Science Advances
AAAS

Supplementary Materials for

The neural basis of delayed gratification

Zilong Gao, Hanqing Wang, Chen Lu, Tiezhan Lu, Sean Froudist-Walsh, Ming Chen, Xiao-Jing Wang*, Ji Hu*, Wenzhi Sun*

*Corresponding author. Email: xjwang@nyu.edu (X.-J.W.); huji@shanghaitech.edu.cn (J.H.); sunwenzhi@cibr.ac.cn (W.S.)

Published 1 December 2021, *Sci. Adv.* **7**, eabg6611 (2021)
DOI: 10.1126/sciadv.abg6611

The PDF file includes:

Figs. S1 to S8

Other Supplementary Material for this manuscript includes the following:

Movies S1 and S2
Other file

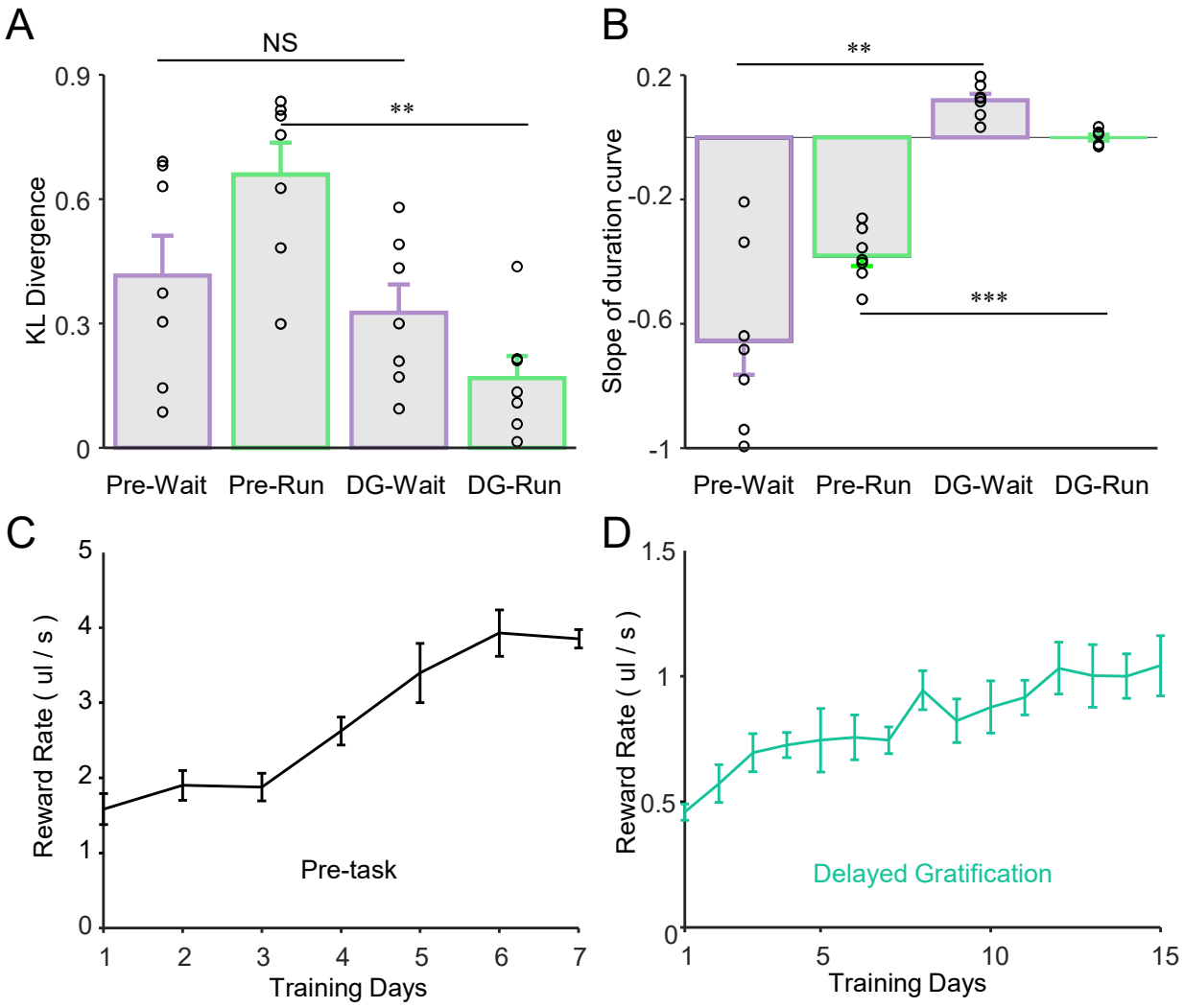

Fig. S1. The change of behavioral performance and reward rate during behavioral training. (**A**) The KL divergence of the distribution of waiting duration and running duration between the first day and the last day in pre-task and delayed gratification (DG). The KL divergence of waiting duration have no difference between pre-task and delayed gratification ($p = 0.53$, paired t-test), but KL divergence of running duration in pre-task was significantly ($p = 0.002$, paired t-test) different with the KL divergence of running duration from (**B**) The slopes of curves (Fig.1 E and Fig.1 H) that waiting duration and running duration changed during behavioral training. The waiting duration and running duration in pre-task were both changed more quickly ($p = 0.004$, waiting duration; $p<0.001$, running duration, paired t-test) than those from delayed gratification. (**C**-**D**) The reward rate both increased in pre-task training (**C**) and delayed gratification training (**D**). All error bars represent the s.e.m..

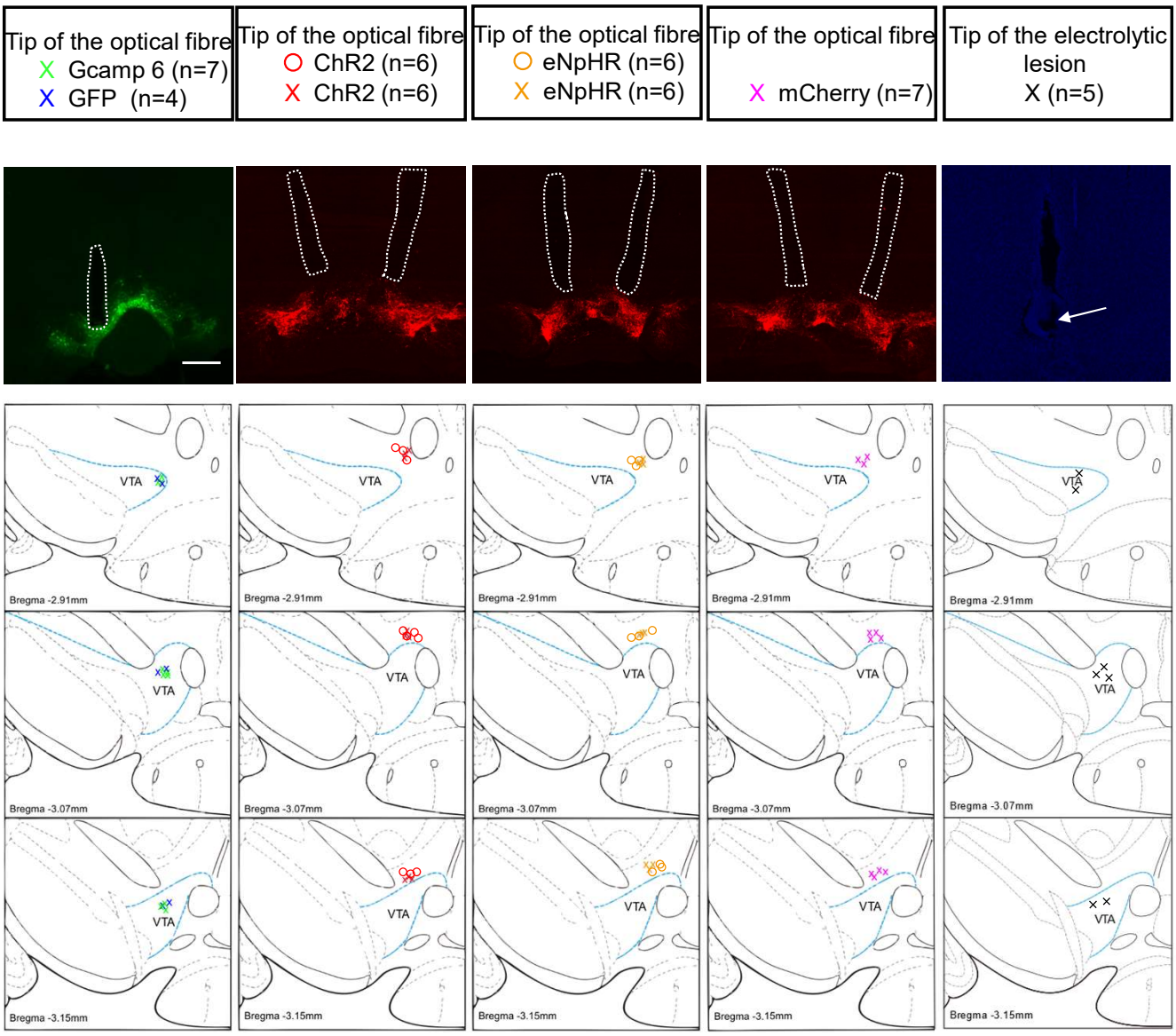

Fig. S2. The histology and tip positions of optical fibre in GCaMP6m (green, n = 7), GFP (blue, n = 4), ChR2 (red, n = 12), ENpHR (yellow, n = 12), mCherry (magenta, n = 7) and electrolytic lesion (black, n = 5) were shown as coordinates in the mouse brain atlas. The circle markers showed the fiber placements from the mouse with optogenetic manipulation in the early training process. Scale bar 500um.

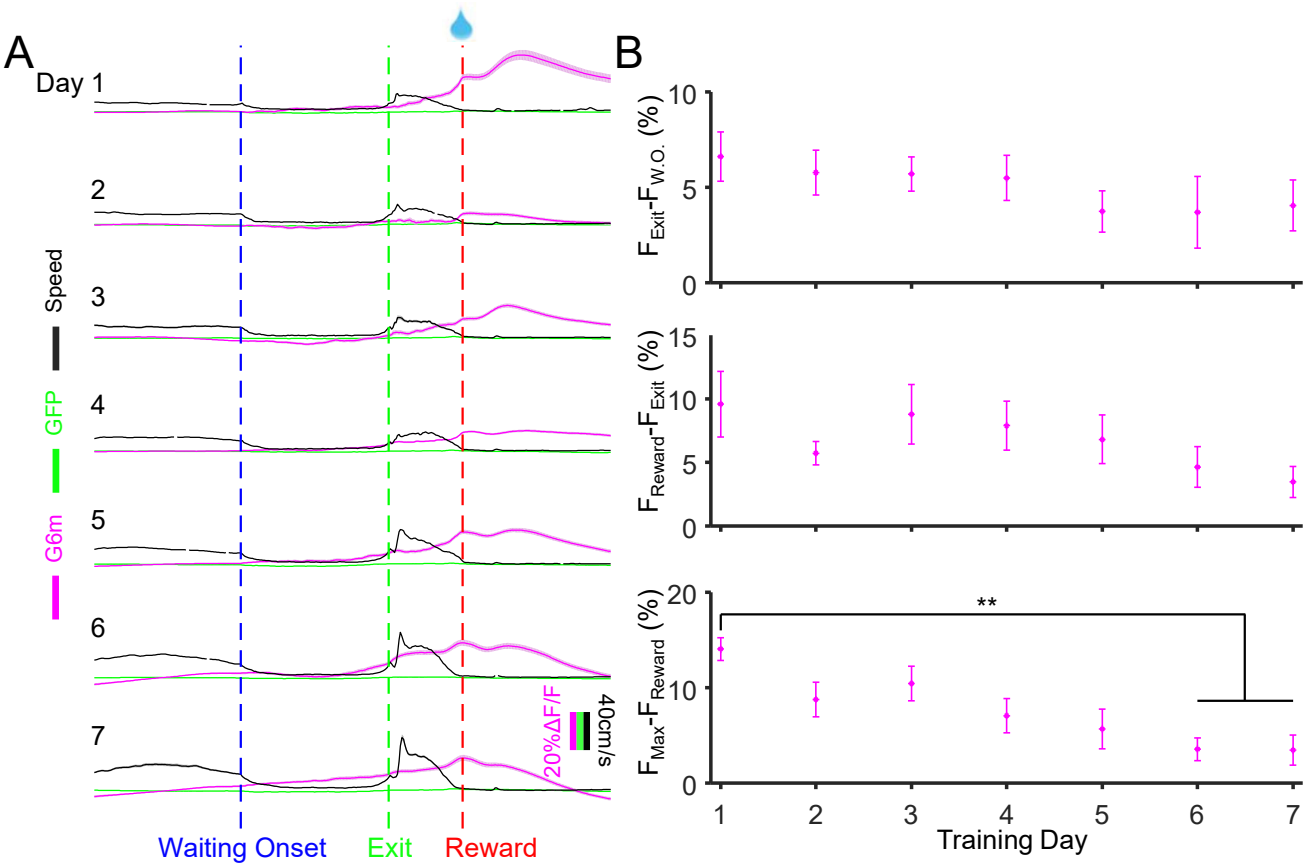

Fig. S3. The calcium and GFP signals of VTA DA neurons in behavioral tasks. (**A**) The GFP (green), calcium signals (magenta) of VTA DA neurons, and mouse running speed (black) in 7 days of pre-training. (**B**) The calcium signals of VTA DA neurons were changing with pre-training progress. The max calcium signals after the mouse received water rewards significantly decreased on days 6&7 ($p<0.01$, Friedman test). All error bars represent the s.e.m..

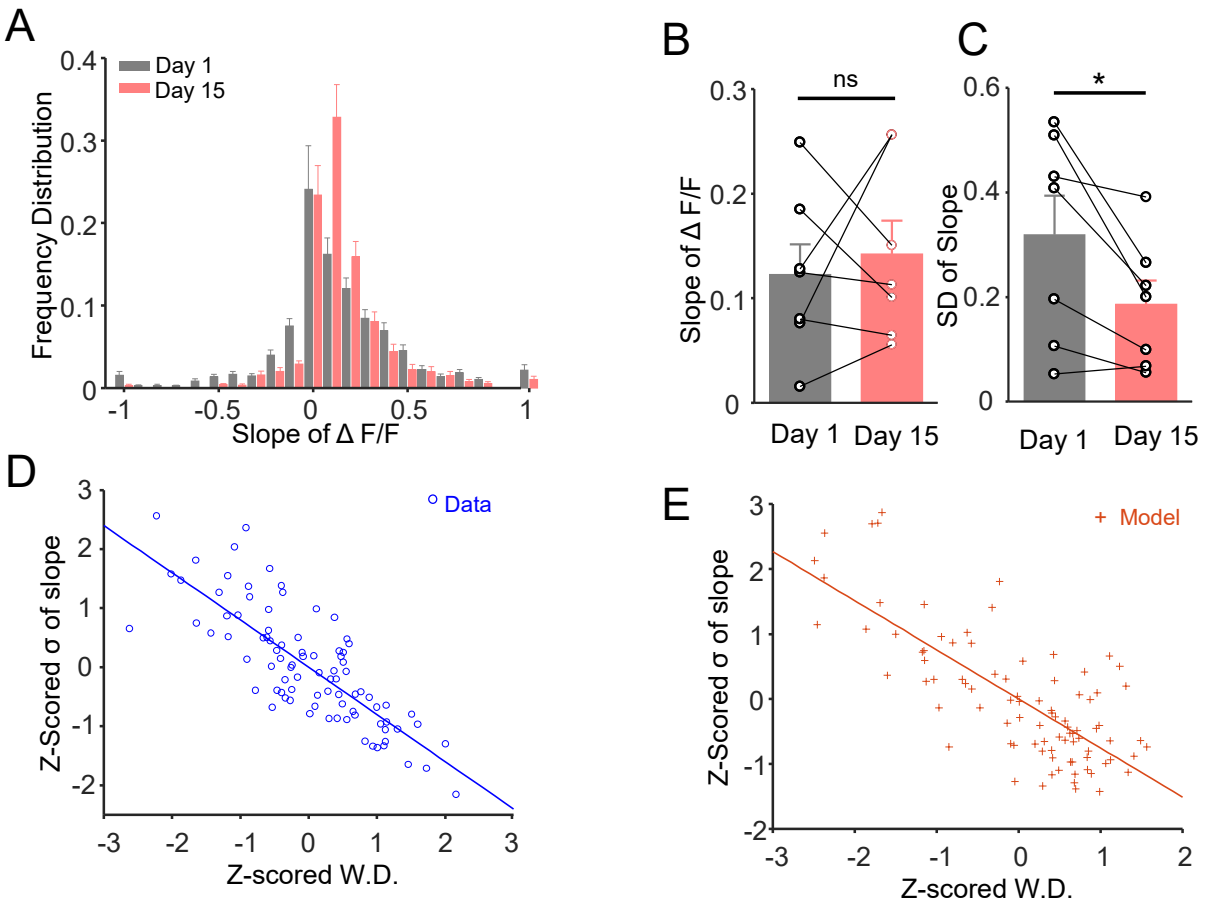

Fig. S4. The ramping dopamine activity became more stable with delayed gratification training. (**A**) The Frequency distribution of the slope of Z-scored ΔF/F during waiting (gray: day 1, red: day 15, n = 7). (**B**) The slope of ΔF/F had no difference between day 1 and day 15 (p=0.63, paired Student's t-test). (**C**) The standard deviation(σ) of the slope of ΔF/F in day 15 (0.32±0.07) was significantly decreased than that in day 1 (0.17±0.05, p=0.03, paired Student's t-test). (**D-E**) The z-scored σ of ΔF/F and z-scored waiting durations were negatively correlated both in the experimental data (d, blue, r = -0.80, p<0.001) and RL model (e, red, r = -0.76, p<0.001). W.D. is for waiting duration. All error bars represent the s.e.m..

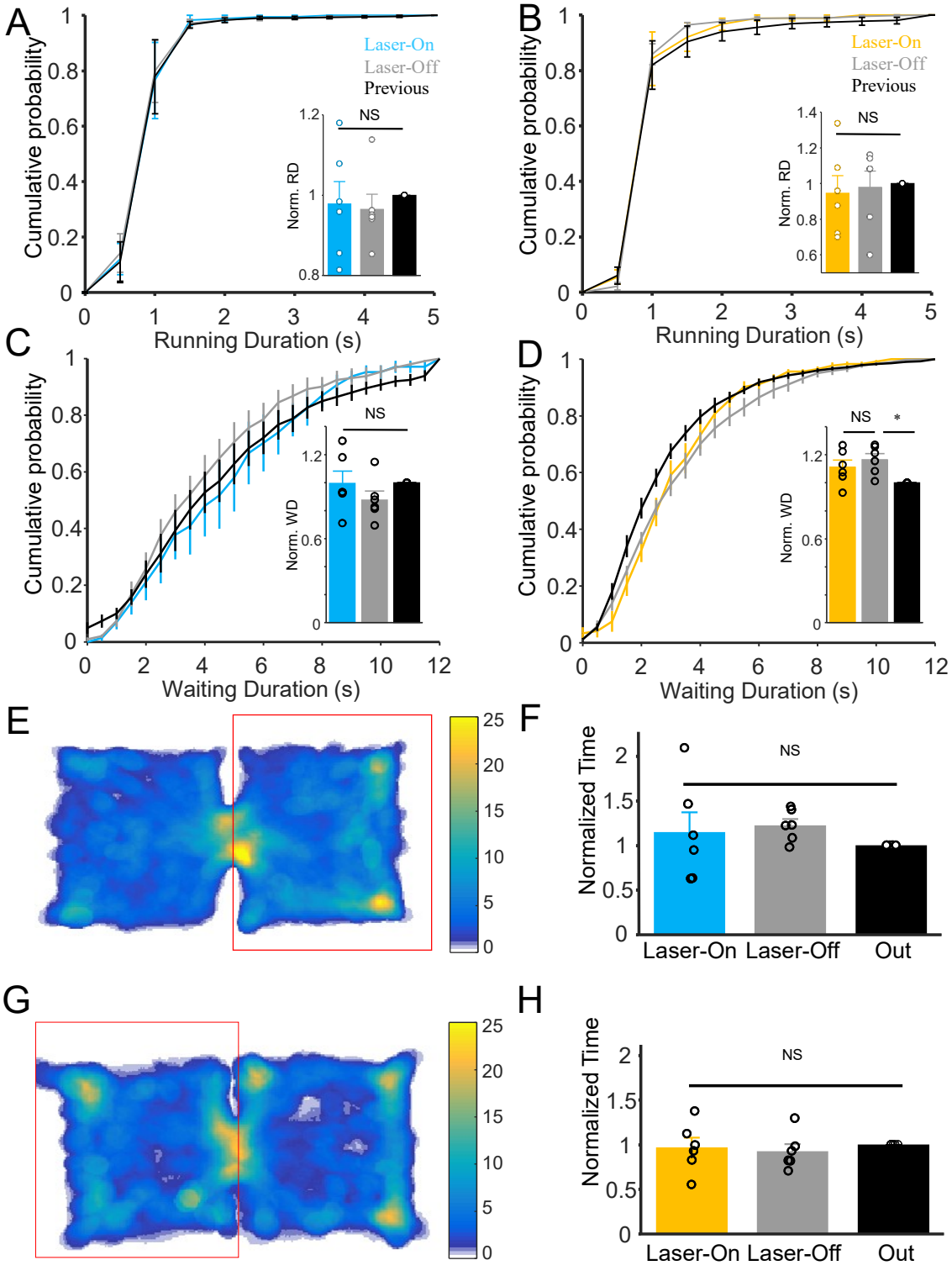
A
Cumulative probability
Running Duration (s)
Laser-On
Laser-Off
Previous
NS
Norm. RD
B
Cumulative probability
Running Duration (s)
Laser-On
Laser-Off
Previous
NS
Norm. RD
C
Cumulative probability
Waiting Duration (s)
NS
Norm. WD
D
Cumulative probability
Waiting Duration (s)
NS
*
Norm. WD
E
25
20
15
10
5
0
F
Normalized Time
NS
Laser-On
Laser-Off
Out
G
25
20
15
10
5
0
H
Normalized Time
NS
Laser-On
Laser-Off
Out

Fig. S5. Optical activation or inhibition of VTA DA activity didn't change the motivation of mouse in the delayed gratification tasks, the waiting duration of mouse in the early process of delayed gratification training, or the duration the mouse stayed in the given box in RPPT. (**A**) The running duration didn't change while the VTA DA neurons were optically activated (F=0.20, p=0.82, one-way ANOVA). (**B**) The result was the same as Figure A while optical inhibiting the VTA DA neurons (F=0.12, p=0.88, one-way ANOVA). (**C**) The waiting duration didn't change (p = 0.37, laser-on vs laser-off) while the VTA DA neurons were optically activated in the early training process (F=1.31, p=0.30, one-way ANOVA). (**D**) The result was the same (p = 0.59, laser-on vs laser-off) as Figure C while optical inhibiting the VTA DA neurons (F=4.97, p=0.22, one-way ANOVA). (**E**) The heatmap of mouse traces in RPPT in which the VTA DA neurons were optically activated pseudo-randomly in 20% probability while the mouse entered into the given box (red rectangle). (**F**) The Z-scored duration that the mouse stayed in the given box while the VTA DA neurons were activated (Laser-On In) had no significant difference (F=0.75, p=0.44, one-way ANOVA, n=6) with the un-inhibited durations (Laser-Off In) and durations in another box (Out). (**G**) The heatmap of mouse traces as shown in Figure C while inhibiting the VTA DA neurons in the given box (red rectangle). (**H**) Optical inhibiting the VTA DA neurons also didn't change the duration the mouse stayed in the given box (F = 0.17, p = 0.73, one-way ANOVA, n=6). All error bars represent the s.e.m..

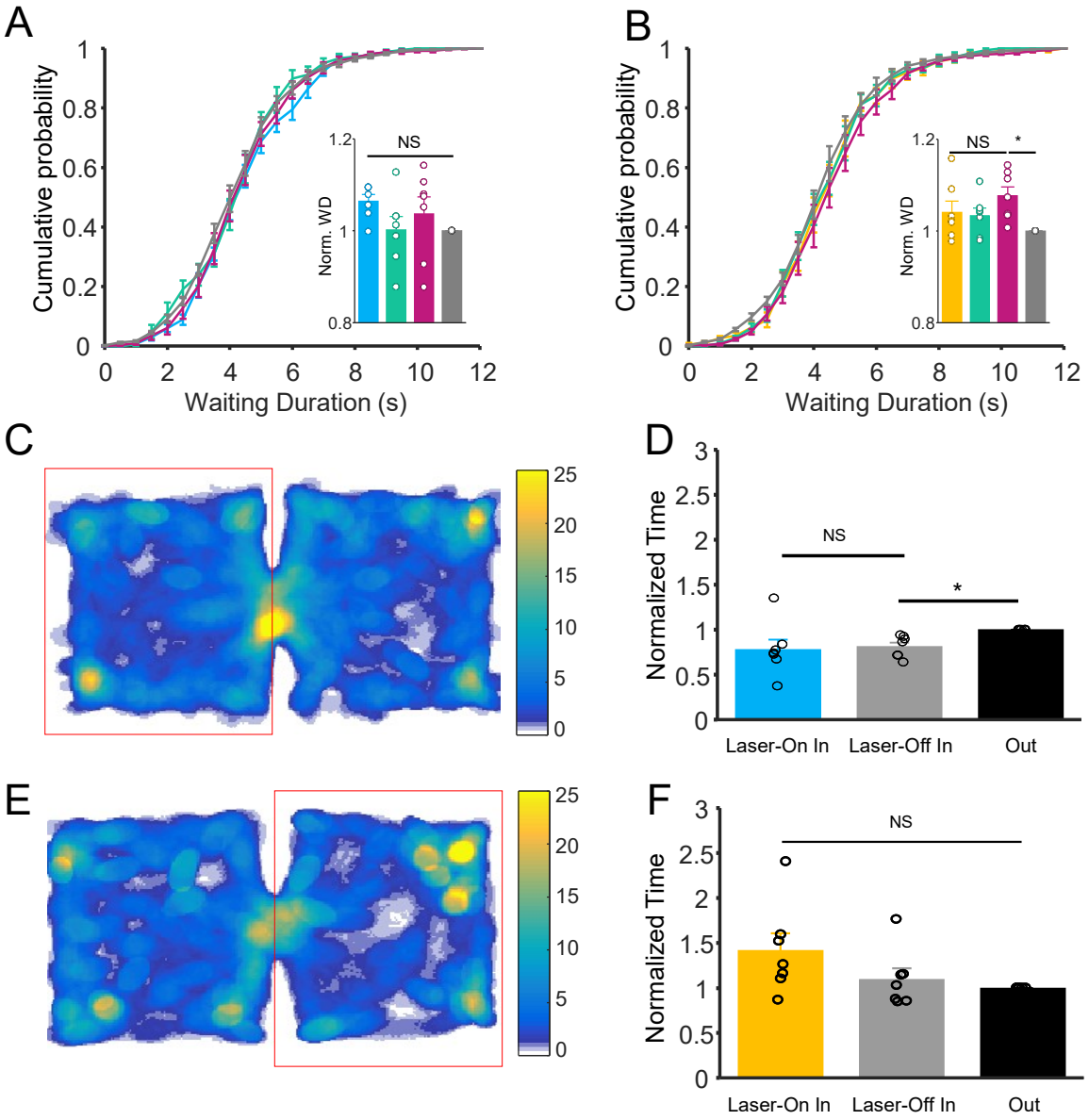

Fig. S6. Optogenetic manipulation of DAT-Cre mouse expressed mCherry in the delayed gratification tasks and RPPT. (**A**) Waiting durations in 473nm laser delivered trials (blue) are not different compared with those of all other trials (p=0.17, Friedman test, n=7). (**B**) Waiting durations of 589nm laser un-delivered trials (magenta) slightly increased compared with the waiting duration of the previous day (p=0.02, Friedman test, n=7). (**C**) Heat-map of mouse traces in RPPT in which the VTA of the mouse was delivered 473nm laser pseudo-randomly in 20% probability while the mouse entered into a randomly chosen box (red rectangle). (**D**) Mean durations that the mouse stayed in the chosen box while the laser delivered (Laser-On In), laser off (Laser-Off In) and the other box. There is no significant difference in waiting duration between Laser-On In, and Laser-Off In ($F=3.54$, $p=0.09$, one-way ANOVA, n=7). (**E**) Heatmap of mouse traces same as shown in (c) while the mouse was delivered 589nm laser in a randomly chosen box (red rectangle). (**F**) Mean durations that the mouse stayed in the chosen box while 589nm laser delivered (Laser-On In), laser off (Laser-Off In), and in the other box. 589nm laser delivering to mCherry mouse didn't alter waiting duration mouse stayed in any boxes under all experimental conditions ($F=2.64$, $p=0.14$, one-way ANOVA, n=7). All error bars represent the s.e.m..

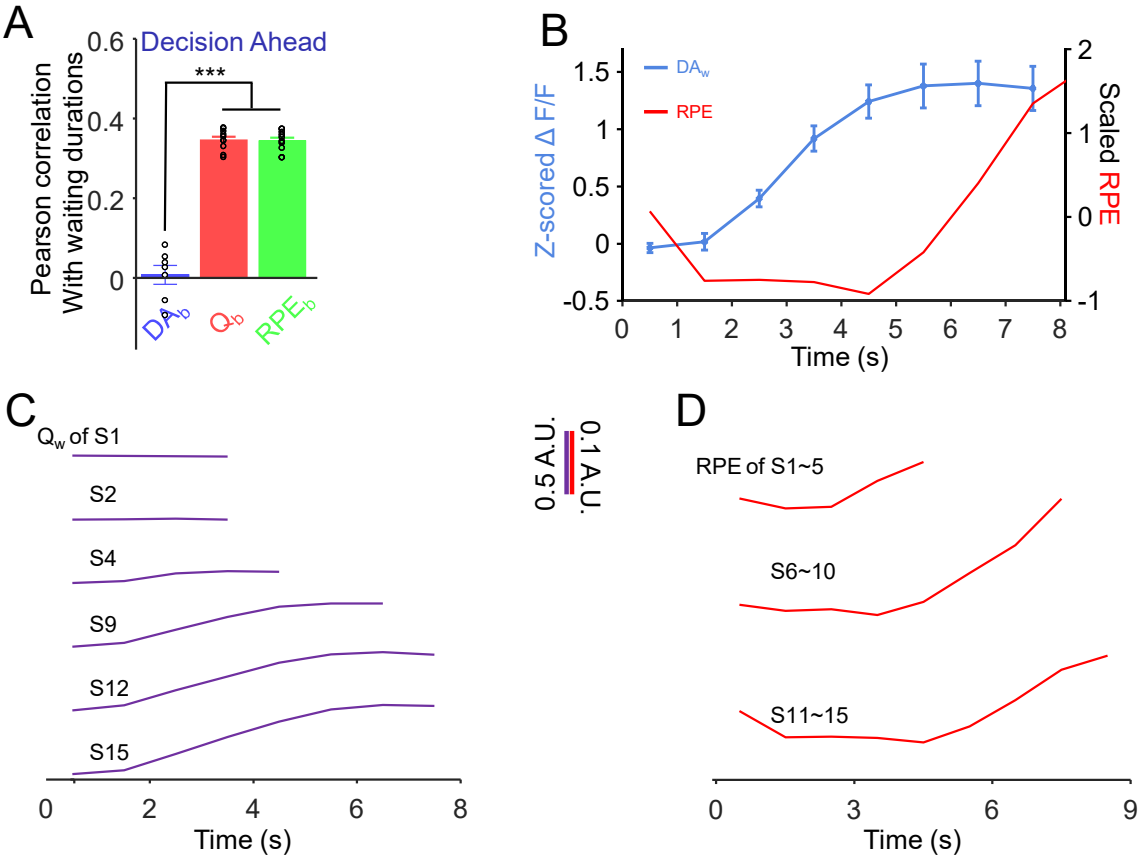

Fig. S7. Reward prediction error is not significantly correlated with ramping dopaminergic activity. (**A**) The correlation coefficient of mean DA activity 1 s before waiting (DA_b), the value of action (Q_b), and RPE of action (RPE_b) before waiting in the Decision Ahead model, with waiting durations. The correlation of V_b (r = 0.35±0.01, p=0.001, n=10, Pearson correlation) and RPE_b (r = 0.34±0.01, p=0.001, n=10, Pearson correlation) with waiting duration were significantly (p<0.001, Kruskal-Wallish test) higher than the CC of DA_b (r=0.01±0.02, p=0.36, n=7) with waiting durations. (**B**) Plots of Z-scored ΔF/F values (DA_w, light blue) at 0.5s before the waiting ended and RPE of waiting (RPE_w). There were no significant correlation between DA_w and RPE_w (r = 0.34, p = 0.41, Pearson correlation). (**C-D**) Value of waiting (Q_w) (**C**) and RPE (**D**) changed in the Continuous Deliberation model. All error bars represent the s.e.m..

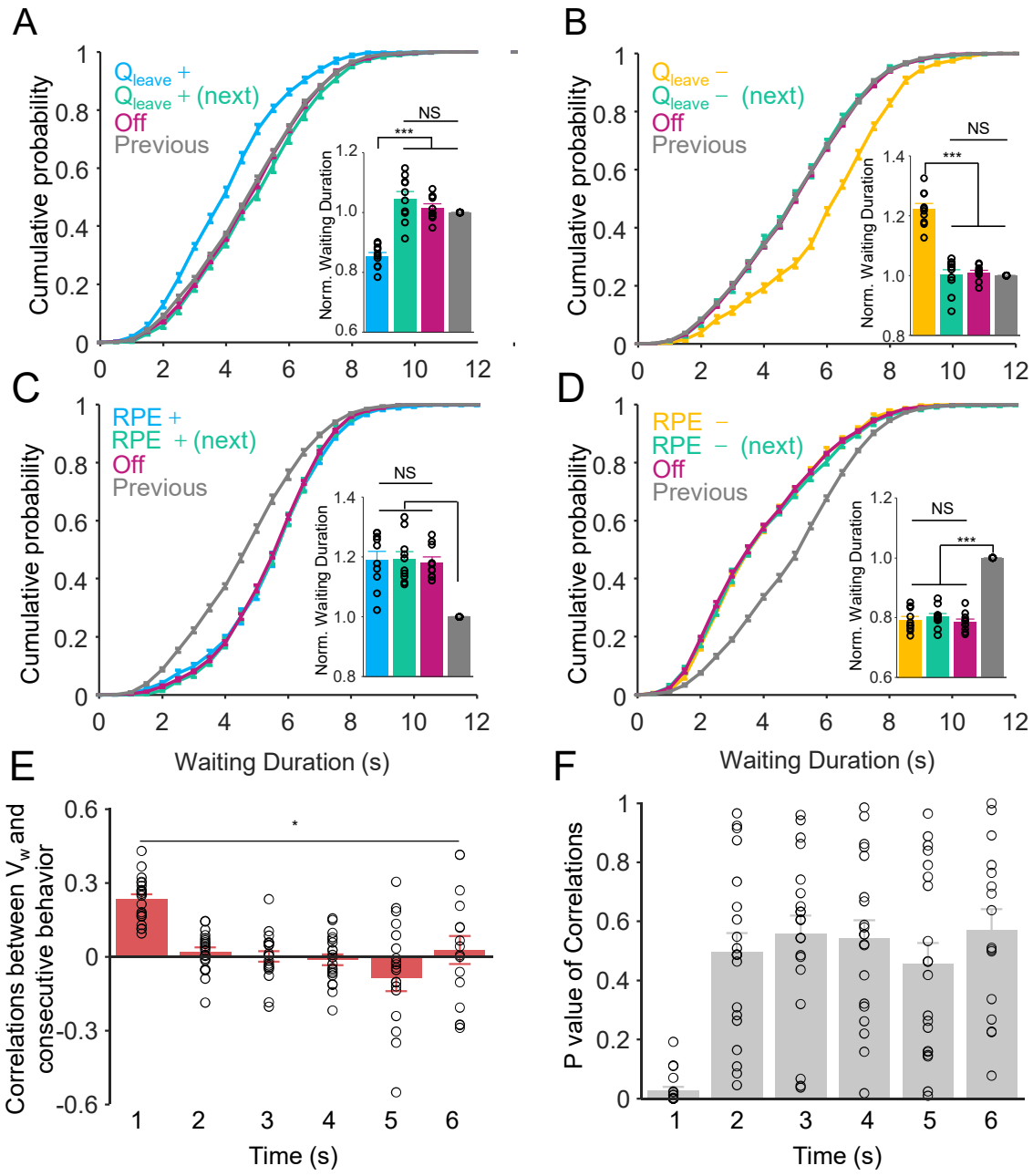

Fig. S8. Manipulation of the value of leaving and RPE in RL model. (**A-B**) Either increasing or decreasing the value of leaving (Q_{leave}) in the continuous deliberation model as with the manipulation of Q_{wait} induced the opposite results compared with the optogenetics manipulating DAergic activity (increasing Q_{leave} in **A**: p<0.001, Friedman test, n=10; decreasing Q_{leave} in **B**: p <0.001, Friedman test, n=10) and had no influences on other trials (**A-B**, p>0.999, Friedman test, n=10). (**C-D**) Either increasing (**C**) or decreasing (**D**) the RPE in the continuous deliberation model as with the experimental data, alters the waiting durations in the same direction in all trials, whether or not the RPEs-manipulation (increasing RPEs in c: p<0.001, Friedman test, n=10; decreasing RPE in (**D**): p <0.001, Friedman test, n=10). (**E**) The value of waiting is only positively correlated (0.23±0.02, p =0.03±0.01, n =20) with the adjacent behavior in the *Continuous Deliberation* RL model. (**F**) The p values of correlation coefficients in **E**. All error bars represent the s.e.m..

MovieS1 The mouse performed pre-training task well after one week's training. After one week's training of pre-training task, the mouse performed the task with short waiting duration and running duration.

MovieS2 The mouse performed delayed gratification task well after one month's training. After one month's training of delayed gratification task, the mouse performed the task with long waiting duration and short running duration.