

1 **Flexible analysis of animal behavior via time-resolved manifold embedding**

2 Ryan A. York^{1*}, Arnaldo Carreira-Rosario¹, Lisa M. Giocomo¹, and Thomas R.

3 Clandinin^{1*}

4

5 **Affiliations**

6 ¹Department of Neurobiology, Stanford University, Stanford, CA 94305

7 *Correspondence: ryanyork@stanford.edu; trc@stanford.edu

8

9

10 **Abstract**

11 Uncovering relationships between neural activity and behavior represents a critical
12 challenge, one that would benefit from facile tools that can capture complex structures
13 within large datasets. Here we demonstrate a generalizable strategy for capturing such
14 structures across diverse behaviors: Time-REsolved BehavioraL Embedding
15 (TREBLE). Using data from synthetic trajectories, adult and larval *Drosophila*, and mice
16 we show how TREBLE captures both continuous and discrete behavioral dynamics, can
17 uncover variation across individuals, detect the effects of optogenetic perturbation in
18 unbiased fashion, and reveal structure in pose estimation data. By applying TREBLE to
19 moving mice, and medial entorhinal cortex (MEC) recordings, we show that nearly all
20 MEC neurons encode information relevant to specific movement patterns, expanding
21 our understanding of how navigation is related to the execution of locomotion. Thus,
22 TREBLE provides a flexible framework for describing the structure of complex behaviors
23 and their relationships to neural activity.

24 **Main**

25 Accurate descriptions of animal behavior are essential to understanding brain
26 function. However, naturalistic behavior in freely moving animals is often continuous,
27 has structure across multiple timescales, and can vary broadly between individuals and
28 contexts. Thus, statistical tools that can capture and visualize the temporal structural of
29 behavior, can operate on large data sets derived from many individuals, and are
30 generalizable across systems and experiments are of central interest.

31 Advances in tracking technology have enabled measurements of animal
32 movement from a wide range of species (Datta et al. 2019; Pereira et al. 2020; Mathis et
33 al. 2020). These datasets are often extremely rich, and include correlated movements
34 across timescales, features that must be accounted for in efforts to link neural activity to
35 behavior. In parallel, a variety of sophisticated methods have emerged to parse such
36 behavioral structure and to measure behavioral changes caused by ever more powerful
37 experimental perturbations (Brown & de Bivort 2018; Datta et al. 2019; Pereira et al.
38 2020; Mathis et al. 2020). However, many of these methods are complex to implement,
39 and require extensive adaptation to specific species, contexts and experimental goals.
40 As a result, many investigators continue to rely upon either instantaneous measures of
41 specific behavioral parameters (such as velocities), or standard dimensionality reduction
42 approaches in which relatively little variance in the behavior is accounted for in the first
43 two dimensions. We therefore reasoned that an accessible framework that could be
44 easily applied to a wide variety of species and contexts, and which would allow the
45 temporal structure of behavior to be embedded and visualized in a low dimensional
46 manifold, would be of widespread utility.

47 Here we describe Time-REsolved BehavioraL Embedding (TREBLE), an easy-to-
48 use method that can describe the structure of behavior, assess the effects of
49 experimental perturbations across the entire space of behavioral measurements, and
50 provide intuitive representations of the relationships between neural activity and
51 behavior. We reasoned that individual movements, by analogy with some genomic
52 analyses, could be treated akin to conserved nucleotide sequence blocks of variable
53 length. In genomics, quantitative and qualitative differences in sequence can be
54 efficiently revealed by shotgunning long sequences into smaller, overlapping blocks that
55 preserve local structure (Staden 1979; Wang et al. 2009). These blocks are then
56 assembled into libraries that densely sample the structure, frequency and surrounding
57 sequence of each block, while smoothly reconstructing the entire sequence.

58 Quantitative variation in these features can then be assessed using standardized
59 computational methods often including dimensionality reduction followed by statistical
60 testing. Building on this conceptual parallel, we developed TREBLE as a method for
61 extracting all behavioral sequences from a dataset (analogous to shotgun sequencing),
62 assembling ‘libraries’, and creating a shared ‘behavior space’ using dimensionality
63 reduction.

64 In the TREBLE framework, behavior is first quantitatively measured (Figure 1A).
65 Relevant measurements such as centroid velocities, or changes in body or limb
66 position, are computed from video data, and segmented into highly overlapped temporal
67 windows, the size of which is constrained by the temporal structure of behavior. These
68 windows are then collected into a large library encompassing all individuals and
69 experimental perturbations and assembled in a low-dimensional behavior space.

70 Depending on window size, this space can flexibly capture a range of temporal
71 dynamics, ranging from unique trajectories to recurring patterns of movement. The
72 resulting space can be leveraged to decode complex patterns of movement in a time-
73 resolved fashion, facilitating analyses such as the rapid comparison of individuals or
74 identification of neural perturbations and stimulus effects.

75 Here, we use TREBLE to analyze a wide variety of data sources (synthetic
76 trajectories, adult and larval fruit flies, mice (with both 2D and 3D pose tracking)) and
77 applications (behavioral repertoire description, experimental perturbation, intersection of
78 behavior with imaging and electrophysiological neuron recordings). We show that
79 TREBLE's parameters can be rationally chosen using empirical analyses and that the
80 resulting behavior spaces are regularly structured and display recurrent dynamics.

81 Individual movement bouts can be easily decoded and classified, the representation of
82 which can be tailored to a user's need and allowing for detailed dissections of behavior
83 into its constituent components. TREBLE can handle data from many individuals and
84 millions of data points simultaneously, making possible the detection of otherwise
85 invisible behavioral variation. Finally, we demonstrate the ability of TREBLE to rapidly
86 uncover neurobehavioral relationships during optogenetic perturbations and neural
87 imaging in flies, as well as electrophysiological recordings in freely moving mice.

88

89 **Results**

90 ***Calibrating TREBLE***

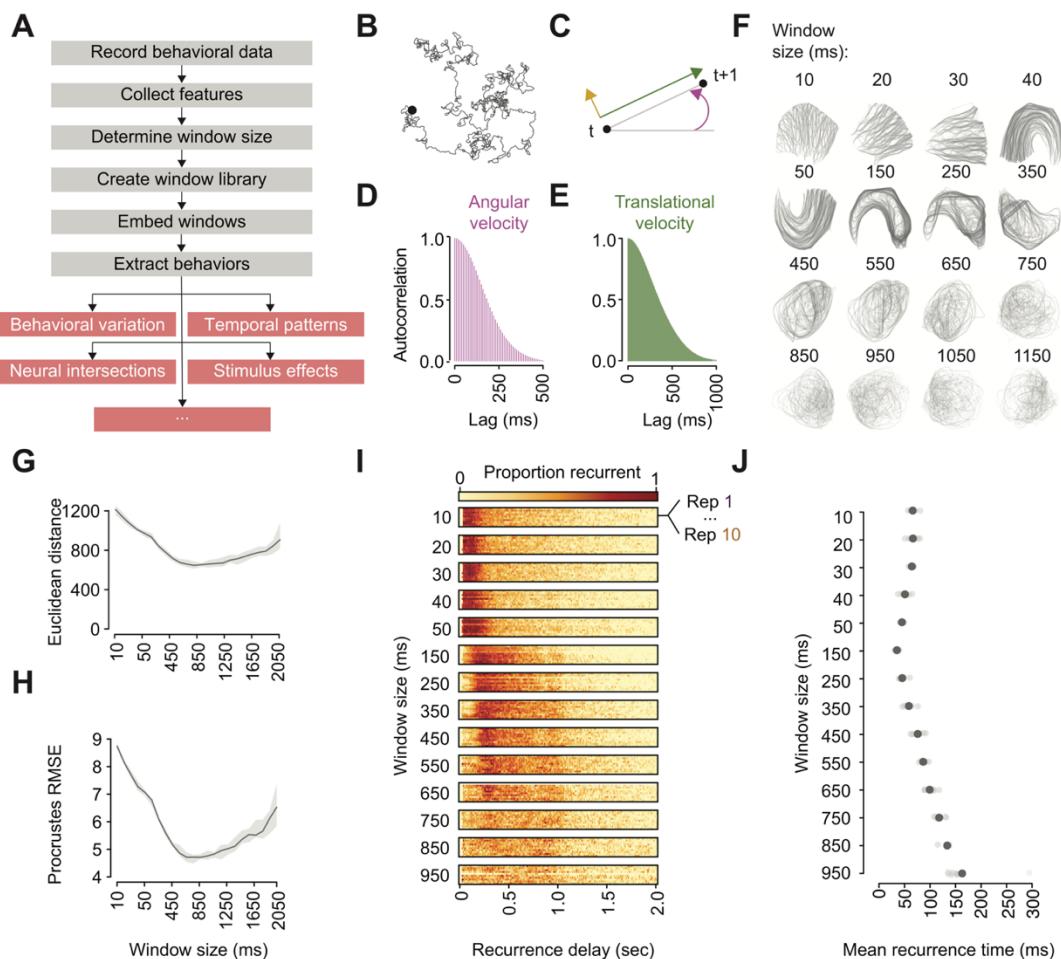
91 To calibrate this approach, we first used a large, synthetic dataset comprised of
92 correlated random walks to explore the relationships between window size and the

93 emergent structure of behavior space (Bovet 1988) (Figure 1B; Figures S1A-B).
94 Individual random walks were completely described by a combination of velocity
95 features (translational velocity, angular velocity, and side slip; Figure 1C) with
96 characteristic temporal dependencies determined by the correlated random walk
97 generator (and highlighted by their autocorrelations; Figures 1D-E). We reasoned that
98 window sizes smaller than the underlying correlations within and between velocities
99 would artificially granularize features, while windows larger than the underlying
100 correlations would combine uncorrelated features. Given this, an ‘ideal’ window size
101 should show less variance across trials and display recurrent dynamics, paths through
102 the space that repeat independently in the dataset since stereotyped movements that
103 repeat would necessarily lead to recurrent paths in such a space.

104 To explore this, we assessed the effect of window size on feature representation,
105 sweeping sizes from 10ms (containing 1 window, reflecting an instantaneous
106 measurement of features) to 2 seconds (see Methods). Windows extracted from all
107 replicate walk trajectories were then embedded into behavior space using the UMAP
108 algorithm, a computationally efficient, non-linear dimensionality reduction approach
109 (McInnes et al. 2018). The structure of behavior space varied broadly as a function of
110 window size; topologies ranged between disordered (10-30ms), recurrent (40~350ms),
111 and unique paths (450ms ~ 2 seconds) (Figure 1F; Figure S1B). These observations
112 confirmed our initial intuition that window size could have a range of effects on the
113 structure of behavior space and that more ordered spaces could be derived by
114 considering timescales longer than raw/instantaneous measurements.

115

115



134

135 **Figure 1 The TREBLE framework and its application to synthetic data**

136 (A) Outline of the TREBLE framework. Behavioral data are recorded. Relevant features (such as velocities) are extracted these
137 recordings following which empirical analyses are used to determine the optimal sampling window size. The chosen window size is
138 then used to create a library spanning all time points in the data set. The resulting window library is embedded into a low-
139 dimensional behavior space from which recurrent behaviors can be decoded and used for a number of analyses (examples labeled
140 in red boxes). (B) Example correlated random walk used for parameter tuning (see also Figure S1). (C) Velocity components that
141 can be calculated from movements in a 2-d plane. The black point at time t denotes the beginning of the trajectory which proceeds
142 to $t+1$. The purple arrow corresponds to the angular velocity of this trajectory while the green and yellow arrows represent
143 translational velocity and side slip, respectively. (D) The observed autocorrelation distribution of angular velocity computed from all
144 correlated random walks. (E) The observed autocorrelation distribution of translational velocity. (F) Example behavior spaces for a
145 range of window sizes. Spaces are plotted here by connecting temporally adjacent points (corresponding to feature windows), gray
146 lines. Darker lines reflect repeated visits along the same pathway. (G) Mean intra-point Euclidean distances as a function of window
147 size. Mean values (dark gray); Standard error of the mean (shading). (H) Procrustes distance RMSE measures as a function of
148 window size. Mean values (dark gray); Standard error of the mean (light gray shading). (I) Recurrence plot of behavior spaces
149 produced from the correlated random walk dataset. The proportion of recurrent points given a range of time delays spanning from 0
150 to 2 seconds, is indicated by the color of the corresponding bins (ranging from light yellow to dark red). Each window size contains
151 the distributions for 10 replicate walks. (J) Mean recurrence times as a function of window size. Population mean (large dark circle);
152 Individual replicates (small light circle).

153 We next sought to quantify this topological variation. We employed two statistics,
154 one targeting local structural differences and the other focused on global variation. To
155 assess local variation, we measured the average Euclidean distance between
156 temporally adjacent points, reasoning that spaces with smoother paths would display
157 smaller distances with less variance. Mean values of Euclidean distance displayed a
158 trough spanning window sizes of between 450 to 1250ms (Figure 1G). The coefficient of
159 variation (CV) highlighted a broad overlapping region of similarity, spanning window
160 sizes of between 10 to 950ms (Figure S1C). To assess global variation, we measured
161 Procrustes distance (McInnes et al. 2018; Dryden & Mardia 1998). This metric
162 compares the difference between configurations of points, used here to quantify the
163 distance between replicate behavior spaces of a given window size. This measure
164 displayed a pattern similar to Euclidean distance, with a large trough between 450 and
165 1250ms, and a consistent CV up to window sizes of 950ms (Figure 1H; Figure S1D).
166 Thus, these metrics allow quantification of the effects of window size, in this case
167 revealing stability in topologies of behavior space across a range of windows.

168 We then explored how the choice of window size affects the temporal
169 progression of trajectories through behavior space by measuring the frequency with
170 which portions of a trajectory were repeated. We delineated a neighborhood around
171 each point in behavior space (see Methods) and determined the amount of time
172 between visits to each neighborhood. This was then represented as the proportion of all
173 neighborhoods that were revisited (Bruno et al. 2017). The distribution of these values
174 was consistent across replicates *within* a given window size but varied across sizes
175 (Figure 1I). Window sizes of 150 to 650ms displayed a dominant return time of

176 approximately 250ms, with a secondary peak at 1 second (Figure 1I). Windows larger
177 than 650ms displayed increasingly noisy return time distributions (Figure 1I). Reflecting
178 this, plots of mean recurrence times revealed a structured tuning curve in which the
179 mean and variance were minimized at a window size of 150ms (Figure 1J). Combining
180 these observations, the broad tuning curves in both the spatial and temporal properties
181 of behavior space indicate that while window size can have appreciable effects, this
182 representation of behavior can be robust across multiple parameter values.

183

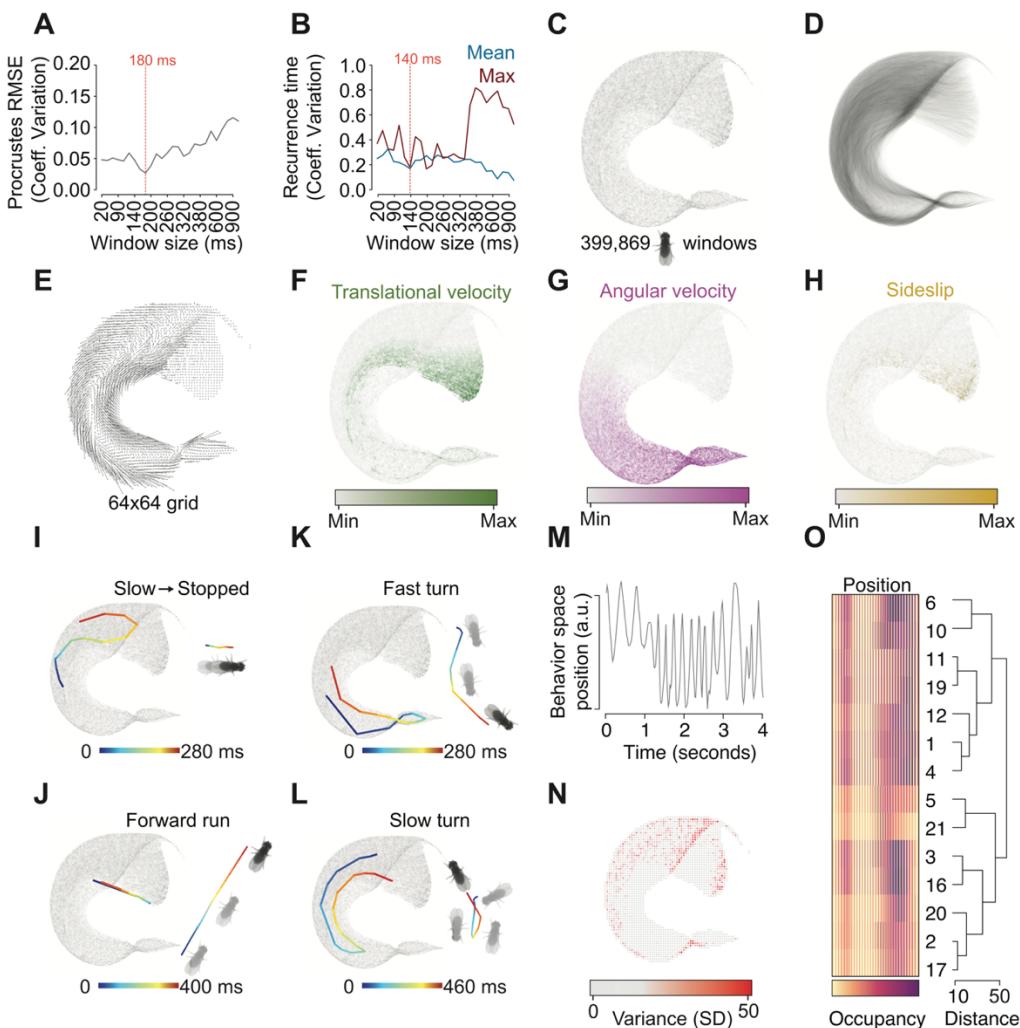
184 ***Analysis of fruit fly locomotion***

185 How does TREBLE perform when applied to biological data? First, we applied
186 the TREBLE to fruit fly locomotor behavior collected from individual animals walking on
187 an air-cushioned ball while exploring a virtual world (Haberkern et al. 2019). We
188 extracted rotational, translational and slip velocity components from these data and
189 applied the empirical window analysis outlined above. Structural variation analysis of
190 behavior space revealed a minima at a window size of 180ms (Figure 2A; Figures S2A-
191 C) while mean recurrence time was minimized at 140ms (Figure 2B; Figures S2D-F).
192 Given that these metrics displayed only a small amount of variation between these two
193 window sizes, we focused further analyses on a behavior space constructed using a
194 window size of 160ms.

195 We constructed a behavior space from the movements of 14 individual walking
196 fruit flies (Figure 2C; 399,869 windows). Individual points in this space corresponded to
197 unique windows with specific combinations of velocity features (Figure 2C).

198

199



228 **Figure 2** Analyzing fruit fly locomotion with TREBLE

229 **(A)** The coefficient of variation of Procrustes distance as a function of window size. The observed minimum at 180ms is denoted by
 230 the red dotted line. **(B)** Coefficient of variation for the mean (blue line) and maximum (i.e. highest proportion of time bins displaying
 231 recurrence; dark red line) recurrence times as a function of window size. The optimal tradeoff between the mean and maximum
 232 times is denoted at 140ms with a red dotted line. **(C)** Fly locomotor behavior space. Each point corresponds to a feature window
 233 (399,869 windows in total) as extracted from the 14 individual flies. **(D)** Pathways through the locomotor behavior space, produced
 234 by connecting temporally adjacent windows with partially transparent lines (as in Figure 1F). **(E)** Walking fruit fly behavior space
 235 represented as a vector field. Arrow direction and magnitude correspond to the angle and mean direction taken after visiting each
 236 bin. **(F)** The distribution of translational velocity across behavior space. Darker green corresponds to larger values of translational
 237 velocity. **(G)** The distribution of angular velocity across behavior space. Darker purple corresponds to larger values. **(H)** The
 238 distribution of side slip across behavior space. Darker gold corresponds to larger values. **(I)** Example bout of stopping. The pathway
 239 through behavior space is represented on the left. XY coordinates of the actual bout are plotted on the right. Both representations
 240 are colored by elapsed time. **(J-L)** Same as **(I)** but for fast turns, forward runs, and slow turns, respectively. **(M)** Sample trajectory
 241 through behavior space as represented by a 1-dimensional coordinate value. **(N)** Distribution of per-bin standard deviation across
 242 locomotor behavior space. Color corresponds to the variance of bin-wise visitation frequency across all 14 flies. **(O)** Hierarchical
 243 clustering of individual fly density maps. The dendrogram on the right represents the relationships between all 14 flies (numbered at
 244 the tips of each branch). Individual fly density maps are presented on the left in 1-dimension by converting the density map matrix
 245 into a single vector (presented in 2-dimensions in Figure S2J). Darker colors reflect more time spent ('occupancy') in a specific
 246 region.

247 Connecting these points based on their temporal order revealed stereotyped paths in
248 which individual flies repeated the same patterns of movement (Figure 2D; Figure S2G).
249 As a result, the mean vector field produced by these was highly structured (Figure 2E).
250 These pathways traversed regions of space defined by the input velocity features,
251 meaning that position within the space could be used to infer the underlying pattern of
252 movement (Figures 2F-H). Analyzing continuous movement through these regions
253 demonstrated that individual behavioral bouts and sequences could be identified
254 (Figures 2I-L; manually chosen). Finally, consistent with the notion that these
255 sequences were recurrent (Figure 2B), plotting the positions of individual fly trajectories
256 in behavior space over time reveals periodicity (Figure 2M, and data not shown). Taken
257 together, these observations show how TREBLE can be used to identify repeated,
258 intuitive, and interpretable patterns of behavior over time.

259 The capacity of TREBLE to co-embed many trials or individuals in the same
260 space facilitates direct measurements of trial-to-trial or individual-to-individual variation.
261 For example, we found that each of the 14 flies moved in grossly similar ways, as
262 indicated by the fact that 83% of the behavior space was explored by all individuals,
263 while less than 1% of the behavior spaces was explored by only one individual (Figures
264 S2H-I). At the same time, individual trajectories varied greatly in how often they
265 traversed different parts of the space. To measure this, we computed a probability
266 density map for each individual (Figures S2J), and identified small, specific regions of
267 behavior space that had the greatest variance between individuals (Figure 2N; Figures
268 S2H-I). Hierarchical clustering of 1-d vectorized versions of these (Methods) maps
269 revealed a variety of behavioral profiles across individuals (Figure 2O). Thus, in this

270 case, behavior space was composed of movement types common to all individuals and
271 could therefore act as template to compare the unique statistics of each.

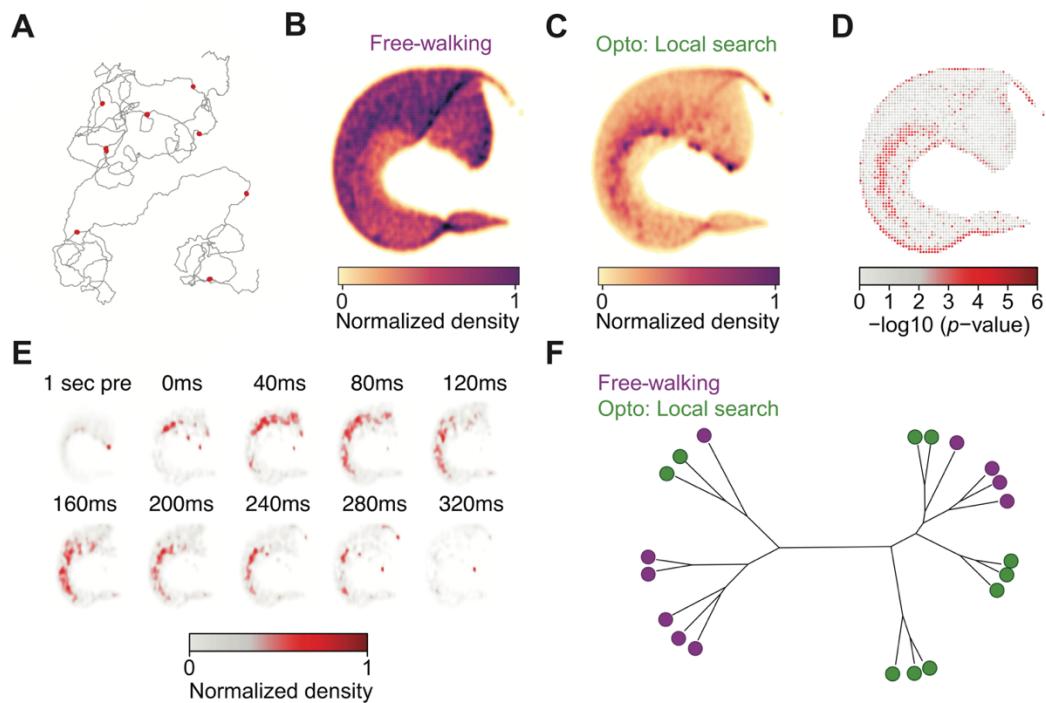
272

273 ***Detecting behavioral changes due to optogenetic manipulation***

274 To test the ability of TREBLE to detect behavioral changes arising from an
275 experimental manipulation, we analyzed locomotor behavior during optogenetic
276 activation of gustatory sensory neurons expressing Gr5a, a sugar receptor (Haberkern
277 et al. 2019) (Figure 3A; see Methods). To induce local search behavior, a 200 ms
278 optogenetic stimulation was delivered every time a fly reached a pre-specified area and
279 was repeated for every return visit. We processed nineteen optogenetic trials
280 (1,110,025 windows) and co-embedded these windows with those from the control flies
281 (described above) in the same behavior space. The movements contained within
282 optogenetically stimulated trials overlapped extensively with those in control flies, with
283 >97% of the space occupied by both datasets (Figures S2K-L). However, on average,
284 the frequencies with which specific patterns emerged in the two datasets diverged
285 dramatically (Figures 3B-C) and were unevenly distributed across behavior space (per
286 bin Kruskal-Wallis test, see Methods; Figure 3D). These observations suggest that,
287 while the structure of locomotion is conserved overall, these two groups display quite
288 different temporal patterns of behavior.

289 To further explore this difference, we leveraged the continuous, time-resolved
290 nature of TREBLE to map the evolution of average responses to neuronal activation
291 over time (Figure 3E). Compared to pre-stimulus epochs, overall behavioral responses
292 displayed an initial period of slowing (spanning from 0 to 120ms after stimulus offset),

293 followed by increased turning (120ms to 280ms) (Figure 3E). These dynamics are
294 consistent with known local search behaviors (Haberkern et al. 2019; Corfas et al.
295 2019). Despite these population-level patterns, we noted a surprising amount of
296 behavioral heterogeneity amongst optogenetically activated flies, such that the behavior
297 of some optogenetically activated flies overlapped with that of controls (Figure 3F;
298 Figures S2J-M). Thus, TREBLE can assess the effects of a behavioral perturbation to
299 detect both population and individual level variation.



320 **Figure 3 Identifying behavioral perturbations via optogenetic manipulation**
321 (A) Example trajectory for a single optogenetic trial. The fly's path through the virtual world is represented by the dark grey line. Red
322 dots correspond to locations in which optogenetic stimulation was presented. (B) Average density map computed from the 14 free-
323 walking control flies (darker color corresponds to more time spent in a specific region). (C) Average density map computed from the
324 19 optogenetically activated flies. (D) Bin-wise differences between control and optogenetically activated flies. Color corresponds to
325 the significance ($-\log_{10}$ transformation of the p -value; Kruskal-Wallis test) of the differences in density between the two groups.
326 Darker red corresponds to increasingly significant differences. (E) Time evolving responses to optogenetic stimulation in behavior
327 space. Each space reflects a specific time window before (first space) and after stimulation (each subsequent space). Color
328 corresponds to the average response to stimulation across all 19 flies (represented as normalized density; see methods). Windows
329 span 1 second of time, beginning at the value represented above each space. (F) The behavioral relationships between all free-
330 walking control and optogenetically activated flies. The tree represents the results of hierarchical clustering on the density maps of
331 each individual fly (as in Figure 2O). Each branch tip is associated with an individual fly, group is represented by color (Controls =
332 purple; Optogenetic trials = dark green).

333 ***Identifying structure in more complex feature sets: Locomotion of Drosophila***

334 ***larvae***

335 Thus far we have used to TREBLE to analyze behaviors as changes in centroid
336 velocity components. However, pose estimation methods have made analyses of other
337 behavioral features, such as posture or limb movement, increasingly common (Pereira
338 et al. 2020; Mathis et al. 2020). As such, pose estimation methods typically represent
339 behavior in multi-dimensional spaces that may or may not include explicit velocities.
340 With this in mind, we next assessed the ability of the TREBLE framework to capture a
341 high-dimensional combination of postural and velocity features describing larval
342 *Drosophila* locomotion.

343 Locomotion in *Drosophila* larvae is characterized by stereotyped changes in size
344 and posture correlated with peristaltic movements and bending that produce forward
345 and backward translation and turning (Clark et al. 2015). To capture this, we tracked
346 *Drosophila* larvae via machine vision and at each time point calculated 11 complex
347 postural and velocity features following established methods (Figure 4A; n = 72) (Risse
348 et al. 2014, 2017). Given that certain of these features may be correlated, we performed
349 principal component analysis and found that 8 components were sufficient to explain
350 over 90% of the observed variance (Figure S3D). These 8 PCs were then used to run
351 the iterative window procedure, resulting in an optimal window size of 800ms (Figures
352 S3A-C). This produced a 2D behavior space that revealed a strongly oscillatory region
353 (Figure 4B) with highly directional movement (Figure 4C). Analyzing the distribution of
354 the input features revealed that this oscillator corresponded strongly with features of
355 peristaltic locomotion (Figures 4D-G). Specifically, the oscillator appeared to switch

356 between a regime of lower velocity paired with an increase in the area of the animal's
357 shape (i.e. scrunching) and higher velocity with lower overall area (stretching) (Figures
358 4D-G). When oscillating, these regimes produced an individual 'run' of crawling (Clark et
359 al. 2018).

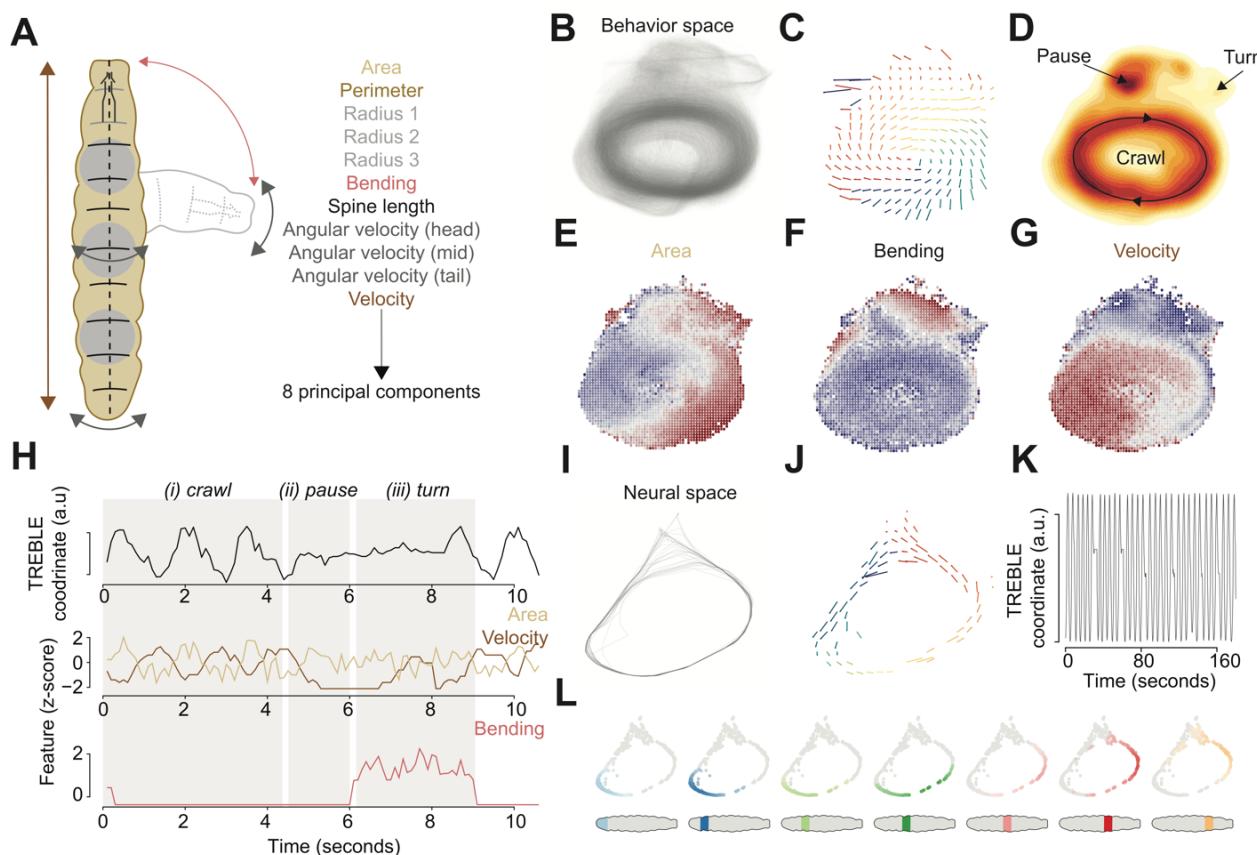
360

361 The upper portion of the space, outside the oscillating domain, was split between
362 a region of increased bending and another associated with pausing (Figure 4D).

363 Comparing the distribution of features over time to movement in the TREBLE space
364 highlighted the structured relationships between these features (Figure 4H). During
365 crawling (denoted by 'i' in Figure 4H) velocity and area oscillate inversely. When a turn
366 is initiated, the animal first pauses (ii), decreasing velocity and stabilizing area, and then
367 begins bending with a corresponding shift in position in behavior space (iii). After the
368 turn is complete, oscillatory crawling begins again (Figure 4H). Overall, oscillatory
369 waves occur with a period of around 1 second (Figure S3E), matching previous
370 observations of *Drosophila* larval crawling (Heckscher et al. 2012). TREBLE therefore
371 captured the common structural and temporal elements of the *Drosophila* larval
372 ethogram (Clark et al. 2016) within a single behavior space.

373 The recurrent properties of crawling mirror waves of motor neuron activity
374 (Tastekin et al. 2018; Clark et al. 2018). We therefore wondered if using TREBLE to
375 analyze the output of motor neurons might also yield a corresponding oscillator
376 describing neural dynamics. We used calcium imaging data from fictively crawling
377 larvae (Tastekin et al. 2018) to construct a 'neural space' from the activity of seven
378 motor neuron classes (Figure 4I; Figures S3F-K).

379



380 **Figure 4 Larval *Drosophila* crawling dynamics**

381 (A) Cartoon of larval *Drosophila* movement and accompanying features used for analysis. (B) Pathways through the larval
 382 locomotor behavior space ($n = 72$ larvae). (C) The mean vector field of larval locomotor space. Direction and size of arrow
 383 correspond to the mean movement through a given bin in space. Color denotes the angle of the arrow. (D) Probability
 384 density function of larval locomotor space plotted as a heatmap. Behaviors annotated qualitatively. (E-G) The distribution of area (E),
 385 bending (F), and velocity (G) as a function of larval behavior space (z-scores). Blue corresponds to negative values, red represents
 386 positive values. (H) Example distribution of behavior space position (TREBLE coordinate; a.u.), area (z-score; yellow), velocity (z-
 387 score; brown), and bending (z-score; red) over a ~10 second epoch of behavior. The approximate durations of behavioral states
 388 are annotated via grey shading and labeled above (i: crawl, ii: pause, iii: turn). (I) Pathways through larval neural space (plotted
 389 here is an example distribution of a single trial; full space in Figure S3I). (J) Larval neural space plotted as a mean vector field. (K)
 390 Example of oscillating position in neural space over a 160 second period of time. (L) Probability density functions of motor neuron
 391 activity across behavior space. The plots proceed from posterior (leftmost) to anterior (rightmost) as denoted by the cartoon larvae
 392 plotted below (with approximate location of motor neuron segment colored).

393

394 As with the behavior space, we found that motor neuron activity yielded an oscillator
 395 (Figure 4I) with directional (Figure 4J) and stereotyped movement (Figure 4K). Bouts of
 396 crawling are associated with waves of neural activity, originating in posterior neurons
 397 and propagating in the anterior direction (Tastekin et al. 2018; Clark et al. 2018).

398 Plotting the peak activity of each motor neuron type in the neural space recapitulated
399 this observation (Figure 4L). Peak activity of the posterior neurons occurred in the lower
400 left-hand portion of the space, followed by spatially sequential peaks of the more
401 anterior motor neurons (Figure 4L) moving in the same direction described the mean
402 vector field of the space (Figure 4J). These observations reveal that both the behavioral
403 output and neural dynamics underlying *Drosophila* larval locomotion can be captured in
404 a common oscillatory framework. In addition, these analyses show that TREBLE can
405 applied to find structure in higher-dimensional neural and behavior data.

406

407 ***Identifying structure in more complex feature sets: Mouse pose dynamics***

408 We next sought to generalize TREBLE to a much more complex form of
409 behavioral data, mouse movements in three dimensions (from Markowitz et al. 2018).
410 To do so, we analyzed mouse behavior measured with 3D imaging via the MoSeq
411 pipeline (Wiltschko et al. 2015; Markowitz et al. 2018). In this pipeline, freely moving
412 mice were imaged in an arena using three orthogonal cameras at 30Hz (Wiltschko et al.
413 2015; Markowitz et al. 2018). These video streams were then processed to produce 17
414 behavioral features (Figure 5A) representing mouse movement in three dimensions
415 (Wiltschko et al. 2015; Markowitz et al. 2018) which we then used as input to the
416 iterative window size procedure. A window size of 130ms was chosen (Figure S4A-C),
417 yielding a behavior space that organized aspects of posture and movement into a
418 recognizable, and recurrent, structure (Figures 5B-J). Analyzing the distribution of input
419 features in behavior space (Figures 5E-J) allowed us to identify portions of the space
420 related to characteristic behaviors such as walking, scrunching, rearing, and pausing

421 (Figures 5C-D). These designations were corroborated by calculating the average 3D
422 pose of animals as a function of position in behavior space (Figure S4D). These
423 observations indicate that TREBLE can be used to identify structure in high-dimensional
424 pose estimation datasets.

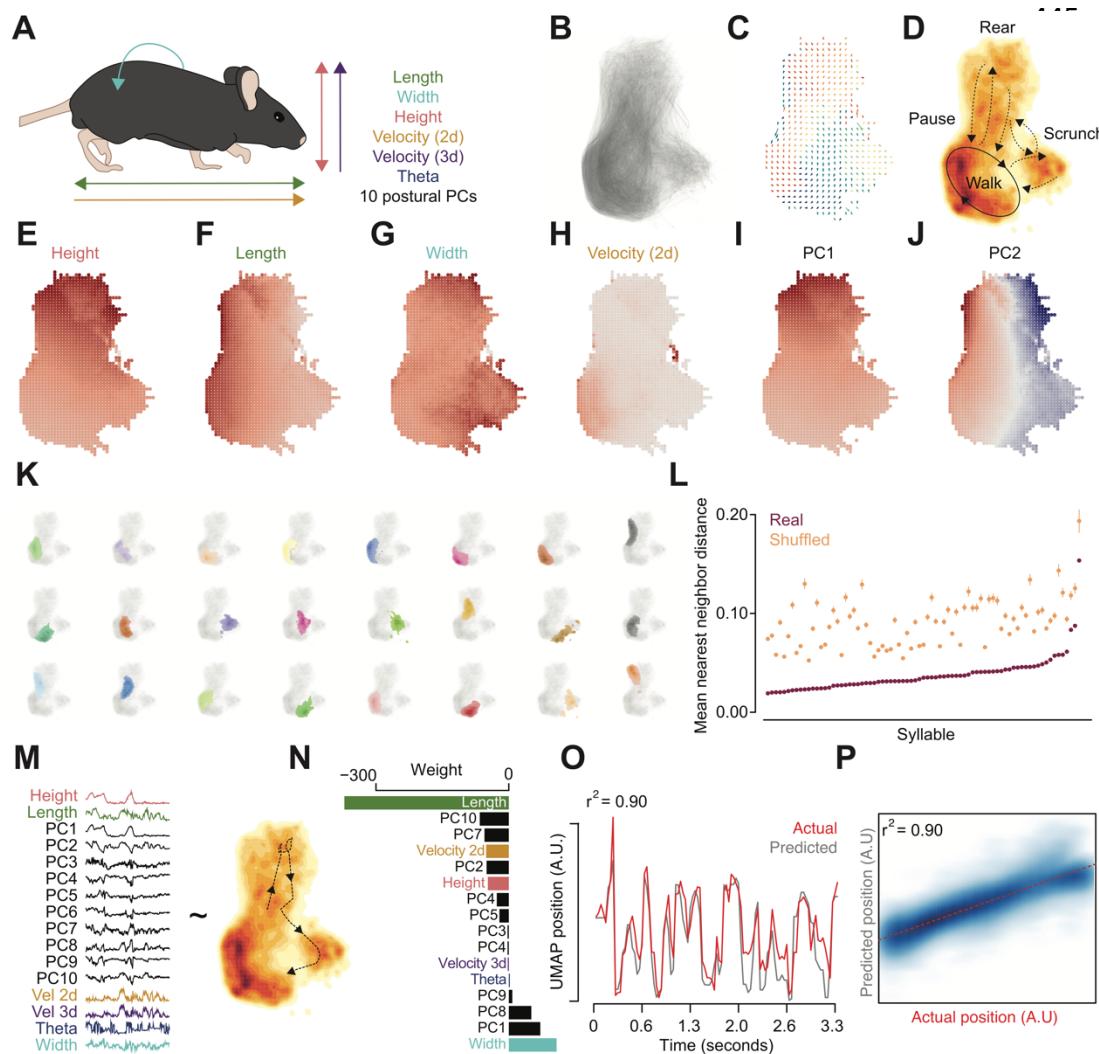
425 In addition to measurements of 3D movement, the MoSeq pipeline also identifies
426 how discrete elements of mouse behavior - behavioral ‘syllables’ – are sequenced over
427 time (Wiltschko et al. 2015; Markowitz et al. 2018). What is the relationship between the
428 continuous representation of 3D behavior provided by TREBLE and the discrete output
429 of the MoSeq pipeline (Markowitz et al. 2018)? Annotating the TREBLE behavior space
430 with the individual syllables identified by MoSeq revealed that the distributions of
431 individual syllables appeared highly compact (Figure 5K; Figure S4E). Syllables
432 produced by MoSeq corresponded to distinct and identifiable behavioral states identified
433 by TREBLE, such as walking (Figure 5K; yellow and blue syllables; first row, fourth and
434 fifth columns) and scrunching (Figure 5K; purple syllable; second row, third column).
435 Furthermore, these distributions were non-random. By comparing the mean nearest
436 neighbor distance for each syllable to a shuffled distribution, we found that the mean
437 distance between points was significantly lower than expected by chance ($p < 0.0001$;
438 permutation test) for all syllables (see Methods). Therefore, TREBLE and MoSeq
439 capture complementary aspects of behavior - both discrete and continuous – albeit
440 using very broadly different statistical frameworks.

441

442

443

444



475 **Figure 5 Analyzing 3D pose dynamics in mice**
476 (A) Cartoon of mouse 3D movement and accompanying features used for analysis. (B) Pathways through mouse 3D pose space (n
477 = 8 mice). (C) The mean vector field of mouse 3D pose space. Direction and size of arrow correspond to the mean movement
478 through a given bin in space. Color denotes the angle of the arrow. (D) Probability density function of mouse 3D pose space plotted
479 as a heatmap. Behaviors annotated qualitatively. (E-J) The distribution of height (E), length (F), width (G), 2D velocity (H), postural
480 PC1 (I), and postural PC2 (J) as a function of larval behavior space. (E-H) Color ranges from grey (minimum value) to red
481 (maximum). (I-J) Blue corresponds to negative values, red represents positive values. (K) The distributions of the 24 most common
482 behavioral syllables (as identified by MoSeq) in behavior space. A probability density function across behavior space was computed
483 for each syllable and then plotted in color on top of the full behavior space (in grey; see Methods). (L) The distribution of mean
484 nearest-neighbors distance between points in behavior space for all syllables (n = 43). Purple denotes the observed values. Orange
485 corresponds to the mean and distribution (lines; lower (25th percentile) and upper (75th percentile) hinges of a boxplot) of shuffled
486 data (10,000 permutations). (M) Visualization of the variables used to construct the regularized generalized linear model. Features
487 are represented by example time series (left) and were compared to movement through behavior space (right). (N) Barplot of
488 coefficient weights from the final model, sorted by weight and colored to match the example time series in (M). (O) Comparison of
489 the actual position in behavior space (grey) to the prediction from the final GLM (red) for an example ~3 second time period. (P)
490 Smoothed scatterplot comparing observed and predicted behavior space positions for the full dataset. Darker blue denotes greater
491 density of points. Dashed red line corresponds to the fit of a regression between observed and predicted values.

492 Finally, we wondered whether TREBLE was truly capturing the majority of
493 behavioral variation present in this complex dataset. To address this, we created a
494 regularized generalized linear model (GLM) comparing the relationship of the input
495 features to the position in TREBLE behavior space over time (Figure 5M; see Methods).
496 We found that the GLM captured over 90% of the variance in the data, and pose
497 features such as length and width, and related postural principal components,
498 contributed strongly to the model (Figures 5N-P). Position in the TREBLE space could
499 be predicted with a substantial degree of accuracy from the feature set (Figures 5O-P),
500 predictability that was consistent across individual trials (Figure S4F). We therefore
501 conclude that TREBLE is able to explain a substantial portion of behavioral variation
502 even in complex feature sets and can robustly represent the 3D dynamics of animal
503 movement in a low-dimensional, continuous framework.

504

505 ***Using TREBLE to characterize neural encoding of behavior***

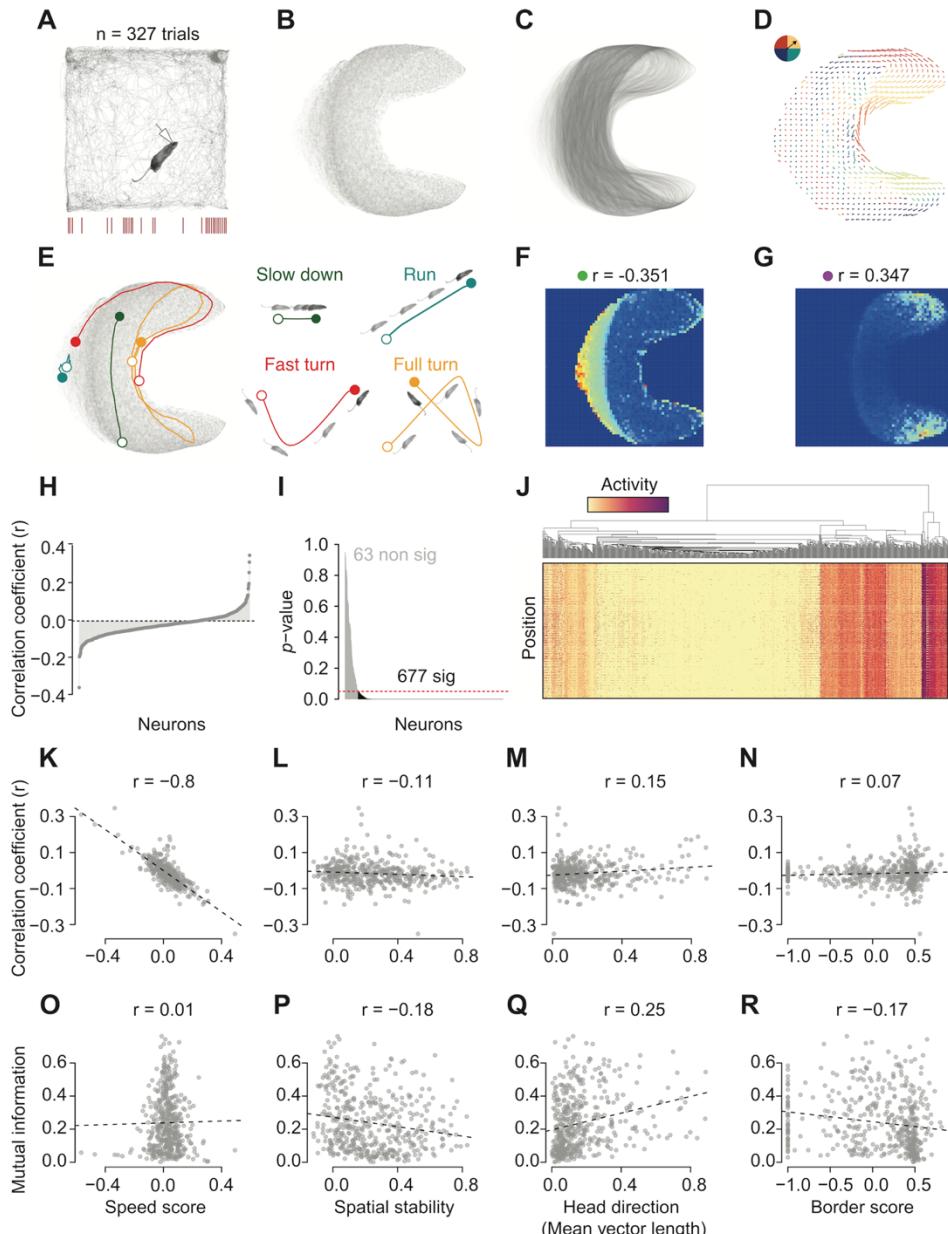
506 A common, and often difficult, goal in neuroscience is to relate neural activity to
507 behavior. We therefore used TREBLE to identify behavioral coding in 794 medial
508 entorhinal cortical (MEC) neurons recorded during the foraging trials presented above
509 (n = 14 mice, 327 trials) (Hardcastle et al. 2017) (Figure 6A). The MEC is hypothesized
510 to support navigation and contains a population of functionally defined neurons that
511 encode a variety of behavioral variables such as an animal's position in external space
512 (e.g. grid and border cells), head direction, and running speed (Kropff et al. 2015;
513 Sargolini et al. 2006; Solstad et al. 2008; Hafting et al. 2005).

514 We extracted time-varying velocity measures from positional data of individual
515 mice as they foraged for randomly scattered food rewards in a 1m by 1m box (Figure
516 6A). We then collected windows from these data using the iterative selection procedure
517 to choose a window size of 400ms (Figure S5A-F). As we previously observed with the
518 correlated random walk and *Drosophila* spaces, the resulting behavior space ordered
519 points by velocity features (Figures S5G-L) and connected them with continuous and
520 directionally recurrent pathways (Figures 6B-D). Moreover, distinct behavioral bouts –
521 such as running, turning, and stopping – could be easily decoded from the space and
522 mapped back onto real XY coordinates (Figure 6E).

523 We related MEC neuron activity to movement in the behavior space defined by
524 TREBLE. To do so, we first created a 2-dimensional behavioral tuning curve for each
525 neuron. Each cell's average activity was mapped as a function of the animal's position
526 in behavior space, revealing a wide variety of patterns (Figure S5M). Correlation
527 coefficients were then calculated by comparing each neuron's activity with a 1-
528 dimensional representation of behavior space position (64 x 64 grid: Methods). Strongly
529 positive or negative correlations thus reflect that a given's neuron activity is increased in
530 different portions of behavior space (Figure 6F-H), of which a number were identified.
531 Consistent with the MEC's role in navigation, permutation tests revealed that these
532 correlations were overwhelmingly non-random, with only 63 of the 794 MEC neurons
533 displaying non-significant relationships with locomotor behavior after multiple test
534 correction (Figure 6I). Furthermore, clustering neural activity patterns across behavior
535 space revealed multiple distinct types of relationships between MEC neurons and
536 locomotor behavior (Figure 6J; Figure S5M). Most neurons were active in relatively

537 small and specific regions of behavior space, suggesting that they encode information
538 relevant to particular locomotor movements (Figure S5M). Other neurons displayed
539 distributed activity across behavior space, consistent with less selective coding of
540 locomotor information (Figure S5M). Thus, these various patterns can be considered 2-
541 d tuning curves, representing the relationship between neural activity and behavior for
542 each neuron.

543 To what extent do these patterns reflect previously described representations of
544 navigational coding in the MEC? To address this, we intersected TREBLE correlations
545 with other commonly calculated MEC coding variables (speed, head direction, spatial
546 stability, border proximity; see Methods). Unsurprisingly, we found that behavior space
547 correlated with speed coding (Figure 6K) compared to the other measures (Figures 6L-
548 N). However, a more precise measure of coding capacity requires a comparison
549 between the amount of information jointly shared between position in behavior space,
550 speed score, and neural activity. To do this, we computed the mutual information (MI)
551 between TREBLE position and neural activity for each neuron and then related these
552 measures to the commonly calculated MEC coding variable (Figures 6O-R). Higher MI
553 values here indicate that spatial coordinates in behavior space contain more information
554 about a given neuron's activity, reflecting stronger behavioral coding. Strikingly, MI was
555 uncorrelated with speed score (Figure 6O), indicating that encoding of locomotor
556 behavior was not solely related to variation in speed. This raised the possibility that
557 these neurons may encode information about other locomotor variables including longer
558 timescale and more complex maneuvers.



592

Figure 6 Mapping mouse navigational coding

593 **(A)** Example trajectory for a foraging mouse with paired spike recordings via an implanted electrode (drawn). **(B)** Mouse
 594 locomotor behavior space, each point corresponds to a temporal window. **(C)** Pathways through mouse locomotor behavior space,
 595 produced by connecting temporally adjacent windows with partially transparent lines. **(D)** Mouse locomotor behavior space
 596 represented as a vector field. Arrow direction and magnitude correspond to the angle and mean direction taken after visiting each
 597 bin. Arrows are colored by the degree of the direction vector (corresponding to circle in upper left-hand corner). **(E)** Example
 598 behavior bouts decoded from behavior space. Pathways through behavior space are plotted on the left. Bout starts are indicated by
 599 the open circle, ends by the closed circle. Positions in real XY space are plotted on the right. **(F)** Example 2-d tuning curve of a
 600 neuron that is negatively correlated with behavior space (activity measured as spikes/second; labeled with a green dot for reference
 601 in later figures). Colors range from blue to red, corresponding to minimum activity (blue) to maximum activity (red). Pearson
 602 correlation (r) is denoted above the plot. **(G)** Example 2-d tuning curve of a neuron that is positively correlated with behavior space
 603 (activity measured as spikes/second; labeled with a purple dot for reference in later figures). Colors range from blue to red,
 604 corresponding to minimum activity (blue) to maximum activity (red). Pearson correlation (r) is denoted above the plot. **(H)**

605 Distribution of correlation coefficients between neural activity and behavior space position for all MEC neurons (activity measured as
606 spikes/second). **(I)** Distribution of p -values (Bonferroni corrected) resulting from permutation tests (10,000 shuffles) of the
607 correlations plotted in **(h)**. The red dotted line indicates $p = 0.05$. **(J)** Clustered 2-d tuning curves for all neurons. The heatmap
608 corresponds to a linearized version of the 2-d tuning curve, colored by average activity as a function of position in behavior space.
609 **(K)** Scatterplot of the relationship between the correlation coefficient of the relationship between behavior space and neural activity
610 (y-axis) and speed score (x-axis). Pearson correlation (r) is denoted above the plot. **(L)** Scatterplot of the relationship between the
611 correlation coefficient and spatial stability. **(M)** Scatterplot of the relationship between the correlation coefficient and head direction
612 (HD) mean vector length. **(N)** Scatterplot of the relationship between the correlation coefficient and border score. **(O)** Scatterplot of
613 the relationship between the mutual information (MI) of behavior space and neural activity (y-axis) and speed score (x-axis). **(P)**
614 Scatterplot of the relationship between MI and spatial stability. **(Q)** Scatterplot of the relationship between MI and head direction
615 (HD) mean vector length. **(R)** Scatterplot of the relationship between MI and border score.

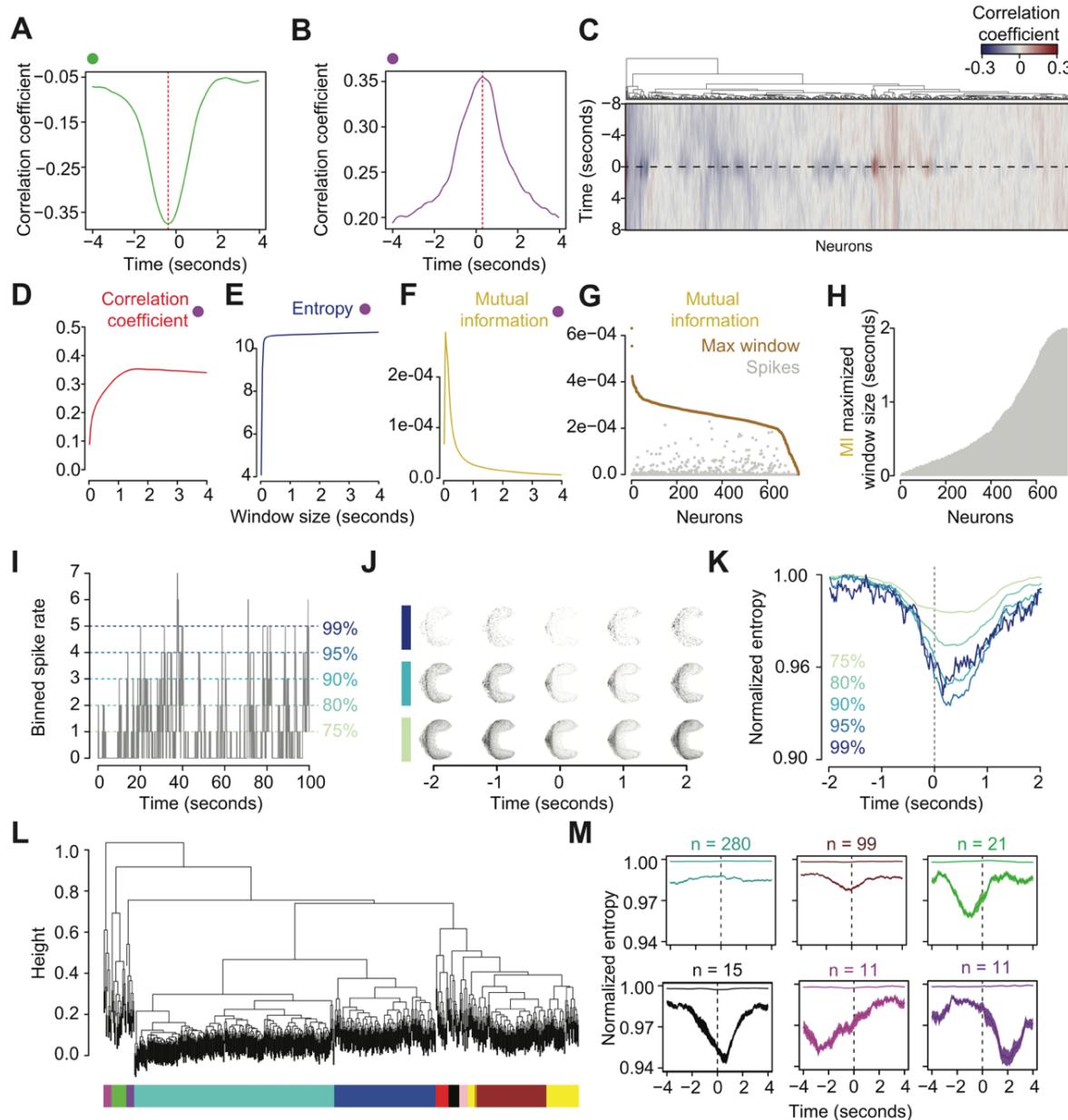
616

617 ***Using TREBLE to identify temporal variation in the neural coding of behavior***

618 Neural activity can relate to continuous behaviors on a variety of timescales, the
619 discovery of which is an area of active interest (Datta et al. 2019; Krakauer et al. 2017;
620 Glaser & Kording 2016). We therefore wanted to test the ability of TREBLE to provide a
621 path for relating neural activity with behavior and took advantage of the apparent rich
622 capacity of MEC neurons to encode locomotor behavior over a variety of timescales.

623 To do this, we first computed the cross-correlation of TREBLE position and
624 activity for each neuron (Figures 7A-C). This analysis identifies when a neuron's activity
625 correlates with position in behavior space by calculating correlations between the two
626 variables using a sliding window centered around a temporal offset of zero (i.e.
627 simultaneity). We found a variety of temporal offsets in the cross-correlation
628 distributions for each neuron. Peak correlations could both precede (Figure 7A) and
629 follow (Figure 7B) changes in behavior. Clustering the distribution of these cross-
630 correlation coefficients revealed a variety of neural classes, including negative, positive,
631 and even mixed negative and positive cross-correlation profiles (Figure 7C; Figure
632 S6A). Notably, cross-correlation peaks could be of varying widths; some were tightly
633 centered (e.g. Figures 7A-B) while others displayed correlations lasting seconds (Figure

634 S6A). Therefore, in addition to their instantaneous relationships, the activity of MEC
 635 neurons can covary with behavior over a range of temporal scales and offsets.



636 **Figure 7 Temporal variation in mouse navigational coding**
 637 **(A)** Cross-correlation distribution of a neuron that is negatively correlated with behavior space (same as in **6F**). **(B)** Cross-correlation
 638 distribution of a neuron that is positively correlated with behavior space (same as in **6G**). **(C)** Hierarchical clustering of cross-
 639 correlation distributions across all neurons. Color corresponds to the correlation coefficient, blue indicating negative and red
 640 indicating positive correlations. The y-axis represents time before and after instantaneity, ranging from 8 seconds before to 8
 641 seconds after. The x-axis corresponds to the relationships of all MEC cells based on their cross-correlations with behavior (which
 642 are shown via the colored heatmap). In essence, each column contains the same information as **7A-B** but here represents the
 643 distributions using variation in color over time, rather than a 1-d tuning curve. **(D)** Correlation coefficient distribution as a function of
 644 varied window sizes for the neuron plotted in **6G** and **7B**. **(E-F)** same as **(D)** but for entropy **(E)** and mutual information **(MI)**. **(G)**
 645 Distribution of MI for all neurons, comparing the maximally informative window size (in gold) and that measured from spikes (grey).

646 (H) Barplot of the distribution of MI maximized window sizes for all neurons. (I) Example neural activity trace (using MI maximized
647 window size) and corresponding percentile cutoffs (at 75%, 80%, 90%, 95%, and 99%). (J) Temporal distributions in behavior space
648 (1 second windows) surrounding peaks in activity (75%, 90%, and 99% cutoffs). (K) Entropy distributions of surrounding activity
649 bouts corresponding to each percentile (labeled and colored on plot). Normalized entropy measures are shown to be able to
650 compare across percentiles with different sample sizes. (L) Hierarchical clustering of entropy distributions across all neurons.
651 Colored bars indicate clusters as measured using dynamic tree trimming (see Methods). (M) Example mean entropy distributions for
652 four clusters identified using dynamic tree trimming. Plotted are the distributions for the 75% (top) and 99% (bottom) cutoffs. Mean
653 and standard error are plotted as well as sample sizes (above).

654

655 Given this observation we hypothesized that individual MEC neurons may most
656 reliably encode behavior over specific timescales. To test this, we calculated each
657 neuron's activity over a range of temporal windows and measured their associations
658 with behavior space via correlation, mutual information (MI), and entropy (i.e. how
659 narrowly neural activity was distributed across behavior space) (Figures 7D-F). For
660 most neurons, the strength of the correlation between activity and behavior as well as
661 entropy plateaued (Figures 7D, 7F) while MI displayed a clear, and at times very
662 specific, maximum value as a function of temporal window size (Figure 7F; Figure S6B).
663 This pattern suggested that, given temporal binning that maximized MI, neural activity
664 could be more specifically associated with distinct regions in behavior space. This was a
665 general pattern: across all neurons the MI maximized window size encoded more
666 information than instantaneous spikes (Figure 7F). The distribution of these MI
667 maximized window sizes varied broadly, ranging from instantaneous spiking to window
668 sizes up to 2 seconds, reflecting the diversity of patterns observed in the cross-
669 correlation analyses (Figure 7G; Figure 7C). MEC neurons thus possess specific and, in
670 some cases, finely tuned temporal scales over which their activity most informatively
671 encodes movement.

672 Finally, we asked whether these temporal relationships might be further refined
673 by considering how the magnitude of each neuron's activity constrains its behavioral
674 coding. Do extrema in neuronal firing rates reflect specific behaviors? To answer this
675 question, we calculated percentile cutoffs for each neuron's activity across the trial
676 using the MI maximized window sizes (Figure 7I). We then examined whether
677 increasing activity rates were associated with consistent changes in movement through
678 behavior space, indicating increasing behavioral specificity. To do so, we analyzed the
679 distribution of movements within behavior space by computing the entropy of the
680 distribution associated with neural activity bouts above each cutoff (Figures 7J-K). In
681 this metric, lower entropy values correspond to increased stereotypy in behavior space.
682 For many neurons, as activity increased the corresponding behavior space entropy
683 decreased, meaning that high levels of neuronal activity tended to be associated with
684 specific behavioral patterns (Figure 7K; Figure S6C). Hierarchical clustering of these
685 entropy distributions separated neurons into coherent groups with temporally similar
686 profiles (Figures 7L-M). Notably, a number of clusters displayed entropy minima that
687 were significantly offset from zero, meaning that peaks in neural activity could either
688 precede or follow changes in behavior (Figure 7M; Figure S6C). Taken together, these
689 results demonstrate that MEC neurons can encode locomotor information on temporally
690 varying scales, both before and after changes in behaviors occur. Moreover, these
691 analyses demonstrate how TREBLE can be used to uncover the relationship between
692 neuron activity and continuous behavior across a rich neural dataset from freely moving
693 animals.

694

695 **Discussion**

696 Technological developments have led to substantial shifts in how researchers are
697 attempting to unlock the development, expression, and evolution of animal behavior.
698 Facilitated by high throughput data analysis, a number of ambitious conceptual projects
699 have recently been proposed. These include considering behavior as a natural
700 extension of physics (Brown & de Bivort 2018), attempting to capture many species'
701 complete behavioral repertoires (Anderson & Perona 2014), or even reconciling the two
702 dominant lineages of behavioral research, comparative psychology and ethology (Datta
703 et al. 2019). Achieving these goals requires the study of individuals across a diversity of
704 animal species, in both controlled and naturalistic contexts, and in conjunction with
705 other data streams, critically including neural activity. Given this, behavioral analytical
706 methods that are statistically robust, broadly applicable, and easily linked to the function
707 of nervous systems are of general interest. Here we show how TREBLE can produce
708 low-dimensional representations of behavior that capture key features of the temporal
709 structure of animal movements and poses, allow one to visualize the phenotypic
710 consequences of perturbations, and reveal relationships to neural activity. By analogy
711 with genomics, these “heatmaps” can constitute analytical endpoints, but can also guide
712 the subsequent development of more targeted hypotheses.

713 All analytical methods for parsing behavior inherently make tradeoffs between
714 computational ease, temporal resolution, and generalizability. TREBLE does not require
715 extensive model fitting, is agnostic to feature type, and can accommodate many millions
716 of data points. Moreover, we describe how the key parameters needed, such as window
717 size, can be systematically explored and rationally chosen to capture behavioral

718 structure of interest. By cutting behavior into defined temporal windows that are
719 extremely densely sampled, TREBLE simultaneously captures both long and short
720 temporal structures in behavior. Furthermore, representation of these can be handled
721 flexibly. For example, here we represent temporal structure in two ways: 1) as
722 continuously varying time series and 2) as information theoretic summary measures of
723 entire distributions over time (using entropy and mutual information). It is also possible
724 for users to discretize TREBLE's output simply by segmenting movement through
725 behavior space into sub regions. Finally, TREBLE is generalizable in three different
726 ways. First, as we demonstrate, the framework can be applied to a variety of movement
727 sources, including synthetic trajectories and multiple animal species. Second, the
728 position of TREBLE in analysis pipelines is adjustable. While TREBLE can be
729 implemented as a standalone method for behavioral analysis from beginning to end, it
730 can also be used as a screening tool for identifying statistically significant behavioral
731 variation to be followed by other approaches specifically tailored to the question of
732 interest. Finally, as we show using optogenetic activation in flies and neural recordings
733 in mouse MEC, TREBLE can be used to find meaningful links between nervous
734 systems and behavior in a variety of contexts. This versatile nature of the TREBLE
735 framework therefore opens up the possibility for customizable high-throughput
736 behavioral analyses.

737 TREBLE can uncover otherwise opaque behavioral variation by directly
738 comparing behavioral data across many individuals and conditions. Here, we show that
739 free walking fruit flies possess substantial variation in locomotor dynamics and that this
740 variation is associated with specific regions of behavior space. Surprisingly, we find that

741 this variation in free walking behavior overlaps with the diversity of responses observed
742 after an optogenetic perturbation. A limited set of pre-determined behavioral
743 measurements, as would have been used typically, may have missed this variation.
744 Similarly, even unbiased statistical methods, without clear grounds for comparison
745 across animals or experimental conditions, would have been unlikely to have detected
746 these patterns. This highlights the benefit of using unsupervised statistical descriptions
747 such as TREBLE for assessing the structure of behavior across diverse individuals.
748 Furthermore, we envision that TREBLE might be employed on control data sets to
749 measure variability across baseline individuals in order to facilitate statistical power
750 calculations across the full range of behavioral dynamics that might be present. Such
751 catalogs of variability could be thus be developed across species' wildtype behavioral
752 repertoires, allowing researchers to better account for inborn variance or threshold
753 effects in applications such as genetic mapping, behavioral manipulations, and neural
754 perturbations (Lopez-Alonso et al. 2015; Ayroles et al. 2015; Buchanan et al. 2015).

755 Associating neural activity with continuous behavior presents two key challenges.
756 First, identifying the specific aspects of behavior that activity corresponds to is difficult to
757 determine in the context of a large behavior repertoire. Second, the temporal scales
758 over which this relationship occurs can be hard to know a priori. Here, we show how
759 TREBLE can be used to address these issues by characterizing the oscillatory patterns
760 of motor neuron activity during larval *Drosophila* crawling and the coding of locomotor
761 behavior in mouse medial entorhinal cortex (MEC) neurons as proofs of concept.

762 In the latter case, we find that TREBLE captures known components of speed
763 coding in MEC neurons while also uncovering extensive variation in temporal and rate-

764 based coding. Previous work has found that MEC neurons prospectively code for
765 changes in an animal's speed and position ~50-80ms in the future (Kropff et al. 2015).
766 Our analyses add to these findings, suggesting that MEC neurons can encode changes
767 both before *and* after behavioral events, doing so across a variety of timescales. These
768 patterns may reflect the presence a multitude of temporal encoding strategies for
769 behavior in the MEC, similar to what has been seen in basal ganglia movement
770 selection (Markowitz et al. 2018; Jin & Costa 2015), orofacial rhythms (Moore et al.
771 2013), and whole-brain activity during *C. elegans* locomotion (Kaplan et al. 2020). We
772 speculate that this may arise from some MEC neurons encoding short time-scale
773 behavioral events (e.g. turning left or right) while others update based on longer-term
774 navigational behaviors such as goal-oriented searching or foraging. These results
775 demonstrate how TREBLE can be leveraged to identify specific aspects of behavior that
776 are associated with the activity of individual neurons and to uncover the temporal
777 structure of such relationships. More broadly, we anticipate that TREBLE may be useful
778 in uncovering the statistical structure of behaviorally relevant activity across many
779 diverse populations of neurons.

780 There are a number of areas in which TREBLE may be further developed and
781 employed. First, movement through behavior space may be classified on timescale that
782 are longer than those presented here, affording descriptions of behavioral state over
783 time or the emergence of behavior in development. Second, the ability of TREBLE to
784 co-embed multiple individuals with temporal resolution may make it particularly
785 amenable to studying social and collective behavioral dynamics. Third, by readily
786 capturing variability in behavior, TREBLE may be amenable to exploring differences

787 across individuals arising from factors imbuing variance such as reaction norms,
788 behavioral syndromes, and environmental or genetic variation. Finally, it is especially
789 intriguing to consider how TREBLE may be further leveraged to jointly infer neural
790 coding principles by applying the framework to the structure of behavioral and neural
791 dynamics in parallel.

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810 **Acknowledgements**

811 We thank members of the Clandinin lab and Kiah Hardcastle for useful discussion and
812 input. We are grateful to Jeffrey Markowitz and Sandeep Robert Datta for providing the
813 mouse 3D pose dataset. This work was supported by the NIH (U19NS104655; T.R.C.)
814 (F31; A.C.R.), the Simons Foundation (T.R.C., L.M.G.), and the Stanford School of
815 Medicine Deans Office (R.A.Y.).

816

817 **Author contributions**

818 R.A.Y. and T.R.C. conceived the study. R.A.Y. wrote the TREBLE code and performed
819 data analysis with input from L.M.G on mouse behavior and MEC neuronal recordings.
820 A.C.R collected the larval *Drosophila* crawling data. R.A.Y., L.M.G., and T.R.C.
821 prepared the manuscript.

822

823 **Competing interests statement**

824 The authors declare no competing interests.

825

826

827

828

829

830

831

832

833 **Methods**

834 **Datasets**

835 Correlated random walks were produced using the `TrajGenerate` function in
836 the `trajr` R package (McLean 2018). Ten replicate walks were produced per
837 parameter tested (e.g. window size), sampled at a rate of 100 frames per second and
838 consisting of 10,000 frames. All walks were generated using the same underlying
839 angular and linear error distributions (Normal distribution; Mean = trajectory length;
840 Standard deviation = 0.5 (angular)/0.2 (linear)).

841 Details of the *Drosophila* walking dataset can be found in Haberkern et al. 2019.
842 Briefly, animals in the free-walking dataset ('WTB hybrid' genotype) were allowed to
843 explore a circular matt acrylic platform (radius 11.4 cm) surrounded by a siliconized
844 acrylic cylinder to prevent climbing. Videos were recorded from 120 cm above the
845 platform using a Flea3 camera at 12.3 Hz with a spatial resolution ~40 pixel/cm. Trials
846 lasted 10 minutes after a 1-2 minute acclimation period.

847 The local search data were produced from female flies possessing
848 optogenetically-accessible sugar receptor neurons (Gr64f-Gal4 > ChrimsonR). Animals
849 were allowed to explore a virtual landscape consisting of distributed 'cone forest' while
850 walking on a circular treadmill. Treadmill movement data (fly's position and velocity)
851 were collected at 360 Hz. Optogenetic stimulation was triggered whenever a fly crossed
852 within a 10mm radius of a given cone and persisted for 200ms.

853 For the *Drosophila* larvae analyses, embryos were collected for 1 h on
854 standard 3.0% agar molasses collection caps covered with a thin layer of wet yeast.
855 Twenty-four hours later, hatched embryos were transferred to standard cornmeal fly

856 food. After forty-eight hours (L2 larval stage), animals were collected and transferred to
857 a Petri dish with 1.2% agar and relocated to a behavioral room kept at 23°C and 60%
858 humidity. Ten to fifteen minutes after acclimation to the room, groups of 5 to 10 larvae
859 were transferred to a 30 x 30 cm 1.2% agar arena. After 15 to 30 seconds, locomotion
860 was recorded using a FIM imaging system (Risse et al 2013, [https://www.uni](https://www.uni-muenster.de)
861 [muenster.de](https://www.uni-muenster.de)) at 10 fps for 5 minutes. The FIM system was equipped with an azA2040-
862 25gm (Basler) camera and a LM16HC-SW (Kowa) lens. Individual larvae were then
863 tracked using FIMtrack software (Risse et al 2013).

864 Details of the mouse 3D pose dataset can be found in Marowitz et al. 2018 and
865 Wiltschko et al. 2015. Depth and position were collected using a Microsoft Kinect V2
866 while individual mice explored a circular arena (collected at 30Hz; 512x424 pixels frame
867 size). The mouse's center and orientation were estimated using an ellipse fit. An 80 x 80
868 pixel box was then drawn around the mouse and used to rotate the frame so that the
869 mouse was always facing the righthand side. These cropped and adjusted frames were
870 then used as input for principal component analysis from which the top 10 of these
871 postural PCs were used for downstream analyses.

872 Details of the mouse behavior and MEC dataset can be found in Hardcastle et al.
873 2017. Behavior and neural recordings were collected from two cohorts of adult wildtype
874 mice: 5 male and 2 female C57BL/6 mice (405 MEC cells recorded) and 7 male
875 C57BL/6J:129SVEV mice (389 MEC cells). Two polymide-coated platinum iridium 90%-
876 10% tetrodes were implanted in each mouse prior to the experimental period.
877 Behavioral data were collected approximately one week after surgery in large open
878 environments with black walls containing chocolate flavored cereal to induce foraging

879 (sizes varying; see Hardcastle et al. 2017 for details). The majority of recording
880 sessions lasted between 30-35 minutes with a small number ranging in time between 12
881 and 122 minutes. During each session, position, head direction, and running speed
882 were recorded every 20ms and single unite spikes were recorded at a 10kHz sampling
883 rate.

884

885 ***Generation of correlated random walk behavior spaces***

886 The trajectories and velocity distributions of replicate correlated random walks
887 were extracted using the custom function `iterative_umap`. First, a novel walk
888 trajectory was generated using `TrajGenerate` (as described above). Instantaneous
889 velocity components were then calculated (translation, angular velocity, sideways
890 velocity) and sampled using windows of a specified width (denoted here as w) and step
891 size (denoted here as s). The sampling procedure was as follows.

892 First, given frame i , the velocity components were extracted for frames $i:i+w$.
893 Angular and sideways velocity values were normalized to the first frame in the window
894 so that each velocity vector to originate from zero. Since we weren't concerned with
895 information pertaining to turn direction (i.e. left vs. right) the proceeding velocity values
896 were then adjusted so that the second frame was always positive. The resulting velocity
897 vectors were linearized and concatenated, resulting in a single vector of length $3w$. This
898 procedure was then repeated iteratively every s frames for the length of the trial (t).
899 After each iteration the concatenated vectors were appended to a matrix with $3w$ rows
900 to ultimately create t/s columns.

901 This procedure yielded a library of densely sampled velocity fragments
902 corresponding to behavioral dynamics for each moment in the trial. The `R`
903 implementation of the `UMAP` algorithm (McInnes et al. 2018) was then used to embed
904 these fragments into a low-dimensional behavior space (each point in space
905 corresponding to a window). The resulting space thus provided a 2-dimensional `xy`
906 position for each point/window in the trajectory. For downstream analyses we computed
907 simplified positional information for each space using the `bin_space` function. This
908 function decomposes behavior space into a grid of desired size (n bins x n bins). For
909 example, for a desired grid of 16 bins x 16 bins, a new set of `xy` coordinates will be
910 calculated corresponding to 16 intervals spanning the minimum/maximum `xy`
911 coordinates in the original space. The position of each point in the space will then be
912 compared to the new coordinates and associated with the closest bin, in this example
913 case generating a reduced set of 16x16, or 256, unique bins.

914

915 ***Analyzing correlated random walk parameter space***

916 We assessed the relationship of parameter choice and behavior space structure
917 by sweeping the two main window sampling parameters: window width and step size.
918 We examined 24 different window widths, ranging from 10ms to 2 seconds (10-40ms
919 sampled at 10ms steps; 50ms-2 seconds sampled at 100ms steps). Ten correlated
920 random walk replicates were generated for each window size and then processed and
921 embedded into behavior space using the `iterative_umap` function. We also
922 explored the effect of step size by producing replicate behavior spaces from trajectories
923 sampled at step sizes varying from 10 ms (i.e. 1 frame) to 1 second, separated by 100

924 millisecond intervals. Again, for each step size 10 correlated random walk replicates
925 were generated, processed, and embedded into behavior space using the
926 `iterative_umap` function.

927 Intra-point Euclidean distance was calculated for each space by comparing the
928 spatial position of temporally adjacent points. Specifically, the position of a given point
929 (corresponding to a unique temporal window) in behavior space (x_t , y_t) was compared
930 to the position of the next point/window in time (x_{t+1} , y_{t+1}). The Euclidean distance
931 between these points was then calculated and stored into a vector, allowing for the
932 distribution of distances to be compared across replicates and parameter conditions.
933 The coefficient of variation was also calculated for these measures to control for
934 substantial variation in the magnitude of effect; this was done by dividing the
935 distribution's standard deviation by its mean.

936 Procrustes distance was used to measure the similarity between replicate
937 behavior spaces of a certain parameter combination (Dryden & Mardia 1998). Briefly,
938 this method treats spaces and their component points as a set of landmarks for pairwise
939 comparison. To do so, pairs of behavior spaces are scaled to be similar sizes then
940 shifted and rotated to possess the same position and orientation in space (Dryden &
941 Mardia 1998). The distance between corresponding points in the two adjusted landmark
942 sets is then calculated using Procrustes distance (reported here as Root Mean Square
943 Error; RMSE). We performed this procedure and calculated Procrustes distance for all
944 pairwise combinations of the 10 replicate behavior spaces within a given parameter set.
945 Coefficient of variation for this measure was also calculated as above.

946 To assess recurrence, we measured the average return time for all points in each
947 space. We first used the binning procedure outlined above to create a 32x32 grid so
948 that the resulting space was composed of 1024 unique bins. We then computed a
949 distance threshold to delineate a “neighborhood” around each bin (5% smallest intra-bin
950 distance). For a given bin, each time the trajectory passed through its threshold the
951 event was recorded and its duration stored. This distribution of return delays was
952 analyzed by computing the proportion of points that displayed a return for a given time
953 delay. For example, given a delay of 150ms, all bins would be scanned and the
954 proportion possessing returns that occurred took between 150 and 160ms would be
955 calculated. We calculated proportions in this manner for delays between 0ms and 2
956 seconds at 10ms intervals (as plotted in Figure 1I). Mean recurrence time was
957 measured by collecting all observed return times for each replicate and calculating the
958 mean.

959

960 ***Generation of free-walking fruit fly behavior spaces***

961 The movement trajectories of 20 female WTB flies walking in a circular area were
962 collected from Haberkern et al. 2019. For each fly we computed velocity components as
963 above and interpolated the resulting vectors from 12.3 Hz to 50 Hz to increase
964 smoothness in the downstream behavior space. Initial analyses suggested the removal
965 of 6 trials due to missing data or lack of substantial movement, resulting in a final
966 dataset of 14 flies.

967 As above, we performed an empirical test to decide on the optimal window size
968 for behavior space creation. To do so, we sampled the first 2000 frames from each trial

969 (to increase computational efficiency) and swept through window sizes ranging from
970 20ms to 1 second (70 ms step size), creating a single behavior space for each of the 14
971 flies per window size. The resulting spaces were then directly compared using the
972 Euclidean distance, Procrustes distance, and recurrence metrics previously described
973 (Figure S2).

974

975 ***Analyzing free-walking fruit fly behavior spaces***

976 After sweeping window parameters, we generated a behavior space composed
977 of all 14 free-walking trials using a window size of 160ms (399,869 windows). We then
978 used `bin_space` to calculate point coordinates in a 64x64 grid. Stereotypy in
979 movement through the space was visualized using a vector field transformation (Figure
980 2D). All instances in which trajectories through the space passed into a given bin were
981 collected and then used to calculate the mean x and y vectors of the trajectory leaving
982 the bin. These mean values were then represented visually using arrows that originated
983 out of the corresponding bin, the direction and magnitude of which were dictated by the
984 mean x and y vectors. Velocity distributions across the space (Figures. 2F-H) were
985 visualized by calculating the mean value for each velocity component per window.
986 These values were then used as input to the color function determining the hue of each
987 point in the space.

988 Intra-fly variation in behavior space was assessed using the reduced 64x64 grid.
989 For each fly we calculated the number of times its trajectory passed through a given bin,
990 in addition to a binary measure of whether that bin was visited at all ('0' if not; '1' if
991 visited at least once). The continuous counts of bin visits were then used to calculate

992 bin-wise mean and variance across all flies, the latter of which is visualized in Figure 3F.
993 The binary measure was used to assess variation in overall space occupancy across
994 flies (Figure S2).

995 We used 2-dimensional histograms to compare the overall behavioral patterns of
996 individual flies. To do so, we created an individual map for each trial comparing the
997 binned distributions of x and y coordinates via 2-dimensional kernel density estimation
998 using the function `kde2d` (bandwidth = 2; 32 grid points in each direction) in the `MASS`
999 `R` package. To facilitate comparisons across trials, the resulting density maps were
1000 normalized to the max density value of each and then linearized so that each was
1001 represented by a single vector of density values. The relationships between trials were
1002 inferred by calculating a distance matrix of these density vectors which was then used
1003 as input for hierarchical clustering (`hclust` function; Figure 3G).

1004

1005 ***Generation of optogenetic local search behavior spaces***

1006 The movement trajectories of 19 Gr64f-Gal4 > ChrimsonR flies in virtual reality
1007 was collected from Haberkern et al. 2019. We processed and calculated velocity from
1008 these files as above, down sampling from 360 to 50 Hz to match the sampling rate of
1009 the free-walking dataset (for downstream comparisons). Windows were extracted from
1010 all 19 trials using the same width (160 ms) as above, yielding a library of 1,110,025
1011 windows. We then used the `predict` function in the `UMAP` `R` package to embed all
1012 windows, one fly at a time, using the free-walking fly behavior space as a template. This
1013 procedure produced a combined behavior space composed of 33 trials (1,509,894
1014 windows) that facilitated comparisons across experimental conditions and individuals.

1015 Variation in space occupancy between free-walking and local search trials was
1016 assessed using 2-d density maps. As before, we used `bin_space` to get new
1017 coordinates for each trial, and then calculated density maps using `kde2d` (bandwidth =
1018 1; 200 grid points in each direction), which were then normalized by dividing each
1019 density estimate by the maximum value in the map. As in Figure 3G these density maps
1020 were then linearized, combined in a matrix, and clustered using Euclidean distance and
1021 hierarchical clustering to produce the tree in Figure 4F.

1022

1023 ***Analyzing optogenetic local search behavior spaces***

1024 We used a bin-wise Kruskal-Wallis test to statistically analyze differences in
1025 space occupancy between the two groups. For a given fly/trial we calculated the percent
1026 occupancy at each bin (number of visits to bin divided by total number of windows). A
1027 Kruskal-Wallis test was then used to compare the percent occupancies of the individual
1028 trials between free-walking and local search flies from which the test statistic and
1029 accompanying *p*-values were collected. We used Bonferroni correction to adjust these
1030 *p*-values, controlling for the number of tests performed (4,096). The adjusted *p*-values
1031 were then used to visually assess regions of greater differentiation between the free
1032 walking and local search trials, plotted as the -log10 transformation of the *p*-values (as
1033 seen in Figure 3F).

1034 Time-evolving responses to optogenetic stimulation were assessed using density
1035 maps. To do so, we sampled the behavior space trajectories of each trial before and
1036 after all bouts of optogenetic stimulation. For each bout of stimulation, we extracted
1037 positions in behavior space for the second before stimulation in order to represent a

1038 baseline behavioral distribution. We then extracted positions immediately after
1039 stimulation using 1 second windows and step size of 100ms, from 0ms to 1 second
1040 afterward (Figure 4E). We combined positions across bouts and individuals within
1041 window to produce an aggregated response profile using kernel density estimation
1042 (bandwidth = 2; 64 grid points in each direction; common x and y limits across
1043 windows). If individuals displayed common behavioral responses to stimulation in a
1044 specific window, then the related density map should show structure in its distribution
1045 (i.e. concentrated red regions in Figure 4E).

1046

1047 **Larval Drosophila analysis**

1048 FIMtrack (Risse *et al.* 2014; Risse *et al.* 2017) was used to track the behavior of
1049 72 *Drosophila* larvae while crawling on an agar surface (collected at 10Hz). FIMtrack
1050 outputs a number of per-frame measurements representation an animal's shape and
1051 orientation. We selected primary measurements reflecting larval size, shape, and
1052 velocity (Figure 4A) in addition to the angular velocity of the head, midpoint, and tail for
1053 analysis. Due to variation in the mean of size measurements (i.e. area, perimeter, radii,
1054 spine length) over individual trials, these measures were detrended using the `ma`
1055 function in the R package `forecast` (window size = 10) after which all measures were
1056 then converted to z-scores. Given that the information of some of these features may be
1057 redundant, a principal component analysis was used to find an appropriate number of
1058 axes that could explain the variation in the dataset. We found that 8 principal
1059 components explained >90% of the variance in the feature set. These top 8 components
1060 were used as input to the iterative window procedure, sweeping a range of windows

1061 between 100ms and 5 seconds. As before the resulting spaces were compared using
1062 Euclidean distance, Procrustes distance, and recurrence metrics (Figures S3A-C). After
1063 analyzing these metrics, a window size of 800ms was chosen for all downstream
1064 analyses.

1065 The full behavior space was plotted as vector field and with features highlighted
1066 in the fashion described above in the analyses of adult *Drosophila* behavior. Behavior
1067 labels and movement patterns through the behavior space (Figure 4D) were
1068 qualitatively assessed. Temporal patterns in the movement through behavior space
1069 were assessed via autocorrelation (Figure S3E). Autocorrelation was measured using
1070 the acf function in R (lag size = 100).

1071 We used calcium imaging data from Tastekin *et al.* 2018 to examine neural
1072 activity during larval locomotion. We analyzed motor neuron activity (7 neurons per
1073 side) of larvae that were performing fictive locomotion (collected at 4-5Hz). Neural
1074 activity was measured using GcamP6f expressed in glutamatergic neurons (CG9887-
1075 lexA > GCamP6f) during optogenetic activation of PDM-DN neurons (PDM-
1076 DN>CsChrimson::mVenus) which induces stopping behavior. For each time point
1077 calcium fluorescence was converted to $\Delta F/F$ and then converted to z-scores for
1078 comparison across trials. A window size of 1 second was chosen for behavior space
1079 creation after the iterative window procedure. All trials were embedded in the same
1080 space and features/vector fields were plotted as previously described.

1081

1082 **Mouse 3D behavior analysis**

1083 Mouse 3D behavior data were previously published in Markowitz et al. 2018. In
1084 each trial, the mouse 3D posture was measured using the MoSeq pipeline (Wiltschko et
1085 al. 2015) from which the following features were calculated and used in this analysis:
1086 height, length, width, velocity (2-dimensional), velocity (3-dimensional), velocity (theta),
1087 and 10 postural principal components calculated from an 80 pixel x 80 pixel
1088 representation of the mouse's position and height in 3D space. These features were
1089 then used an input to the iterative window procedure and a range of window sizes
1090 between 33ms and 1.66 seconds was examined (Figures S4A-C). A final window size of
1091 133ms was chosen.

1092 The full behavior space was plotted as vector field and with features highlighted
1093 as described above. Behavior labels and movement patterns through the behavior
1094 space (Figure 5D) were qualitatively assessed. The distributions of MoSeq syllables in
1095 behavior space (Figures 5K, S4E) were assessed by associating the timing of each
1096 syllable's occurrence with the corresponding xy positions in behavior space. These
1097 were then used to calculate a 2-dimensional probability density function of the xy
1098 coordinates in space (100x100 grid). The density at each point in the 100x100 grid was
1099 then dividing by the maximum value so that the distribution varied between 0 and 1. The
1100 top 95% of these values were then plotted as a heatmap over the full behavior space
1101 distribution (Figures 5K, S4E).

1102 The dispersion of syllables in behavior space was assessed via nearest neighbor
1103 distance using the function `nndist` ($k = 1$) in the R package `spatstat` (Baddeley et
1104 al. 2015). The significance of per-syllable dispersion was then computed via
1105 permutation tests. For each syllable, the mean nearest neighbor (nn) distance was

1106 calculated. The timing of syllable occurrence was then randomly shuffled 10,000 times.
1107 During each permutation the distribution of the shuffled data in behavior space was
1108 measured and used to calculate mean nn distance. Significance was assessed by
1109 computing *p*-values comparing the number of occurrences in which the shuffle nn
1110 distances were smaller than the observed mean nn distance, divided the number of
1111 permutations. Bonferroni correction was then used to adjust the *p*-values given the
1112 number of syllables tested.

1113 A regularized generalized linear model was used to examine the relationship
1114 between mouse 3D pose behavior space and the original input features. The model was
1115 created with the R package `glmnet` (Friedman et al. 2010) using position in behavior
1116 space as the outcome variable and the features (height, length, width, velocity (2-
1117 dimensional), velocity (3-dimensional), velocity (theta), 10 PCs) as predictors. The set
1118 training set was composed of 75% of the data. We used 10-fold cross-validation via the
1119 `caret` (Kuhn 2016) package to compute the optimal alpha and lambda values for
1120 regularizing coefficient weights. The fit of the final model (R^2) was computed by
1121 comparing the predicted behavior space positions in the remaining 25% of the data to
1122 the actual values. To measure variability across mice, individual models were also
1123 created for each trial in the same fashion, the fits of which are compared in Figure S4F.

1124

1125 ***Generation of mouse 2D locomotor behavior space***

1126 Behavior data and neural recordings were previously published in from
1127 Hardcastle et al. 2017. For each cell, full trial spike event data (collected at 10kHz) and
1128 positional coordinates (collected at 50Hz) were extracted. As above, velocity

1129 components over the course of the whole trial were computed from the positional
1130 coordinates. We then used the empirical window-size test to identify the optimal window
1131 size for behavior space creation. 30 trials were randomly sampled for the iterative
1132 window test. The first 3000 frames of each trial were used to sweep through window
1133 sizes ranging from 20ms to 1 second (40ms step size; Figures S5A-F). The resulting
1134 spaces were then directly compared using the Euclidean distance, Procrustes distance,
1135 and recurrence metrics (Figures 6A-F). After analyzing these metrics a window size of
1136 400ms was chosen for all downstream analyses.

1137 Given the large size of the dataset (41,850,995 windows) we opted to perform a
1138 seed-embedding procedure to produce individual behavior spaces for each trial. To do
1139 so, we extracted and combined together the first 5,000 windows from each trial and
1140 created a seed behavior space using UMAP. We then individually embedded the
1141 remaining windows for each trial using the `predict` function in the R implementation of
1142 UMAP. Time points and XY coordinates corresponding to movement through these
1143 individual behavior spaces were then combined and used jointly for overall annotation of
1144 the mouse locomotion (As seen in Figures 6B-E). This all-trial behavior space was used
1145 to calculate 64x64 bin coordinates, allowing all mouse trials to be directly compared
1146 within the same behavior space architecture. Vector field representation (Figure 6D)
1147 and physical XY space movement decoding (Figure 6E) were computed as previously
1148 described.

1149 The spike times for all cells in a given trial were then associated with the
1150 corresponding behavioral timepoints. To do so, spike rates were calculated over 20ms
1151 bins corresponding to the sampling rate of mouse positional data (50Hz) and, thus, the

1152 rate of movement through behavior space. The rate of activity was also calculated over
1153 1 second (1Hz) scales to explore the extent to which broad temporal differences might
1154 be present between neural activity and behavior. For each cell/trial pair the Pearson
1155 correlation between behavior space position and firing rate was calculated. The
1156 significance of these correlations was measured using permutation tests. For each cell,
1157 the spike rate data were shuffled 10,000 times and then correlated with position in
1158 behavior space. These correlations were used to calculate *p*-values by comparing the
1159 number of times shuffled correlations were greater than the observed value, divided by
1160 the number of permutations (10,000; Figure 5I). 2-dimensionsal tuning curves (as in
1161 Figures 6F-G and Figure S5M) were computed for each cell by calculating the mean
1162 spike rate in each bin using the 64x64 representation previously calculated. Mutual
1163 information between the cells and behavior space was calculated (in bits) by comparing
1164 the resulting 2-d tuning curves to the distribution of average occupancy time per bin for
1165 each trial using the function `mi.plugin` from the R package `entropy`. The canonical
1166 MEC coding variables (speed score, spatial stability, head direction, border score;
1167 Figures 6K-R) used were previously computed following Hardcastle *et al.* 2017. Cross-
1168 correlations between neural activity and behavior space position were calculated with
1169 the function `ccf` in R (lag of 16 seconds as seen in Figure 7C).

1170 Temporal variation in the association of neural activity and behavior space (as in
1171 Figures 7D-H) was assessed by calculating the correlation coefficient, entropy, and
1172 mutual information between the two across a range of windows (0-4 seconds, 20ms
1173 steps). Correlations and mutual information were calculated for each window size as
1174 above. Shannon entropy (in bits) was calculated using the distribution of mean firing

1175 values from 2-d neural tuning curve for each window (using the `entropy` function from
1176 the `entropy` R package). MI maximized window sizes were chosen for each cell using
1177 the timescale at which the maximum mutual information between behavior space
1178 position and neural activity occurred.

1179 Rate-based differences in coding were assessed using the MI maximized window
1180 sizes identifying above. For each cell, we calculated the activity rates corresponding to
1181 the 75th, 80th, 90th, 95th, and 99th percentiles. We then identified the time points at
1182 which each cell's activity occurred above the respective percentile. The corresponding
1183 positions in behavior space were then extracted in 8 second windows surrounding each
1184 timepoint (from 4 seconds before to 4 seconds after; 20ms bins). 2-d tuning curves were
1185 then constructed from the distribution of points in behavior space for all events that
1186 occurred across the corresponding 20ms bins. For example, the behavior space
1187 positions occurring 100-80ms before all events above the 99th percentile would be
1188 compared, followed by the same calculation for all events occurring 80-60ms before,
1189 then 60-40ms before, etc. This procedure would thus produce 400 2-d tuning curves per
1190 percentile, corresponding to a sampling rate of 50Hz over a course of 8 seconds. The
1191 Shannon entropy (in bits) of each 2-d tuning curve was then calculated and plotted as a
1192 curve over time to detect consistent changes in its distribution (Figure 7K). To
1193 compare these entropy distributions across all cells each was divided by its maximum
1194 value. The normalized entropy measures for the 75th and 99th percentile values were
1195 then used as input to hierarchically cluster all cells. Clusters were identified using a
1196 dynamic tree cutting algorithm via the `cutreeHybrid` function in the R package
1197 `dynamicTreeCut` (minimum cluster size of 10) (Langfelder et al. 2007).

1198 **References**

1199 1. Allen, B.D., Singer, A.C., and Boyden, E.S. (2015). Principles of designing
1200 interpretable optogenetic behavior experiments. *Learn Mem* 22, 232-238.

1201 2. Anderson, D.J., and Perona, P. (2014). Toward a science of computational
1202 ethology. *Neuron* 84, 18-31.

1203 3. Ayroles, J.F., Buchanan, S.M., O'Leary, C., Skutt-Kakaria, K., Grenier, J.K.,
1204 Clark, A.G., Hartl, D.L., and de Bivort, B.L. (2015). Behavioral idiosyncrasy
1205 reveals genetic control of phenotypic variability. *Proc Natl Acad Sci U S A* 112,
1206 6706-6711.

1207 4. Baddeley, A., Rubak, E., Turner, R. (2015). *Spatial Point Patterns: Methodology*
1208 and *Applications with R*. (Chapman and Hall/CRC Press).

1209 5. Berman, G.J., Choi, D.M., Bialek, W., and Shaevitz, J.W. (2014). Mapping the
1210 stereotyped behaviour of freely moving fruit flies. *J R Soc Interface* 11.

1211 6. Bovet, P.B., S. (1988). Spatial analysis of animal's movements using a correlated
1212 random walk model. *Journal of Theoretical Biology* 131, 419-433.

1213 7. Brown, A.E., Yemini, E.I., Grundy, L.J., Jucikas, T., and Schafer, W.R. (2013). A
1214 dictionary of behavioral motifs reveals clusters of genes affecting *Caenorhabditis*
1215 *elegans* locomotion. *Proc Natl Acad Sci U S A* 110, 791-796.

1216 8. Brown, A.E.X., and de Bivort, B. (2018). Ethology as a physical science. *Nature*
1217 *Physics* 14, 653-657.

1218 9. Bruno, A.M., Frost, W.N., and Humphries, M.D. (2017). A spiral attractor network
1219 drives rhythmic locomotion. *Elife* 6.

1220 10. Buchanan, S.M., Kain, J.S., and de Bivort, B.L. (2015). Neuronal control of
1221 locomotor handedness in *Drosophila*. *Proc Natl Acad Sci U S A* 112, 6700-6705.

1222 11. Calhoun, A.J., and Murthy, M. (2017). Quantifying behavior to solve sensorimotor
1223 transformations: advances from worms and flies. *Curr Opin Neurobiol* 46, 90-98.

1224 12. Calhoun, A.J., Pillow, J.W., and Murthy, M. (2019). Unsupervised identification of
1225 the internal states that shape natural behavior. *Nat Neurosci* 22, 2040-2049.

1226 13. Cande, J., Namiki, S., Qiu, J., Korff, W., Card, G.M., Shaevitz, J.W., Stern, D.L.,
1227 and Berman, G.J. (2018). Optogenetic dissection of descending behavioral
1228 control in *Drosophila*. *Elife* 7.

1229 14. Clark, M.Q., McCumsey, S.J., Lopez-Darwin, S., Heckscher, E.S., and Doe, C.Q.
1230 (2016). Functional Genetic Screen to Identify Interneurons Governing
1231 Behaviorally Distinct Aspects of *Drosophila* Larval Motor Programs. *G3*
1232 (Bethesda) 6, 2023-2031.

1233 15. Clark, M.Q., Zarin, A.A., Carreira-Rosario, A., and Doe, C.Q. (2018). Neural
1234 circuits driving larval locomotion in *Drosophila*. *Neural Dev* 13, 6.

1235 16. Corfas, R.A., Sharma, T., and Dickinson, M.H. (2019). Diverse Food-Sensing
1236 Neurons Trigger Idiothetic Local Search in *Drosophila*. *Curr Biol* 29, 1660-1668
1237 e1664.

1238 17. Datta, S.R., Anderson, D.J., Branson, K., Perona, P., and Leifer, A. (2019).
1239 Computational Neuroethology: A Call to Action. *Neuron* 104, 11-24.

1240 18. Dryden, I.L., and Mardia, K.V. (1998). Statistical shape analysis (Chichester:
1241 John Wiley and Sons).

1242 19. Duistermars, B.J., Pfeiffer, B.D., Hooper, E.D., and Anderson, D.J. (2018). A
1243 Brain Module for Scalable Control of Complex, Multi-motor Threat Displays.
1244 *Neuron* 100, 1474-1490 e1474.

1245 20. Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization Paths for
1246 Generalized Linear Models via Coordinate Descent. *J Stat Softw* 33, 1-22.

1247 21. Glaser, J.I., and Kording, K.P. (2016). The Development and Analysis of
1248 Integrated Neuroscience Data. *Front Comput Neurosci* 10, 11.

1249 22. Haberkern, H., Basnak, M.A., Ahanonu, B., Schauder, D., Cohen, J.D., Bolstad,
1250 M., Bruns, C., and Jayaraman, V. (2019). Visually Guided Behavior and
1251 Optogenetically Induced Learning in Head-Fixed Flies Exploring a Virtual
1252 Landscape. *Curr Biol* 29, 1647-1659 e1648.

1253 23. Hafting, T., Fyhn, M., Molden, S., Moser, M.B., and Moser, E.I. (2005).
1254 Microstructure of a spatial map in the entorhinal cortex. *Nature* 436, 801-806.

1255 24. Hardcastle, K., Maheswaranathan, N., Ganguli, S., and Giocomo, L.M. (2017). A
1256 Multiplexed, Heterogeneous, and Adaptive Code for Navigation in Medial
1257 Entorhinal Cortex. *Neuron* 94, 375-387 e377.

1258 25. Heckscher, E.S., Lockery, S.R., and Doe, C.Q. (2012). Characterization of
1259 Drosophila larval crawling at the level of organism, segment, and somatic body
1260 wall musculature. *J Neurosci* 32, 12460-12471.

1261 26. Jin, X., and Costa, R.M. (2015). Shaping action sequences in basal ganglia
1262 circuits. *Curr Opin Neurobiol* 33, 188-196.

1263 27. Johnson, R.E., Linderman, S., Panier, T., Wee, C.L., Song, E., Herrera, K.J.,
1264 Miller, A., and Engert, F. (2020). Probabilistic Models of Larval Zebrafish
1265 Behavior Reveal Structure on Many Scales. *Curr Biol* 30, 70-82 e74.

1266 28. Kaplan, H.S., Salazar Thula, O., Khoss, N., and Zimmer, M. (2020). Nested
1267 Neuronal Dynamics Orchestrate a Behavioral Hierarchy across Timescales.
1268 *Neuron* 105, 562-576 e569.

1269 29. Kaplan, H.S., Salazar Thula, O., Khoss, N., and Zimmer, M. (2020). Nested
1270 Neuronal Dynamics Orchestrate a Behavioral Hierarchy across Timescales.
1271 *Neuron* 105, 562-576 e569.

1272 30. Katsov, A.Y., Freifeld, L., Horowitz, M., Kuehn, S., and Clandinin, T.R. (2017).
1273 Dynamic structure of locomotor behavior in walking fruit flies. *Elife* 6.

1274 31. Krakauer, J.W., Ghazanfar, A.A., Gomez-Marin, A., MacIver, M.A., and Poeppel,
1275 D. (2017). Neuroscience Needs Behavior: Correcting a Reductionist Bias.
1276 *Neuron* 93, 480-490.

1277 32. Kropff, E., Carmichael, J.E., Moser, M.B., and Moser, E.I. (2015). Speed cells in
1278 the medial entorhinal cortex. *Nature* 523, 419-424.

1279 33. Kuhn, M. (2008). Building Predictive Models in R Using the caret Package.
1280 *Journal of Statistical Software* 28, 1-26.

1281 34. Langfelder, P., Zhang, B., and Horvath, S. (2008). Defining clusters from a
1282 hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* 24,
1283 719-720.

1284 35. Lopez-Alonso, V., Cheeran, B., Rio-Rodriguez, D., and Fernandez-Del-Olmo, M.
1285 (2014). Inter-individual variability in response to non-invasive brain stimulation
1286 paradigms. *Brain Stimul* 7, 372-380.

1287 36. Markowitz, J.E., Gillis, W.F., Beron, C.C., Neufeld, S.Q., Robertson, K., Bhagat,
1288 N.D., Peterson, R.E., Peterson, E., Hyun, M., Linderman, S.W., et al. (2018). The
1289 Striatum Organizes 3D Behavior via Moment-to-Moment Action Selection. *Cell*
1290 174, 44-58 e17.

1291 37. Marques, J.C., Lackner, S., Felix, R., and Orger, M.B. (2018). Structure of the
1292 Zebrafish Locomotor Repertoire Revealed with Unsupervised Behavioral
1293 Clustering. *Curr Biol* 28, 181-195 e185.

1294 38. Mathis, A., Schneider, S., Lauer, J., and Mathis, M.W. (2020). A Primer on
1295 Motion Capture with Deep Learning: Principles, Pitfalls, and Perspectives.
1296 *Neuron* 108, 44-65.

1297 39. McInnes, L., Healy, J., and Melville, J. (2018). UMAP: Uniform Manifold
1298 Approximation and Projection for Dimension Reduction. In *arXiv e-prints*, pp.
1299 arXiv:1802.03426.

1300 40. McLean, D.J.S.V., M.A. (2018). trajr: An R package for characterisation of animal
1301 trajectories. *Ethology* 00, 1-9.

1302 41. Moore, J.D., Deschenes, M., Furuta, T., Huber, D., Smear, M.C., Demers, M.,
1303 and Kleinfeld, D. (2013). Hierarchy of orofacial rhythms revealed through
1304 whisking and breathing. *Nature* 497, 205-210.

1305 42. Pereira, T.D., Shaevitz, J.W., and Murthy, M. (2020). Quantifying behavior to
1306 understand the brain. *Nat Neurosci* 23, 1537-1549.

1307 43. Risse, B., Berh, D., Otto, N., Klambt, C., and Jiang, X. (2017). FIMTrack: An
1308 open source tracking and locomotion analysis software for small animals. PLoS
1309 Comput Biol 13, e1005530.

1310 44. Risse, B., Otto, N., Berh, D., Jiang, X., and Klambt, C. (2014). FIM imaging and
1311 FIMtrack: two new tools allowing high-throughput and cost effective locomotion
1312 analysis. J Vis Exp.

1313 45. Robie, A.A., Hirokawa, J., Edwards, A.W., Umayam, L.A., Lee, A., Phillips, M.L.,
1314 Card, G.M., Korff, W., Rubin, G.M., Simpson, J.H., et al. (2017). Mapping the
1315 Neural Substrates of Behavior. Cell 170, 393-406 e328.

1316 46. Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B.L., Witter, M.P., Moser, M.B.,
1317 and Moser, E.I. (2006). Conjunctive representation of position, direction, and
1318 velocity in entorhinal cortex. Science 312, 758-762.

1319 47. Solstad, T., Boccara, C.N., Kropff, E., Moser, M.B., and Moser, E.I. (2008).
1320 Representation of geometric borders in the entorhinal cortex. Science 322, 1865-
1321 1868.

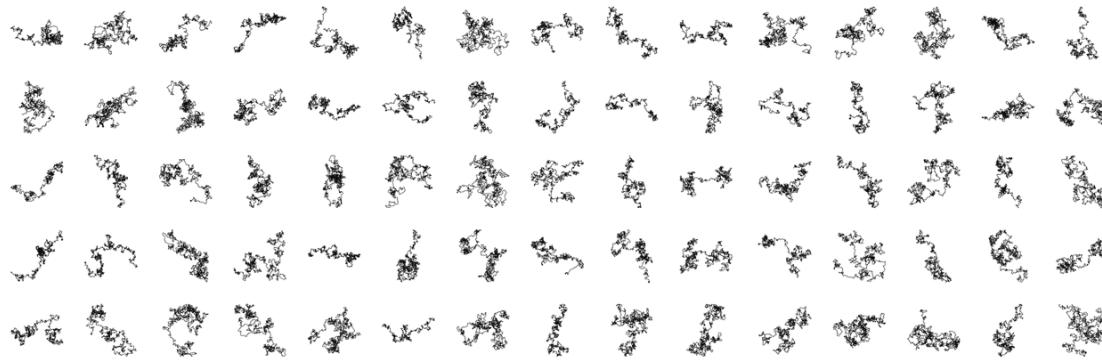
1322 48. Staden, R. (1979). A strategy of DNA sequencing employing computer programs.
1323 Nucleic Acids Res 6, 2601-2610.

1324 49. Tastekin, I., Khandelwal, A., Tadres, D., Fessner, N.D., Truman, J.W., Zlatic, M.,
1325 Cardona, A., and Louis, M. (2018). Sensorimotor pathway controlling stopping
1326 behavior during chemotaxis in the *Drosophila melanogaster* larva. Elife 7.

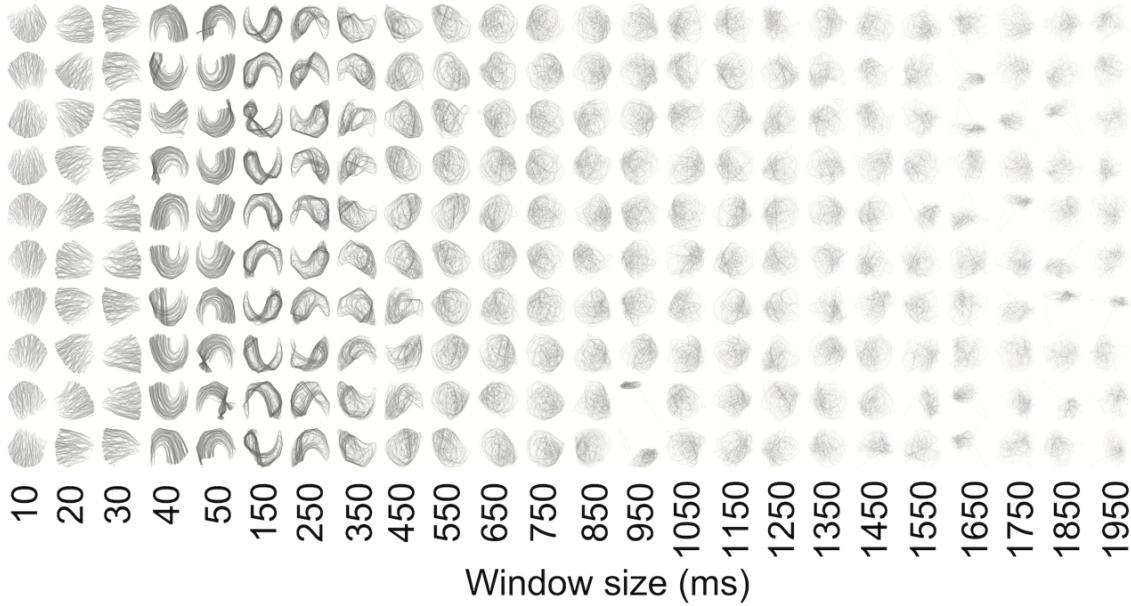
1327 50. Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for
1328 transcriptomics. Nat Rev Genet 10, 57-63.

1329 51. Wiltschko, A.B., Johnson, M.J., Iurilli, G., Peterson, R.E., Katon, J.M.,
1330 Pashkovski, S.L., Abraira, V.E., Adams, R.P., and Datta, S.R. (2015). Mapping
1331 Sub-Second Structure in Mouse Behavior. *Neuron* 88, 1121-1135.
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351

A

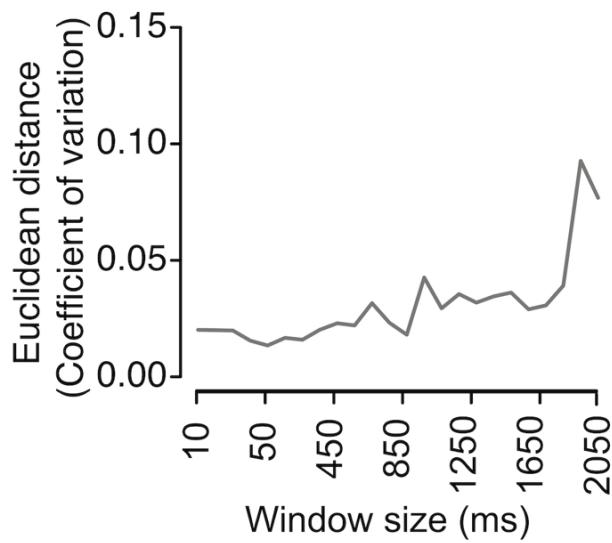


B

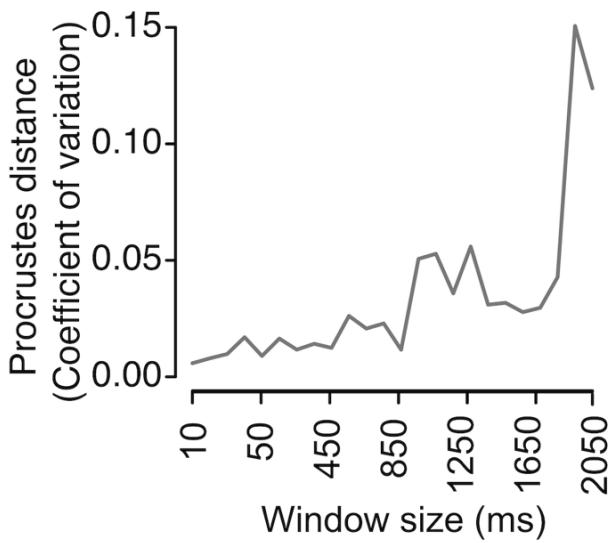


Window size (ms)

C



D



1353 **Figure S1 Correlated random walk iterative window tests**

1354 **(A)** A set of 100 example correlated random walks used for parameter tuning. Each is
1355 10,000 frames long and sampled at 100 FPS.

1356 **(B)** Trajectories through behavior space are plotted for each of the 10 correlated
1357 random walk replicates per window size tested. Replicates are arranged in rows while
1358 window sizes are arranged in columns. Trajectories were visualized by connecting
1359 temporally adjacent points in behavior space with lines. Recurrent dynamics are
1360 represented by highly overlapping lines, reflecting repeated excursions through those
1361 specific paths.

1362 **(C)** Coefficient of variation of mean intra-point Euclidean distance as a function of
1363 window size.

1364 **(D)** Coefficient of variation of Procrustes distance RMSE as a function of window size.

1365

1366

1367

1368

1369

1370

1371

1372

1373

1374

1375

1376

1377

1378

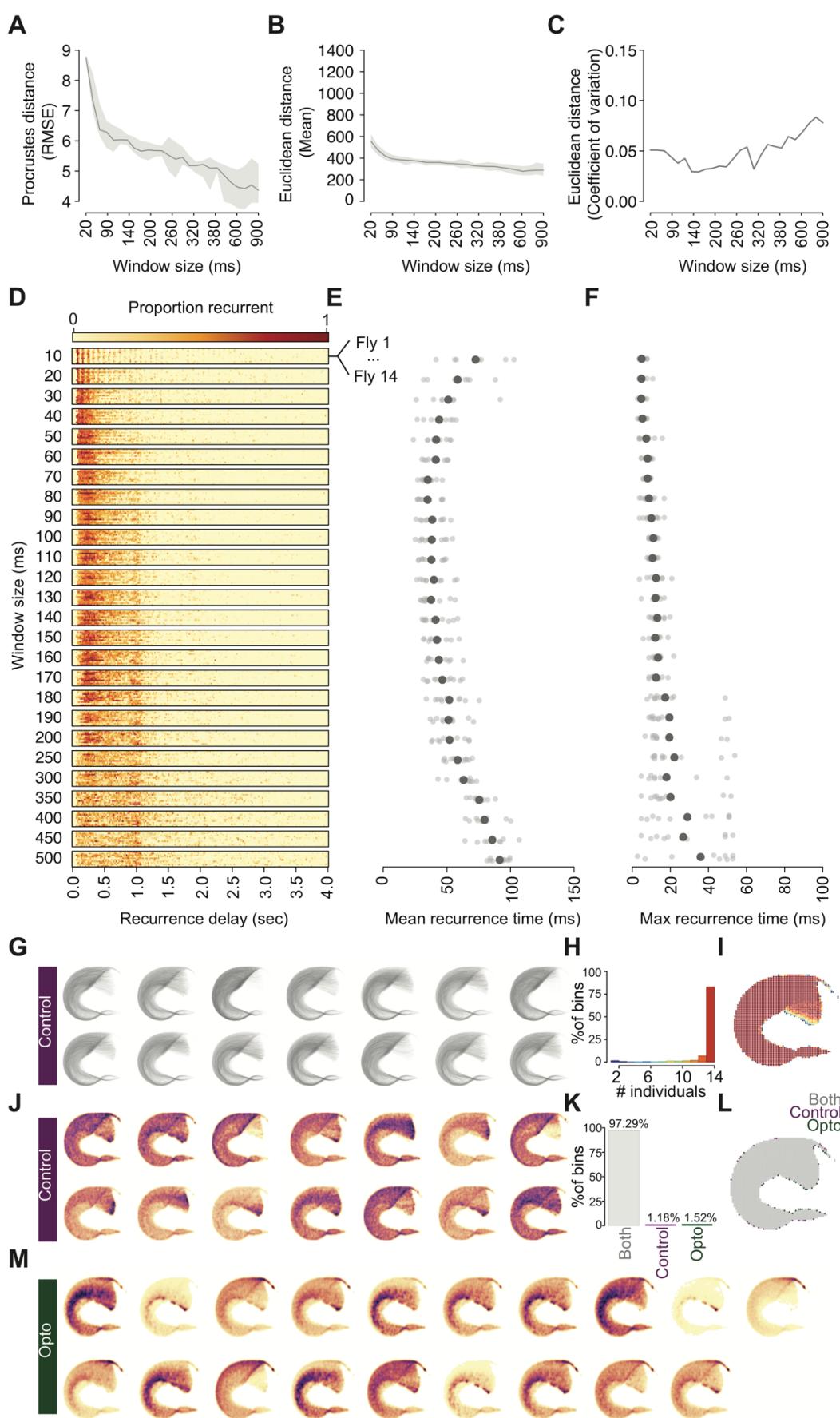
1379

1380

1381

1382

1383



1385 **Figure S2 Assessing structural and temporal components of the free-walking fly
1386 dataset**

1387 **(A)** Procrustes distance RMSE measures as a function of window size. The darker line
1388 corresponds to the mean value while the shading reflects standard error.

1389 **(B)** Mean intra-point Euclidean distances. The darker line corresponds to the mean
1390 value while the shading reflects standard error.

1391 **(C)** Coefficient of variation of mean intra-point Euclidean distance as a function of
1392 window size.

1393 **(D)** Recurrence plot of behavior spaces produced from the free-walking dataset. The
1394 proportion of recurrent points given a range of time delays ranging from 0 to 4 seconds
1395 is indicated by the color of the corresponding bins (from light yellow to dark red). Each
1396 horizontal window size bar includes all 14 flies tested.

1397 **(E)** Mean recurrence times for all 14 flies as a function of window size. The larger circle
1398 corresponds to the population mean while smaller circles correspond to each fly.

1399 **(F)** Maximum recurrence times for all 14 flies as a function of window size. Each point
1400 reflects the bin in which the largest proportion of points displayed recurrence. The larger
1401 circle corresponds to the population mean while smaller circles correspond to each fly.

1402 **(G)** Pathways through the locomotor behavior space for all 14 control flies, produced by
1403 connecting temporally adjacent windows with partially transparent lines.

1404 **(H)** Bar plot comparing the percent of behavior space bins that were visited by at least n
1405 individuals. Bars are color coded from blue to red to correspond to **(C)**.

1406 **(I)** Locomotor behavior space binned to a 64x64 grid. Each bin is colored corresponding
1407 to the number of individuals that visited it.

1408 **(J)** Density maps for all control flies. Each behavior space corresponds to an individual
1409 and is colored by the kernel density estimate generated for that fly's trajectory in
1410 behavior space. Darker color corresponds to a greater density in a given region.

1411 **(K)** Bar plot comparing the percentage of bins in a 64x64 gridded behavior space that
1412 were visited by control flies (purple), optogenetically activated flies (green), or both
1413 (grey).

1414 **(L)** Behavior space colored by the distribution of overlapping and unique bins. Each bin
1415 is colored by the designations in the bar plot.

1416 **(M)** Density maps for optogenetically activated flies. Each behavior space corresponds
1417 to a unique fly ($n = 19$) and is colored by the kernel density estimate generated for that
1418 fly's trajectory through behavior space. Darker colors correspond to a greater density in
1419 a given region.

1420

1421

1422

1423

1424

1425

1426

1427

1428

1429

1430

1431

1432

1433

1434

1435

1436

1437

1438

1439

1440

1441

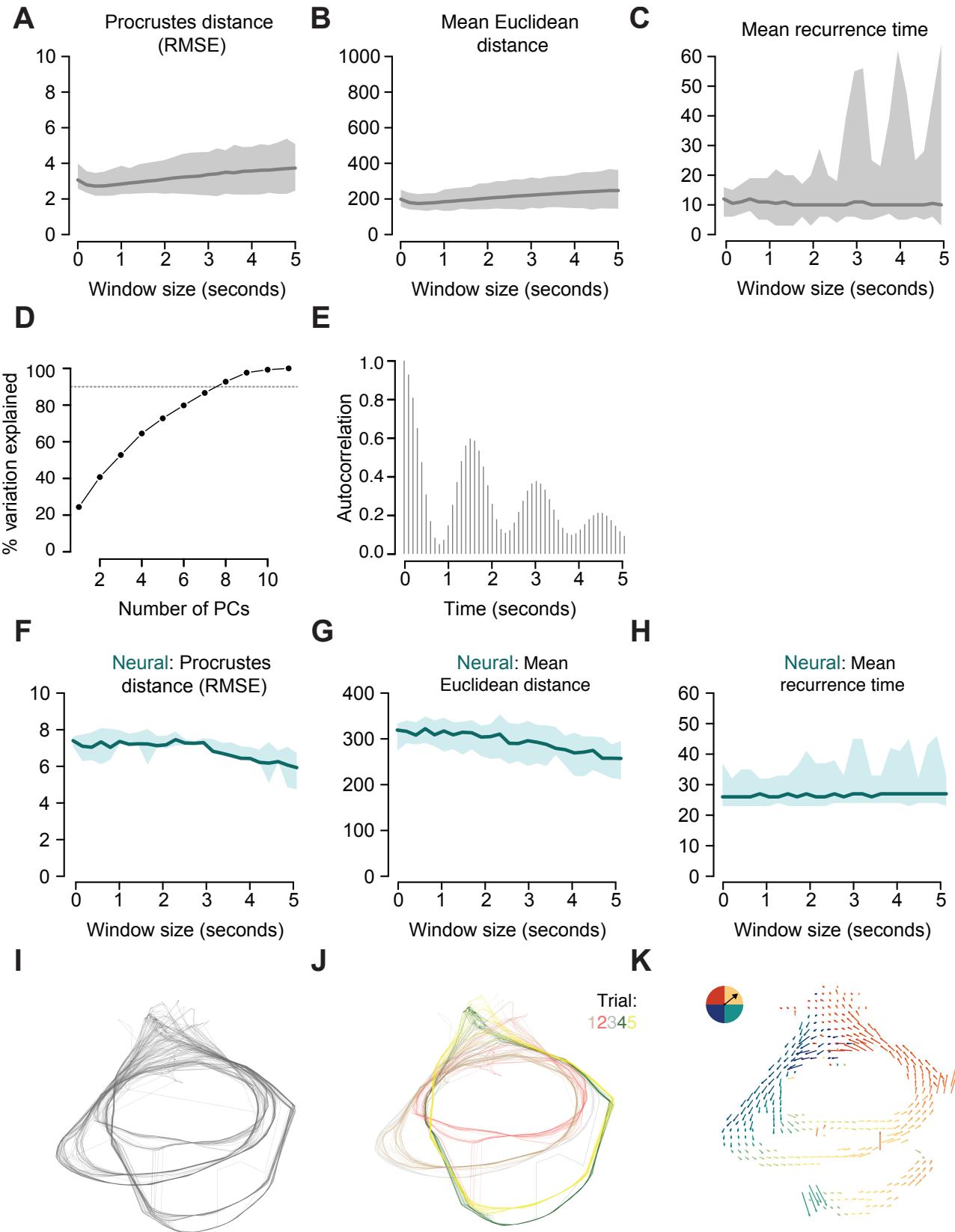
1442

1443

1444

1445

1446



1448 **Figure S3 *Drosophila* larvae behavioral and neural spaces**

1449 **(A-C)** Procrustes distance **(A)**, mean Euclidean distance **(B)**, and mean recurrence time
1450 **(C)** as a function of window size for the crawling *Drosophila* larvae behavior space. The
1451 darker line corresponds to the mean value while the shading reflects standard error.
1452 **(D)** Scatterplot of the percent variation explained by increasing numbers of principal
1453 components representing the 11 input features used. Dashed line corresponds to 90%
1454 variation explained.
1455 **(E)** Example distribution of the autocorrelation of position in larval behavior space over
1456 200 seconds. Position was represented by single value encoding of the 64x64 binned
1457 behavior larval space and then used as input to the autocorrelation calculation.
1458 **(F-H)** Procrustes distance **(F)**, mean Euclidean distance **(BG**, and mean recurrence
1459 time **(H)** as a function of window size for the crawling *Drosophila* larvae neural space. In
1460 this case, the iterative windows procedure was run using fluorescence traces from 7
1461 motor neurons as input. The darker line corresponds to the mean value while the
1462 shading reflects standard error.
1463 **(I)** The full larval *Drosophila* neural space, containing 5 independent trials.
1464 **(J)** The full larval *Drosophila* neural space, colored by individual trials (denoted in
1465 legend in upper right hand corner).
1466 **(K)** Larval *Drosophila* neural space represented as a mean vector field. Color
1467 corresponds to the angle of each vector (reflected by colored circle in the upper left
1468 hand corner).

1469

1470

1471

1472

1473

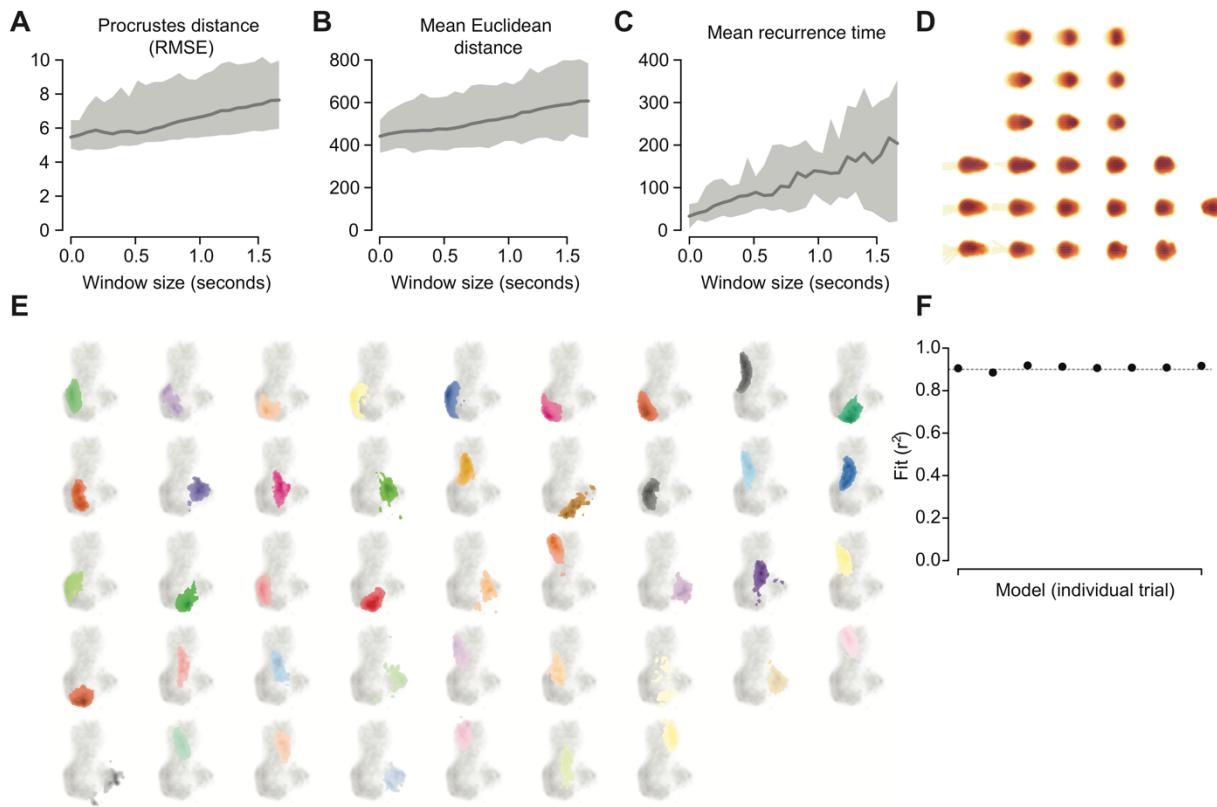
1474

1475

1476

1477

1478



1479

1480 **Figure S4 Mouse 3D pose behavior space**

1481 **(A-C)** Procrustes distance **(A)**, mean Euclidean distance **(B)**, and mean recurrence time **(C)** as a function of window size. The darker line corresponds to the mean value while the shading reflects standard error.

1482 **(D)** Heatmaps representing mean 3D pose as a function of behavior space position (grouped into 25 unique bins). Time points in which each bin was visited were extracted 1483 and then associated with the corresponding moments in the raw 3D imaging data. 1484 These instances were then averaged, producing a mean 3D posture per bin, 1485 represented here as a heatmap (yellow = further from imaging camera; darker red = 1486 closer/higher).

1487 **(E)** The distributions of the full set of behavioral syllables (identified by MoSeq) in 1488 behavior space. A probability density function across behavior space was computed for 1489 each syllable and then plotted in color on top of the full behavior space (in grey; as in 1490 Figure 5K). 1491

1494 (F) Fits (r^2) of generalized linear models comparing input features and behavior space
1495 position for individual trials/mice. The fit of the model using all trials is denoted by the
1496 grey dashed line.

1497

1498

1499

1500

1501

1502

1503

1504

1505

1506

1507

1508

1509

1510

1511

1512

1513

1514

1515

1516

1517

1518

1519

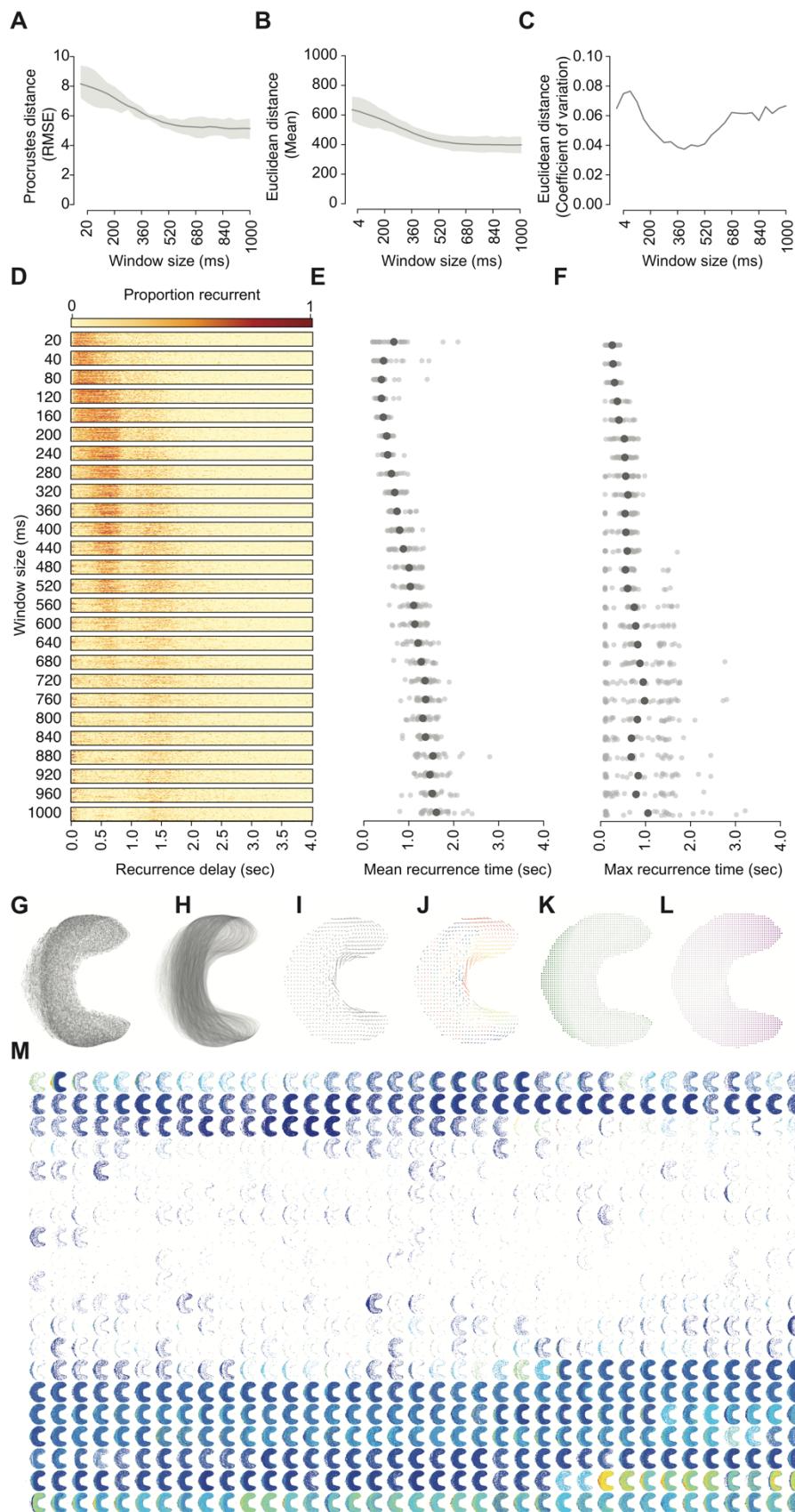
1520

1521

1522

1523

1524



1526 **Figure S5 Mouse locomotor behavior space and its association with MEC activity**

1527 **(A)** Procrustes distance RMSE measures as a function of window size. The darker line
1528 corresponds to the mean value while the shading reflects standard error.

1529 **(B)** Mean intra-point Euclidean distances. The darker line corresponds to the mean
1530 value while the shading reflects standard error.

1531 **(C)** Coefficient of variation of mean intra-point Euclidean distance as a function of
1532 window size.

1533 **(D)** Recurrence plot of behavior spaces taken from 30 random mouse trials. The
1534 proportion of recurrent points given a range of time delays ranging from 0 to 4 seconds
1535 is indicated by the color of the corresponding bins (from light yellow to dark red). Each
1536 horizontal window size bar includes all 30 mice tested.

1537 **(E)** Mean recurrence times as a function of window size. The larger circle corresponds
1538 to the population mean while smaller circles correspond to each mouse trial.

1539 **(F)** Maximum recurrence times as a function of window size. Each point reflects the bin
1540 in which the largest proportion of points displayed recurrence. The larger circle
1541 corresponds to the population mean while smaller circles correspond to each mouse
1542 trial.

1543 **(G)** Mouse locomotor behavior space, each point corresponds to a temporal window.

1544 **(H)** Pathways through mouse locomotor behavior space, produced by connecting
1545 temporally adjacent windows with partially transparent lines.

1546 **(I)** Mouse locomotor behavior space represented as a vector field. Arrow direction and
1547 magnitude correspond to the angle and mean direction taken after visiting each bin.

1548 **(J)** Arrow direction and magnitude correspond to the angle and mean direction taken
1549 after visiting each bin. Arrows are colored by the degree of the direction vector
1550 (corresponding to circle in upper left-hand corner).

1551 **(K)** Distribution of translational velocity across mouse behavior space (darker green
1552 corresponds to higher values).

1553 **(L)** Distribution of angular velocity across mouse behavior space (darker purple
1554 corresponds to higher values).

1555 (M) 2-d tuning curves of neural activity (1 second bins) for all MEC neurons. The
1556 ordering of tuning curve position was determined by hierarchical clustering (i.e. more
1557 similar tuning curves are placed adjacent to each other).

1558

1559

1560

1561

1562

1563

1564

1565

1566

1567

1568

1569

1570

1571

1572

1573

1574

1575

1576

1577

1578

1579

1580

1581

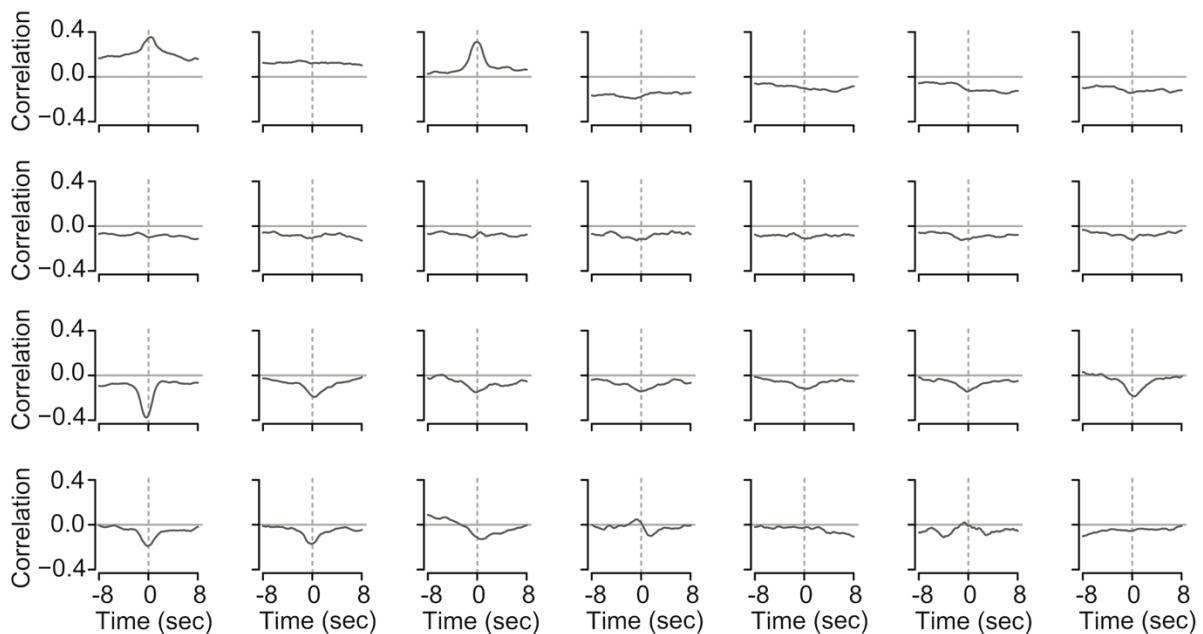
1582

1583

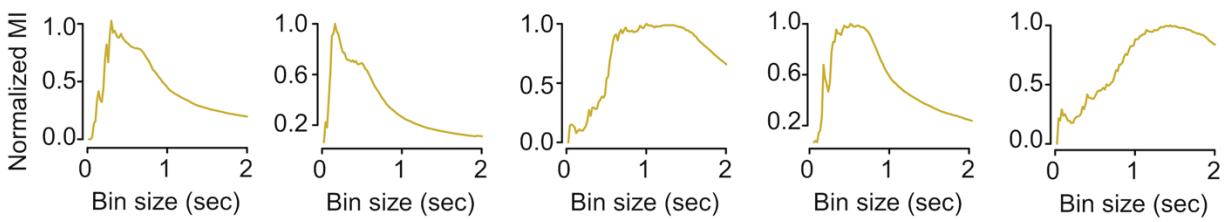
1584

1585

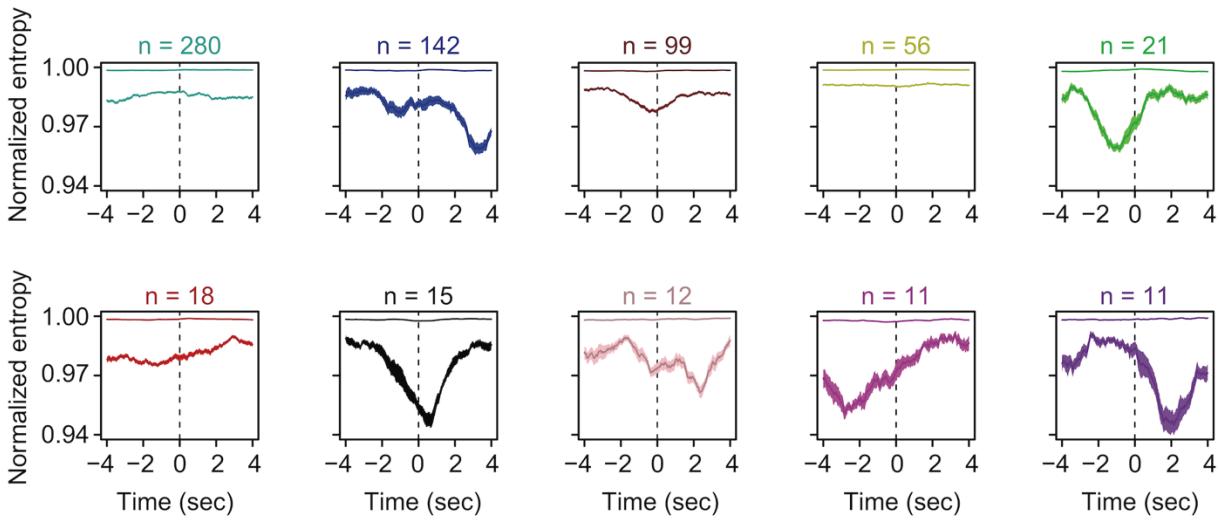
A



B



C



1586

1587

1588

1589 **Figure S6: Variation in temporal and rate-based associations between MEC
1590 neurons and behavior**

1591 **(A)** Example cross-correlations between behavior space position and MEC neuron
1592 activity over a 16 second time window. The dotted grey line corresponds to time point
1593 zero. The solid grey line reflects a correlation coefficient of zero.

1594 **(B)** Example mutual information (MI) tuning curves as a function of binning size of
1595 neural activity. For comparison purposes, MI is presented as a normalized value in
1596 which each distribution has been divided by its maximum value.

1597 **(C)** Mean entropy distributions of ten clusters identified using dynamic tree trimming.
1598 Plotted are the distributions for the 75% (top) and 99% (bottom) cutoffs. Mean and
1599 standard error are plotted as well as sample sizes, as in Figure 7E.

1600

1601

1602

1603

1604

1605

1606

1607

1608

1609

1610

1611

1612

1613

1614

1615

1616

1617

1618

1619