Encoding of spatial location by frontal orienting field neurons

Liujunli Li[1,2], Ce Ma[1,2], Jingjie Li[1,2,4], Jeffrey C. Erlich[1,2,3,4*]

1 NYU-ECNU Institute of Brain and Cognitive Science at NYU Shanghai, China
2 NYU Shanghai, Shanghai, China
3 Shanghai Key Laboratory of Brain Functional Genomics (Ministry of Education), East China Normal University, Shanghai, China
4 Sainsbury Wellcome Centre, University College London, London, UK

*** j.erlich@ucl.ac.uk**

Abstract

The neural mechanisms of motor planning have been extensively studied in rodents. Frontal cortical areas seem to encode upcoming choice, but limitations of typical tasks make it challenging to determine whether activity represents a planned movement direction in a self-centered reference frame or a goal position in a world-centered reference frame. Here, we trained rats to make delayed visually-guided orienting movements to six different directions, with four different target positions for each direction, which allowed us to disentangle position versus direction tuning in neural activity. We recorded single unit activity from the rat frontal orienting field (FOF) in secondary motor cortex, a region involved in planning orienting movements. We found neurons in the FOF that were tuned for specific directions of movement and also neurons that fired while animals were at specific port positions. Interestingly, after the visual cue onset, movement direction tuning emerged earlier than target position tuning. At the level of individual neurons, the current head position modulated the planned movement direction as a gain field. These results suggest the FOF participates not only in the motoric processes of sensorimotor behavior, which could happen strictly in egocentric coordinates, but in more complex aspects like reference frame transformation or maintaining a stable model of the world during movements.

Introduction

We use multiple reference frames to represent the world. For example, as you plan a movement to reach for you morning coffee, the arm region of motor cortex may represent the goal in an arm-centered reference frame and your frontal eye field

represents the goal in an eye-centered reference frame. Other areas of your brain may represent the cup relative to the room or table or the milk or sugar. During these kinds of sensorimotor behaviors, sensory information initially is represented in sensor reference frames and motor commands are finally represented in muscle reference frames. Since those sensors and effectors are embodied, we can think of these representations as being in egocentric (or 'self-centered') reference frames, i.e. they are reference frames that move around with the subject as they move through the world. However, our lived experience is in a allocentric (or 'world-centered') reference frame: We feel as if we move around and make decisions in a stable world. Moreover, allocentric representations are found all over the brain (Hafting et al., 2005, O'Keefe and Nadel, 1978, Taube et al., 1990, Wilber et al., 2014). Thus, a full understanding of the neurobiology of motor-planning needs to address the question of where and how these reference frame transformations take place (Andersen et al., 1990, 1985, Andersen and Mountcastle, 1983, Cohen and Andersen, 2002).

The neural mechanisms of motor planning in rodents have been extensively studied in two-alternative forced choice (2AFC) and go-nogo tasks (Chen et al., 2017, Erlich et al., 2011, Sul et al., 2011). Using these tasks, detailed circuit mechanisms have been mapped (Guo et al., 2017, 2014, Li et al., 2015) and models of neural computation have been developed (Inagaki et al., 2019, Kopec et al., 2015, Li et al., 2016) to explain neural mechanisms underlying the preparation and execution of movements in the time scale of seconds. Converging evidence has implicated the frontal orienting field (FOF), a subregion of the secondary motor cortex (M2), as a cortical substrate for planning orienting movements (Erlich et al., 2011, Hanks et al., 2015, Olson et al., 2020), especially when those plans require flexible sensorimotor processes (Erlich et al., 2015, Siniscalchi et al., 2016, Zhu et al., 2021). FOF receives inputs from numerous cortical and thalamic sources, including the posterior parietal cortex (Reep et al., 1994) and the retrosplenial cortex (Yamawaki et al., 2016), both of which exhibit egocentric as well as allocentric spatial representations (Wang et al., 2020). One would naturally ask: does FOF represent the action goals in an egocentric or allocentric reference frame? Given the limitations of typical behavioral paradigms for investigating motor planning, the answer is largely unknown. If FOF is closer in the motor-planning circuit to the effector, we would expect its representations to be egocentric. If FOF is further upstream in the circuit, we would expect its representations to be allocentric or see a mixture of representations. We previously speculated that FOF might be a rodent analogue of the primate frontal eye field (FEF; Erlich et al., 2011). Since the FEF is generally believed to encode space in retinal or gaze centered coordinates (Schall, 2009), our main hypothesis was that the rat FOF would encode movement in an egocentric frame.

To test the reference frame of action plan representation in the FOF, we designed a multi-directional, multi-positional orienting task that could distinguish allocentric versus egocentric reference frames. Consistent with our hypothesis, we found egocentric movement direction related activity during the delay. In contrast with

our expectations, allocentric target positions were also encoded in the FOF. Interestingly, the position encoding only represented the *current* position of the animals, rather than participate in planning. These two reference frames were multiplexed at the single-neuron level: task-related activity was best described as a gain-field modulation of movement direction tuning by position. The presence of both ego and allocentric representations suggests that FOF participates in reference frame shifts (Caruso et al., 2018, Cassanello and Ferrera, 2007, Pouget and Sejnowski, 1997) or contributes to error monitoring (Sommer and Wurtz, 2008). Our results are largely consistent with the neurophysiological observations made during motor planning in primates, suggesting shared circuit mechanism for spatial cognition and sensorimotor computation across species, despite the substantial differences in anatomical structures (Wise, 2008), sensory systems (i.e. differences in visual acuity) and behavioral output (saccades vs. orienting) (Ebbesen et al., 2018).

Results

We trained rats to perform visually-guided orienting movements to 6 directions, with each direction starting from 4 different start positions (Figure 1, Figure S1). This design allowed us to dissociate neural activity associated with movement direction from the potential influence of start and target position. In other words, by having multiple start positions for each direction and vice-versa, we could distinguish between self-centered and world-centered movement plans.

Each trial in the task began with illuminating a 'start port', randomly chosen from one of the 7 operant ports, with both yellow and blue LEDs. Rats fixated in the start port until a go sound. The start port LEDs extinguished at the start of fixation, and the target port was illuminated with blue LEDs shortly after fixation onset (0 - 0.29 s delay). For trials that started in the center, one of the six remaining operant ports was chosen at random as the target. For the other start positions, one of three adjacent ports was chosen as the target (Figure S1). Throughout the paper we indicate directions (and direction tuning) with a blue-red colormap where leftward directions are blue and rightward are red. Downward directions are darker and upward directions are lighter (Figure 1C). For positions we use a green-orange colormap with green associated with left-side ports and orange with right-side ports. Lower ports are darker and higher port are lighter (Figure 1D).

After the go sound, rats left the start port and moved to the target port. The target port LEDs extinguished once the animal arrived at the target port (or poked in another port in error). If the rat poked in the correct port, the water delivery port LED illuminated, a "correct" sound was played, and reward could be collected at the reward port. If the rat poked into the wrong port, an "error" sound was played and there was no reward (Figure 1A&B). Animals kept still during the fixation period (Figure 1E; for more rigorous treatment for video analysis, see later sections). A trial was considered incomplete if the animal did not poke into any port after the start port within 15 seconds. Unless otherwise specified, fixation

violations or incomplete trials were not included in analyses. Rats performed 318.89 ± 7.76 (mean ± s.e.) trials in each 1.5 hour recording session (n = 104 sessions). As expected from a visually-guided task, performance was good (% Correct = 94.05 ± 0.53%, mean ± s.e., n=104 sessions; FIGURE S2A).

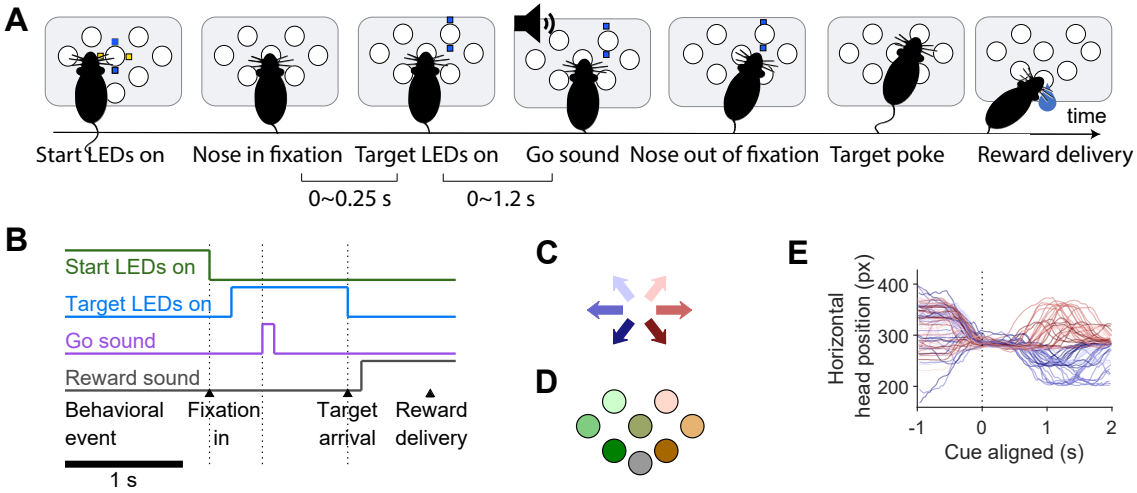

Figure 1. A visually-guided multi-directional orienting task in rats. **A.** Schematic of the task. Each trial began with the onset of a pair of blue and yellow LED, cuing the rat to nose poke into the start port. The start LED turned off upon arrival in the start port, and after a short delay, a blue LED illuminated, indicating the target port. After a go sound, the rat left the start port and poked into the target port. Water reward was delivered for correctly performed trials. **B.** Timeline of a trial in a typical session. **C.** The color scheme of the 6 movement directions. **D.** The color scheme of the 7 port positions. **E.** An example of head position tracking data extracted from video using DeepLabCut. Each line is the horizontal head coordinate during a trial, in unit of pixels from video frames. Only trials starting from the central port are shown. During the fixation period the head position does not predict their upcoming movement.

FOF neurons encode both egocentric movement directions and allocentric head positions.

Consistent with previous findings, there were neurons selective to movement direction during the planning phase (FIGURE 2A) and the execution phase (FIGURE 2D). When firing rates were conditioned on both direction and target position, direction tuning of the three example neurons in the left column seemed to be modulated by target position (FIGURE 2A-I). Surprisingly, we also found neurons that were tuned to the position of the animal (FIGURE 2J-L) as well as neurons that seems to code the conjunction of position and direction (FIGURE 2M-R). To quantify the relative strength of tuning, we compared generalized linear models (GLMs) of each neurons activity fit to each of the task variables. We fit spike counts in a 500 ms window to 3 GLMs, where the independent variable was

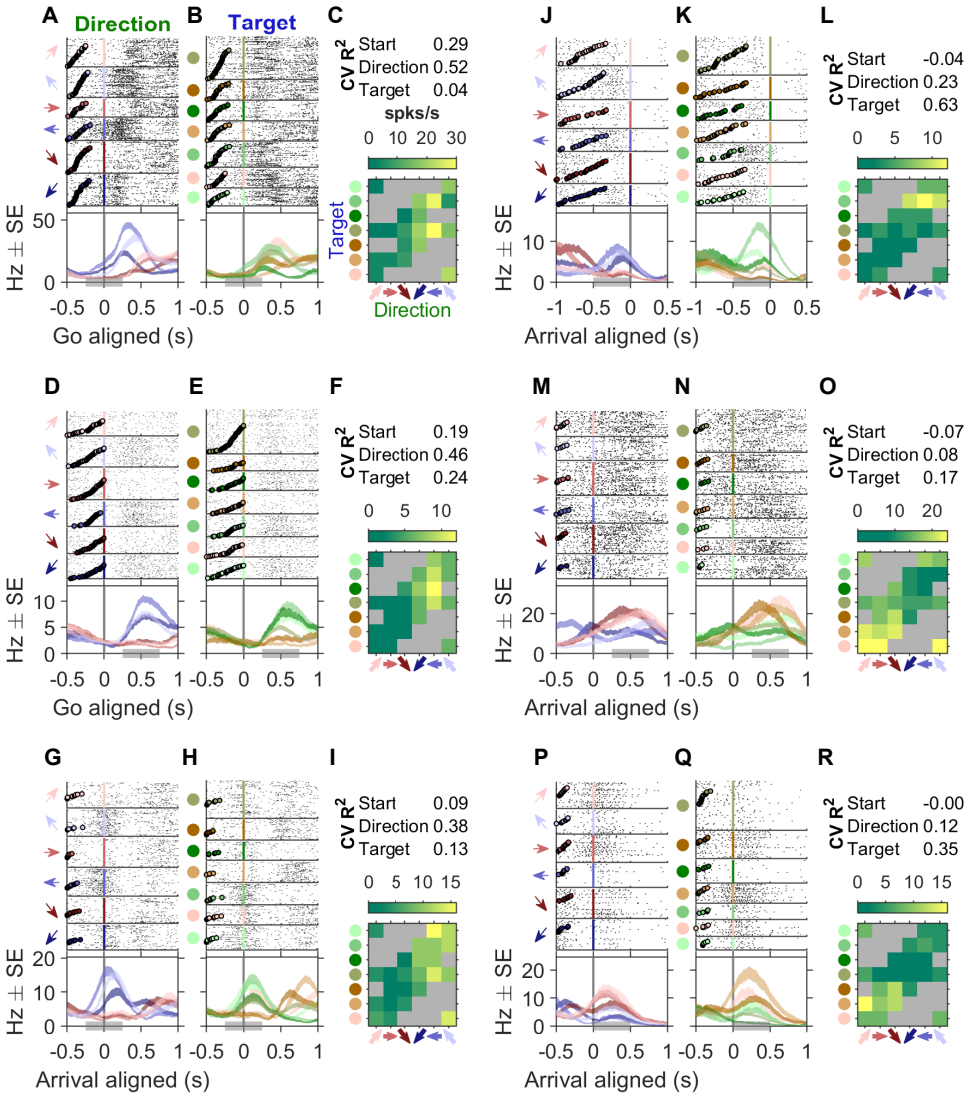

Figure 2. Example neurons with egocentric and allocentric spatial representation. **A-C** An example neuron more modulated by the direction than by the target position. **A**. Raster plots and PETHs aligned to the go sound and sorted by movement directions. The top 6 panels show spike rasters grouped by the 6 movement direction, indicated by arrows on the left side of the panels. Circles in each raster panel indicate the time of the visual cue onset on each trial. The bottom panel shows the PETHs of the spikes generated by a causal half-Gaussian kernel with an SD of 200 ms. The shaded areas of the PETHs indicate ± s.e.. The grey bar at the bottom of the panel indicates the time window to compute the firing rate in each trial type in C. **B**. Raster plots and PETHs of the same cell and same alignment as in **A**, but sorted by target position. Target positions are color coded as in FIGURE 1D. Circles in each raster panel indicate the time of visual cue onset on each trial. **C.** The maximum *a posteriori* estimated firing rate for each trial type, where the prior was a Poisson distribution whose mean was estimated from all trials. **D-F** and **G-I**, two more example cells as in A-C. **J-L**, **M-N** and **P-R**, example neurons more modulated by the target position than direction. Neural activity was aligned to the target poke. Circles in each raster panel indicated the time of the go sound.

the start position, the direction or the target position. The three example neurons in the right column were more tuned to target position than movement direction (FIGURE 2J-R).

Egocentric direction was encoded earlier than the allocentric target position.

From visual examination of neural activity, it is apparent that there is considerable heterogeneity in both the tuning to different task variables and also in the dynamics of the tuning across the trial. In order to get a holistic view of the tuning across the trial at the single neuron level, we used Poisson generalized linear models (GLMs) to fit the spike counts of each neuron using either start position, direction or target position, in four 300 ms time windows as follows: "pre-cue", -300 ms to 0 ms aligned to the visual cue onset; "post-cue", 0 ms to 300 ms aligned to the visual cue onset; "go", 0 ms to 300 ms aligned to the go sound; "arrival", -150 ms to 150 ms aligned to the target poke. The spatial variables (*start, direction, target*) were coded as factors to avoid assuming any specific functional form of tuning. Of the 1224 neurons (from 132 sessions in 4 rats), 541 (44%) were selective to one or more task variable in at least one time window. In our task, the three variables (start, direction and target) are correlated. For example, if the animal starts from a port on the left, it will be likely to move to the right. In other words, the start position conveys information about the set of possible targets and directions before the onset of the visual cue. As such, a neuron purely tuned for direction could be spuriously found to be tuned for start or target position. Our first approach in addressing this issue was to estimate which of the three variables *best* explained the neural activity. We found most neurons best tuned to the start position early in the trials. After the visual cue onset, direction tuning increased. Target position tuning emerged later than movement direction, and peaked at target poke (FIGURE 3A). This result is striking, given that the appearance of a visual target cue provides information about the direction required to move to it, *as well as* the position of the target. To validate that this method was effective, we generated surrogate spike counts (matched to the tuning and firing rates of real neurons) and found that the false positive rate (e.g. incorrectly labeling a 'start' neuron as a 'direction' or 'target' neuron) to be less than 10% (FIGURE S6C). However, this 'best variable' approach fails to accurately describe neurons that have mixed selectivity, an oversight that will be redressed later in the paper. We then extended the GLM analysis by sliding time windows aligned to task events, and the same temporal trend was captured by the R^2s of the corresponding GLMs across time (FIGURE S6A).

We further examined the geometry of spatial representation on a continuous scale using pseudopopulation decoding. We pooled all the neurons across sessions where there were at least 8 trials for each start position, direction or target position (1197 neurons, 99 sessions, 3 rats) to construct the pseudopopulation. A substantial portion of the variance in the FOF population activity was related to the spatial

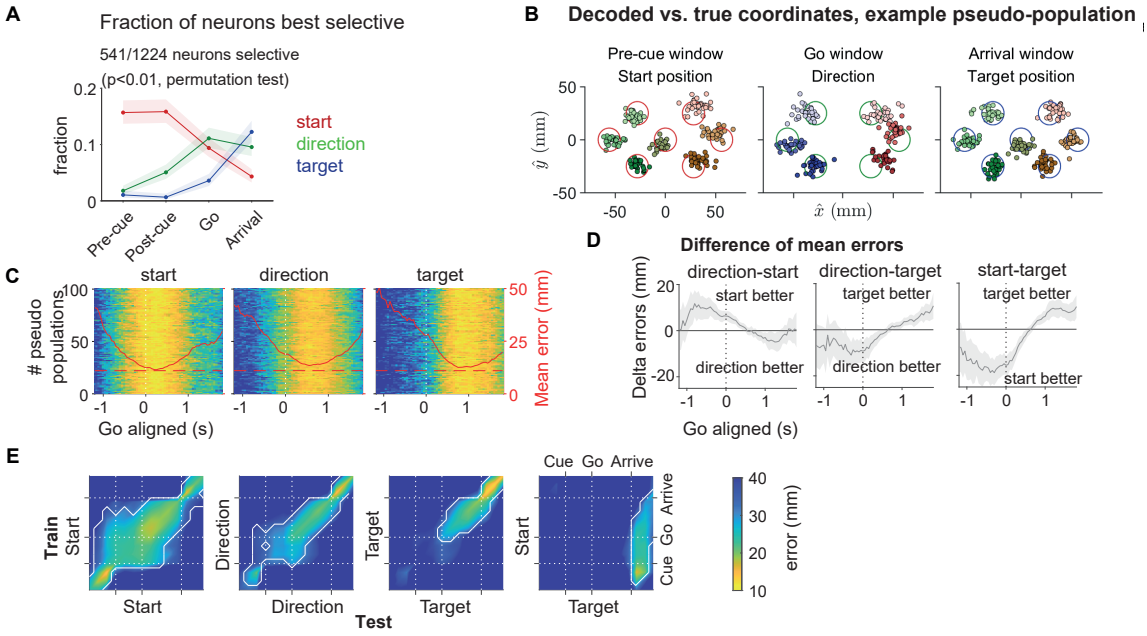

Figure 3. Temporal profile of start position, direction and target position encoding in the FOF population. **A.** The fraction of neurons best selective to the start position (red), the direction (green) or the target position (blue) in 4 time windows (see main text or Methods for definition). Shaded areas indicate the 95% binomial confidence intervals. "Pre-cue", -300 to 0 ms to visual cue onset. "Post-cue", 0 to 300 ms to visual cue onset. "Go", 0 to 300 ms to go sound. "Arrival", -150 to 150 ms to target poke. **B.** Predicted coordinates of start position, movement vector and target position in an example pseudopopulation decoding using all neurons with enough trials for each type (n = 1194). Each small circle indicates the predicted coordinates in a pseudo trial, where the color indicates the pseudo trial class. Each large circle indicates the coordinates and the diameter of a port on the port wall. **C.** Decoding errors for each pseudopopulation across time aligned to the go sound for the 3 spatial variables. Decoding error is defined as the Euclidean distance between the predicted and the actual coordinates. Each row is a different pseudopopulation where neurons were resampled with replacement. Red lines indicate the mean Euclidean distance across the 100 pseudopopulations. **D.** Mean ± s.d. of the difference of decoding errors between two spatial variables across the 100 peudopopulations. Positive difference indicates the better decoding of the second variable, and vice versa. **E.** Decoding errors with cross-window decoding. Colors of the heat-maps indicate the mean Euclidean distance between the decoded and true spatial coordinates, averaged across 100 pseudopopulations. The decoders were trained at one time window and tested at another. In the last panel the multivariate linear model was trained for start position and decoded for target position. Contours, $p < 0.01$ (extreme pixel-based test).

variables: The projection of the population activity onto its first 4 principle components represented the horizontal and vertical spatial coordinates (FIGURE S7). We decoded the coordinates of each spatial variable with a multivariate linear decoder, using only the first 4 principle components of the pseudopopulation activity patterns in 300 ms time windows with two-fold cross validation (FIGURE 3C-E, FIGURE S7D). We defined the decoding error as the Euclidean distance between the predicted and the actual spatial coordinates. The decoding was faithful: The errors were as small as the radius of the port (around 11 mm) at the best time window for each spatial variable. Put in other words, the geometry of the port wall was embedded in a linear subspace of the FOF activity (FIGURE 3C&D).

On the population level, the mean errors across pseudo-trials decreased sequentially for start position, direction and target position decoding (FIGURE 3C&D). The quality of decoding varied across different pseudopopulations, which can obscure the relative latency of decoding of different task variables. To better assess the latency, we took the difference in decoding between variables for *each* pseudopopulation, and then examined the average of these differences across the pseudopopulations (FIGURE 3D). This analyses demonstrates that at the time of the go cue, direction decoding from the population is significantly better than target decoding, consistent with results from single neuron analysis.

To investigate the stability of representation across the trial, we decoded pseudopopulation activity at one time window with decoders trained at a different time window (FIGURE 3E). We found that start position tuning was stable across much of the trial. Interestingly, decoders trained with **start positions** during fixation could accurately decode **target positions** around target poke, suggesting a consistent coding for the current head position throughout the trial (FIGURE 3E).

Single FOF neurons tracked the allocentric current head position.

We reasoned that single neurons that had consistent tuning for start position and target position should underlie the "current position" coding observed in the pseudopoulation analysis. In fact, many neurons tuned to the target position around target poke were also tuned to the start position early in the trials, and the tuning was consistent (FIGURE 4A-B). We quantified this consistency using the Pearson correlation between start position tuning in the "pre-cue" window and target position tuning in the "arrival" window, denoted as the "start-target tuning correlation" (FIGURE 4C). Among neurons selective to both the start and the target position ($p < 0.05$ for both the start and the target GLMs, n = 174), the mean start-target tuning correlation was 0.656, [0.607,0.703] (mean, [95% CI], $p < 10^{-}3$, permutation test) (FIGURE 4D).

To investigate the temporal dynamics of the correlation across the trial, we split the trials into two halves and computed the correlation between start tuning in one half of the trials and target tuning in the other half (FIGURE 4E). Then we

computed the cross-temporal correlation of start position tuning and target position tuning respectively in a similar fashion (FIGURE 4F&G). Among neurons selective to both the start and the target position ($p < 0.05$ for both the start and the target GLMs, n = 174), start position tuning early in the trial and target position tuning around target poke were positively correlated (FIGURE 4E). The peak of mean start-target tuning correlation among these neurons was 0.555, on the same scale as the target-target tuning correlation, which was 0.579. The consistency of start position and target position tuning were not limited to this selected group of 'significant' neurons. Among all the neurons with spatial selectivity ($p < 0.05$ for any one of the GLMs, n = 808; note, we used a less strict threshold here than we did for FIGURE 3), the start tuning correlation, the target tuning correlation, and the start-target tuning correlation were all positively correlated (FIGURE S8). In other words, neurons tuned to the start position were also likely to be tuned to the target position, and the start and target tuning were likely to be consistent. These observations collectively indicates a place-cell like encoding of current head position in the FOF. One would be surprised about a significant correlation between start position tuning late in the trial with target position tuning early in the trial. This seemed strange, since target information was not available before the visual cue onset. This was due to the correlation between start position and target position in our task, actually mirroring the strong correlation between early start tuning and late target tuning.

To identify the time when start position encoding transit into target position encoding, we fit the neural spike counts across time to the start and the target GLMs, and defined the time of transition as the time when the R^2 of the start position GLM first became smaller than the target position GLM (see Methods for details). For most of these neurons, the switch time lied between the go sound and the target poke time, although a few neurons switched before the go sound or after target poke (FIGURE 4H).

Spatial preference of FOF neurons.

The numbers of neurons preferring each start position and target position spanned across all the ports (FIGURE 4I-J). The preferred positions were not uniformly distributed ($\chi^2(6, N = 202) = 17.35$, $p = 0.009$ for start position, $\chi^2(6, N = 209) = 48.39$, $p = 9.86 \times 10^{-9}$ for target position). Consistent with the current head position coding, the distribution of the preferred start position among start position selective neurons were similar to the distribution of preferred target position among target position selective neurons ($\chi^2(6, N = 411) = 9.67$, $p = 0.139$, FIGURE 4I&J).

To our surprise, the number of neurons with preferred directions on the left-right axis was significantly larger than those with preferred directions on the up-down directions ($\chi^2(1, N = 274) = 76.91$, $p < 10^{-5}$, FIGURE 5A&B). Interestingly, rats mostly made errors into the same left/right side as instructed (FIGURE S2B). For

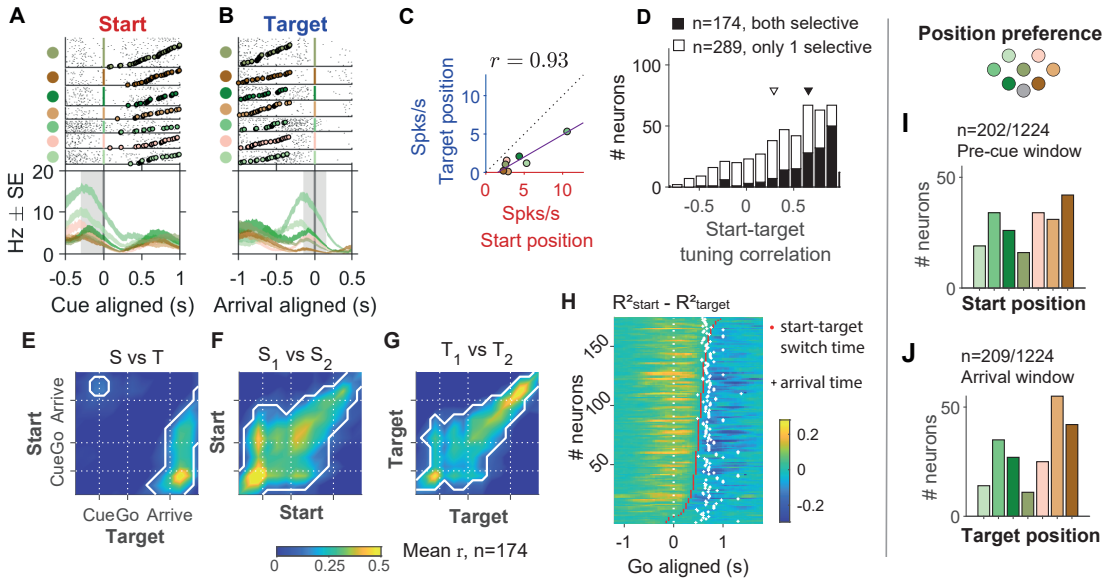

Figure 4. FOF neurons encoded the current head position. **A.** Raster plots and PETHs of an example neuron aligned to the cue, grouped by start position. The shaded grey area indicates the time window to calculate the x-axis firing rate in C. **B.** The same neuron aligned to target poke, grouped by target position. The shaded grey area indicates the time window to calculate the y-axis firing rate in C. **C.** The correlation between start position tuning and y position tuning in the example neuron, where r denotes start-target tuning correlation of this neuron as in D. Purple line denotes the total least square fit. **D.** The distribution of the start-target tuning correlation. Black bars are for neurons selective to both start position and target position, and white bars are for neurons selective to only one of the two variables (black, 0.66, $[0.61, 0.70]$; white, 0.29, $[0.25, 0.34]$, 95% CI of the mean). Triangles indicate the means. **E.** The start-target tuning correlation in warped time windows aligned to the visual cue, the go sound and the arrival, averaged across neurons with both start and target selectivity ($n = 174$). The white contours indicate the areas where correlation is significantly larger than 0 ($p < 0.05$ with Bonferonni correction). Different from C and D, these correlations were calculated between start tuning in half of the trials and target tuning in the other half of trials, and vice versa, then took the average. **F.** Similar to E, but for the mean Pearson correlation between pairs of time windows for start position tuning in one half of trials versus the other half. **G.** Similar to F, but for target position tuning. **H.** Time of transition from start position coding to target position coding in single neurons. The color of the heat-map indicate the difference between R^2s of the start position GLM and the target position GLM, calculated at each 300 ms time window aligned to the go sound. Each row was a neuron (n = 174). The red dots indicate the time of switching from R^2_{start} higher to R^2_{target} higher (see Methods for details). The white crosses indicate the averaged time of target poke for that session. **I.** The number of neurons preferring each start position in the "pre-cue" time window, among cells that had significant start position selectivity in the "pre-cue" window ($p < 0.01$, permutation test for the start position GLM). **J.** The number of cells preferring each target position in the "arrival" time window, among cells that have significant target position selectivity in that window ($p < 0.01$, permutation test for the target position GLM).

example, when instructed to top-left, errors were usually to the middle-left or bottom-left ports but rarely to the right. The coincidence between neural tuning bias and behavior bias tempts us to speculate that rats may relatively under-represent vertical directions compared to horizontal directions. Consistent with previous findings, there was no significant difference between the numbers of neurons preferring the ipsilateral and the contralateral side ($\chi^2(1, N = 274) = 0.88$, $p = 0.34$, Erlich et al., 2011).

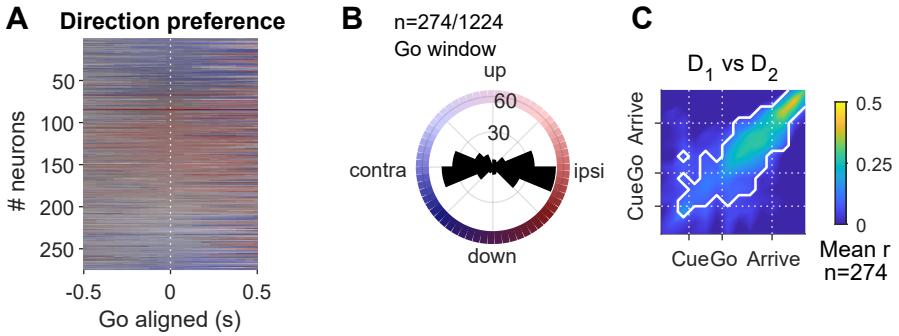

Figure 5. Direction preference of FOF neurons. **A.** The preferred direction of 274 cells that had significant direction selectivity in either the "post-cue" or "go" time windows ($p < 0.01$, permutation test for the GLM), in sliding windows aligned to the go sound (50 ms per bin, 300 ms bin size). The alignment was causal, i.e. time 0 indicated -300 to 0 ms. The saturation of the colors indicate the relative amplitude of the R^2 of the direction GLM, and the neurons were sorted by the preferred direction in the "go" time window. **B.** The distribution of preferred direction in the "go" time window (0 to 300 ms after the go sound). **C.** Pearson correlation of direction tuning curves at one time versus another, among same neurons as in A and B. Color indicate the mean correlation across these neurons. White contour indicates the area with where correlation was significantly larger than zero with Bonferroni correction.

Mixed selectivity of movement direction and head position.

In previous analyses, we examined the encoding and decoding of one spatial variable at a time. Nonlinear mixed selectivity supports flexible readout by allowing high-dimensional representation of information from multiple sources (Rigotti et al., 2013). For example, a nonlinear transformation is required to compute the length of a 2-D vector ($\sqrt{A^2 + B^2}$), and a linear readout of the distance is only feasible if A and B are encoded in a space with 3 or more dimensions. In primates gaze shifts, nonlinear interactions of the saccade direction and the initial eye-in-orbit position in the frontal eye field has been observed (Andersen et al., 1990, 1985, Andersen and Mountcastle, 1983). The nonlinear coding was specified as the gain field model, where the initial position multiplicatively influenced the bell-shaped direction tuning. The gain field model has been suggested as a mechanism for reference frame transformation (Pouget and Sejnowski, 1997, Zipser and Andersen, 1988). Thus, to

compare the FOF activity with previous results from FEF (and other regions) we tested whether FOF neurons encoded spatial variables with pure, additive or multiplicative tuning.

To determine which model best described single neuron activity in the FOF, we fit spike counts of each neuron in the 0 to 500 ms time window aligned to the visual cue onset to 4 functional forms: pure direction tuning (Eq. 1), pure position tuning (Eq. 2), an additive model (Eq. 4), or a gain-field model (Eq. 3)(FIGURE 6A-C). The fraction of neurons that had larger $CV\ R^2$ for the gain field model than any other model was the highest (FIGURE 6D). Pairwise model comparison revealed that the cross-validated R^2s ($CV\ R^2$s) for the gain field model were significantly larger than other models, among neurons whose average $CV\ R^2$s for the 4 models were larger than 0.05 (permutation test, FIGURE 6E, TABLE 1). The result was qualitatively the same for spike counts in the -300 to 500 ms window aligned to the go sound, and in the 500 ms window with maximum cross-trial-type variance aligned to the visual cue onset. Thus the majority FOF neurons had nonlinear mixed selectivity to the self-centered and the world-centered spatial variables.

M_1	M_2	p value	$atanh(M_1) - atanh(M_2)$ Mean, [95% CI]
Gain field	Additive	2×10^{-5}	0.0129, [0.0073, 0.0229]
Gain field	Direction	2×10^{-5}	0.0617, [0.0544, 0.0774]
Gain field	Start	2×10^{-5}	0.0753, [0.0650, 0.0960]
Additive	Direction	2×10^{-5}	0.0488, [0.0430, 0.0588]
Additive	Start	2×10^{-5}	0.0624, [0.0529, 0.0775]
Direction	Start	0.02	0.0136, [0.0025, 0.0259]

Table 1. Difference of the Fisher-transformed CVR^2s between model pairs for neurons whose mean $CVR^2 > 0.05$ for all four models. p values were computed by bootstrapping the $atanh(M_1) - atanh(M_2)$ distribution using 10^5 bootstraps.

Head position explained FOF activity better than the body posture.

Neurons in M2 have been reported to encode the body posture rather than the upcoming movement in rats foraging in large arena (Mimica et al., 2018). We thus tested whether the spatial encoding we observed were actually representing body posture. We analysed video and extracted the coordinates of the animals' head, ears and the hip in 58 sessions, which had 638 single units. We used the angle of the head-hip axis relative to the behavior box as a proxy for posture, denoted as the body angle (FIGURE 7A). Unsurprisingly, we found that body angle was strongly correlated with the current head position. When a rat's head was in one of the right-side ports, their body was also angled to the right, and likewise for the left-side ports (FIGURE 7B, 0.616, [0.578, 0.648],mean, [95% CI] for the linear model of $\theta \sim port$ in 59 sessions).

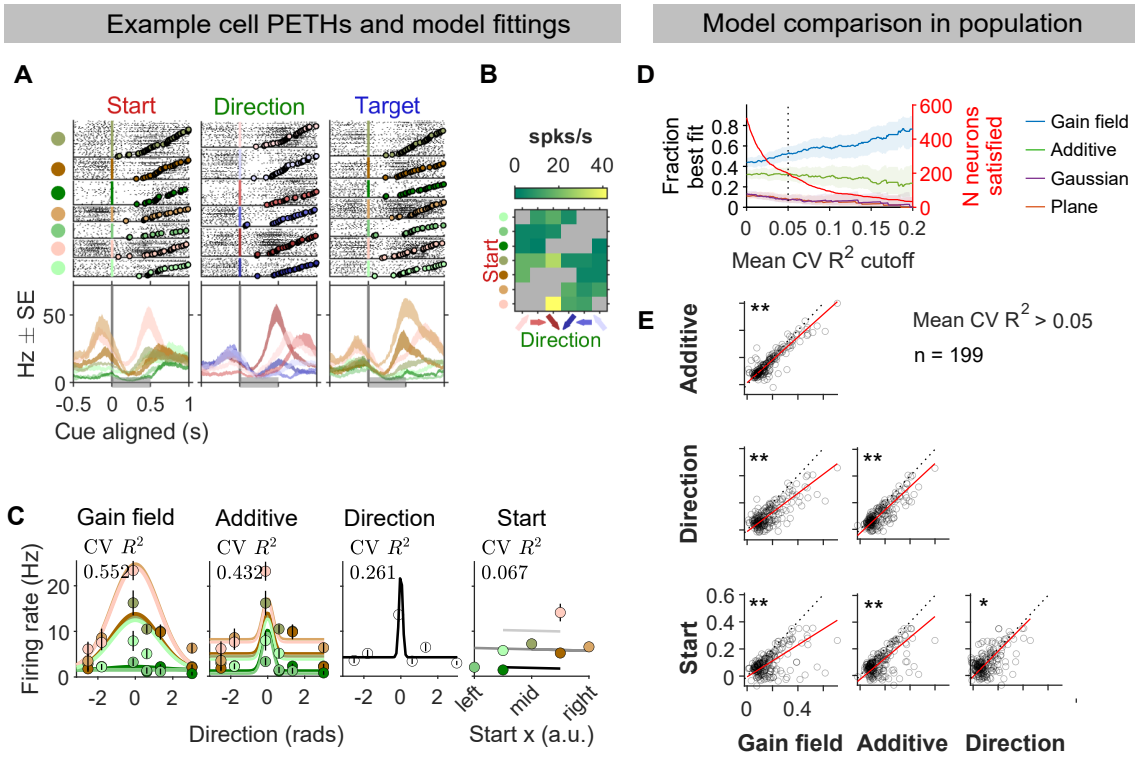

Figure 6. The spatial selectivity of FOF neurons were best explained by the gain field model. **A.** Raster plots and PETHs of an example neuron. Trials were grouped by the start position (left panel), the direction (middle panel) and the target position (right panel). Spikes were aligned to the visual cue onset. PETHs were generated by a causal half-Gaussian Kernel with an SD of 400 ms. The gray bar at the bottom indicates the 500 ms time window after the visual cue onset, the time window used for other panels in this figure. **B.** Estimated mean firing rate in each movement trajectory using a maximum a posterior estimator, same as in FIGURE 2. **C.** Predicted firing rates of four fit models (lines) and the mean and s.e.(circles and error bars) of firing rates in each trial condition. $CV R^2$, cross-validated R^2. **D.** Left axis, fraction of neurons best fit by the model, among neurons whose mean cross-validated R^2s over the 4 models was larger than the x-axis indicated value. Best fit was defined as having the largest cross-validated R^2 among the 4 models. Error bars were 95% confidence intervals of the binomial distribution. Right axis, neurons that crossed the mean $CV R^2$ criteria for each x-axis value. **, $p < 0.001$. *, $p < 0.05$. **E.** Each panel plots the cross-validated R^2s of the x-axis model versus the y-axis model. Each circle indicates a neuron. The red line indicates the total least square fit to the data. The dashed black line marks the diagonal. All single units were included in the analysis (n = 1224), but each panel only showed neurons whose mean of cross-validated R^2s in the 4 models were larger than 0.05 (n = 199). Statistical test and total least square fitting were also based on these neurons. P values indicates the significance against the null hypothesis that the difference between the Fisher-z transformed cross-validated R^2s in the x-axis model and the y-axis model was not significantly different from zero (permutation test of the mean). The mean R^2 in the x-axis model was larger than the y-axis model in all of these panels.

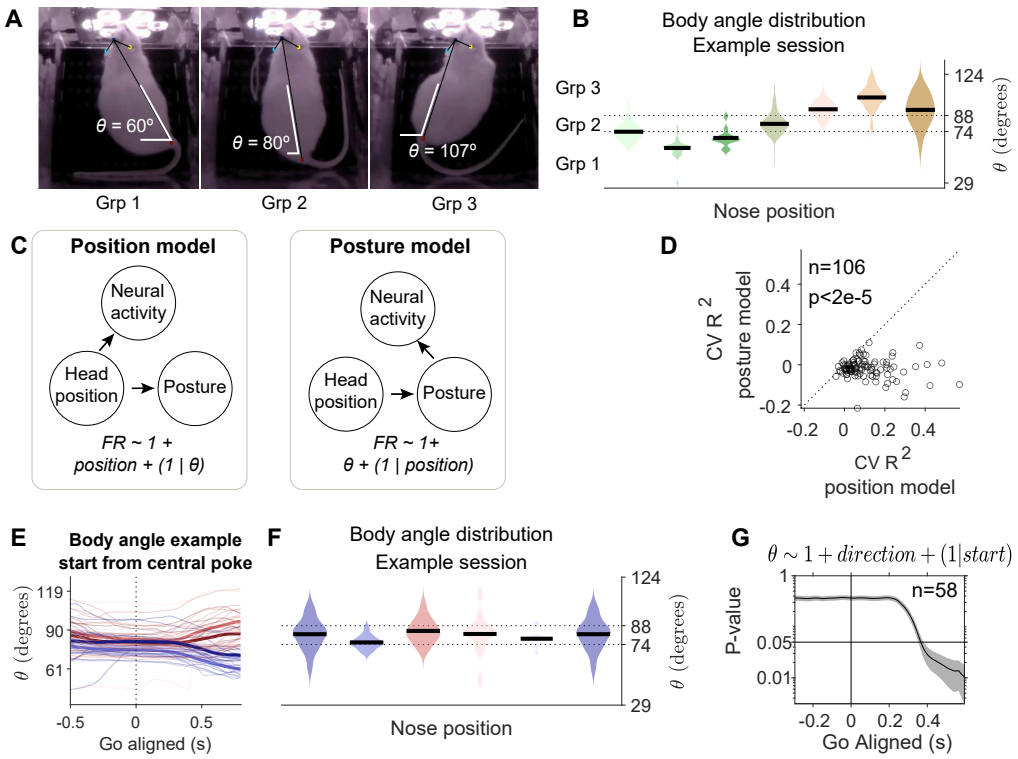

Figure 7. The Head position explained FOF activities better than the body posture. **A.** The body angle in video frames from the fixation period in 3 trials starting from the central port in the same session. **B.** Distribution of body angles during fixation for each start port in the same session as in A. The start ports are color coded. Black lines in the violin plots indicate the group means. Dashed lines indicate the separations of the 3 groups of body angles with equal numbers of trials. **C.** Two possible models of the causal relationship among the neural activity, the head position and the body posture. In the *position model* (left), both neural activity and posture were causally modulated by the experimentally manipulated head position. In the *posture model* (right), neural activity was causally modulated by the head position only through the mediation of posture. The two causal graphs were formulated as two generalized linear mixed effect models, where posture was quantified with the body angle θ. **D.** The cross-validated R^2s of the position model versus the posture model. Each circle is one neuron that had significant start position among sessions with video tracking. P-values were against the null hypothesis that the mean (position R^2- posture R^2) across neurons was not significantly different from zero. **E.** The body angle when moving from the central port towards different directions across time aligned to the go sound. Each thin line was a trial (n = 76). Each thick line was the average of trials into the 6 directions. Data was from the same example session as in A and B. **F.** The distribution of body angles when the nose was in the central port (n = 76) plotted against the movement directions (color coded). Data was from the same example session as in A, B and F. **G.** The significance level of the movement direction modulating the body angle across time aligned to the go sound, where the effect of start position was captured in the random effect. Line and error bars, mean ± s.d. of the p values over the 58 sessions.

To identify whether neural activity was tuned to the head position versus posture we took a causal inference approach (Neal, 2020). Note that the head position is under experimental control in our task (in contrast to Mimica et al., 2018). Also, we have already established that there is a correlation between head position and neural activity in the FOF. Thus, there are really only two relevant causal models to consider. First, the 'position model', where neural activity is causally affected by the head position, and correlation with posture is mediated by a 'backdoor path' through head position (FIGURE 7C, LEFT). In this model, conditioning on the body posture will not alter the correlation between neural activity and the head position. In the second 'posture model', the correlation of neural activity with position is mediated by the influence of posture on neural activity (FIGURE 7C, RIGHT). In this scenario, the effect of posture on neural firing does not depend on head position. Conditioning on head position will not remove the correlation between posture and neural activity, but conditioning on posture will remove the correlation between head position and neural activity. We found that the position model was as good or better than the posture for almost all neurons with position selectivity (FIGURE 7D, $p = 2 \times 10^{-5}$, n=106, permutation test). In contrast to the correlation between body angle and position, the body angle was not significantly correlated with the movement direction until around 350 ms after the go sound (FIGURE 7E-G). Given that the upcoming movements were encoded before the go sound (FIGURE 2A), spatial encoding in the planning phase could not be explained by body posture.

Discussion

Motor planning in rodents has previously been studied with tasks where the action space is either very high dimensional (such as navigating a maze or an open field) or very low dimension (such as 2AFC or go-nogo tasks). We took an intermediate approach: We studied planning in an orienting task that involved movements across 6 directions and 7 head position, and also provided precise experimental control of the timing of the task. We observed encoding of both egocentric and allocentric spatial parameters at the single-neuron and population level in the FOF. The encoding of allocentric start position, egocentric movement direction and allocentric target position emerged sequentially over the trial. This is striking, given that the visual target cue simultaneously conveys egocentric and allocentric information. In single neurons, tuning to the start position during fixation was correlated with target position tuning at the time of target poke, suggesting that those neurons tracked the animals' current head position. The activity of the majority of cells were best explained by a gain field model, where Gaussian direction tuning was modulated by position, compared to single-task-variable models and an additive model.

Subjects made few errors in the task, but when they did, they mostly made up/down errors (FIGURE S2B). We also observed an over-representation of left versus right than up versus down direction coding in the FOF (FIGURE 5).

Although we do not have evidence for a causal relationship between these two observations, one might speculate that the left/right over-representation in the FOF might be behaviorally relevant (Jovalekic et al., 2011). In contrast to our findings, there were more pitch (up/down) tuned neurons than azimuth (left/right) tuned neurons reported in the M2 in rats foraging in a large arena (Mimica et al., 2018). There were two key differences between our experiment and theirs: they recorded from a larger anterior-posterior range of M2; and our rats were under tight experimenter control whereas their rats were foraging freely. Also, in our task, all visually-guided movements had a horizontal component which might have biased our results. It would be interesting to record the same neural populations in two tasks: one like ours and one like theirs to see if neurons in M2 can dynamically shift their tuning based on task demands.

Our task was inspired by paradigms widely used in non-human primates to study the neural mechanism of saccadic eye movements (Bruce and Goldberg, 1985). Rodent head orienting is functionally similar to the primate gaze orienting as a form of overt attention in that it redirects the sensory fields of vision, audition, olfaction and whiskers (Bush et al., 2016, McCluskey and Cullen, 2007, Monteon et al., 2010). Based on similarities in connectivity and function, the rodent FOF is considered be analogous to the primate FEF (Ebbesen et al., 2018, Erlich et al., 2011, Reep et al., 1987, 1990), but a strict correspondence between rodent and primate frontal regions may not exist (Barthas and Kwan, 2017, Wise, 2008). That said, our finding of gain field modulation of movement direction by initial head position is similar to the observation that initial gaze direction gain modulates the planned saccade direction in FEF (Cassanello and Ferrera, 2010, 2007).

Neurons in the primate FEF have mostly been reported to have egocentric spatial representations (but see Bharmauria et al., 2020). However, the body positions in primate experiments were typically restricted by a primate chair, whereas our rats moved around in the box during the task, making the world reference frame more stable than the body reference frame. Recent work in freely moving monkeys found widespread world-centered coding across frontal and prefrontal cortex (Maisson et al., 2022), including supplementary motor area and dorsolateral prefrontal cortex. Although FEF was not one of the regions they recorded from, the finding suggests that allocentric representations are common across the brain, when animals' movements are not restricted.

Our result builds on a recent interesting finding of allocentric modulation of egocentric movement planning in rat M2 (overlapping anatomically with the region we define as FOF; Olson et al., 2020). There are three main advances of our result, compared to theirs. First, we use a vertically oriented port wall instead of maze. This allowed us to demonstrate that the FOF contains 2-dimensional maps (horizontal and vertical planes) for both position and directions (Figure 3B). Second, their animals were goal directed, but were not explicitly cued on each trial about the goal location or direction while they navigated along a triple T-maze. In our task, the target was visually cued on each trial during a fixation period, which

gives us good experimental control over planning versus execution phases. This allowed us to demonstrate that planning is done preferentially in egocentric coordinates (FIGURE 3). Finally, we determined that modulation of direction tuning by position was via a multiplicative gain field (FIGURE 6).

Gain modulation is a nonlinear process in which neurons combine information from two or more sources, evident in a plethora of primate cortical and subcortical brain regions (for review, see Salinas and Sejnowski, 2001). It is a powerful computational mechanism for coordinate transformation, such as vector summation and subtraction (Pouget and Sejnowski, 1997, Salinas and Abbott, 1995, Zipser and Andersen, 1988). In our task, summation of the start position coordinates and the egocentric movement vector gives the target position coordinates. We observed all these variables to be coded in the FOF. In some neurons, target position coding emerged before the rat arrived at target, suggesting that it may have been internally computed rather than being due to sensory feedback related to the target poke (i.e. proprioceptive, visual and whisker inputs related to poking in the target). However, whether the target position coding was locally computed in FOF from the start position and movement vector coding is unknown. It could also be computed in other brain areas, such as the posterior parietal cortex or retrosplenial cortex, which contain mixtures of egocentric and allocentric encoding (Wang et al., 2020). Error trial analysis would be helpful at determining whether the target encoding in FOF was computed from the start and direction coding since it would allow us to ask whether target encoding was more like where the animal went vs. where we would expect the animal to go based on start and position decoding. However, our visually-guided task had a low error rate which hampered our ability to perform this analysis.

Our results, which were quite surprising to us, are largely consistent with decades of primate research into the neurobiology of motor planning (Andersen et al., 1985, Bharmauria et al., 2020, Caruso et al., 2018, Salinas and Sejnowski, 2001, Wang et al., 2007, Zipser and Andersen, 1988) suggesting conserved neural computations for motor planning across species. However, despite decades of research, a full understanding of phenomena like gain fields and reference-frame shifts has been hampered by a lack of tools for precise perturbations for circuit dissection in primates (cited from Shenoy et al., 2013). In contrast, the neural circuits underlying basic phenomena of motor planning (i.e. spatial memory and movement initiation in 2AFC tasks), have been revealed in rodents models (Bharmauria et al., 2020, Duan et al., 2021, Guo et al., 2017, Inagaki et al., 2019, Kopec et al., 2015, Li et al., 2016, Yang, 2022). Our behavioral paradigm and neurophysiological observations provide a basis for employing the latest tools of rodent systems neuroscience to make progress in understanding all the phenomena associated with ethological motor planning.

Materials and Methods

Subjects

Three adult male Sprague-Dawley rats and one adult male Brown Norway rat (Vital River, Beijing, China) was used in this study. For a portion of the experiments presented here, rats were placed on a controlled-water schedule and had access to free water 20 minutes each day in addition to the water they earned in the task. For some experiments, rats were given free access to a 4% citric acid solution (Reinagel, 2018), in addition to the normal water they earned in the task. They were kept on a reversed 12 hour light–dark cycle and were trained during their dark cycle. Animal use procedures were approved by New York University Shanghai International Animal Care and Use Committee following both US and Chinese regulations.

Behavior

Rats were trained in custom behavioral chambers, located inside sound- and light-attenuated boxes. Each chamber ($23 \times 23 \times 23$ cm) was fit with a vertical 2-D port wall that had 7 operant ports and 1 reward delivery port, with speakers located on the left and right side (FIGURE 1). Each operant port contained a pair of blue and a pair of yellow light emitting diodes (LED) as visual cues, as well as an infrared LEDs and photo-transistors for detecting rats' interactions with the ports. The reward delivery port contained a stainless steel tube for delivering water rewards.

The task timeline is described in detail in Results and FIGURE 1. In one rat (2147), in addition to the main type of timeline, which we denote as "target during fixation", there were two other types of trial timelines: "target before fixation" and "target after go". These trial types are described in detail in FIGURE S4.

The duration of the fixation period was dynamically adjusted for each animal, ranging between $0 \sim 1.2$ s (FIGURE S2). A trial was considered a fixation violation if the rat withdrew from the start port before the go sound. In fixation violation trials, an "error" sound was delivered and the trial was aborted.

In the final behavioral stage, three rats (subject ID 2068, 2095, 2134) moved in 6 directions and 30 movement trajectories, and one rat (subject ID 2147) moved in 4 directions and 16 movement trajectories (FIGURE S1).

In rat 2068, 2095 and 2134, there were two session types interleaved across days: the "reference" sessions and the "distance" sessions. In the "reference‘' sessions, each of the 6 directions had 4 movement trajectories of the same distance. In the "distance" sessions, each of the 6 directions had 3 movement trajectories involving 3 ports, where the distance of one movement trajectory is twice of the movement of the other two(FIGURE S1).

In rat 2147, there were 4 movement directions, from and to 7 operant ports. Each direction had 4 trajectories of the same distance, summing up to a total of 16 trajectories (FIGURE S1).

Behavior training pipeline

Rats went through a series of training stages, which had mainly two phases: (1) the operant conditioning phase, and then (2) the multi-directional orienting phase.

In the operant conditioning phase, rats became familiar with the training apparatus and learned to do a one step poke into the illuminated choice port. The first stage was to learn to collect reward from the reward delivery port. Each trial began with the illumination of the reward port, and water reward was immediately delivered upon port entry, followed by a short timeout period before the start of next trial. After the rats learned to poke into the reward port reliably (not missing any reward for 6 trials in a row), they proceeded to the next training stage. In the second stage, we turned on the LED for several random ports at the beginning of each trial. Rats had to first poke into any illuminated choice port before gaining water reward from the reward delivery port. The number of the illuminated port will gradually decrease to one after several trials when animals started to learn. After animals were able to poke the only illuminated choice port successfully for 6 trials in a row, we will upgrade them to the second training phase.

In the first stage of the multi-directional orienting phase, the start port was always the central port, and the target port was one of the 6 surrounding ports. Rats needed to poke into the start port to trigger the target port light, and then poke into the target port after a delay. Trials of the same movement trajectory was repeated until the animal could do several correct trials in a row. The training of "fixation" at the start port was introduced in this phase. Fixation means the rat had to keep its nose in the start port for a given time period (typically $> 0.5s$). Fixation duration was initially 0.02 s early in the training, and was gradually increased based on an adaptive steps method: the fixation duration would increase on a successful fixation, and decrease when the fixation failed. We trained subjects to perform fixation for at least 0.6 s before the surgery, and fixation duration always jittered across trials in recording sessions. However, the speed to recover fixation after the surgery varied across subjects, thus we manually adjusted the fixation duration for each subject. In 2095, the mean fixation period in each session was shorter than 0.2 s for around 30% of sessions. In other subjects, the mean fixation period was typically longer than 0.4 s (Figure S2 G). In the second stage of the training phase, the start port could be any one of the 7 ports, and the target port was one of the 6 remaining ports. Rats were trained on the "target during fixation" trial class. Rat 2147 was then introduced to the "target before fixation" and "target after go" trial classes, described in Figure S4.

Electrophysiology

Rats were unilaterally implanted with Cambridge Neurotech microdrives and 64 channel silicon probes. To target the frontal orienting field (FOF), silicon probes were placed at anteroposterior (AP) and mediolateral (ML) coordinates as following: rat 2068, AP +2.2, ML +1.5; rat 2095, AP +2.0, ML +1.5; rat 2134, AP +2.5, ML

-1.5; rat 2147, AP +2.2, ML -1.4 (FIGURE S3). The probes were placed initially 0.5 mm below the brain surface and were advanced 50 to 100 um every 1 to 4 days, after each recording session. The same neurons could be sampled across sessions.

Analysis of neural data

Spike sorting Spike sorting was performed with automatic clustering and manual curation in JRClust. Single units were defined as those with signal-to-noise ratio larger than 5, fraction of inter-spike interval smaller than 1 ms less than 1% , and within-trial firing rate larger than 1 Hz. However, our main results were robust to different single neuron criteria (FIGURE S5).

Data inclusion For a unit to be included in the main figures, the session must have at least 8 trials for each of the 7 start positions, 4 or 6 movement directions and 7 target positions. The sessions could be "reference", "distance", or sessions from rat 2147. As described above, there were 3 types of trial timeline in rat 2147, but only "target during fixation" trials are included in the main figures. This resulted in 104 sessions and 1224 single cells in 4 animals.

In pseudopopulation decoding, we included cells from sessions that had at least 8 trials for each of the 6 movement directions, 7 start positions, and 7 target positions. The sessions could be "reference" or "distance" sessions. This resulted in 1197 cells from 99 sessions, 3 animals.

Video tracking was available in 58 sessions in 3 rats, and there were 638 single neurons in these sessions. In posture tuning analysis, we only included neurons with start position selectivity during fixation ($p < 0.01$, permutation test) from sessions that had video tracking, resulting in 106 cells, 30 sessions from 3 rats.

For neuron inclusion criteria for all the main figures, see TABLE S1.

Unless otherwise specified, only correctly performed trials were included in the analysis.

Key time windows In single-neuron level analysis for single spatial variables, we focused on 4 key time windows: *pre-cue*, -300 to 0 ms aligned to target cue onset; *post-cue*, 0 to 300 ms aligned to target cue onset; *go*, 0 to 300 ms aligned to the go sound; *arrival*, -150 to 150 ms aligned to arrival at the target.

When comparing the pure and mixed selective models (FIGURE 6), we used the 0 to 500 ms time window aligned to the visual cue onset in.

For continuous time windows, neural response was quantified by counting the number of spikes in sliding windows of 300 ms width and 50 ms step, aligned causally to task events, unless otherwise specified. Causal alignment means that the value at time 0 refers to the neural activity in a time bin between -300 ms and 0 ms. The responses could be aligned to the time of visual target cue onset, the go sound onset, the time of "fixation out" (the time when the nose left the start port,

detected by the IR sensors) or target poke (the time when the nose arrived at the target port, detected by the IR sensors).

Cross-validated R^2s and log likelihood The cross-validated R^2s (denoted as CVR^2s in figures) was defined as

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{Var(y)}$$

$\hat{y}_i$ was the predicted mean spike count in the ith trial from the test set, and y_i was the observed spike count in this trial. In the best case, y is equal to $\hat{y}$, and the cross-validated R^2 is 1. In the worst case, $\hat{y}$ is uncorrelated with y, and $\frac{\sum_i (y_i - \hat{y}_i)^2}{Var(y)}$ can be larger than 1 by chance, thus CVR^2s can be negative.

The cross-validated log likelihood is defined as

$$LL = \sum_{i=1}^{N} log(P(y_i|\hat{y}_i))$$

$P(y_i|\hat{y}_i)$ denote the probability of observing the spike count on the ith trial being y_i, given the predicted mean spike count being $\hat{y}_i$. N denote the number of trials. We assumed spike counts to follow the Poisson distribution.

PETHs For each neuron, we combined the spike trains in correctly performed trials by the movement directions or target positions, and generated smoothed PETHs with an half-Gaussian kernel of 200 ms standard deviation. The kernel was causal, such that selectivity at time t was contributed by neural activity at or before time t.

Generalized linear models for single neuron selectivity We fit the neural spike counts in specific time windows to 3 generalized linear models (GLMs) with Poisson distributions:

$$Spikes \sim \beta_0 + \beta_1 X_{start}$$

$$Spikes \sim \beta_0 + \beta_1 X_{direction}$$

$$Spikes \sim \beta_0 + \beta_1 X_{target}$$

The spatial variables were all included as dummy variables. Start position had 7 levels, direction had 6 levels, and target position had 7 levels.

In the main text, we labeled neurons as "having significant selectivity to a spatial variable". This significance is derived from the permutation test of the GLM, where the test statistic was the leave-one-out cross-validated log-likelihood, against the null hypothesis that the log-likelihood was not significantly different from when trial labels were shuffled. For each spatial variable, the trial labels were shuffled 1000 times to obtain a distribution of goodness-of-fits, then the p value of the GLM was

the fraction of log-likelihoods from shuffling that is greater than the log-likelihood from actual data.

The best selectivity of a neural response (FIGURE 3A) was assigned to the spatial variable with the smallest p value of the GLM. When there were ties in p values, we additionally compared the leave-one-out cross-validated log-likelihoods, and the best selectivity was assigned to the spatial variable with the largest sum of log-likelihoods. The reason for possible ties is that we are using permutation tests with 5000 permutations.

In FIGURE 3 B, the R^2s were derived from GLMs fit similarly, but without cross-validation. Only the model with the largest R^2 was plotted with the corresponding color at each time point. Both the best GLM measurement and the R^2 measurement for the relative strength of spatial selectivity were robust to the task-induced correlation between spatial variables, which was verified with surrogate data (FIGURE S6, also see Surrogate data paragraph below).

Cross-validated R^2s of the GLMs in FIGURE 2 were derived from 10-fold cross validation.

Spatial preference of single neurons The preferred direction of a neuron in a specific time window was defined as the direction of the weighted vector sum of the coordinates of the 6 movement directions, where the weight for each direction was the mean firing rate in that direction. Horizontal directions were defined as directions between $-45°$ and $45°$ around the horizontal directions, and similar for vertical directions. The preferred position of a neuron at a specific time window was simply denoted as the port with the highest mean firing rate.

Start-target tuning correlation The start-target tuning correlation in FIGURE 4D was defined as the Pearson correlation between the tuning curve for start position in the "pre-cue" window and target position in the "arrival" window:

$$\rho = cov(FR_{start}, FR_{target})/\sigma(FR_{start})\sigma(FR_{target})$$

FR_{start} is a vector where each element is the mean firing rate for a specific start position, and FR_{target} is a vector where each element is the mean firing rate for a specific target position. Similarly, we calculated the start-target tuning correlations for different pairs of time windows and the tuning correlation between time t_i and t_j for the same variable. The time windows were warped to align to the time of the visual cue, the go sound and the target poke. The 300 ms "pre-cue" and "arrival" time windows were preserved. For these time windows, we calculated the tuning correlations between to halves of trials that were randomly split, that is, we used start tuning in the first half versus target tuning in the second half, and vice versa, and then took the average of the two correlation coefficients.

The tuning correlation was subject to the Fisher-z transformation, and then tested against the null hypothesis that the mean value across a specific neural population was not significantly larger than zero. The p value indicated the fraction

of mean lying below zero among 10^4 bootstraps, and was corrected with Bonforroni correction.

Timing of start-target switching In FIGURE 4 I, spike counts in 300 ms time windows of 50 ms steps were fit to the Poisson GLM of either start or target position. The R^2s of each GLM across time was smoothed with a moving average kernel with the size of 3 bins, and then the start GLM R^2 was subtracted by the target GLM R^2. The time of switching from start encoding to target encoding was defined as the first time window where $R^2_{start} - R^2_{target} < 0$, after the positive peak of $R^2_{start} - R^2_{target}$.

Mixed selectivity To detect mixed selectivity for each neuron, we compared between 4 models of the neural firing rate.

In the Gaussian direction tuning model, the firing rate was a Gaussian function centered by the preferred direction, defined as in the "Spatial preference of single neurons" section in Methods.

$$f(\theta) = b_0 + b_1 \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{\theta^2}{2\sigma^2}) \tag{1}$$

In the start position plane model, the firing rate was modulated linearly by the horizontal and vertical coordinates of the start position.

$$f(x, y) = b_0 + b_1 x + b_2 y \tag{2}$$

In the gain field model, the firing rate was a multiplicative combination of a Gaussian tuning centered by the preferred direction and a start position modulation.

$$f(\theta, x, y) = b_0 + (b_1 x + b_2 y) \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{\theta^2}{2\sigma^2}) \tag{3}$$

In the additive model, the firing rate was an additive combination of a Gaussian tuning centered by the preferred direction and a start position modulation.

$$f(\theta, x, y) = b_0 + (b_1 x + b_2 y) + \frac{b_3}{\sigma\sqrt{2\pi}} \exp(-\frac{\theta^2}{2\sigma^2}) \tag{4}$$

In these functions, f was the firing rate, θ was the movement direction relative to the preferred direction, and x and y are the horizontal and vertical coordinates of the start positions in the port wall. $b_{0,1,2,3}$ and σ were fit for each model with maximum likelihood estimation assuming Poisson spike counts, using the Matlab *fmincon* function. The models were fit with 20-fold cross validation.

For each pair of models, we compared the cross-validated R^2s among neurons whose average of cross-validated R^2s of the 4 compared models was larger than 0.05. Significance level was tested against the null hypothesis that the mean difference between the Fisher-z transformed cross-validated R^2s in the x-axis model and the

y-axis model was not significantly larger than zero, quantified as the fraction of the mean value smaller than zero among 10^5 bootstraps.

For a given movement direction, the target position coordinates were simply an additive translation of the start position, thus the start position and target position modulations were the same in the gain field model and the additive model.

Pseudopopulation decoding Pseudopopulations were constructed separately when decoding each spatial variable. For example, when decoding the start position, we constructed the pseudopopulation by resampling trials for each start position. For each neuron, trials were randomly split into 2 folds for each condition of the spatial variable, and 32 trials were resampled for each fold in each condition, yielding 224 trials per fold for position decoding and 192 trials per fold for direction decoding. These resampled trials were termed "pseudo-trials". Pseudo-trials were resampled randomly with a different seed for each neuron, so as to remove trial-by-trial correlations between neurons in the same session.

We included single neurons from sessions that had at least 8 trials for each of the 7 start positions, 6 directions and 7 target positions (1197 neurons, 99 sessions, 3 rats). In each pseudopopulation, neurons were also resampled with a different seed, so that only about 63.2% of the neurons were included in each pseudopopulation. We generated 100 pseudopopulations for the decoding each spatial variable.

The error of decoding was defined as the Euclidean distance between true and predicted coordinates. The goodness of prediction was measured as the mean error between the predicted and the actual coordinates across all the pseudo-trials in that pseudo session. For start and target position, we removed the trials with [0 0] coordinates from the quantification of decoding accuracy. This is because predictions from unsuccessful decoding tend to cluster at the [0 0] coordinates, so decoding accuracy for the [0 0] coordinates in positions (i.e. the central port) will always seem "good".

Decoding was performed with multivariate regression models (*mvregress* function in Matlab Statistics and Machine Learning Toolbox). The spike counts of the training set and the test set were first combined, z-scored, and applied to principle component analysis, and then split again for training and decoding. We found that the cross-validated decoding accuracy was the best when including only the first 4 PCs when decoding start position from neural activity in the "pre-cue" time window (Figure S7G), so we used 4 PCs for all subsequent pseudopopulation decoding.

In Figure 3B, multivariate regression models were trained and tested with spike counts in 3 key time window as indicated in the plot (see "Key time windows" section for details). The plots showed the result from one example pseudo-session, the same pseudo-session as the first row in Figure 3C. In Figure 3C, the time windows were 300 ms width with 50 ms step, causally aligned to the go sound. Each row was a different pseudo-session. Each row of the heat map was a pseudopopulation with a different resampling of neurons, and the color indicated the mean error with two-fold cross-validation. In Figure 3D, the mean error of one

spatial variable was subtracted by the mean error of another spatial variable in each pseudopopulation, termed the Delta error, which represented the relative goodness of decoding between the two spatial variables. The distribution of Delta errors of the 100 pseudopopulations was then illustrated with the mean and standard deviation.

In FIGURE 3E, models were trained at one time window and tested at another also with two-fold cross validation. In the last panel, a multivariate regression model was trained to decode start position coordinates from the start position pseudopopulation data, and then tested its decoding of target position coordinates from the target position pseudopopulation data. The heat-map showed the average of mean errors across 100 pseudopopulations. P values for multiple comparisons were obtained with extreme pixel-based permutation test (Cohen, 2014). One dummy heat-map was generated for each pseudopopulation by shuffling trial labels, and the maximum of each heat map was gathered to construct a null distribution. The white contours enclosed the area in which the averaged mean error was smaller than the minimum value of this distribution ($p < 0.01$).

Posture analysis

Video tracking and body posture estimation Videos were acquired at 30 frames per second with one Raspberry Pi Camera at the top of the rig.

We estimated the coordinates of the head, the ears and the hip of the rat in the video frames using DeepLabCut (Mathis et al., 2018). Coordinates were in the unit of pixels. The head position was approximated as the Intan chip plugged on the animal's head. The body angle was estimated as the angle between the head and the tail relative to the horizontal direction in the video frame. For continuous sampling of body coordinates aligned to task events, we linearly interpolated to achieve the sampling rate of 100 frames per second.

The body angle was defined as the angle between the head-hip axis and the horizontal axis. In FIGURE 7B,D&F, the body angle was at the interpolated frame 150 ms after the nose arrived at the start port.

Posture tuning In FIGURE 7G, we fit the body angle as a linear mixed effect model of start position and movement direction:

$$\theta \sim 1 + direction + (1|position) \tag{5}$$

where θ represents the continuous variable of the body angle at each interpolated video frames aligned to the go sound. *Direction* was a 6-level categorical variable of the movement directions. *Position* was a 7-level categorical variable for the start positions. This design allowed the inclusion of all possible trials whereas factoring out the effect of start position on body angle.

In FIGURE 7B-D, we compared the fitting of neural spike rates to two Generalized Linear Mixed Effect models of Poisson link function:

$$Spikes \sim 1 + \theta + (1|Position) \tag{6}$$

and

$$Spikes \sim 1 + Position + (1|\theta) \quad (7)$$

Here θ was a 3-level categorical variable of the body angle from the median frame in this time window. Each level was a group of equal number of trials with body angles that were small, medium or large within the session. Modeling the body angle as a categorical variable allowed for arbitrary tuning functions for body angles. *Spikes* were the spike counts in the $0 \sim 300$ ms time window aligned to the arrival of the start port. *Position* was the 7-level categorical variable of the start position. For each neuron, the cross-validated R^2s and the log-likelihoods of spike count observation under each model were calculated with leave-one-out cross validation.

Surrogate data

To verify the reliability of our model comparison approaches, we performed the same analysis on surrogate data which were known to have specific spatial tuning as on real neurons. These surrogate spike counts were generated to have specific spatial tuning, to match the overall firing rate of a real neuron, and as if sampled from a real session.

To generate surrogate data matching one neuron but tuned to a specific model, we first fit the spike counts of the real neuron to that model, so as to obtain the predicted firing rate at each trial condition. We then randomly selected a real session and used a Poisson random process to generate spike counts that followed these predicted firing rates. In the surrogate data for GLMs, we only used neurons that have significant selectivity to that GLM to generate surrogate data. In surrogate data to compare between pure and mixed selectivity models, we used all neurons to generate surrogate data.

Following stages of analyses was identical for surrogate data and real data. If our methods were reliable, the surrogate data designed to have a specific functional form would be best fitted by the same functional form.

Code and data sharing

Please visit `https://github.com/erlichlab/fof-visually-guided` to access the code used for analyses and to generate figures. Links to the data are available from the github repository.

Author Contributions

JCE and LL conceptualized and designed the experiments. LL, CM, JCE and JL collected the experimental data. LL, CM and JL curated the spike sorting and contributed to analysis pipelines for electrophysiology. LL analysed the behavioral

and neural data under supervision of JCE. CM extracted tracking information from videos. LL and JCE wrote the manuscript with comments from CM and JL.

Acknowledgements

The authors thank Yingkun Li, Nengneng Gao and Anyu Fang for their assistance with behavioral training and animal care. We thank Hungtu Chen, Yizhou Chen, Xuan Wen, Cequn Wang, Yidi Chen, Josh Moller-Mara, and Sylvain Dubroqua for their help with developing hardware and software for the training facility. In particular Hungtu Chen helped to develop earlier versions of the behavioral task. LLJL and JL were supported by the NYU Shanghai Doctoral Fellowship, Winston Foundation Fund, and BOCO Fund for Science and Research. The research was also supported by grants to JCE from: the Program of Shanghai Academic/Technology Research Leader (15XD1503000); the Science and Technology Commission of Shanghai Municipality (15JC1400104); the 111 Project, Base B16018; the National Natural Science Foundation of China (NSFC); NYU-ECNU Institute of Brain and Cognitive Science at NYU Shanghai. JCE also acknowledges the funders of the Sainsbury Wellcome Centre (The Wellcome Trust and the Gatsby Charitable Foundation).

References

Andersen, R., Bracewell, R., Barash, S., Gnadt, J., and Fogassi, L. (1990). Eye position effects on visual, memory, and saccade-related activity in areas LIP and 7a of macaque. *The Journal of Neuroscience*, 10(4):1176–1196.

Andersen, R., Essick, G., and Siegel, R. (1985). Encoding of spatial location by posterior parietal neurons. *Science (New York, N.Y.)*, 230(4724):456–458.

Andersen, R. and Mountcastle, V. (1983). The influence of the angle of gaze upon the excitability of the light- sensitive neurons of the posterior parietal cortex. *The Journal of Neuroscience*, 3(3):532–548.

Barthas, F. and Kwan, A. C. (2017). Secondary Motor Cortex: Where 'Sensory' Meets 'Motor' in the Rodent Frontal Cortex. *Trends in Neurosciences*, 40(3):181–193.

Bharmauria, V., Sajad, A., Li, J., Yan, X., Wang, H., and Crawford, J. D. (2020). Integration of Eye-Centered and Landmark-Centered Codes in Frontal Eye Field Gaze Responses. *Cerebral Cortex*, page bhaa090. 4 citations (Crossref) [2022-11-03].

Bruce, C. J. and Goldberg, M. E. (1985). Primate frontal eye fields. I. Single neurons discharging before saccades. *Journal of Neurophysiology*, 53(3):603–635. 938 citations (Crossref) [2022-11-03] tex.eprint: 3981231 tex.eprinttype: pmid.

Bush, N. E., Solla, S. A., and Hartmann, M. J. (2016). Whisking mechanics and active sensing. *Current Opinion in Neurobiology*, 40:178–188.

Caruso, V. C., Pages, D. S., Sommer, M. A., and Groh, J. M. (2018). Beyond the labeled line: Variation in visual reference frames from intraparietal cortex to frontal eye fields and the superior colliculus. *Journal of Neurophysiology*, 119(4):1411–1421.

Cassanello, C. and Ferrera, V. P. (2010). Vector subtraction and eye position gainfields in macaque frontal eye field. *Journal of Vision*, 5(8):580–580. 0 citations (Crossref) [2022-11-03].

Cassanello, C. R. and Ferrera, V. P. (2007). Computing vector differences using a gain field-like mechanism in monkey frontal eye field: Eye position signals in FEF. *The Journal of Physiology*, 582(2):647–664.

Chen, T.-W., Li, N., Daie, K., and Svoboda, K. (2017). A Map of Anticipatory Activity in Mouse Motor Cortex. *Neuron*, 94(4):866–879.e4. 136 citations (Crossref) [2022-11-03].

Cohen, M. X. (2014). *Analyzing Neural Time Series Data: Theory and Practice*. The MIT Press.

Cohen, Y. E. and Andersen, R. A. (2002). A common reference frame for movement plans in the posterior parietal cortex. *Nature Reviews. Neuroscience; London*, 3(7):553–62. 00532.

Cowey, A. and Ek, T. B. (1973). CONTRALATERAL 'NEGLECT' AFTER UNILATERAL DORSOMEDIAL P R E F R O N T A L LESIONS 1N RATS. *Brain Research*, page 11.

Crowne, D. P. and Pathria, M. N. (1982). Some attentional effects of unilateral frontal lesions in the rat. *Behavioural Brain Research*, 6(1):25–39.

Duan, C. A., Pan, Y., Ma, G., Zhou, T., Zhang, S., and Xu, N.-l. (2021). A cortico-collicular pathway for motor planning in a memory-dependent perceptual decision task. *Nature Communications*, 12(1):2727. tex.ids= duan2021corticocolliculara number: 1 publisher: Nature Publishing Group.

Ebbesen, C. L., Insanally, M. N., Kopec, C. D., Murakami, M., Saiki, A., and Erlich, J. C. (2018). More than Just a "Motor": Recent Surprises from the Frontal Cortex. *The Journal of Neuroscience*, 38(44):9402–9413. 43 citations (Crossref) [2022-11-03].

Erlich, J. C., Bialek, M., and Brody, C. D. (2011). A cortical substrate for memory-guided orienting in the rat. *Neuron*, 72(2):330–343. 212 citations (Crossref) [2022-11-03] tex.eprint: 22017991 tex.eprinttype: pmid tex.pmcid: PMC3212026.

Erlich, J. C., Brunton, B. W., Duan, C. A., Hanks, T. D., and Brody, C. D. (2015). Distinct effects of prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat. *eLife*, 4. 00000.

Guo, Z. V., Inagaki, H. K., Daie, K., Druckmann, S., Gerfen, C. R., and Svoboda, K. (2017). Maintenance of persistent activity in a frontal thalamocortical loop. *Nature*, 545(7653):181–186. 00003.

Guo, Z. V., Li, N., Huber, D., Ophir, E., Gutnisky, D., Ting, J. T., Feng, G., and Svoboda, K. (2014). Flow of Cortical Activity Underlying a Tactile Decision in Mice. *Neuron*, 81(1):179–194. 458 citations (Crossref) [2022-11-03].

Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., and Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436(7052):801–806. 2335 citations (Crossref) [2022-11-03].

Hanks, T. D., Kopec, C. D., Brunton, B. W., Duan, C. A., Erlich, J. C., and Brody, C. D. (2015). Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature*, 520(7546):220–223. 337 citations (Crossref) [2022-11-03] tex.eprint: 25600270 tex.eprinttype: pmid tex.pmcid: PMC4835184.

Inagaki, H. K., Fontolan, L., Romani, S., and Svoboda, K. (2019). Discrete attractor dynamics underlies persistent activity in the frontal cortex. *Nature*, 566(7743):212–217. tex.ids= inagakiDiscreteAttractorDynamics2019a.

Jovalekic, A., Hayman, R., Becares, N., Reid, H., Thomas, G., Wilson, J., and Jeffery, K. (2011). Horizontal biases in rats' use of three-dimensional space. *Behavioural Brain Research*, 222(2):279–288.

Kopec, C. D., Erlich, J. C., Brunton, B. W., Deisseroth, K., and Brody, C. D. (2015). Cortical and Subcortical Contributions to Short- Term Memory for Orienting Movements. *Neuron*, 88(2):367–377. 83 citations (Crossref) [2022-11-03] tex.eprint: 26439529 tex.eprinttype: pmid.

Leonard, C. M. (1969). The prefrontal cortex of the rat. I. Cortical projection of the mediodorsal nucleus. II. Efferent connections. *Brain Research*, 12(2):321–343.

Li, N., Chen, T.-W., Guo, Z. V., Gerfen, C. R., and Svoboda, K. (2015). A motor cortex circuit for motor planning and movement. *Nature*, 519(7541):51–56.

Li, N., Daie, K., Svoboda, K., and Druckmann, S. (2016). Robust neuronal dynamics in premotor cortex during motor planning. *Nature*, 532(7600):459–464. 00035.

Maisson, D. J.-N., Voloh, B., Cervera, R. L., Conover, I., Zambre, M., Zimmermann, J., and Hayden, B. Y. (2022). Widespread coding of navigational variables in prefrontal cortex. preprint, Neuroscience.

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., and Bethge, M. (2018). DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, 21(9):1281–1289.

McCluskey, M. K. and Cullen, K. E. (2007). Eye, Head, and Body Coordination During Large Gaze Shifts in Rhesus Monkeys: Movement Kinematics and the Influence of Posture. *Journal of Neurophysiology*, 97(4):2976–2991. 40 citations (Crossref) [2022-11-03].

Mimica, B., Dunn, B. A., Tombaz, T., Bojja, V. P. T. N. C. S., and Whitlock, J. R. (2018). Efficient cortical coding of 3D posture in freely behaving rats. *Science*, 362(6414):584–589. tex.ids: mimicaEfficientCorticalCoding2018a.

Monteon, J. A., Constantin, A. G., Wang, H., Martinez-Trujillo, J., and Crawford, J. D. (2010). Electrical Stimulation of the Frontal Eye Fields in the Head-Free Macaque Evokes Kinematically Normal 3D Gaze Shifts. *Journal of Neurophysiology*, 104(6):3462–3475.

Neal, B. (2020). Introduction to Causal Inference from a Machine Learning Perspective. Publisher: Unpublished.

O'Keefe, J. and Nadel, L. (1978). *The hippocampus as a cognitive map*. Clarendon Press ; Oxford University Press, Oxford : New York.

Olson, J. M., Li, J. K., Montgomery, S. E., and Nitz, D. A. (2020). Secondary Motor Cortex Transforms Spatial Information into Planned Action during Navigation. *Current Biology*, page S0960982220303547. 18 citations (Crossref) [2022-11-03].

Pouget, A. and Sejnowski, T. J. (1997). Spatial Transformations in the Parietal Cortex Using Basis Functions. *Journal of Cognitive Neuroscience*, 9(2):222–237.

Reep, R., Chandler, H., King, V., and Corwin, J. (1994). Rat posterior parietal cortex: topography of corticocortical and thalamic connections. *Experimental Brain Research*, 100(1).

Reep, R., Corwin, J., Hashimoto, A., and Watson, R. (1987). Efferent Connections of the Rostral Portion of Medial Agranular Cortex in Rats. *Brain Research Bulletin*, 19(2):203–221. 175 citations (Crossref) [2022-11-03].

Reep, R. L., Goodwin, G. S., and Corwin, J. V. (1990). Topographic organization in the corticocortical connections of medial agranular cortex in rats. *The Journal of Comparative Neurology*, 294(2):262–280.

Reinagel, P. (2018). Training Rats Using Water Rewards Without Water Restriction. *Frontiers in Behavioral Neuroscience*, 12:84. 12 citations (Crossref) [2022-11-03].

Rigotti, M., Barak, O., Warden, M. R., Wang, X.-J., Daw, N. D., Miller, E. K., and Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature*, 497(7451):585–590.

Salinas, E. and Abbott, L. (1995). Transfer of coded information from sensory to motor networks. *The Journal of Neuroscience*, 15(10):6461–6474. 175 citations (Crossref) [2022-11-03].

Salinas, E. and Sejnowski, T. J. (2001). Gain Modulation in the Central Nervous System: Where Behavior, Neurophysiology, and Computation Meet. *The Neuroscientist*, 7(5):430–440. 169 citations (Crossref) [2022-11-03].

Schall, J. D. (2009). Frontal Eye Fields. In Squire, L., editor, *Encyclopedia of Neuroscience*, volume 4, pages 367–374. Oxford: Academic Press.

Shenoy, K. V., Sahani, M., and Churchland, M. M. (2013). Cortical Control of Arm Movements: A Dynamical Systems Perspective. *Annual Review of Neuroscience*, 36(1):337–359. tex.ids: shenoyCorticalControlArm2013a.

Siniscalchi, M. J., Phoumthipphavong, V., Ali, F., Lozano, M., and Kwan, A. C. (2016). Fast and slow transitions in frontal ensemble activity during flexible sensorimotor behavior. *Nature Neuroscience*, 19(9):1234–1242.

Sinnamon, H. M. and Galer, B. S. (1984). Head movements elicited by electrical stimulation of the anteromedial cortex of the rat. *Physiology & Behavior*, 33(2):185–190. tex.ids: sinnamon1984head.

Sommer, M. A. and Wurtz, R. H. (2008). Brain Circuits for the Internal Monitoring of Movements. *Annual Review of Neuroscience*, 31(1):317–338. 00324.

Sul, J. H., Jo, S., Lee, D., and Jung, M. W. (2011). Role of rodent secondary motor cortex in value-based action selection. *Nature Neuroscience*, 14(9):1202–1208. 145 citations (Crossref) [2022-11-03].

Taube, J., Muller, R., and Ranck, J. (1990). Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *The Journal of Neuroscience*, 10(2):420–435. 1038 citations (Crossref) [2022-11-03].

Wang, C., Chen, X., and Knierim, J. J. (2020). Egocentric and allocentric representations of space in the rodent brain. *Current Opinion in Neurobiology*, 60:12–20. 51 citations (Crossref) [2022-11-03].

Wang, W., Chan, S. S., Heldman, D. A., and Moran, D. W. (2007). Motor Cortical Representation of Position and Velocity During Reaching. *Journal of Neurophysiology*, 97(6):4258–4270.

Wilber, A. A., Clark, B. J., Forster, T. C., Tatsuno, M., and McNaughton, B. L. (2014). Interaction of Egocentric and World-Centered Reference Frames in the Rat Posterior Parietal Cortex. *Journal of Neuroscience*, 34(16):5431–5446.

Wise, S. P. (2008). Forward frontal fields: phylogeny and fundamental function. *Trends in Neurosciences*, 31(12):599–608.

Yamawaki, N., Radulovic, J., and Shepherd, G. M. (2016). A Corticocortical Circuit Directly Links Retrosplenial Cortex to M2 in the Mouse. *The Journal of Neuroscience*, 36(36):9365–9374. 77 citations (Crossref) [2022-11-03] tex.eprint: 27605612 tex.eprinttype: pmid tex.ids: yamawaki2016Corticocortical tex.pmcid: PMC5013186.

Yang, W. (2022). Thalamus-driven functional populations in frontal cortex support decision-making. *Nature Neuroscience*, 25:39.

Zhu, X., Bao, C., Moller-Mara, J., Li, J., Dubroqua, S., and Erlich, J. C. (2021). The rat frontal cortex encodes a value map for economic decisions under risk.

Zipser, D. and Andersen, R. A. (1988). A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature*, 331(6158):679. 746 citations (Crossref) [2022-11-03].

1 Supplemental Figures

Figure	Panel(s)	N	Criteria
FIGURE 3	A	1224	Single neurons in 4 animals that had at least 8 trials for each condition. 771 neurons from reference sessions, 426 from distance sessions, 27 from rat 2147.
FIGURE 3	B-E	1197	Single neurons that had at all 6 directions, and at least 8 trials in each group of trial conditions. 771 neurons from reference sessions, 426 from distance sessions. 3 rats.
FIGURE 4	D (black bars)-H	174	Single neurons that have significant selectivity to start port during the "precue" time window and target port during the "arrival" time window(GLM, both $p < 0.05$), among the 1224 neurons in FIGURE 3A.
FIGURE 4	D (white bars)	289	Single neurons that have significant selectivity to either the start port during the "precue" time window or the target port during the "arrival" time window (GLM, only one $p < 0.01$), among the 1224 neurons in FIGURE 3A.
FIGURE 4	I	202	Single neurons that have significant selectivity to start port during the "pre-cue" time window(GLM, $p < 0.01$), among the 1224 neurons in FIGURE 3A.
FIGURE 4	J	209	Single neurons that have significant selectivity to target port during the "arrival" time window(GLM, $p < 0.01$), among the 1224 neurons in FIGURE 3A.
FIGURE 5	A-C	274	Single neurons that have significant selectivity to direction during the"post-cue" or "go" time window (GLM, $p < 0.01$), among the 1224 neurons in FIGURE 3A.
FIGURE 6	E	199	Single neurons that have the average of CVR^2s larger than 0.05. Effectively from 4 rats.
FIGURE 7	D	106	Single neurons that meet two criteria: (1) Having significant start position selectivity during the "pre-cue" time window (GLM, $p < 0.01$). (2) Having video recordings, which resulted in 638 out of 1224 neurons in 58 out of 133 sessions. Effectively from 3 rats.

Table S1. Criteria for the inclusion of neurons in each analysis. Single neurons: SNR>5, in-trial firing rate > 1 Hz, fraction of inter-spike interval less than 2 ms < 1%. All the cells are from sessions where the target cue illuminated during the fixation period.

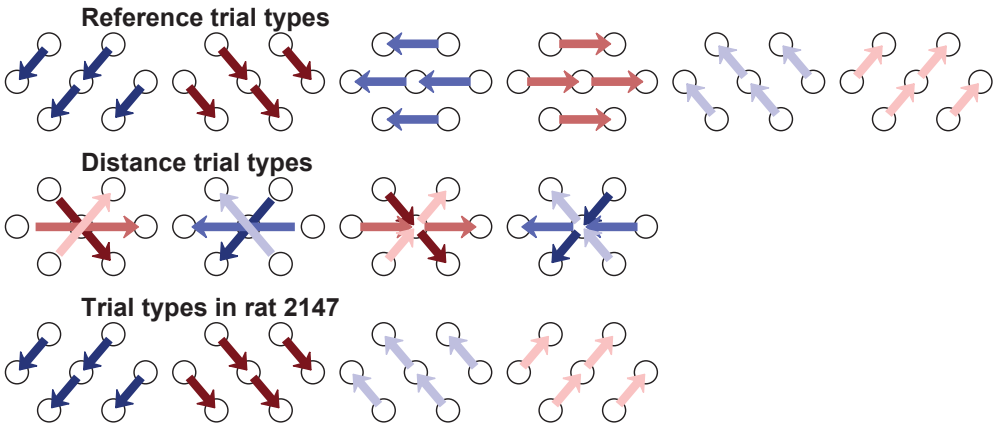

Figure S1. Movement trajectories configurations in each type of sessions. There were 24 configurations in the "reference" sessions. There were 18 configurations in the "distance" sessions, where each movement direction has two short distance configurations and one long distance configuration. In subject 2147, there were 16 configurations.

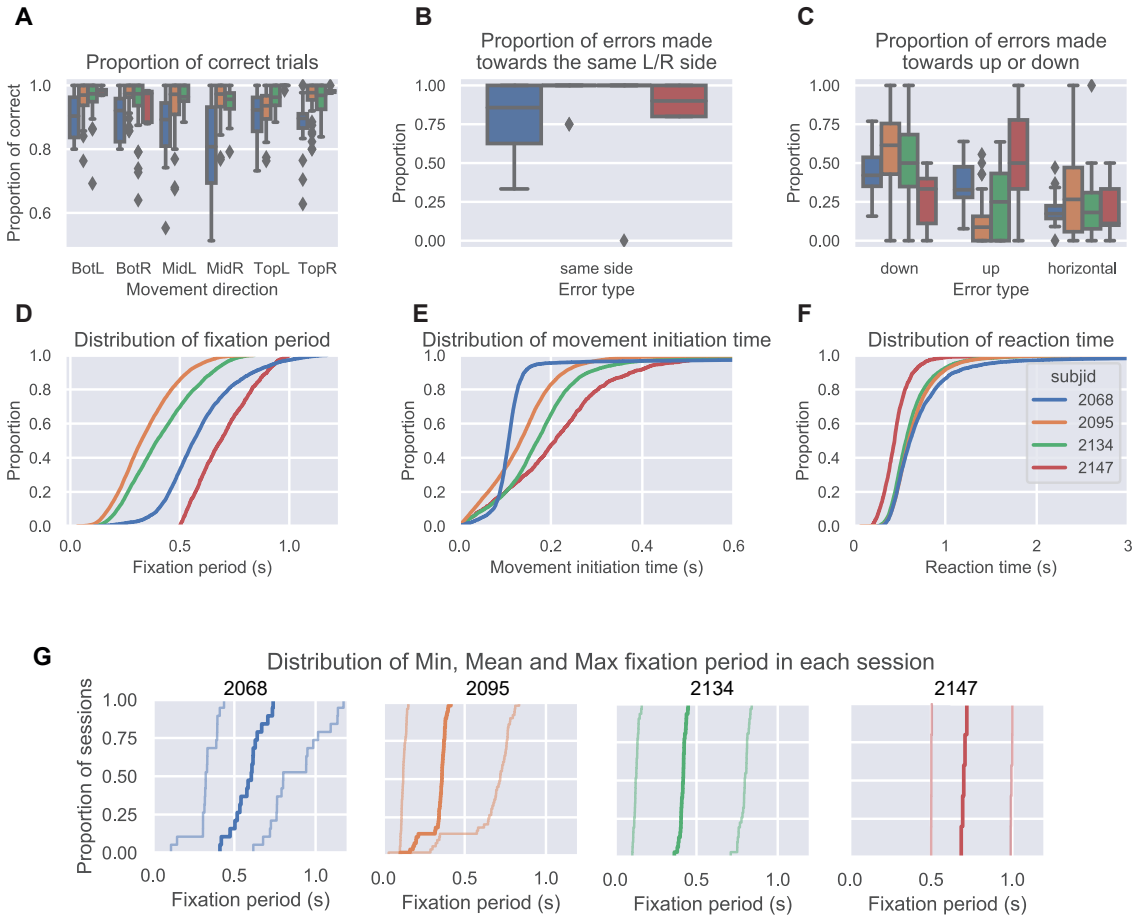

Figure S2. Behavioral measurements of the 4 animals. **A**. The proportion of correct trials among all the finished trials in each movement direction and each animal. The proportion of correct movements among all finished trials in each session was summarized by subjects. **B**. The proportion of errors made into the same side among trials starting from the central port. Errors were made into the same side, and animal 2095 and 2134 almost made no error into the opposite side. The percentage of same side errors were significantly larger than 50% (91.87% $\pm$ 2.20% across all sessions, mean $\pm$ s.e., $p < 10^{-3}$, permutation test for the mean over sessions, n = 288 sessions). **C**. The proportion of errors that were relatively up, down or in the same vertical level, compared to the correct target port position among all error trials. In A-C, points indicate outliers and box plot indicates the quartiles. **D**. The distribution of instructed fixation period in correctly performed trials, defined as the time between nose arrival at the start port and the go sound. **E**. The distribution of movement initiation time, defined as the time between the go sound and the nose departure from the start port. **F**. The distribution of reaction time, defined as the time between the go sound and the nose arrival at the target port. The nose departure or arrival time were detected by the infra-red sensors in each port. **G**. The distribution minimum, mean and maximum of instructed fixation period in each session among correctly performed trials. Each panel was an animal. Thick line was the ECDF of the mean fixation period in each session. Thin lines were the ECDFs of the minimum and maximum fixation periods, respectively.

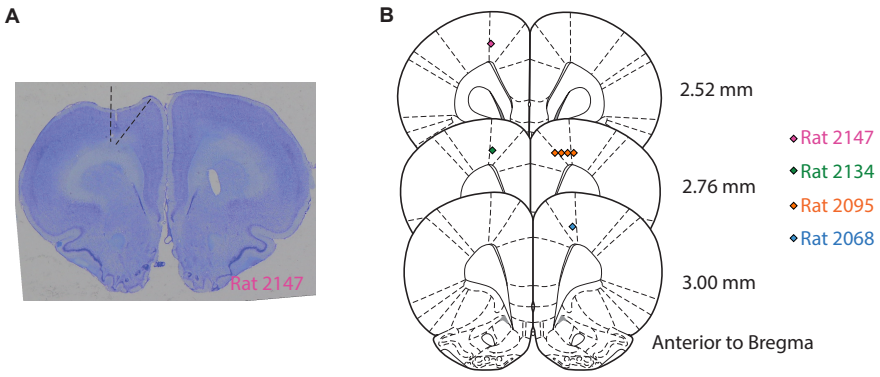

Figure S3. Histology. **A.** The coronal section of an example rat brain (2147) showing the placements of the silicon probes. Dashed lines indicate the estimated area of M2 in this brain section. **B.** The lesion marks matched to the coronal sections of the rat brain atlas(Paxinos and Watson, 2004). Lesions were made at the end of all the recording sessions with 200 uA for 3s relative to the GND. Colored marks indicate the lesion marks in **A**, and colors indicate subject ID. Recording sites in this manuscript varied across subjects in the anterior-posterior axis, ranging from AP 2.5 to 3.5 relative to the Bregma. These recording sites were within an head-orienting-related area based on previous anatomical, lesion, and microstimulation studies (Cowey and Ek, 1973, Crowne and Pathria, 1982, Leonard, 1969, Sinnamon and Galer, 1984), although more anterior compared to a previous electrophysiological study recording from the FOF literature where rats performed an auditory-guided 2AFC head orienting task (Erlich et al., 2011). Despite the variability in recording targets, the main findings in this manuscript were qualitatively consistent across subjects (FIGURE S5). Further experiments will be required to determine if there is a functional gradient over the anterior-posterior axis in the FOF.

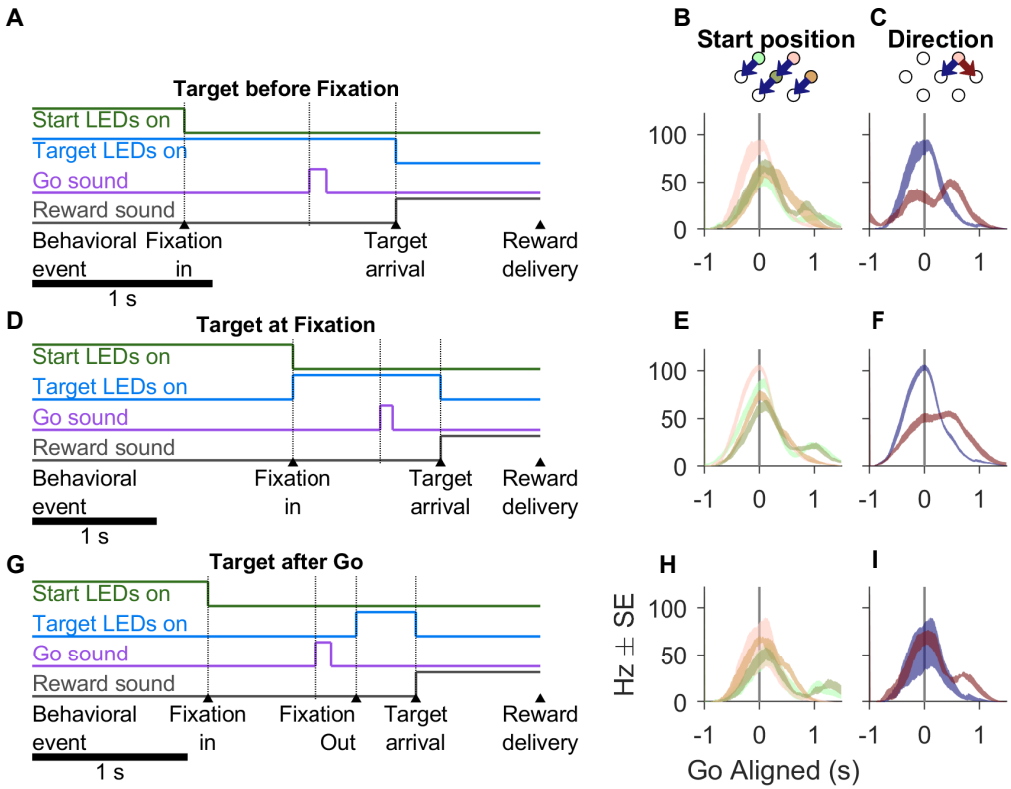

Figure S4. 3 types of trial timelines in rat 2147, and the corresponding PETHs of an example neuron. In rat 2147, there were 2 types of sessions interleaved across days. In one session type, the trial timeline was similar to other animals, where the target LED was always illuminated when the rat fixated into the start port. In the other session type, there were 2 randomly interleaved types of timeline: "target before fixation" and "target after go". The example neuron in B-C and H-I were recorded from the same session, and in E-F from another session. These units were considered as the same neuron based on the recording site, the waveform and the PETHs. **A.** The timeline of the **"target before fixation"** trial class. In these trials, the target LEDs and the start port LEDs illuminated at the same time, thus the target position information was available to the rat before fixation. **B.** The PETHs of an example neuron in "target before fixation" trials that were grouped by the start position and were into the same direction. **C.** The PETHs of the same example neuron as in B, but grouped by the movement direction and were from the same start position. **D-F** Similar to A-C, but for the **"target at fixation"** trials. In these trials, the target LEDs were illuminated when the animal poked into the start port. The neuron was selective to both the start position and the movement direction before the go sound. Note, that the neuron did not encode the movement direction earlier in C than in F, indicating that direction representation was related to the movement planning and not the visual cue. **G-I** Similar to A-C, but for the **"target after go"** trials. In these trials, the target LEDs illuminated after the animal has left the start port, which was detected by the IR sensors. Since the animal didn't know the target port during fixation, the neuron was not selective to the movement direction before the go sound.

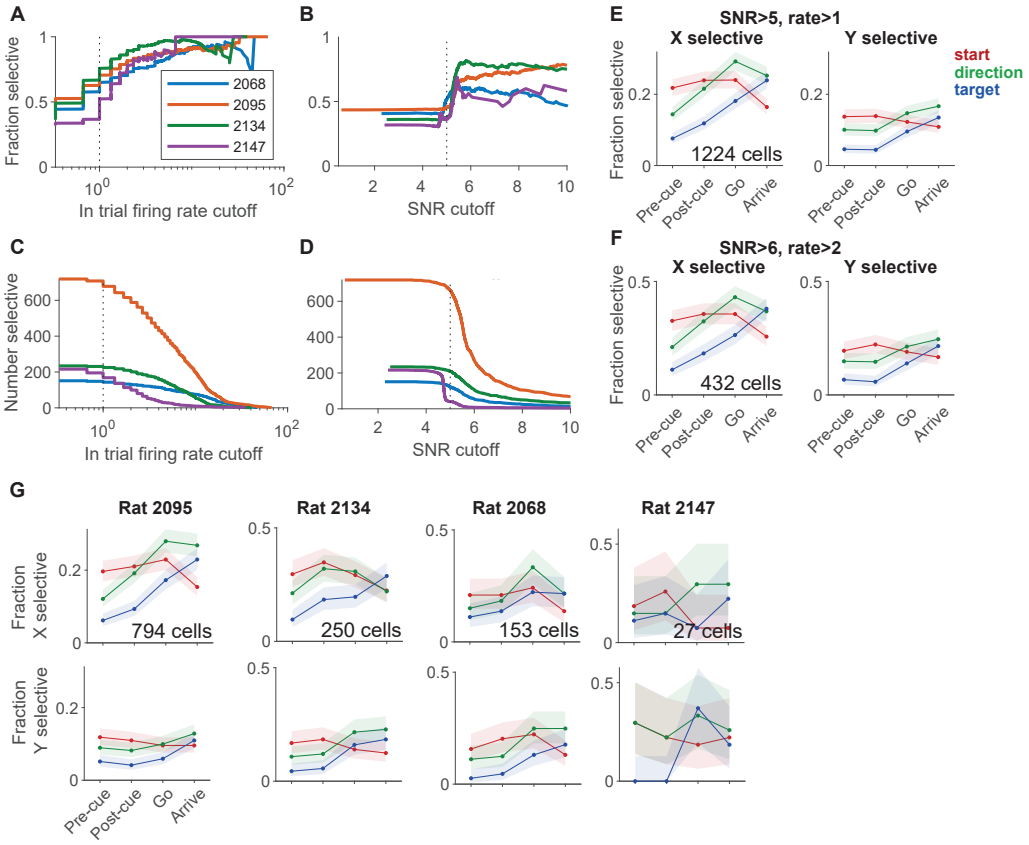

Figure S5. The sequential encoding of start position, direction and target position was consistent across subjects and single neuron selection criteria. In this figure, the selectivity of a neural response to a task variable was measured at each time point for the x and y coordinates separately using the area under the empirical ROC curve (the Wilcoxon-Mann-Whitney U-statistic). The p value was calculated by permutation test, where the null distribution was generated by shuffling the category labels 2000 times. **A.** The fraction of single neurons with spatial selectivity at different cutoff criteria for the in-trial firing rate. The signal-to-noise ratio cut-off was fixed at 5. A neuron is classified as spatially selective if the area under the empirical ROC curve is significantly larger than chance for any one of the 3 spatial variables (start position, direction or target position) during any one of the 4 time windows ("pre-cue", "post-cue", "go" or "arrival")($p < 0.05$, permutation test). **B.** The fraction of single neurons with spatial selectivity at different cutoff criteria for the signal-to-noise ratio (SNR) of the waveform. The cut-off criteria for in-trial firing rate was fixed at 1 Hz ($p < 0.05$, permutation test). **C.** The number of single neurons with spatial selectivity at different in-trial firing rate cutoff. **D.** The number of single neurons with spatial selectivity at different SNR cutoff. for the signal-to-noise ratio of the spike waveform. **E.** The fraction of neurons selective to the x coordinates (left panel) or y coordinates (right panel) of each spatial variable at the 4 time windows based on ROC analysis ($p < 0.05$, permutation test), when the SNR cutoff was 5 and in-trial firing rate cutoff was 1 Hz. Error bars were 95% confidence intervals of the binomial distribution. **F.** Similar to E, but the cut-off criteria is SNR > 6 and in-trial firing rate > 2. **G.** Similar to E, but for each subject.

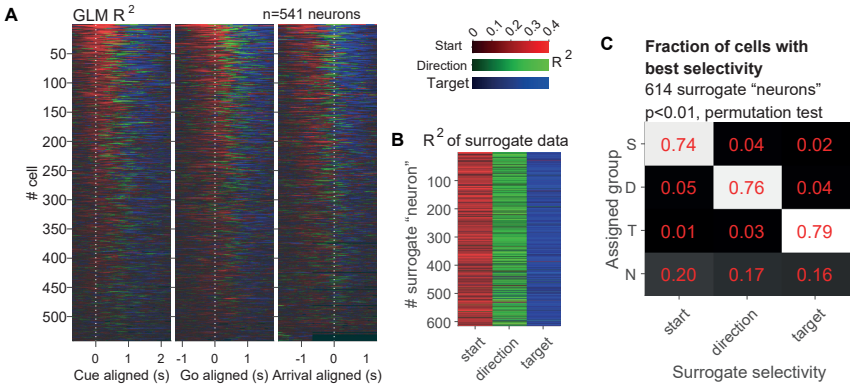

Figure S6. Validation of GLM analysis with surrogate data. **A.** The R^2 of GLMs among neurons that were selective to at least one spatial variable in at least one of the 4 time windows (n = 541). R^2s of GLMs are plotted over the right edge of the window (causal) of 300 ms with 50 ms steps. At each time window, only the GLM with the largest R^2 is shown. Neurons are sorted by the total mass of the R^2s of start position, direction and target position for the 3 alignments respectively. **B.** The R^2s of GLMs in surrogate data with specific spatial selectivity. Each column was a group of surrogate data designed to have start, direction or target selectivity, and each row was a surrogate neuron. The color indicated the R^2 of the model with the maximum R^2, same as in FIGURE 3B. If the surrogate "neuron" had the largest R^2 in the start position model, the corresponding cell was colored red, etc. Note, that the best model according to the R^2s was visually consistent with the surrogate selectivity. **C.** Fraction of cells best selective to each spatial variable in surrogate data with specific spatial selectivity. There were 614 "surrogate neurons" in each column, designed to selective to the start position, the direction and the target position, respectively. The number in each row represented the fraction of "surrogate neurons" assigned as best selective to each model. S, start position model. D, direction model. T, target position model. N, non-selective to any one of the models. Surrogate data was generated to match the overall firing rates and the behavioral sessions of real neurons, but were known to selective to only one spatial variable (Methods). The best selectivity of a surrogate "neuron" was assigned in the same way as in FIGURE 3A. The assignment of "best selectivity" was correct in around 75% of surrogate "neurons".

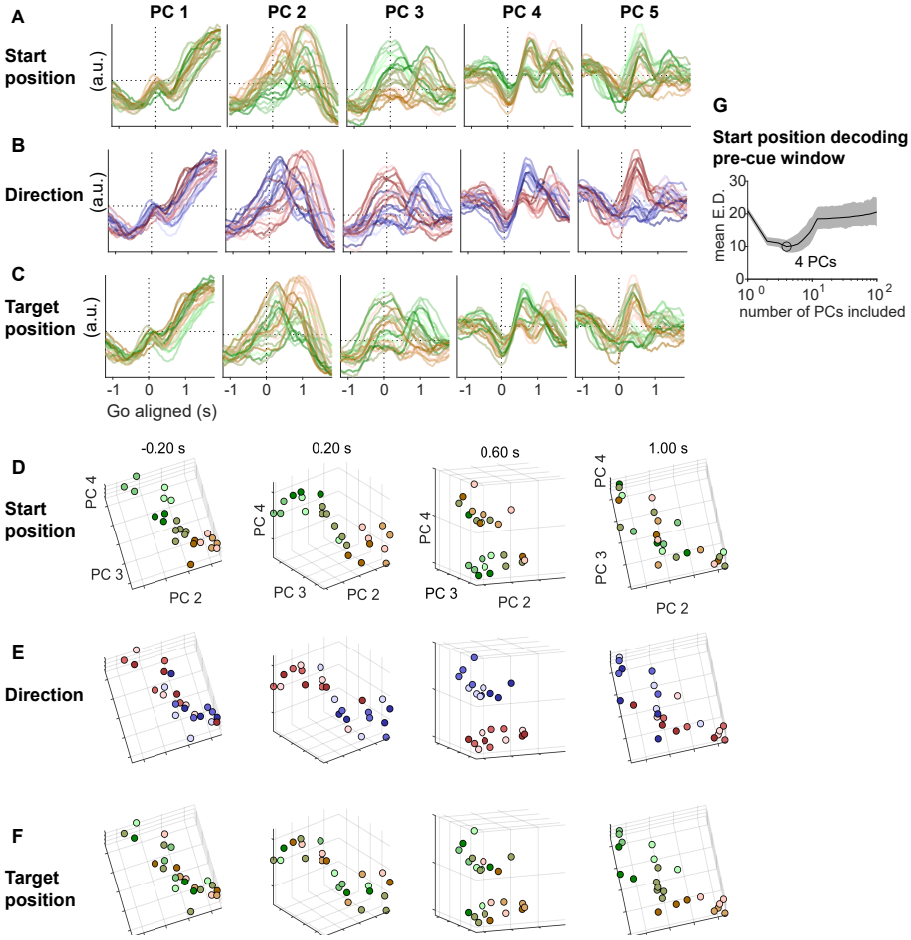

Figure S7. Principle components across time and trial types in the FOF population. **A.** The first 5 principle components of FOF population activity in each movement configuration across time aligned to the go sound, colored by the start position. The population was consisted of neurons with at least 8 trials in each of the 6 directions, 7 start positions and 7 target positions in "reference" sessions (771 neurons). Principle components were computed from a matrix of N by $C \times T$, where N was the number of neurons, C was the number of trial configurations, and T was the number of time windows. **B-C.** Similar to A, but the PCs were colored by the direction and the target position, respectively. **D-F.** The coefficients of PC 2, PC 3 and PC 4 in each movement configuration at different time windows aligned to the go sound, colored by the start position (D), the direction (E) and the target position (F), respectively. The coefficients were rotated at each time window to aid visualization. In early stages of the trial (e.g. -0.20 s aligned to the go sound), trials with different start positions were well separated. As the movement initiates, direction became well separated in these PCs (0.20 s and 0.60 s). In later phases of the trial, the target position was well separated in these PCs (1.00 s). **G.**Number of principle components included versus pseudopopulation decoding accuracy. Mean error for start position decoding using spike count data in the pre-cue time window, among the same neurons and using same methods as in FIGURE 3 C, but including different numbers of principle components for the decoding. Thin line and shaded areas indicated the mean and the 5% and 95% intervals, as a result of re-sampling neurons with replacement for constructing the pseudopopulation.

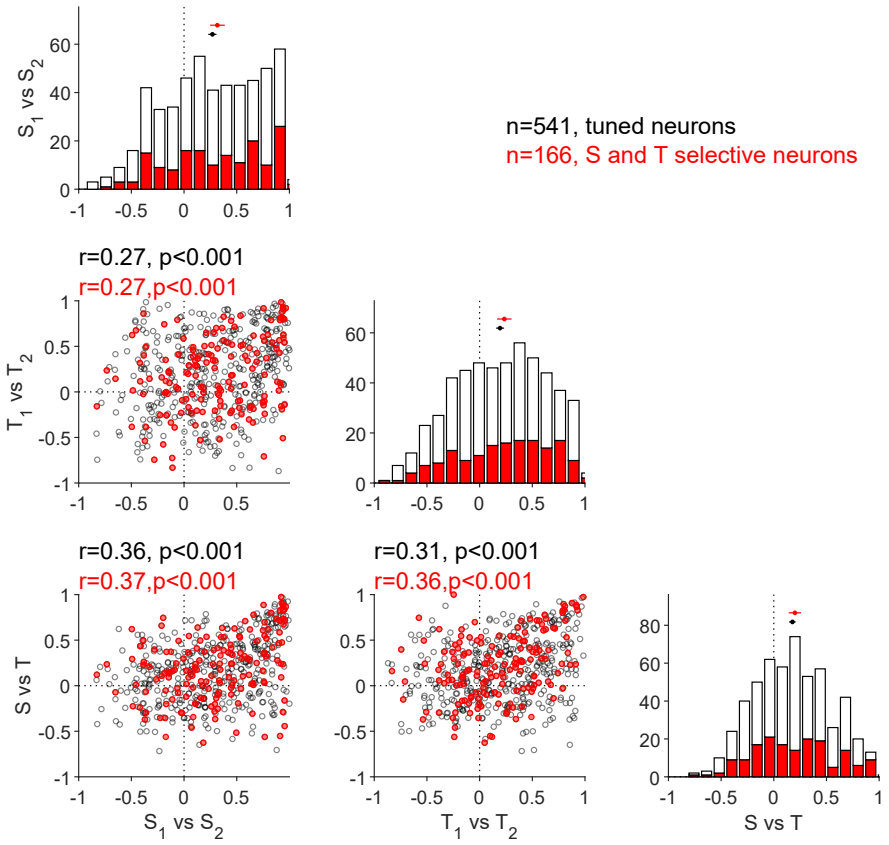

Figure S8. Start-target tuning correlation, start-start tuning correlation, and target-target tuning correlation were correlated. **Diagonal.** The three panels illustrate the number of neurons with different levels of start tuning correlation (at the pre-cue time window), target tuning correlation (at the arrival time window), or the start-target tuning correlation (pre-cue time window for start and arrival time window for target), respectively. The white bars indicate all the neurons with any spatial selectivity in any one of the "pre-cue", "post-cue", "go" or "arrival" time windows ($p < 0.05$, n= 808). The black bars indicate the neurons that were selective to the start position during the "pre-cue" window as well as the target position during the "arrival" window ($p < 0.05$, n=174). The black dot and horizontal line at the top of a panel indicate the mean ± 95% CI of the correlation coefficients among all the cells. **Scatter plots.** The tuning correlation of one pair of spatial variables versus the other. Black circles indicate all the spatially selective cells, matching white bars in the diagonal panels (n=541). Red circles indicate start and target selective cells, matching black bars in the diagonal panels (n=174). Texts at the top of each panel were the Spearman rank correlation coefficients of one *tuning correlation* versus the other, and text color matches the dot color. Start-start tuning correlation, target-target tuning correlation and start-target tuning correlation were significantly correlated with each other, i.e. neurons with consistent position selectivity tend to also be consistent between start and target selectivity.

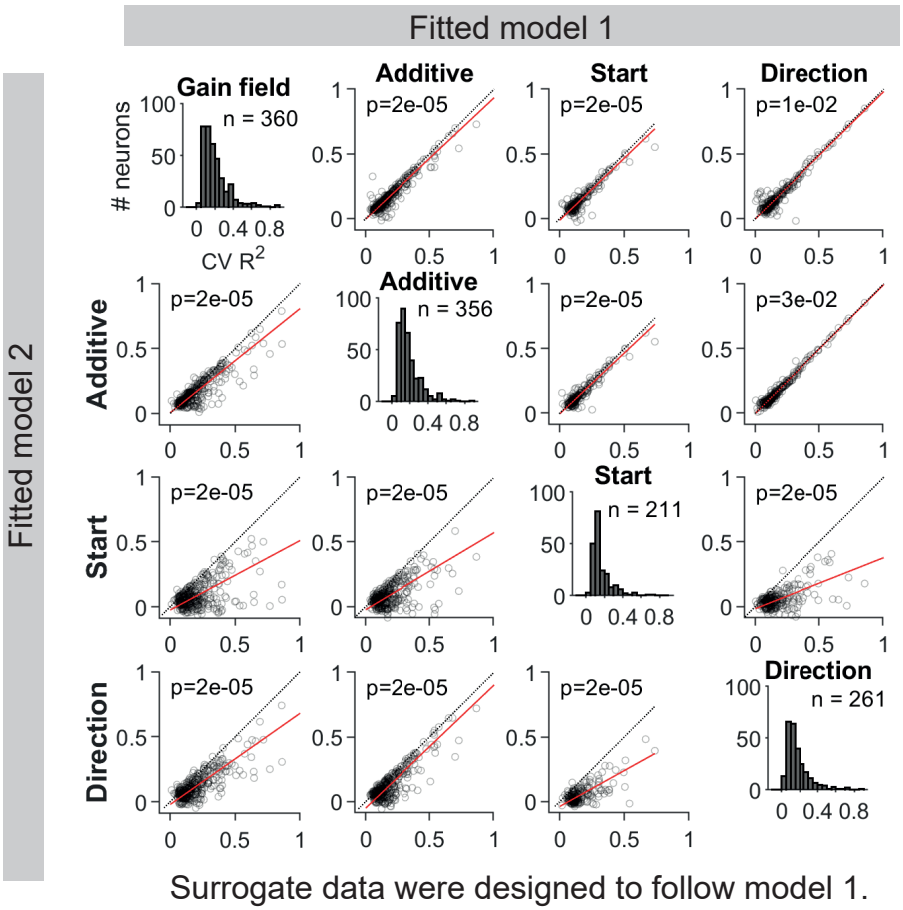

Figure S9. Detection of mixed selectivity with surrogate dataset. Scatter plots. Each point is one surrogate "neuron" in each panel. The position on the x and y axes indicates the cross-validated R^2s of one model versus the other. In each column of panels, the spike counts of surrogate "neurons" were designed to follow the same model as the fit model in the x axis. To generate the surrogate spike counts that followed the specific model, we fit the real neurons (n = 1224) to that model, and then generated Poisson distributed spike counts according to the predicted mean firing rates in each trial condition. We fit the surrogate spike counts to the spatial tuning models as in FIGURE 6, obtained 20-fold cross-validated R^2s, and only plotted those surrogate "neurons" whose average of R^2s in the 4 models were larger than 0.05. The number of included surrogate "neurons" in each surrogate dataset was indicated in the histograms. The p value in each panel was calculated with permutation test against the null hypothesis that the mean difference between the Fisher-z transformed R^2s of the x axis model and the y axis model was not significantly different from zero (permuted 2000 times, n indicated in each panel). In each panel except the diagonal panels, the mean difference across surrogate "neurons" was significantly larger than 0, indicating that the model comparison with cross-validated R^2s well captured the true functional form of spike count modulation. Notably, the method distinguished between the gain field and the additive model.

Histograms. The distribution of cross-validated R^2s for the 4 models, among surrogate "neurons" whose average of R^2s in the 4 models were larger than 0.05 in each dataset.