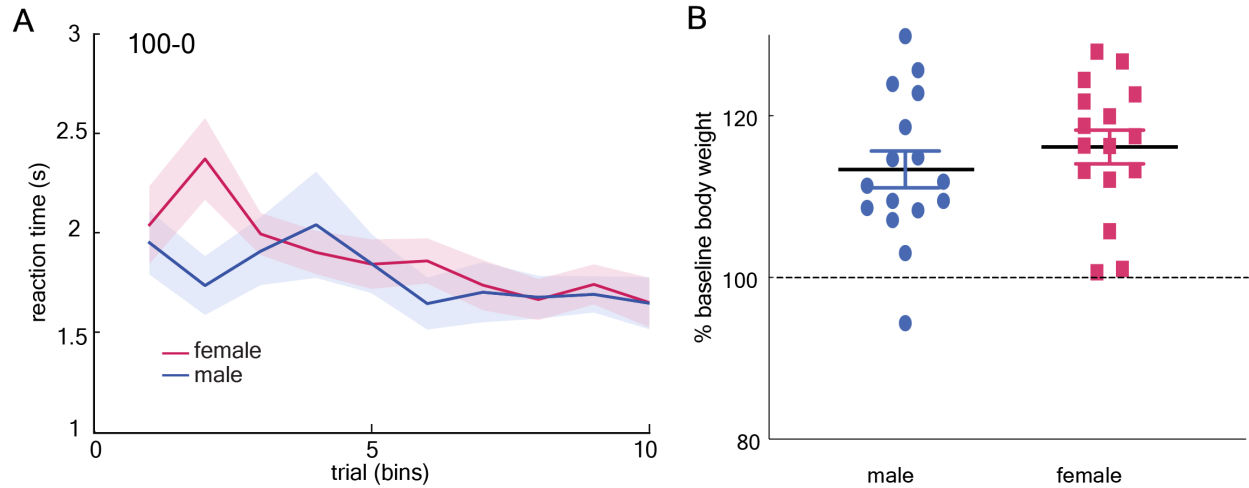
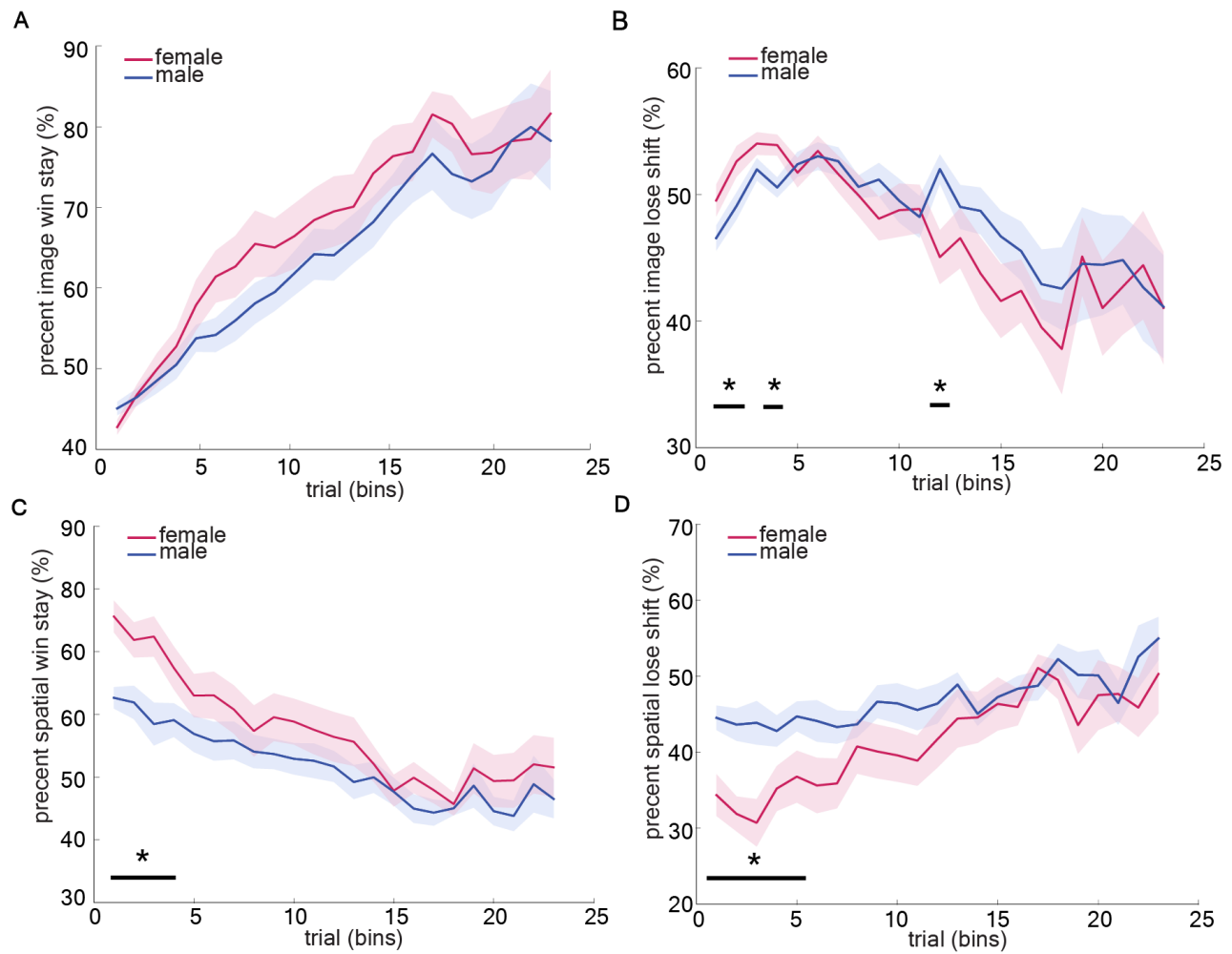


**Figure S1. Individual PC2 plots. Related to Figure 2e.** PC scores for individual male and female animals for PC2. Data shown as bins of 150 trials.

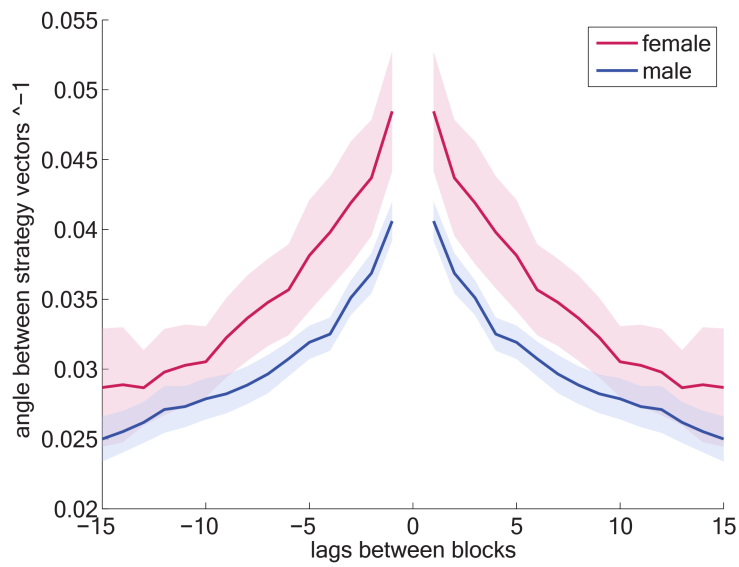


**Figure S2. No sex difference in reaction time across sexes in a deterministic schedule (100-0) and average percent free feeding weight during all behavioral tasks. Related to Figure 3a.** A) There was no significant difference in reaction time between sexes (GLM, main effect of sex,  $\beta_1 = -1.79$ ,  $p = 0.07$ ; interaction term,  $\beta_3 = 1.239$ ,  $p = 0.22$ ). Reaction time of both sexes decreased over time (main effect of number of trials,  $\beta_2 = -3.84$ ,  $p < 0.0002$ ). Data shown as bins of 50 trials. Graphs depict mean  $\pm$  SEM. B) The average percent free feeding weight during behavioral tasks is 113% for males and 116% for females and was not significantly different across sex ( $t(30) = 0.899$ ,  $p = 0.38$ ). It is unlikely that the higher acquisition rate in females was due to different motivations or hunger levels.

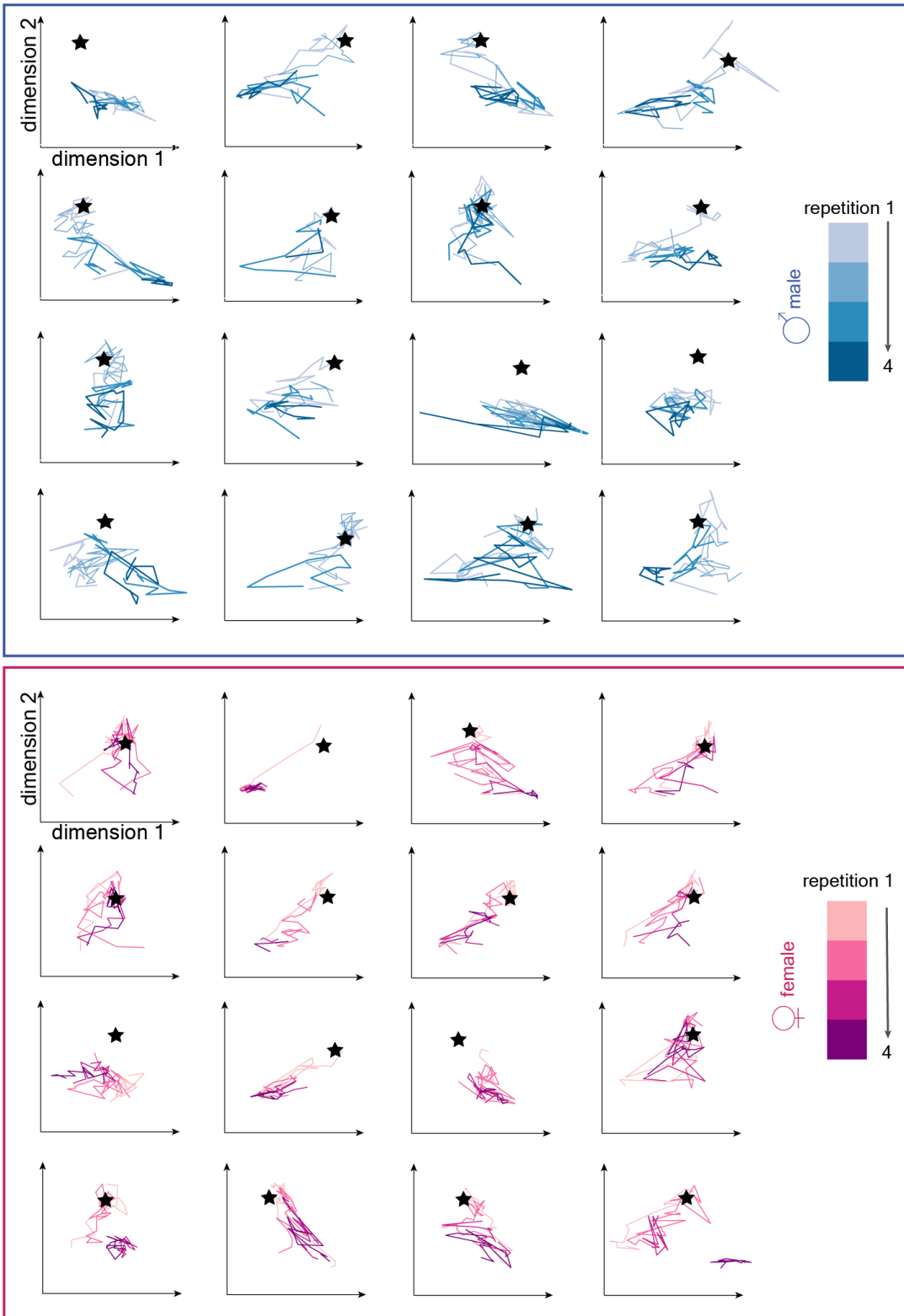


**Figure S3. Win-stay lose-shift in image and spatial dimension separately in both sexes.**

**Related to Figure 4a.** A) Percent win-stay in the image dimension across sexes over time. There was no significant sex difference in percent image win stay and both sexes image win stayed more across trial bins (GLM, main effect of number of trials,  $\beta_2 = 0.02$ ,  $p < 0.0001$ ). B) Percent lose-shift in the image dimension across sexes over time. There was a main effect of number of trials and an interaction effect (GLM, main effect of number of trials,  $\beta_2 = -0.01$ ,  $p < 0.0001$ ; interaction effect,  $\beta_3 = -0.003$ ,  $p < 0.029$ ). C) Percent win-stay in the spatial dimension across sexes over time. Females stayed more on the same side when rewarded (GLM, main effect of sex,  $\beta_1 = 0.104$ ,  $p < 0.0001$ ; main effect of number of trials,  $\beta_2 = 0.008$ ,  $p < 0.0001$ ; interaction effect,  $\beta_3 = -0.004$ ,  $p < 0.002$ ). D) Percent lose-shift in the spatial dimension across sexes over time. Females shifted more to the other side when not rewarded (GLM, main effect of sex,  $\beta_1 = -0.09$ ,  $p < 0.0001$ ; main effect of number of trials,  $\beta_2 = -0.01$ ,  $p < 0.0001$ ; interaction effect,  $\beta_3 = -0.003$ ,  $p < 0.029$ ). Data shown as bins of 150 trials. \* indicates  $p < 0.05$ . Graphs depict mean  $\pm$  SEM.



**Figure S4. Inverse angle between choice patterns across blocks as a function of the time between blocks. Related to Figure 4D.** Nearby blocks are more similar to each other in both males and females, but the block-by-block strategies become more distinct as blocks are further apart in both males and females.



**Figure S5. Individual MDS plots. Related to Figure 4f.** Each graph depicts a different animal, with the colors representing repetitions of the task. The star represents the optimal strategy for each projection. Data shown as bins of 150 trials.

<i>Direct effect</i>	$\beta l$	<i>p value</i>	
NAc → PC2	0.72	0.022	
PFC → PC2	0.75	0.019	
<i>Mediator (sex) effect</i>	$\alpha$	<i>p value</i>	
NAc → sex	0.20	0.024	
PFC → sex	0.26	<0.004	
<i>Indirect effect</i>	$\beta'$	<i>p value</i>	
NAc → PC2	0.38	0.20	
PFC → PC2	0.33	0.31	
<i>Mediation effect</i>	$\alpha\beta'$	<i>z score</i>	<i>p value</i>
NAc → PC2	0.38	1.836	<0.039
PFC → PC2	0.33	2.035	<0.026

**Table S1. Mediation coefficients and p values of mediation analysis. Related to Figure 5f.**