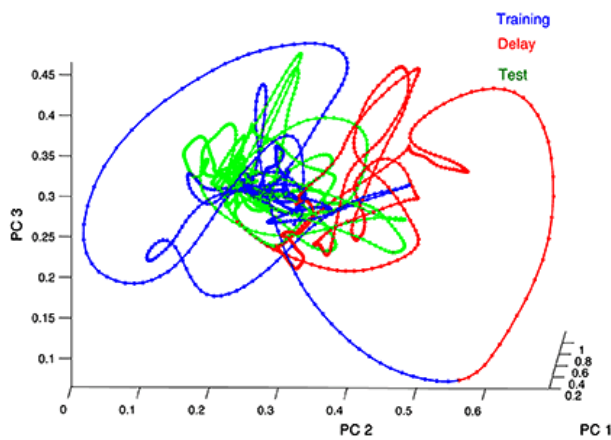
Movie 3. 3D state-space visualizations obtained by kernel-PCA for the optimally expanded space ($O = 4$) following an injection of 1.0 mg/kg AMPH.
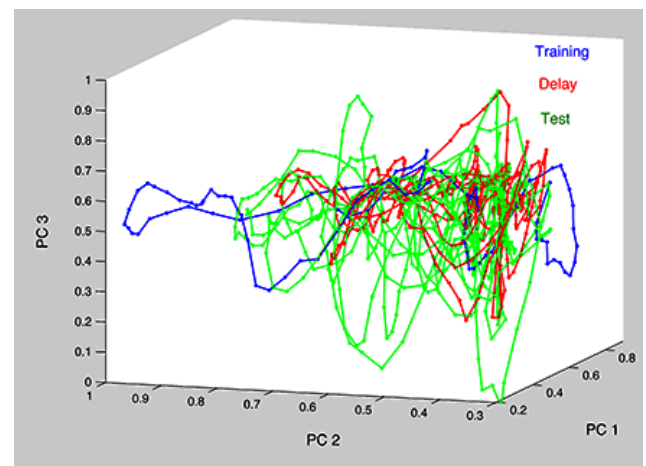
Movie 4. 3D state-space visualizations obtained by kernel-PCA for the optimally expanded space ($O = 4$) following an injection of 3.3 mg/kg AMPH. Notice that trajectories cannot be disentangled as in Movies 1–3.

shows that, for all correlation orders tested ($q = 2, 3, 4$), these were much weaker if the animals commit many behavioral errors (at least three) than if they commit relatively few (no more than two; rank-sum test, $p < 0.005$). These observations are consistent with those in Figure 3F1 and suggest that the worse state space discriminability in the high error group and, in particular the weaker dependence on expansion order, is at least partly rooted in a breakdown of cross-correlations or coordinated activity among units when animals commit many errors, a tendency that we had observed previously (Lapish et al., 2008).

AMPH alters the attractor dynamics of task-epoch-specific states in a dose-dependent and bidirectional manner

Next we studied the differences in flow fields of the task-epoch-specific trajectories for each dose of AMPH. The purpose of these analyses was to characterize attractor-like properties of task epochs by the patterns of convergence or divergence of neural trajectories toward or away from task-epoch-specific population states or orbits. More specifically, task-epoch state boundaries were defined through the optimal kernel-FDA classifier (see above), and convergence or divergence was analyzed with respect to the so-defined states.

Figure 7A shows a schema of different kinds of neural trajectories converging toward their respective task-epoch clusters. A trajectory corresponding to a certain task epoch that terminates in a correctly classified sequence of population vectors (i.e., up to the time point that corresponds to the end of that behavioral epoch) provides an indication that this trajectory converges toward or cycles within the proper task-epoch boundaries, i.e., it indicates "attracting" behavior (red). However, when the final section of the trajectory is incorrectly classified, i.e., when it leaves the task-epoch boundary as defined by the FDA classifier, it is considered to be divergent (yellow). It is important to note that the definition of these task-state boundaries is purely statistical (i.e., based on the discriminant analysis) and does not include any knowledge about convergence or divergence of trajectories. It is also crucial to note that the number of such convergent trajectories does not necessarily imply a low SE. For instance, as shown in Figure 7A, a high SE (40%, 8 of 20 points in the schema) is consistent with a scenario in which all trajectories converge. Conversely, even if all trajectories are divergent, the SE can still be low. Thus, separation and trajectory convergence are complementary views on the neural dynamics.

The analysis can be illustrated in the 3D projections obtained by penalized kernel-FDA from the optimally expanded space:
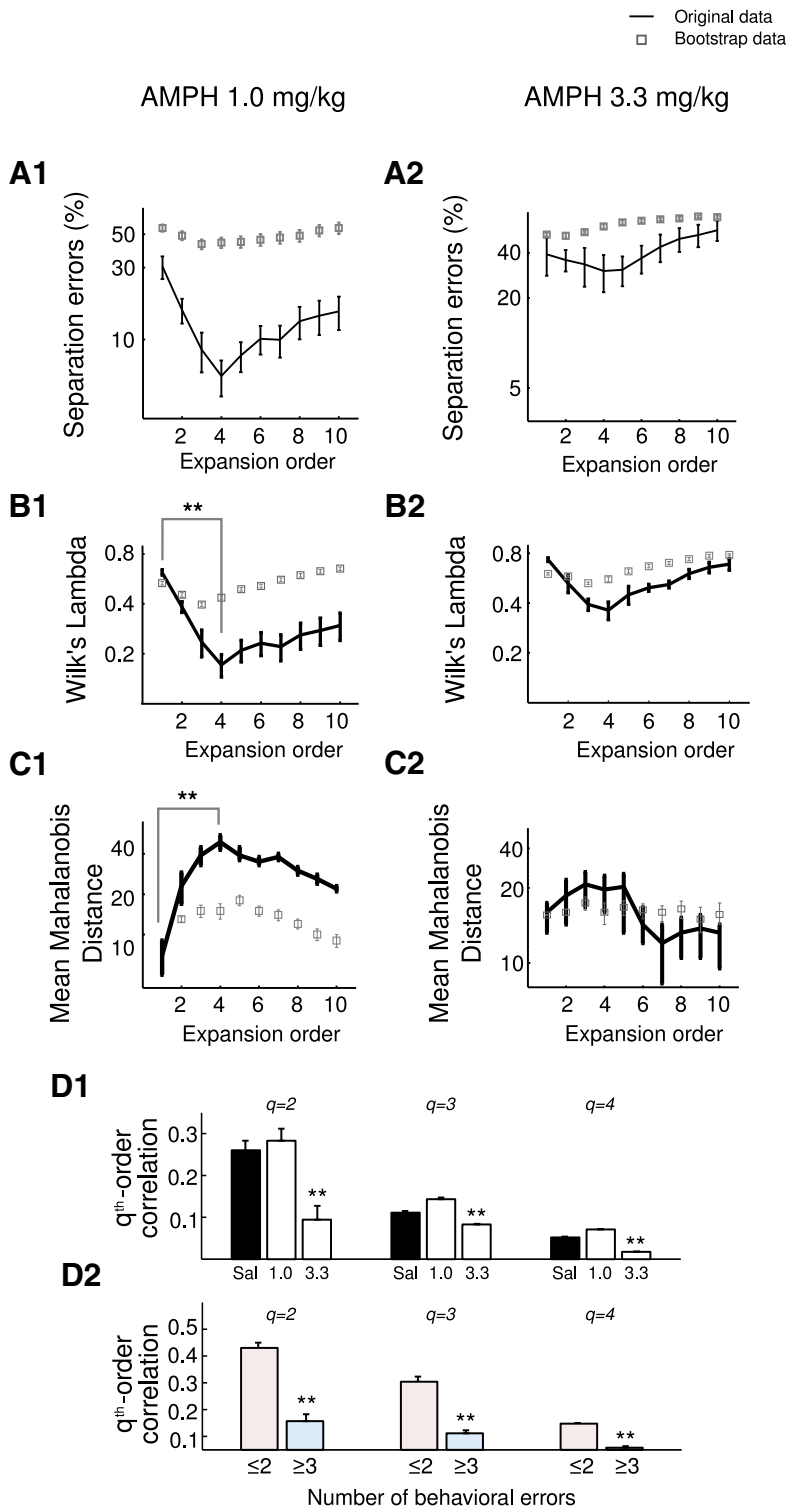
Figure 5. Neural state-space structure is enhanced across expansion orders after a low dose of AMPH but destroyed after a high dose. *A1, A2*, SEs (mean ± SEM) as function of expansion order *O* for the 1.0 mg/kg (*A1*) and 3.3 mg/kg (*A2*) AMPH-treated animals for both the original data and corresponding bootstraps (gray squares). The same analyses for Wilk's lambda (*B1, B2*) and mean Mahalonobis distance (*C1, C2*). Note the clear local minimum in the 1.0 mg/kg AMPH condition compared with the weak minimum in the 3.3 mg/kg AMPH condition. For related analysis, see Movies 1–4. *D1*, Q [th]-order correlations among neurons are diminished after 3.3 mg/kg AMPH relative to a SAL (**$p < 0.01$, rank-sum test). *D2*, Q [th]-order correlations are decreased in animals that make at least three behavioral errors regardless of treatment group (**$p < 0.01$, rank-sum test).

Figure 7*B* shows, for one SAL-treated animal, only convergent neural trajectories, i.e., trajectories that ultimately cycle into or return to a task-epoch-specific population cluster. In this figure, population vectors that are misclassified are marked by a black circle. They tend to be the ones with higher velocities and correspond to transitions of the neural trajectory from one cognitive epoch to another.

We then quantified, within the full high-dimensional spaces, the degree to which the conditions defining attracting states were met. Figure 7*C* shows the fraction of trajectories that escape from task-epoch-specific clusters without returning to them (i.e., divergent trajectories). As before (Fig. 4*C*), a significant dose × treatment interaction was observed for the fraction of divergent trajectories within the optimally expanded spaces (two-way ANOVA for $O = 4$, interaction $F_{(1,116)} = 4.3$, $p < 0.05$). For the two SAL control groups, only a small fraction (~15–20%) of all trajectories were divergent with respect to their task-epoch-specific clusters. Similar to the results for task-phase separation (Fig. 6), the level of divergence was even lower in the 1.0 mg/kg AMPH group than in SAL controls (Fig. 7*C,D*), but it was consistently higher (going up to ~30–40%) for the 3.3 mg/kg AMPH group (Fig. 7*C,E*), regardless of expansion order (Fig. 7*C*). These results support the hypothesis that the task-epoch clusters constitute attractor-like regions within the expanded neural state spaces, with 80–90% of trajectories returning to these states or bound within them after treatment with either SAL or 1.0 mg/kg AMPH (Fig. 7*C,D*). Furthermore, they suggest that, although a low dose of AMPH enhances attractor-like properties, the 3.3 mg/kg dose of AMPH may lead to "flatter" attractor landscapes.

The altered convergence/divergence of neural trajectories after AMPH treatment is further supported by regressing the velocity of the neural trajectory on the distance from the centroid. When the low and the high doses of AMPH are directly compared, differences are observed in both the variance explained by the regression model (r^2, Wilcoxon's signed-rank test, $W = 40$, $p < 0.02$) and its slope (normality accepted, Lilliefor's test, $p > 0.05$, two-sided t test, $t_{(8)} = 4.1$, $p < 0.004$). Thus, a low dose of AMPH increases trajectory velocity at points farther from the centroid, in contrast to the much slower dynamics after treatment with the higher dose of AMPH (Fig. 8*A*). To further back up these observations, we also split veloc-
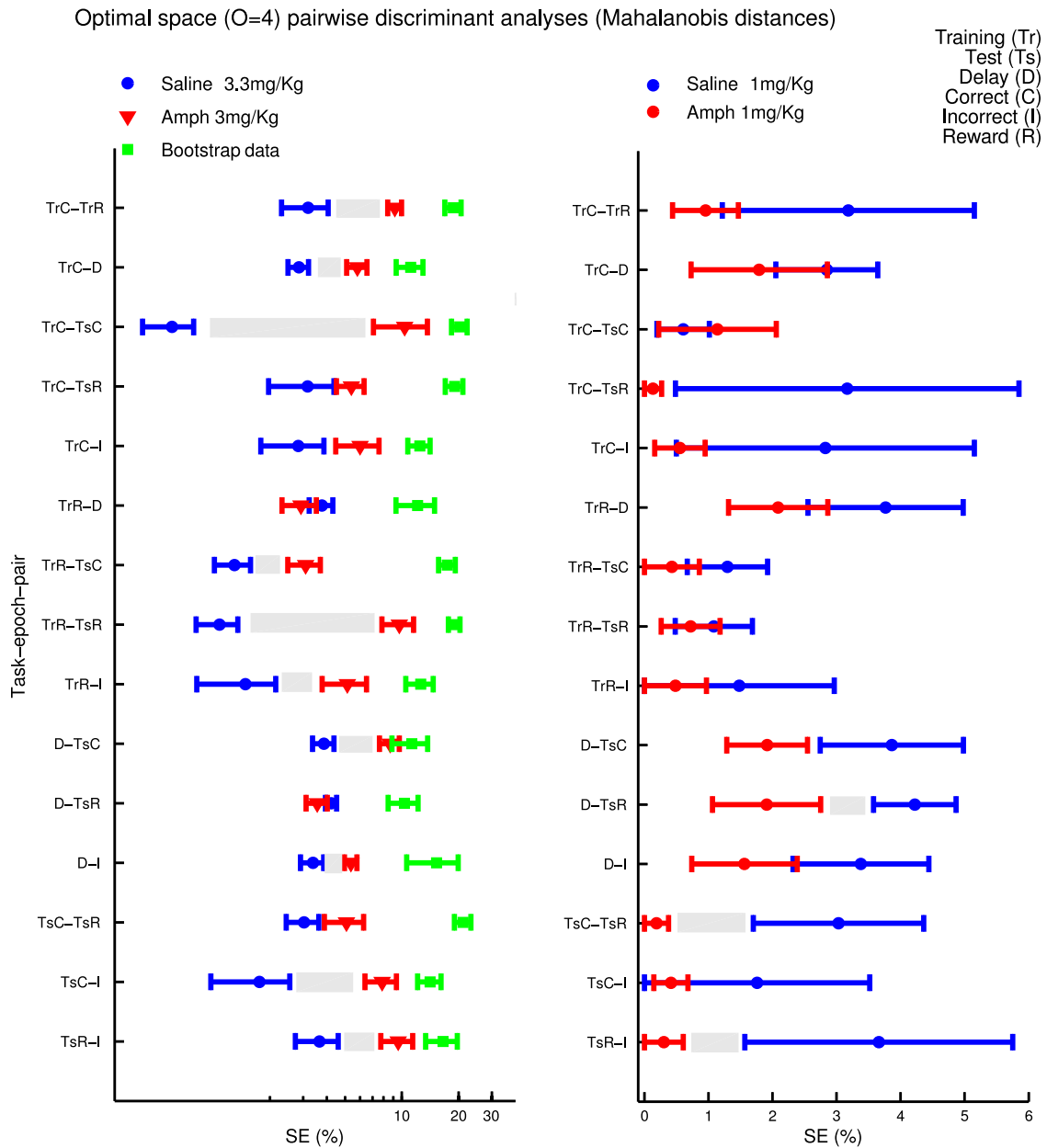
Figure 6. Differences in SEs are stable across task-epoch comparisons. Animals treated with 3.3 mg/kg AMPH exhibit an increase in SE across all pairwise task-epoch comparisons except for two epochs involving reward (left). A consistent tendency toward a reduction in SE was observed in animals treated with 1.0 mg/kg AMPH (right, bootstraps omitted for clarity).

ity vectors into two groups according to whether they were immediately preceding or entering a task-epoch ("boundary" vectors) or whether they ranged well within a task-epoch window ("core" vectors). As shown in Figure 8B1, differences between the low and high AMPH-dose groups were overall significant (rank-sum test, $U > 27,129$, $p < 0.03$) for both core and boundary vectors. When examining these results in more detail as a function of task epoch, we found that there were no differences in vector velocities for the delay and incorrect choice epochs (Fig. 8B2, $p > 0.11$), whereas for correct choice and reward epochs, the differences between core and boundary velocities were significantly larger in the 1.0 mg compared with the 3.3 mg AMPH treatment group (Fig. 8B3; rank-sum test, $U > 20,285$, $p < 0.013$). Thus, these results further strengthen the hypothesis that low AMPH conditions are associated with steeper attractor valleys (larger differences in flow) than high AMPH conditions, specifically during decision-making and reward encoding.

Discussion

In the present study, we examined the hypothesis that low AMPH concentrations would enhance whereas high AMPH concentrations would diminish attractor-like properties in prefrontal network dynamics (Durstewitz and Seamans 2008). We addressed this hypothesis by constructing state spaces from the recorded MSUA through statistical and nonlinear techniques in which the flow of neural trajectories and thus attractor-like properties could be discerned (Balaguer-Ballester et al., 2011). Bidirectional changes in neural dynamics were observed that were dependent on the specific dose of AMPH delivered. Specifically, a low 1.0 mg/kg dose of AMPH increased the extent to which neural trajectories and task-epoch-specific population states could be discerned, whereas the 3.3 mg/kg dose had a deleterious effect on these features. Moreover, compared with a SAL control, the lower AMPH dose enhanced attractor-like properties of task-epoch-
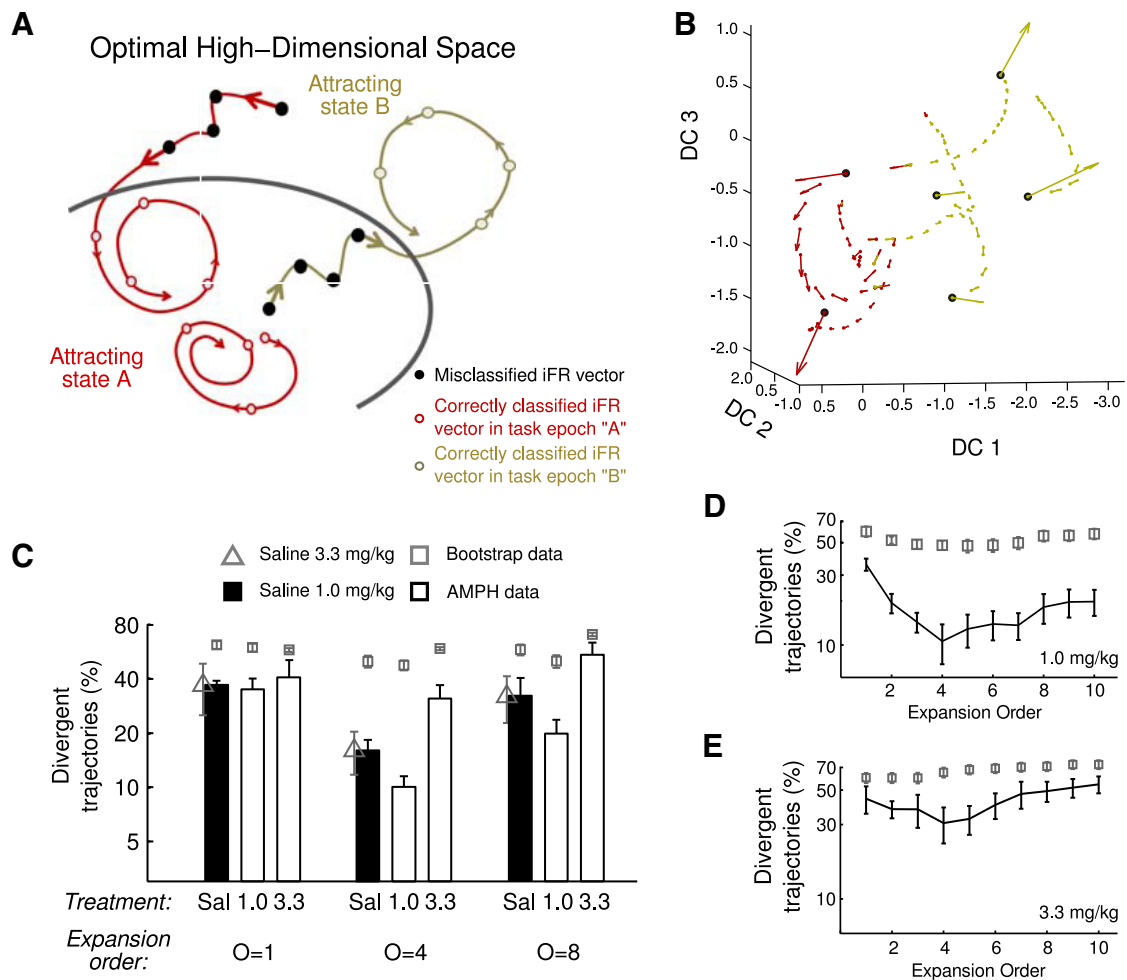
Figure 7. Attraction and divergence in expanded state spaces. **A**, Schema illustrating sections of trajectories both entering and exiting two putative task-epoch-specific attracting states. In this example, the SE is 40%, whereas the percentage of divergent trajectories is 0%. **B**, Projection onto the three maximal discriminant coordinates (DC) derived from the optimally expanded space from a SAL animal. Black circles mark points of transition from one task state to another. **C**, Percentage of neural trajectories that escape from an attractor-like state associated with an individual cognitive epoch (as defined by kernel-FDA; see Results) is reduced in optimally expanded spaces in the 1.0 mg/kg conditions but increased in the 3.3 mg/kg AMPH condition. Bidirectional and dose-dependent effects on trajectory convergence are observed when comparing SAL- and AMPH-treated animals (see Results). The percentage of divergent trajectories at each expansion order is shown for 1.0 mg/kg (**D**) or 3.3 mg/kg (**E**) treated animals (black lines) compared with block-permutation bootstrap data (gray boxes).

specific states as supported by the stronger convergence of task-epoch trajectories and the increased vector velocities toward the class centers. These observations indicate "steeper valleys" of attraction under low AMPH. In contrast, in rats given a 3.3 mg/kg dose of AMPH, attractor states strongly deteriorated with trajectories more easily escaping from task-epoch states and with much weaker separation among population activity states associated with different task epochs. Hence, as suggested previously (Aultman and Moghaddam, 2001; Shoblock et al., 2003; Durstewitz and Seamans, 2008), AMPH appears to regulate neural dynamics in a bidirectional, dose-dependent manner, in which a lower dose (1.0 mg/kg) enhances attractor states, whereas a higher dose (3.3 mg/kg) diminishes them.

Relevance to computational theories of monoaminergic neuromodulation of the PFC

Computational theories of monoamine modulation of cortical dynamics have long postulated that monoamine release might influence the stability of cortical representations in a concentration- and task-dependent manner (Durstewitz et al., 2000; Seamans and Yang, 2004; Aston-Jones and Cohen, 2005a,b; Durstewitz and Seamans, 2008). Accordingly, moderate changes in DA/NE concen-

tration, caused by treatment with low doses of AMPH, may enhance attractor states, whereas high DA concentrations, as measured after higher doses of AMPH, have been proposed to diminish attractor dynamics (Durstewitz and Seamans, 2008). These inferences were drawn from biophysical computational modeling, based on *in vitro* recordings, that have shown that these effects on attractor dynamics are a consequence of the changes in voltage-gated and synaptic ionic conductances associated with different DA concentrations (Durstewitz et al., 2000; Seamans et al., 2001a,b; Durstewitz and Seamans 2002; Seamans and Yang 2004). For example, the simultaneous increase of NMDA and $GABA_A$ synaptic currents evoked by moderate levels of bioavailable DA can increase the "energy barrier" between attractor states and enhance the degree of convergence toward these states, thereby making it more difficult to switch between states (Durstewitz et al., 2000; Durstewitz and Seamans, 2008; Durstewitz, 2009). This may facilitate cognitive functions, such as working memory, decision-making, or more generally any goal-directed activity, via the active and stable maintenance of information that corresponds to a system that locks transiently into a specific attractor state. Conversely, higher DA concentrations would diminish the energy barrier between attractor states,
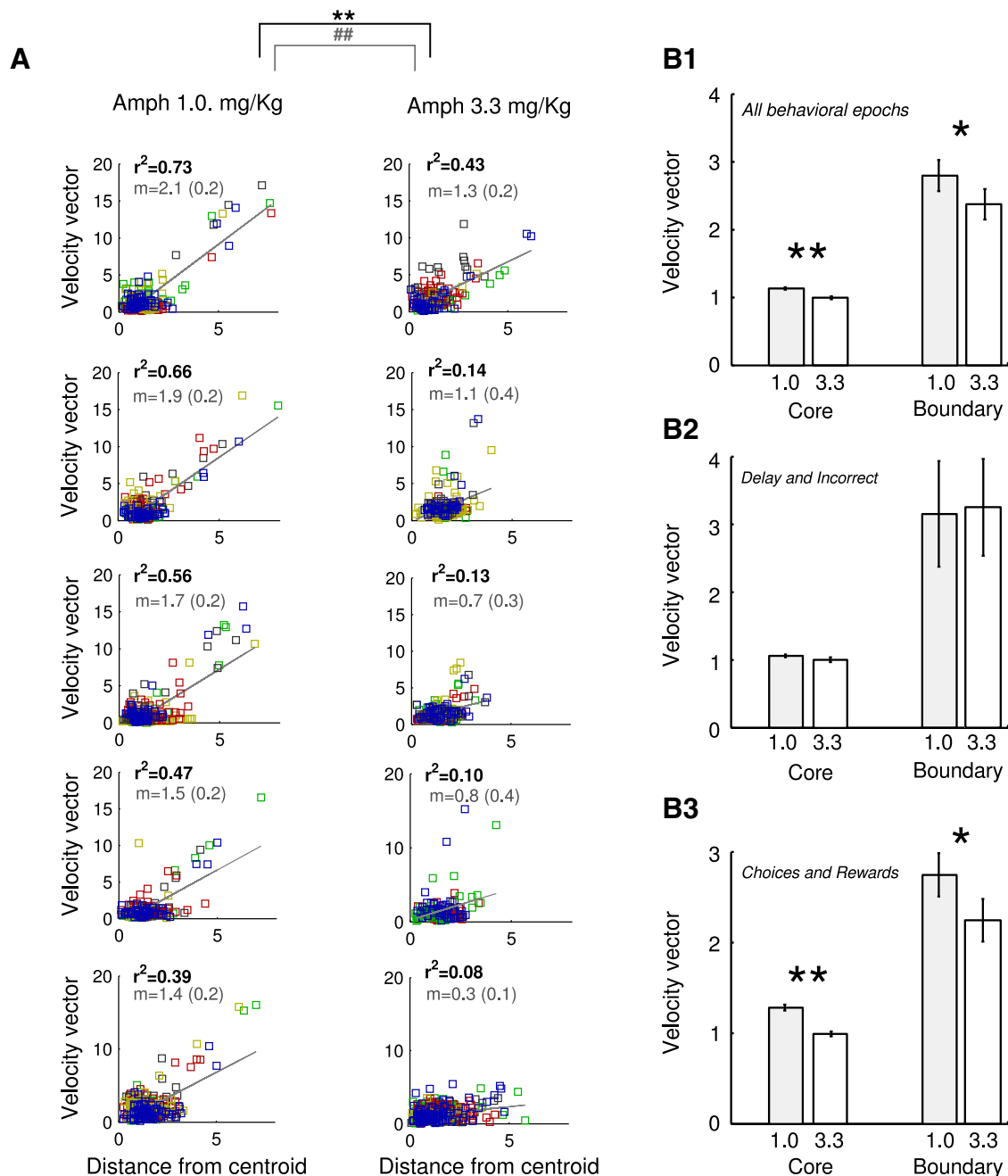
Figure 8. A high dose of AMPH reduces trajectory velocity in expanded neural state space. ***A***, Regressions of the vector velocity on the distance from the centroid for each AMPH treatment reveal differences in the variance explained (inset, black text, r^2) and slope [inset, gray text, slope mean slope (95% confidence interval width)]. **$p < 0.01$, rank-sum test for r^2 comparisons; ##$p < 0.01$, t test for slope comparisons. ***B***, Flow vectors were classified into those immediately preceding or entering a task-epoch (boundary vectors) and those within a task-epoch window (core vectors), and the velocity of the vector was measured for each. Collapsing across all behavioral epochs (***B1***), a significant difference in vector velocity was observed between the 1.0 and 3.3 mg/kg treated animals that was not observed when only the delay and incorrect epochs were considered (***B2***). Instead, the differences between the 1.0 and 3.3 mg/kg AMPH conditions could be attributed mainly to the choice and reward epochs (***B3***). *$p < 0.05$, **$p < 0.01$, rank-sum test.

thereby reducing the degree and strength of convergence toward these states and thus the force required to switch among them. This in turn may facilitate cognitive functions, such as flexibility and task switching, as supported by recent *in vivo* measures of DA during cognitive tasks (St Onge et al., 2011) and behavioral observations (Floresco and Magyanar, 2006; Armbruster et al., 2012; Stelzel et al., 2013). Thus, by differentially modulating attractor landscapes in mPFC, monoamines may regulate the balance between opposing cognitive requirements (Durstewitz and Seamans, 2008). Our current experimental findings provide sup-

port for these theoretical ideas by confirming that low doses of AMPH (1.0 mg/kg) enhance separation among task-phase-specific attracting states (Fig. 4) and convergence toward them (Figs. 7, 8), whereas higher doses of AMPH (>3.0 mg/kg) strongly diminish separation among states (Fig. 3) and the degree of convergence (Figs. 7, 8).

Consistent with our previous study (Balaguer-Ballester et al., 2011), neural trajectories were optimally unfolded when iFR interactions up to some specific order ($O = 3$–5) were used in the state space expansion but did not increase any further from that

optimal order. It is tempting to interpret these numbers in terms of the higher-order interactions that are relevant for describing the joint probability distribution of the spiking activities of the neurons. The relevance of higher-order neural interactions for neural information processing and coding has been reported extensively and debated in a number of recent studies (Riehle et al., 1997; Averbeck et al., 2006; Schneidman et al., 2006; Montani et al., 2009; Quian Quiroga and Panzeri, 2009; Ohiorhenuan et al., 2010). Some of these studies have suggested that up to second-order dependencies among units are sufficient to capture neural population dynamics (Schneidman et al., 2006), whereas others argued that the relevant order may depend on the spatial (columnar) distance between neurons in the neocortex or the brain area studied (Ohiorhenuan et al., 2010). As pointed out previously (Balaguer-Ballester et al., 2011), the optimal order may simply pick out the point at which further increasing the dimensionality of the space compared with the number of data available will start to add noise rather than information, thus misleading the classifier (i.e., the "bias–variance tradeoff"; Hastie et al., 2009). However, previous observations (Balaguer-Ballester et al., 2011) seem to disprove this interpretation because the optimal expansion order appears to be rather robust with respect to variations in dataset size or numbers of recorded neurons. Although the multinomial basis expansions were used primarily as a statistical tool to discern attractor behavior, we observed that second- and higher-order unit correlations were indeed differentially affected by AMPH (Fig. 5). This supports the idea that the state space expansions used detect biologically meaningful differences in network structure and coordination.

Parallels between rodent and human data

A main reason AMPH was used in the present study was to gain insight into the neural dynamics that may underlie the cognitive and behavioral effects of psychostimulant drugs observed in the numerous human studies. Although AMPH is a drug of abuse and can induce mania and psychosis at high doses, it has also been shown to improve attention or vigilance and as therapeutic intervention in the treatment of a number of neuropsychiatric disorders (Sulzer et al., 2005). In humans, AMPH improves working memory, recall, and attention at low doses (Mattay et al., 1996, 2000, 2003; Servan-Schreiber et al., 1998; Barch and Carter, 2005) but causes memory impairments at higher doses (Krystal et al., 2005; Tipper et al., 2005). Furthermore, similar dose-dependent effects have been observed in working memory and attention tasks in rodents (Aultman and Moghaddam, 2001; Shoblock et al., 2003; Chudasama et al., 2005). The 3.3 mg/kg dose of AMPH used in the present study disrupted the memory-based choice of arms on the radial maze. In contrast, in animals already performing at criterion levels, performance was not improved further by the 1.0 mg/kg dose that enhanced attractor dynamics. However, it would be of interest to determine whether the lower dose of AMPH could restore optimal performance under conditions in which memory for the location of food was degraded by longer delay periods.

Mattay et al. (2003) and Tipper et al. (2005) observed a bidirectional effect of AMPH on both performance of the *N*-back working memory task and the BOLD signal, which was related to allelic variations in catechol-*O*-methyl transferase, leading to alterations in the bioavailability of basal PFC DA. In the study by Mattay et al. (2003), when working memory required the greatest effort, a decrease in the BOLD response and impairment in behavioral performance were observed under conditions when PFC catecholamine efflux was presumably the highest. Conversely,

under conditions in which PFC DA levels were presumably in an intermediate range, no measurable change in behavior was detected, but an increase in cortical efficiency was observed in the BOLD response. These findings fit surprisingly well with our current data, because stability and convergent activity in the population dynamic was enhanced after the 1.0 mg/kg AMPH dose, with minimal changes in behavior, whereas the higher AMPH dose both impaired coherent population dynamics and adversely affected task performance. Although additional experiments will be required to causatively link attractor dynamics to specific cognitive processes, according to the computational theories discussed above (Durstewitz et al., 2000; Durstewitz and Seamans, 2002, 2008), the enhanced attractor properties in the 1.0 mg/kg AMPH condition support a more robust representation of task-relevant information (or signal) at the expense of non-task-related information (or noise). The relationship between the 3.3 mg/kg dose used in the current study and the decrease in the BOLD response under high PFC monoamine bioavailability is perhaps even more straightforward because a reduction in neural firing would be expected to lead to a diminished BOLD response. Thus, these very different experimental approaches ultimately arrive at similar conclusions with regards to the powerful bidirectional modulation of prefrontal information processing by monoamines. The dose–concentration effect also appears to exhibit an inverted U shape, consistent with many past studies on catecholaminergic modulation of cognition (Arnsten, 2007, 2011). Hence, the present observations provide a neural attractor-based explanation for the effects of catecholamine modulation of cognitive processing in humans.

References

Armbruster DJ, Ueltzhöffer K, Basten U, Fiebach CJ (2012) Prefrontal cortical mechanisms underlying individual differences in cognitive flexibility and stability. J Cogn Neurosci 24:2385–2399. CrossRef Medline

Arnsten AF (2011) Catecholamine influences on dorsolateral prefrontal cortical networks. Biol Psychiatry 69:e89–e99. CrossRef Medline

Arnsten AFT (2007) Catecholamine and second messenger influences on prefrontal cortical networks of "representational knowledge": a rational bridge between genetics and the symptoms of mental illness. Cereb Cortex 17 [Suppl 1]:i6–i15. CrossRef

Arnsten AF, Cai JX, Murphy BL, Goldman-Rakic PS (1994) Dopamine D1 receptor mechanisms in the cognitive performance of young adult and aged monkeys. Psychopharmacology (Berl) 116:143–151. CrossRef Medline

Aston-Jones G, Cohen JD (2005a) Adaptive gain and the role of the locus coeruleus-norepinephrine system in optimal performance. J Comp Neurol 493:99–110. CrossRef Medline

Aston-Jones G, Cohen JD (2005b) An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annu Rev Neurosci 28:403–450. CrossRef Medline

Aujla H, Beninger RJ (2001) Hippocampal-prefrontocortical circuits: PKA inhibition in the prefrontal cortex impairs delayed nonmatching in the radial maze in rats. Behav Neurosci 115:1204–1211. CrossRef Medline

Aultman JM, Moghaddam B (2001) Distinct contributions of glutamate and dopamine receptors to temporal aspects of rodent working memory using a clinically relevant task. Psychopharmacology (Berl) 153:353–364. CrossRef Medline

Averbeck BB, Latham PE, Pouget A (2006) Neural correlations, population coding and computation. Nat Rev Neurosci 7:358–366. CrossRef Medline

Balaguer-Ballester E, Lapish CC, Seamans JK, Durstewitz D (2011) Attracting dynamics of frontal cortex ensembles during memory-guided decision-making. Plos Comput Biol 7:e1002057. CrossRef Medline

Barch DM, Carter CS (2005) Amphetamine improves cognitive function in medicated individuals with schizophrenia and in healthy volunteers. Schizophr Res 77:43–58. CrossRef Medline

Berridge CW, Stalnaker TA (2002) Relationship between low-dose amphetamine-induced arousal and extracellular norepinephrine and do-

pamine levels within prefrontal cortex. Synapse 46:140–149. CrossRef Medline

Brown MF (1992) Does a cognitive map guide choices in the radial-arm maze? J Exp Psychol Anim Behav Process 18:56–66. Medline

Brunel N, Wang XJ (2001) Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. J Comput Neurosci 11:63–85. CrossRef Medline

Cano-Colino M, Almeida R, Compte A (2013) Serotonergic modulation of spatial working memory: predictions from a computational network model. Front Integr Neurosci 7:71. Medline

Chudasama Y, Nathwani F, Robbins TW (2005) D-Amphetamine remediates attentional performance in rats with dorsal prefrontal lesions. Behav Brain Res 158:97–107. CrossRef Medline

Compte A, Brunel N, Goldman-Rakic PS, Wang XJ (2000) Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. Cereb Cortex 10:910–923. CrossRef Medline

Duong T, Hazelton ML (2005) Cross-validation bandwidth matrices for multivariate kernel density estimation. Scand J Stat 32:485–506.

Durstewitz D (2009) Implications of synaptic biophysics for recurrent network dynamics and active memory. Neural Netws 22:1189–1200. CrossRef Medline

Durstewitz D, Seamans JK (2002) The computational role of dopamine D1 receptors in working memory. Neural Netw 15:561–572. CrossRef Medline

Durstewitz D, Seamans JK (2008) The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-o-methyltransferase genotypes and schizophrenia. Biol Psychiatry 64:739–749. CrossRef Medline

Durstewitz D, Kelc M, Güntürkün O (1999) A neurocomputational theory of the dopaminergic modulation of working memory functions. J Neurosci 19:2807–2822. Medline

Durstewitz D, Seamans JK, Sejnowski TJ (2000) Neurocomputational models of working memory. Nat Neurosci [Suppl] 3:1184–1191. CrossRef Medline

Efron B, Tibshirani R (1993) An introduction to the bootstrap. New York: Chapman and Hall.

Floresco SB, Magyar O (2006) Mesocortical dopamine modulation of executive functions: beyond working memory. Psychopharmacology (Berl) 188:567–585. Medline

Goldman-Rakic PS, Castner SA, Svensson TH, Siever LJ, Williams GV (2004) Targeting the dopamine D1 receptor in schizophrenia: insights for cognitive dysfunction. Psychopharmacology (Berl) 174:3–16. Medline

Grilly DM, Loveland A (2001) What is a "low dose" of d-amphetamine for inducing behavioral effects in laboratory rats? Psychopharmacology (Berl) 153:155–169. CrossRef Medline

Gulley JM, Stanis JJ (2010) Adaptations in medial prefrontal cortex function associated with amphetamine-induced behavioral sensitization. Neuroscience 166:615–624. CrossRef Medline

Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning. New York: Springer.

Homayoun H, Moghaddam B (2006) Progression of cellular adaptations in medial prefrontal and orbitofrontal cortex in response to repeated amphetamine. J Neurosci 26:8025–8039. Medline

Howes OD, Kapur S (2009) The dopamine hypothesis of schizophrenia: version III–the final common pathway. Schizophr Bull 35:549–562. CrossRef Medline

Kantz H, Schreiber T (2004) Nonlinear time series analysis, 2 edition. Cambridge: Cambridge UP.

Krystal JH, Perry EB Jr, Gueorguieva R, Belger A, Madonick SH, Abi-Dargham A, Cooper TB, Macdougall L, Abi-Saab W, D'Souza DC (2005) Comparative and interactive human psychopharmacologic effects of ketamine and amphetamine: implications for glutamatergic and dopaminergic model psychoses and cognitive function. Arch Gen Psychiatry 62:985–994. CrossRef Medline

Lapish CC, Durstewitz D, Chandler LJ, Seamans JK (2008) Successful choice behavior is associated with distinct and coherent network states in anterior cingulate cortex. Proc Natl Acad Sci U S A 105:11963–11968. CrossRef Medline

Mattay VS, Berman KF, Ostrem JL, Esposito G, Van Horn JD, Bigelow LB, Weinberger DR (1996) Dextroamphetamine enhances "neural network-specific" physiological signals: a positron-emission tomography rCBF study. J Neurosci 16:4816–4822. Medline

Mattay VS, Callicott JH, Bertolino A, Heaton I, Frank JA, Coppola R, Berman

KF, Goldberg TE, Weinberger DR (2000) Effects of dextroamphetamine on cognitive performance and cortical activation. Neuroimage 12:268–275. CrossRef Medline

Mattay VS, Goldberg TE, Fera F, Hariri AR, Tessitore A, Egan MF, Kolachana B, Callicott JH, Weinberger DR (2003) Catechol O-methyltransferase val158-met genotype and individual variation in the brain response to amphetamine. Proc Natl Acad Sci U S A 100:6186–6191. CrossRef Medline

Mika S, Rätsch G, Weston B, Schölkopf B, Smola AJ, Müller K (2000) Advances in neural information processing systems. In: Advances in neural information processing systems, pp 526–542. Cambridge, MA: MIT.

Montani F, Ince RAA, Senatore R, Arabzadeh E, Diamond ME, Panzeri S (2009) The impact of high-order interactions on the rate of synchronous discharge and information transmission in somatosensory cortex. Philos Trans A Math Phys Eng Sci 367:3297–3310. CrossRef Medline

Ohiorhenuan IE, Mechler F, Purpura KP, Schmid AM, Hu Q, Victor JD (2010) Sparse coding and high-order correlations in fine-scale cortical networks. Nature 466:617–621. CrossRef Medline

Omi T, Shinomoto S (2011) Optimizing time histograms for non-Poissonian spike trains. Neural Comput 23:3125–3144. CrossRef Medline

Pum M, Carey RJ, Huston JP, Müller CP (2007) Dissociating effects of cocaine and d-amphetamine on dopamine and serotonin in the perirhinal, entorhinal, and prefrontal cortex of freely moving rats. Psychopharmacology (Berl) 193:375–390. CrossRef Medline

Quian Quiroga R, Panzeri S (2009) Extracting information from neuronal populations: information theory and decoding approaches. Nat Rev Neurosci 10:173–185. CrossRef Medline

Rabinovich MI, Varona P (2011) Robust transient dynamics and brain functions. Front Comput Neurosci 5:24. CrossRef Medline

Rabinovich M, Huerta R, Laurent G (2008) Neuroscience. Transient dynamics for neural processing. Science 321:48–50. CrossRef Medline

Riehle A, Grün S, Diesmann M, Aertsen A (1997) Spike synchronization and rate modulation differentially involved in motor cortical function. Science 278:1950–1953. CrossRef Medline

Saadi K, Talbot NLC, Cawley GC (2007) Optimally regularised kernel Fisher discriminant classification. Neural Netw 20:832–841. Medline

Sauer T, Yorke JA (1993) How many delay coordinates do you need? Int J Bifurc Chaos 03:737–744. CrossRef

Sawaguchi T, Goldman-Rakic PS (1994) The role of D1-dopamine receptor in working memory: local injections of dopamine antagonists into the prefrontal cortex of rhesus monkeys performing an oculomotor delayed-response task. J Neurophysiol 71:515–528. Medline

Sawaguchi T, Matsumura M, Kubota K (1990a) Catecholaminergic effects on neuronal activity related to a delayed response task in monkey prefrontal cortex. J Neurophysiol 63:1385–1400. Medline

Sawaguchi T, Matsumura M, Kubota K (1990b) Effects of dopamine antagonists on neuronal activity related to a delayed response task in monkey prefrontal cortex. J Neurophysiol 63:1401–1412. Medline

Schneidman E, Berry MJ 2nd, Segev R, Bialek W (2006) Weak pairwise correlations imply strongly correlated network states in a neural population. Nature 440:1007–1012. CrossRef Medline

Schölkopf B, Smola AJ, Müller K (1998) Kernel principal component analysis. Neural Comput 10:1299–1316. CrossRef

Seamans JK, Yang CR (2004) The principal features and mechanisms of dopamine modulation in the prefrontal cortex. Prog Neurobiol 74:1–58. CrossRef Medline

Seamans JK, Floresco SB, Phillips AG (1995) Functional differences between the prelimbic and anterior cingulate regions of the rat prefrontal cortex. Behav Neurosci 109:1063–1073. CrossRef Medline

Seamans JK, Floresco SB, Phillips AG (1998) D1 receptor modulation of hippocampal-prefrontal cortical circuits integrating spatial memory with executive functions in the rat. J Neurosci 18:1613–1621. Medline

Seamans JK, Durstewitz D, Christie BR, Stevens CF, Sejnowski TJ (2001a) Dopamine D1/D5 receptor modulation of excitatory synaptic inputs to layer V prefrontal cortex neurons. Proc Natl Acad Sci U S A 98:301–306. CrossRef Medline

Seamans JK, Gorelova N, Durstewitz D, Yang CR (2001b) Bidirectional dopamine modulation of GABAergic inhibition in prefrontal cortical pyramidal neurons. J Neurosci 21:3628–3638. Medline

Seeman P (2011) Schizophrenia diagnosis and treatment. CNS Neurosci Ther 17:81–82. CrossRef Medline

Servan-Schreiber D, Carter CS, Bruno RM, Cohen JD (1998) Dopamine and the mechanisms of cognition: Part II. D-amphetamine effects in human subjects performing a selective attention task. Biol Psychiatry 43: 723–729. CrossRef Medline

Shoblock JR, Maisonneuve IM, Glick SD (2003) Differences between d-methamphetamine and d-amphetamine in rats: working memory, tolerance, and extinction. Psychopharmacology (Berl) 170:150–156. CrossRef Medline

Stelzel C, Fiebach CJ, Cools R, Tafazoli S, D'Esposito M (2013) Dissociable fronto-striatal effects of dopamine D2 receptor stimulation on cognitive versus motor flexibility. Cortex 49:2799–2811. CrossRef Medline

Stone TW (1976) Responses of neurones in the cerebral cortex and caudate nucleus to amantadine, amphetamine and dopamine. Br J Pharmacol 56:101–110. CrossRef Medline

St Onge JR, Abhari H, Floresco SB (2011) Dissociable contributions by prefrontal D1 and D2 receptors to risk-based decision making. J Neurosci 31:8625–8633. CrossRef Medline

Sulzer D, Sonders MS, Poulsen NW, Galli A (2005) Mechanisms of neurotransmitter release by amphetamines: a review. Prog Neurobiol 75: 406–433. CrossRef Medline

Takens F (1981) Detecting strange attractors in turbulence. In: Dynamical systems and turbulence, Warwick 1980 (Rand D, Young LS, eds), pp 366–381. Lecture Notes in Mathematics. Berlin: Springer.

Tipper CM, Cairo TA, Woodward TS, Phillips AG, Liddle PF, Ngan ETC (2005) Processing efficiency of a verbal working memory system is modulated by amphetamine: an fMRI investigation. Psychopharmacology (Berl) 180:634–643. CrossRef Medline

White NM, McDonald RJ (2002) Multiple parallel memory systems in the brain of the rat. Neurobiol Learn Mem 77:125–184. CrossRef Medline

Williams GV, Rao SG, Goldman-Rakic PS (2002) The physiological role of 5-HT2A receptors in working memory. J Neurosci 22:2843–2854. Medline

Witten DM, Tibshirani R (2011) Penalized classification using Fisher's linear discriminant. J R Stat Soc Series B Stat Methodol 73:753–772. CrossRef Medline

Yang S, Smit AF, Schwartz S, Chiaromonte F, Roskin KM, Haussler D, Miller W, Hardison RC (2004) Patterns of insertions and their covariation with substitutions in the rat, mouse, and human genomes. Genome Res 14: 517–527. Medline

Yu BM, Cunningham JP, Santhanam G, Ryu SI, Shenoy KV, Sahani M (2009) Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. J Neurophysiol 102:614–635. CrossRef Medline