Title: Nucleus accumbens neurons encode the kinematics of reward approach locomotion

David Levcik[1,4#], Adam H. Sugi[1,2,3#], Marcelo Aguilar-Rivera[5#], José A. Pochapski[1,2], Gabriel Baltazar[1,2,3], Laura N. Pulido[1,2], Cyrus Villas Boas[1], Romulo Fuentes-Flores[6], Saleem M. Nicola[7,8], Claudio Da Cunha[1,2,3*]

1. Laboratório de Fisiologia e Farmacologia do Sistema Nervoso Central, Universidade Federal do Paraná, 81531-980, Curitiba, Brazil.

2. Department of Pharmacology, Universidade Federal do Paraná, Curitiba, Brazil

3. Department of Biochemistry, Universidade Federal do Paraná, Curitiba Brazil

4. Institute of Physiology of the Czech Academy of Sciences, Videnska 1083, 142 20, Prague, Czech Republic

5. Department of Bioengineering, University of California, 9500 Gilman Drive MC 0412, La Jolla, San Diego, 92093, USA

6. Departamento de Neurociencia, Facultad de Medicina, Universidad de Chile, Av. Independencia 1027, Independencia 8380453, Santiago, Chile

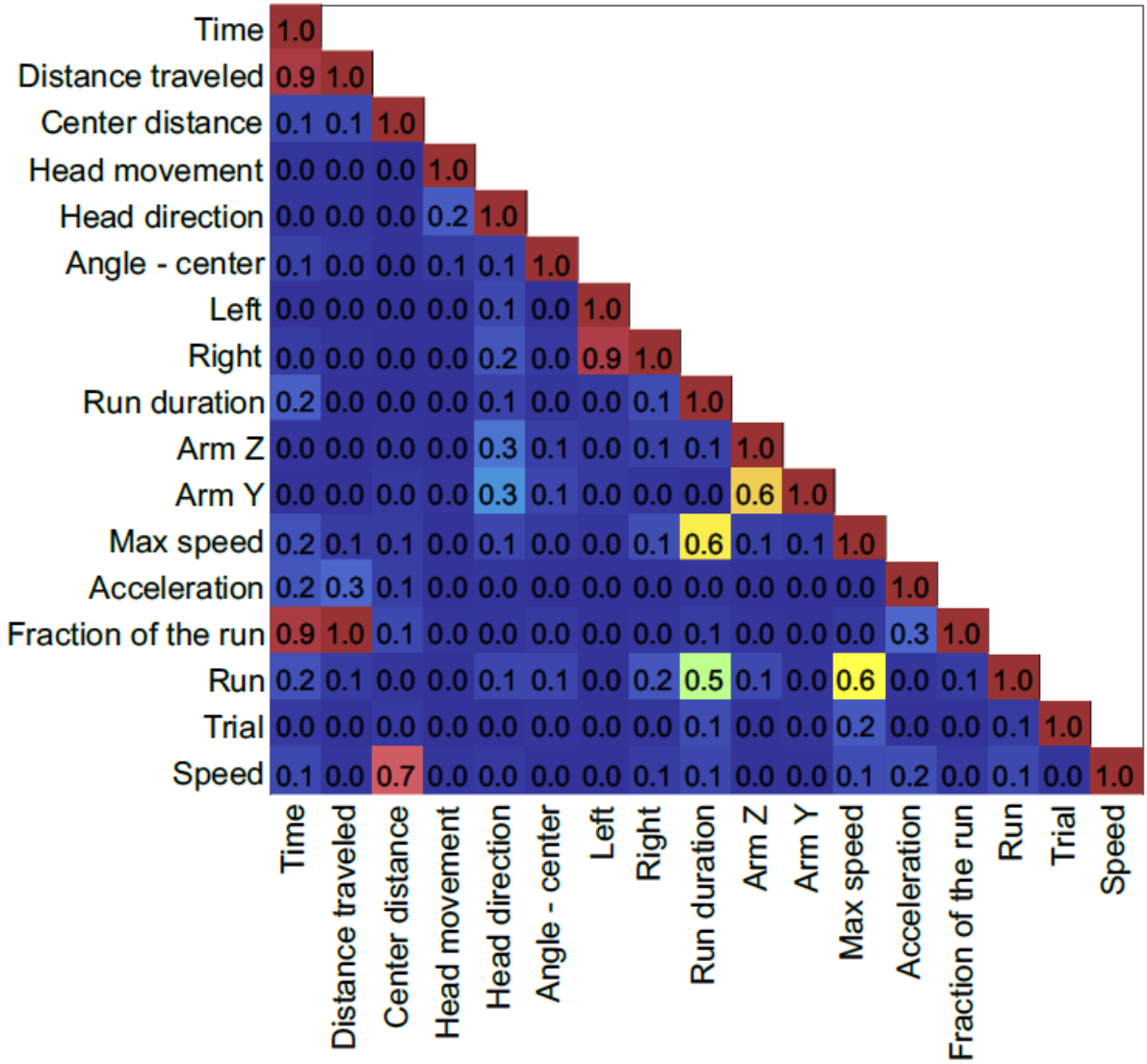
7. Department of Neuroscience, Albert Einstein College of Medicine, 1300 Morris Park Ave, Bronx, New York, 10461 USA

8. Department of Psychiatry, Albert Einstein College of Medicine, New York, USA

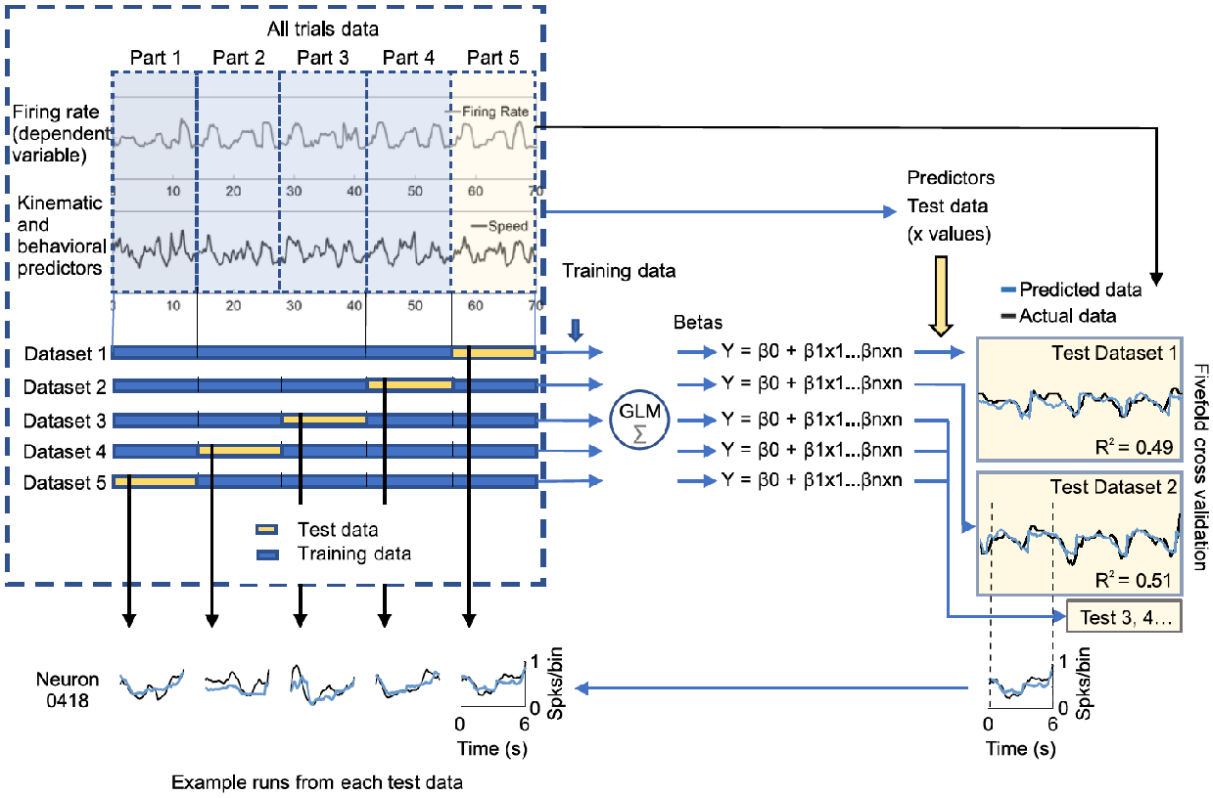
The authors presented equally important contributions.

* Corresponding author email address: dacunha.claudio@gmail.com

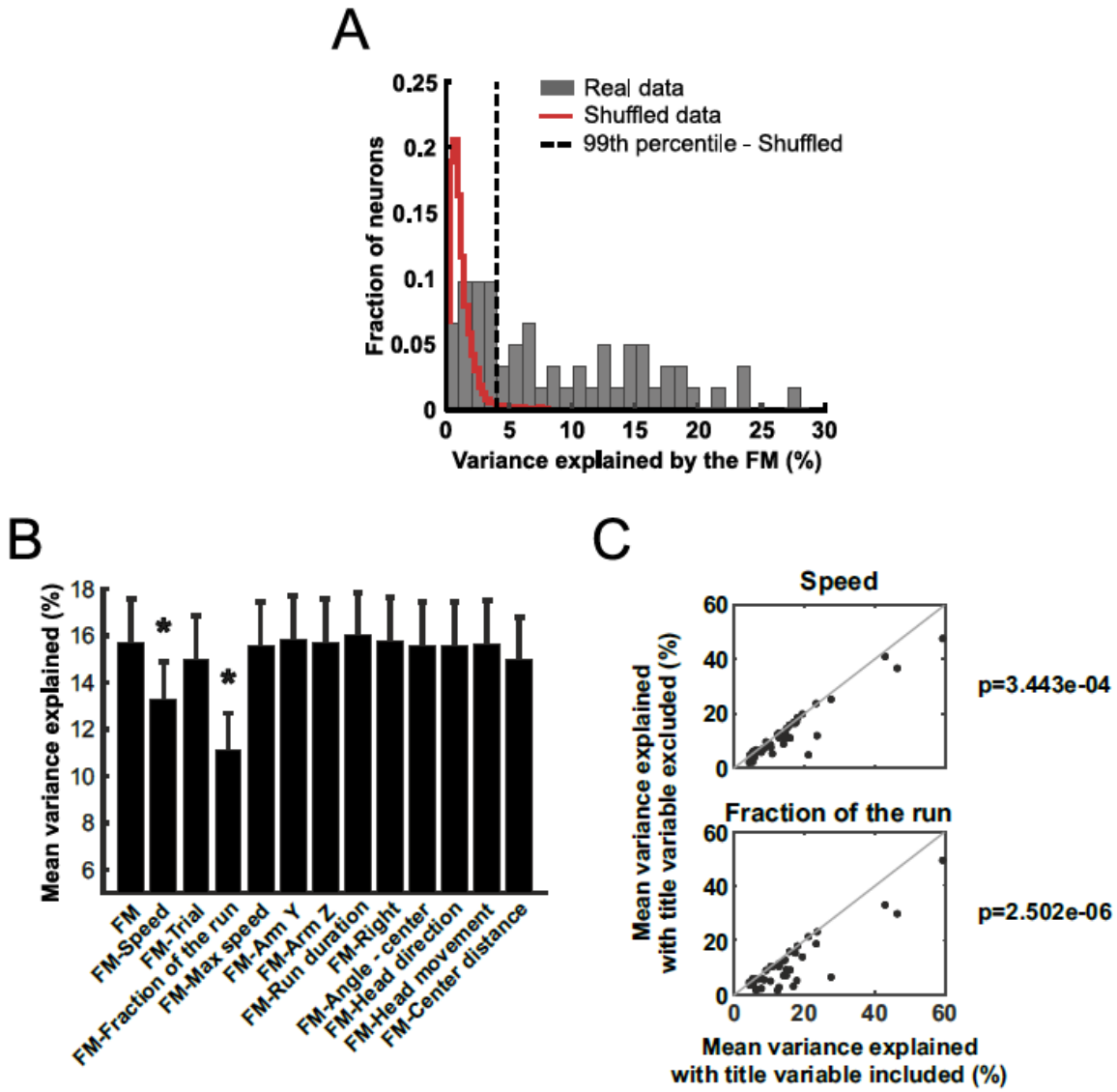
Supplemental figures



Supplemental Figure S1: Correlation table for predictor variables considered for use in the GLM encoding model. For the pairs with high correlation ($r > 0.8$), one of the two variables was chosen to be included in the model.

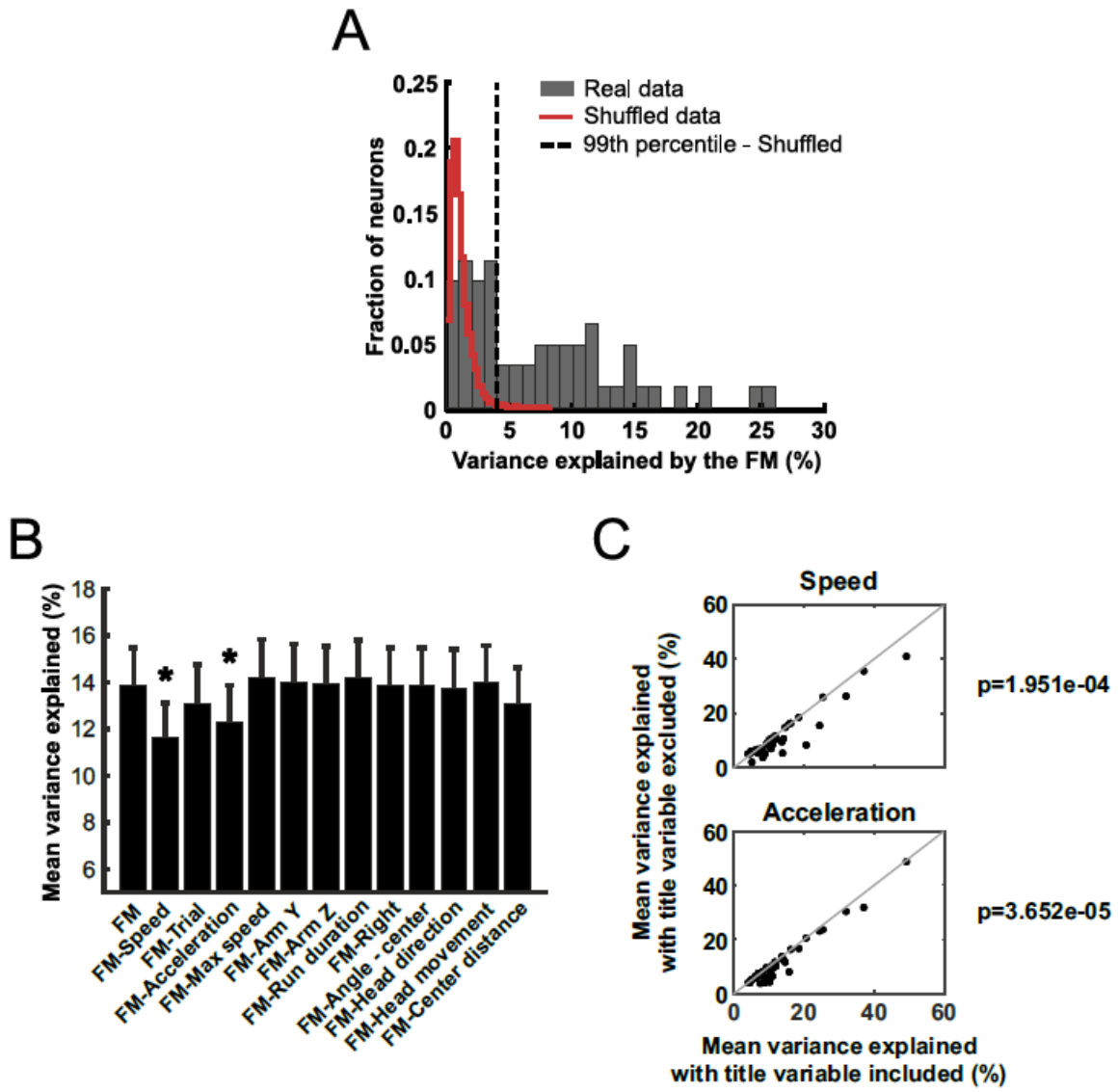


Supplemental Figure S2: Fivefold cross-validation method for evaluating the GLM encoding model. Data (firing rate and predictor variables) from a whole session were divided into five parts (datasets). Each dataset had the same number of runs (selected randomly). Four of the five datasets were used as training datasets to calculate the β values for each predictor. Then, using the predictors data from the remaining dataset (the test dataset) as input to the encoding model, predicted firing rates (Y) were calculated. The predicted firing rates were compared to the actual firing rates from the test data set. The variance explained by the GLM was calculated as the r^2 of the correlation between the predicted and actual data. This process was repeated 5 times, with each dataset used only once as test data. The variance explained by the model was the mean of the 5 r^2 values (r^2_{mean}).

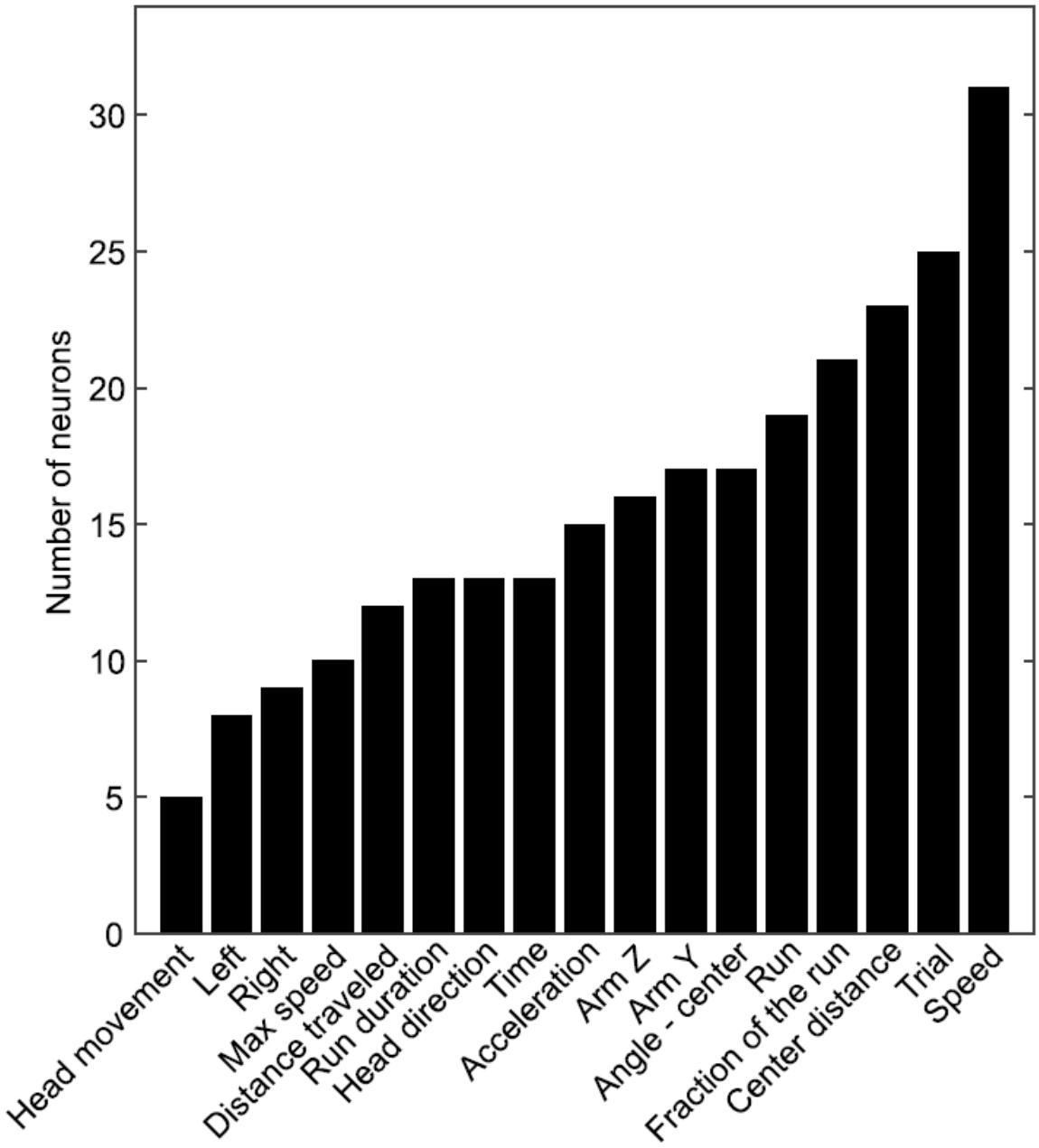


Supplemental Figure S3: GLMs without the *acceleration* variable. **(A)** Distribution of the variance explained by a GLM full model (FM) that used all variables shown in (B) as predictors. Red line delimits the distribution of the variance explained by a GLM calculated with shuffled neuronal data (100 shuffled models per neuron). Neurons to the right of the dashed line (99th percentile of the shuffled models distribution) were classified as significantly explained by the GLM (63% of recorded neurons). In this and subsequent graphs, variance explained is depicted as a percent (100 x r^2_{mean}). **(B)** Percentage of mean variance explained (mean ± SEM of r^2_{mean} across neurons) of the firing rate of neurons with activity significantly explained by the FM. The mean variance explained by the FM is shown in the first columns. The other columns show the variance explained by the FM excluding the indicated variable. * $p < 0.01$, paired t-test (after Bonferroni correction) comparing the indicated model to the FM. **(C)** Significant comparisons of mean variance explained by the FM (X-axis) and by the FM excluding the indicated variable (Y-axis) for all neurons (black

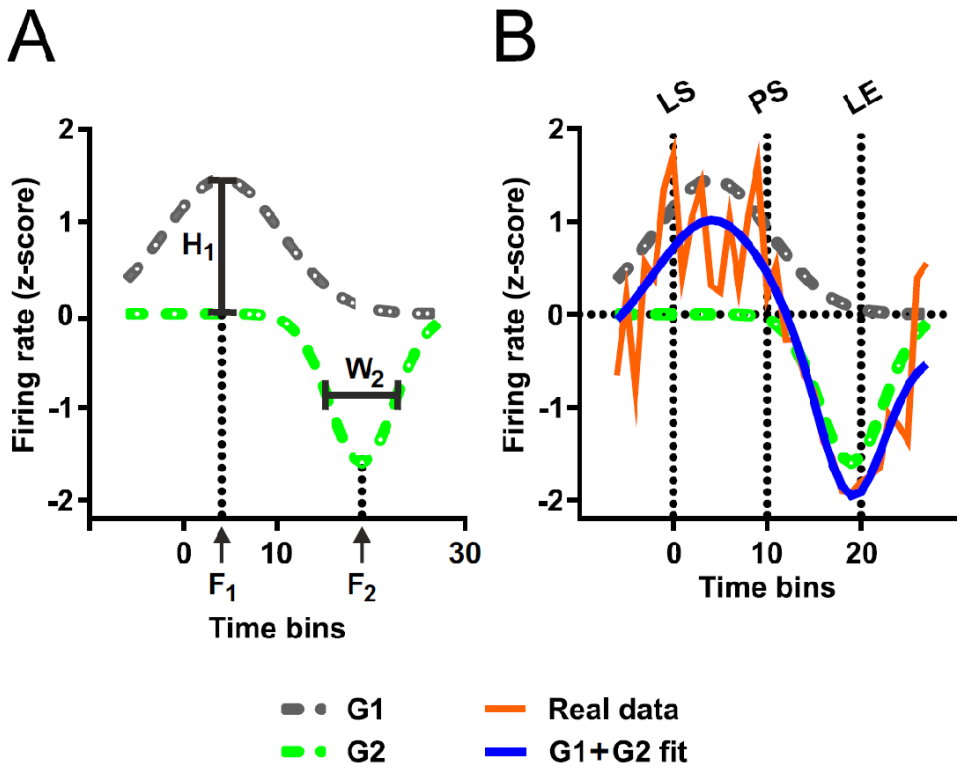
dots) explained significantly by the FM. The p-values represent comparison of the mean variance explained by the FM and a model without the particular predictor.



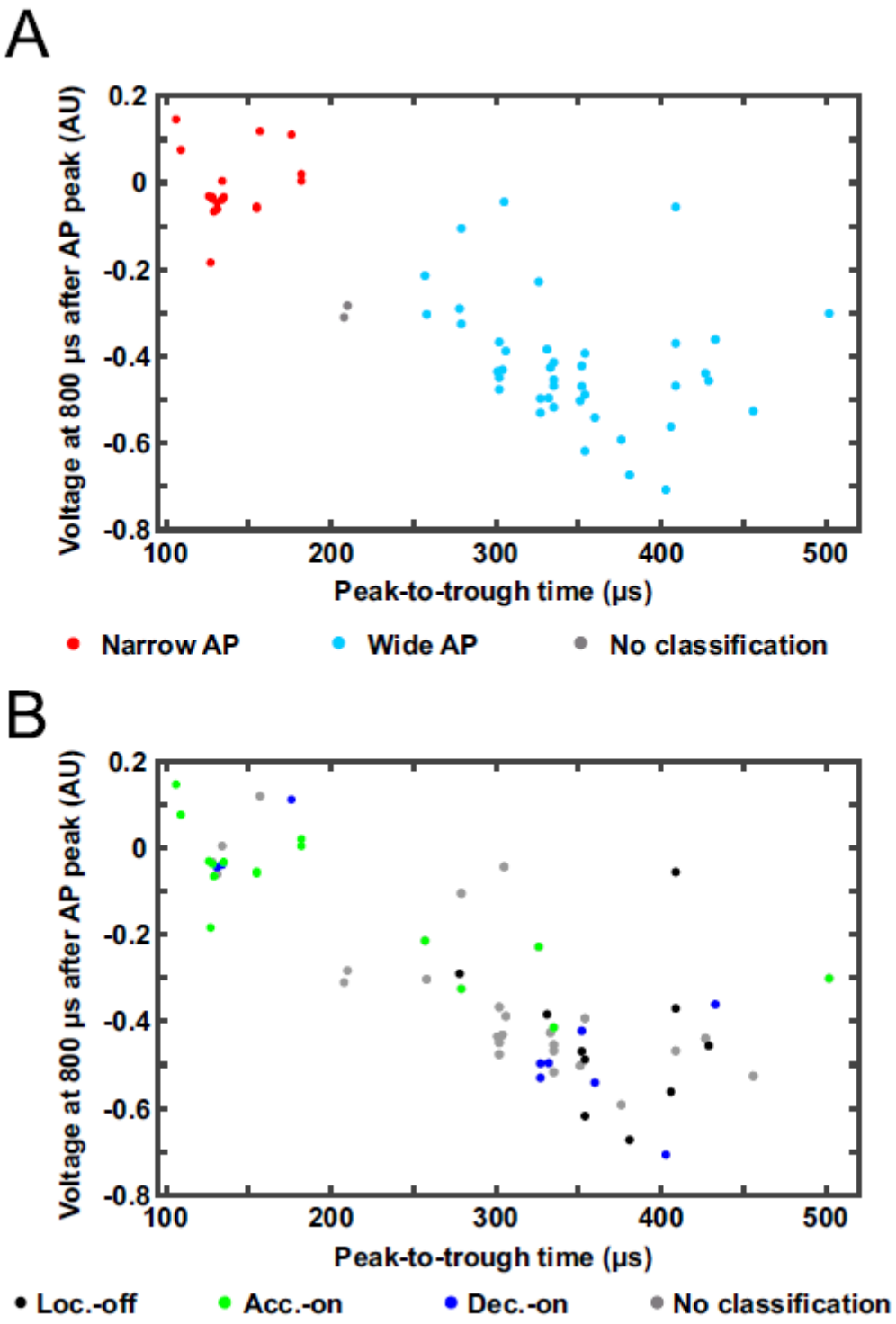
Supplemental Figure S4: GLMs without the *fraction of the run* variable. Panels are as described for Supplemental Figure S3, except that these GLM predicted the changes in the firing rate of 58% of the neurons.



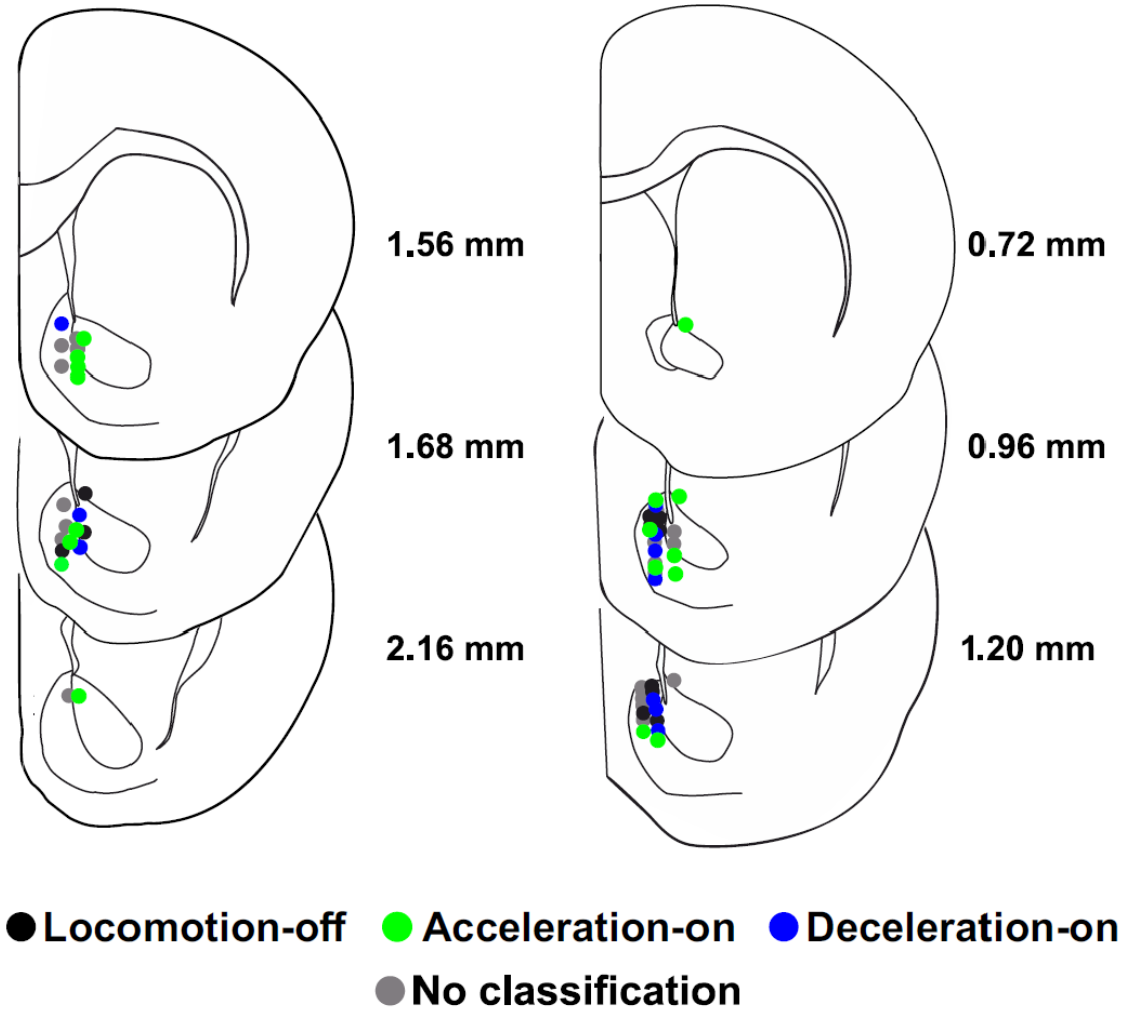
Supplemental Figure S5: Histogram summarizing the significant variable predictors for all the recorded neurons as selected by the LASSO method. This technique selected the variables speed and fraction of the run as relevant for approximately half of the neurons.



Supplemental Figure S6: The sum of two Gaussians (2G) model. A non-linear model that consists of the sum of two Gaussian curves was used to model the firing rate during the approach run. (A) The two modeled Gaussian curves. (B) The same two curves as well as the mean firing rate data (z-score) and the sum of the two Gaussian curves. The only predictor used in this model is the fraction of time. The period of time from the locomotion start (LS) to the peak of speed (PS) was divided into 10 bins, and the time from the peak of speed and the locomotion end (LE) was also divided into 10 bins. These bins were ~150 ms each. The 7 bins preceding locomotion start (LS) and following locomotion end (LE) were 150 ms each. Changes in z-scored firing rate as function of normalized time ($Y_{(t)}$) was modeled as $Y_{(t)} = G1_{(t)} + G2_{(t)} + BL$, where $G1$ was the equation for the peak part of the curve, $G2$ was the equation for the valley part of the curve and BL was the baseline activity. The Gaussian equations ($G_{(t)}$) used were: $G1_{(t)} = H1e^{-\pi(\frac{X_{(t)}-F1}{W1})^2}$, $G2_{(t)} = -H2e^{-\pi(\frac{X_{(t)}-F2}{W2})^2}$, where H, F, and W are constants with the following meaning: H represents the amplitude of the Gaussian curve, which reflects the height ($H1$) and depth ($H2$) of the activity peak and the valley, respectively. F is the bin where the Gaussian is centered. Therefore, $F1$ is the bin in which the neuron is most active, and $F2$ is the bin in which the neuron is least active. W is the standard deviation of each Gaussian, which reflects the durations of the activation peak ($W1$) and valley ($W2$).



Supplemental Figure S7: Separation of recorded cells (A) into two clusters based on their waveform characteristics (narrow action potential (AP) - red, wide AP - turquoise). Narrow AP neurons are putative fast-spiking interneurons, whereas some wide AP neurons are putative spiny projection neurons. The same data is plotted (B) with the neurons' LOff (black), /AOn (green), and /DOn (blue) classifications.



Supplemental Figure S8: Placement of tips of recording electrodes. Neurons located outside of the nucleus accumbens were excluded from the analysis.

Supplemental tables

	Rat #04	Rat #08	Rat #09	Rat #11	Rat #19	**Total**	**%**
Number of recorded neurons	18	4	25	11	4	**62**	**100**
Number of neurons with signifcant GLM result	16	1	12	9	1	**39**	**63**
Number of locomotion-off cells (LOff)	1	2	7	1	0	**11**	**18**
Number of acceleration-on cells (AOn)	10	0	4	3	0	**17**	**27**
Number of deceleration-on cells (DOn)	2	1	4	3	0	**10**	**16**
Number of unclassified cells	5	1	10	4	4	**24**	**39**
Total	18	4	25	11	4	**62**	**100**

Supplemental Table S1: Summary of recorded cells in individual rats.

	Variable type	In GLM	Unit or value	Description
Speed	Continuous	Yes	cm/s	Movement speed of the animal during the task
Fraction of the run*	Continuous	Yes	from 0 to 1	Fraction of the total distance traveled during the reward approach
Acceleration	Continuous	Yes	cm^2/s	Acceleration of the animal during the task
Max speed	Constant	Yes	cm/s	Maximum speed reached during the run
Arm Y	Dummy	Yes	1 or 0	1 if the animal is approaching reward arm Y
Arm Z	Dummy	Yes	1 or 0	1 if the animal is approaching reward arm Z
Right	Dummy	Yes	1 or 0	1 if the upcoming turn direction was right
Left	Dummy	No	1 or 0	1 if the upcoming turn direction was left
Run duration	Constant	Yes	s	Total duration of the run
Trial	Constant	Yes	from 1 - 12	Number of the trial in a session of 12 trials
Run	Constant	No	from 1 - 36	Number of the run in a session of 12 trials
Angle - center	Continuous	Yes	degrees (0 - 360)	Angle formed by the center of the maze and the animals location
Head direction	Continuous	Yes	degrees (0 - 360)	Animal's head direction
Head movement	Continuous	Yes	degrees (-360 to +360)	Changes in head direction (counter-clockwise or clockwise)
Distance traveled*	Continuous	No	cm	Distance traveled since the locomotion onset of each run
Time*	Continuous	No	s	Time since the locomotion onset of each run
Center distance	Continuous	Yes	cm	Distance between the animal location and the center of the maze

*The values of these variables are reseted (to zero) at the beginning of each new run

Supplemental Table S2: Description of individual variables considered for use as predictors in the GLM encoding model. Variable types include Continuous (continuous variables that could change throughout each run), Constant (continuous variables that were constant throughout each run, but could change across runs), and Dummy (0 or 1). Variables marked Yes under In GLM were included in the GLMs; those marked No were not. “Arm Y” and “Arm Z” were both 0 in cases where the animal entered Arm X.

Rat	Rat #04	Rat #08	Rat #09	Rat #11	Rat #19
reference memory errors (%)	0.883	0.935	7.631	11.864	8.911
working memory errors (%)	0.442	0.000	9.237	21.469	28.713
error runs per session	0.462	0.333	7.000	14.857	19.000
reward run latency (sec)	2.920	3.256	3.106	3.282	3.017

Supplemental Table S3: Summary of behavioral performance of individual rats during the recording sessions. Data are shown as means.

Supplemental Experimental Procedures

Surgery. The rats were anesthetized with isoflurane (2%), placed in a stereotaxic frame (David Kopf, USA), and a customized microdrive with 2 x 4 stereotrode arrays (17 μm Ni-Chrome wire with insulation; A-M Systems, USA) was chronically implanted unilaterally above NAc shell at the following coordinates: AP +1.2 to +2.2 mm from bregma, ML +0.6 to +1.0 mm from midline, DV -5.0 mm from dura. The coordinates were determined according to the Paxinos and Watson Atlas (Paxinos G and Watson C, 2007). Six anchoring stainless steel screws were placed in the skull to attach the implant using acrylic dental cement. One of the screws also served

as a ground. After surgery, enrofloxacin (20 mg/kg; i.m.) was injected, neomycin gel was applied to the tissue around the implant, and ibuprofen (50 mg/500 ml) was added to the drinking water for the three following days. The rats were allowed to recover for seven days after the surgical procedures.

Electrophysiological recording. After one week of recovery, the stereotrodes were lowered in 40 μm increments (up to 160 μm/day) until single-unit activity was reliably detected and isolated. All stereotrodes were previously gold-plated with a nanoZ impedance tester (Multi Channel Systems, USA) to impedances between 100 - 300 kOhms. Two light-emitting diodes attached to the 16-channel headstage (20 x gain) signaled the position of the animal and allowed video tracking (30 frames/s) with the Cineplex system (Plexon, USA). Neural activity was amplified (1000 x), filtered (300 - 6000 Hz) and sampled at 40 kHz with the OmniPlex D Neural Data Acquisition system (Plexon, USA) to acquire and record single unit activity.

Spike sorting. Recorded spikes were manually sorted offline into clear clusters of putative cells with Offline Sorter software (Plexon, USA). To be accepted for spike sorting, spike amplitudes had to be at least three standard deviations higher than the background activity. Clusters were based on waveform properties identified through principal component analysis. Inter-spike interval histograms were calculated, and those clusters with similar waveform shape that showed a clearly recognizable refractory period (> 1.5 ms) were considered as originating from a single neuron.

LASSO. Beacause some predictors are highly correlated with one another, we used the LASSO (MATLAB R2021b) technique for feature selection or regularization. We identified predictive and redundant variables among the 17 independent variables by using a 10-fold cross-validation. We applied this technique for each neuron. First, we find the smallest Lambda value from the LASSO 10-fold cross validation plot that had the smallest mean squared error (MSE) of the models. We used the first minimum since some neurons showed an asymptotic curve to the X-axis (Lambda values) on such Lambda plots. Second, we identified the variables for the sparsest model using a Lambda value at one standard error to the left of the minimum MSE. This Lambda provides the closest model to the real neural firing rate.

Neuronal classification based on electrophysiological properties. The average action potential waveform was normalized for each neuron by the maximum voltage value corresponding to the peak of the inverted waveform resulting in an arbitrary unit (AU) of voltage. The normalized averaged waveform for all the single neurons were aligned at the peak. We plotted the peak-to-trough time vs. voltage values (AU) at 800 μs after the peak as at this time the waveforms of individual neurons showed the best separation. We used KS-means to locate the peak-to-trough time boundary between the potential different phenotype (MSNs, fast-spiking interneurons) groups corresponding, respectively, to wide and narrow action potential waveform clusters. To avoid overlap between points with similar values we added a random number between 0 to 10 to the peak to trough variable. Then the waveforms in the wide action potential cluster were sorted based on the proportion of ISIs that a single neuron shows over 2 s. All the single neurons with a proportion of ISIs over 2 s greater than 0.4 were considered as putative MSNs, while the remaining cells with a wide action potential were considered as not classified. The cells in the narrow action potential cluster were classified as interneurons.

Histology. Electrolytic microlesions (12 μA negative current between one wire of each stereotrode and the ground, 30 s duration) were made at the tips of all stereotrodes using the nanoZ (White Matter LLC, USA). The next day, the rats were deeply anesthetized with ketamine (75 mg/kg; i.p.) and xylazine (10 mg/kg; i.p.), and then transcardially perfused with 300 ml saline followed by 300 ml 4% paraformaldehyde. The brains were removed and 40 μm slices were processed for cresyl violet staining. The final locations of the electrodes were verified.

Supplemental References

Paxinos G, Watson C (2007) The rat brain in stereotaxic coordinates. London: Elsevier Academic.