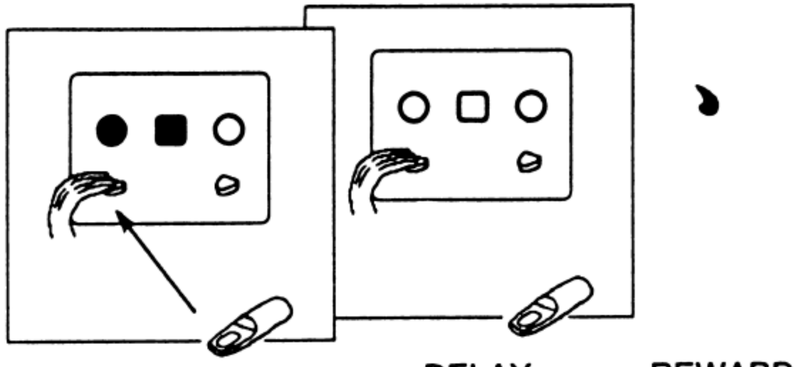
Models of Basal Ganglia in Decision Making

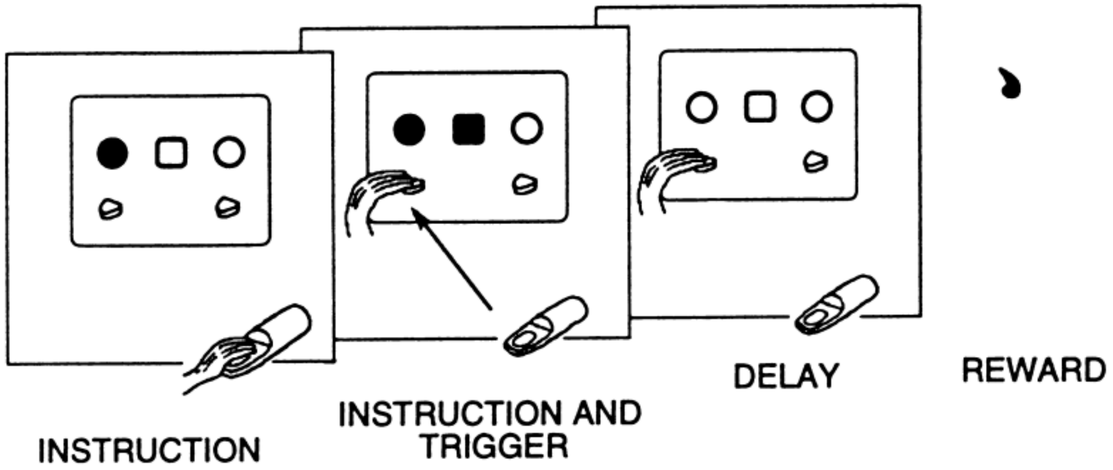
4/23/2021

Brabeeba Wang

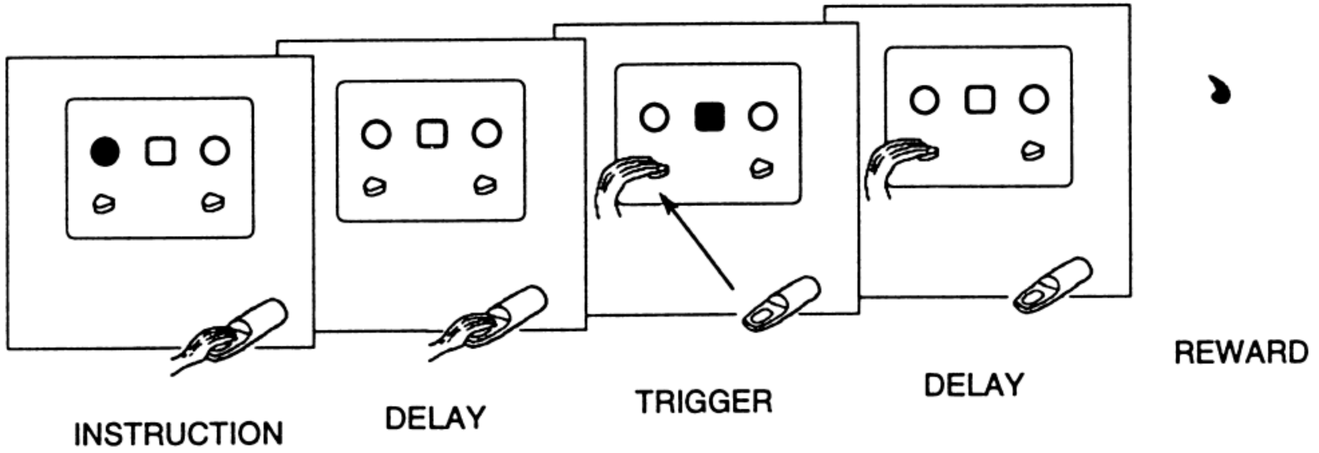
A Spatial choice task

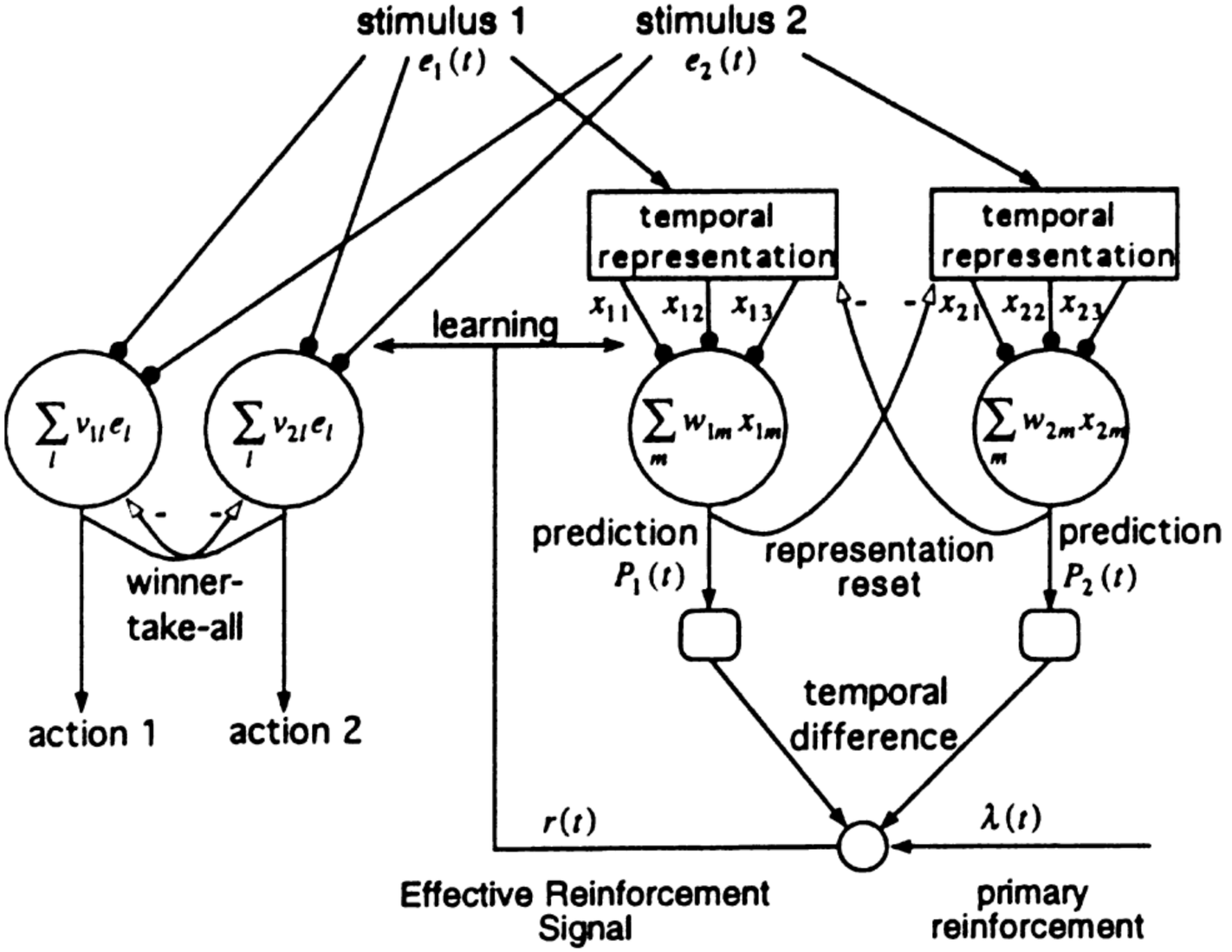


B Instructed spatial task



C Spatial delayed response task





$$P_l(t) = \sum_m W_{lm} X_{lm}(t).$$

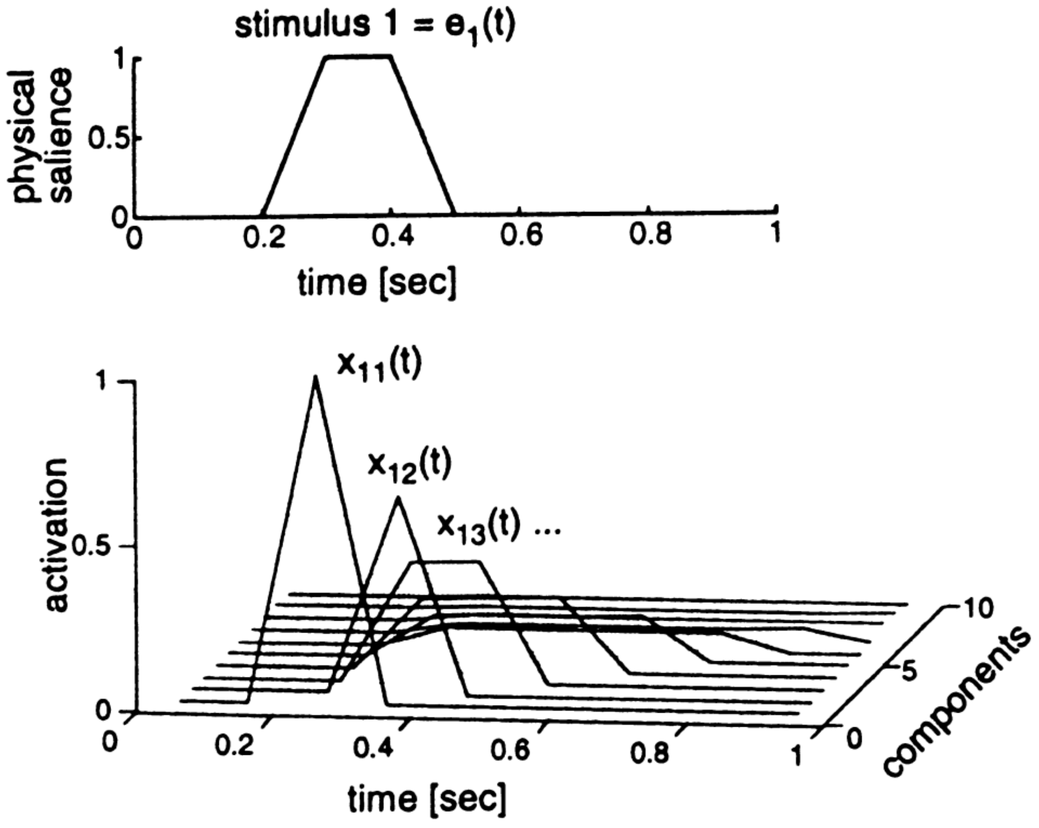
$$P(t) = \sum_l P_l(t).$$

$$r(t) = d + \lambda(t) + \gamma P(t) - P(t-1).$$

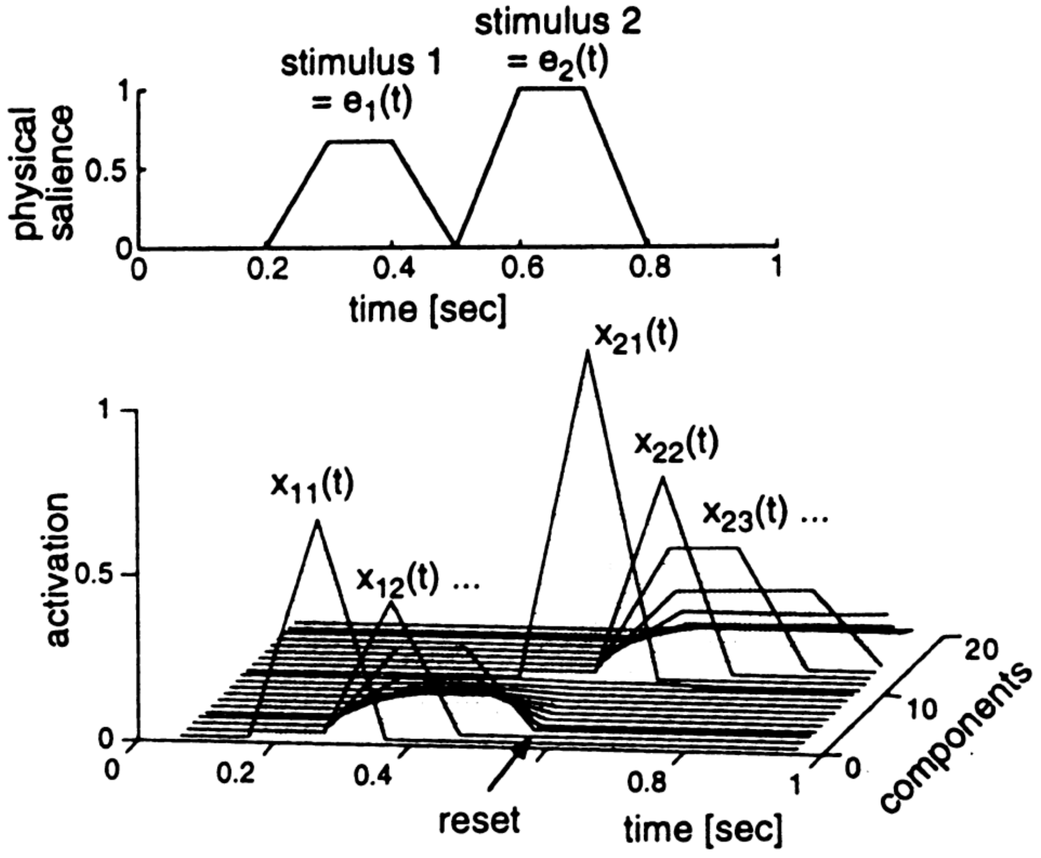
Actor **Critic**

$$W_{lm}(t) = W_{lm}(t-1) + \eta_c (r(t) - d) \lfloor x_{lm}(t-1) - \gamma x_{lm}(t) \rfloor_+$$

A Temporal representation of stimulus 1



B Stimulus 2 resets representation of stimulus 1



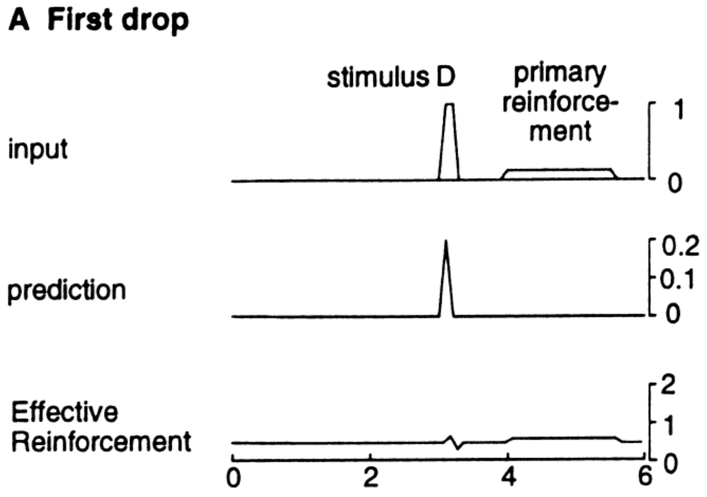
$$\text{If} P_{stimulus_onset}(t) = \sum_{l} w_{l1} X_{l1}(t) > P(t-1),$$

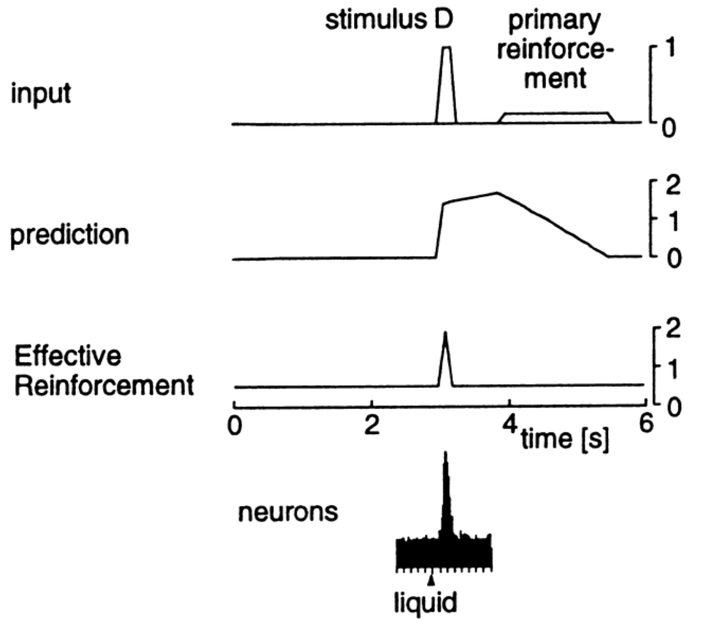
$$\text{then} x_{l,m\neq 1}(t) = 0.$$

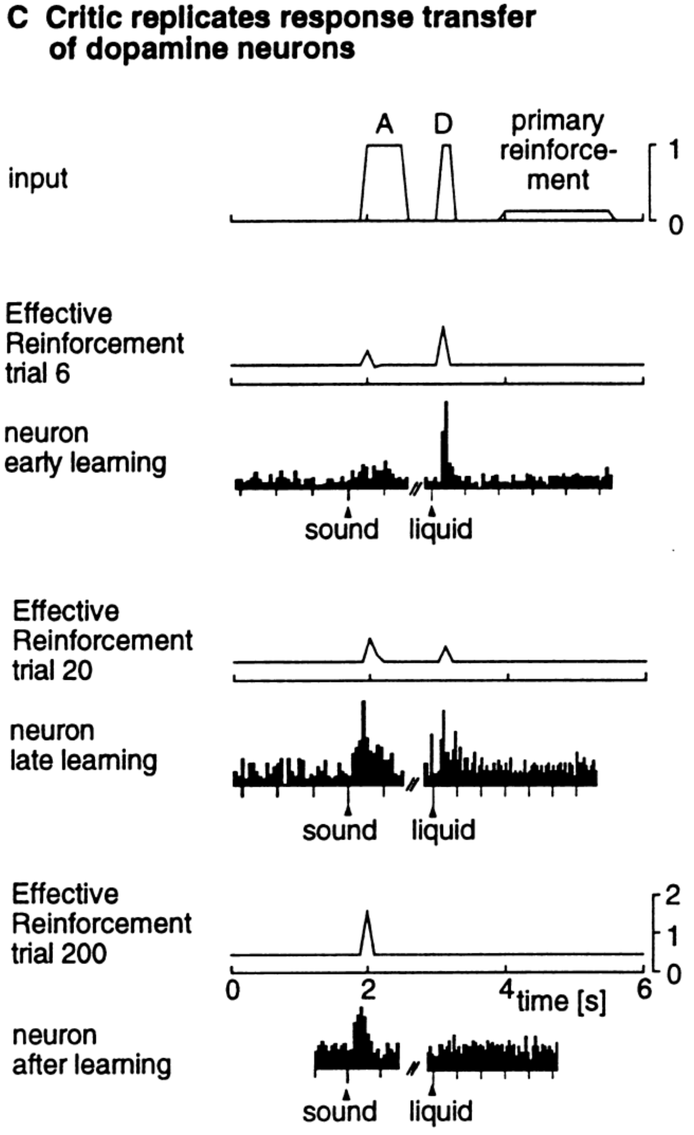
$$a'_n(t) = \left(\sum_l \nu_{nl} \bar{e}_1(t) - \sigma_n(t) \right).$$

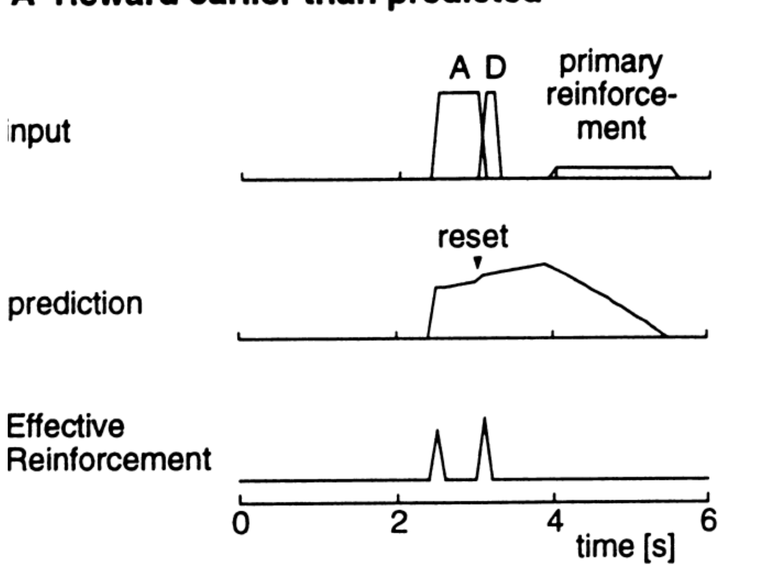
$$a_n(t) = \begin{cases} 1 \ \textit{if} \ a'_n(t) > 0 \ \textit{and} \ a'_n(t) > a'_m(t) \\ 0 \qquad\qquad\qquad\qquad \textit{else} \end{cases}.$$

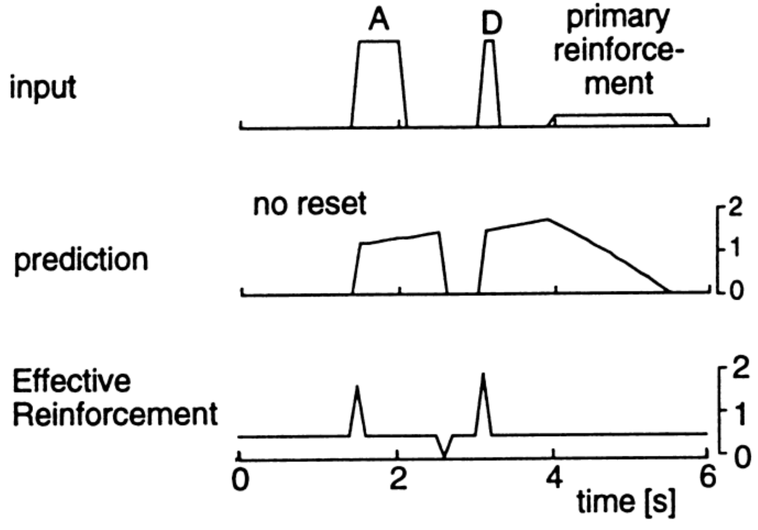
$$\nu_{nl}(t) = \nu_{nl}(t-1) + \eta_a [r(t) - d] \bar{a}_n(t) \bar{e}_l(t),$$

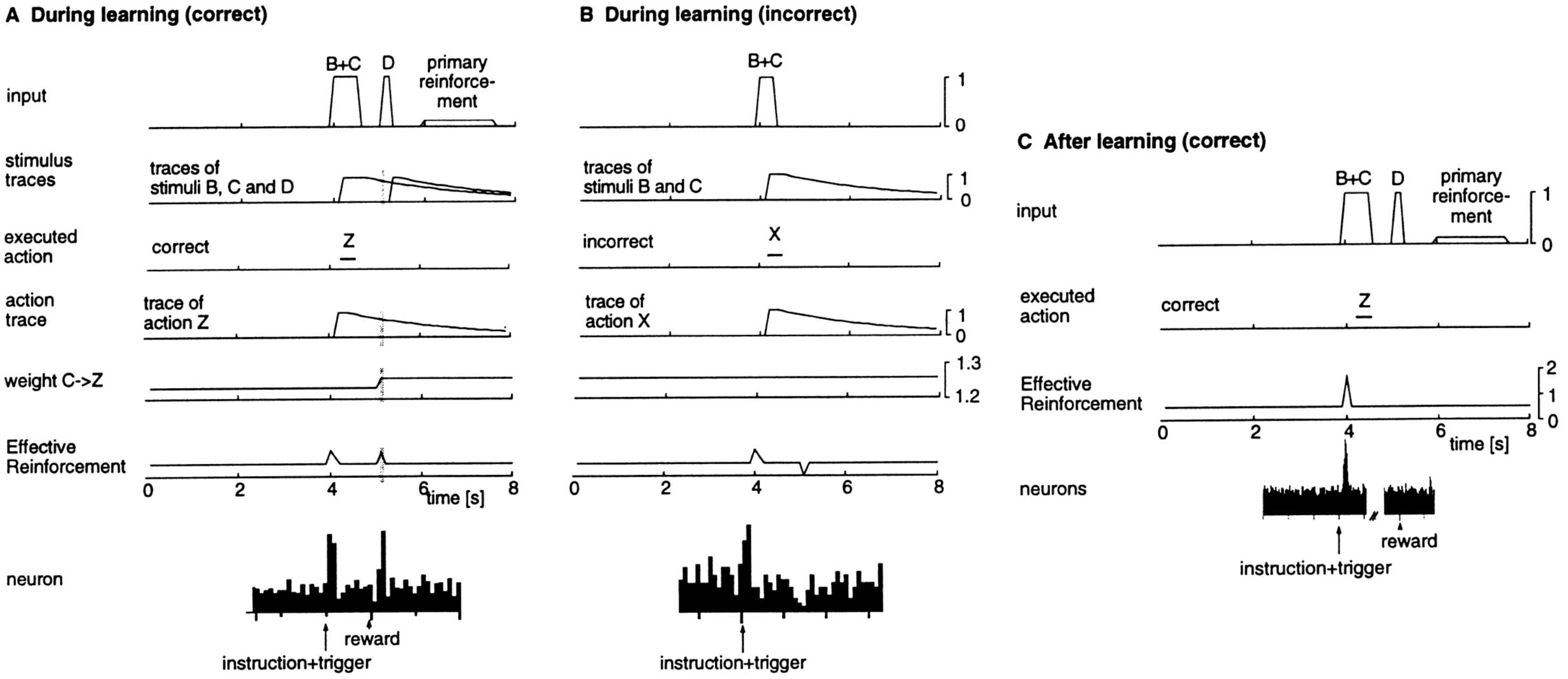

A First drop
stimulus D
primary reinforce-ment
input
prediction
Effective Reinforcement
1
0
0.2
0.1
0
2
1
0
0
2
4
6

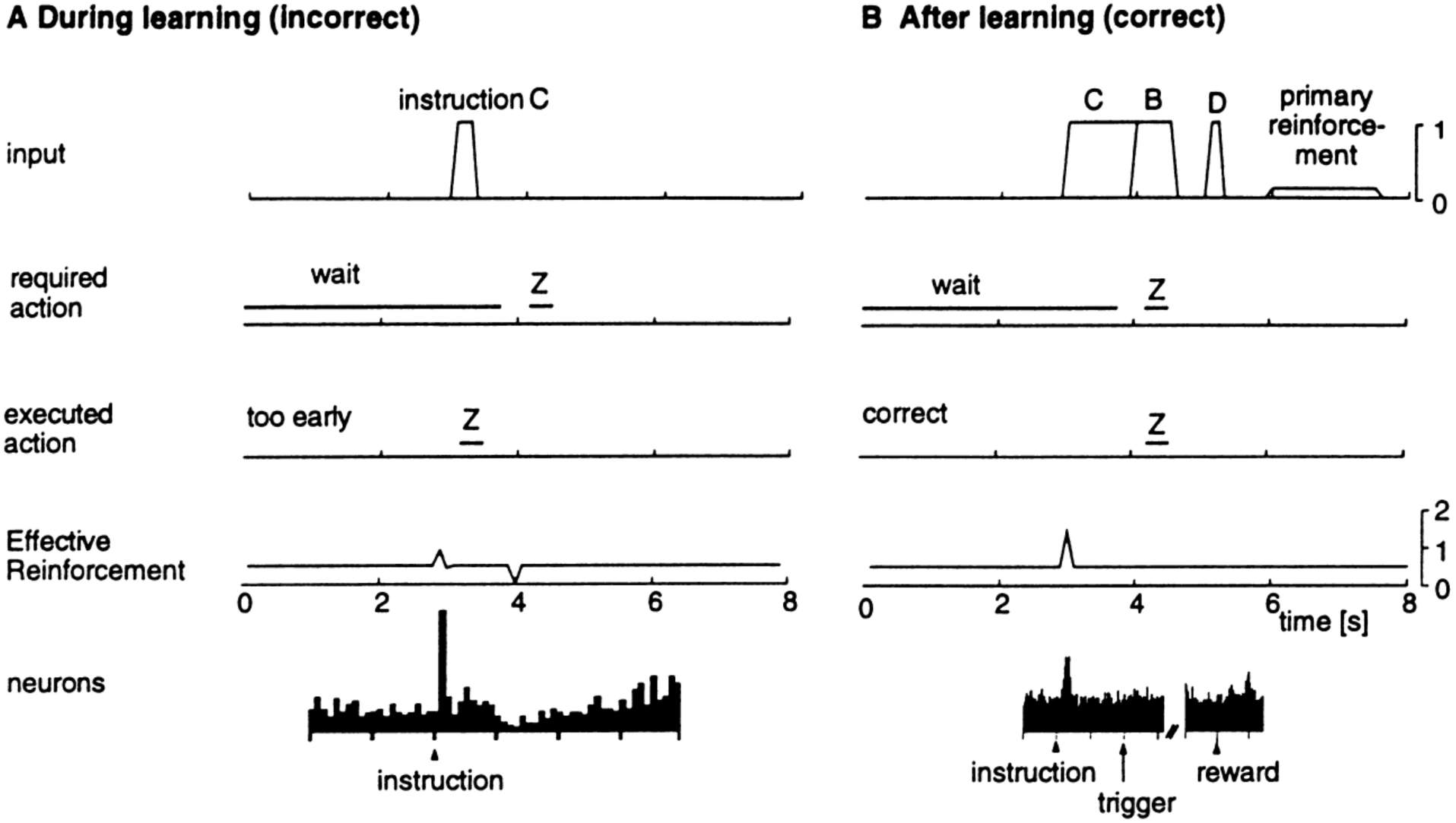

B Experienced with liquid
stimulus D
primary reinforce-ment
input
prediction
Effective Reinforcement
neurons
liquid
time [s]

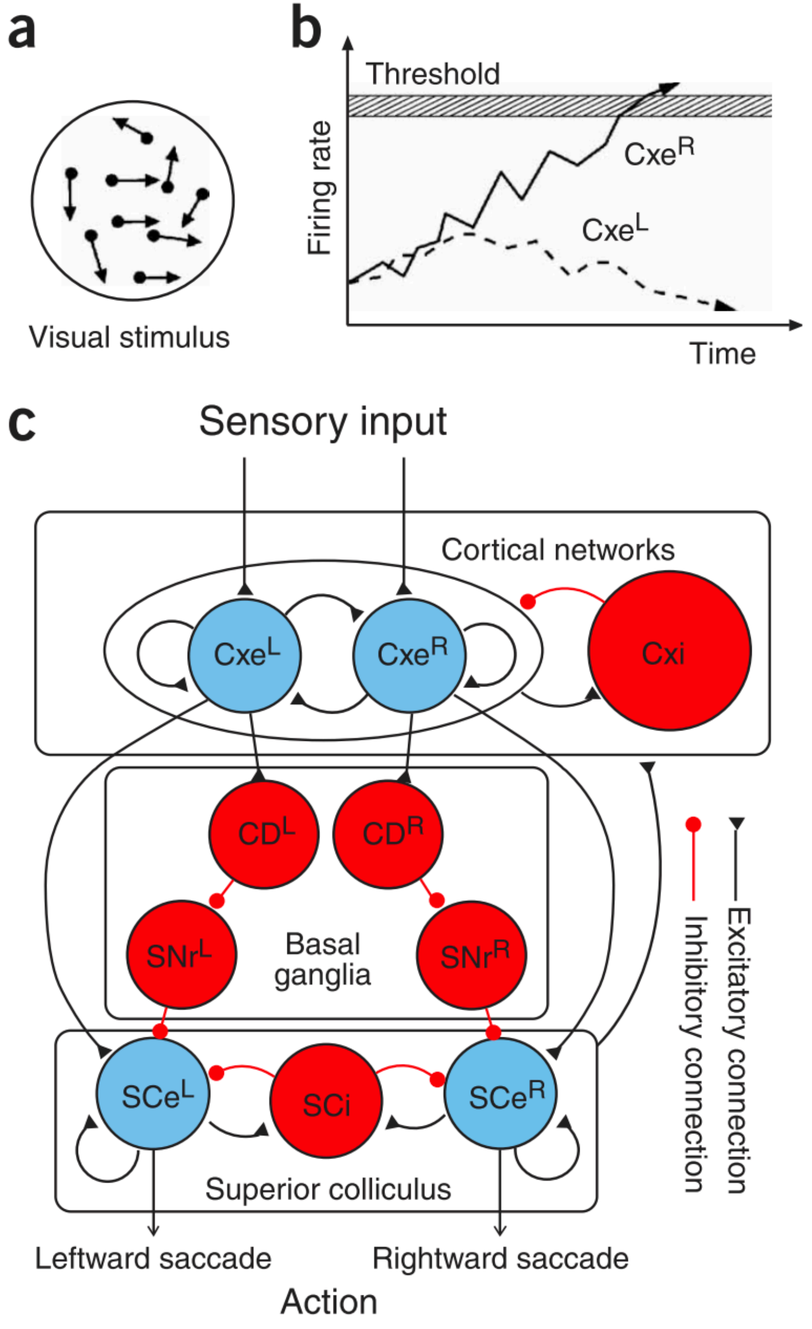

C Critic replicates response transfer of dopamine neurons
A
D
primary reinforce-ment
input
Effective Reinforcement trial 6
neuron early learning
sound
liquid
Effective Reinforcement trial 20
neuron late learning
Effective Reinforcement trial 200
neuron after learning
time [s]

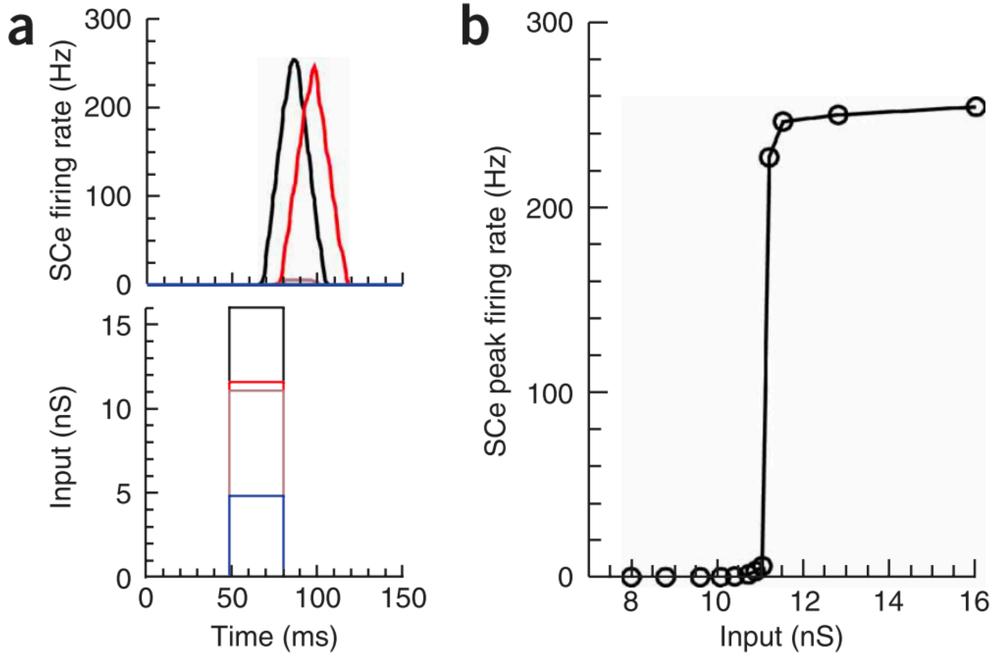

A Reward earlier than predicted
A
D
primary reinforce-ment
input
reset
prediction
Effective Reinforcement
time [s]

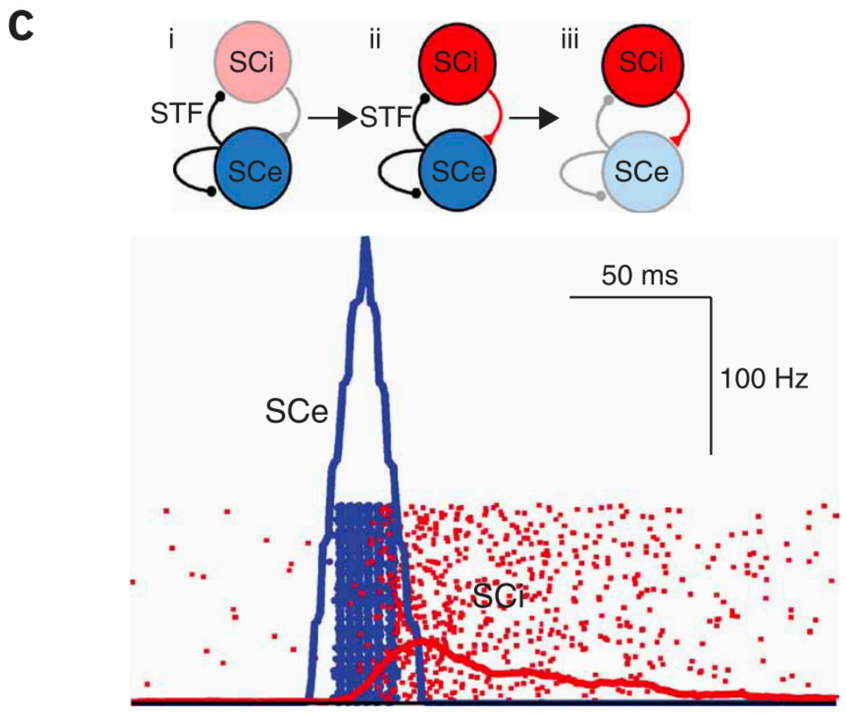

B Reward later than predicted
A
D
primary reinforce-ment
input
no reset
prediction
Effective Reinforcement
time [s]

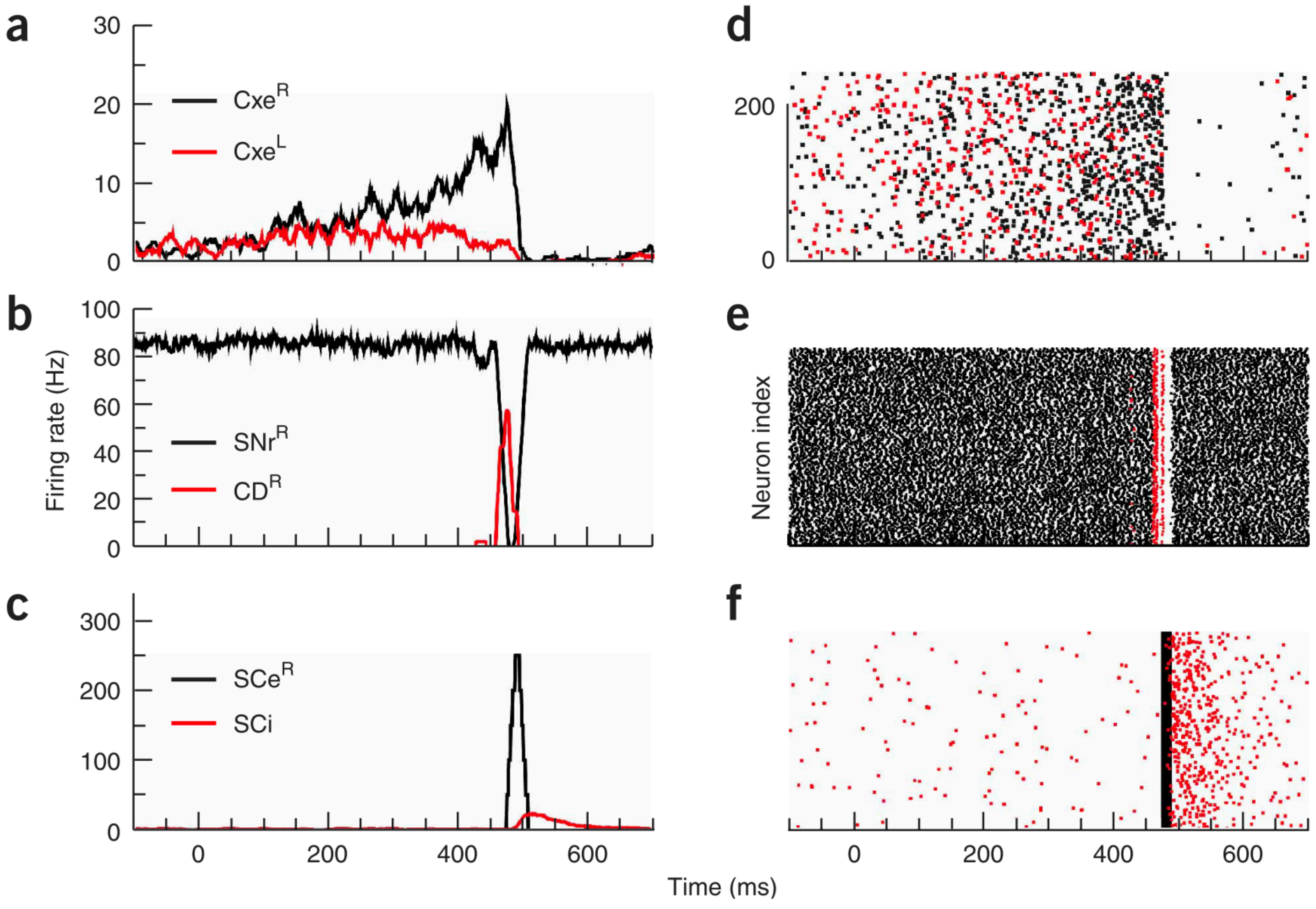

A During learning (correct)
input
B+C
D
primary reinforce-ment
stimulus traces
traces of stimuli B, C and D
executed action
correct
Z
action trace
trace of action Z
weight C->Z
Effective Reinforcement
0
2
4
6
8
time [s]
neuron
reward
instruction+trigger
B During learning (incorrect)
B+C
1
0
traces of stimuli B and C
incorrect
X
trace of action X
1.3
1.2
instruction+trigger
C After learning (correct)
input
B+C
D
primary reinforce-ment
executed action
correct
Z
Effective Reinforcement
2
1
0
time [s]
neurons
reward
instruction+trigger

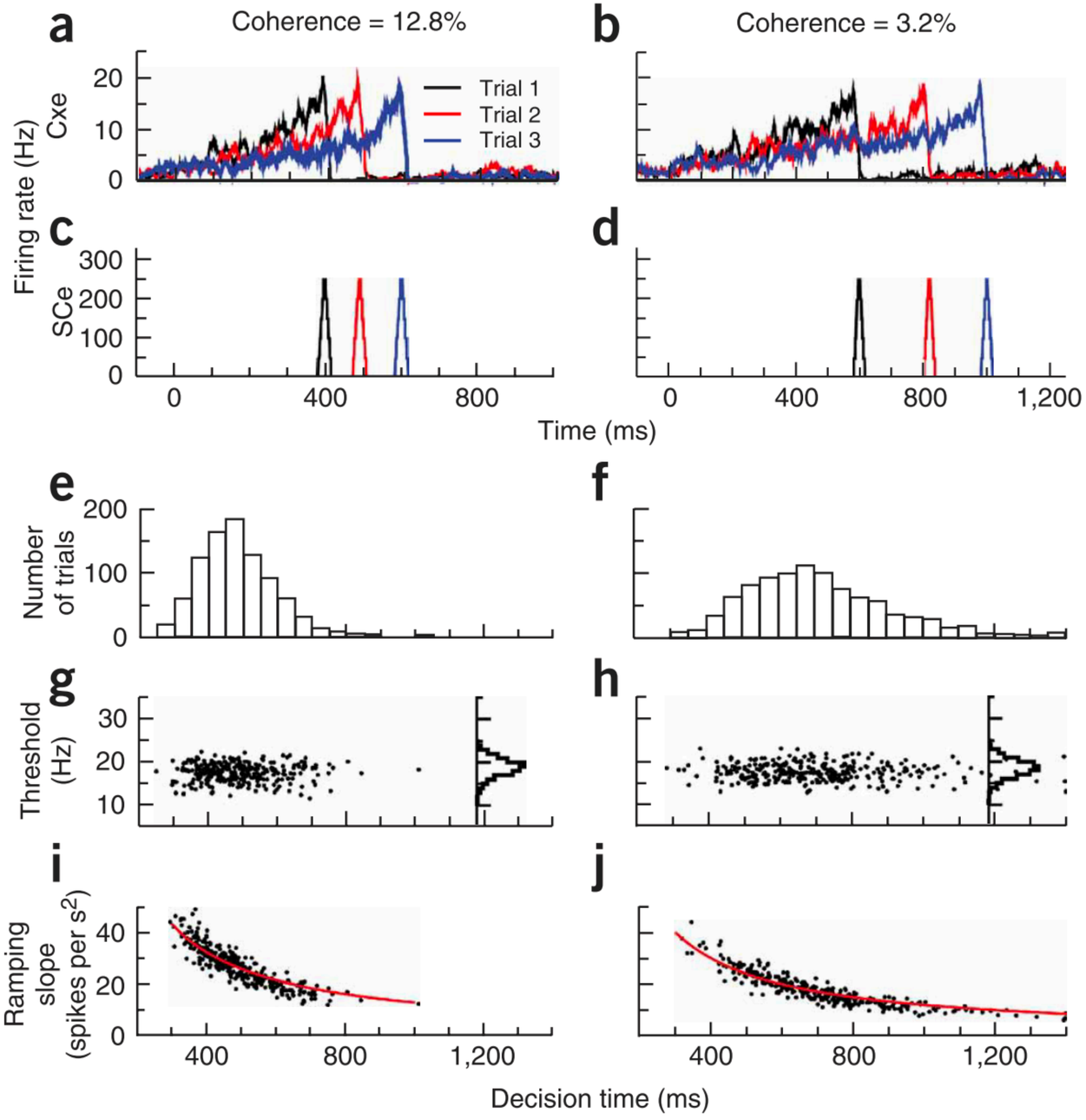

A During learning (incorrect)
B After learning (correct)
input
instruction C
C
B
D
primary reinforce-ment
1
0
required action
wait
Z
wait
Z
executed action
too early
Z
correct
Z
Effective Reinforcement
2
1
0
0
2
4
6
8
0
2
4
6
8
time [s]
neurons
instruction
instruction
trigger
reward

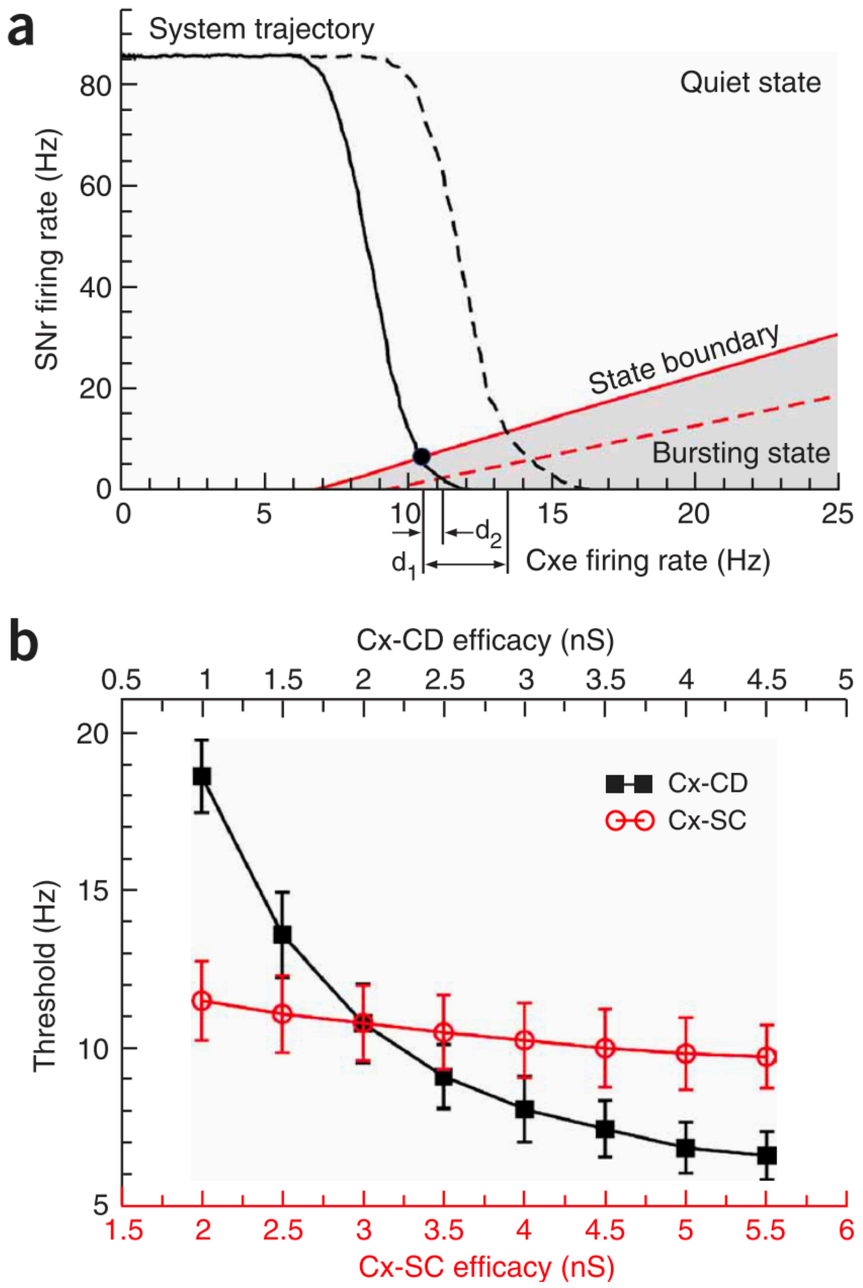

a
b
Threshold
Firing rate
CxeR
CxeL
Time
Visual stimulus
c
Sensory input
Cortical networks
CxeL
CxeR
Cxi
CDL
CDR
SNrL
SNrR
Basal ganglia
SCeL
SCi
SCeR
Superior colliculus
Inhibitory connection
Excitatory connection
Leftward saccade
Rightward saccade
Action

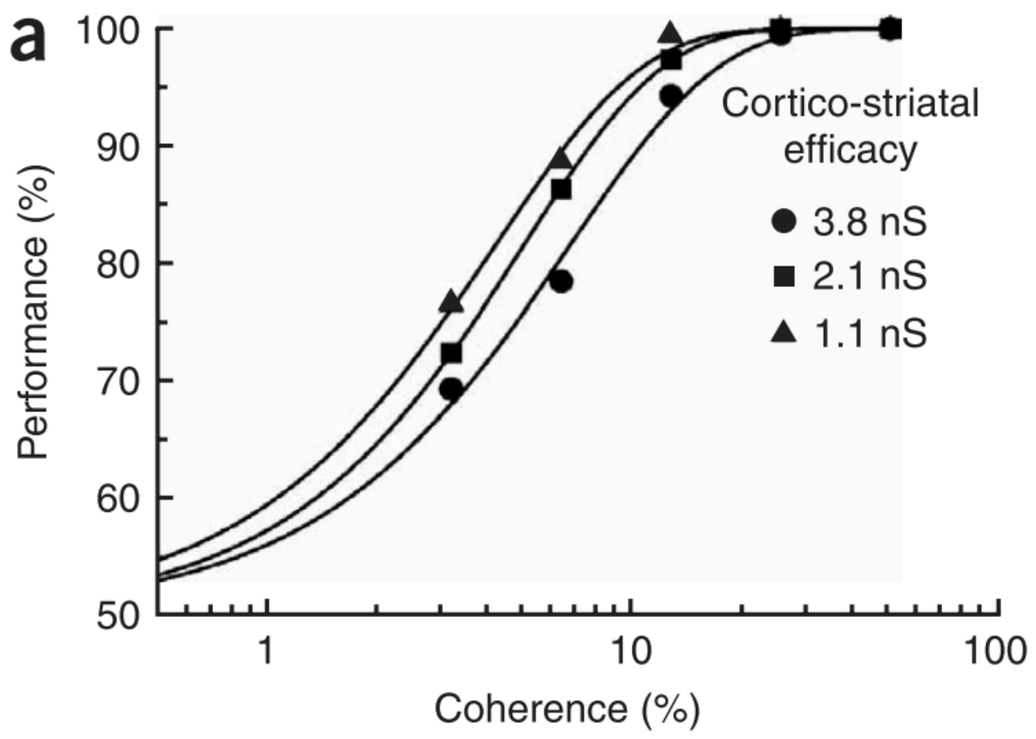

a
SCe firing rate (Hz)
300
200
100
0
Input (nS)
15
10
5
0
0
50
100
150
Time (ms)
b
SCe peak firing rate (Hz)
300
200
100
0
8
10
12
14
16
Input (nS)

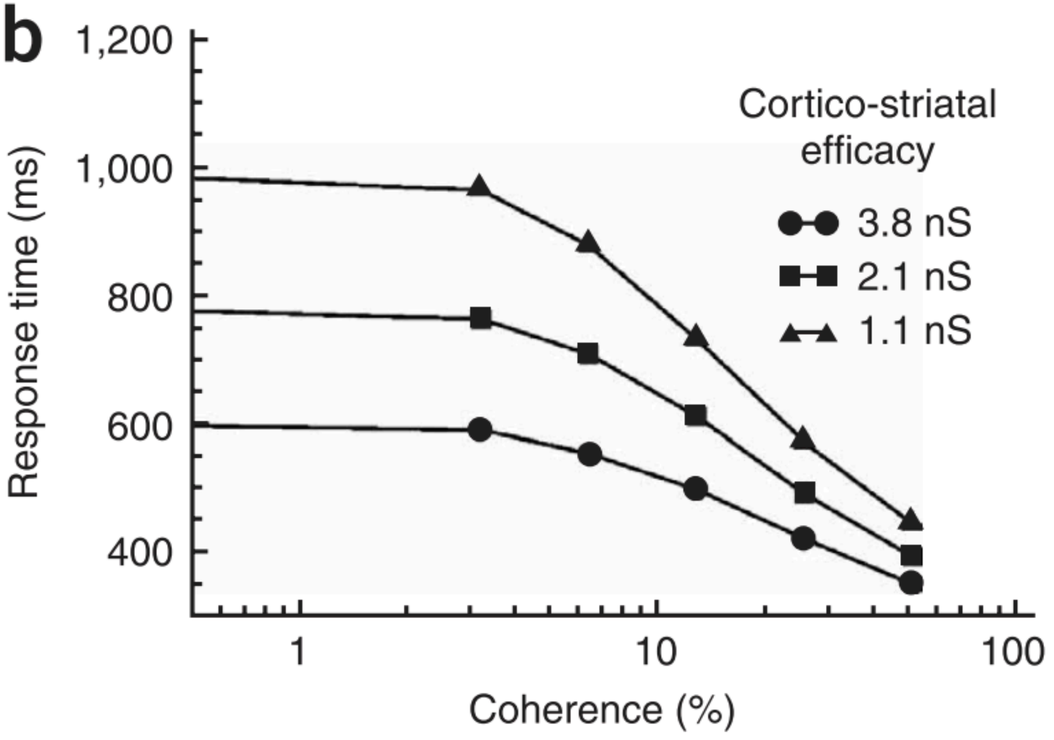

c
i
ii
iii
SCi
STF
SCe
50 ms
100 Hz
SCe
SCi

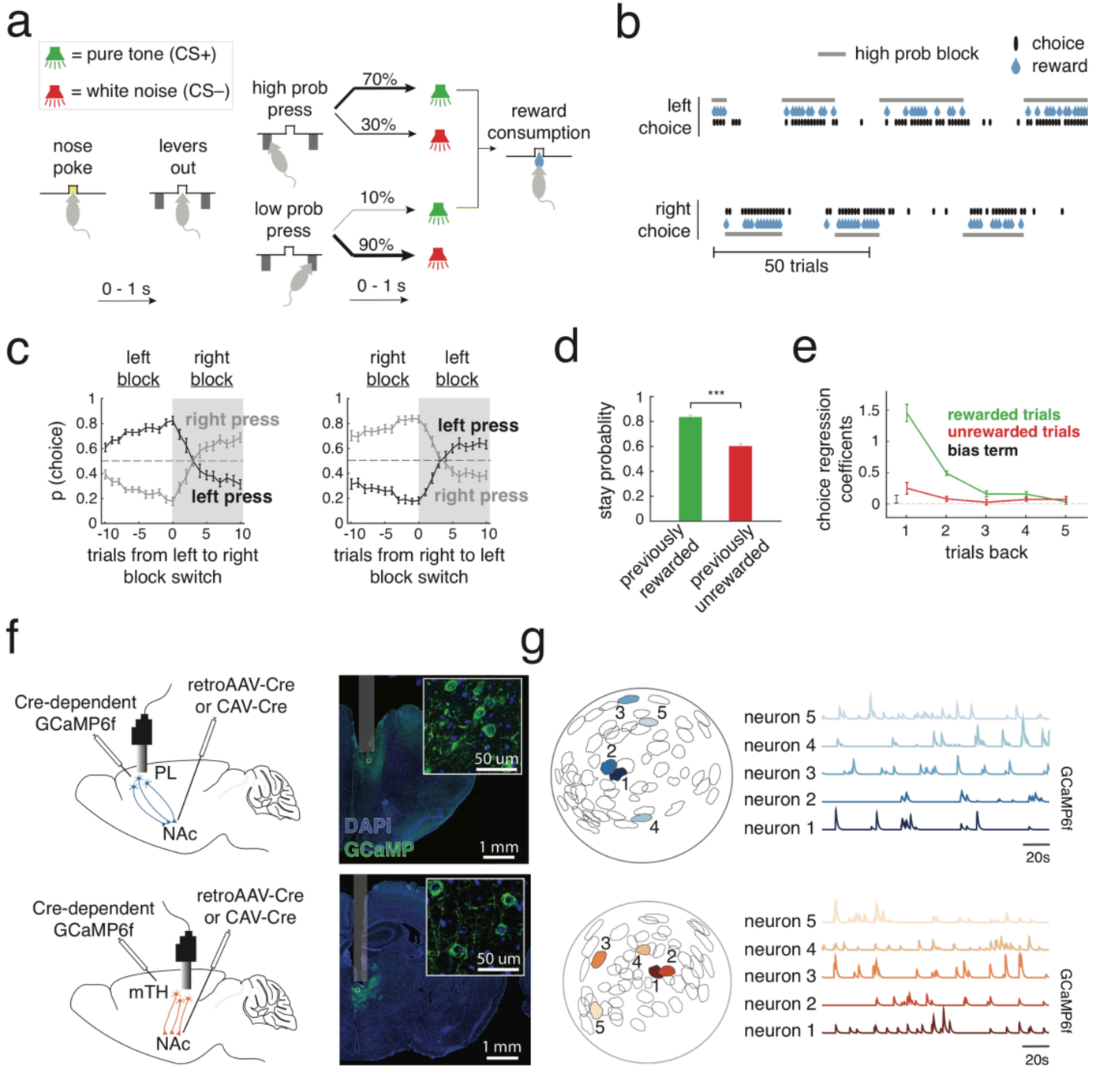

a
30
20
10
0
Cxe^R
Cxe^L
b
100
80
60
40
20
0
Firing rate (Hz)
SNr^R
CD^R
c
300
200
100
0
SCe^R
SCi
0
200
400
600
Time (ms)
d
200
0
e
Neuron index
f
0
200
400
600

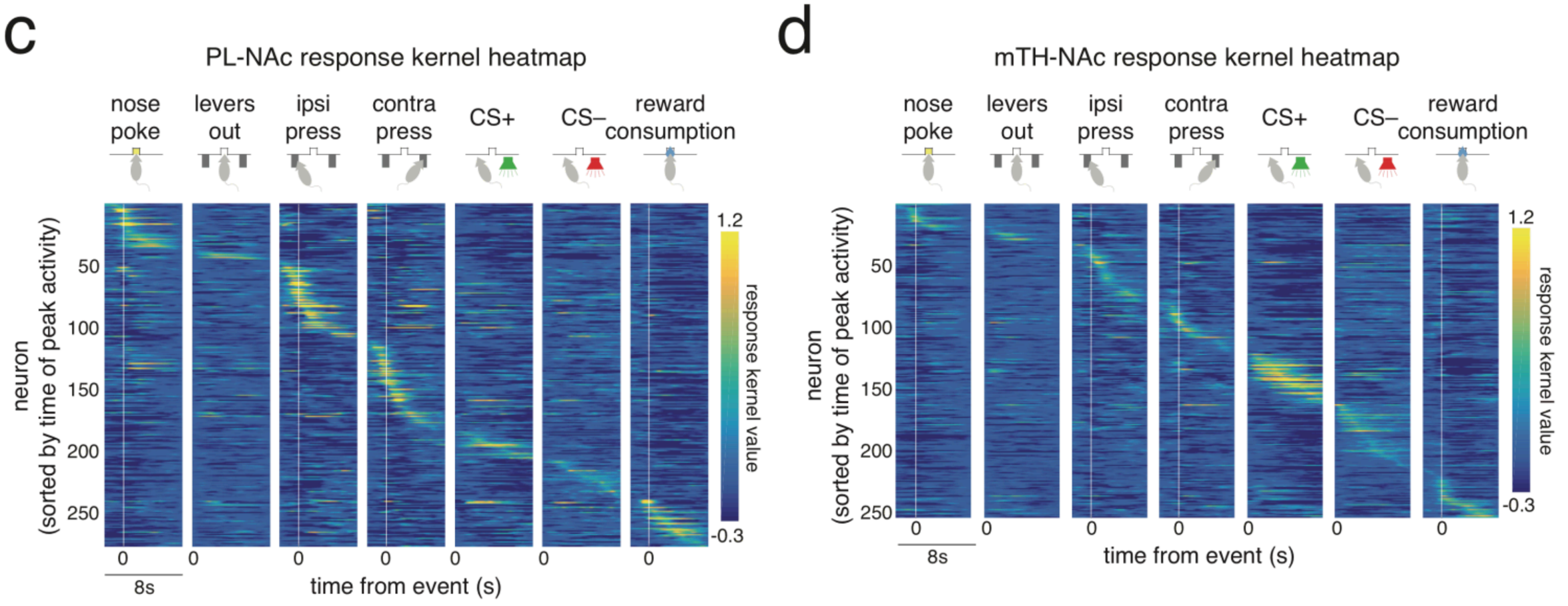

a
Coherence = 12.8%
b
Coherence = 3.2%
Trial 1
Trial 2
Trial 3
Firing rate (Hz)
Cxe
SCe
c
d
Time (ms)
e
f
Number of trials
g
h
Threshold (Hz)
i
j
Ramping slope (spikes per s²)
Decision time (ms)


a
System trajectory
Quiet state
SNr firing rate (Hz)
State boundary
Bursting state
d1
d2
Cxe firing rate (Hz)
b
Cx-CD efficacy (nS)
Cx-CD
Cx-SC
Threshold (Hz)
Cx-SC efficacy (nS)

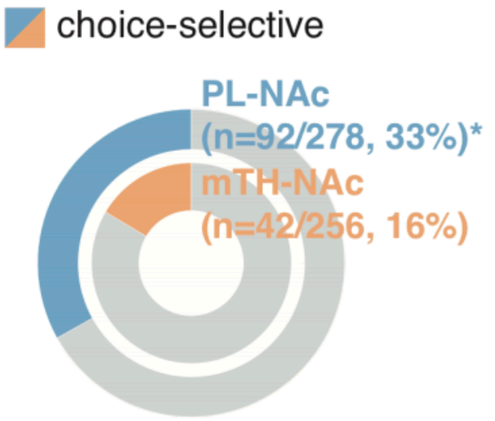

a
100
90
80
70
60
50
Performance (%)
1
10
100
Coherence (%)
Cortico-striatal efficacy
3.8 nS
2.1 nS
1.1 nS


b
1,200
1,000
800
600
400
Response time (ms)
1
10
100
Coherence (%)
Cortico-striatal efficacy
3.8 nS
2.1 nS
1.1 nS

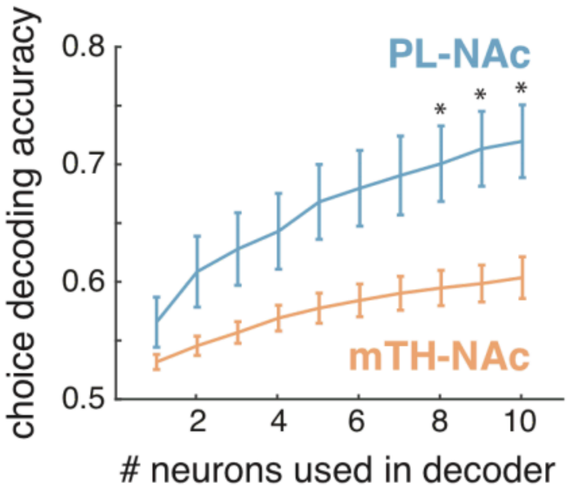

a
= pure tone (CS+)
= white noise (CS–)
nose poke
levers out
high prob press
70%
30%
low prob press
10%
90%
reward consumption
0 - 1 s
0 - 1 s
b
high prob block
choice
reward
left choice
right choice
50 trials
c
left block
right block
right press
left press
p (choice)
trials from left to right block switch
right block
left block
left press
right press
trials from right to left block switch
d
stay probability
***
previously rewarded
previously unrewarded
e
choice regression coefficents
rewarded trials
unrewarded trials
bias term
trials back
f
Cre-dependent GCaMP6f
retroAAV-Cre or CAV-Cre
PL
NAc
50 um
DAPI
GCaMP
1 mm
Cre-dependent GCaMP6f
retroAAV-Cre or CAV-Cre
mTH
NAc
50 um
1 mm
g
neuron 5
neuron 4
neuron 3
neuron 2
neuron 1
GCaMP6f
20s
neuron 5
neuron 4
neuron 3
neuron 2
neuron 1
GCaMP6f
20s


c
PL-NAc response kernel heatmap
nose poke
levers out
ipsi press
contra press
CS+
CS–
reward consumption
neuron (sorted by time of peak activity)
50
100
150
200
250
1.2
-0.3
response kernel value
0
8s
time from event (s)
d
mTH-NAc response kernel heatmap
nose poke
levers out
ipsi press
contra press
CS+
CS–
reward consumption
neuron (sorted by time of peak activity)
50
100
150
200
250
1.2
-0.3
response kernel value
0
8s
time from event (s)

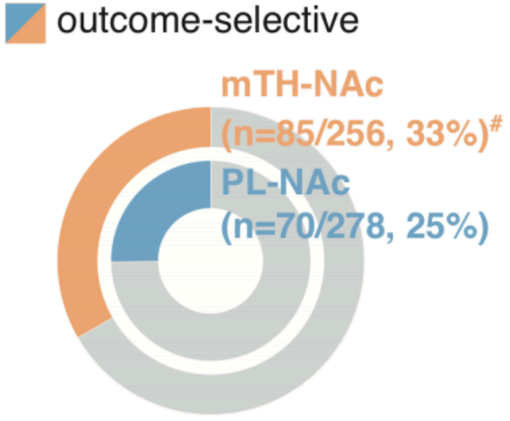
a


choice-selective
PL-NAc
(n=92/278, 33%)*
mTH-NAc
(n=42/256, 16%)

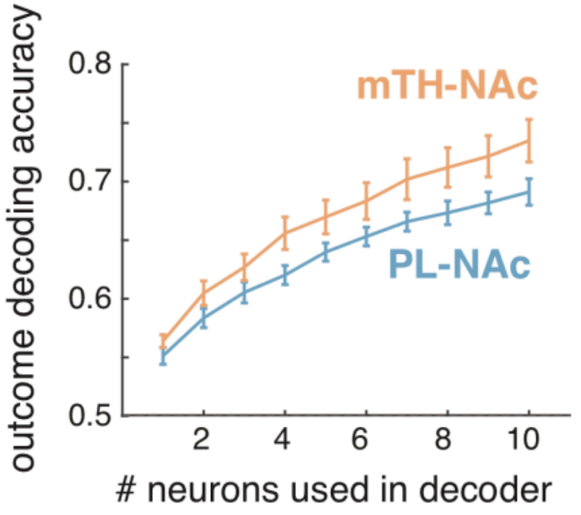
b

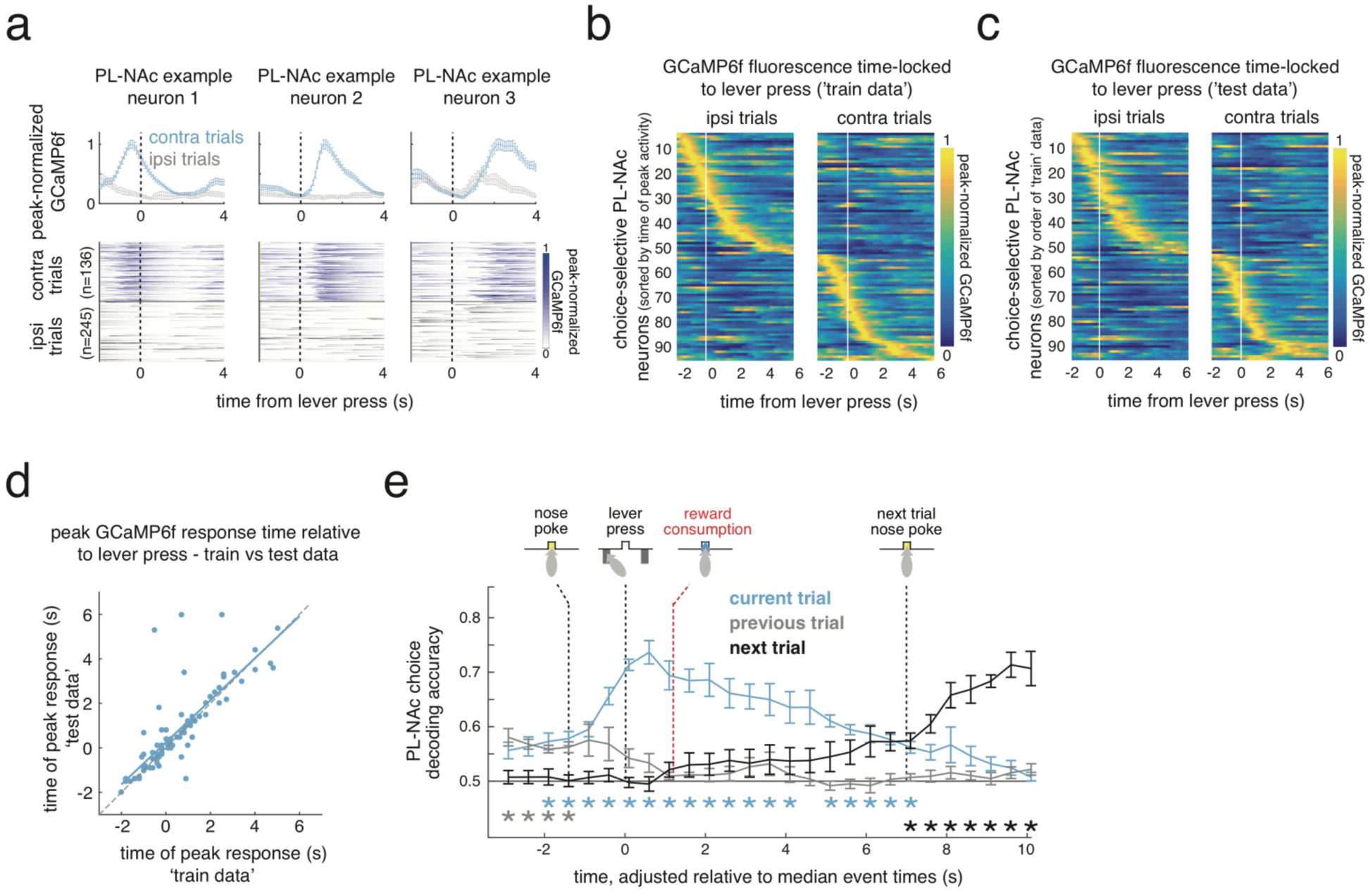

PL-NAc
mTH-NAc
choice decoding accuracy
0.8
0.7
0.6
0.5
2
4
6
8
10
neurons used in decoder

c

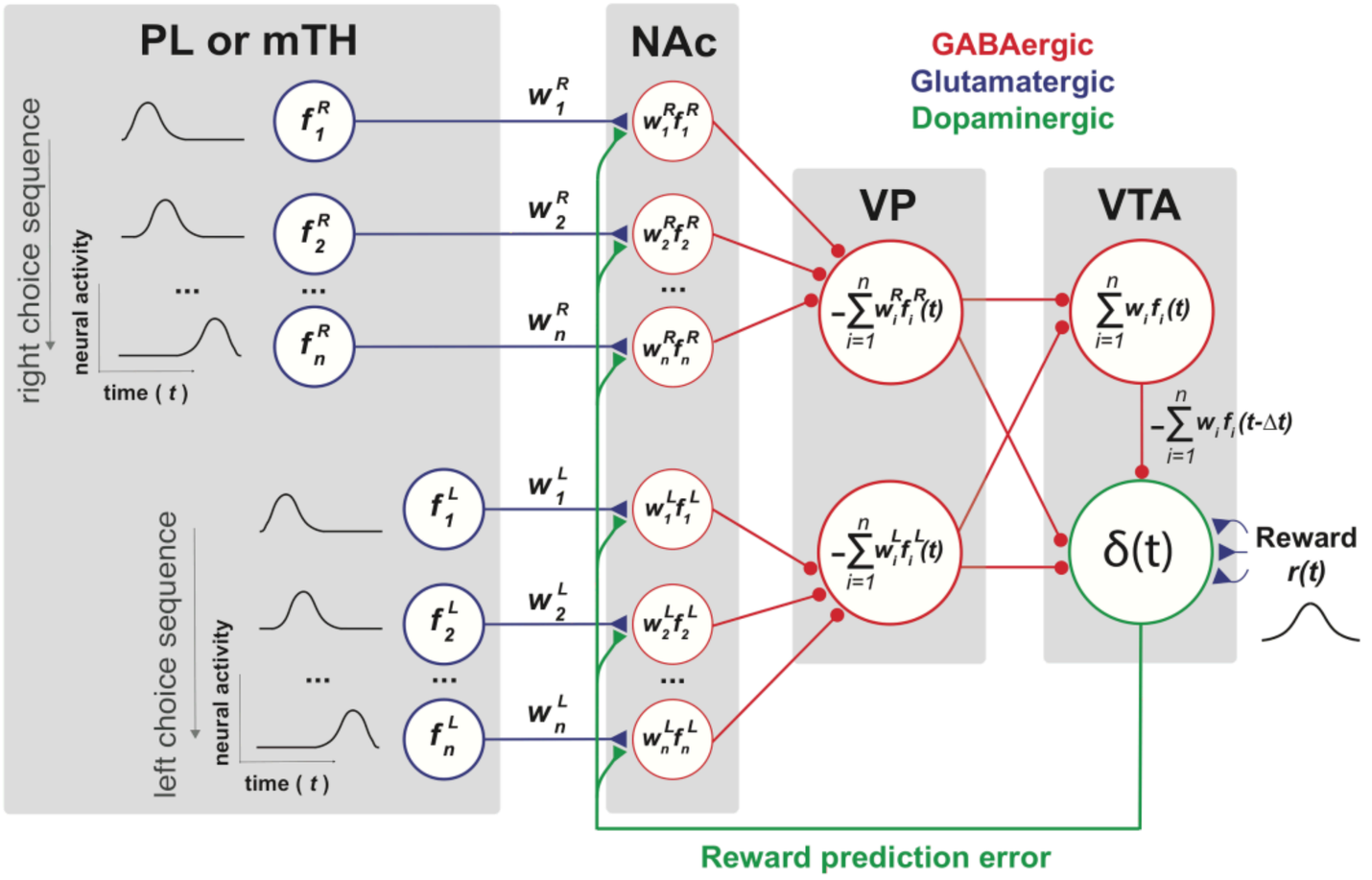

outcome-selective
mTH-NAc
(n=85/256, 33%)#
PL-NAc
(n=70/278, 25%)

d


mTH-NAc
PL-NAc
outcome decoding accuracy
0.8
0.7
0.6
0.5
2
4
6
8
10
neurons used in decoder


a
PL-NAc example neuron 1
PL-NAc example neuron 2
PL-NAc example neuron 3
peak-normalized GCaMP6f
contra trials
ipsi trials
contra trials (n=136)
ipsi trials (n=245)
peak-normalized GCaMP6f
time from lever press (s)
b
GCaMP6f fluorescence time-locked to lever press ('train data')
ipsi trials
contra trials
choice-selective PL-NAc neurons (sorted by time of peak activity)
peak-normalized GCaMP6f
time from lever press (s)
c
GCaMP6f fluorescence time-locked to lever press ('test data')
ipsi trials
contra trials
choice-selective PL-NAc neurons (sorted by order of 'train' data)
peak-normalized GCaMP6f
time from lever press (s)
d
peak GCaMP6f response time relative to lever press - train vs test data
time of peak response (s) 'test data'
time of peak response (s) 'train data'
e
nose poke
lever press
reward consumption
next trial nose poke
current trial
previous trial
next trial
PL-NAc choice decoding accuracy
time, adjusted relative to median event times (s)

a



b

$$V_{L,R}(t) = \sum_{i=1}^{n_{LR}} w_i^{L,R}(t) f_i^{L,R}(t)$$

$$V(t) = V_L(t) + V_R(t)$$

$$\delta(t) = r(t) + \frac{\gamma V(t) - V(t-\Delta)}{\Delta}$$

$$\frac{\mathrm{d}w_i(t)}{\mathrm{d}t} = \alpha \delta(t) E_i(t)$$

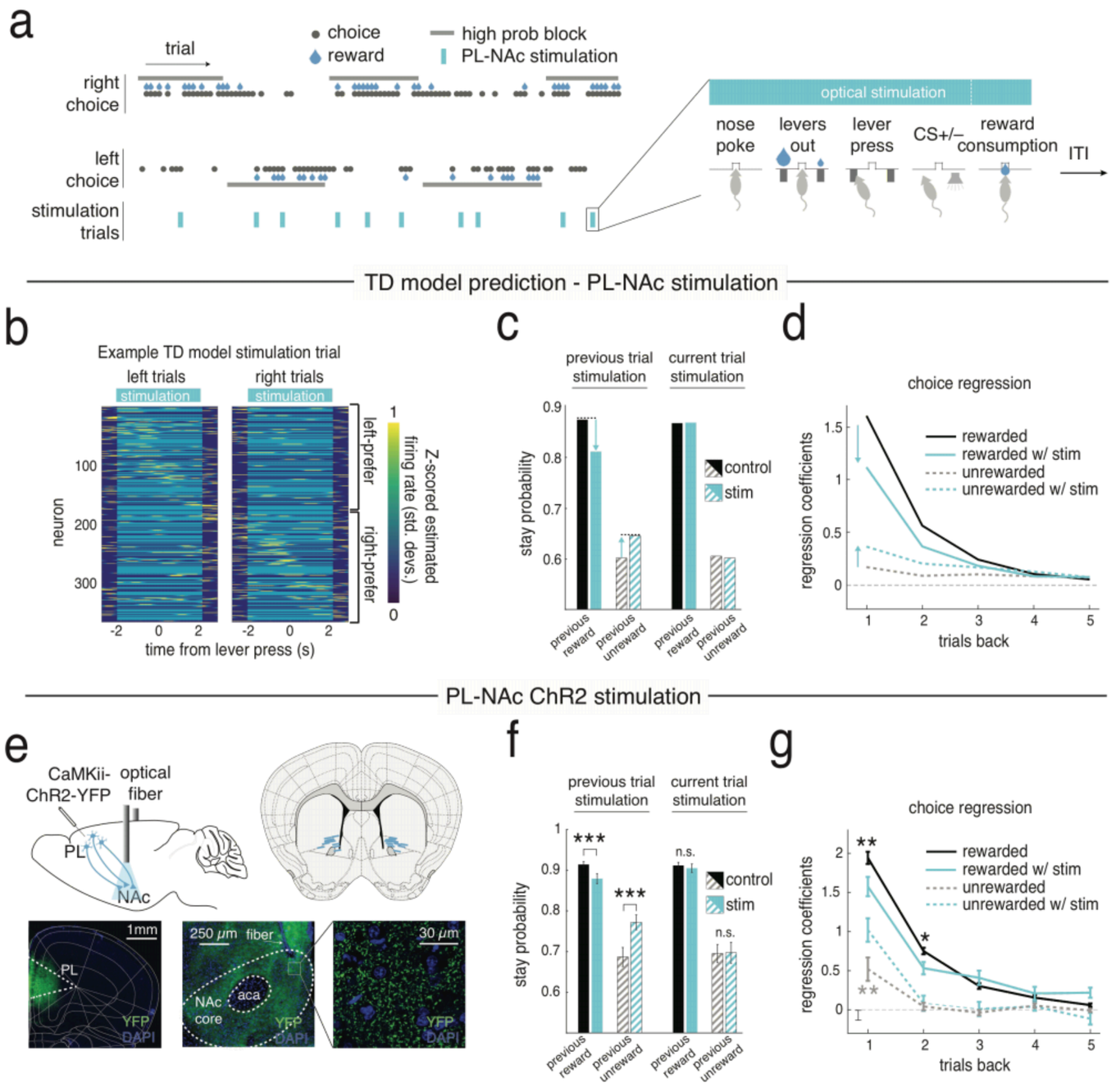
$$\frac{\mathrm{d}E_i(t)}{\mathrm{d}t} = \frac{-E_i(t)}{\tau_e} + f_i(t)$$

$$d_{left} = \mathbb{E}_t \left[\sum_{i=1}^{60} w_i^{left} n_i^{left}(t) \right]$$

$$d_{right} = \mathbb{E}_t \left[\sum_{i=1}^{60} w_i^{right} n_i^{right}(t) \right]$$

$$Prob(left) = \frac{exp(\beta_{value} d_{left} + \beta_{stay} I_{left})}{exp(\beta_{value} d_{left} + \beta_{stay} I_{left}) + exp(\beta_{value} d_{right} + \beta_{stay} I_{right})}$$

$$Prob(right) = \frac{exp(\beta_{value} d_{right} + \beta_{stay} I_{right})}{exp(\beta_{value} d_{left} + \beta_{stay} I_{left}) + exp(\beta_{value} d_{right} + \beta_{stay} I_{right})}$$


a
trial
choice
reward
high prob block
PL-NAc stimulation
right choice
left choice
stimulation trials
optical stimulation
nose poke
levers out
lever press
CS+/–
reward consumption
ITI
TD model prediction - PL-NAc stimulation
b
Example TD model stimulation trial
left trials
right trials
stimulation
left-prefer
right-prefer
neuron
Z-scored estimated firing rate (std. devs.)
time from lever press (s)
c
previous trial stimulation
current trial stimulation
stay probability
control
stim
previous reward
previous unreward
d
choice regression
regression coefficients
rewarded
rewarded w/ stim
unrewarded
unrewarded w/ stim
trials back
PL-NAc ChR2 stimulation
e
CaMKii-ChR2-YFP
optical fiber
PL
NAc
1mm
250 μm fiber
30 μm
aca
NAc core
YFP
DAPI
f
***
n.s.
g
**
*