Conserved structures of neural activity in sensorimotor cortex of freely moving rats allow cross-subject decoding

Svenja Melbaum
University of Freiburg
David Eriksson
University of Freiburg
Thomas Brox
University of Freiburg
Ilka Diester (ilka.diester@biologie.uni-freiburg.de)
University of Freiburg https://orcid.org/0000-0001-9362-0986

Article

Keywords: motor cortex, premotor cortex, sensorimotor cortex

Posted Date: March 9th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-223097/v1

Conserved structures of neural activity in sensorimotor cortex of freely moving rats allow cross-subject decoding

Svenja Melbaum[1,2], David Eriksson[2,3], Thomas Brox[1,2], Ilka Diester[2,3,4*]

[1]Computer Vision Group, Dept. of Computer Science,
University of Freiburg, 79110 Freiburg, Germany.
[2]BrainLinks-BrainTools Cluster of Excellence,
University of Freiburg, Georges-Köhler-Allee 201, IMBIT, 79110 Freiburg, Germany.
[3]Optophysiology Lab, Faculty of Biology,
University of Freiburg, 79110 Freiburg, Germany.
[4]Bernstein Center Freiburg, University of Freiburg, 79104 Freiburg, Germany.

*To whom correspondence should be addressed; E-mail: ilka.diester@biologie.uni-freiburg.de

Our knowledge about neuronal activity in the sensorimotor cortex relies primarily on stereotyped movements which are strictly controlled via the experimental settings. It remains unclear how results can be carried over to less constrained behavior, i.e. freely moving subjects. Towards this goal, we developed a self-paced behavioral paradigm which encouraged rats to conduct different types of movements. Via bilateral electrophysiological recordings across the entire sensorimotor cortex and simultaneous paw tracking, we identified behavioral coupling of neurons with lateralization and an anterior-posterior

gradient from premotor to primary sensory cortex. The structure of population activity patterns was conserved across animals, in spite of severe undersampling of the total number of neurons and variations of electrode positions across individuals. Via alignments of low-dimensional neural manifolds, we demonstrate cross-subject and cross-session generalization in a decoding task arguing for a conserved neuronal code.

One-sentence summary Similarity of neural population structures across the sensorimotor cortex enables generalization across animals in the decoding of unconstrained behavior.

Introduction Humans and animals are capable of generating a vast array of behaviors, a feature dependent on the brain's ability to generate a wide repertoire of neural activity patterns which have been suspect of relying on subsets of general motifs (*1*). Experimental, computational, and theoretical work has identified the underlying rich structures within neural populations in movement control, decision making and memory tasks (*2*). Similarities of population structures across different modalities like speech and arm movements (*3*), as well as the relevance of population-level phenomena for learning (*4*), hint at the existence of general principles that could be shared also across subjects. For simplistic, constrained behavior like running on a linear track, population structures in some brain regions like the hippocampus seem to be conserved even across subjects (*5*). Similarities of neural population structures in freely roaming animals, during different classes of natural occurring behaviors, have so far not been shown. Whether population structures are sufficiently conserved across subjects to allow cross-subject decoding of the behavioral categories, remains an open question in systems neuroscience with large implications for neuroprosthetic approaches.

We addressed this question with nonlinear mapping applied to electrophysiological recordings across the entire bilateral rat sensorimotor cortex. Neural trajectories of dynamical systems

have been suggested as a method to understand neural activity (*6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 4, 17, 18, 19, 20, 21*). We built on Laplacian Eigenmaps (LEM) (*22, 5*), which map high-dimensional data via the data affinity to a low-dimensional manifold. With affinities defined on neuronal population activity, it can be seen as a tool to visualize structures and relationships between population activities at different time points of a recording session in a low-dimensional space. This can potentially reveal conserved structures across sessions as well as across animals (*5*).

To investigate how conserved the low-dimensional structures are, it is necessary to involve several different types of behavior. In principle, it is possible to train animals in different tasks, but this comes with several limitations: (1) training animals is time consuming, especially if multiple behaviors are involved; (2) the trained behavior often results in stereotyped movements and, due to the plasticity of the mammalian brain, to changes of neuronal representations; (3) frequent transitions between behaviors are not feasible. Furthermore, spontaneous movements influence neuronal activity even in well-controlled tasks (*23*). Therefore, we refrained from controlling the behavior from the start and rather allowed the rats to roam freely in a plexiglas box. Consequently, the animals showed the full array of natural behavior, such as rearing, grooming, turning, stepping, drinking, and resting, in an unbiased manner.

To verify this unconventional approach, we first compared the neuronal activity with previously reported results from more constrained behaviors by focusing on step- and swing-like paw movements. The study confirmed that the quality of information conveyed in our recorded data was comparable to conventional controlled settings. In addition, we report a strong anterior-posterior bilateral gradient from premotor over primary motor to primary sensory cortex, emphasising the strong involvement of more posterior regions in the encoding of step-like behavior.

After this validation, we focused on analyzing the population code on more complex behaviors. We conducted a normal within-session decoding experiment to show that the neuronal

code comprises enough information about the behavior class. Across sessions the signal of individual neurons is not comparable, since neurons typically cannot be traced over multiple days. Across subjects, even the electrode positions vary. However, we found evidence that the signal from the *population* of neurons shares a common structure across sessions and even subjects. In particular, decoding behavioral categories from the neuronal population activity is possible across different subjects.

Results Rats moved unconstrained in a rectangular arena and conducted movements in different behavioral categories (i.e. stepping, turning, drinking, grooming, and rearing), while mainly searching for water drops, which where occasionally delivered by a robot arm positioned under a mesh (Fig. 1a). We recorded the neuronal activities with electrodes covering sensorimotor cortex on both hemispheres (Fig. 1b). Two cameras videotaped the behavior of the rats for simultaneous 3D tracking. Recording sessions ($n = 106$ in total) were distributed over three months and varied in length between 30 and 60 min ($\mu = 36.06$ min, $\sigma = 5.23$ min). In total, we identified 3723 single units ($\mu = 35.12, \sigma = 20.71$ across sessions) which we used for further analysis: 730, 896, and 230 in left M2, M1, and S1, and 432, 793, 642 in right M2, M1, and S1, respectively (*24*).

In order to extract behavioral components from the movements, we focused on step-like behavior. To extract the steps, we binarized the movements of the paws into swing (moving) and stance (not moving) according to a horizontal velocity threshold (0.03 mm/ms). With each paw, rats performed on average one step per second ($\mu = 1.22, \sigma = .29$).

Strongest paw coupling in contralateral S1 Since classical methods like peristimulus time histograms (PSTHs) are not applicable for behavior without trial structure, we computed spike-triggered averages to investigate the relationship between neuronal activity and unconstrained movements (*25*). We defined the spike-triggered average paw swing-stance status

(STAPSSS) as a rough measure for the coupling of individual neurons to paw movements. For each neuron and each paw, we calculated STAPSSS by averaging the swing-stance status in the time period ± 1s around the spikes (Fig. 1c). As statistical control, we randomly shifted the spike train 1000 times to calculate 1000 control STAPSSS waveforms. We considered STAPSSS to be significant if their standard deviation over time exceeded the .99 quantile standard deviation of the control STAPSSS waveforms. Only neurons which spiked temporally more systematically in relation to movement parameters than expected by chance can pass this test. Significantly coupled neurons were characterized by clear peaks in STAPSSS (Fig. 1d). In total, 54% (2029/3723) of all neurons were significantly coupled to at least one paw. These were 45% (534/1162) of all neurons in M2, compared to 53% (908/1689) in M1 and 67% (587/872) in S1.

To take into account the strength of coupling, we defined a continuous measure for paw coupling as the quotient of the STAPSSS standard deviation and the control standard deviation (> 1 for significant neurons). Using this quotient as a dependent variable, we calculated three-way ANOVAs (factors hemisphere, area, rat) for all four paws separately (detailed results in Table S2). In summary, we found for all four paws a stronger coupling on the contralateral side ($p = .04$), arguing for lateralization during locomotion. The coupling increased from anterior to posterior areas ($p < 1\mathrm{e}-11$). For all four paws, the highest mean coupling was localized in contralateral S1 (Fig. 1e). In three out of four paws, the interaction between area and hemisphere was also significant, i.e. the differences between contralateral and ipsilateral hemisphere increased from anterior to posterior areas ($p = .02$). To further investigate the difference in magnitude between contralateral and ipsilateral paw coupling, we defined the contralateral bias as the ratio between the coupling of the contralateral and ipsilateral paw, i.e. $b = c_r/c_l$ for left-hemispheric neurons and $b = c_l/c_r$ for right-hemispheric neurons ($b \approx 1$ for non-biased neurons), with bias denoted as b, coupling as c, right paw as r and left paw as l.

We calculated this bias separately for the front and hind paws. A two-way ANOVA on the contralateral bias of individual neurons revealed a significant effect of the brain area for the front paws ($F_{2,3715} = 44.66, p < 1\text{e}{-19}$) and the hind paws ($F_{2,3715} = 54.56, p < 1\text{e}{-23}$). This confirmed that single neurons had a larger contralateral bias from anterior to posterior areas for front and hind paws (Fig. S1).

Single-unit activity allows decoding of paw movements within sessions Due to the strong paw coupling we hypothesized that it is possible to decode paw movements of the freely moving rats from neuronal activity. To test this hypothesis, we applied feed-forward neural networks to decode the swing-stance status of the right front paw, posed as a two-class classification problem. For each time point we fed in the spike trains ± 400 ms of all units, in time bins of 10 ms duration. The deep neural networks were trained and evaluated separately for each recording session. We chose this approach here since single-neuron activity does not generalize over sessions, in contrast to our population-level decoding approach in the following section. The mean per-class decoding accuracies were well above chance level ($\mu = 71.47\,\%, \sigma = 9.98\,\%$; chance level 50 %). There was no significant correlation between accuracy and train set sizes (Spearman's $\rho = .17, p = 0.07$). However, we found a significant correlation between the accuracy and the percentage of coupled neurons according to our STAPSSS analysis per session (Spearman's $\rho = .63, p < 1\text{e}{-12}$, Fig. 1f), which confirms that STAPSSS is a reliable measure for the correlation between neuronal activity and movement.

Similar structure of population activities Due to the promising decoding results, we sought to determine whether population activities during unconstrained movements contained structures that were conserved across recording sessions or even across different animals. This would potentially allow cross-subject decoding. This analysis was based on population responses, thus allowing us to include multi-units in this analysis. To reveal structures in the

population activities, we used Laplacian Eigenmaps (LEM) (*22, 5*). LEM is a non-linear dimensionality reduction method for extracting low-dimensional manifolds in high-dimensional data with spectral techniques. We applied LEM on neighborhood graphs of neuronal activity vectors to visualize structures and relationships between population activities at different time points of a recording session in a low-dimensional space. While it is close to impossible to find correspondences on single-neuron level across animals, we hypothesized that the underlying structures might be similar for different animals and sessions, given that the same aspects of the behavioral repertoire were covered. Indeed, the resulting projections showed striking similarities across animals and sessions when visualized in three dimensions, namely a clear saddle-like shape (Fig. 2a, 52/95, roughly 55% of session structures had a similar shape as classified by eye), although there were also random-like population structures differing from the majority (14/95, about 15% of the sessions, Fig. S2; the remaining 35% had intermediate levels of structuredness). Sessions with clear saddle-like shape were characterized by a larger number of neurons which were significantly coupled to at least one paw, compared to sessions with intermediate or low level of structuredness (23.57 ± 14.71 vs 16.62 ± 12.33 neurons, $t_{93} = 2.43, p = .016$). To ensure that the saddle-like structures were not a simple artifact of the dimensionality reduction method, we also performed time- and neuron-shuffled control reductions (*5*) which did not lead to any apparent structure (Fig. S3).

To investigate the relationship between the population structures and the corresponding behavior, we proceeded by manually labeling sessions in 500ms snippets into six behavioral classes (Fig. 2b, step/paw movement, turn/head movement, drinking, grooming, rearing, resting). We included all sessions with clear saddle-like shape and with at least five significantly coupled neurons, which resulted in a total of 48 sessions (13 for Rat A, 16 for Rat B, 7 for Rat C, 6 for Rat D, 3 for Rat E, 3 for Rat F). While each session contained at least some samples of each behavior, the occurrences of behaviors still differed considerably across sessions and rats

(Fig. S4). In contrast, the distributions of behaviors across the neural structures revealed clear similarities across rats, which was surprising assuming a sampling of approximately .005% of all neurons[1] on average in only roughly overlapping recording sites (Fig. 1B). For example, the second eigenvector (here: first dimension), the so-called Fiedler vector, clearly represented the difference between movement and rest (Fig, 3a upper panel). For some animals, a clear distinction between more paw-related (paw movement, rearing) and head-related behavior (head movement, drinking) was observable in the third and fourth eigenvector (here: second/third dimension, Fig. 3a lower panel). For all rats, a neural network was able to classify the behaviors above chance level, with the neural activity reduced to ten dimensions as input (mean per-class accuracy $\mu = 47.11\%, \sigma = 9.62\%$; chance level 16.66%, Fig. 2c). The accuracies were correlated to the number of significantly coupled neurons ($n = 48$, Pearson's $\rho = .59, p < 1\text{e}{-5}$, Fig. 2d) and the total number of units (Pearson's $\rho = .54, p < 1\text{e}{-4}$, Fig. S5a) as well as the SNR averaged over units (Pearson's $\rho = .49, p < .001$, Fig. S5b). Common classification mistakes consisted of confusing rearing and stepping, as well as slight head movements (turn) with resting (Fig. 2e). We observed the lowest accuracy for Rat D, Rat E and Rat F. Those rats had a low mean signal-to-noise ratio (Rat E-F, Fig. S5b) or no electrode coverage of posterior areas (Rat D/ Rat F, Fig. 1B). This last aspect made us hypothesize that more posterior regions are primarily involved in the encoding of behavioral classes. To test this hypothesis, we investigated the influence of the different sensorimotor areas on the neural population structures. To this end, we conducted dimensionality reductions with equal numbers of neurons (i.e., 20 randomly chosen units) from M2, M1 or S1 as input. With this subset, we trained artificial neural networks to decode the behavioral classes with the neural activity in a given area reduced to five dimensions as input. Decoding accuracies from M1 were significantly better than M2 (paired t-test, $t_{40} = 4.18, p < .001$) and slightly, but not significantly, better than those

[1]Quotient of recorded cells and estimated total number of cells (approximated for the area covered by the implanted electrodes by assuming a cortical thickness of 2mm and a density of 90k neurons per mm^3 (*26*)).

from S1 ($t_{41} = 1.90, p = .06$). In total, the accuracies were highest in M1 for 28 out of 48 sessions, compared to 15 for S1 and 5 for M2 (Fig. 2f, accuracies $\mu = 25.80 \pm 4.92\%$ in M2, $\mu = 28.60 \pm 5.34\%$ in M1, $\mu = 26.66 \pm 5.41\%$ in S1). The low relevance of anterior sensorimotor regions is in line with the STAPSSS results as well as with the lower decoding accuraries in Rat D and Rat F.

Cross-subject and cross-session decoding For our subsequent cross-subject and cross-session decoding task, we excluded Rat F because of low recording quality, which could have been due to the long delay between implantation and measurements compared to the other rats (see Table S1). For the decoding analysis, we divided the six behavioral classes in two disjoint sets: one "align set" that was used to align the neural structures, and one "decode set" that was used for training and testing a classifier. The mean neural vectors (four dimensions) corresponding to the behavioral classes in the "align set" were used to compute a Procrustes transformation between two sessions to align the population activity structures (*27, 28*). Procrustes transformations involve translation, scaling, reflection and rotation and thus preserve the shape of a set of points. For decoding, we trained a classifier on samples from the "decode set" of one session of one rat, using the activity in the dimensionality-reduced neural space as input, and tested the generalization on another session of the same (cross-session decoding) or a second rat (cross-subject decoding) (Fig. 3b-c). In the first experiment, the "align set" consisted of four behavioral classes with two other classes remaining for the "decode set". This led to a total of 15 possible splits in two sets. Classifiers trained on sessions with good decodability also successfully generalized to other sessions from the same or other rats (Fig. 3d-e, Fig. S7a). These were sessions of Rat A, B, and C with sufficient recording quality and a sufficiently high number of units for a robust estimation of the underlying population structures. The correlation between within-session and between-session accuracies was high (Fig. 3e, $n = 45$, Pearson's

$\rho = .68, p < 1\text{e}{-6}$). We defined the "generalization accuracy" of a session as the average test accuracy across all sessions (mean value per row of Fig. 3d). These generalization accuracies were correlated to the total number of units (Pearson's $\rho = .38, p < .01$), with a higher number of units leading to a better estimation of the population structure. The generalization accuracies were also correlated to the session length (Pearson's $\rho = .37, p < .05$), since the number of samples used for LEM (only time points with sufficient activity were used) varied across sessions and rats. Finally, the recording quality, namely the SNR averaged over units, was correlated with generalization (Pearson's $\rho = .38, p < .01$).

For a more systematic test of the relation between the number of units and generalization, we took all sessions with a generalization accuracy of at least 55% (19 sessions from Rat A, B, C in total) and conducted an ablation study with LEM reductions on reduced number of units (20,40,60,80 units removed per session). We then repeated the generalization experiment on aligned LEM structures. The accuracies steadily decreased with less units (Fig. S6), confirming the high relevance of the number of units for a robust estimate of the population structure. To determine which sensorimotor areas were most relevant for generalization, we again took the best 19 sessions and conducted LEM reductions after removing M1, S1 or M2. For a fair comparison, we additionally removed a random portion of all other units for underrepresented areas, such that the number of units after removal of M1, M2 and S1 neurons remained constant per session. Generalization accuracies on aligned LEM structures considerably broke down after removal of M1 (accuracies $\mu = 53.91 \pm 7.48\%$) and were slightly but significantly lower than after removal of units from S1 (accuracies $\mu = 55.94 \pm 7.73\%$, paired t-test, $t_{5414} = -16.07, p < 1\text{e}{-56}$), M2 (accuracies $\mu = 54.99 \pm 8.32\%$, $t_{5414} = -8.14, p < 1\text{e}{-15}$) or of the same number of units distributed over all areas (accuracies $\mu = 55.55 \pm 7.84\%$, $t_{5414} = -13.89, p < 1\text{e}{-42}$).

In a second experiment, we used only three classes in the "align set" and the three remaining

in the "decode set" to test the generalization under more difficult conditions, resulting in 20 possible splits of the total of six classes. The general pattern of the generalization matrix stayed the same (Fig. S7b-d). To verify that the classifiers did not only learn to discriminate the simplest difference between rest and movement, we conducted another experiment without the class "rest". Although the accuracies were lower in these setting, the general pattern remained the same (Fig. S7e-f). For an assessment of the relevance of alignment of the neural structures, we also tested the generalization on neural structures without explicit alignment as a control. For most cases, accuracies on aligned structures were much higher than those on unaligned structures (Fig. S8).

Discussion We have investigated single-neuron activity as well as population activity patterns in the rat sensorimotor cortex during unconstrained and self-paced behavior. The behavior was as closely related as possible to naturally occurring behavior, as it was based on foraging while still being performed in a limited arena to allow reliable movement tracking. The first analyses represented sanity checks to validate our unconventional approach of studying freely moving animals without clear trial structure as well as the analysis methods. Based on the chosen measure STAPSSS, 54 % of all neurons were significantly coupled to paw movements. This fraction of coupled neurons is in the range of previously reported numbers; i.e. 60 % of neurons in hindlimb motor cortex reacted to different locomotion scenarios (*29*) and 44 % in M1 were body-coupled in freely moving rats (*25*).

Our multi-side recording approach allowed us to comprehensively test for differences in neuronal activity across the entire sensorimotor cortex. Previously, it has been described that the laterality of forelimb representations increases from M2 to M1 in a pedal task for head-restrained rats (*30*). Here, we extend this laterality gradient to more posterior regions, i.e. S1. As we targeted the output layer of cortex (layer V), we putatively biased our recordings

towards pyramidal tract neurons which have been described to be predominately involved in laterality (*30*).

While the above described findings refer to general features of the sensorimotor cortex, the main finding of our study is based on conserved neuronal population structures. The concept of systematic structures in neuronal data is not new and we provide here a brief overview about previous work in this field to contrast our work with. Experimental, computational, and theoretical work has identified rich structure within the coordinated activity of interconnected neural populations in movement control, decision making and memory tasks. These findings are conceptualized within the framework of neural population dynamics which can reveal general motifs (*2*). To reveal structure and geometric properties, recurrent neural networks (RNNs) can be applied to neural data (*31*). Multiple tasks can be represented in different RNN models. In these networks clusters of units have been identified that are specialized for a subset of tasks (*1*). Alternatively, methods like PCA and its variants dPCA and jPCA have been applied to identify stability of motifs across modalities, e.g. arm and speech control (*3*), as well as within and across brain areas (*4*).

In contrast to the described previous studies, we focused on the existence of conserved neuronal structures across animals and without any clear instructed task line but with several behavioral classes. These two points differentiate our study from previous publications in the field. We investigated population activity patterns, which are commonly assumed to reside on low-dimensional manifolds in the full neural state space (*14, 16, 32, 33, 34*). In contrast to the global principal component analysis (PCA) used by most studies (*12, 20, 15, 16, 4, 17, 18, 35*), we rather assume the preservation of local neighborhood relations in the data. Therefore, we employed Laplacian Eigenmaps (LEM) (*22, 5*) to reveal the presumed preserved low-dimensional structures. Remarkably, neuronal population activity during unconstrained behavior contained similar structures across animals and sessions, visibly already by eye in the first three dimen-

sions. Furthermore, the distribution of different behaviors across low-dimensional neural structures was systematic, which we confirmed with above-chance decoding results. The allocation of different behaviors on the population structures revealed strong similarities across rats. Particularly, movement and rest could be visually distinguished very clearly in the first dimension. This is in line with results on clear separations in the neural state space for output-potent and output-null (e.g. preparatory) neural activity (*12, 20*). Given that we are dealing with a relative structure it might have been suspected that rotation and stretching allows revealing these stable structures. However, it is one thing to suspect and another thing to show. Taken together, to our knowledge this is the first time that the conservation of neural structure across animals and different spontaneously occurring behavioral classes has been shown. This finding implies that conserved neuronal structures occur without training, suggesting that the neuronal computations underlying these structures might be similarly realized across individuals either from birth or during development.

To close the loop to our initial analysis, we found a low relevance of anterior motor cortex for information on behavioral categories which is in line with our STAPSSS results. In contrast to the encoding of paw movements though, our results on population decoding of the more high-level behavioral categories hint on major contributions from M1, not mainly S1. Thus, our results closes a gap of a previous study, investigating postural and behavioral encoding in posterior parietal cortex and M2 (*36*).

To support our main claim that low-dimensional neural manifolds are comparable across sessions and animals, even in the case of unconstrained behavior, we demonstrated that simple classifiers trained on one session can generalize to a different session. Since LEM delivers eigenvectors ordered by eigenvalue magnitude, neural manifolds needed to be aligned. A simple supervised, shape-preserving alignment procedure, namely a Procrustes transformation between mean population vectors for different behavioral classes in the dimensionality-reduced

neural space, sufficed for successful cross-rat generalization in a decoding task with different, but related, behavioral classes. Our procedure was applicable to sessions with sufficient recording quality (indicated by a high SNR of the recorded units) and a number of units sufficient for robust population estimation. Further, the generalization accuracies of the sessions were tightly related to the accuracy for a session itself. Generalization considerably worsened with population structure estimates based on fewer units. In line with the within-session decoding results, we also found that generalization significantly decreased after removal of M1, which indicated consistent population responses especially in this area. While we used LEM as a dimensionality-reduction method with solid theoretical basis, we would like to stress that our main focus was to show that neural manifolds during unconstrained behavior are comparable across animals and sessions. Certainly, there exist other non-linear dimensionality reduction methods that could also be applied.

It is remarkable that sampling as little as approximately 0.005% of all neurons, in only roughly overlapping electrode positions, sufficed for the estimation of population structures of sufficiently close similarity to allow cross-subject generalization, at least for sessions with a sufficient number of units to allow for robust neural manifold estimation. Internal states (such as thirst, attention, or motivation) which we did not analyze here, might have influenced the neuronal activities further (*37*). It has been hypothesized that across-individuals decoding might not be possible with increasing task complexity (*5*). In contrast, our results indicate that even during unconstrained behavior the relationships between neural activity patterns are conserved across different animals. This conservation of population-level neural phenomena paves the way for cross-subject decoding, even in the difficult case of unconstrained behavior.

References

1. G. R. Yang, M. R. Joglekar, H. F. Song, W. T. Newsome, and X.-J. Wang, "Task representations in neural networks trained to perform many cognitive tasks," *Nature neuroscience*, vol. 22, no. 2, pp. 297–306, 2019.

2. S. Vyas, M. D. Golub, D. Sussillo, and K. V. Shenoy, "Computation through neural population dynamics," *Annual Review of Neuroscience*, vol. 43, pp. 249–275, 2020.

3. S. D. Stavisky, F. R. Willett, G. H. Wilson, B. A. Murphy, P. Rezaii, D. T. Avansino, W. D. Memberg, J. P. Miller, R. F. Kirsch, L. R. Hochberg, *et al.*, "Neural ensemble dynamics in dorsal motor cortex during speech in people with paralysis," *Elife*, vol. 8, p. e46015, 2019.

4. M. G. Perich, J. A. Gallego, and L. E. Miller, "A neural population mechanism for rapid learning," *Neuron*, vol. 100, no. 4, pp. 964–976, 2018.

5. A. Rubin, L. Sheintuch, N. Brande-Eilat, O. Pinchasof, Y. Rechavi, N. Geva, and Y. Ziv, "Revealing neural correlates of behavior without behavioral measurements," *Nature communications*, vol. 10, no. 1, pp. 1–14, 2019.

6. M. Rabinovich, R. Huerta, and G. Laurent, "Transient dynamics for neural processing," *Science*, vol. 321, no. 5885, pp. 48–50, 2008.

7. K. V. Shenoy, M. Sahani, and M. M. Churchland, "Cortical control of arm movements: a dynamical systems perspective," *Annual review of neuroscience*, vol. 36, 2013.

8. J. A. Michaels, B. Dann, and H. Scherberger, "Neural population dynamics during reaching are better explained by a dynamical system than representational tuning," *PLoS computational biology*, vol. 12, no. 11, p. e1005175, 2016.

9. D. V. Buonomano and W. Maass, "State-dependent computations: spatiotemporal processing in cortical networks," *Nature Reviews Neuroscience*, vol. 10, no. 2, pp. 113–125, 2009.

10. E. E. Fetz, "Are movement parameters recognizably coded in the activity of single neurons?," *Behavioral and Brain Sciences*, vol. 15, no. 4, pp. 679–690, 1992.

11. M. M. Churchland, J. P. Cunningham, M. T. Kaufman, J. D. Foster, P. Nuyujukian, S. I. Ryu, and K. V. Shenoy, "Neural population dynamics during reaching," *Nature*, vol. 487, no. 7405, p. 51, 2012.

12. M. T. Kaufman, M. M. Churchland, S. I. Ryu, and K. V. Shenoy, "Cortical activity in the null space: permitting preparation without movement," *Nature neuroscience*, vol. 17, no. 3, p. 440, 2014.

13. P. T. Sadtler, K. M. Quick, M. D. Golub, S. M. Chase, S. I. Ryu, E. C. Tyler-Kabara, M. Y. Byron, and A. P. Batista, "Neural constraints on learning," *Nature*, vol. 512, no. 7515, p. 423, 2014.

14. J. A. Gallego, M. G. Perich, L. E. Miller, and S. A. Solla, "Neural manifolds for the control of movement," *Neuron*, vol. 94, no. 5, pp. 978–984, 2017.

15. A. Miri, C. L. Warriner, J. S. Seely, G. F. Elsayed, J. P. Cunningham, M. M. Churchland, and T. M. Jessell, "Behaviorally selective engagement of short-latency effector pathways by motor cortex," *Neuron*, vol. 95, no. 3, pp. 683–696, 2017.

16. J. A. Gallego, M. G. Perich, S. N. Naufel, C. Ethier, S. A. Solla, and L. E. Miller, "Cortical population activity within a preserved neural manifold underlies multiple motor behaviors," *Nature communications*, vol. 9, no. 1, p. 4233, 2018.

17. J. A. Hennig, M. D. Golub, P. J. Lund, P. T. Sadtler, E. R. Oby, K. M. Quick, S. I. Ryu, E. C. Tyler-Kabara, A. P. Batista, M. Y. Byron, *et al.*, "Constraints on neural redundancy," *Elife*, vol. 7, p. e36774, 2018.

18. A. A. Russo, S. R. Bittner, S. M. Perkins, J. S. Seely, B. M. London, A. H. Lara, A. Miri, N. J. Marshall, A. Kohn, T. M. Jessell, *et al.*, "Motor cortex embeds muscle-like commands in an untangled population response," *Neuron*, vol. 97, no. 4, pp. 953–966, 2018.

19. C. Pandarinath, K. C. Ames, A. A. Russo, A. Farshchian, L. E. Miller, E. L. Dyer, and J. C. Kao, "Latent factors and dynamics in motor cortex and their application to brain–machine interfaces," *Journal of Neuroscience*, vol. 38, no. 44, pp. 9390–9401, 2018.

20. G. F. Elsayed, A. H. Lara, M. T. Kaufman, M. M. Churchland, and J. P. Cunningham, "Reorganization between preparatory and movement population responses in motor cortex," *Nature communications*, vol. 7, p. 13239, 2016.

21. C. Pandarinath, D. J. O'Shea, J. Collins, R. Jozefowicz, S. D. Stavisky, J. C. Kao, E. M. Trautmann, M. T. Kaufman, S. I. Ryu, L. R. Hochberg, *et al.*, "Inferring single-trial neural population dynamics using sequential auto-encoders," *Nature methods*, vol. 15, no. 10, pp. 805–815, 2018.

22. M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural computation*, vol. 15, no. 6, pp. 1373–1396, 2003.

23. S. Musall, M. T. Kaufman, A. L. Juavinett, S. Gluf, and A. K. Churchland, "Single-trial neural dynamics are dominated by richly varied movements," *Nature neuroscience*, vol. 22, no. 10, pp. 1677–1686, 2019.

24. G. Paxinos and C. Watson, *The rat brain in stereotaxic coordinates: hard cover edition*. Elsevier, 2006.

25. P. A. Kells, S. H. Gautam, L. Fakhraei, J. Li, and W. L. Shew, "Strong neuron-to-body coupling implies weak neuron-to-neuron coupling in motor cortex," *Nature communications*, vol. 10, no. 1, p. 1575, 2019.

26. V. Braitenberg and A. Schüz, *Cortex: statistics and geometry of neuronal connectivity*. Springer Science & Business Media, 2013.

27. J. V. Haxby, J. S. Guntupalli, A. C. Connolly, Y. O. Halchenko, B. R. Conroy, M. I. Gobbini, M. Hanke, and P. J. Ramadge, "A common, high-dimensional model of the representational space in human ventral temporal cortex," *Neuron*, vol. 72, no. 2, pp. 404–416, 2011.

28. H.-T. Chen, J. R. Manning, and M. A. van der Meer, "Between-subject prediction reveals a shared representational geometry in the rodent hippocampus," *bioRxiv*, 2020.

29. J. DiGiovanna, N. Dominici, L. Friedli, J. Rigosa, S. Duis, J. Kreider, J. Beauparlant, R. Van Den Brand, M. Schieppati, S. Micera, *et al.*, "Engagement of the rat hindlimb motor cortex across natural locomotor behaviors," *Journal of Neuroscience*, vol. 36, no. 40, pp. 10440–10455, 2016.

30. S. Soma, A. Saiki, J. Yoshida, A. Ríos, M. Kawabata, Y. Sakai, and Y. Isomura, "Distinct laterality in forelimb-movement representations of rat primary and secondary motor cortical neurons with intratelencephalic and pyramidal tract projections," *Journal of Neuroscience*, vol. 37, no. 45, pp. 10904–10916, 2017.

31. A. A. Russo, R. Khajeh, S. R. Bittner, S. M. Perkins, J. P. Cunningham, L. F. Abbott, and M. M. Churchland, "Neural trajectories in the supplementary motor area and motor cortex exhibit distinct geometries, compatible with different classes of computation," *Neuron*, vol. 107, no. 4, pp. 745–758, 2020.

32. C. J. MacDowell and T. J. Buschman, "Low-dimensional spatio-temporal dynamics underlie cortex wide neural activity," *bioRxiv*, 2020.

33. P. Gao and S. Ganguli, "On simplicity and complexity in the brave new world of large-scale neuroscience," *Current opinion in neurobiology*, vol. 32, pp. 148–155, 2015.

34. J. P. Cunningham and M. Y. Byron, "Dimensionality reduction for large-scale neural recordings," *Nature neuroscience*, vol. 17, no. 11, p. 1500, 2014.

35. J. A. Gallego, M. G. Perich, R. H. Chowdhury, S. A. Solla, and L. E. Miller, "Long-term stability of cortical population dynamics underlying consistent behavior," *Nature Neuroscience*, pp. 1–11, 2020.

36. B. Mimica, B. A. Dunn, T. Tombaz, V. S. Bojja, and J. R. Whitlock, "Efficient cortical coding of 3d posture in freely behaving rats," *Science*, vol. 362, no. 6414, pp. 584–589, 2018.

37. W. E. Allen, M. Z. Chen, N. Pichamoorthy, R. H. Tien, M. Pachitariu, L. Luo, and K. Deisseroth, "Thirst regulates motivated behavior through modulation of brainwide neural population dynamics," *Science*, vol. 364, no. 6437, pp. 253–253, 2019.

38. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

39. U. Von Luxburg, "A tutorial on spectral clustering," *Statistics and computing*, vol. 17, no. 4, pp. 395–416, 2007.

40. D. Eriksson, M. Heiland, A. Schneider, and I. Diester, "Slow and fast cortical dynamics distinguish motor planning and execution," *bioRxiv*, 2019.

Acknowledgments

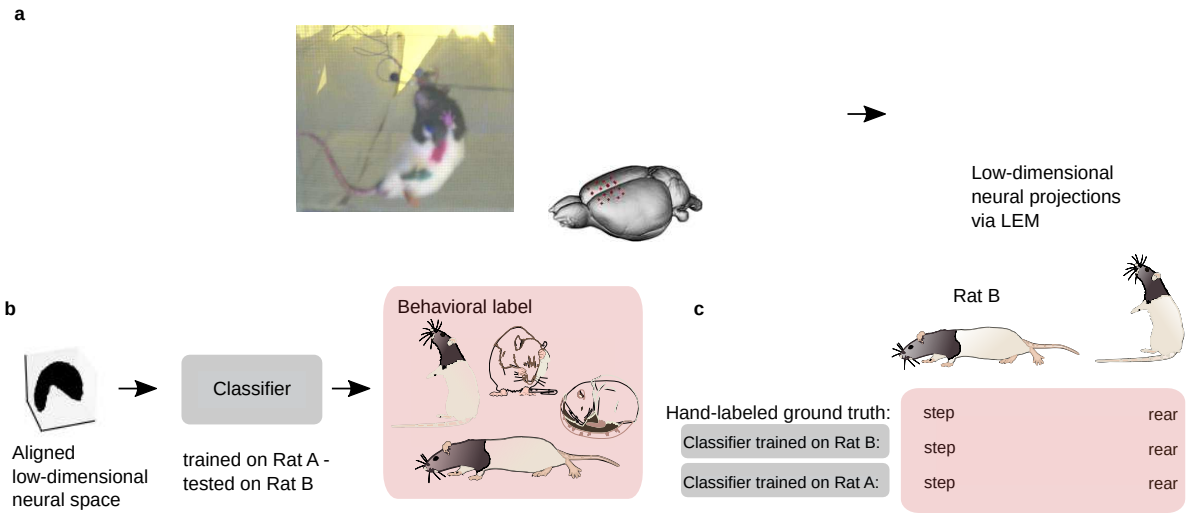
We thank Krishna Shenoy, Artur Schneider, Philipp Schröppel and Christian Zimmermann for comments on the manuscript. **Funding:** This work was supported by BrainLinks-BrainTools, Cluster of Excellence funded by the German Research Foundation (DFG, grant number EXC 1086) and Bernstein Award 2012. **Author contributions**: T.B. and I.D. designed the study. S.M. analyzed the data. S.M., T.B., and I.D. wrote the manuscript. D.E. conceived and recorded the data set. **Competing Interests:** The authors declare that they have no competing financial interests. **Data and materials availability:** Code and data are available from the authors upon reasonable request.

Supplementary materials

Methods

Figs. S1 to S8

Table S1 to S2


a
Low-dimensional
neural projections
via LEM
b
Aligned
low-dimensional
neural space
Classifier
trained on Rat A -
tested on Rat B
Behavioral label
c
Rat B
Hand-labeled ground truth:
Classifier trained on Rat B:
Classifier trained on Rat A:
step
rear
step
rear
step
rear

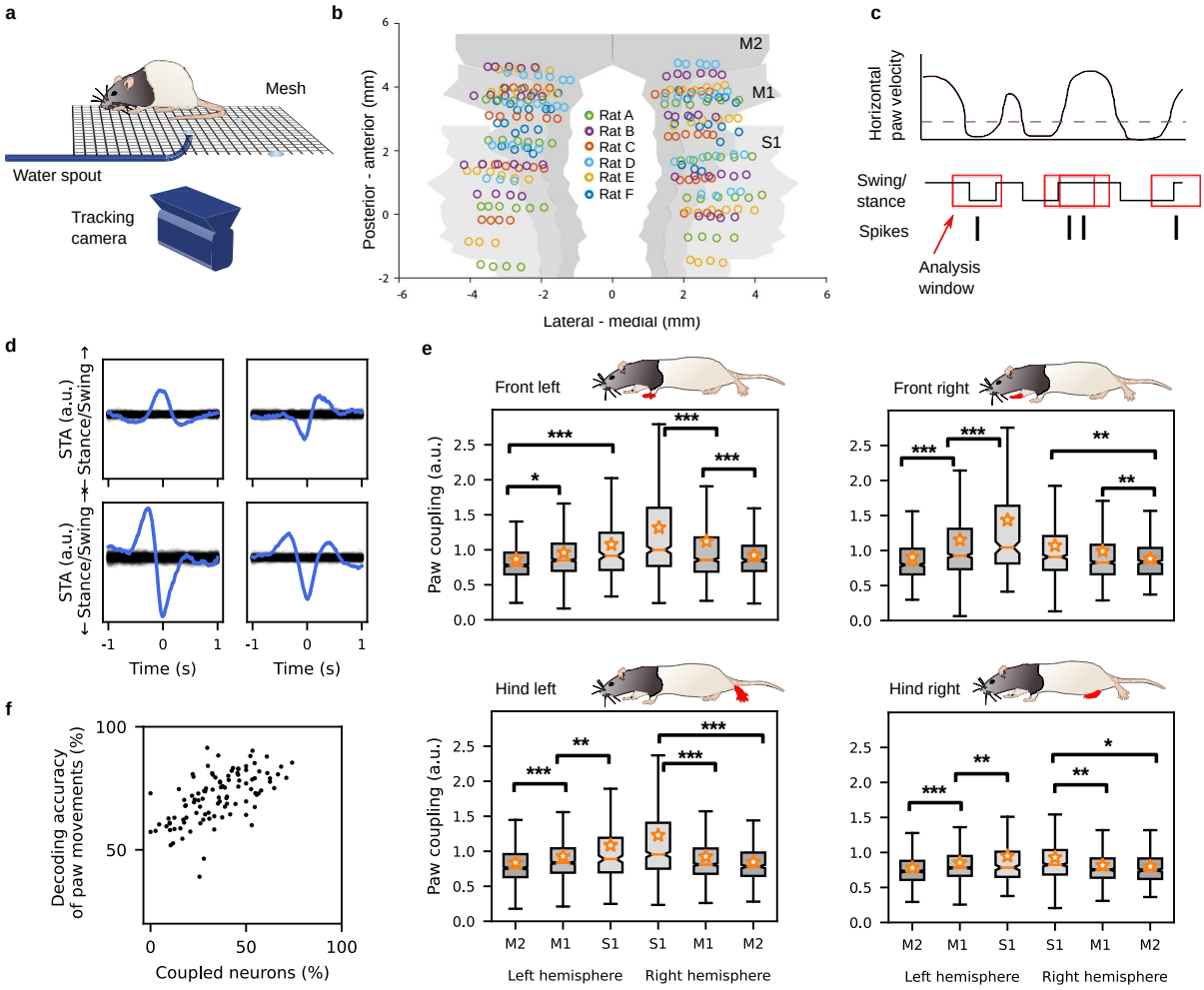


Figure 1: **Spike-triggered Average Paw Swing-Stance Status (STAPSSS) during unconstrained movements.** (a) Behavioral setup with ground mesh, camera, and robot arm delivering water drops, adapted from (*40*). (b) Locations of electrodes of the six implanted rats, adapted from (*40*). (c) Paw movements were binarized into swing (moving) and stance (not moving). STAPSSS was calculated by averaging the swing-stance status in windows $\pm$ 1s (indicated with red boxes) around each spike. (d) STAPSSS for the right front paw of four example single units in left and right S1 (upper panel) and left and right M1 (lower panel). Black lines refer to the statistical control waveforms. (e) Coupling for each paw, brain area, and hemisphere, averaged over neurons. Stars denote results of post-hoc Tukey-Kramer tests (only intra-hemispheric results are indicated). Orange stars denote mean values, notches the 95% confidence intervals for the median. See main text for definition of paw coupling. $*p < .05, **p < .01, ***p < .001$. (f) Accuracies of neural networks trained to predict the status of the right front paw from the neural data were strongly correlated to the percentage of significantly coupled neurons.

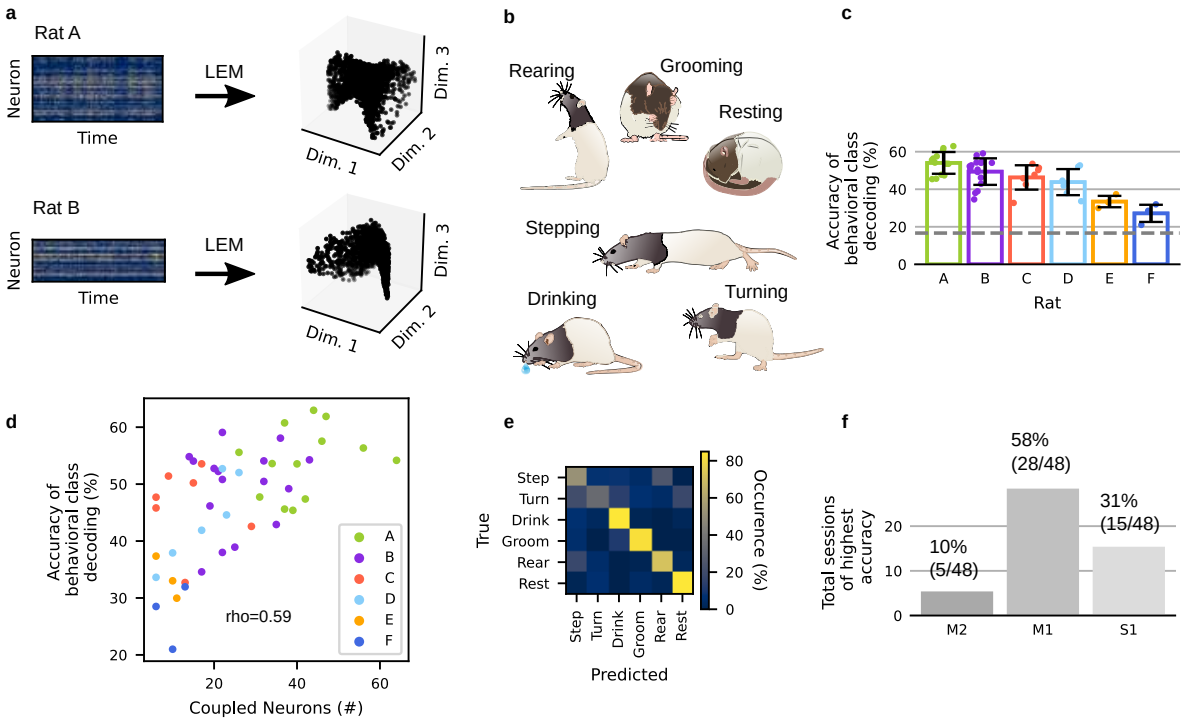


Figure 2: **Behavioral within-session decoding from low-dimensional neural structures.** (a) Non-linear dimensionality reduction through Laplacian Eigenmaps was performed on the neural data of each session separately. (b) The animals' behavior was separated into six different behavioral classes. (c) Classification accuracies for the six behavioral classes given low-dimensional neural input were above chance level for sessions from all six rats. The grey dashed line refers to chance level, errorbars to the standard deviations. (d) Accuracies were correlated to the number of significantly coupled neurons (neurons coupled to at least one paw according to the STAPSSS measure). (e) One exemplary confusion matrix for the test set of one session of Rat A, mean per-class accuracy 68.46%. (f) For most of the sessions, classification accuracies for the six behavioral classes where highest given dimensionality-reduced neural activity from M1 as input, followed by S1 and M2.

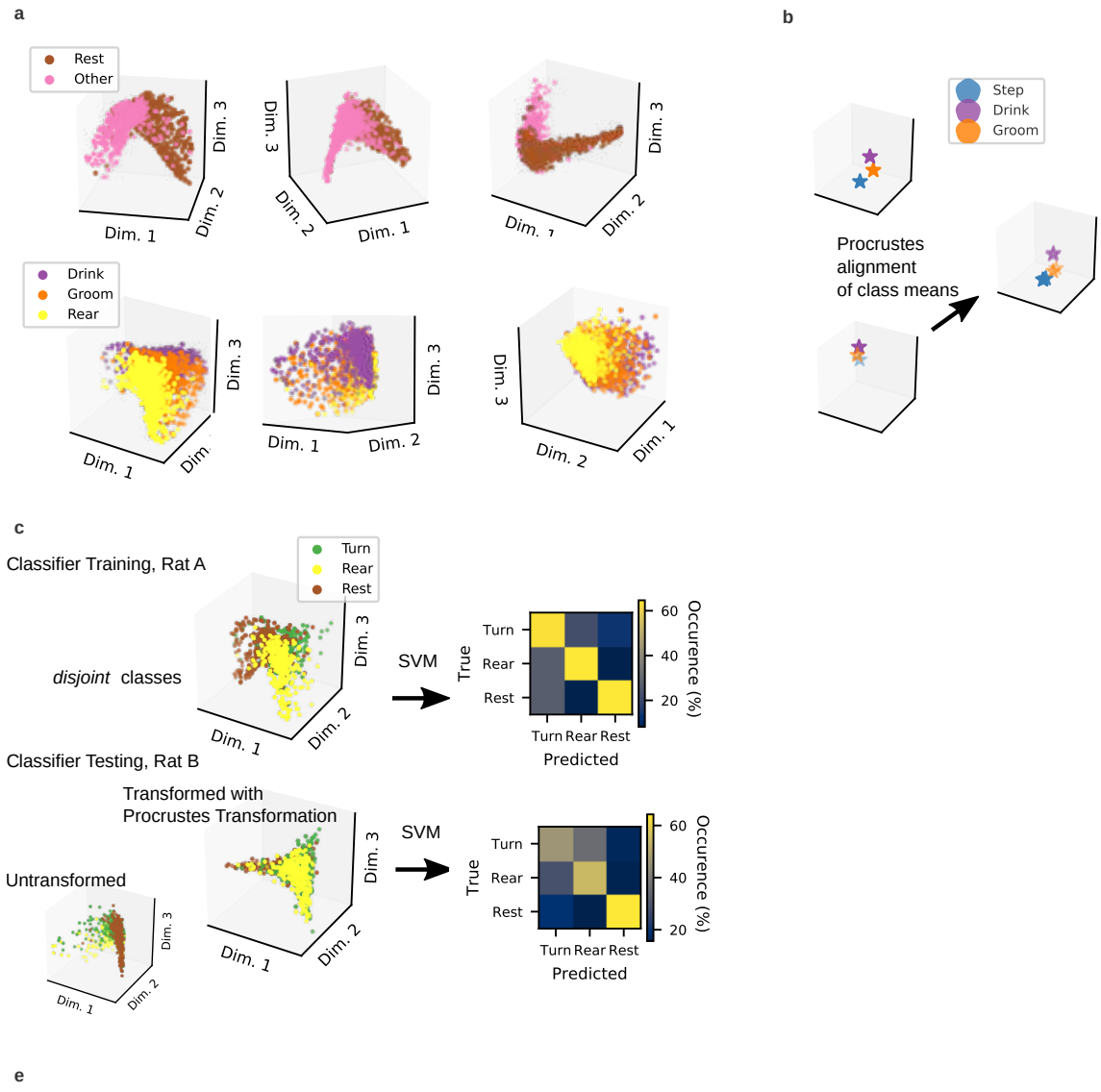


Figure 3: **Similar structure of population activities allows cross-subject decoding.** (Continued on the following page.)

Figure 3: **Similar structure of population activities allows cross-subject decoding.** (a) In the low-dimensional space, different behaviors were distinguishable in as little as three neural dimensions. Upper panel: The first dimension clearly differentiated between rest and movement (all other behavioral classes). Lower panel: Mainly the second/third dimension coded a difference between paw- and head-related behavior (rear vs drink). One session from Rat A, B, and C are depicted. (b) As a first step for our cross-subject generalization method, neural structures were aligned by using a Procrustes transformation between class means for four and three behavioral classes, respectively. (c) A classifier was trained on the two, respective three remaining classes of one rat. The classifier could generalize to the transformed neural data from another rat. (d) Mean per-class accuracies across training and test sessions when aligning on four and testing on two classes. The chance level was 50%. Lower panel: On the diagonal, training and test data came from the same session. Off-diagonal entries refer to tests on data sets which were not identical to the training session. Values are averaged over 20 runs and 15 possible splits of the six behavioral classes into align/classify classes. Upper panel: Averages across train sessions of each rat for tests on all other rats (block-wise averages of the matrix in the lower panel). Error bars provide the standard error of the mean. (e) Within-session and between-session accuracy were highly correlated.

Conserved structures of neural activity in sensorimotor cortex of freely moving rats allow cross-subject decoding – Supplementary Materials

Svenja Melbaum[1,2], David Eriksson[2,3], Thomas Brox[1,2], Ilka Diester[2,3,4*]

[1]Computer Vision Group, Dept. of Computer Science,
University of Freiburg, 79110 Freiburg, Germany.
[2]BrainLinks-BrainTools Cluster of Excellence,
University of Freiburg, Georges-Köhler-Allee 201, IMBIT, 79110 Freiburg, Germany.
[3]Optophysiology Lab, Faculty of Biology,
University of Freiburg, 79110 Freiburg, Germany.
[4]Bernstein Center Freiburg, University of Freiburg, 79104 Freiburg, Germany.

*To whom correspondence should be addressed; E-mail: ilka.diester@biologie.uni-freiburg.de

Methods

Animal surgery We implanted six male Long Evans rats at the age of eight weeks with 22 tungsten electrodes (200 to 600 kOhm impedance, polyimide insulation, WHS Sondermetalle, Grünsfeld, Germany) in 1.2 mm implantation depth in each hemisphere (implantation January 2017 for Rat F, April 2017 for Rat A-E). Electrode locations spanned from -2 to +5mm in anterior/posterior direction and from 1 to 4mm in lateral/medial direction: three medial-lateral rows of six electrodes each, plus one row of four electrodes (see Fig. 1B). Details of the procedure are described elsewhere (*40*). All animal procedures were approved by the Regierungspräsidium Freiburg, Germany.

Behavioral task The rats were kept water-restricted for the time course of the experiments (free access to water for two days per week). For the experiments, the rats moved unconstrained on a mesh of 30×40 cm in a closed arena. Every 10 to 30s a waterspout positioned by two servo motors in a pseudo-randomized way released a drop of water in the mesh, which the animals could find and consume. To prevent the rats from merely following the movements of the waterspout, we included dummy movements of the waterspout that were not followed by a release of water. Even experienced animals were not able to predict the position of water drops without an active search, which was indicated by the fact that animals did not find all water drops throughout a session. This task has been previously described (*40*). Here, we only used part of that data set, i.e. only sessions with a minimum duration of 30 minutes.

Data acquisition and preprocessing of extracellular recordings Extracellular signals were recorded at 30 kHz and band-pass filtered, amplified and digitized using the INTAN (Intan Technologies, Los Angeles, California) head stage situated at the head of the animal. Spike sorting was conducted on high-pass filtered signals (cut-off at 300 Hz) and for each each electrode

separately. Spikes were detected as amplitude threshold crossings of four times the standard deviation of the signals. For each spike, the window of -0.5 to 2 ms around the peak amplitude was extracted (resulting in 76 values per spike). Spike sorting consisted of two phases for each unit. First, a seed spike was estimated. This was accomplished by calculating the spike neighborhoods (spikes within average noise level, half a ms before the spike, across all units) for 500 randomly chosen spikes. The spike with the most neighbors was chosen as seed spike. Second, an iterative procedure optimized the spike waveform. This was done by alternating the calculation of a new noise level for the neighboring spikes, the update of the neighborhood (spikes within the new noise level), and the update of the average waveform. This iterative procedure ended when the neighborhood assignments remained constant. The algorithm proceeded with the remaining spikes by choosing a new seed spike. Details of the offline denoising and spike sorting procedure have been described elsewhere (*40*). For our single-unit analysis, we only kept single units according to the distribution of inter-spike intervals. Single units with a firing rate lower than 0.1 Hz were not included in the analysis. Movements of the colored paws were tracked by two cameras (Stingray, F033C IRF CSM, Allied Vision Technologies) positioned below the mesh. The videos were taken with a 80 Hz frame rate and smoothed with a Gaussian filter before analysis.

Single-unit analysis Paw movements were labeled as 'swing' for a horizontal velocity larger than 0.3 mm per 10 ms, the bin size we used for our analysis, and 'stance' otherwise. Spikes were also binned with 10 ms bin size. For each neuron and each paw, we defined the spike-triggered average paw swing-stance status (STAPSSS) as the behavioral average over all windows ± 1 s around the spikes. We normalized each STAPSSS waveforms by the mean. We defined the paw coupling of a neuron as the ratio of the standard deviation of the STAPSSS waveform and the statistical control standard deviation, which was defined as the .99 quan-

tile standard deviation of a distribution constructed out of standard deviations of the STAPSSS waveforms of 1000 randomly shifted spike trains. If a neuron was not related to a paw's movement, its STAPSSS waveform would be flat, with the standard deviation not exceeding the control standard deviation. We defined the contralateral bias as the ratio of contra- and ipsilateral paw coupling. Statistical analyses were done with Matlab *anovan*, *multcompare* and *ttest* functions. ANOVA tests always included the rat's id as additional factor.

Decoding from spike trains We used fully-connected neural networks with three hidden layers with 500 units each for decoding. Input were the Gaussian-smoothed ($\sigma = 20$ms) binned spikes in ± 400 ms, resulting in 81 input bins for each neuron. In contrast to the STAPSSS analysis, where only single units were considered, we used all units as input for decoding. Each session was split into training, validation and test set (70/15/15 %). Two of the 106 sessions were excluded in the decoding because of too few data. Training was conducted with the Adam optimizer (*38*), batch size 64 and an initial learning rate of 0.0001. A dropout rate of 75 %, l2-regularization ($\lambda = 1e-4$) and early stopping were applied to prevent overfitting. To deal with class imbalance, we used a weighted cross-entropy loss to put more weight on the less frequent class (swing). Reported accuracies are mean per-class accuracies. The decoding accuracies of the deep neural network were significantly better than a baseline linear classifier (two-sided paired t-test, $t = 6.55, p < 1\text{e}{-8}$). For the baseline, we used logistic regression with 3-fold cross-validation of the l2 regularization strength on the concatenated train and validation set. Test sets for each session were the same as for the artificial neural network. Class weights were adjusted inversely proportional to class frequencies, as for the artificial neural network. The artificial neural network was implemented in Tensorflow. For the linear baseline we used Python's scikit-learn function *LogisticRegressionCV*.

Dimensionality reduction We used Laplacian Eigenmaps (*22, 5*) to investigate the low-dimensional structure of the population activity. For each session, spike counts were binned in 100 ms bins and then binarized (1 for at least one spike per bin, 0 for no spikes). Single and multi units were used. Only time points with at least 15 active units were kept. Since we restricted the further analysis to sessions with at least 5000 valid time points, we considered only 95 of the 106 sessions. For each session, we constructed an unweighted, mutual knn-graph based on the Hamming distance on the columns of the $n \times t$ matrix (n units, t time points). Our code for LEM built on code from recent work (*5*). Two iterations of the Laplacian eigenmap algorithm were performed, but in contrast to Rubin et al. we used the Hamming distance in the first iteration and reduced to 20 dimensions in the first step. Furthermore, we applied the random walk normalized Laplacian instead of the symmetric normalized Laplacian as proposed in a previous study (*39*). In detail, we constructed the unnormalized graph Laplacian as $L = D - W$, with D the diagonal degree matrix and W the adjacency matrix of the knn graph. Solving the generalized eigenvalue problem $Lv = \lambda Dv$ corresponded to finding the first eigenvectors of the random walk normalized graph Laplacian $L_n = D^{-1}L$ (*39*). For the LEM reductions on units from different sensorimotor areas, we randomly chose 20 units from each area as input (if less than 20 units for an area were available, the analysis was omitted) and reduced to six dimensions in the first step. For the ablation study on sessions with 20/40/60/80 units removed, we reduced to 20 dimensions in the first two and 10 dimensions in the second two cases. For the study on LEM reductions after removal of sensorimotor areas, we removed n_{max} =max(#M1, #M2, #S1 units) from each area, for each session. For underrepresented areas, we additionally discarded $n_{max} - n_{area}$ randomly chosen units. As before with a lower number of neurons, we also reduced to 10 dimensions. Since the eigenvector corresponding to the smallest eigenvalue (zero) is constant, we discarded the first dimension of LEM for all analyses and decoding studies.

Behavioral Labeling Behavioral labeling of the videos was conducted with the freely available tool MuViLab. Two human annotators that were blinded to the neural data manually labeled 48 sessions, cut in 500 ms snippets. The 48 sessions were chosen based on a clear saddle-like shape and at least five significantly coupled units: Rat A - 13 sessions recorded between 2017/06/08 and 2017/08/03, Rat B - 16 sessions between 2017/06/01 and 2017/08/21, Rat C - 7 sessions between 2017/06/01 and 2017/06/29, Rat D - 6 sessions between 2017/06/08 and 2017/07/11, Rat E - 3 sessions between 2017/06/08 and 2017/06/22, Rat F - 3 sessions between 2017/06/07 and 2017/06/30. Criteria for the behavioral classes were: Step - the rat moved at least one paw, but did not drink or reared at the same time; Turn - the rat moved its head; Drink - the rat drank from the spout or collected water drops from the mesh with its mouth; Groom - the rat performed typical grooming movements; Rear - the rat stood on its hind paws; Rest - no obvious movements. Seldom, samples were excluded from labeling when the behavior of the rat was not clearly visible because it was located near the borders of the arena. Examples for the different behaviours can be found at `https://www.dropbox.com/sh/4uu3cmmmnnovqmb/AABWaTv9H_0MPgHOpx4tPOXwa?dl=0`.

Population-level decoding We trained one deep neural network per session to classify the six behavioral classes given the ten-dimensional neural data in seven bins with 100 ms each as input. The data was min-max normalized (min and max calculated only on training sets). The deep network architecture and training were almost identical to the network used for the decoding task above, yet we used only 200 units per layer, a dropout rate of 25% and chose a cross-validation strategy to deal with unbalanced classes. For this, the available data was split in four equally large parts. Four runs were conducted per session, using two parts as training set, one as validation set for the early stopping, and the fourth as test set. The final test results were calculated as mean over all four test sets and runs. For the area-specific dimensionality-reduced

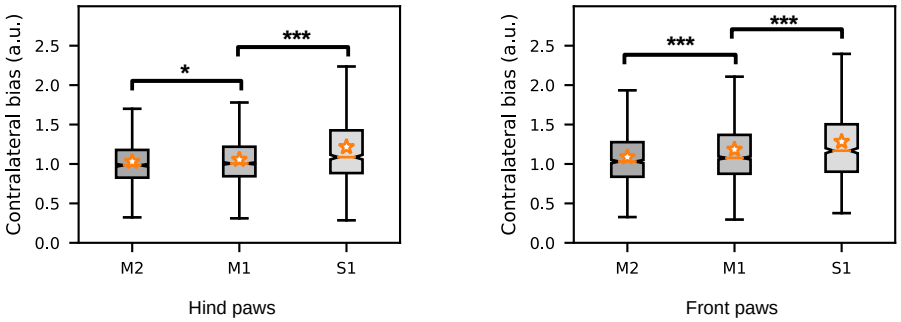


Figure S1: **Contralateral bias largest in S1.** Contralateral bias for the front and hind paws depending on area, averaged over neurons. The bias increased from anterior to posterior regions for both the front and the hind paws. Stars denote results of post-hoc Tukey-Kramer tests. Orange stars denote mean values, notches the 95% confidence intervals for the median. See main text for definitions of paw coupling and bias. $*p < .05, ***p < .001$.

Table S1: **Statistics of the recording sessions**. Dates of implantations as well as recording periods per animal.

	Rat A	Rat B	Rat C	Rat D	Rat E	Rat F
Implantation	20/04/2017	19/04/2017	27/04/2017	11/04/2017	25/04/2017	01/01/2017
First recording	01/06/2017	01/06/2017	01/06/2017	01/06/2017	01/06/2017	07/06/2017
Last recording	15/08/2017	21/08/2017	08/07/2017	21/08/2017	25/08/2017	22/08/2017

data, only five dimensions were used as input for the neural networks. For the supervised alignment procedure, we restricted the analysis to four neural dimensions to avoid underdetermination. We used the Matlab function *Procrustes* to find a transformation between class means. Proper transformation was important because of the sign ambiguity of eigenvectors, which might otherwise lead to different orientations of the neural structures. Before aligning, both neural structures were normalized to the [0-1] range. A SVM with Gaussian kernel (Matlab *fitcecoc*) was used as classifier. Training was conducted with an equalized number of samples per class and default parameters.

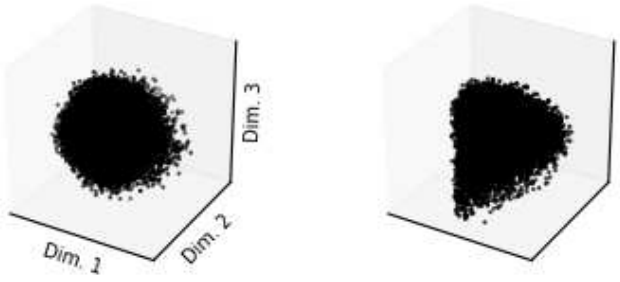


Figure S2: **Some population structures did not show any apparent structure.** Two example sessions (from Rat A and Rat B, respectively) with random-like low-dimensional neural projections. Refers to main paper, Fig. 2a.

Table S2: **ANOVA results for paw coupling**. Paw coupling was defined as the quotient between the STAPSSS standard deviation and the control standard deviation (see main text). Three-way ANOVAs were calculated separately for each paw, on all recorded neurons ($n = 3723$, main effects area, hemisphere, rat; interaction effect area and hemisphere). The table contains the corresponding F and p values.

Paw	Area	Hemisphere	Area x Hemisphere	Rat
Right front	$66.77, p < 1\text{e}{-28}$	$108.85, p < 1\text{e}{-24}$	$18.24, p < 1\text{e}{-07}$	$28.59, p < 1\text{e}{-27}$
Left front	$41.61, p < 1\text{e}{-17}$	$17.15, p < 1\text{e}{-4}$	$2.22, p = .10$	$37.82, p < 1\text{e}{-37}$
Right hind	$25.73, p < 1\text{e}{-11}$	$4.16, p = .04$	$5.63, p = .003$	$23.63, p < 1\text{e}{-22}$
Left hind	$67.47, p < 1\text{e}{-28}$	$6.38, p = .01$	$3.82, p = .02$	$13.52, p < 1\text{e}{-12}$

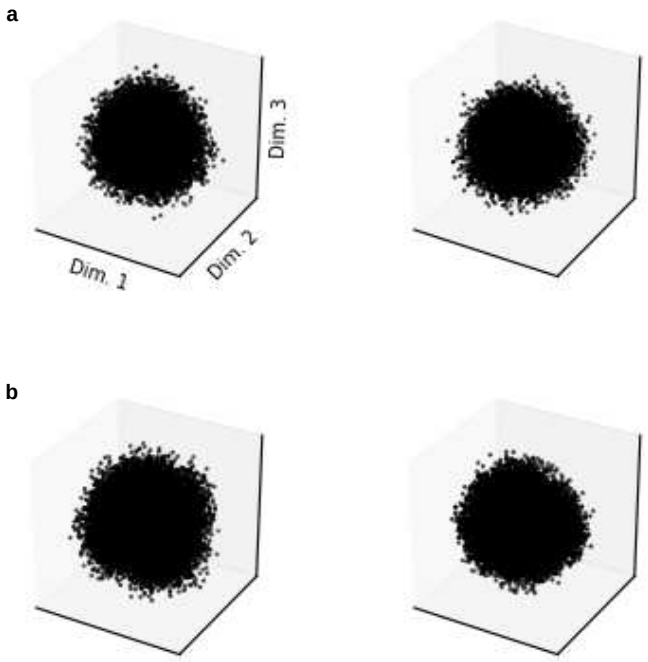


Figure S3: **Control dimensionality reductions with shuffled neuronal activity did not show any apparent structure.** LEM projections for neuron-shuffled data (a) respective time-shuffled data (b) for one session of Rat F (left) and Rat A (right). For neuron-shuffling, units were permuted randomly for each timepoint. For time-shuffling, timepoints were permuted randomly for each neuron.

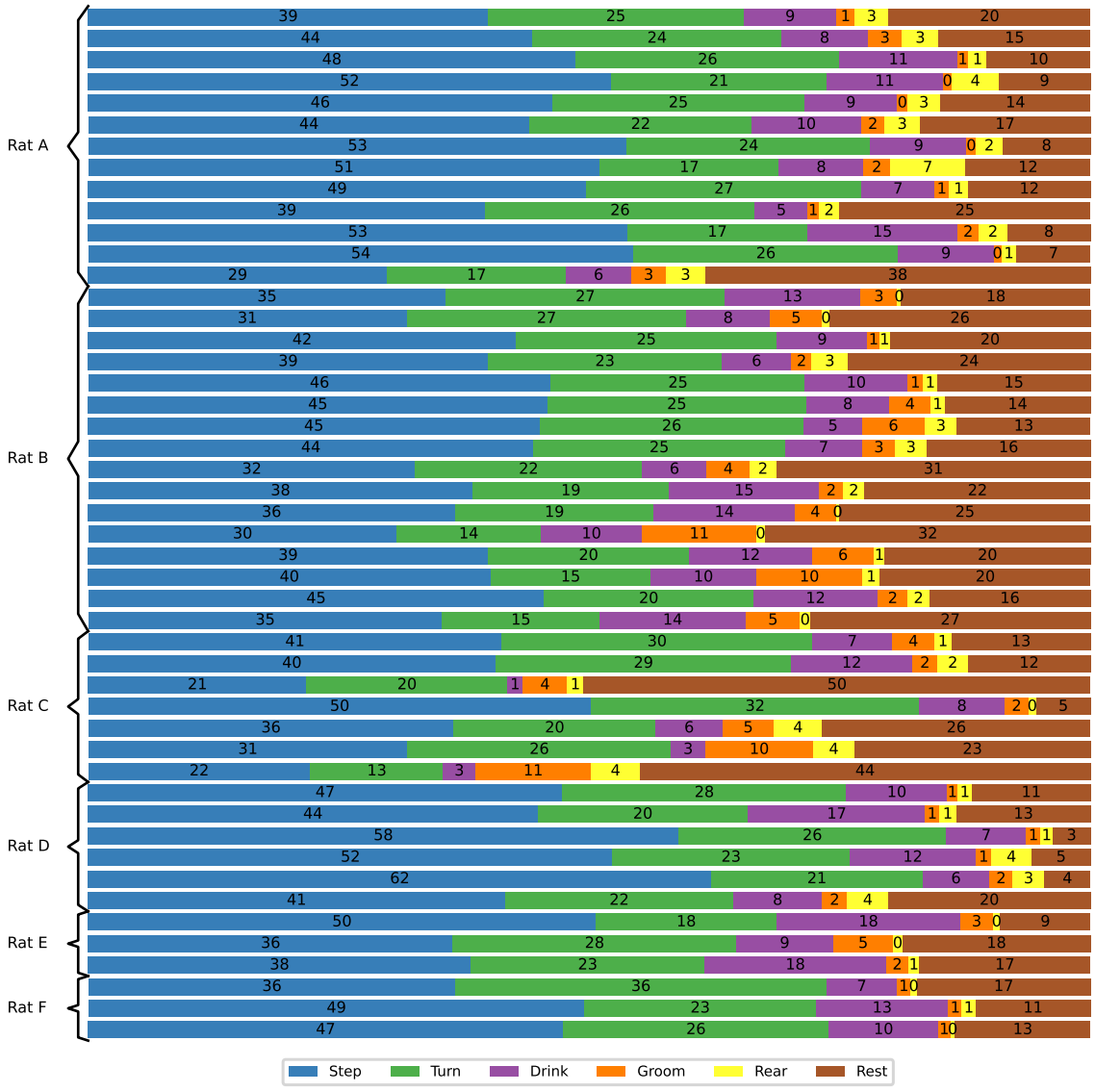


Figure S4: **Histograms of behavior illustrating the distribution of behavioral classes.** One row per session. Provides background for Fig. 2 in main paper.

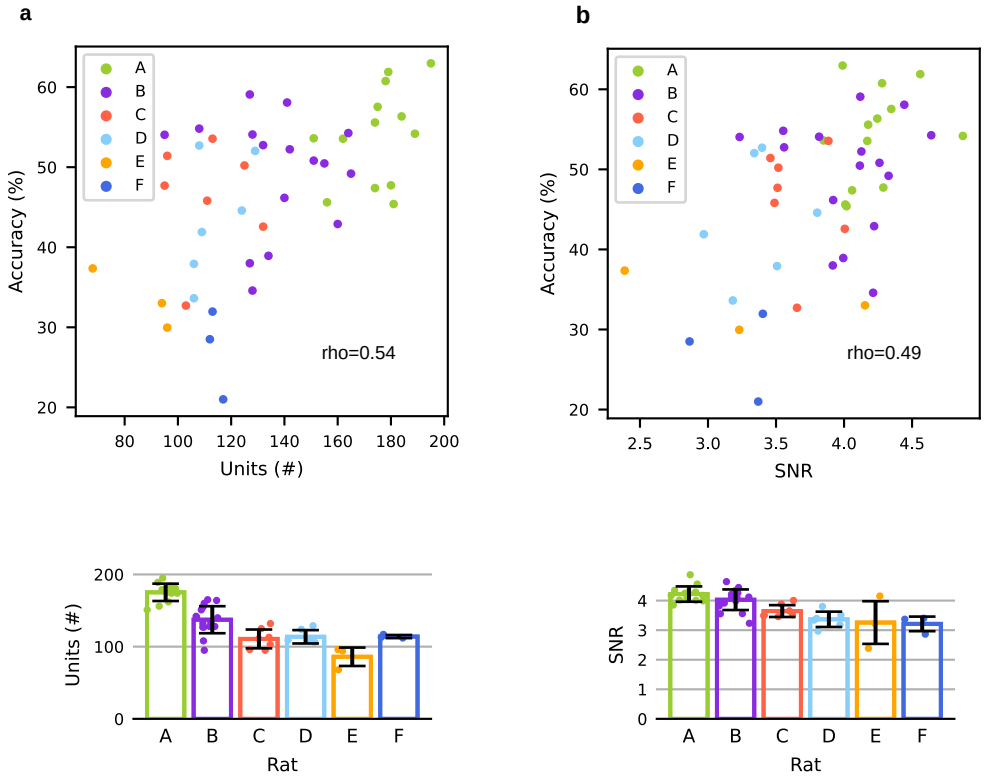


Figure S5: **Correlation between accuracies and units/SNR.** (a) Top: Accuracies vs. number of units for sessions from the six rats. Bottom: Average number of units per rat, errorbars for standard deviation across sessions. (b) Top: Accuracies vs. mean SNR for sessions from the six rats. Bottom: Average SNR per rat, errorbars for standard deviation. Refers to Fig. 2c-d.

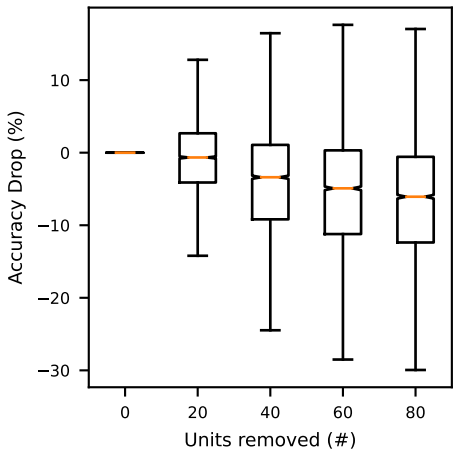


Figure S6: **Generalization worsened with less units.** We repeated the generalization experiment from Fig.3d for all sessions with a generalization accuracy of at least 55% (19 sessions). Accuracy dropped for LEM structures which were computed after removing 20,40,60 or 80 units from each session compared to accuracies with the full number of units. The fewer units, the lower the accuracies.

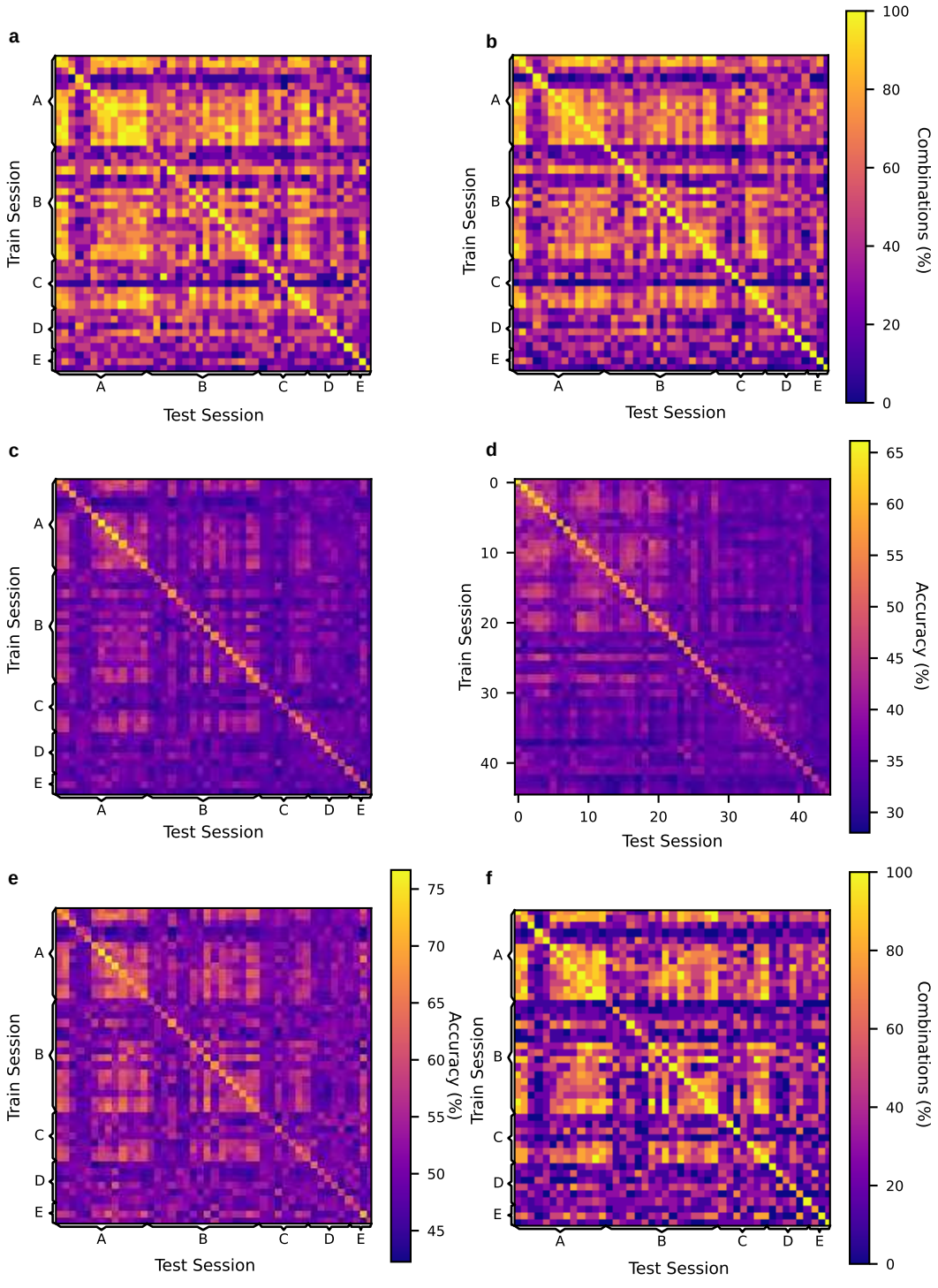

a
b
c
d
e
f
Train Session
Test Session
A
B
C
D
E
Combinations (%)
Accuracy (%)
0
10
20
30
40
20
35
45
50
55
60
65
75
80
100

Figure S 7: **Further cross-subject and cross-session generalization experiments.** We conducted further generalization experiments with a more difficult setting (align on three classes, classify three classes, b-d; experiment without the class "rest", e-f). For all plots, training and test data on the diagonal was from the same session. Off-diagonal entries show testing on data different from the training session. (a) Percentage of splits of the six behavioral classes into "align" and "decode" set (out of 15) that were classified with a significantly higher mean per-class decoding accuracy than chance (50%). Significance was calculated over 20 different training runs, on a .05 significance level with Bonferroni correction, using a one-tailed sign test. Refers to the experiment in main paper, Fig. 3e. (b) Same as in (a), but for the experiment with aligning on three classes and testing with three classes, with 20 combinations in total and chance level of 33.33%. (c) Mean per-class accuracies across training and test sessions when aligning on three and testing on three classes. The chance level was 33.33%. Values were averaged over 20 runs and 20 possible splits of the six behavioral classes into align/classify classes. (d) Same data as in (c), but sorted by accuracy on the diagonal to illustrate the relation between in-session and across-session accuracy. (e) Same as (c) for aligning on three and testing on two classes, without class "rest". The chance level is 50%. Values are averaged over 20 runs and 10 possible splits of the six behavioral classes into align/classify classes. (f) Same as (a) and (b) for aligning on three and testing on two classes, without class "rest". Extends Fig. 3 in main paper.

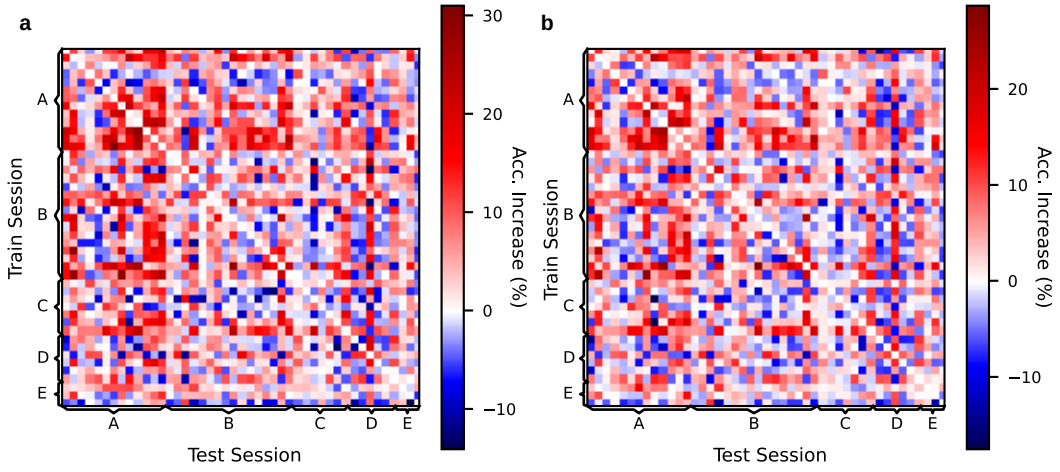


Figure S8: **Decoding accuracy gain through neural manifold alignment.** Accuracy gains for aligned versus unaligned neural structures, for decoding of two (a) respective three (b) classes. Values are averaged over all class combinations. In most cases, the accuracies were higher after alignment (red color spectrum), up to 20-30%. Refers to Fig. 3d in main paper and Fig. S7c.

Figures

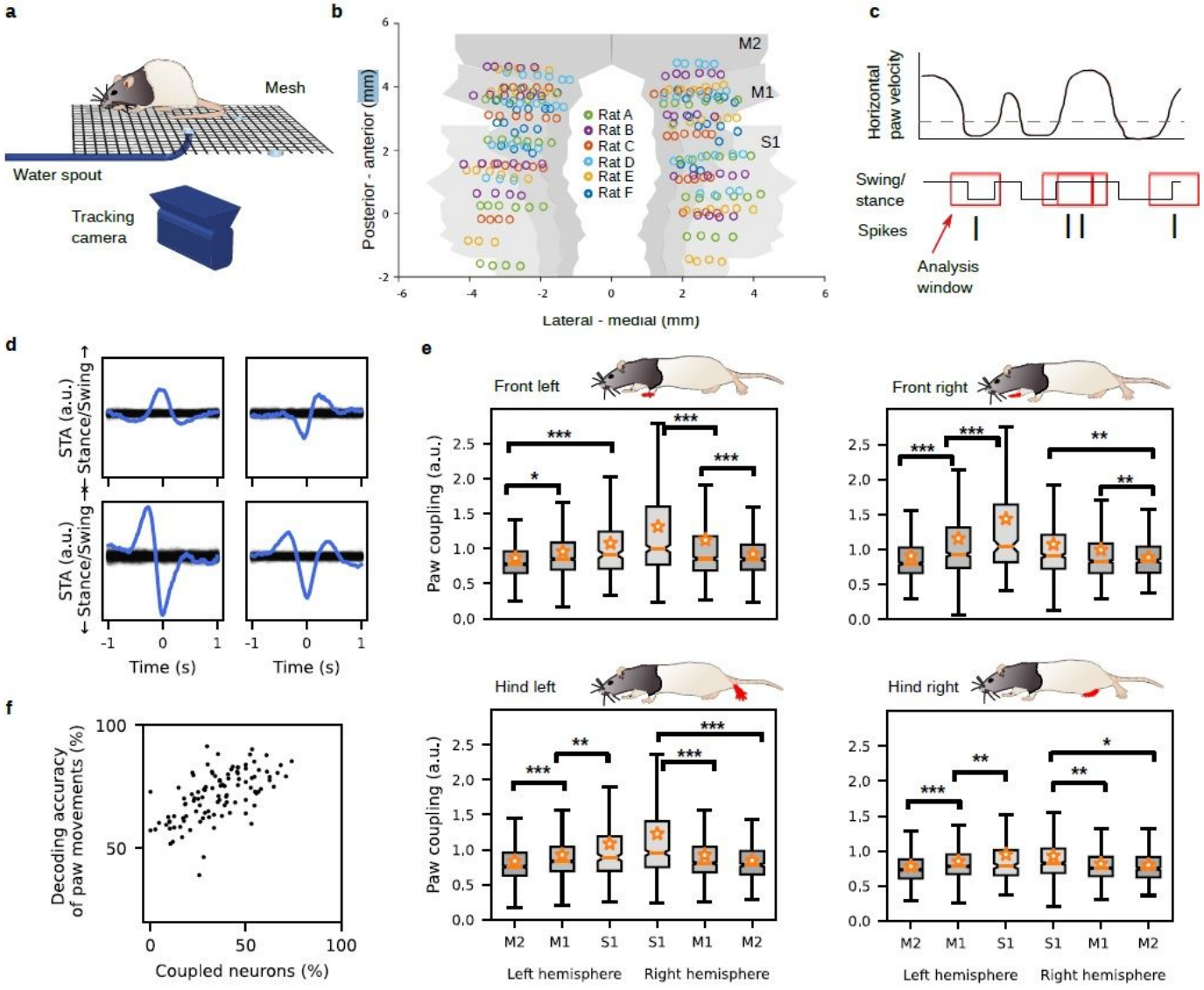


Figure 1

Spike-triggered Average Paw Swing-Stance Status (STAPSSS) during unconstrained movements. (a) Behavioral setup with ground mesh, camera, and robot arm delivering water drops, adapted from (40). (b) Locations of electrodes of the six implanted rats, adapted from (40). (c) Paw movements were binarized into swing (moving) and st + 1s (indicated with red boxes) around each spike. (d) STAPSSS for the right front paw of four example single units in left and right S1 (upper panel) and left and right M1 (lower panel). Black lines refer to the statistical control waveforms. (e) Coupling for each paw, brain area, and hemisphere, averaged over neurons. Stars denote results of post-hoc Tukey-Kramer tests (only intra-hemispheric results are indicated). Orange stars denote mean values, notches the 95% confidence intervals for the median. See main text for definition of paw coupling. *p < :05; **p < :01; ***p < :001. (f)

Accuracies of neural networks trained to predict the status of the right front paw from the neural data were strongly correlated to the percentage of significantly coupled neurons.

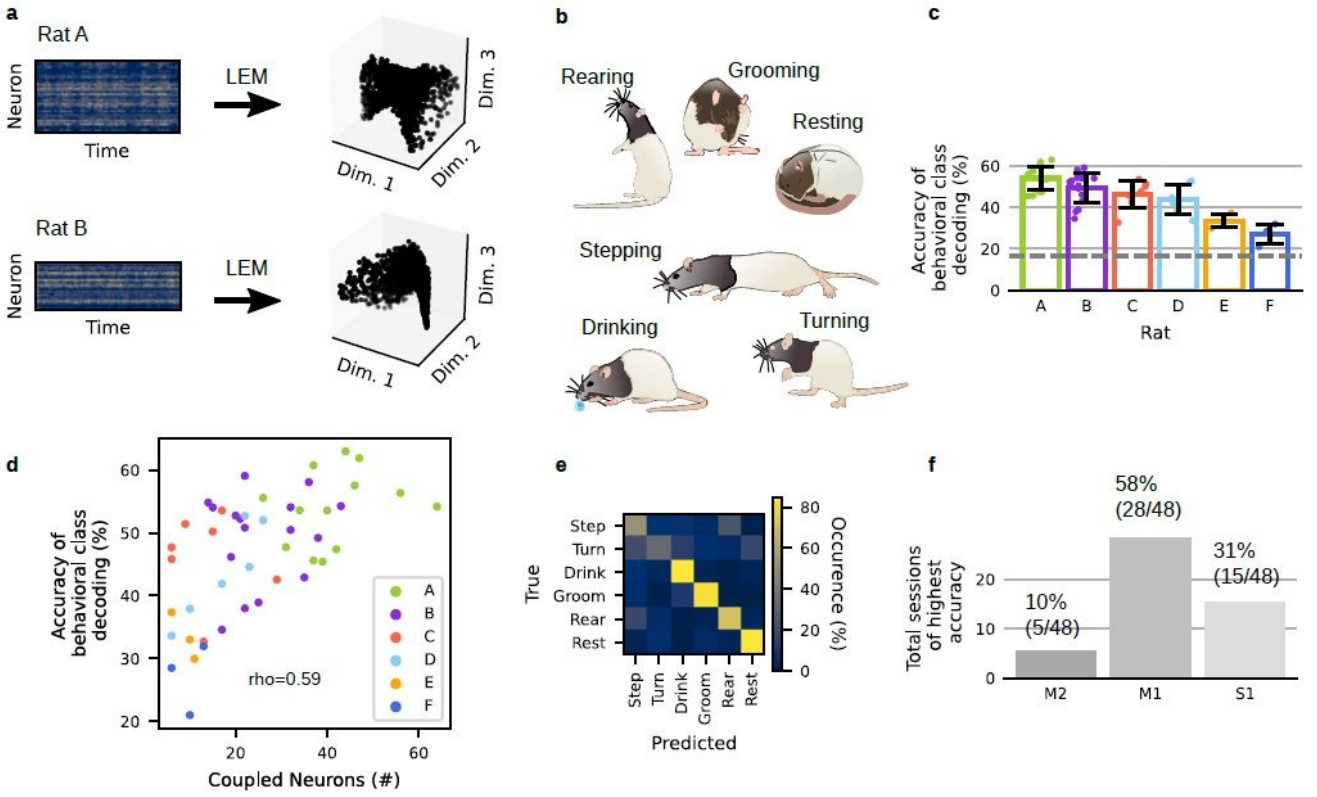


Figure 2

Behavioral within-session decoding from low-dimensional neural structures. (a) Non-linear dimensionality reduction through Laplacian Eigenmaps was performed on the neural data of each session separately. (b) The animals' behavior was separated into six different behavioral classes. (c) Classification accuracies for the six behavioral classes given lowdimensional neural input were above chance level for sessions from all six rats. The grey dashed line refers to chance level, errorbars to the standard deviations. (d) Accuracies were correlated to the number of significantly coupled neurons (neurons coupled to at least one paw according to the STAPSSS measure). (e) One exemplary confusion matrix for the test set of one session of Rat A, mean per-class accuracy 68.46%. (f) For most of the sessions, classification accuracies for the six behavioral classes where highest given dimensionality-reduced neural activity from M1 as input, followed by S1 and M2.

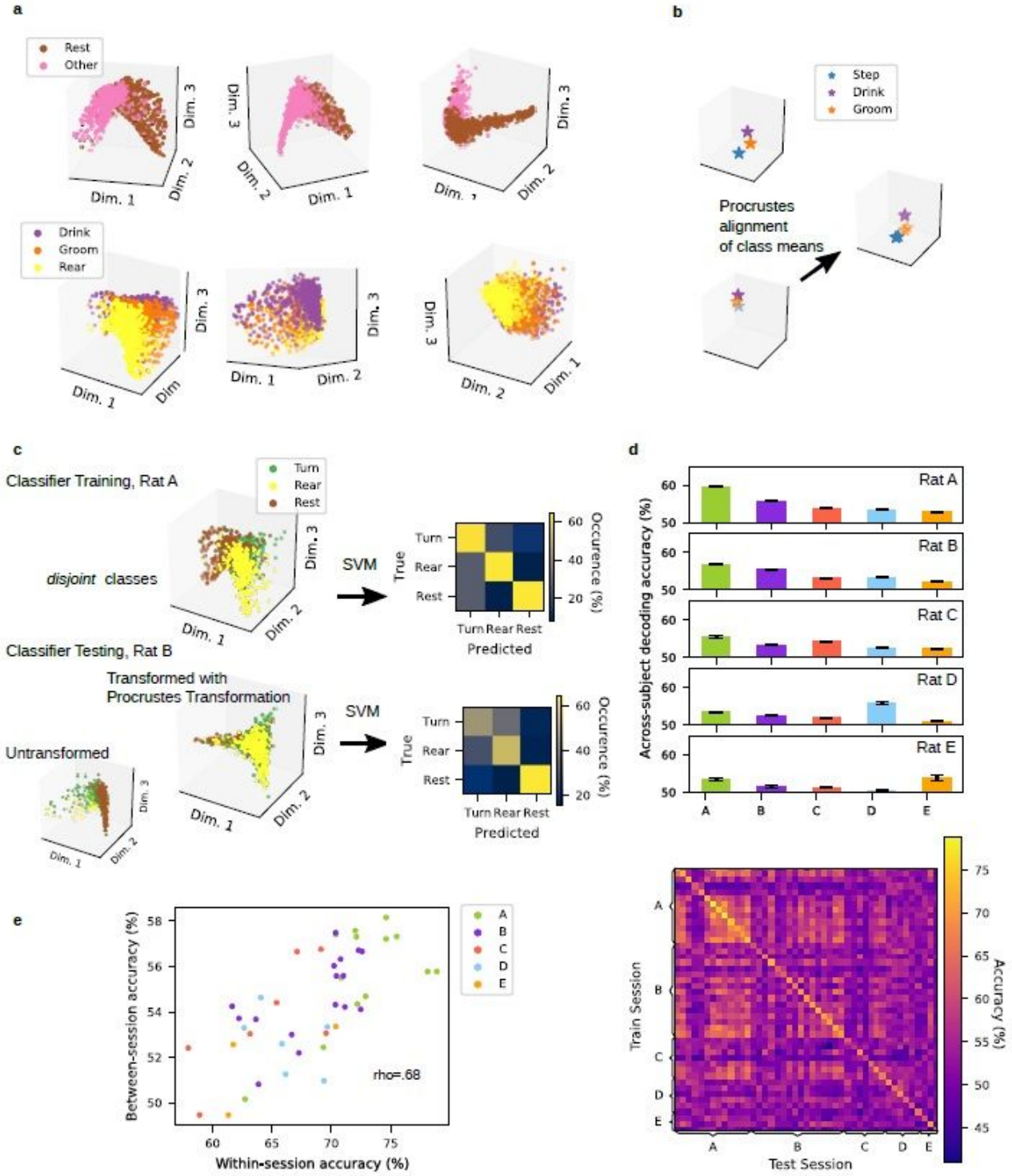


Figure 3

Similar structure of population activities allows cross-subject decoding. (a) In the low-dimensional space, different behaviors were distinguishable in as little as three neural dimensions. Upper panel: The first dimension clearly differentiated between rest and movement (all other behavioral classes). Lower panel: Mainly the second/third dimension coded a difference between paw- and head-related behavior (rear vs drink). One session from Rat A, B, and C are depicted. (b) As a first step for our cross-subject

generalization method, neural structures were aligned by using a Procrustes transformation between class means for four and three behavioral classes, respectively. (c) A classifier was trained on the two, respective three remaining classes of one rat. The classifier could generalize to the transformed neural data from another rat. (d) Mean per-class accuracies across training and test sessions when aligning on four and testing on two classes. The chance level was 50%. Lower panel: On the diagonal, training and test data came from the same session. Off-diagonal entries refer to tests on data sets which were not identical to the training session. Values are averaged over 20 runs and 15 possible splits of the six behavioral classes into align/classify classes. Upper panel: Averages across train sessions of each rat for tests on all other rats (block-wise averages of the matrix in the lower panel). Error bars provide the standard error of the mean. (e) Within-session and between-session accuracy were highly correlated.