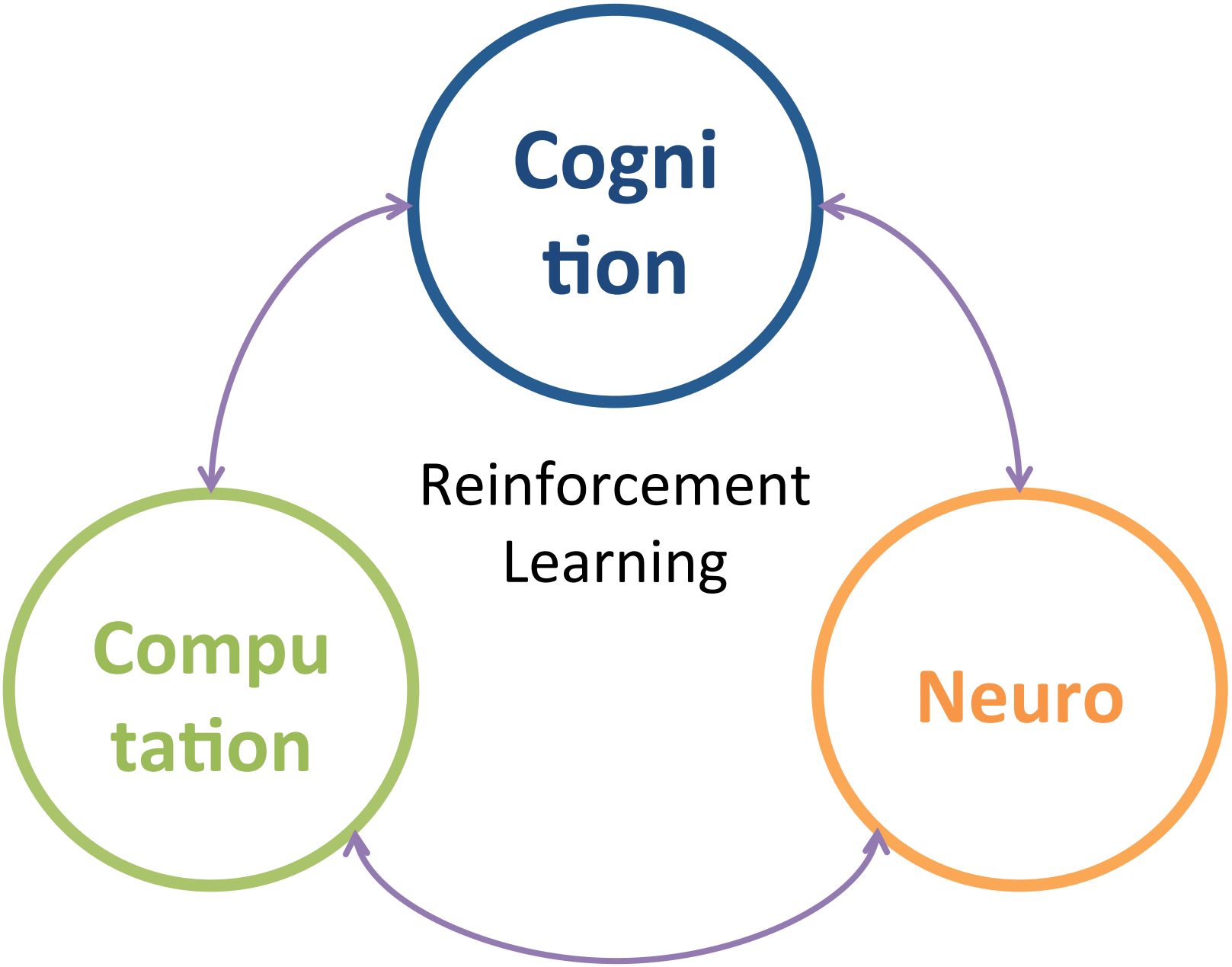
Working Memory Influences Reinforcement Learning Computations in Brain and Behavior

Anne Collins - UC Berkeley

2018/4 Simons Institute


Cogni
tion
Reinforcement
Learning
Compu
tation
Neuro

RL algorithms

Estimate value of a choice: V

prediction error: [Observed - predicted] reward

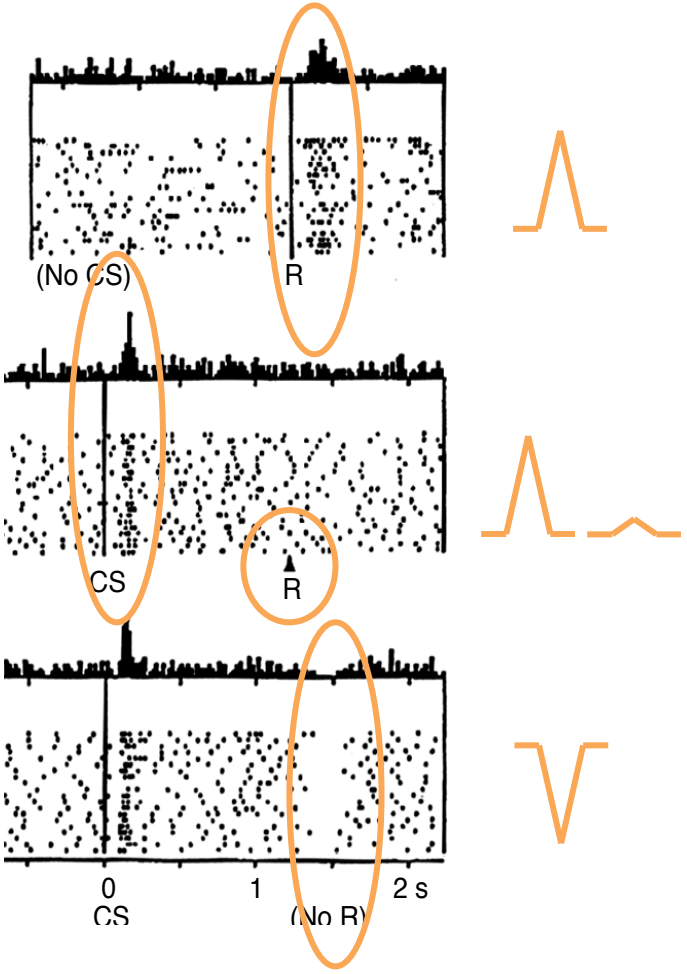
Update estimate:

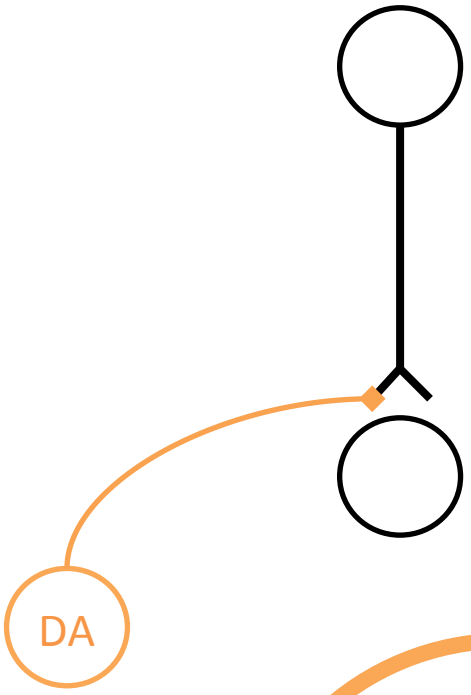
Computation

$$V_{t+1} \leftarrow V_t + \alpha(r_t - V_t)$$

Sutton, Barto 1998

Dopamine neurons encode reward prediction error





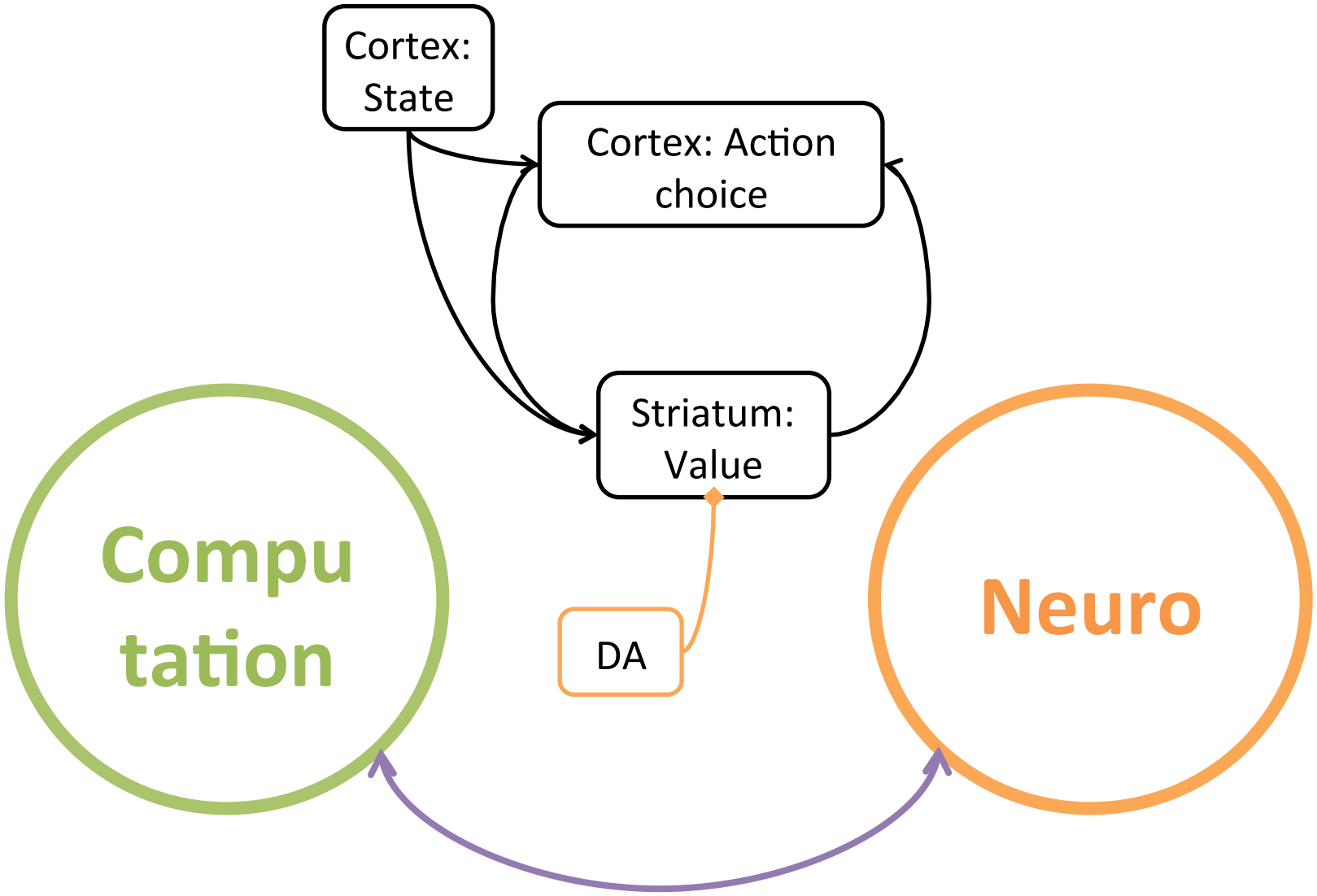
Cortex: state s

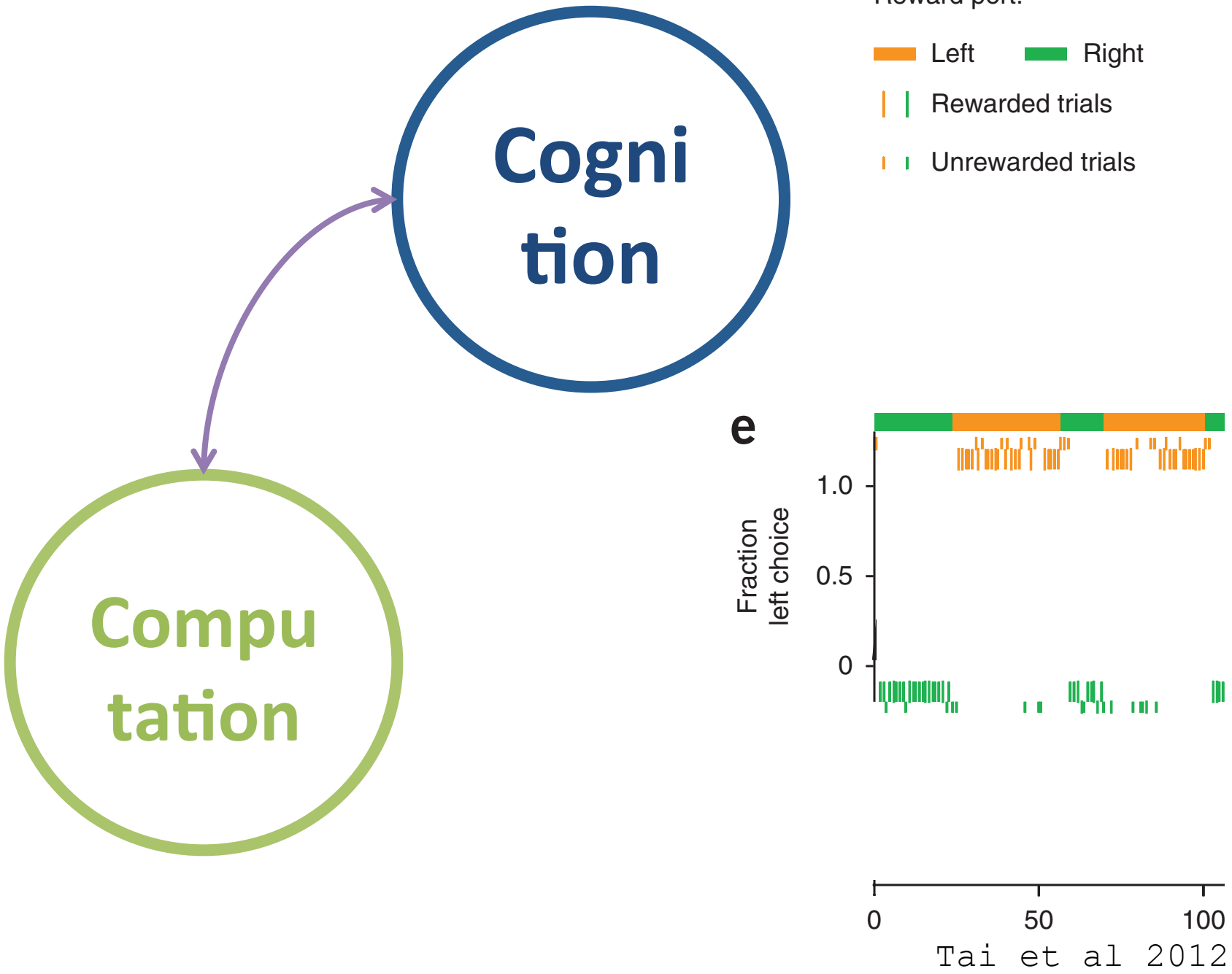
Striatum: choice a

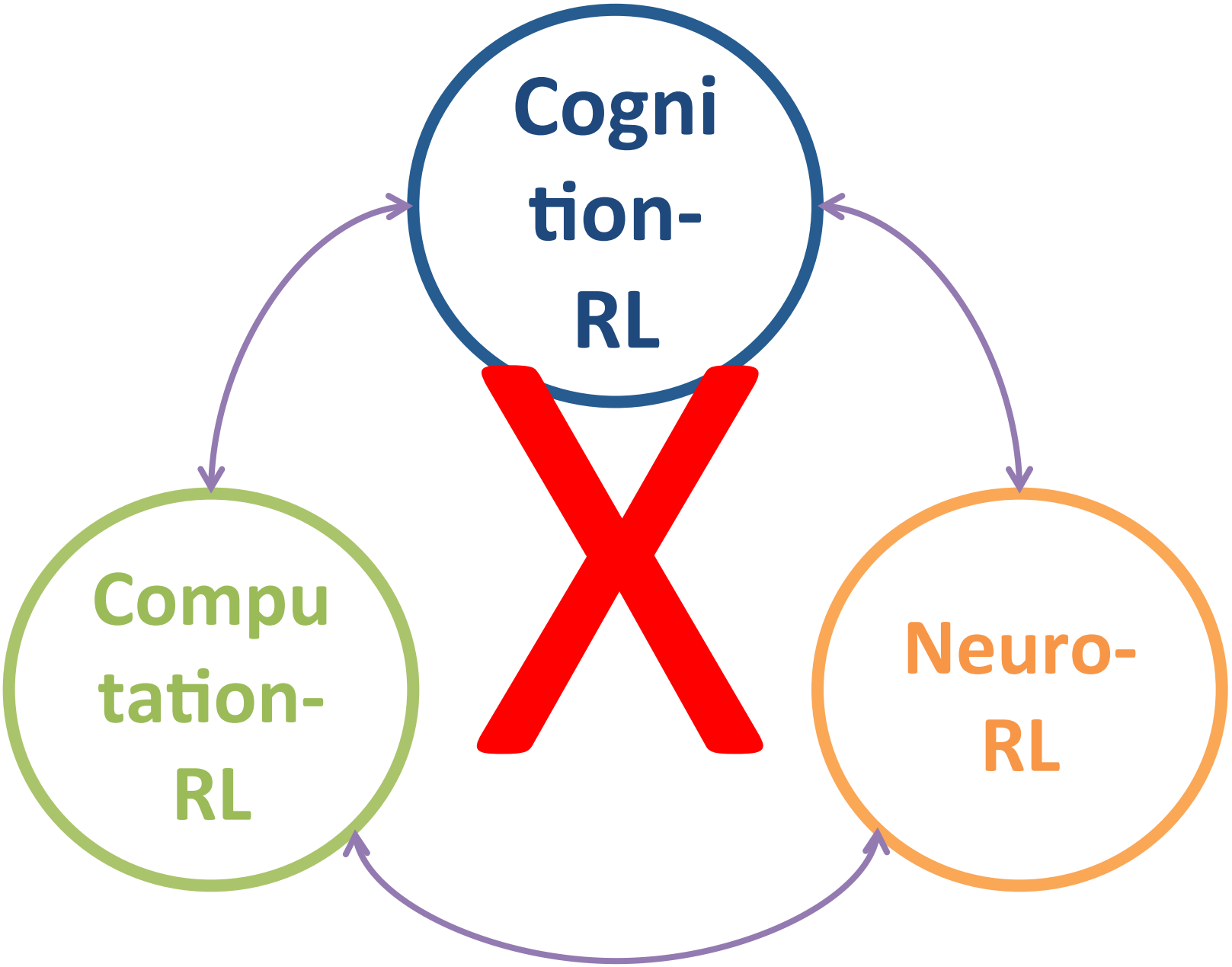
Neuro

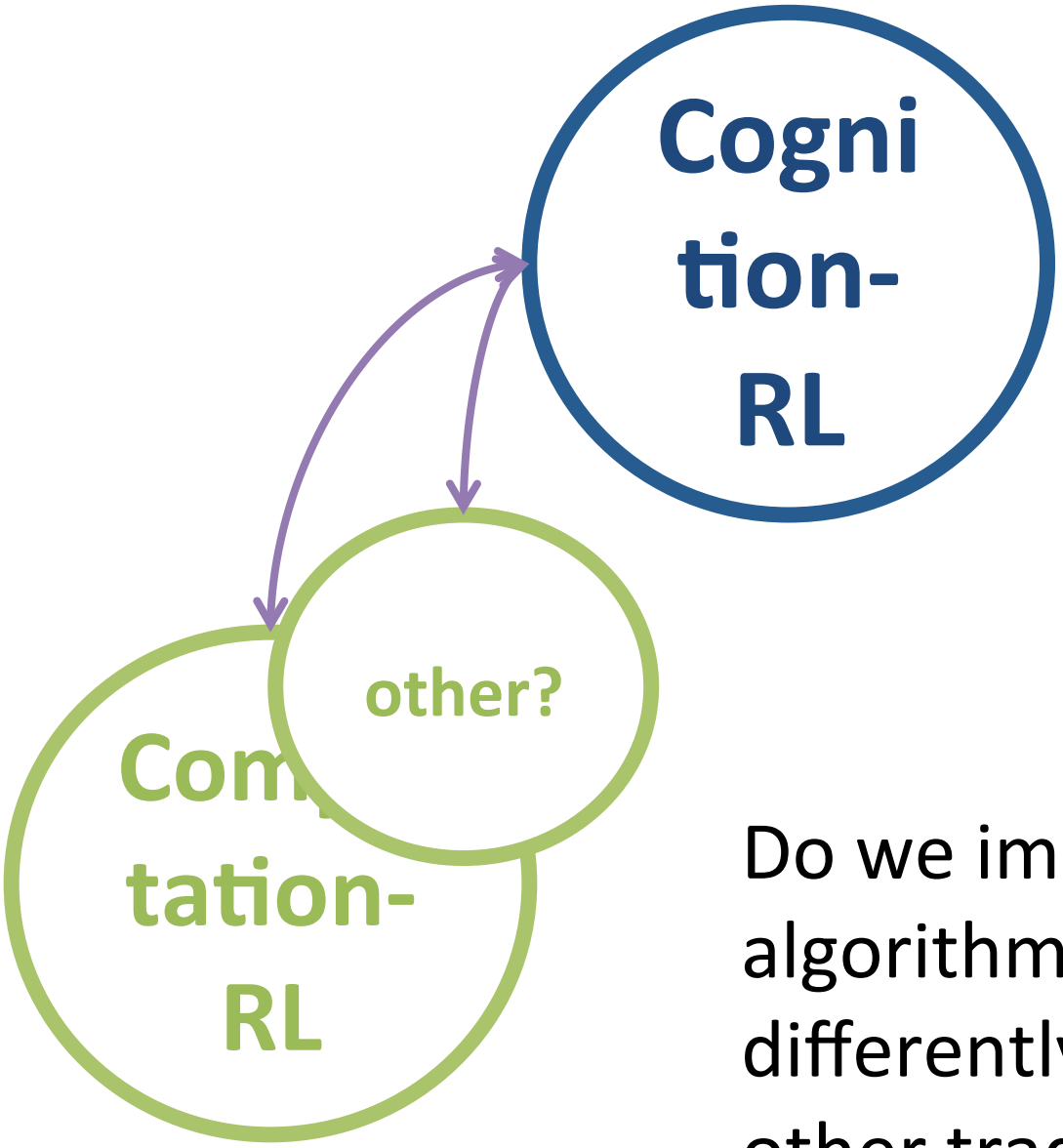
Montague et al 1996, Schultz et al 1997

$$V_{t+1} \leftarrow V_t + \alpha(r_t - V_t)$$

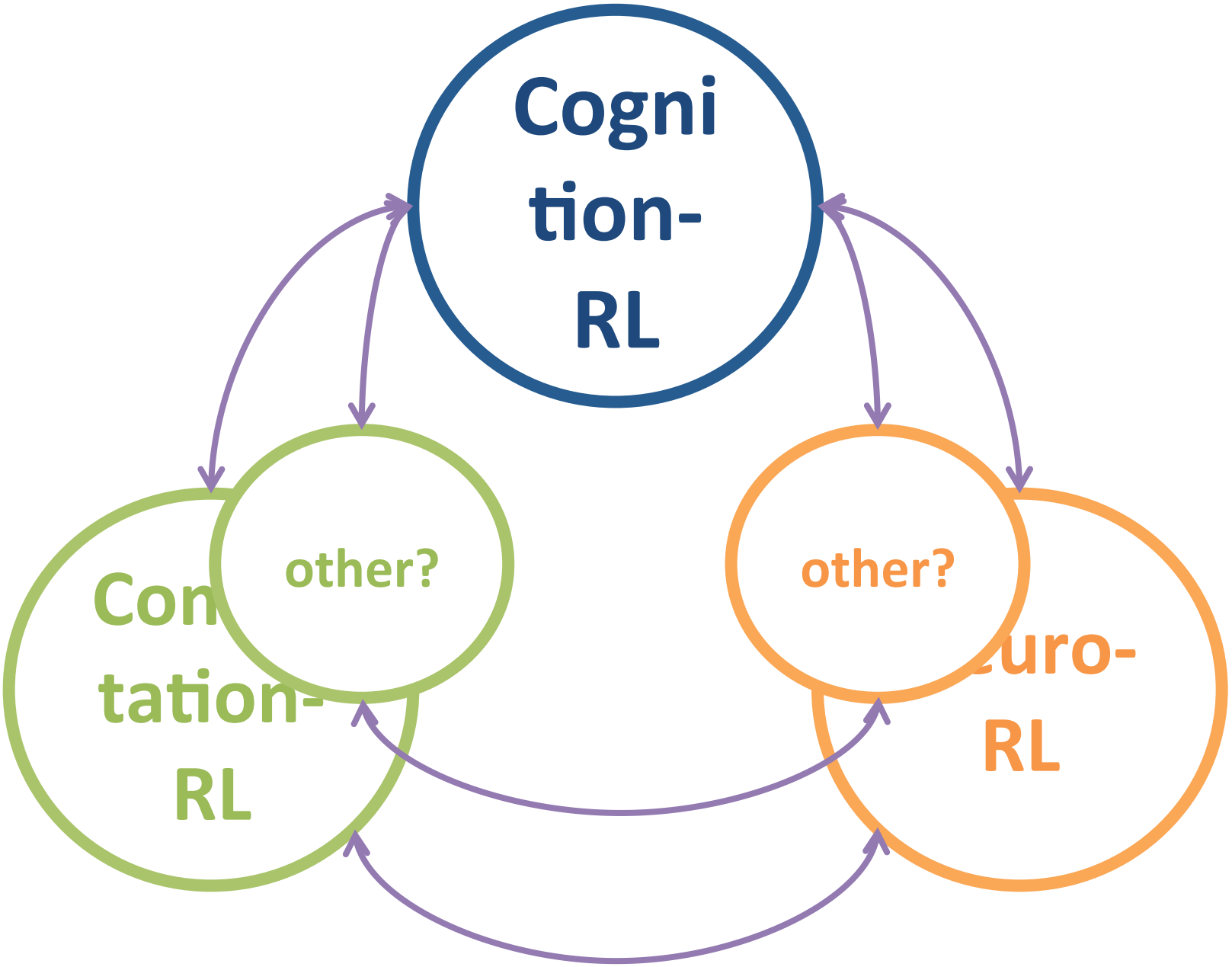



Cogni
tion
Compu
tation
Reward port:
Left
Right
Rewarded trials
Unrewarded trials
e
Fraction
left choice
1.0
0.5
0
0
50
100
Tai et al 2012


Cogni
tion-
RL
Compu
tation-
RL
Neuro-
RL



Do we implement other algorithms that learn differently and optimize other tradeoffs?


Cogni
tion-
RL
other?
other?
Con
tation-
RL
uro-
RL

Two components to learning:

working memory

reinforcement learning

Why it's important:

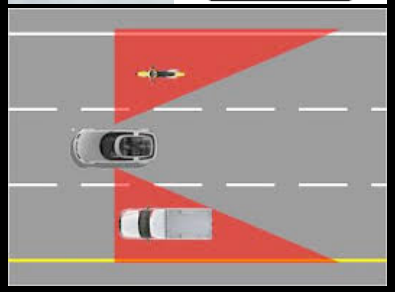
Understanding Schizophrenia

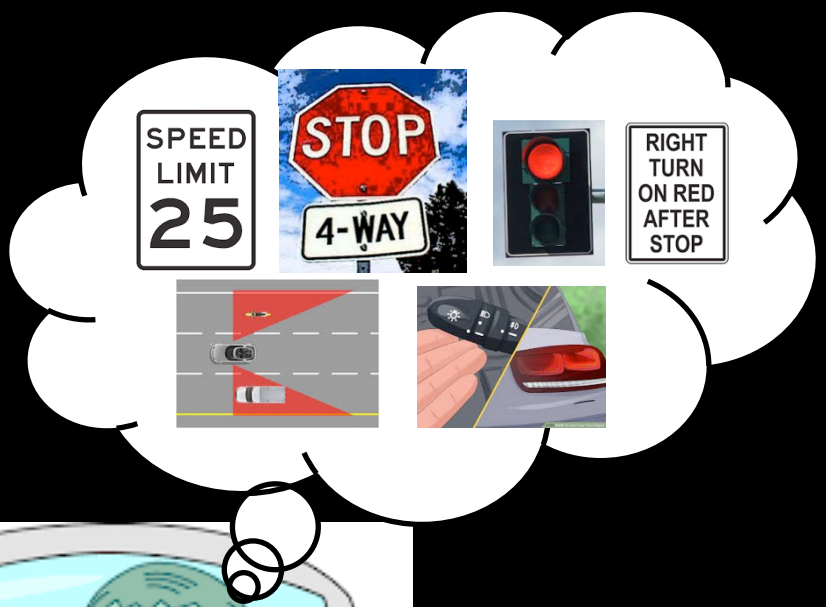
Are WM and RL's computations independent?

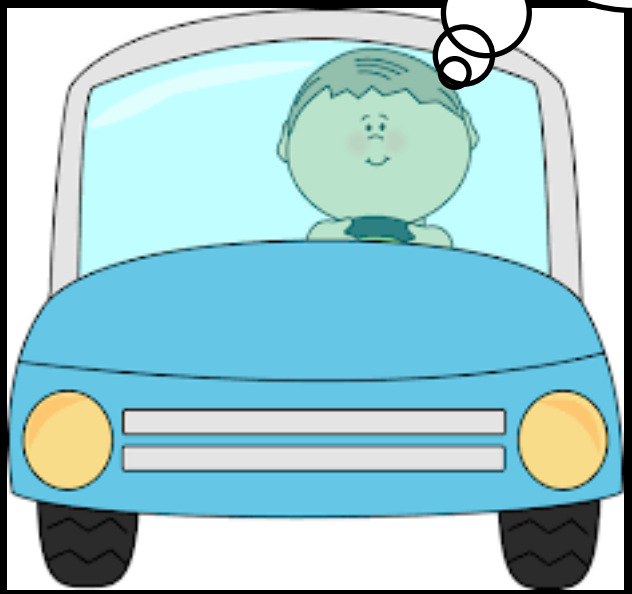

SPEED
LIMIT
25


STOP
4-WAY


RIGHT
TURN
ON RED
AFTER
STOP




SPEED
LIMIT
25
STOP
4-WAY
RIGHT
TURN
ON RED
AFTER
STOP

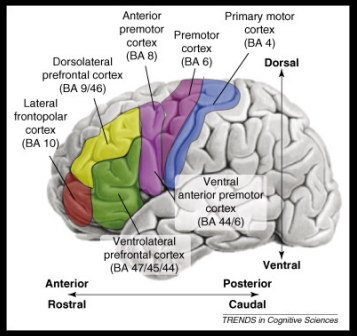



YOUR
SPEED
458

(at least) Two components for learning: optimized for different trade-offs

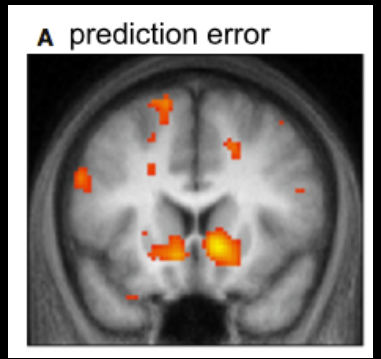
Working Memory (WM)

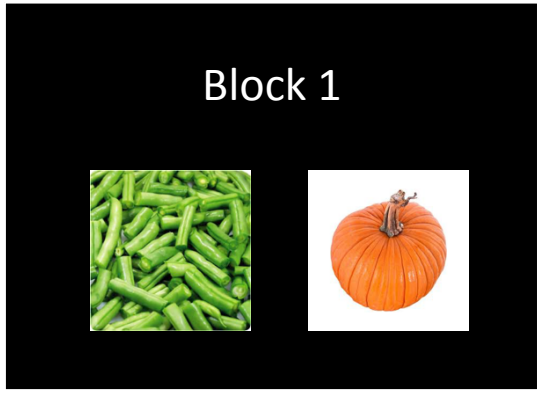
- Fast
- Precise event information storage
- Flexible
- Resource/capacity limited
- Short term
- Costly



Instrumental value learning (RL)

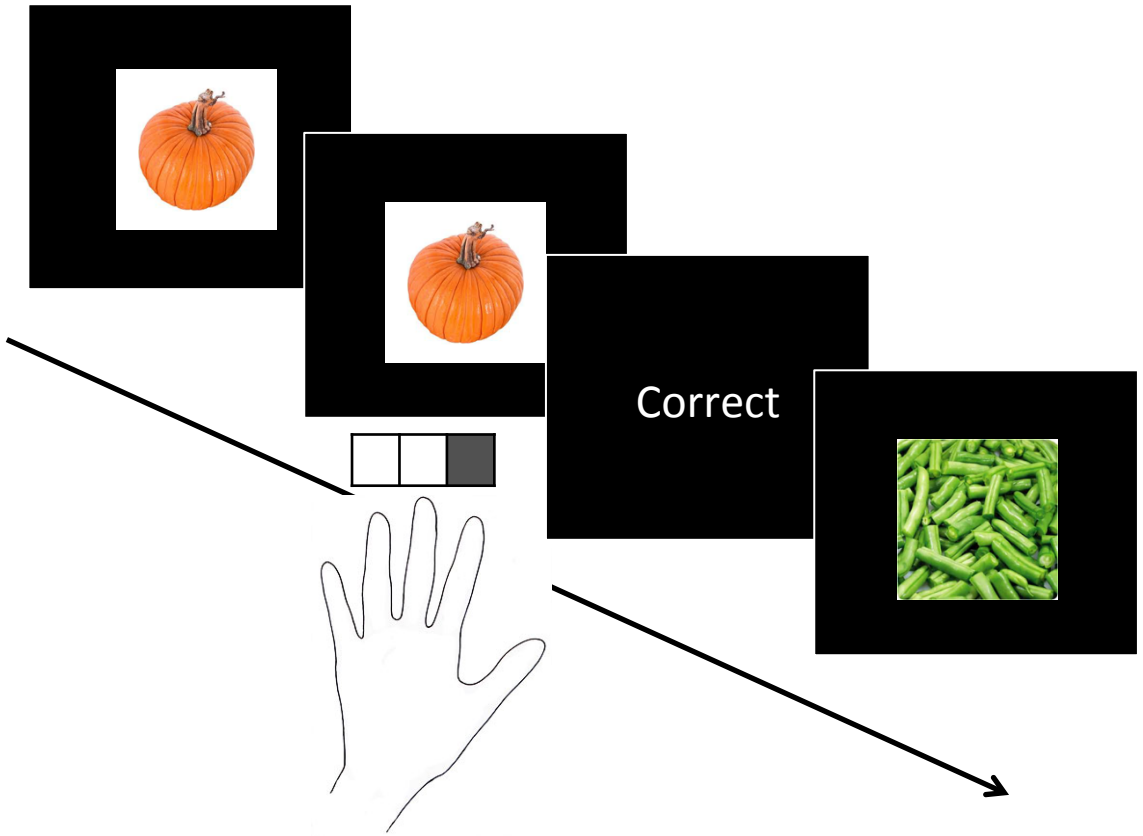
- Incremental
- Integrated value information storage
- Inflexible
- Broad capacity
- Long term robust
- Effortless





Beginning of block 1.
n_S = **2** stimuli to learn.

Beginning of block 2.
n_S = **6** stimuli to learn.



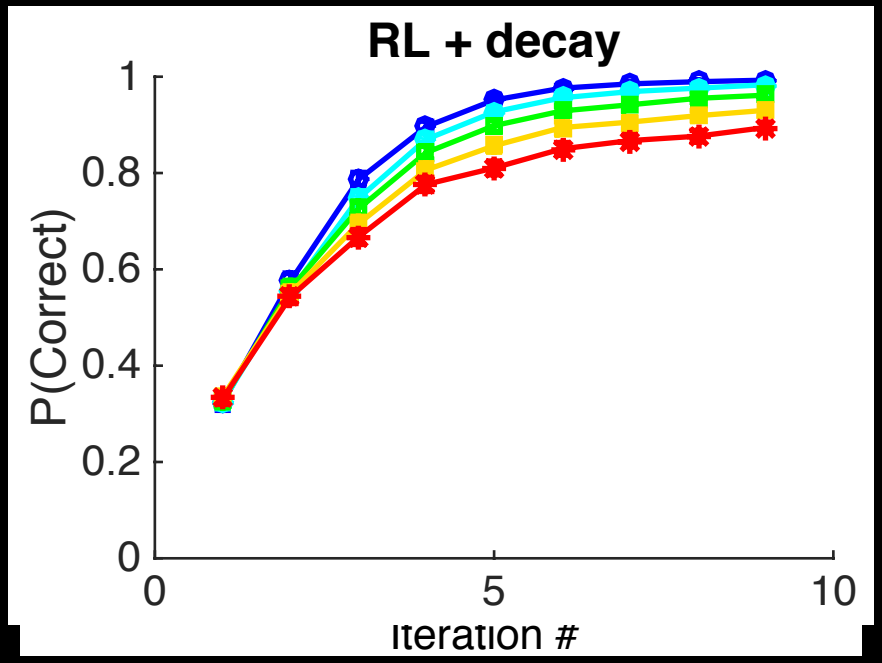
Reinforcement learning:

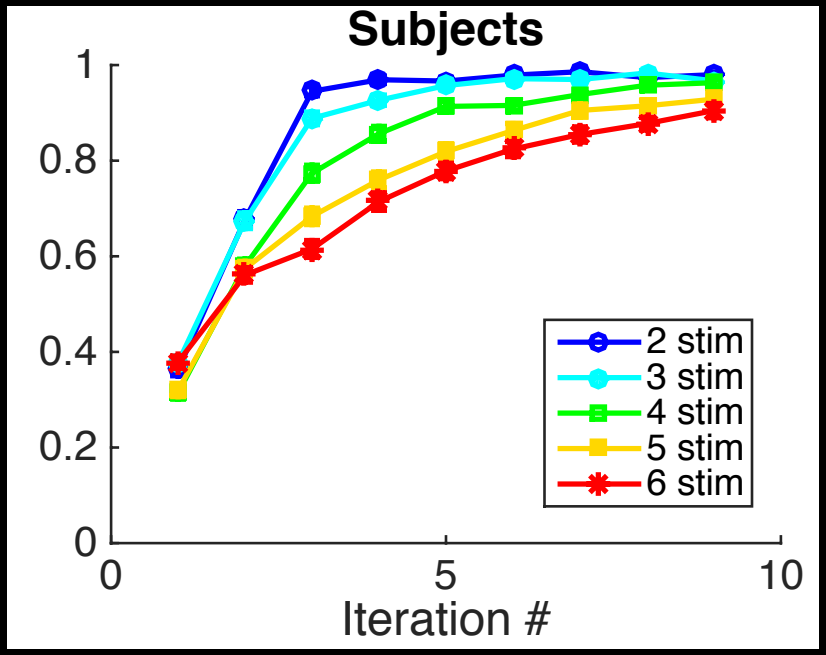
- test effect of reward history [how many past correct choices?]

Set-size manipulation:

- test WM **load** effect [**how many** items to remember?]
- test WM **decay** effect [**how long ago** did I store an item?]

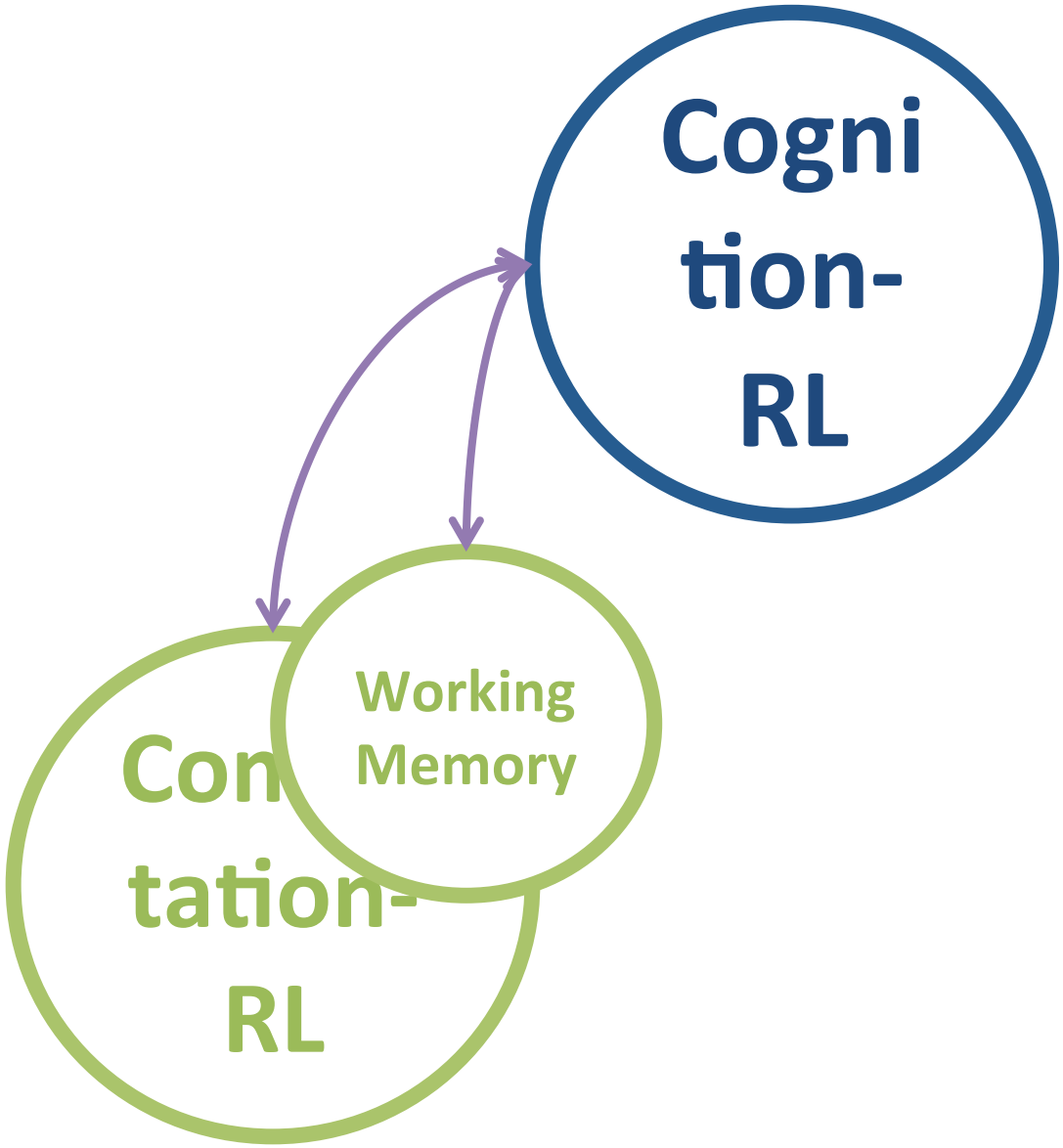
Experimental results: n=78



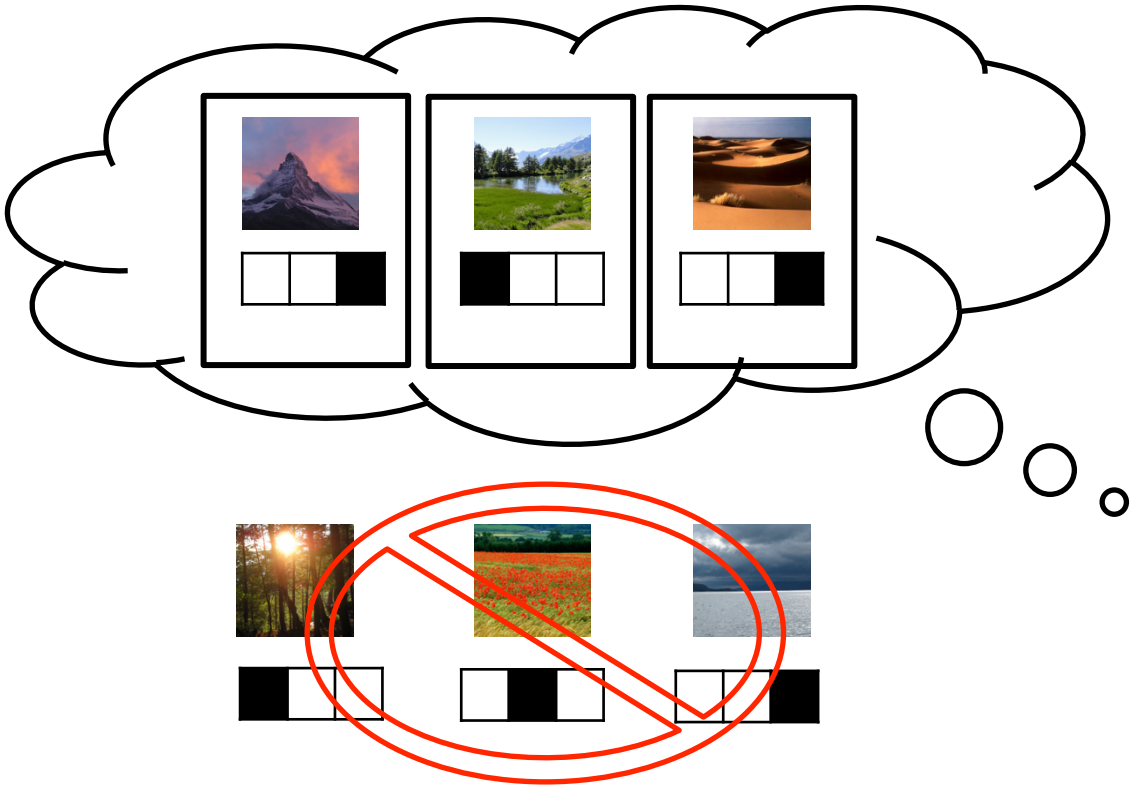


Collins & Frank, 2012

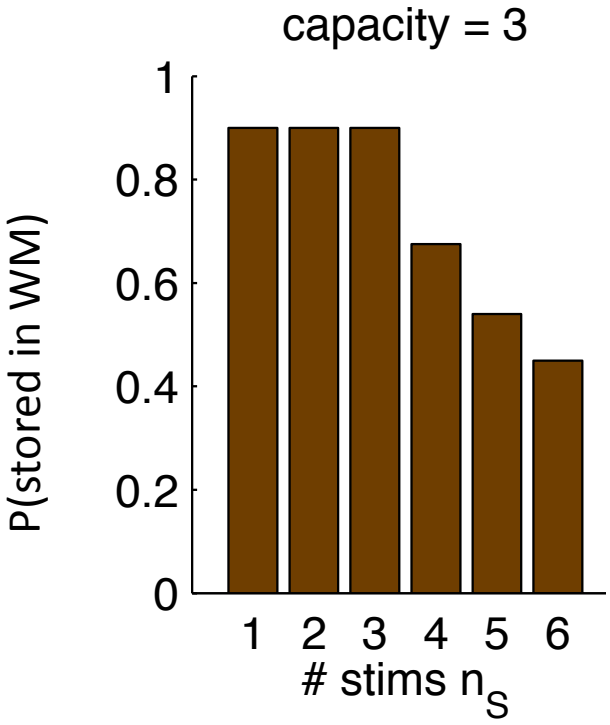
Effect of set-size on learning performance is not accounted for by classic RL


Cogni
tion-
RL
Working
Memory
Con
tation-
RL

Working memory (WM)



1) O'Reilly & Frank 2006
2) Luck & Vogel 2013
3) Bays & Husain 2008
4) Baddeley 2012
5) Collins & Frank 2012



WM: fast storage, reuse of limited amount of information

WM+RL model

RL module:

Learning rate α

WM module:
Fast storage
Decay over time
Capacity **K**

1-w

w

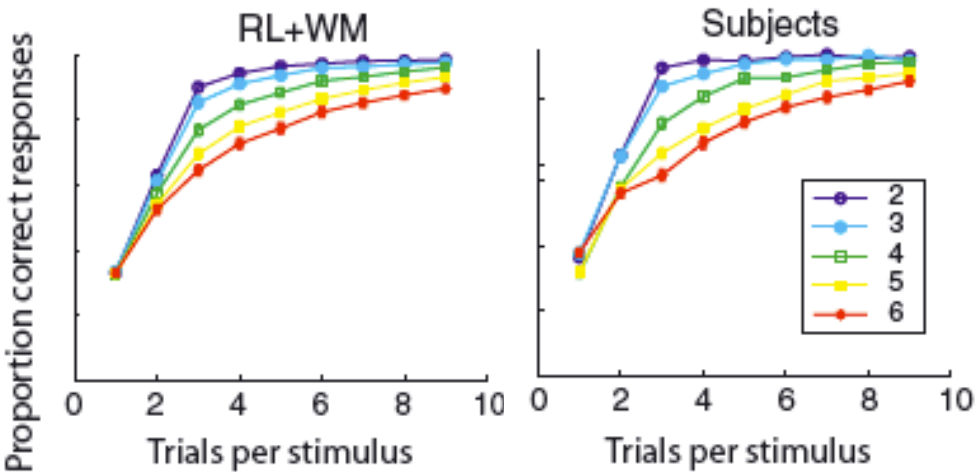
RL behavior:

Key choice
for each image

Mixture policy:

w(t) = weight of WM vs. RL

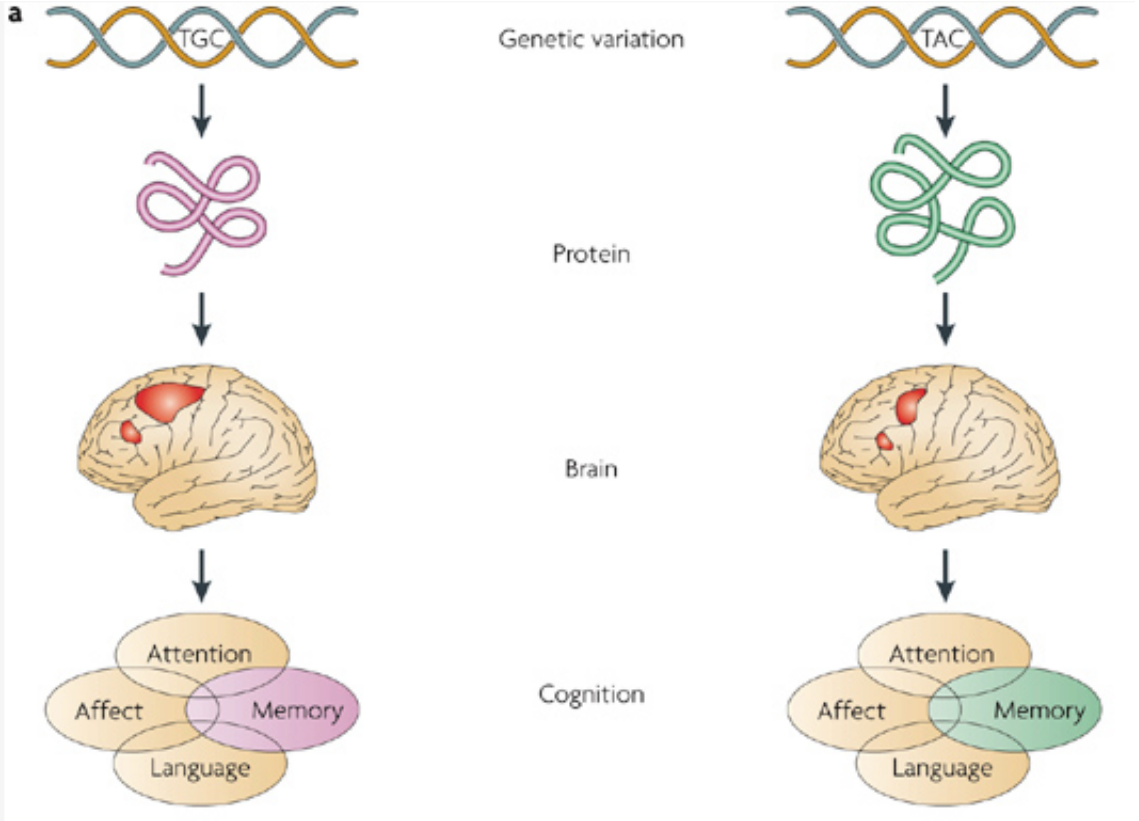
Best fitting model behaviors:



Collins & Frank, 2012

Genes dissociate prefrontal and N-RL contributions to learning.

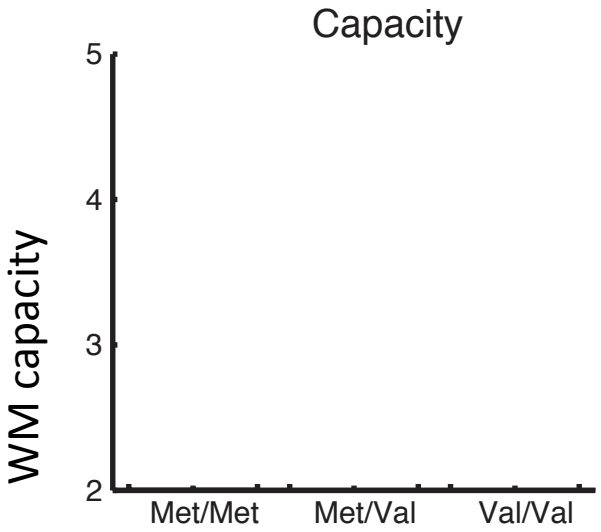
- **COMT**: gene linked to prefrontal WM function [2]
- **GPR6**: gene expressed only in striatum [3], linked to plasticity.



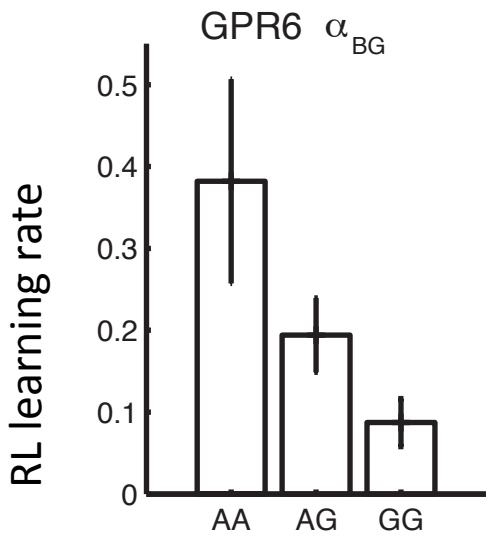
1) Collins & Frank, 2012
2) Frank et al, 2007
3) Lobo et al, 2007
4) Green et al 2008

Genes dissociate prefrontal and N-RL contributions to learning.

Do differences in genotype predict differences in distinct components of behavior, as infered by model fit parameters?

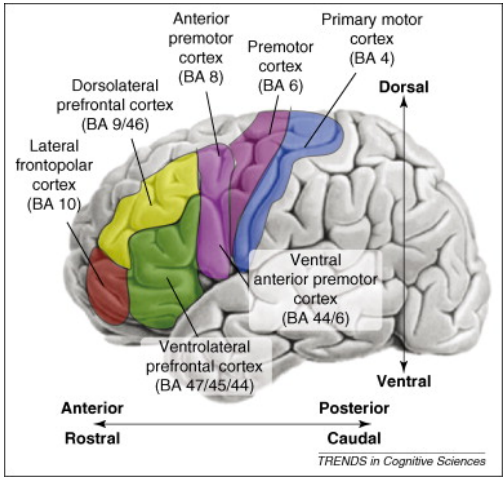


COMT genotype:
Prefrontal function



GPR6 genotype:
N-RL function

1) Collins & Frank, 2012
2) Frank et al, 2007
3) Lobo et al, 2007



Cog-RL

Learning rate α

Capacity K

Working Memory

AI-RL

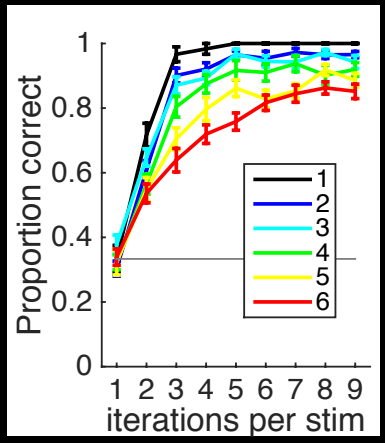
Prefrontal

N-RL

Indicator of Prefrontal function

Indicator of striatal N-RL function

Learning is a mixture of multiple neuro-cognitive processes that implement different computational trade-offs and rely on different neural mechanisms



Two components to learning:

working memory

reinforcement learning

Why it's important:

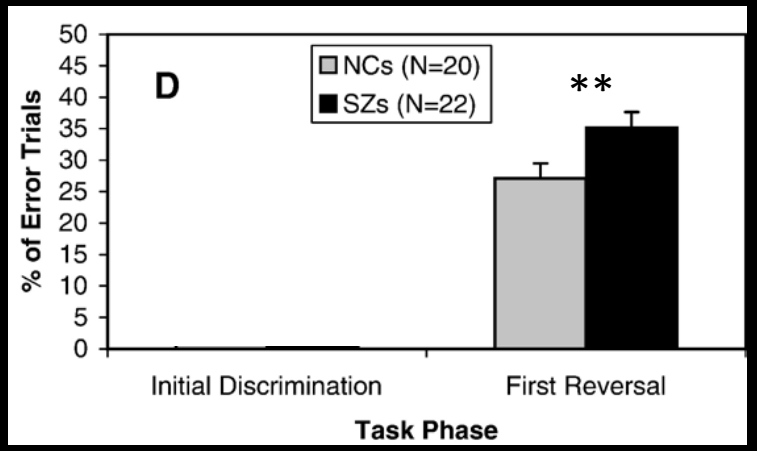
Understanding Schizophrenia

Are WM and RL's computations independent?

Working Memory and learning impairments in Schizophrenia

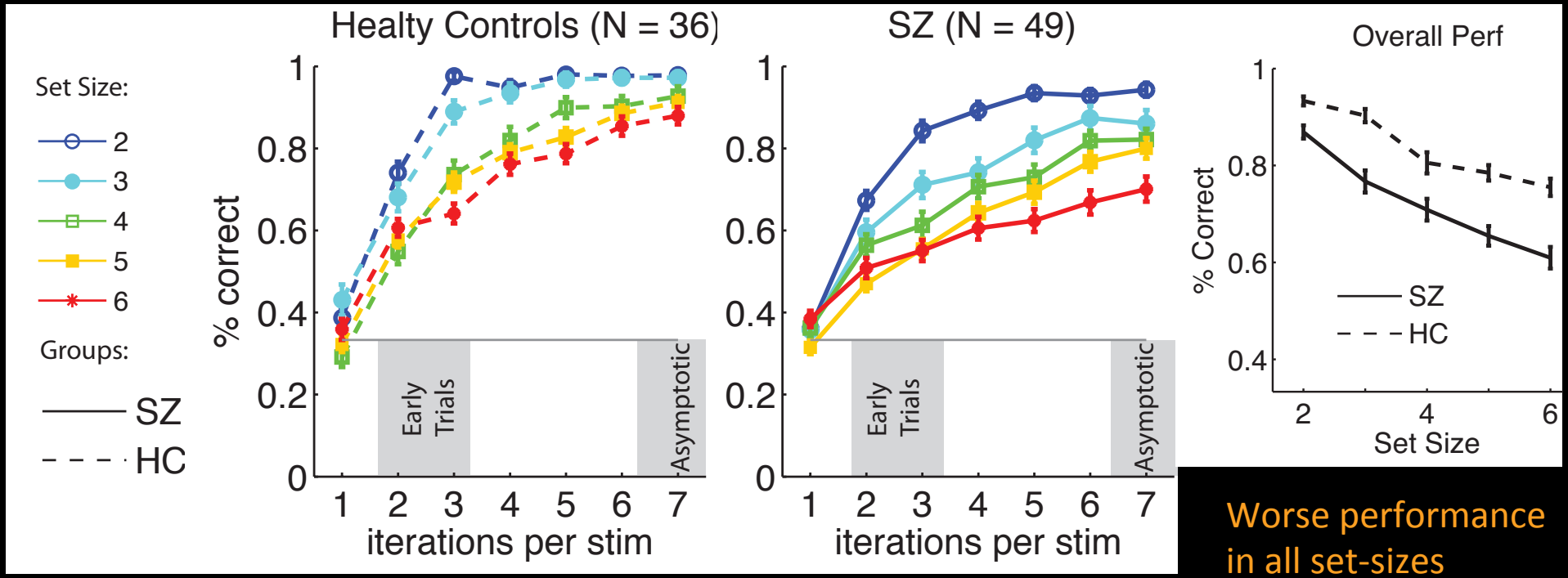
- Prefrontal cortex-dependent impairment: eg.
  - Lee, J. and Park, S. (2005) J. Abnorm. Psychol.
  - Forbes, N.F. et al. (2009) Psychol. Med.
  - Barch & Caeser (2012) TRiCS
- Learning impairments:
  - Ambiguous litterature

Waltz & Gold 2007



Deserno et al 2013 for review

What causes learning deficits in schizophrenia?

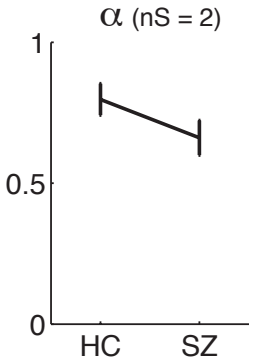


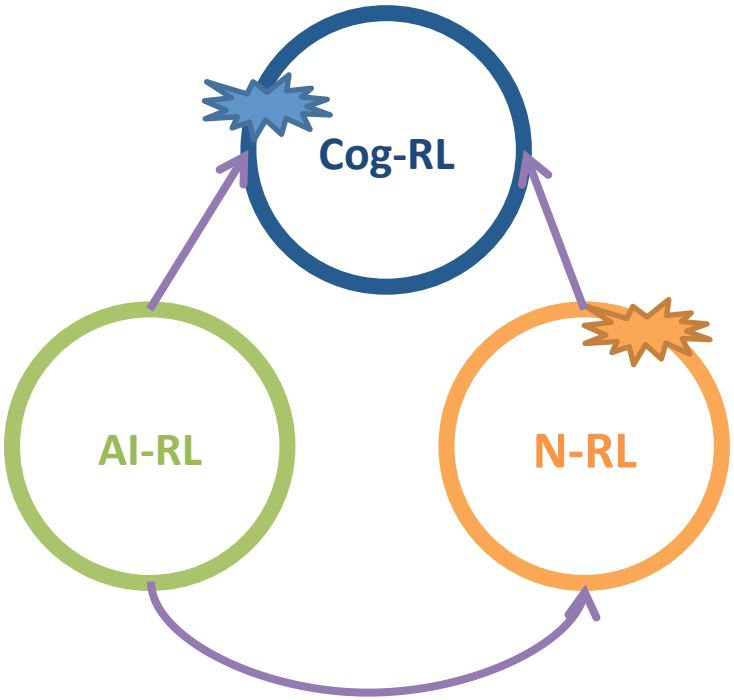
Worse performance in all set-sizes

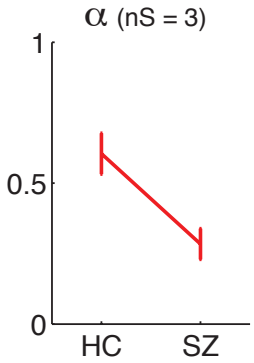
RL-only model leads to ambiguous results

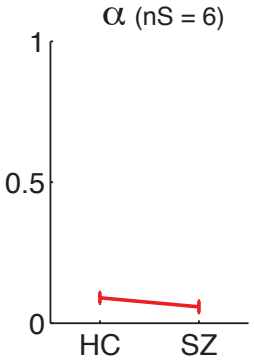
RL model only:

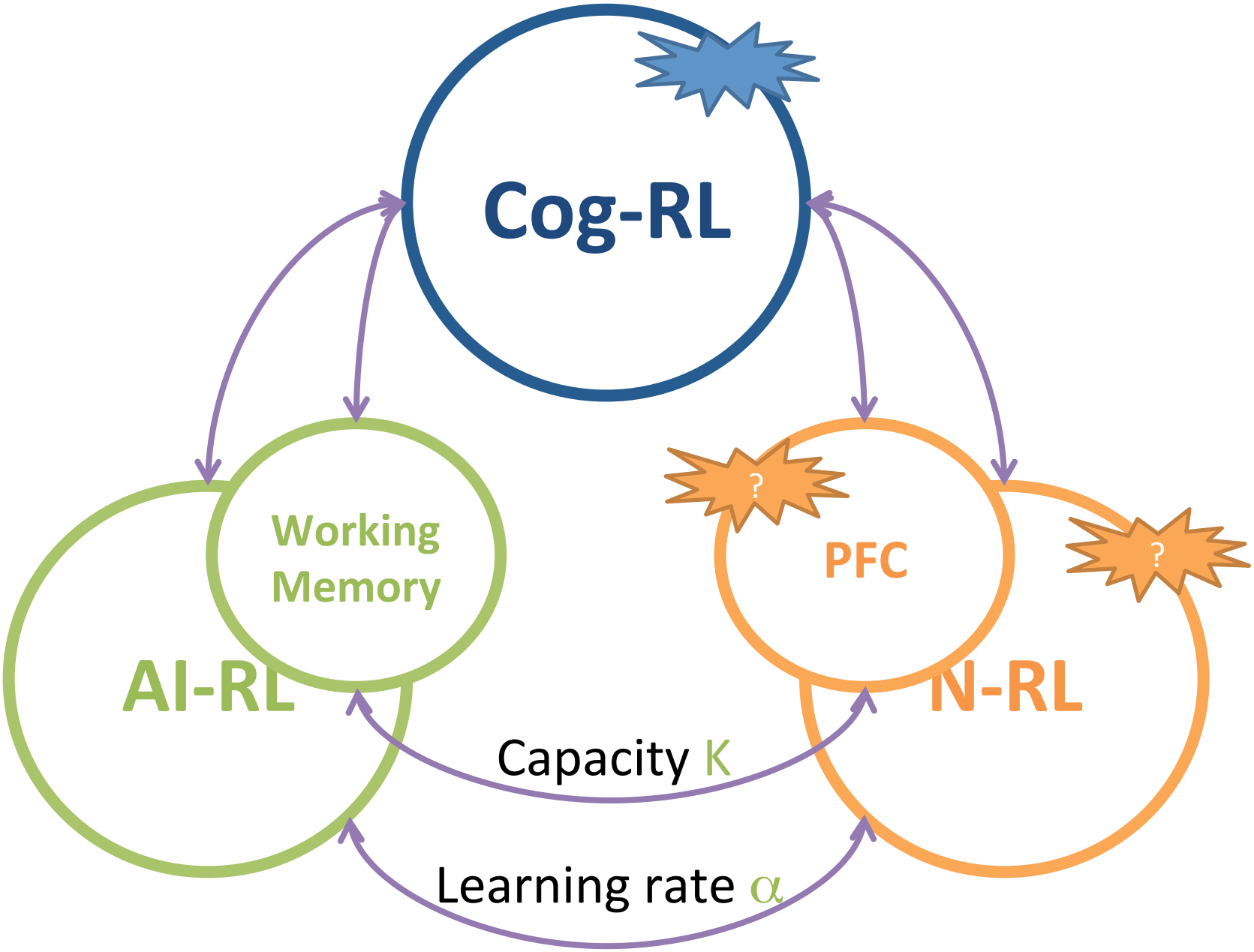
- Set size two: no impairment in RL
- High set size: impairment in RL





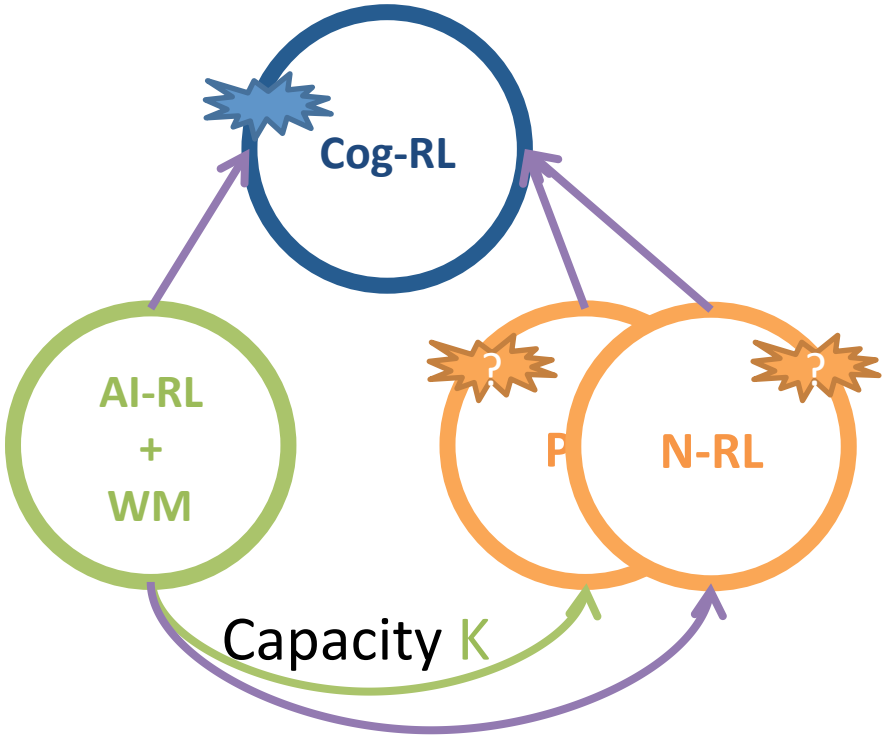


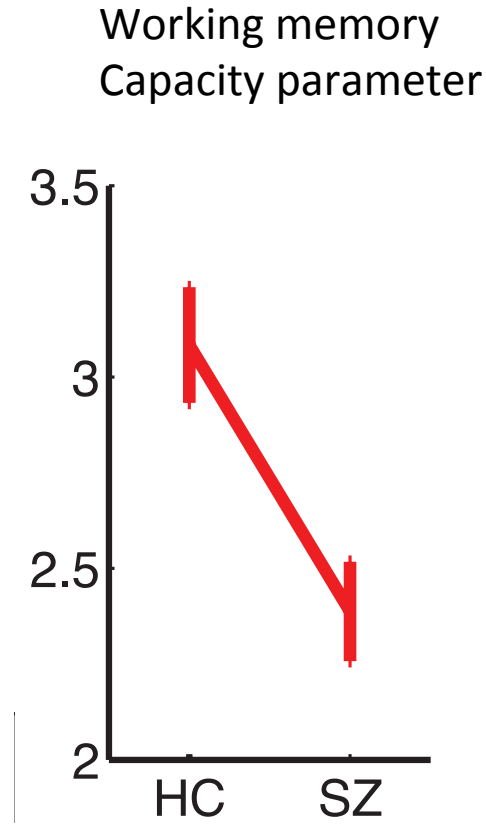



Cog-RL
Working Memory
AI-RL
PFC
N-RL
?
?
Capacity K
Learning rate α

Are deficits due to prefrontal working memory impairment?

- RLWM model fit: Working memory parameters impaired in patients



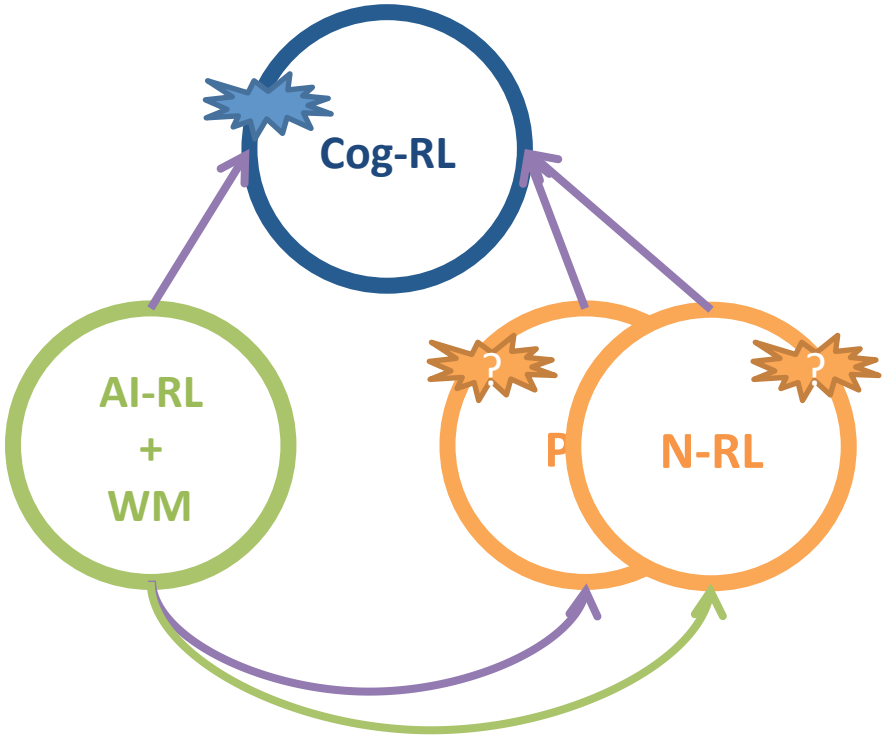


HC: healthy controls
SZ: patients

Collins et al, 2014, JoN

Are deficits due to N-RL impairment?

RLWM model fit: No effect of condition on RL parameters

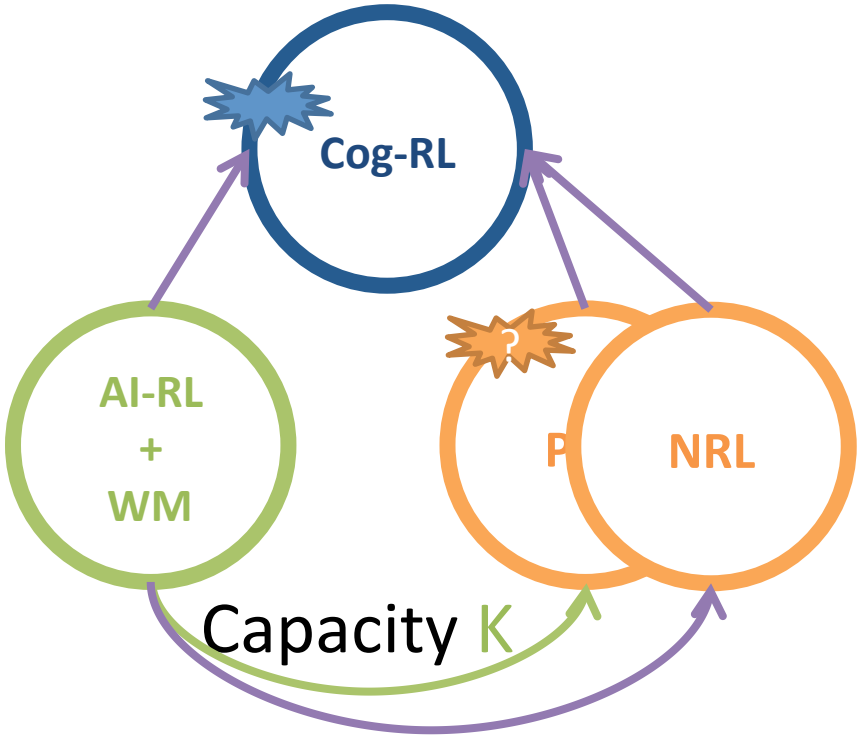


HC: healthy controls
SZ: patients

Collins et al, 2014, JoN

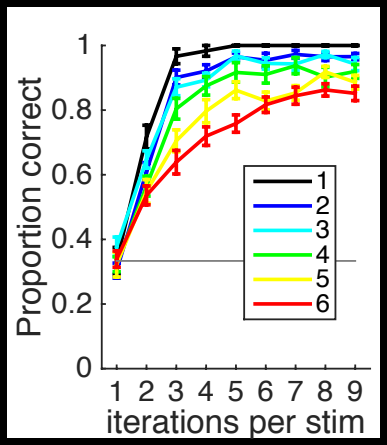
Learning deficits in patients are not due to N-RL system impairment

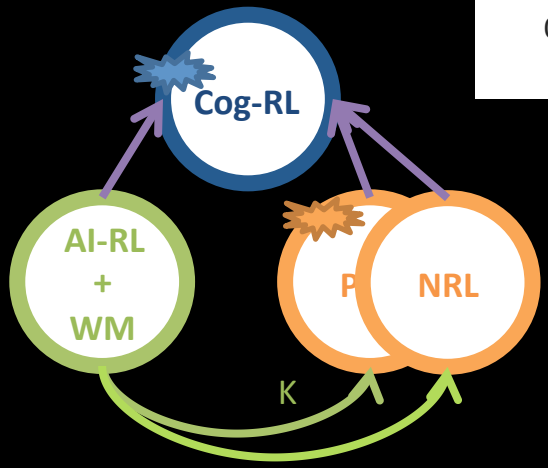
They are due to PFC-WM impairment.



Learning is a mixture of multiple neuro-cognitive processes that implement different computational trade-offs and rely on different neural mechanisms

We can isolate SZ learning dysfunction as WM, not RL





Two components to learning:

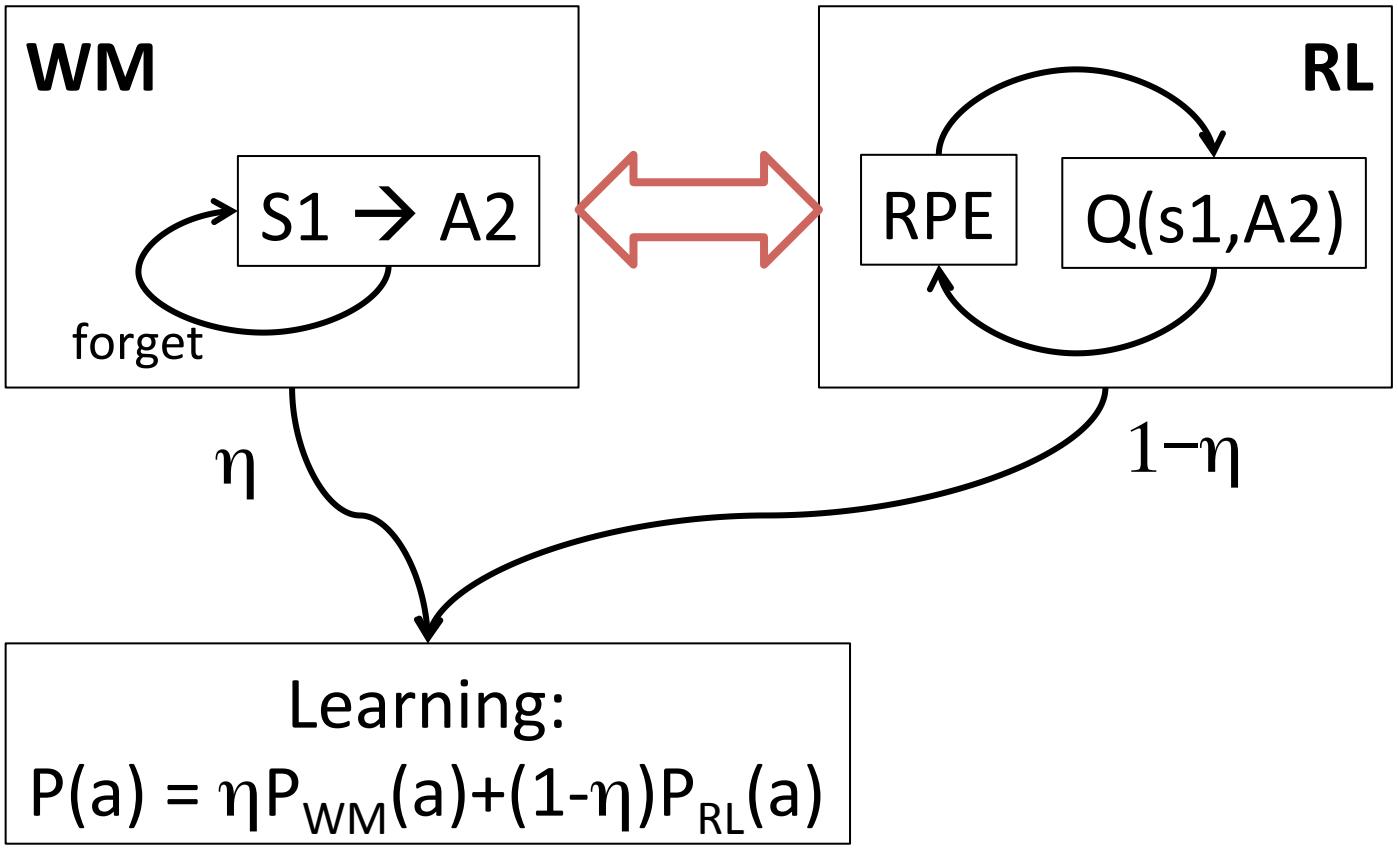
working memory

reinforcement learning

Why it's important:

Understanding Schizophrenia

Are WM and RL's computations independent?


WM
S1 → A2
forget
RL
RPE
Q(s1,A2)
η
1−η
Learning:
P(a) = ηP_WM(a)+(1-η)P_RL(a)

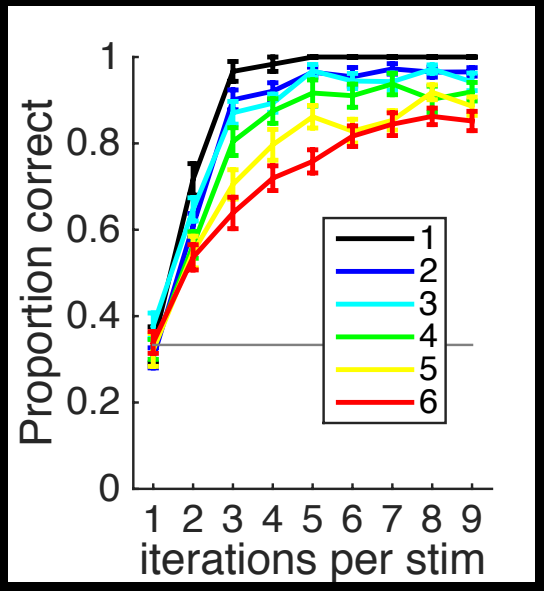
EEG experiment

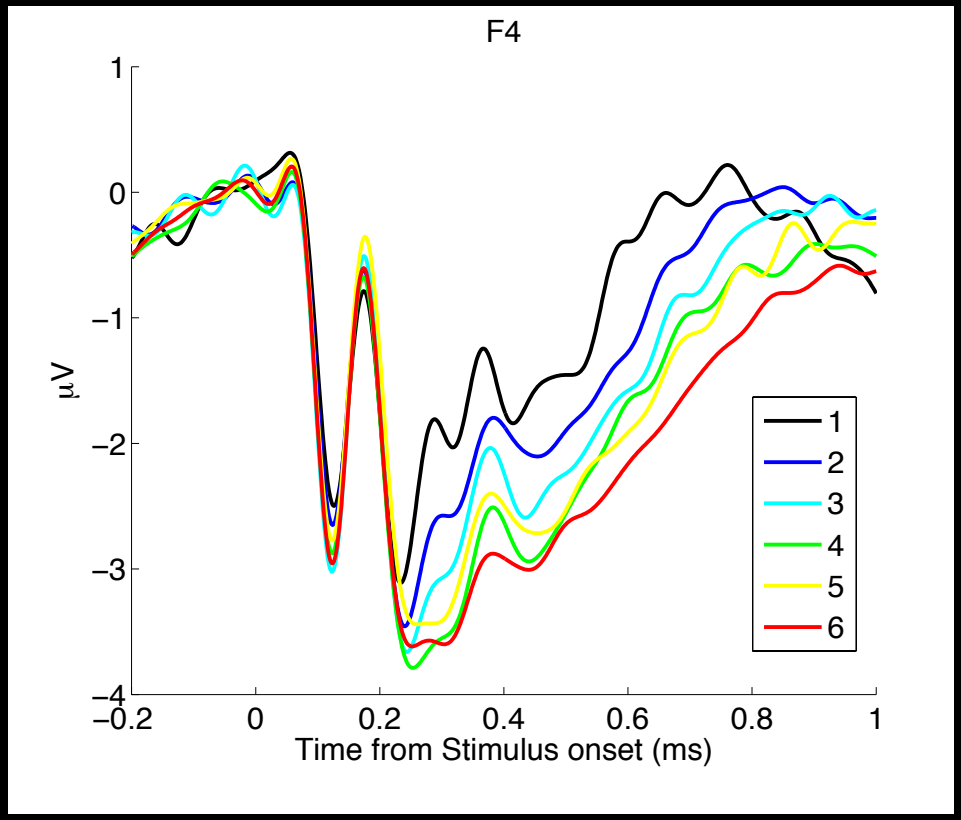
- Measures electrical activity on the scalp resulting from neural activity

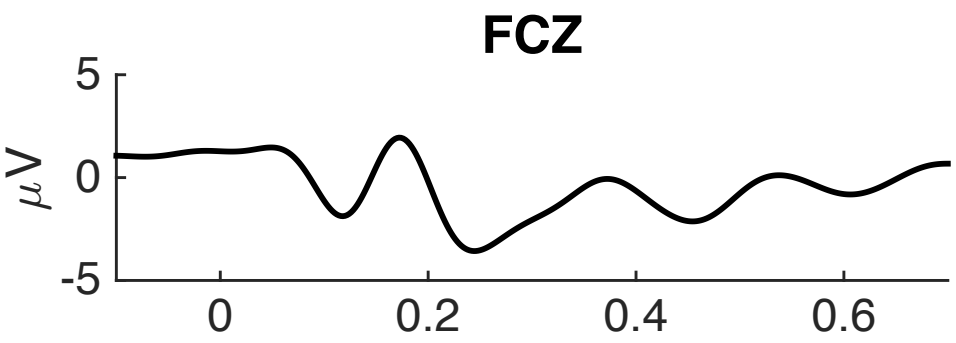
- Can we identify dissociable WM and RL components?
- Temporal interactions?

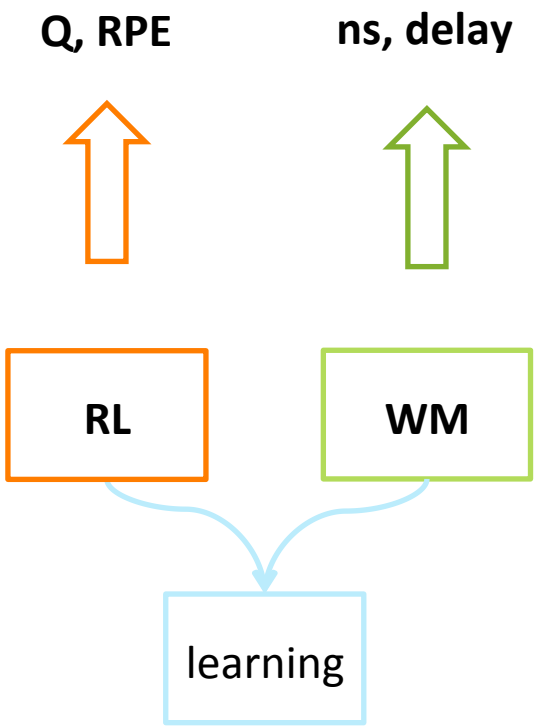
RLWM - EEG

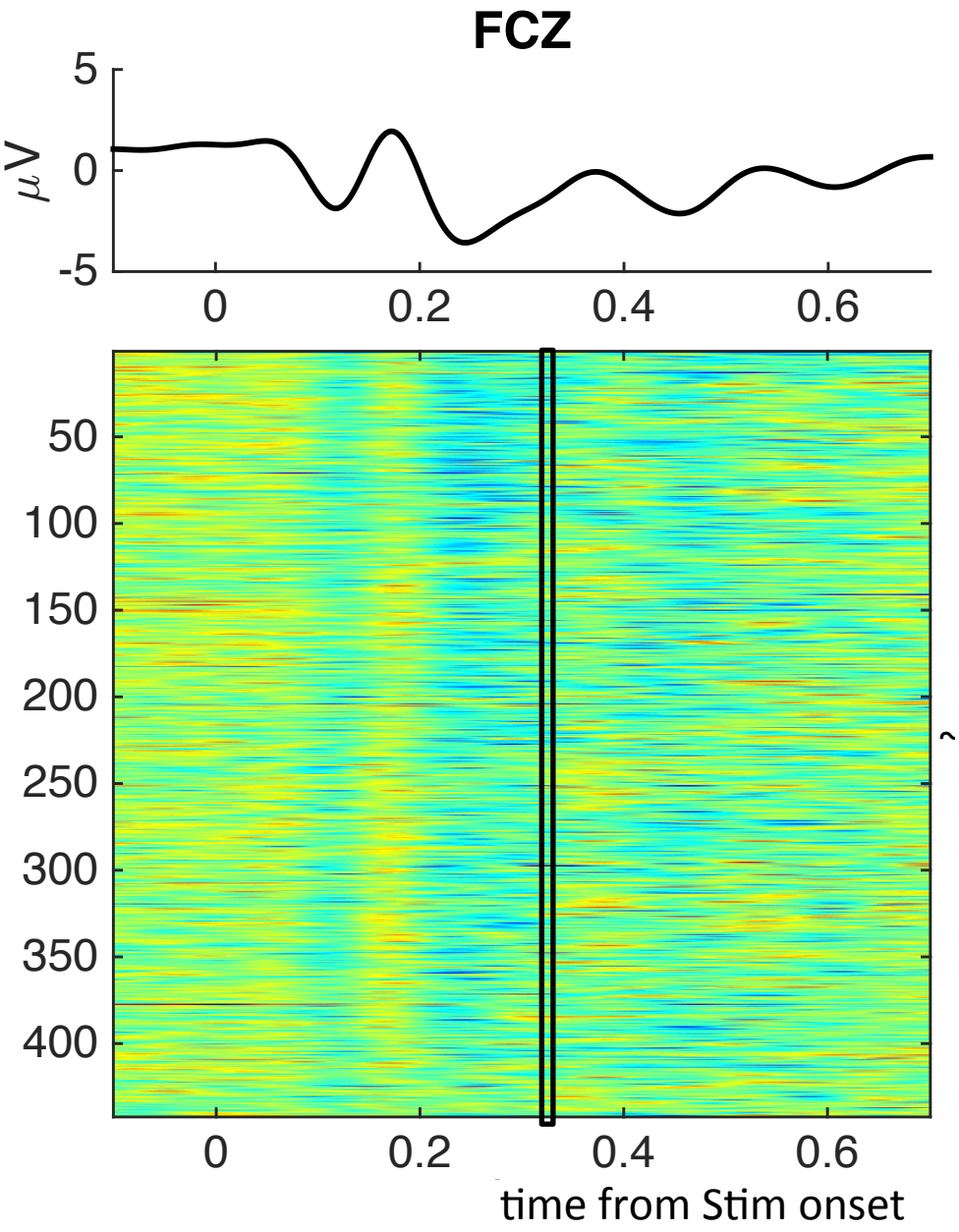
- Stimulus-locked event-related potentials (ERPs)
- Correct trials, as a function of set size.






FCZ
5
0
-5
μV
0
0.2
0.4
0.6

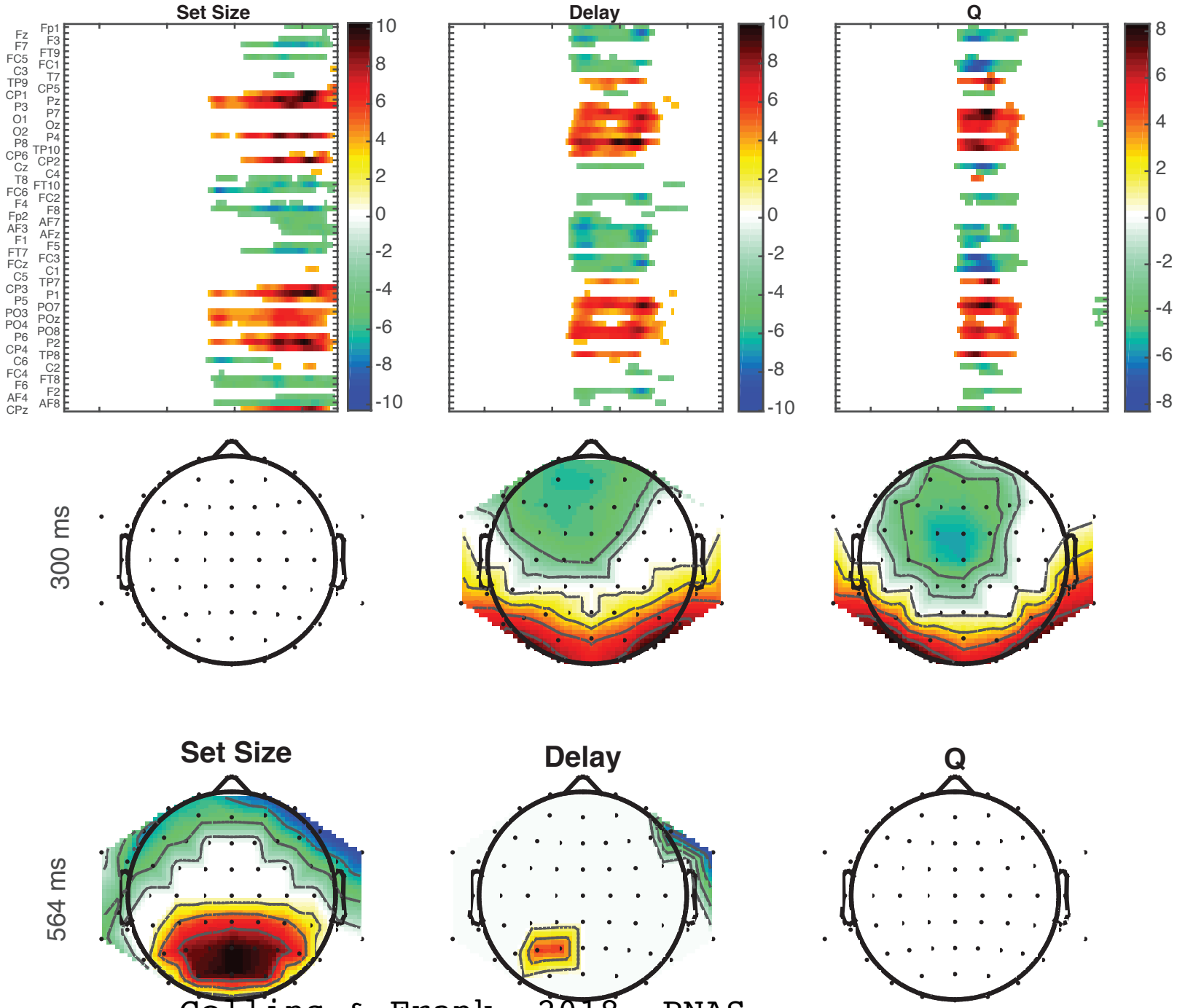

Q, RPE
ns, delay
RL
WM
learning



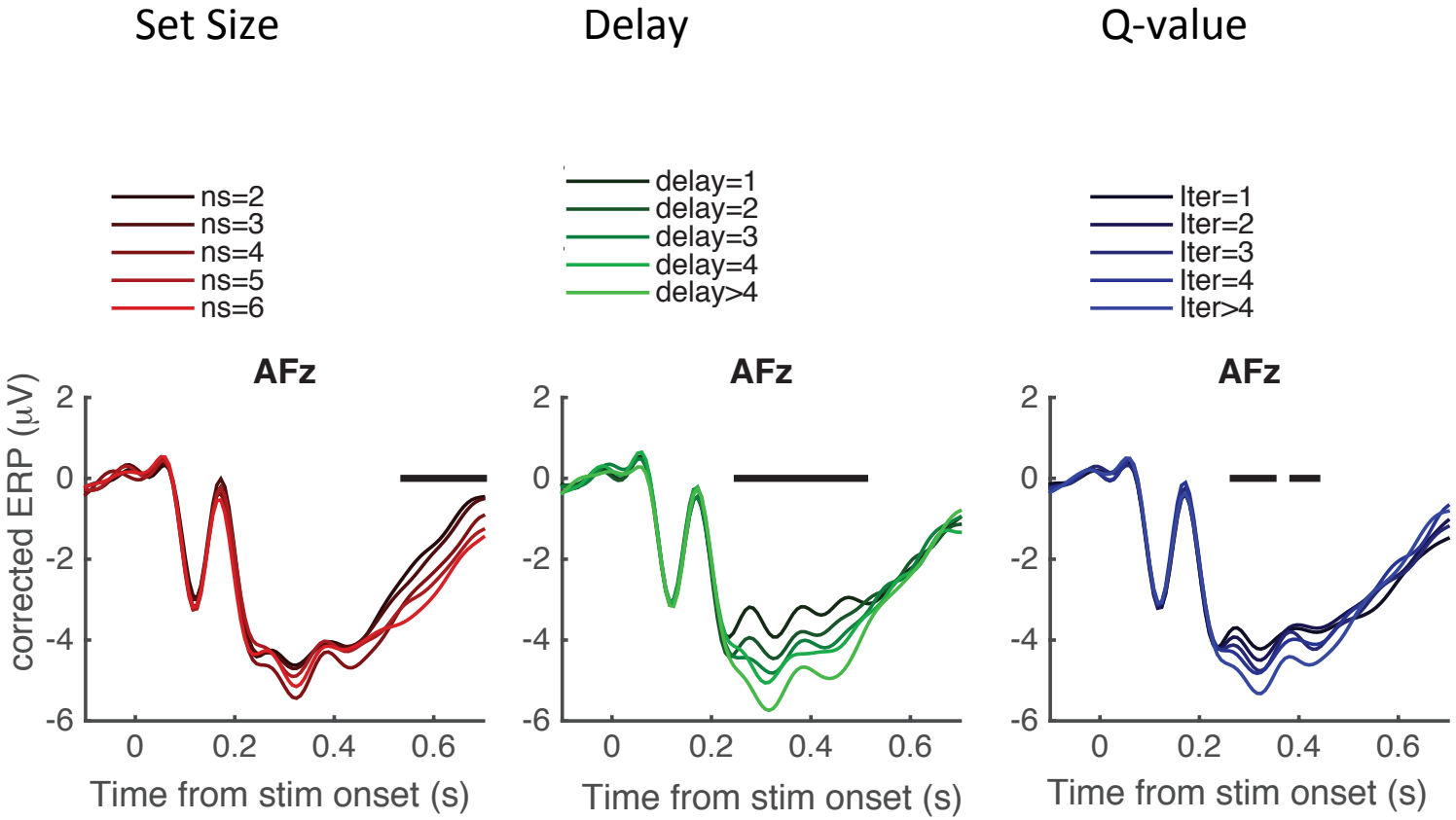
For each subject: β_{NS}(electrodes, time)

…

- Analyze weights across subjects.
- plot ERPs corrected for other factors

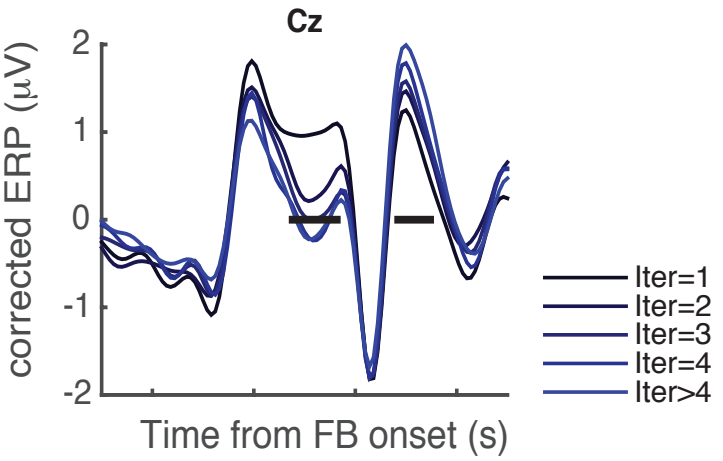


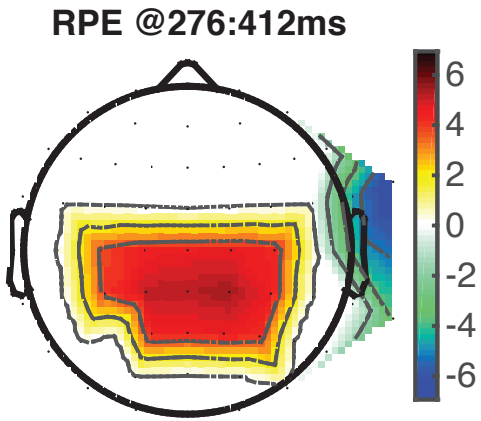
Collins & Frank, 2018, PNAS

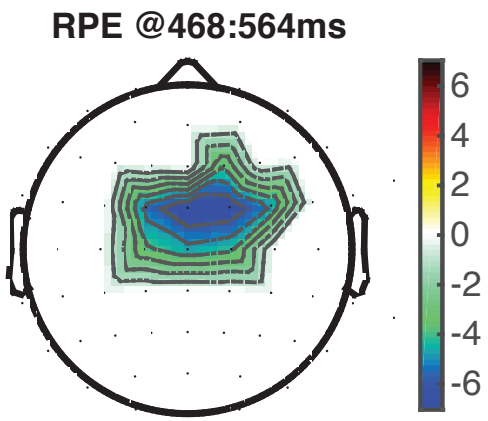


Collins & Frank, 2018, PNAS

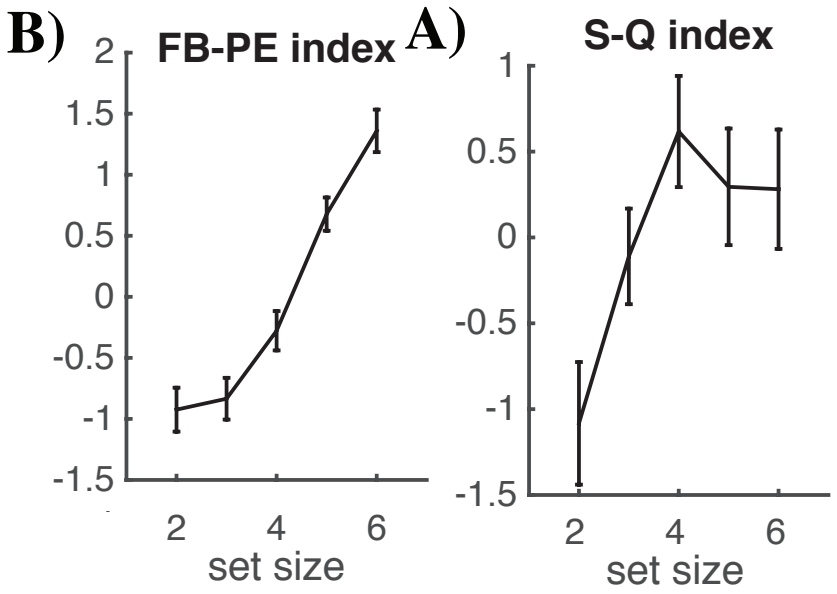
FB-locked: Reward prediction errors?



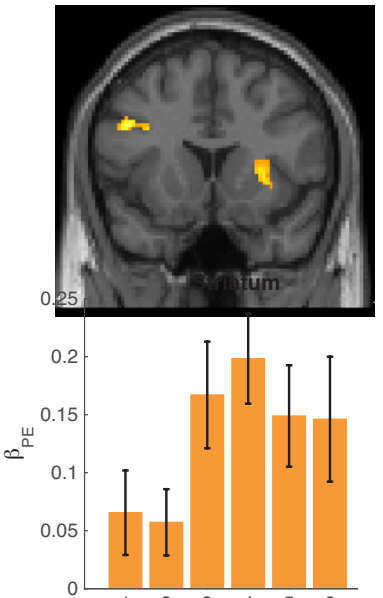




Stronger representation of RPE in higher set-size.



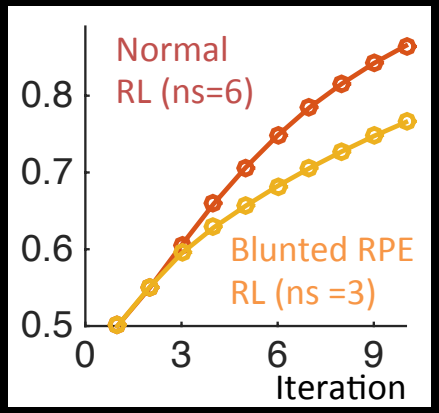
PE x nS

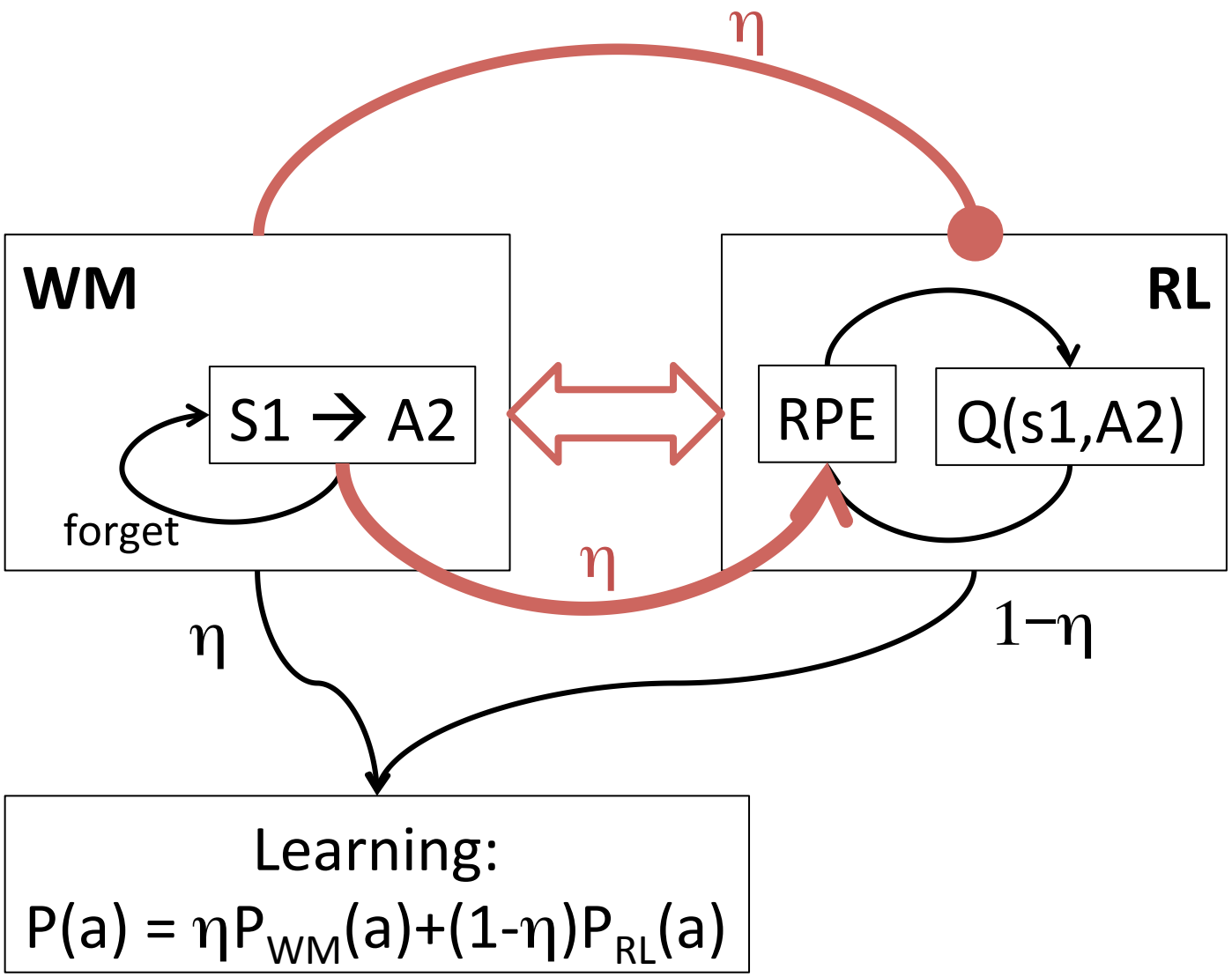


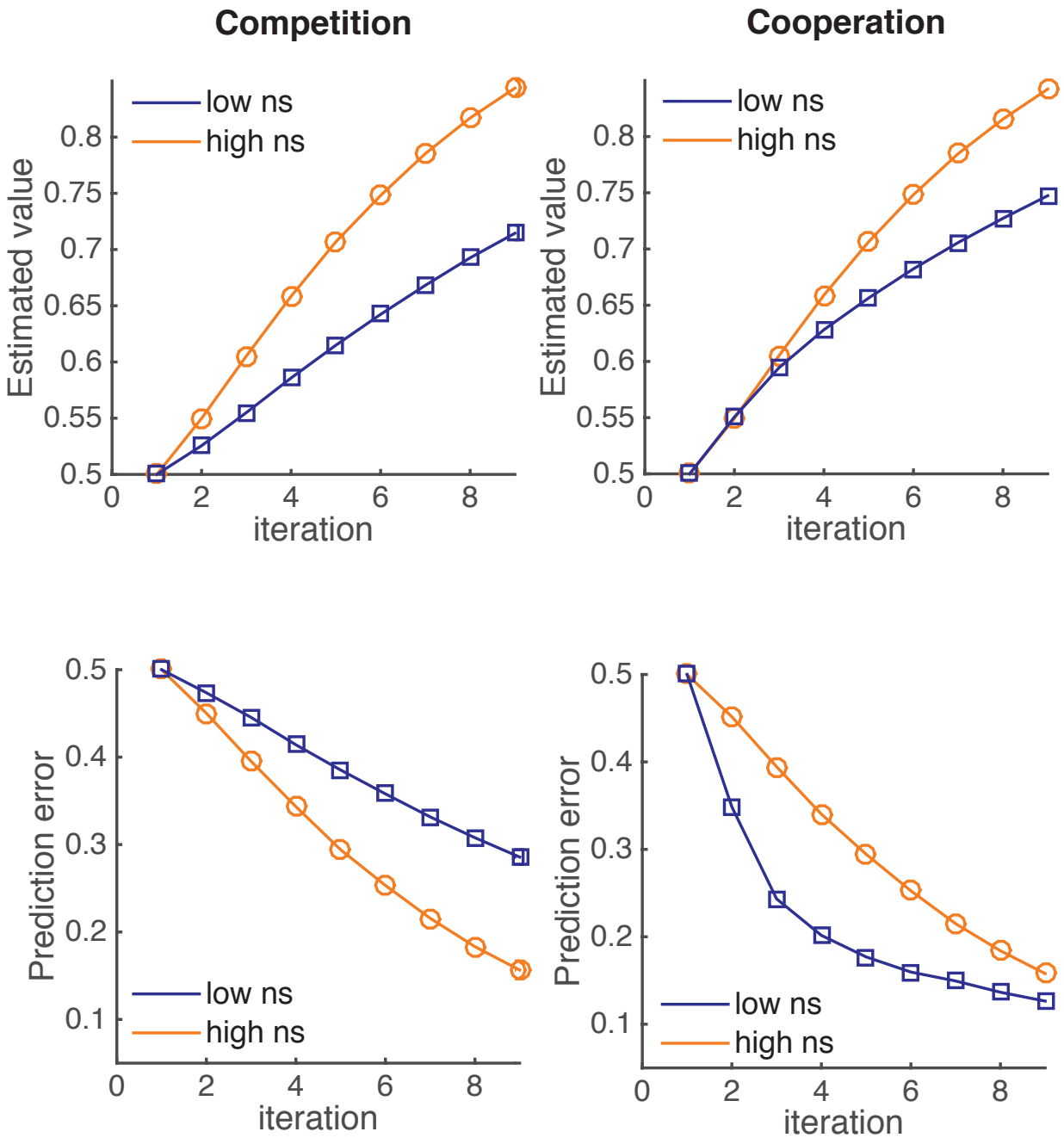
Collins & Frank, 2018, PNAS

Neural signals encoding reward prediction errors/reward expectations are sensitive to working memory function.

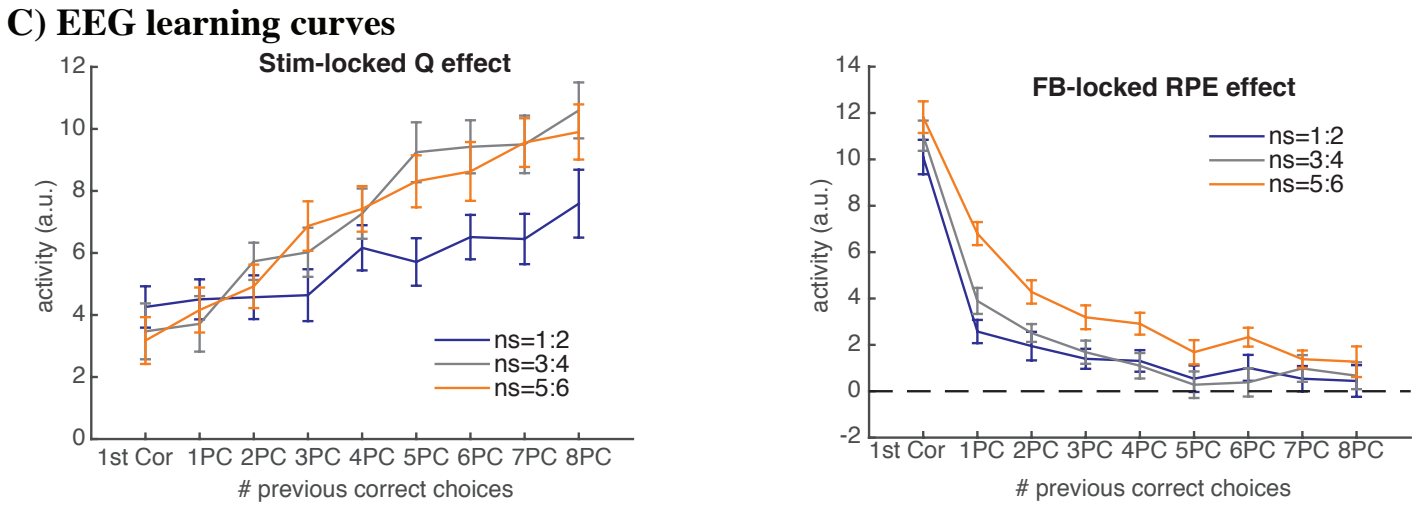
This indicates a possible interaction between RL and WM process.




η
WM
RL
S1 → A2
RPE
Q(s1,A2)
forget
η
η
1−η
Learning:
P(a) = ηP_WM(a)+(1-η)P_RL(a)


Competition
Cooperation
low ns
high ns
Estimated value
iteration
Prediction error
0.5
0.55
0.6
0.65
0.7
0.75
0.8
0.1
0.2
0.3
0.4
0
2
4
6
8

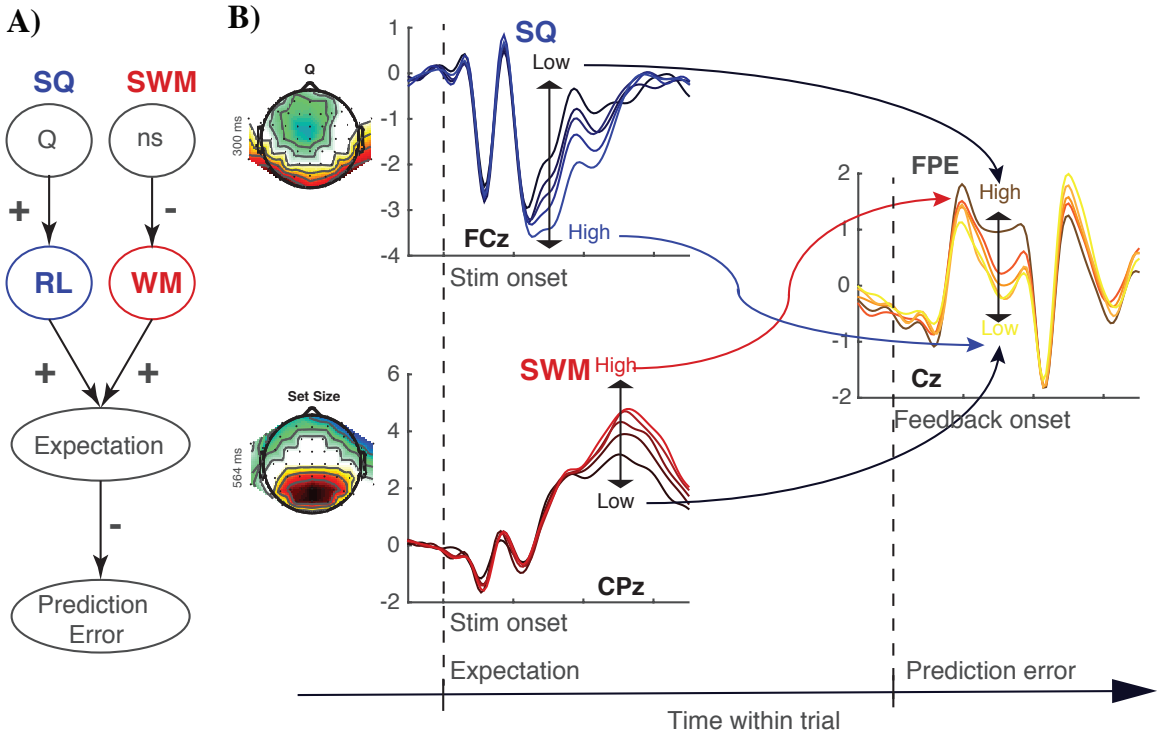
EEG learning curves



EEG RPE signal drops faster for low set-sizes

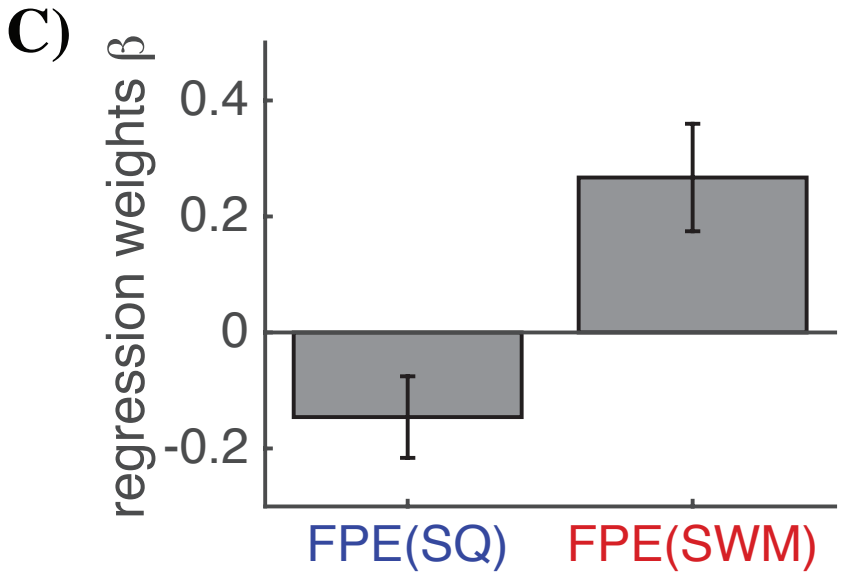
→ WM contribution?

Collins & Frank, 2018, PNAS



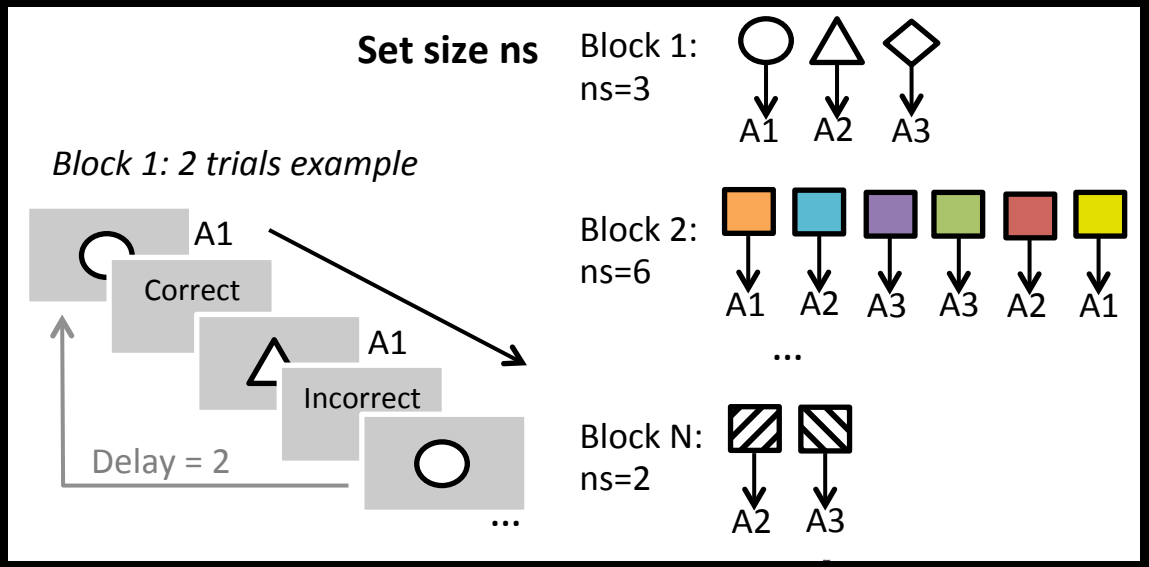
EEG WM signal influences EEG RPE signal.

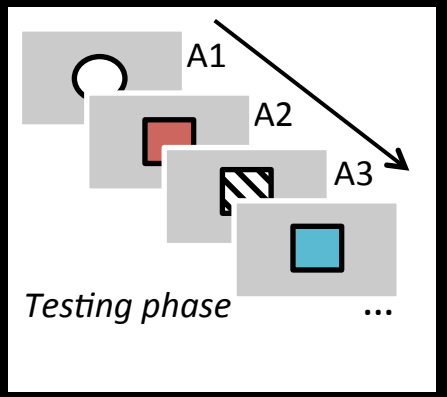
→WM contribution

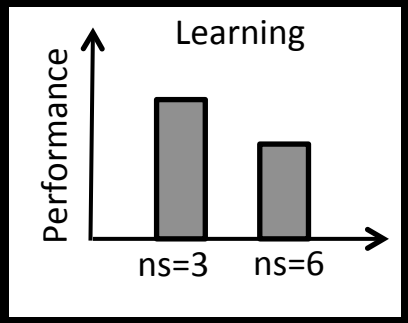


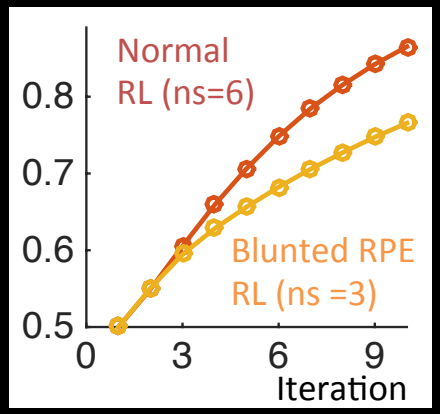
Collins & Frank, 2018, PNAS

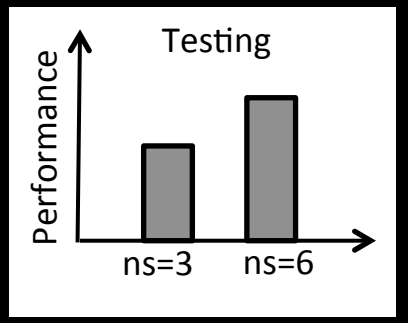
Prediction: Long term associations are learned better in high than low set sizes

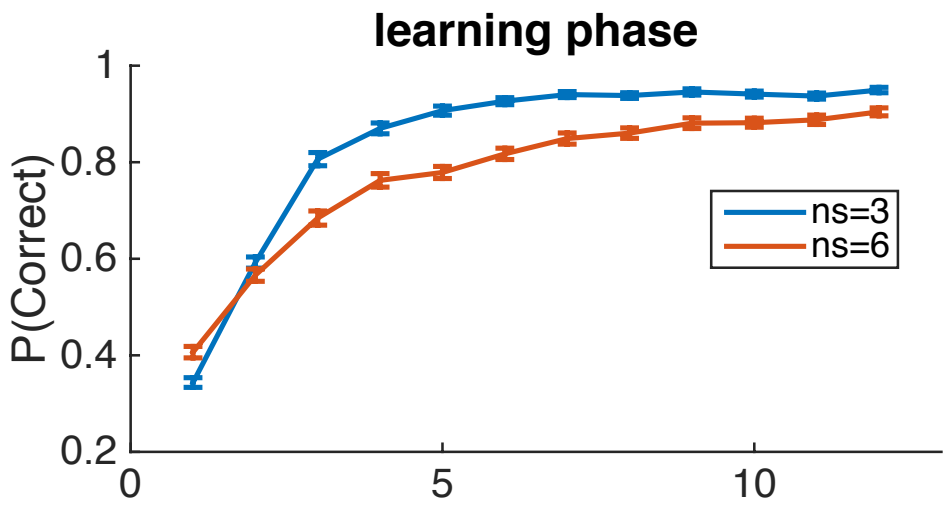


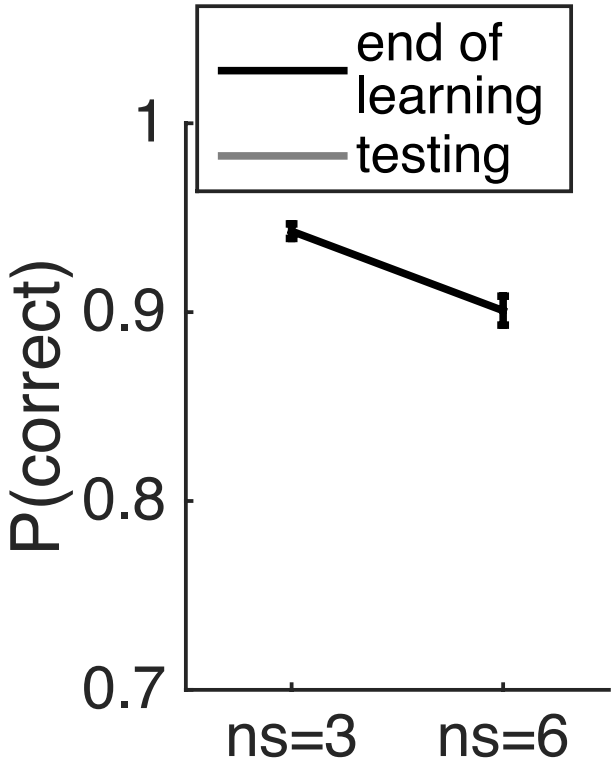


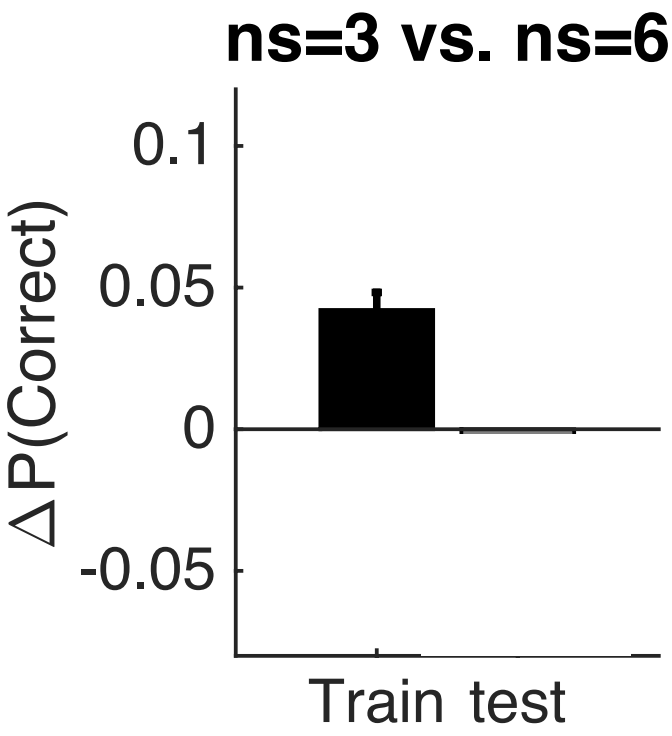












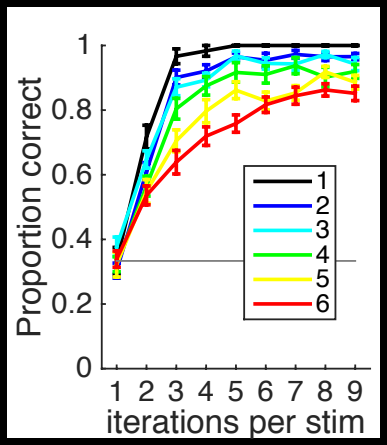
Collins, 2018, JoCN

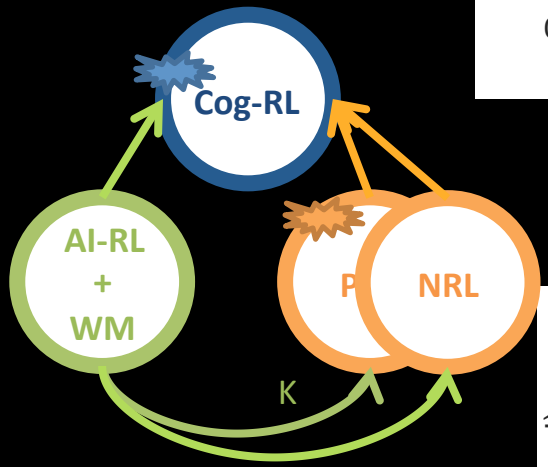
Learning is a mixture of multiple neuro-cognitive processes that implement different computational trade-offs and rely on different neural mechanisms

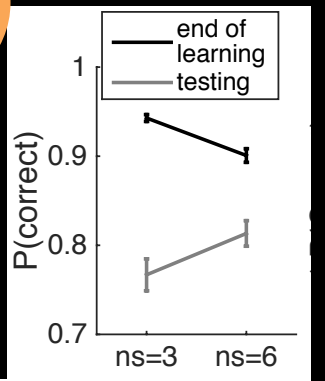
We can isolate SZ learning dysfunction as WM, not RL

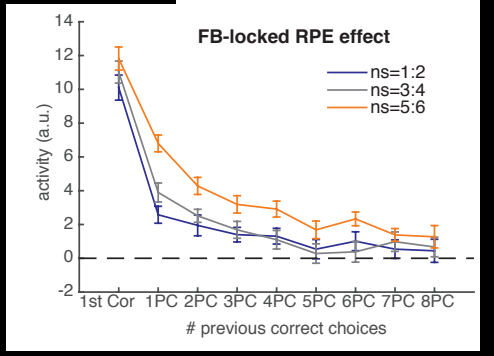
WM interferes in RL computations of value: what is learned faster is learned worse

Interaction is cooperative









Acknowledgements

- Brown
  - Michael Frank (Post-doc advisor)
  - David Badre (fMRI)
  - Frank and Badre labs
- Maryland (SZ studies)
  - Jim Waltz
  - Jim Gold
  - Matthew Albrecht
- Data Collection
  - Chris Gagne
  - Brittany Ciullo
  - Julie Helmers
  - Nora Harhen
  - Sarah Master
  - Frank lab RAs
- Funding
  - NSF
- UC Berkeley CCN lab and collaborators:

Thank you!

Contact me with any questions, comments!

annecollins@berkeley.edu