Development of dopaminergic neurophysiology supports improvements in the use of optimal reward learning strategies through adolescence

1/27/2023

Take homes

- Reward learning shows developmental improvements through adolescence, driven by a change in the ability to identify an optimal learning rate, and apply it during goal oriented decision making

1. Task & RL Modeling

- Striatal dopamine contributes to the development of not only momentary identification of rewards, but generation of task heuristics and learning strategies

2. PET/Tissue Iron

- DA physiology modulates both cortical and sub-cortical functional circuitry, in particular in the vmPFC, to support changes in learning heuristics

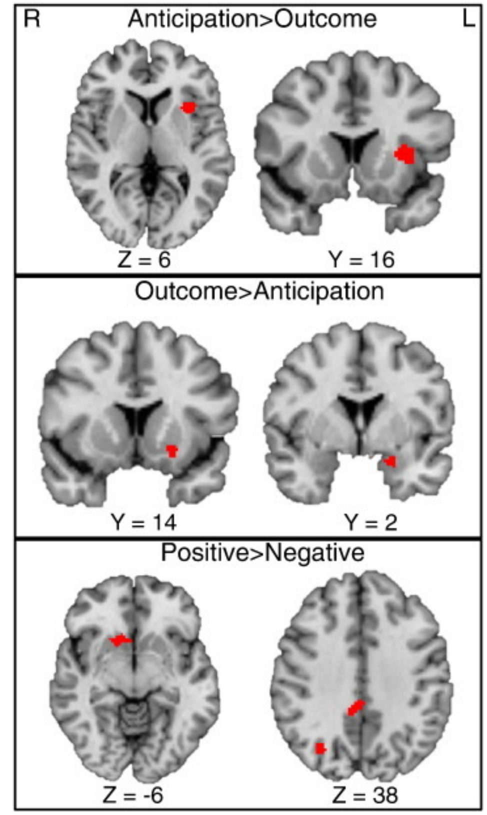
3. fMRI

Adolescence is a time of heightened reward sensitivity

- Rewards have a disproportionate effect on cognitive performance and decision making in adolescence compared to adulthood (e.g., Geier et al, 2012)

- These differences are supported by heightened activity of the ventral striatum (VS) (Silverman et al, 2015; Padmanabhan et al, 2011)

- While heightened reward reactivity may have maladaptive consequences (risky behaviors, experimental substance use and abuse, etc), it is also critical for learning about the structure of action-outcome associations, developing social interactions, and more



Silverman et al 2015 meta-analysis

Introduction

Reward learning matures through adolescence

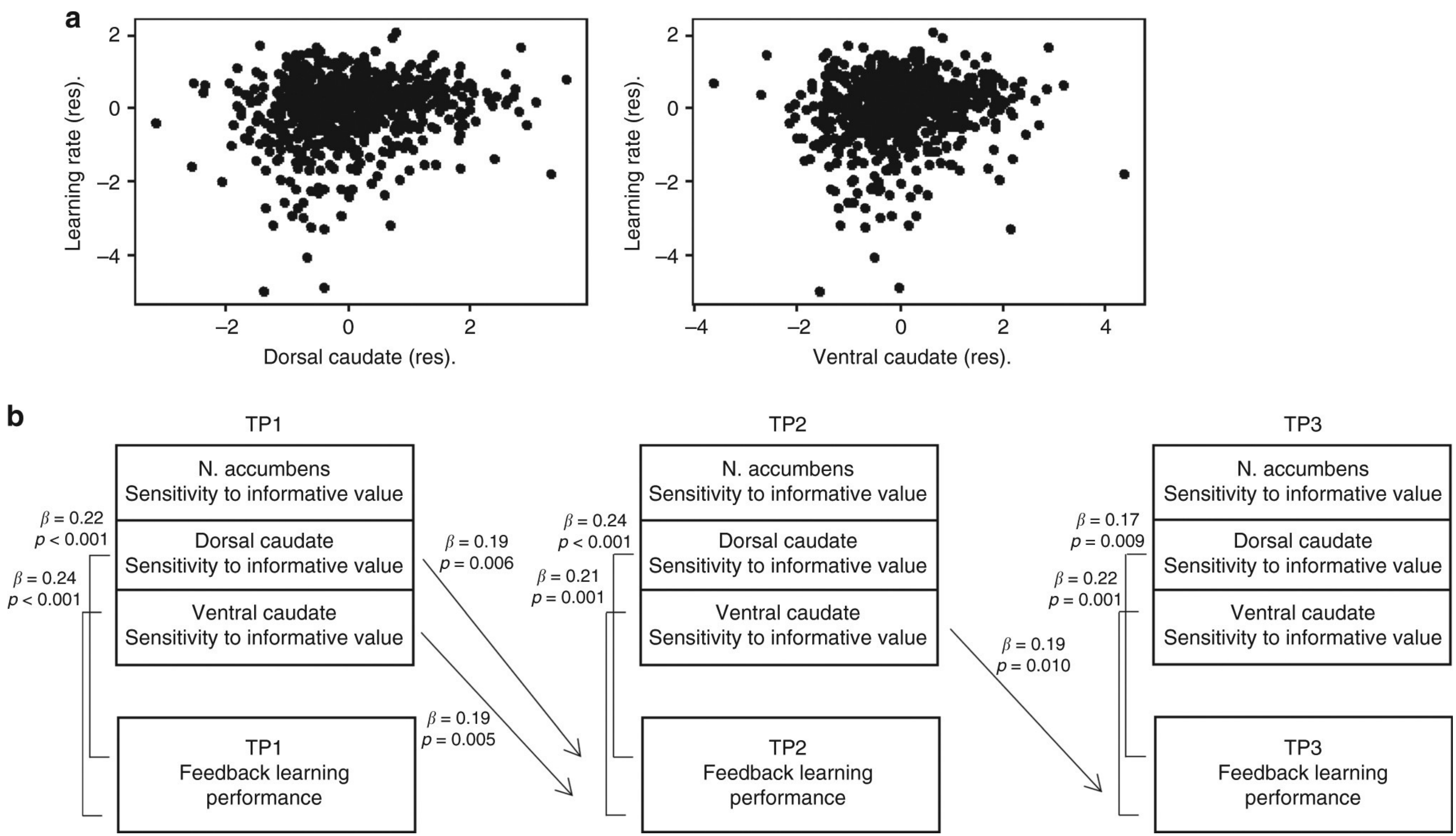
- Development of learning rates is highly task-dependent, potentially reflecting a shift towards more task-optimal learning (Decker et al 2015, Master et al 2019)

- Age-related decreases in RL *temperature* (i.e., undirected exploration) have been somewhat more consistently reported (Christakou et al., 2013; Decker et al., 2015; Javadi et al., 2014; Palminteri et al., 2016; Rodriguez Buritica et al., 2019)

- Changes in reward learning may reflect shifts in learning strategies, not just quantitative changes in RL parameters:
  - Increasing use of model-based learning strategies (Raab & Hartley, 2019)
  - Increased "metacontrol", i.e., dynamic adaptation to task demands (Bolenz & Eppinger, preprint)
  - Increased valence-independent learning (Hauser et al., 2015; Rodriguez Buritica et al., 2019; van den Bos et al., 2012)

Contribution of striatal activation to reward learning

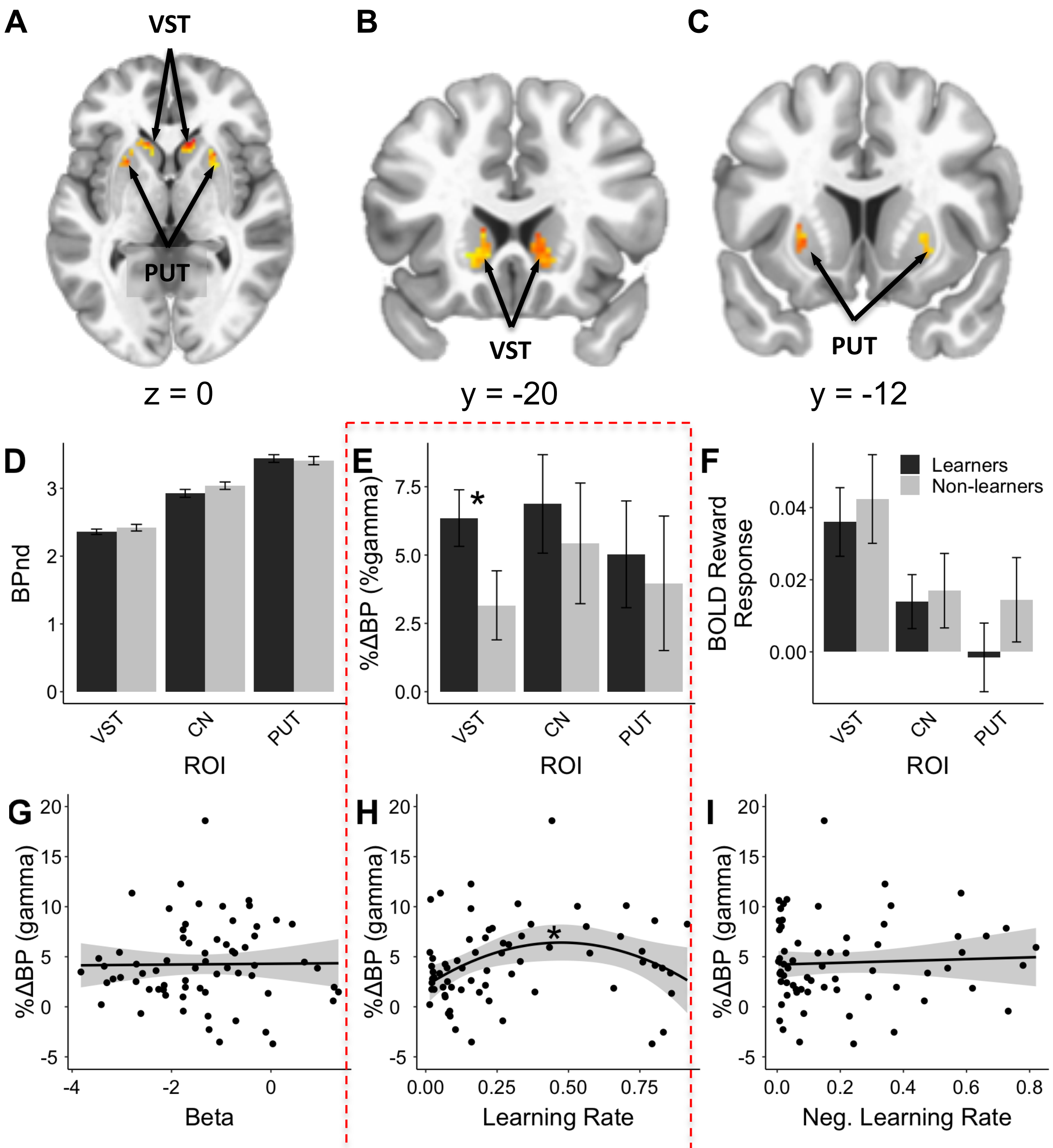
- Striatal (Peters & Crone, 2017) and hippocampal (Davidow et al, 2016) reward-related activity supports reward learning

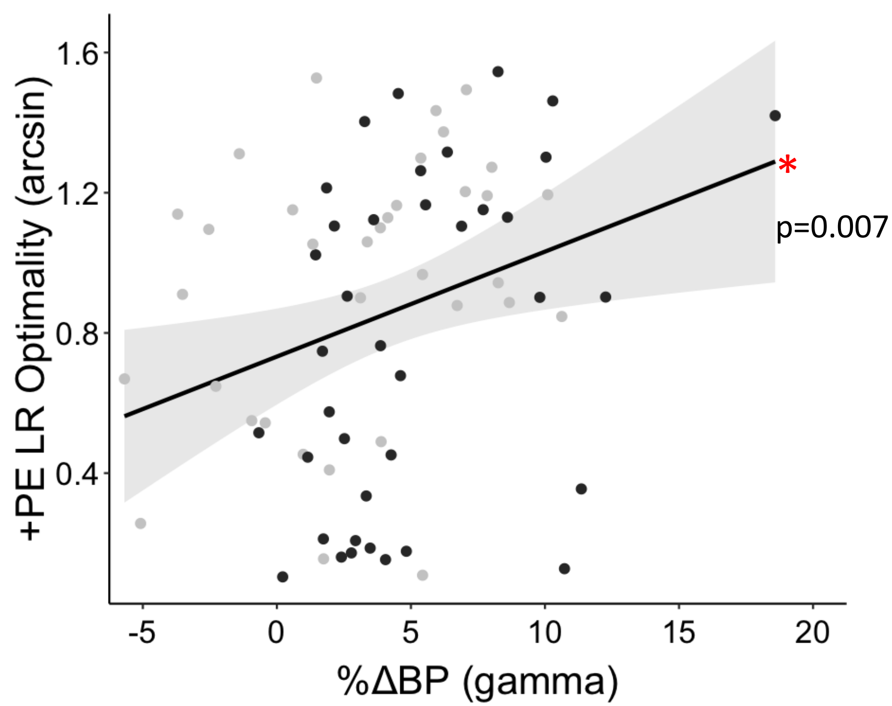


Peters & Crone, 2017

Contribution of striatal dopamine to reward learning

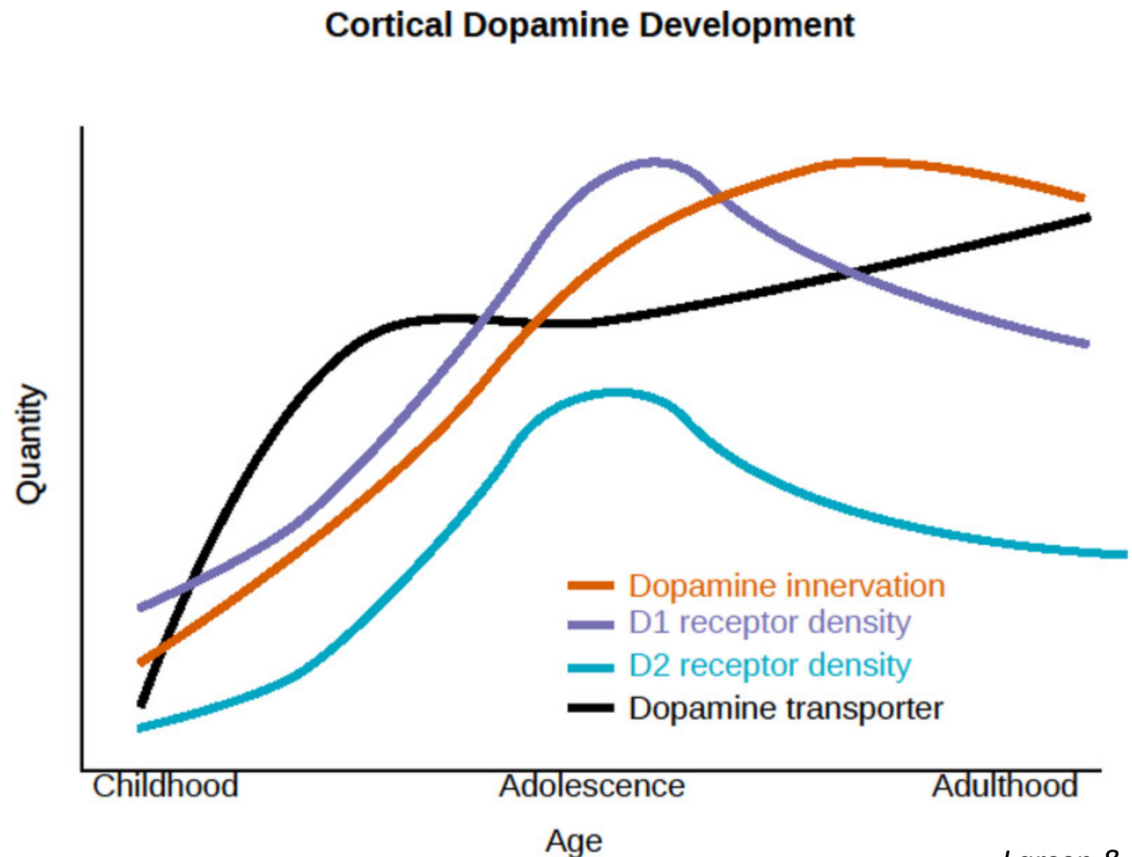
- Task-related changes in [^{11}C]Raclopride binding are associated with learning in adults





Calabro et al, 2022

- Heterogeneous pattern of DA development through adolescence

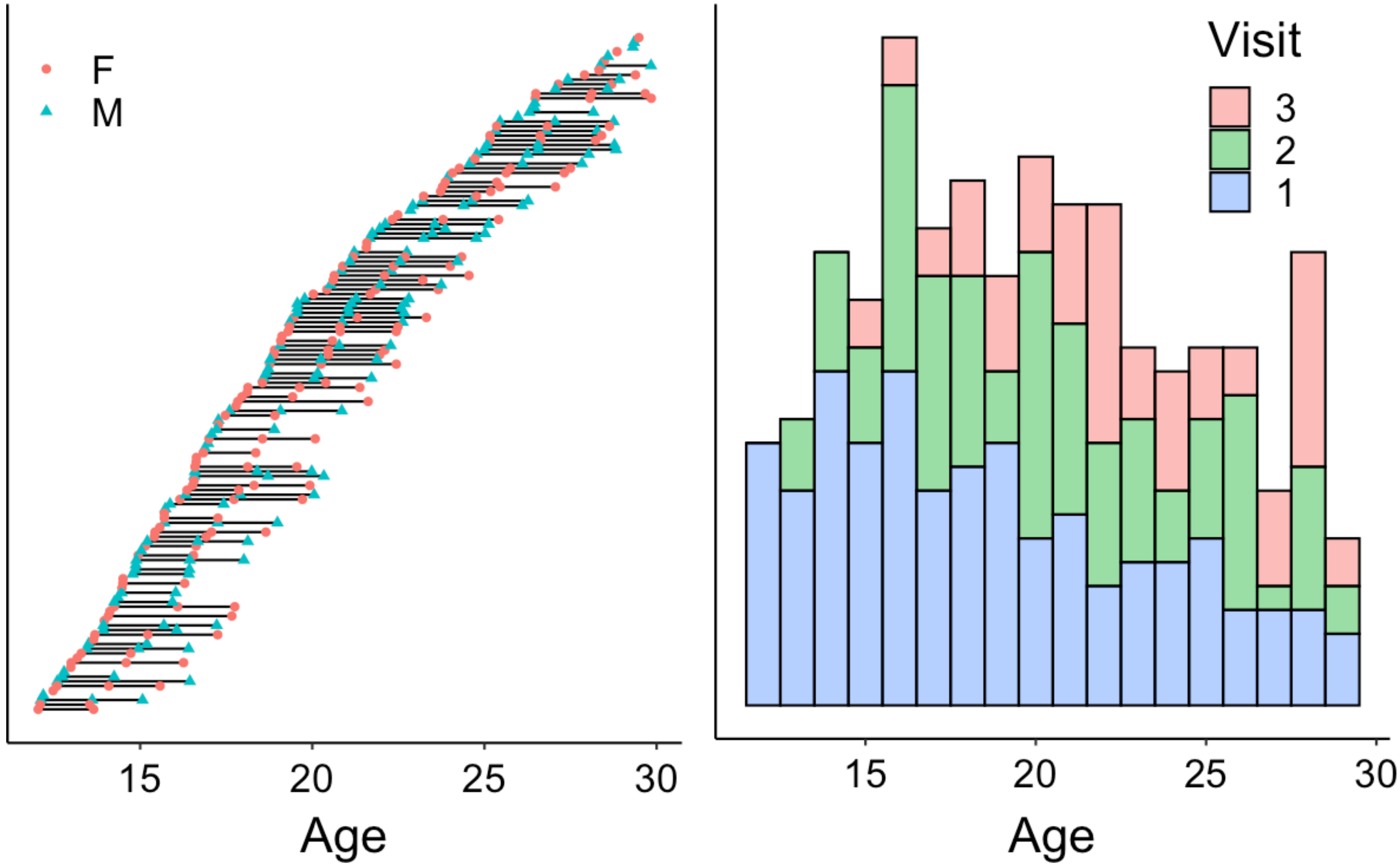


Larsen & Luna, 2018

How do developmental changes in dopamine neurophysiology drive functional changes supporting the development of reward learning behaviors?

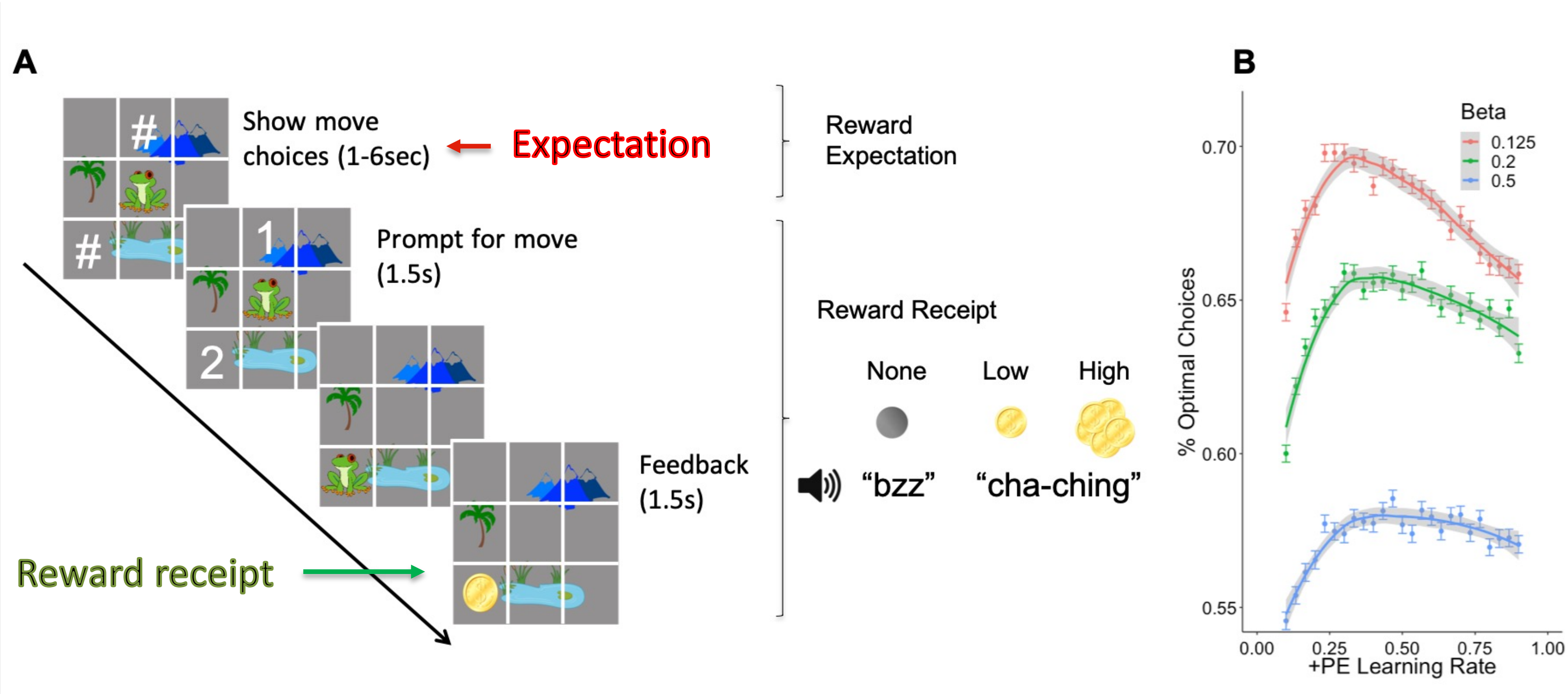
Subjects

- Full sample
  - 145 subjects (77 AFAB, 306 total visits, 1-3 visits per participant)
  - Ages 12.0-29.8 (mean 20.5±4.7)



Methods

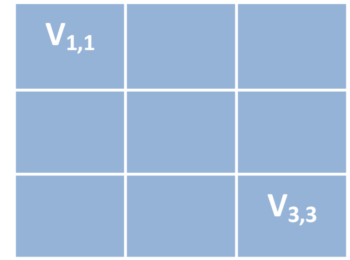
- Task



Calabro et al, 2022

RL model

- What is the per-trial reward expectation/prediction error?
  - Reinforcement learning (RL) model to predict trial-wise responses
  - Maintain an internal state value ($V_{i,j}$) of expected value for each location

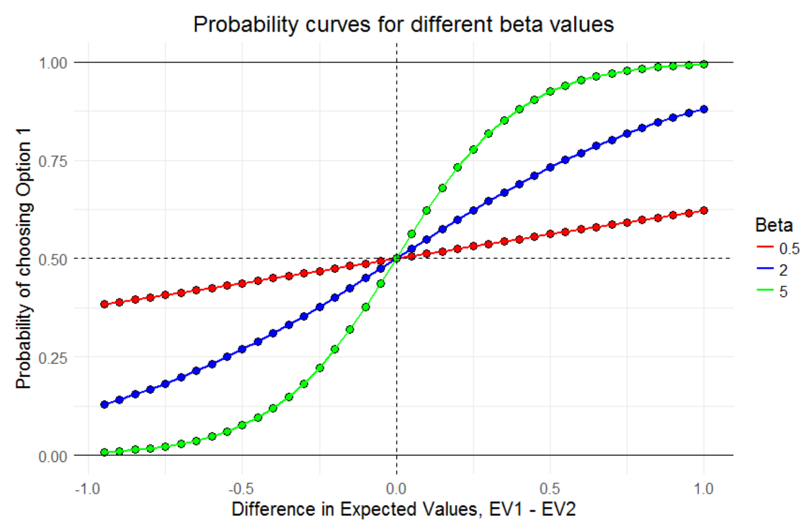


  - After each movement choice S, update internal expected value (V) based on learning rate (v) and prediction error

$$V_{i,j}(t+1) = V_{i,j}(t) + \upsilon \cdot (R - V_{i,j}(t))$$
$$(i,j) \in S$$

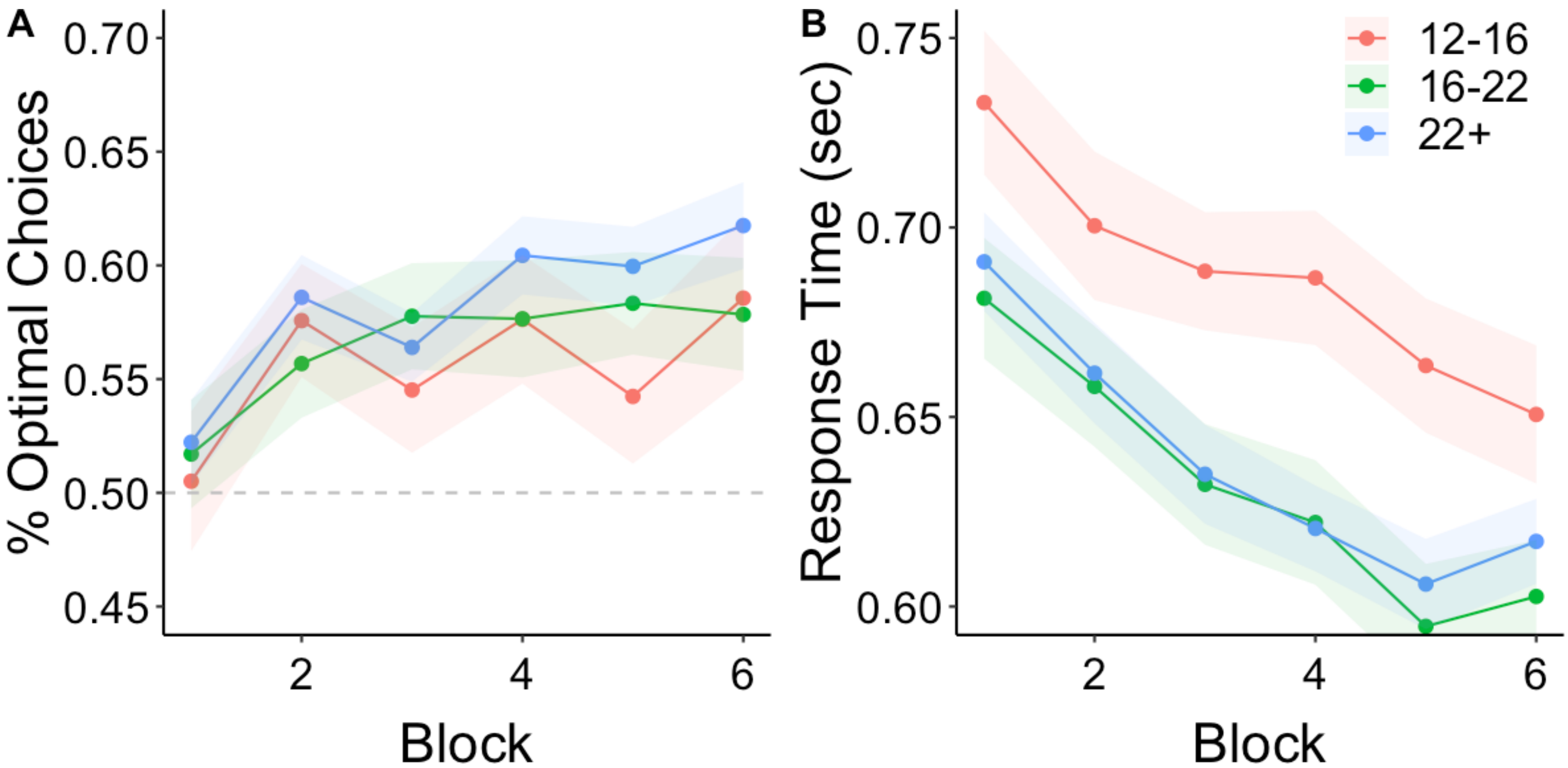
  - Select next move based on a softmax function of the two expected values

$$P_1 = \frac{1}{1 + e^{-\beta(V1 - V2)}}$$



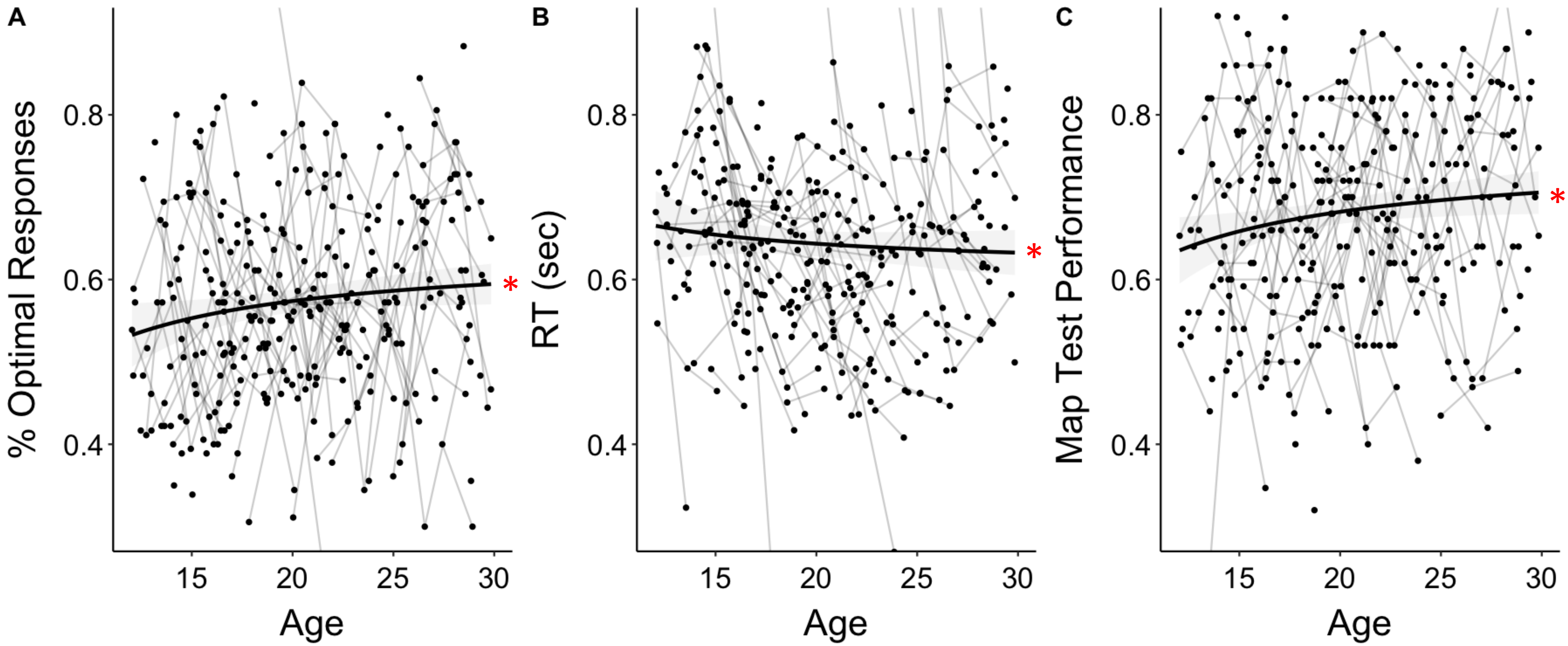
Results

- Developmental improvements in reward learning performance

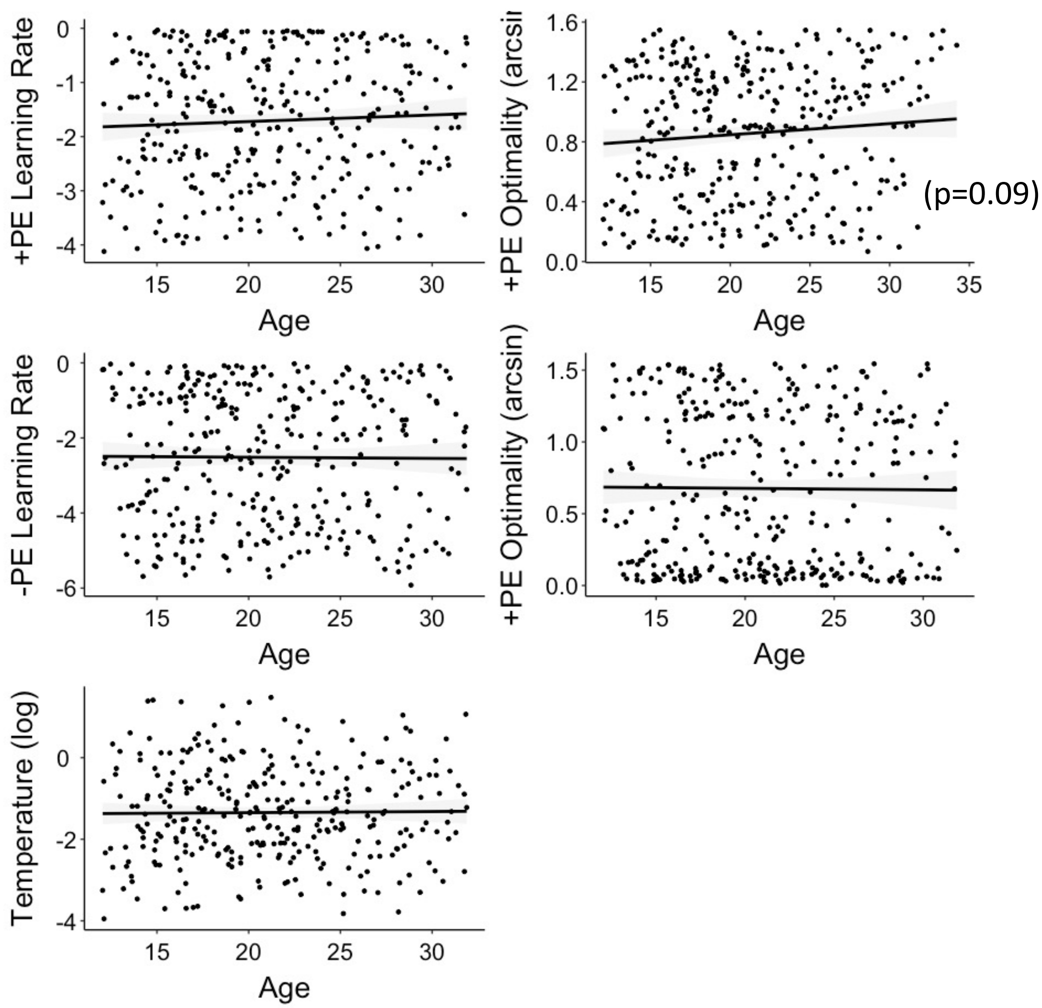


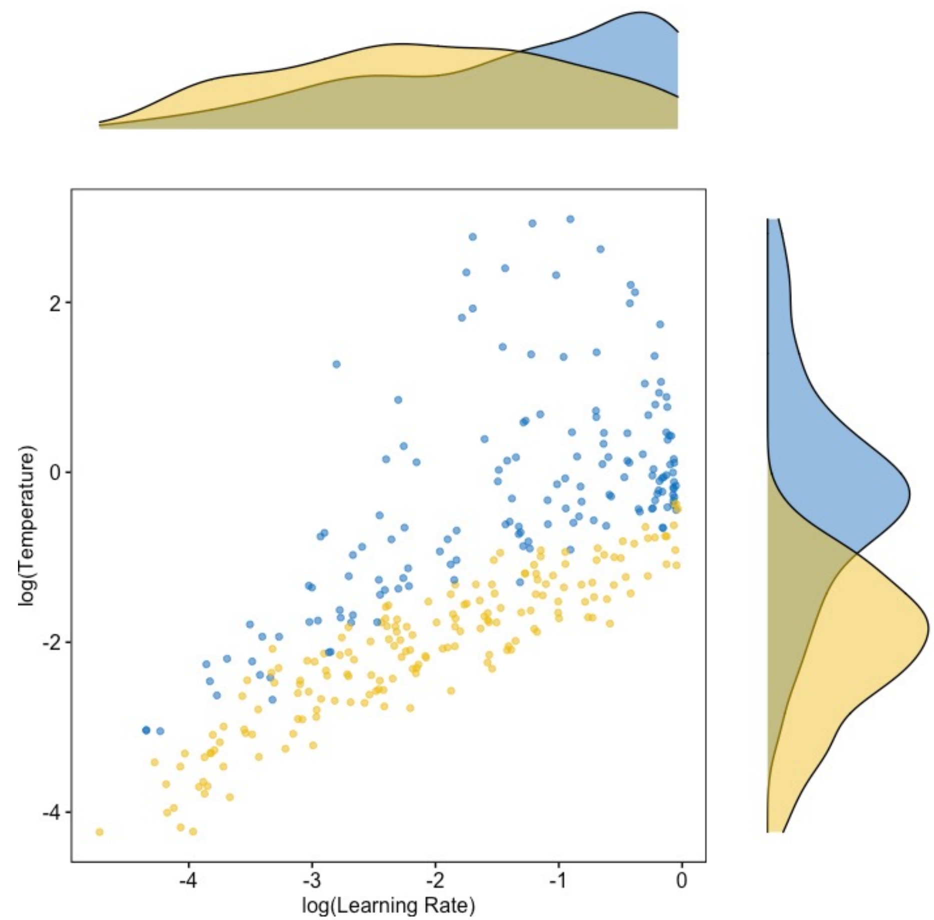
Results

- Developmental improvements in reward learning performance



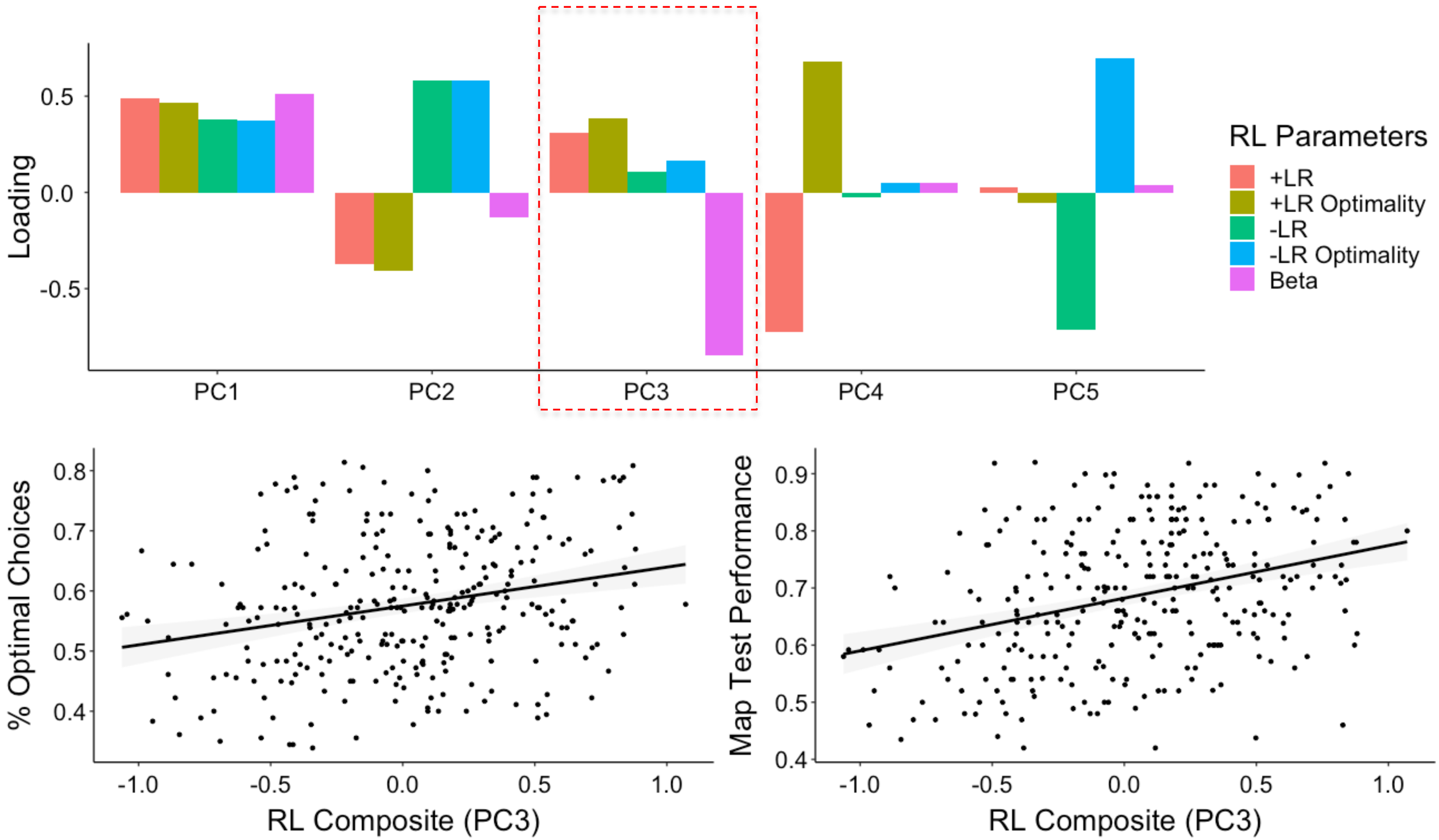
- RL model parameters do not show univariate changes with age



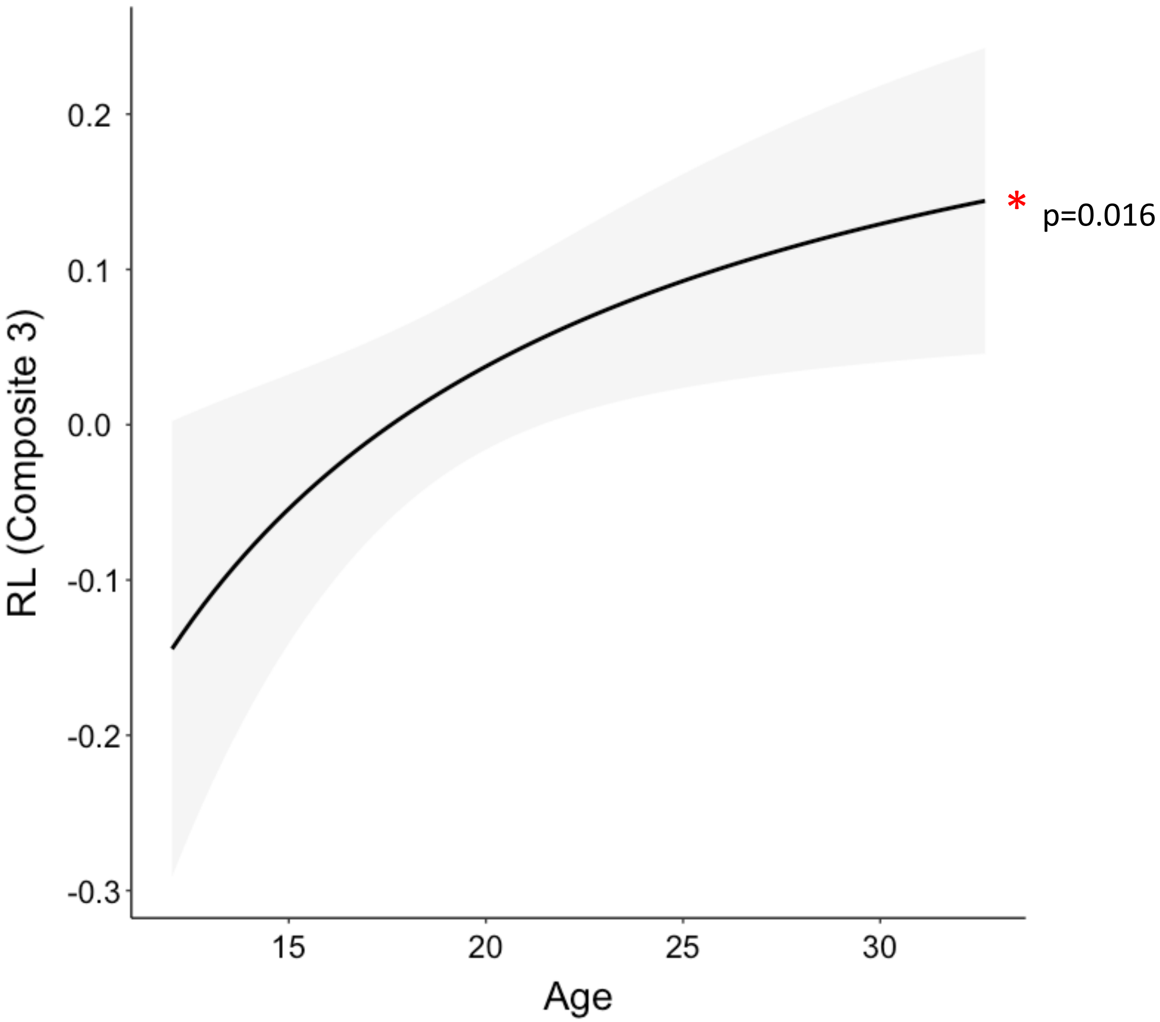


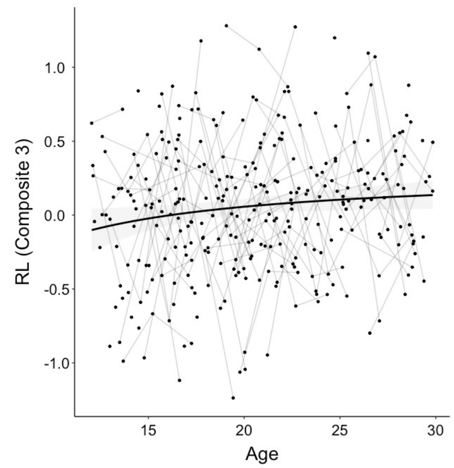
- But, parameters are highly correlated, and what makes someone a good at the task appears to be multivariate in nature

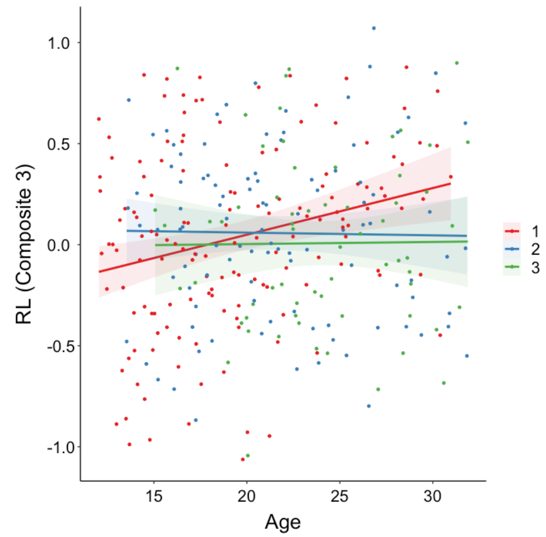
- RL composite parameters predict learning



- Composite parameter PC3 increases significantly with age

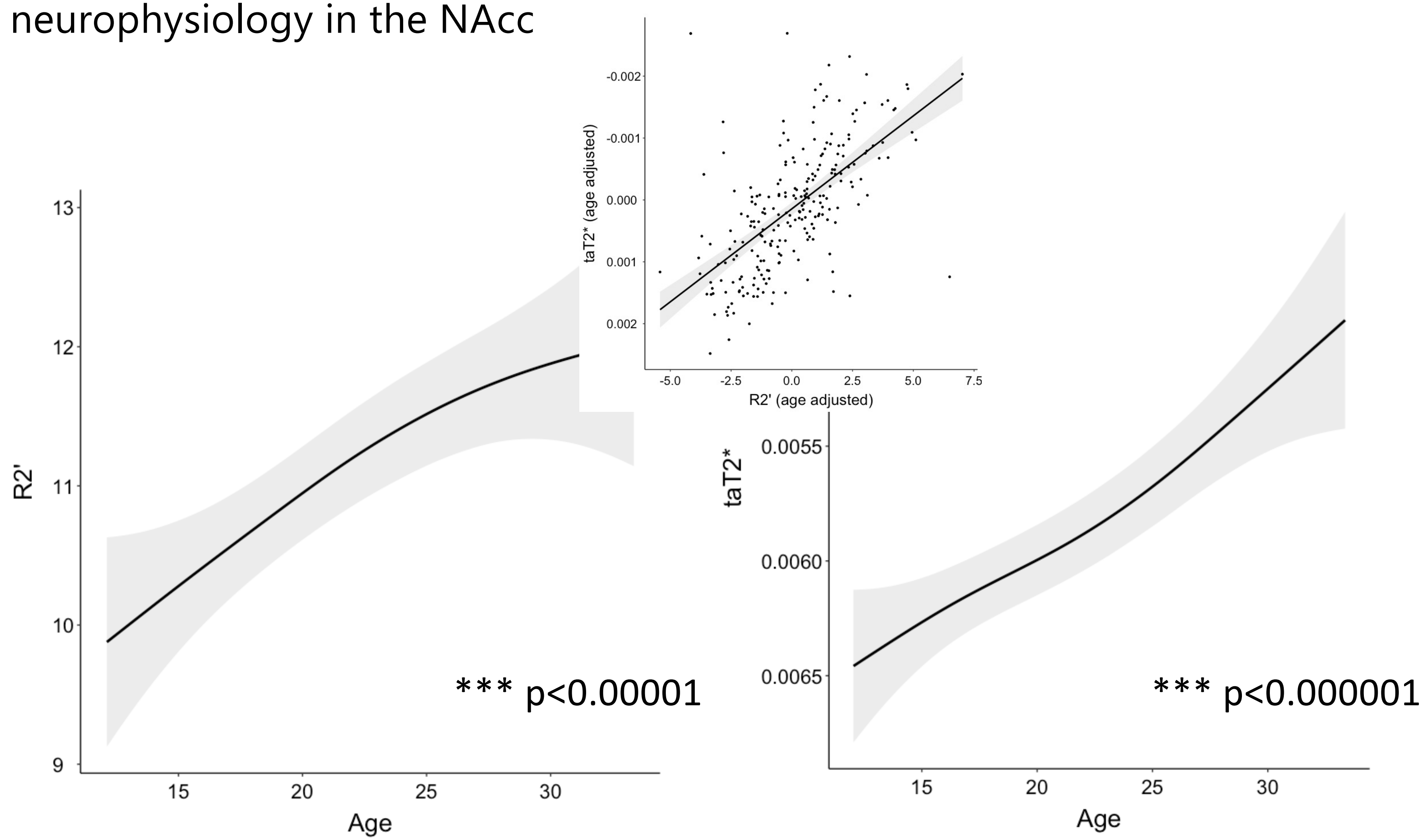






- Effect is most dramatic at time 1 (novelty/task familiarity effect?), but main effect of age persists when controlling for visit

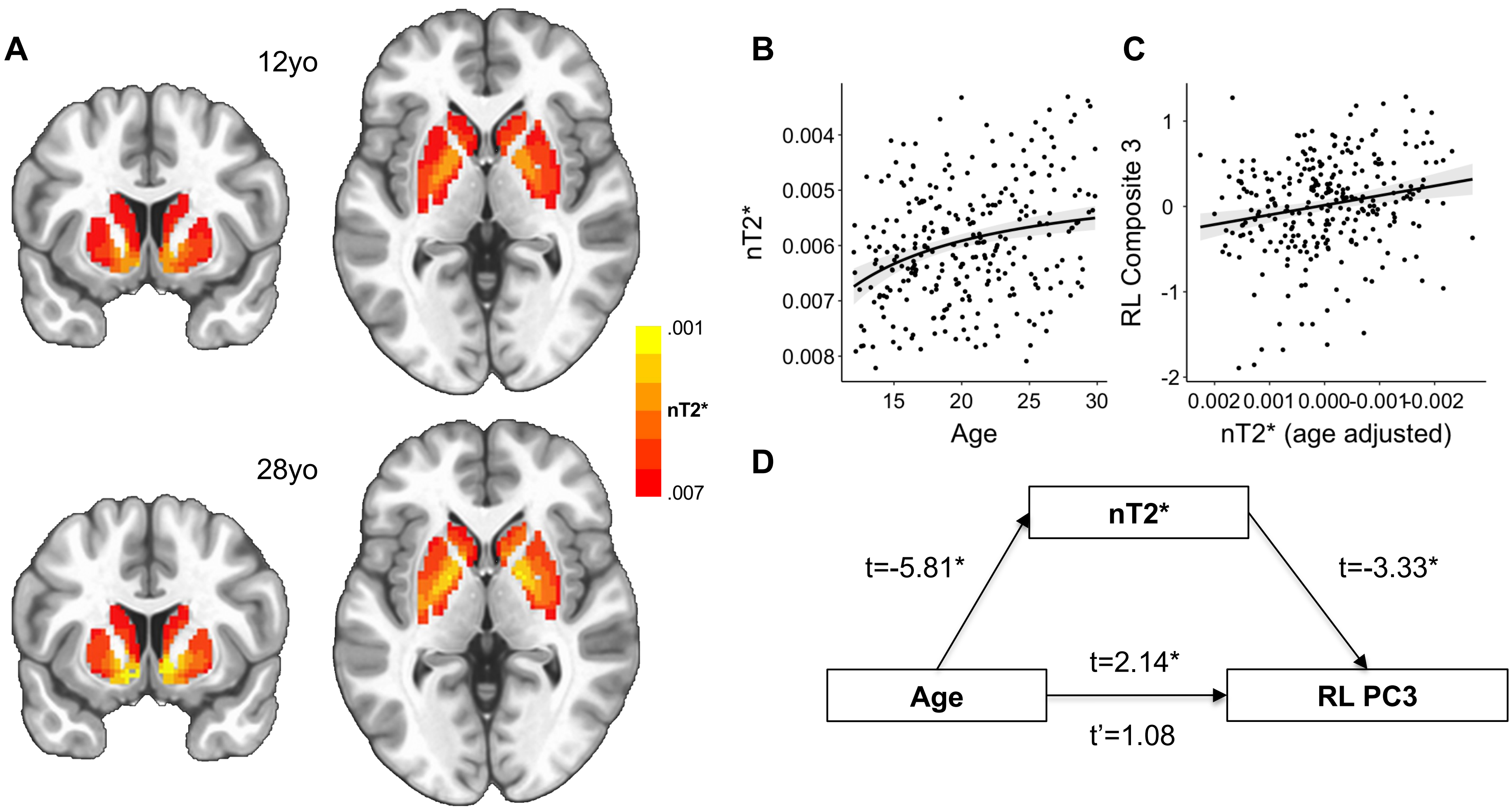
- Tissue iron as a developmentally sensitive indirect marker or striatal dopaminergic neurophysiology in the NAcc



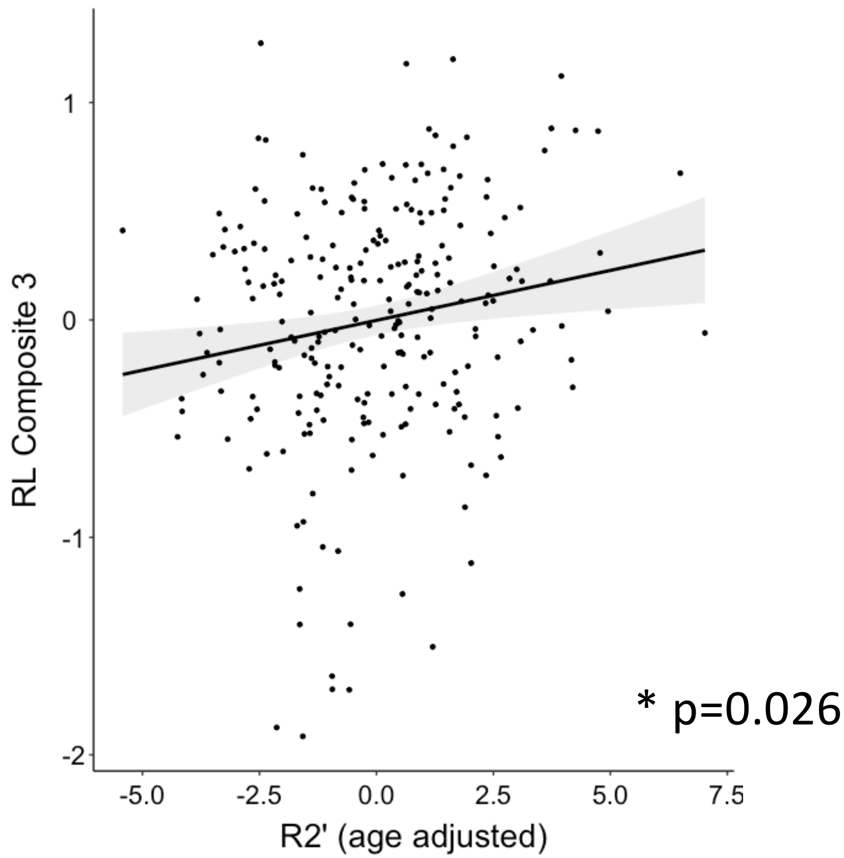
n=130 subjects,
n=236 scans

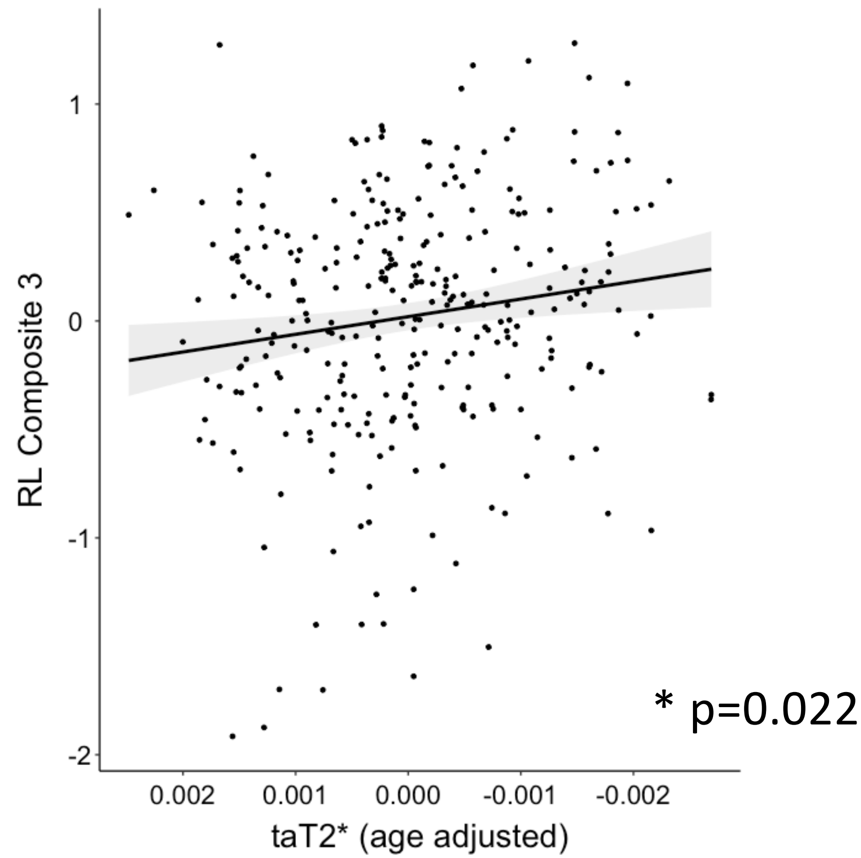
n=149 subjects,
n=308 scans

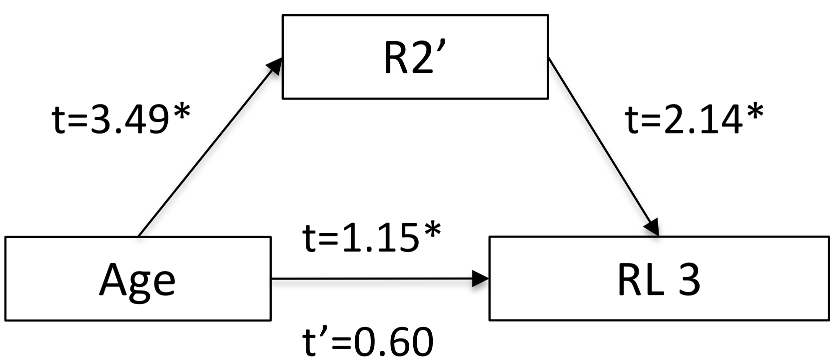
- Developmental increases in tissue iron mediate increases in the use of optimal RL learning strategies (via PC3)

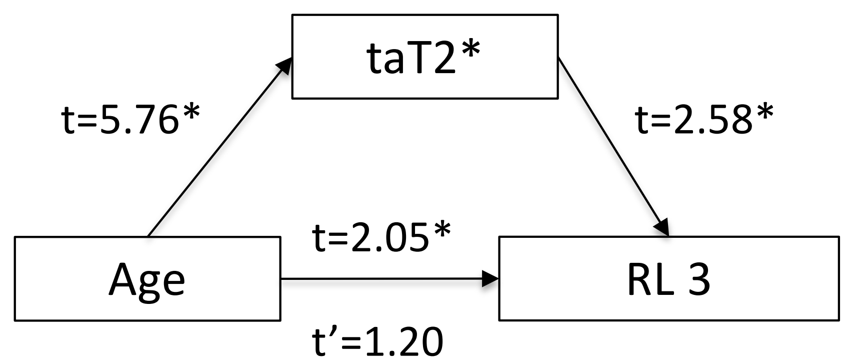


- Replicates in both taT2* and R2'









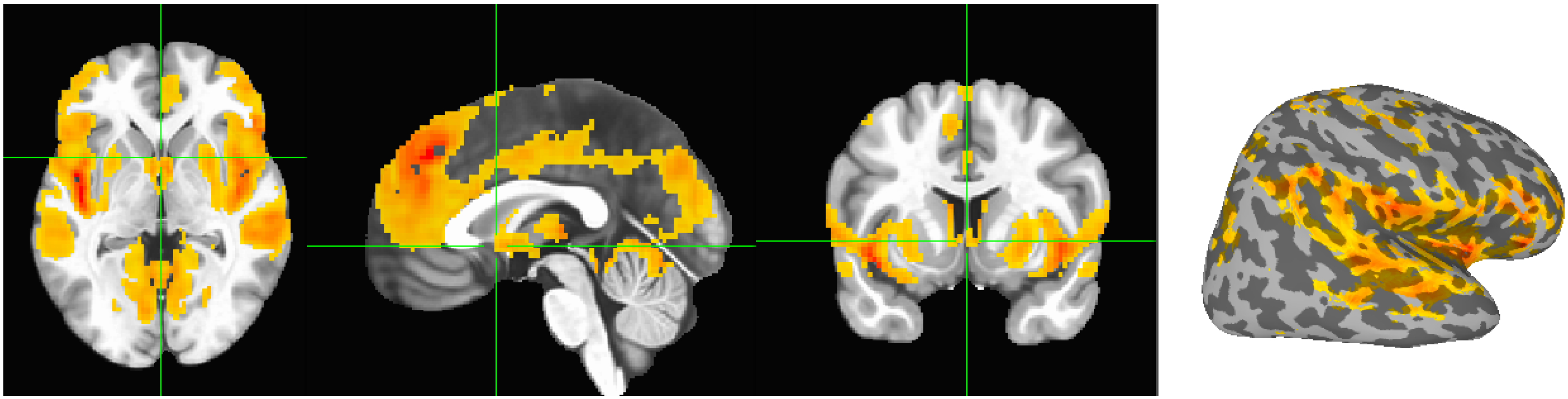
Interim summary

- Development of striatal dopaminergic neurophysiology through adolescence contributes not just to heightened reward sensitivity, but to developmental increases in the ability to make reward-driven choices based on a task-optimal learning strategy.

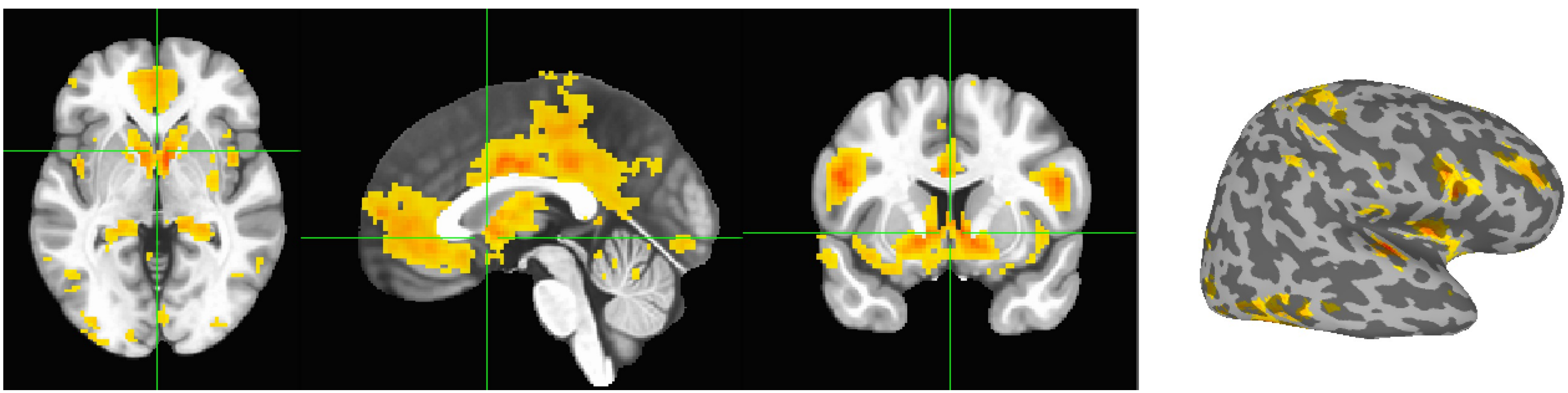
- *What neural computations/activity support this?*

- Two ways to analyze fMRI data:
  - #1: Non-model based

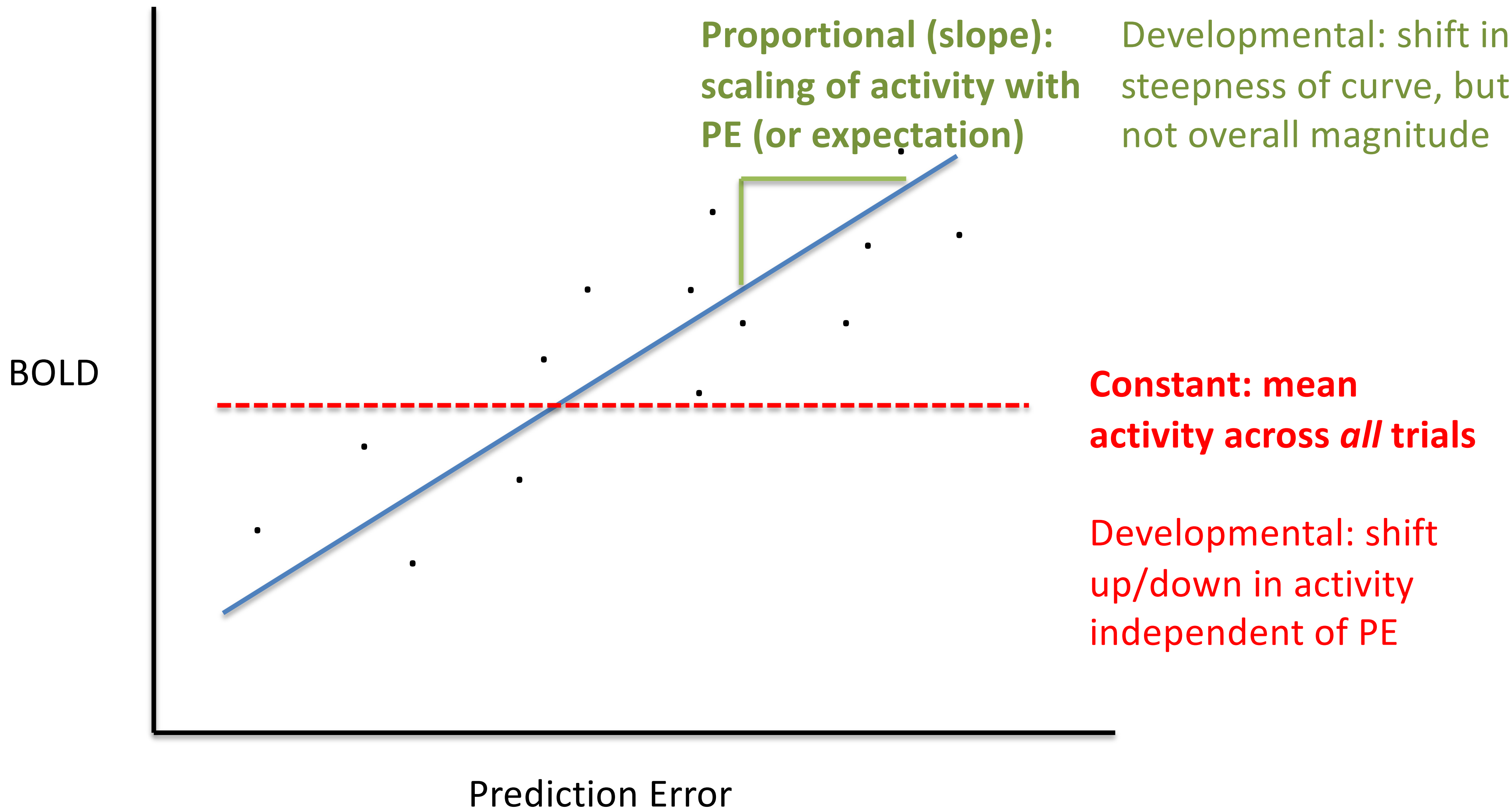
Expectation (hash mark, all trials):



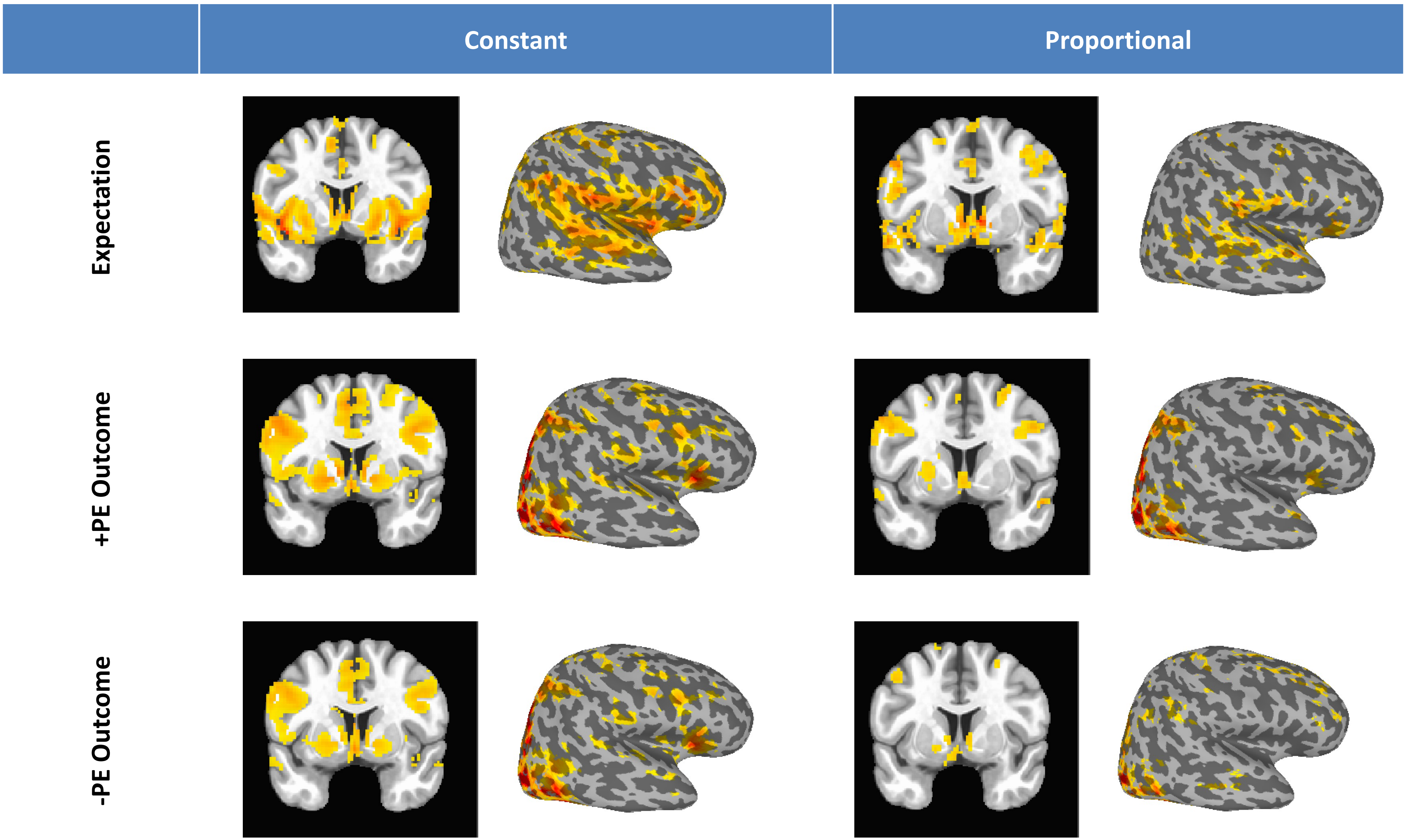
Feedback (rewarded vs non-rewarded):



- Two ways to analyze fMRI data:
  - #2: Parametric (model-based)
  - Example voxel:

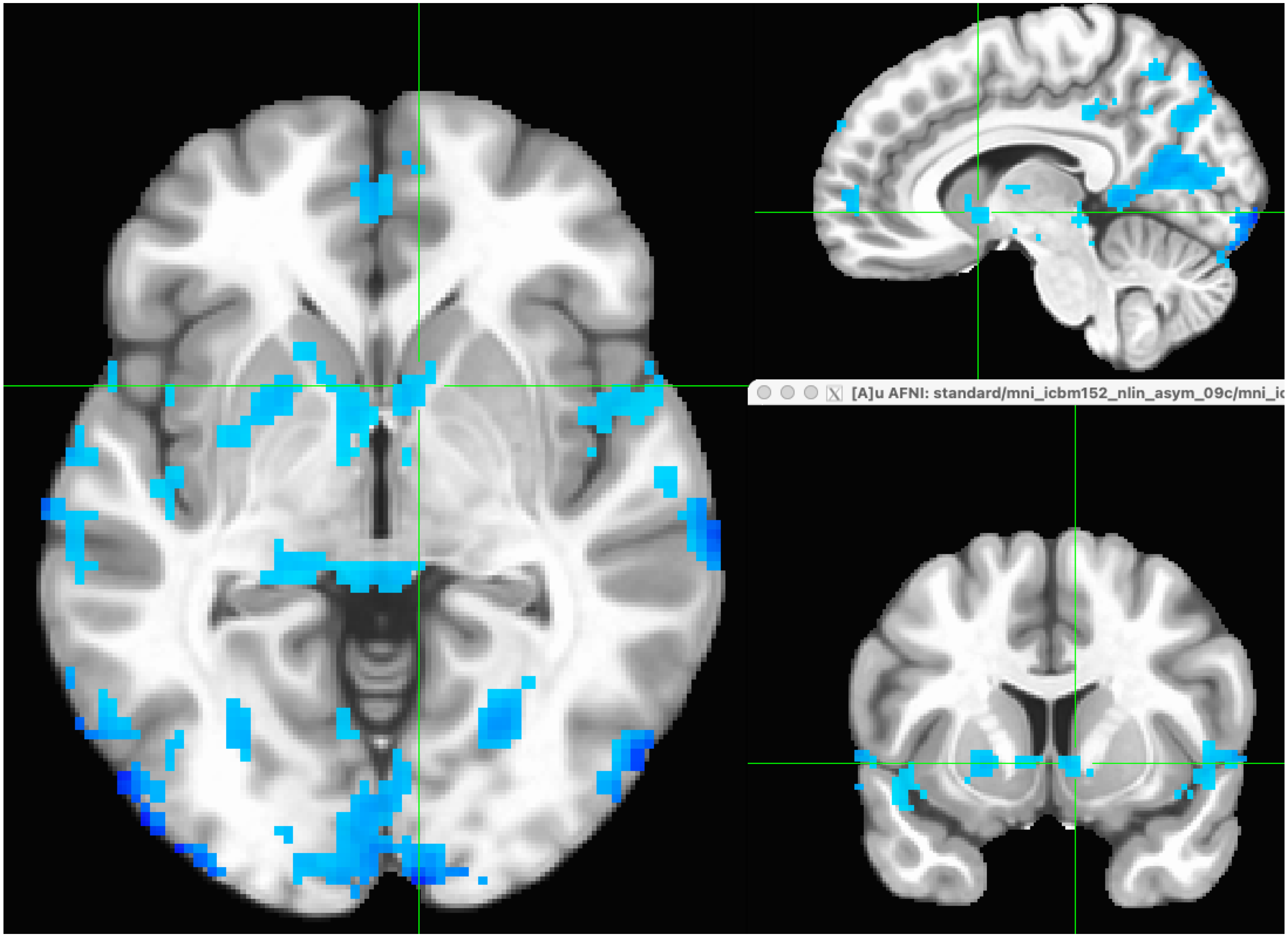


- Two ways to analyze fMRI data:
  - #2: Parametric (model-based)



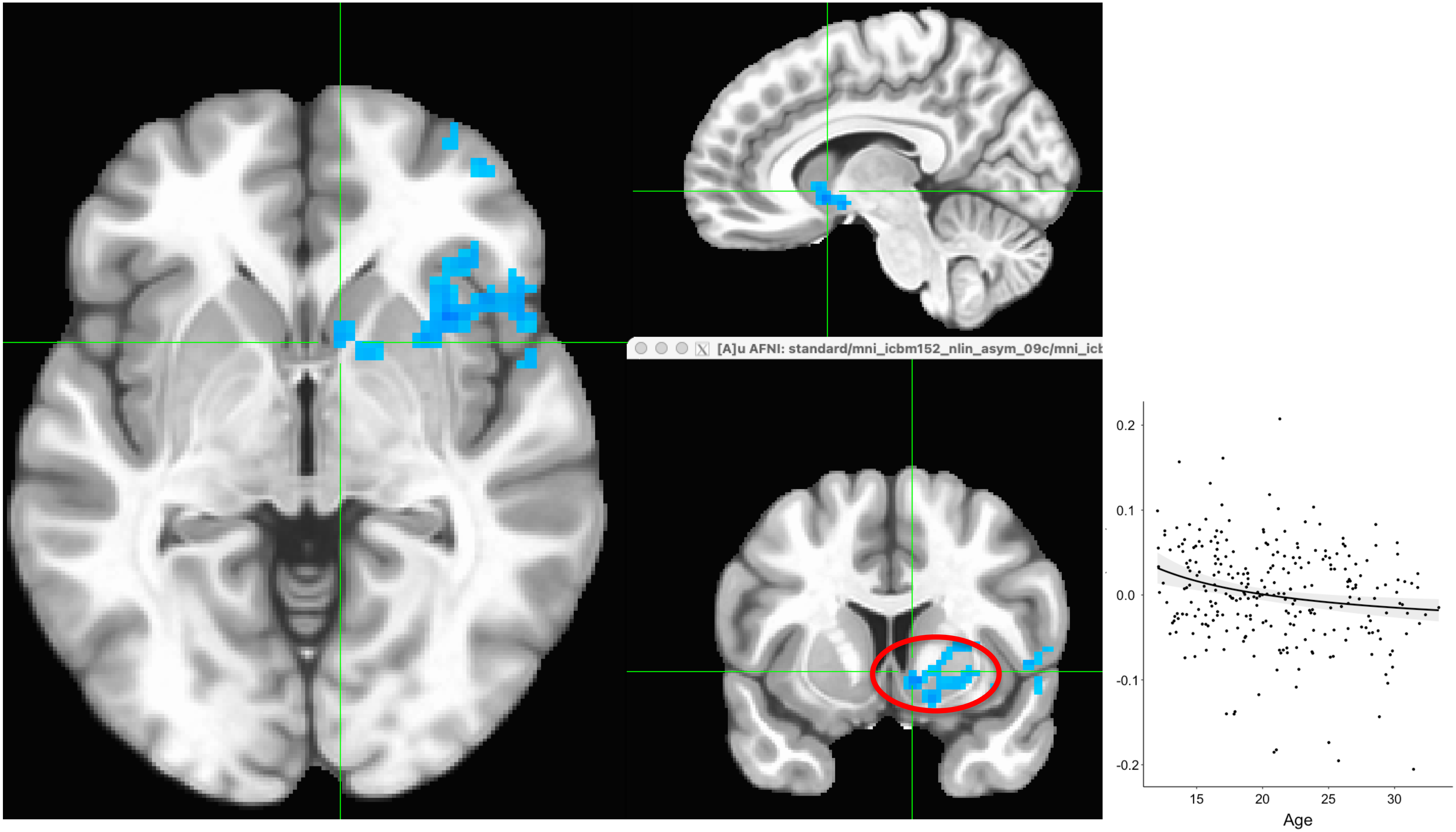
- Many comparisons possible:
  - 8 contrasts of interest (2 traditional + 6 model-based)
  - For each, we are potentially interested in:
    - Main effect
    - Association with
      - Age
      - PC3
      - NAcc tissue iron
    - Interactions
      - Age*PC3
- Group analysis performed using 3dLME
- Cluster correction (ongoing) using ACF-corrected smoothness estimates

- Age-related changes in reward **expectation**



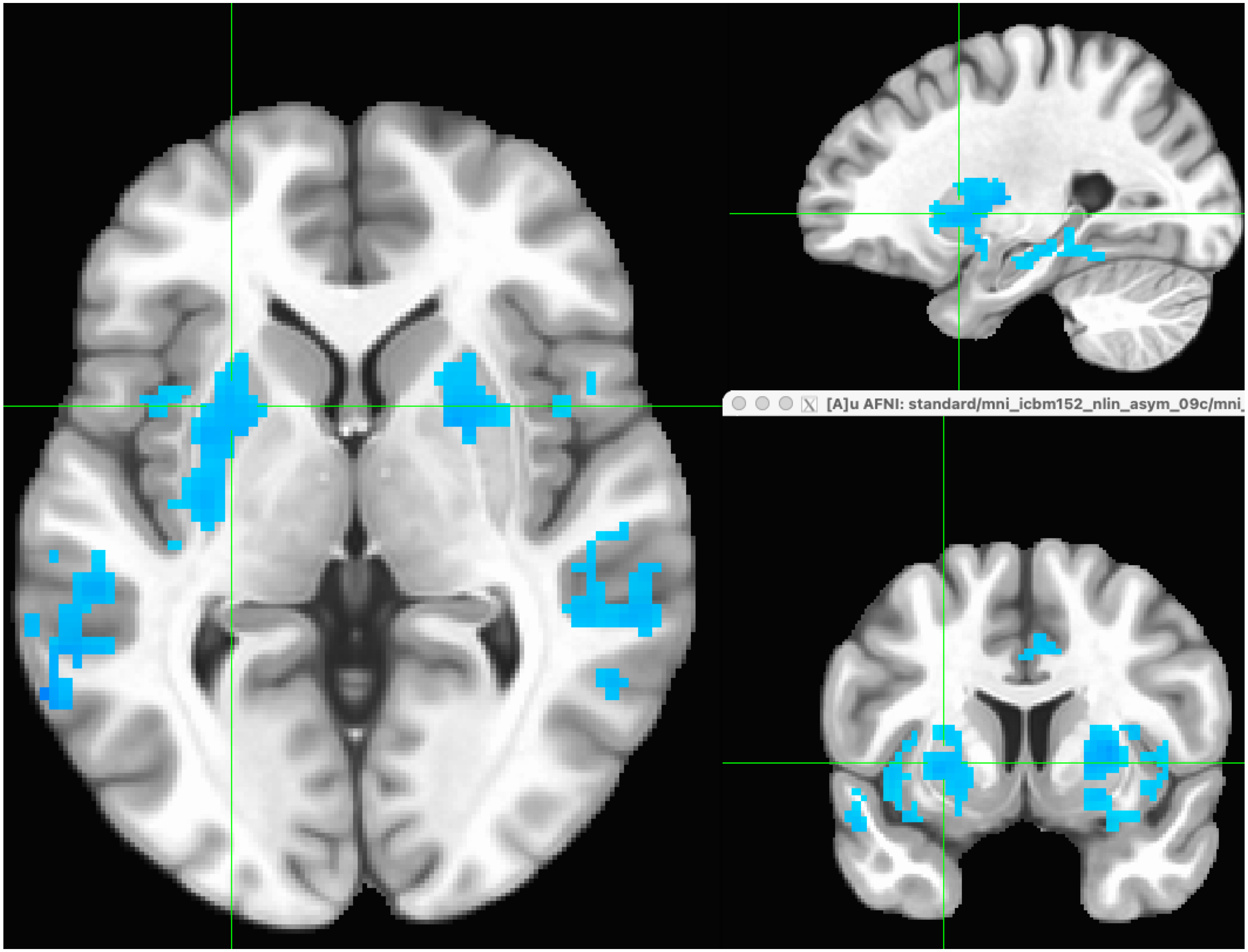
- Predominantly age-related decreases in BOLD expectation response, widely distributed
- Driven mostly by decreases in the mean (not proportional) activation across all trials

- Age-related changes in reward **response**



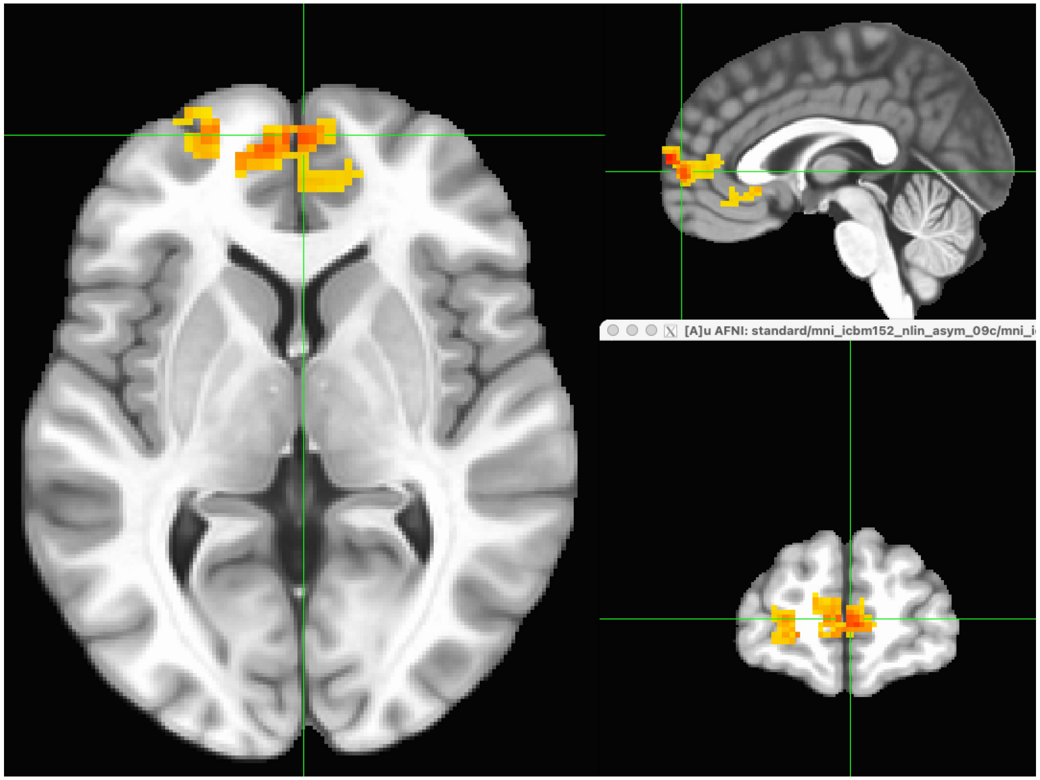
- Predominantly age-related decreases in BOLD reward response, including Nacc
- Most closely associated with changes in the mean response on +PE (i.e., rewarded) outcome

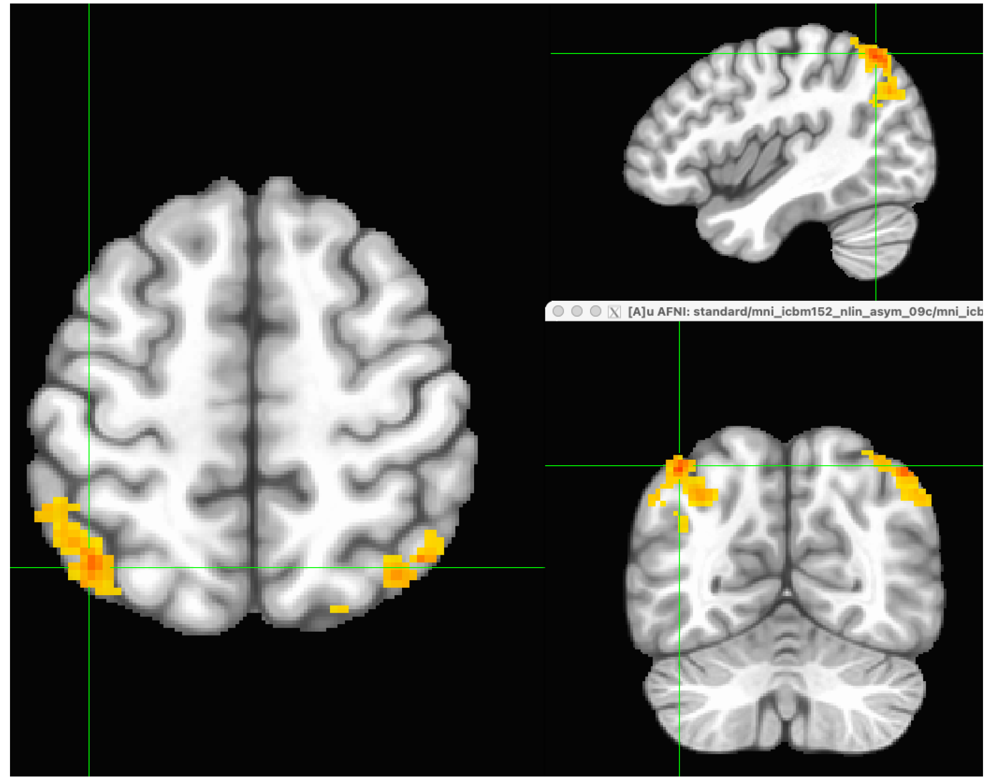
- Associations with PC3 (age independent) during expectation (only):

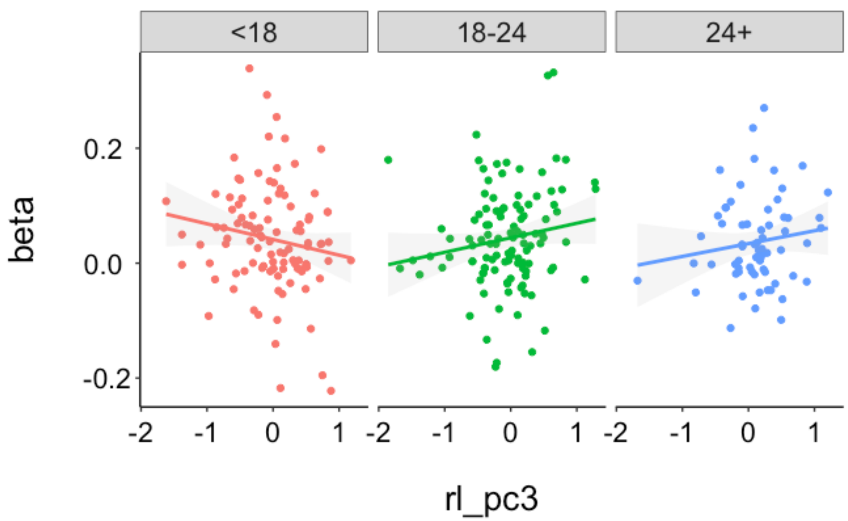


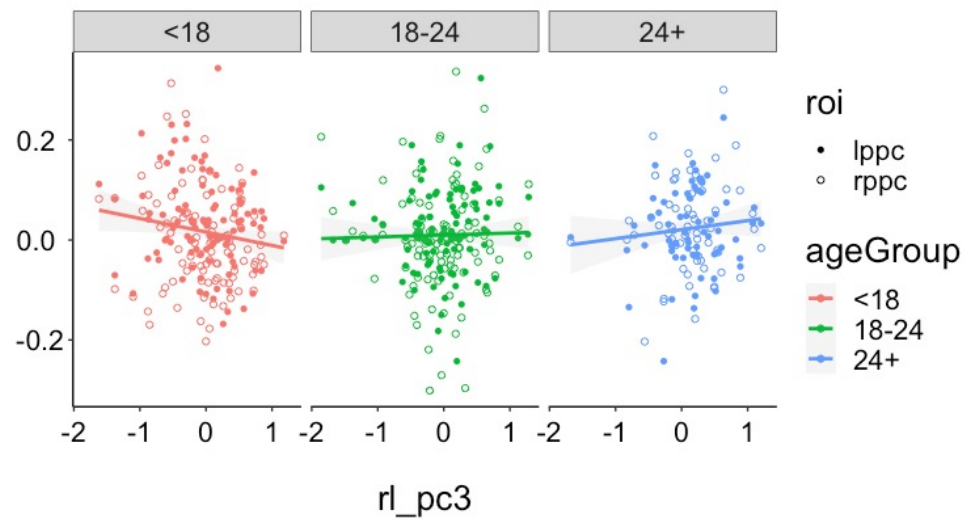
- Increased use of optimal RL strategies is associated with lower ***expectation***-related activation of the putamen, L hippocampus & amygdala, and posterior (STS, PCC) regions

- Age*PC3 interaction in VMPFC, PPC reward response









Summary

- Reward learning improves through adolescence, driven in part by the use of more optimal & reliable (and less exploratory) learning strategies
- Use of these strategies is associated with age-related increases in striatal tissue iron, suggesting a link to DAergic neurophysiology
- Functionally, development of these RL strategies in adulthood is supported by age-dependent activation of the vmPFC & PPC
  - Activation of these regions in adolescence may instead promote more exploratory strategies?
- Combined with decreased VST reward responses, this may represent a shift from subcortical to cortically-dependent processing supporting the transition of reward learning from adolescence to adulthood