

# Human development as stochastic optimization in the space of learning strategies

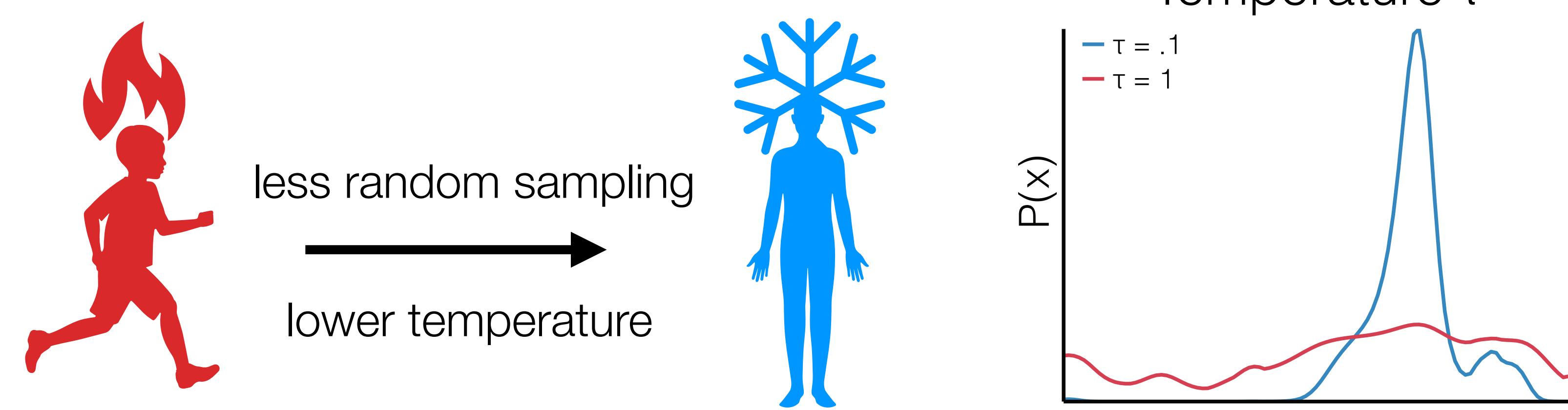
Anna Giron<sup>1\*</sup>, Simon Ciranka<sup>2,3\*</sup>, Eric Schulz<sup>4</sup>, Wouter van den Bos<sup>5</sup>, Azzurra Ruggeri<sup>2,6</sup>, Björn Meder<sup>2,7</sup>, Charley M. Wu<sup>1,2</sup> ([charley.wu@uni-tuebingen.de](mailto:charley.wu@uni-tuebingen.de))

<sup>1</sup>University of Tübingen; <sup>2</sup>Max Planck Institute for Human Development; <sup>3</sup>Max Planck UCL Centre for Comp2Psych; <sup>4</sup>Max Planck Institute for Biological Cybernetics; <sup>5</sup>University of Amsterdam; <sup>6</sup>Technical University Munich; <sup>7</sup>Health and Medical University Potsdam; \* Equal contributions

## How does learning change over the lifespan?

Human development is often compared to *stochastic optimization*, since children exhibit high variability in their choice behavior, which gradually “cools off over age” [1]. Yet is this simply a reduction in randomness or a more complex search process in the space of learning strategies?

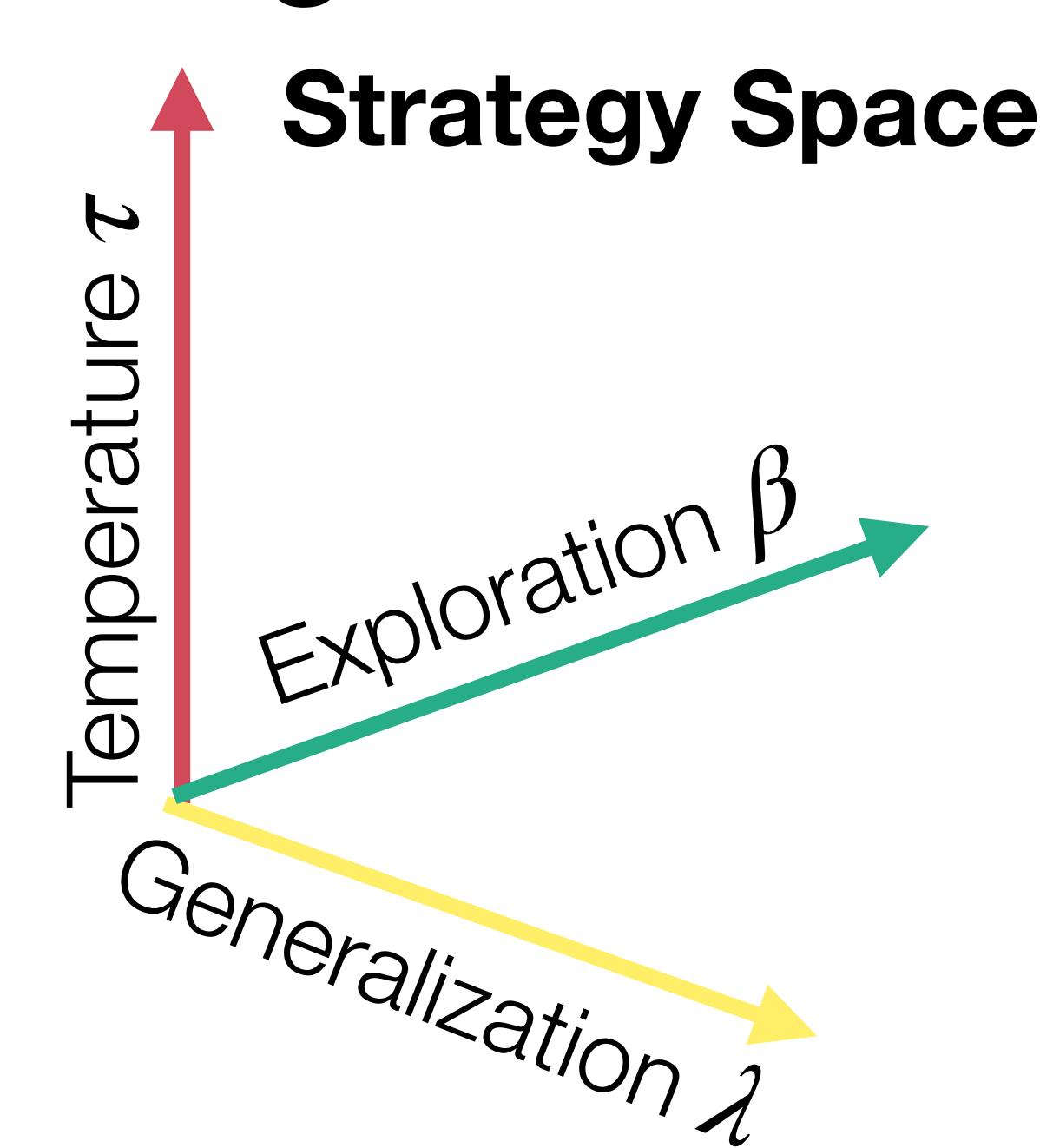
### H1: Randomness reduces over the lifespan



### H2: “Cooling off” applies to changes in the latent space of learning strategies

In addition to **temperature  $\tau$** , learning can be characterized by:

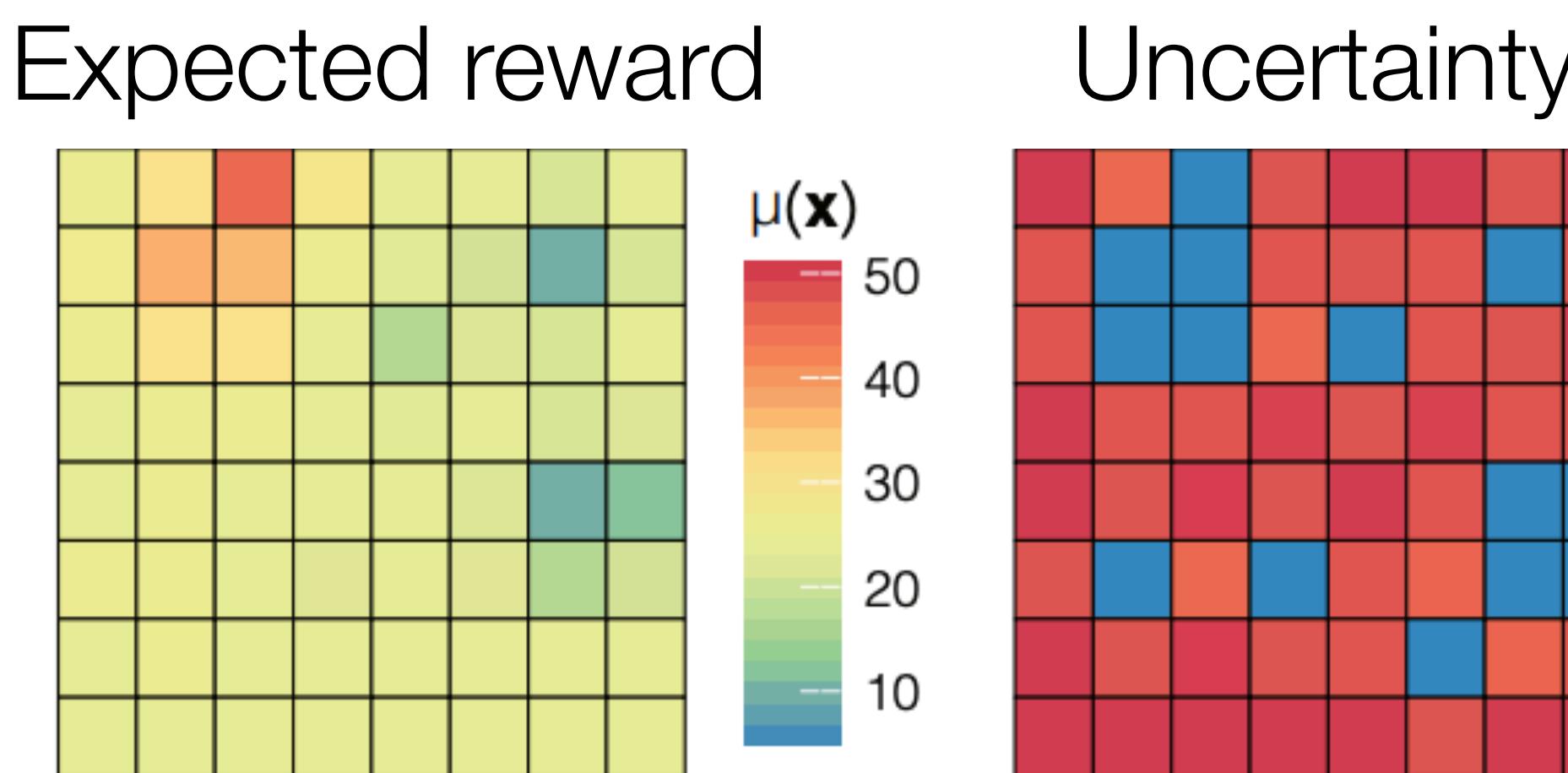
- changes in uncertainty-directed **exploration  $\beta$**  [2-3]
- refinement of **generalization  $\lambda$**



### GP-UCB model captures all 3-dimensions

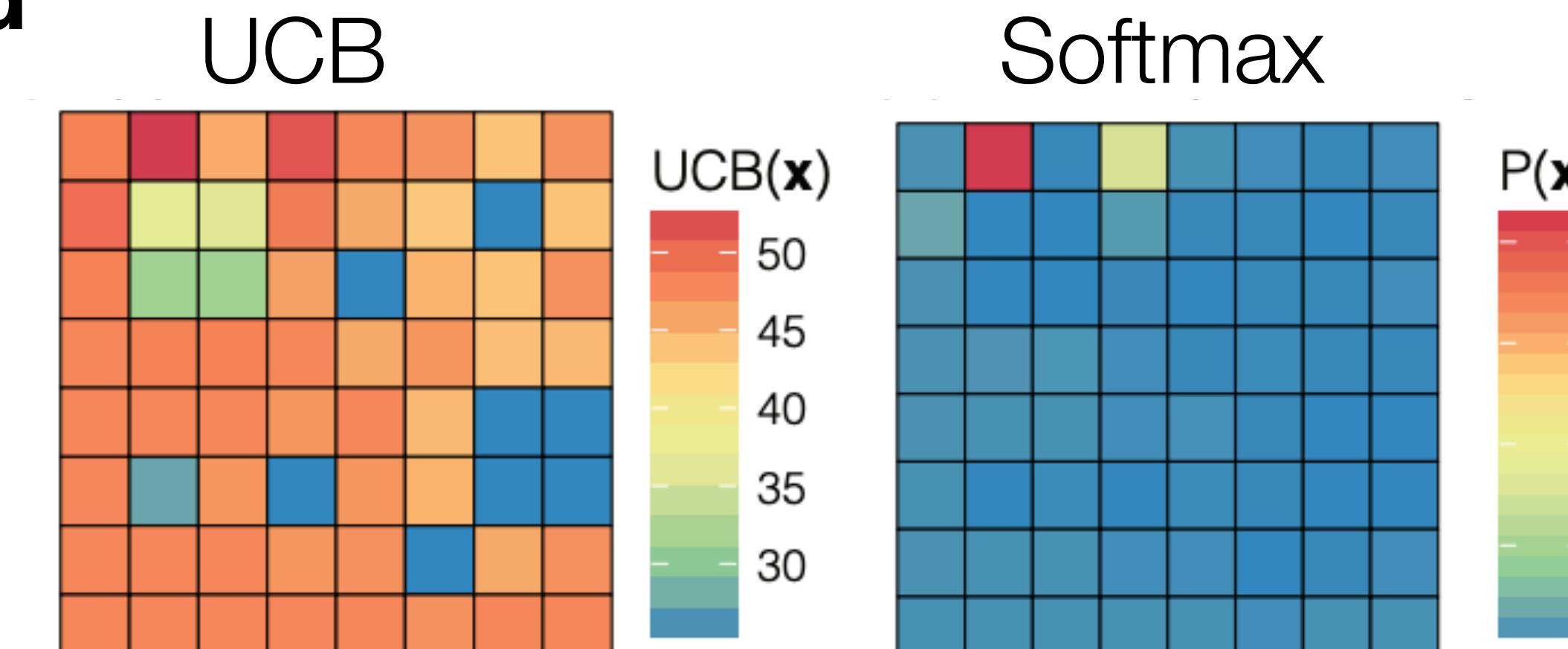
**Gaussian Process (GP)** regression with an RBF kernel predicts reward, where  $\lambda$  captures generalization

$$k_{RBF}(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\lambda^2}\right)$$

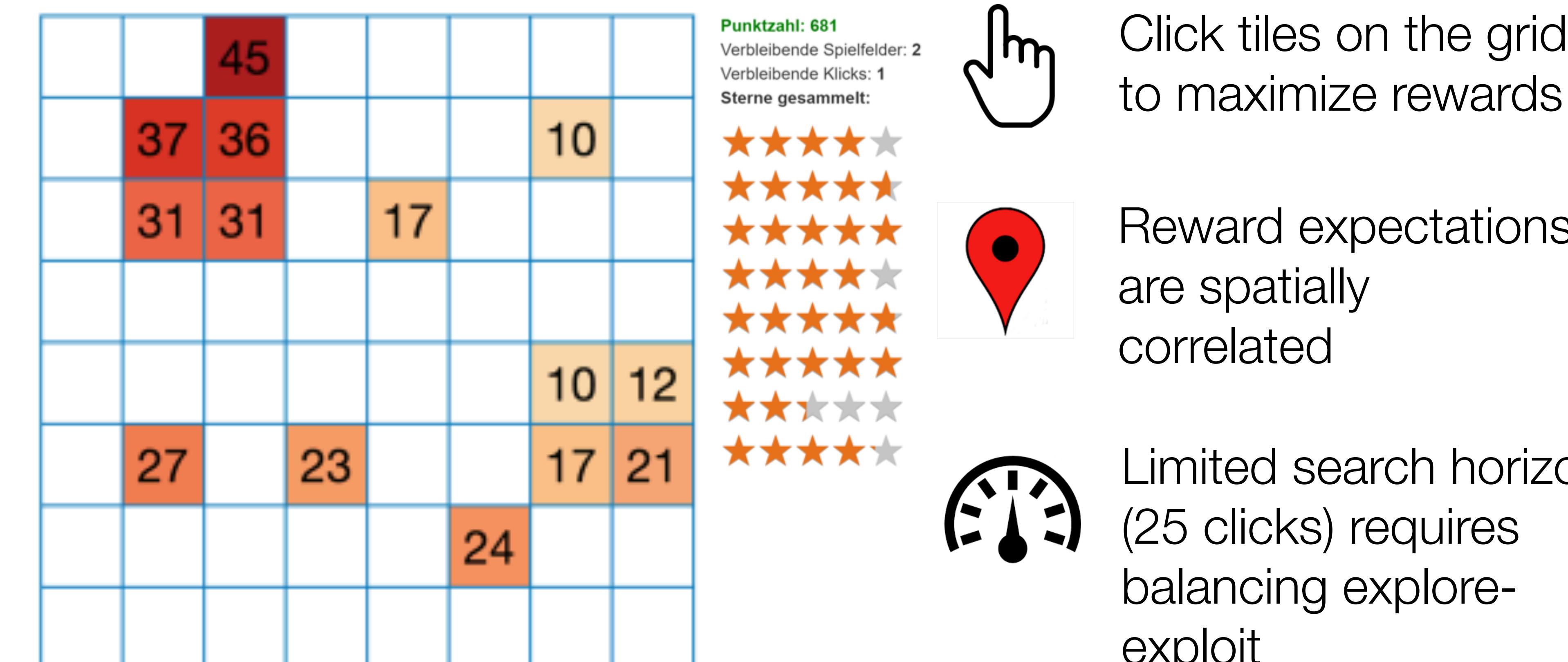


**Upper confidence bound (UCB)** sampling uses  $\beta$  to define the value of exploration relative to exploitation

$$UCB(\mathbf{x}_i) = \mu(\mathbf{x}_i) + \beta\sigma(\mathbf{x}_i)$$



## Spatially correlated bandit task

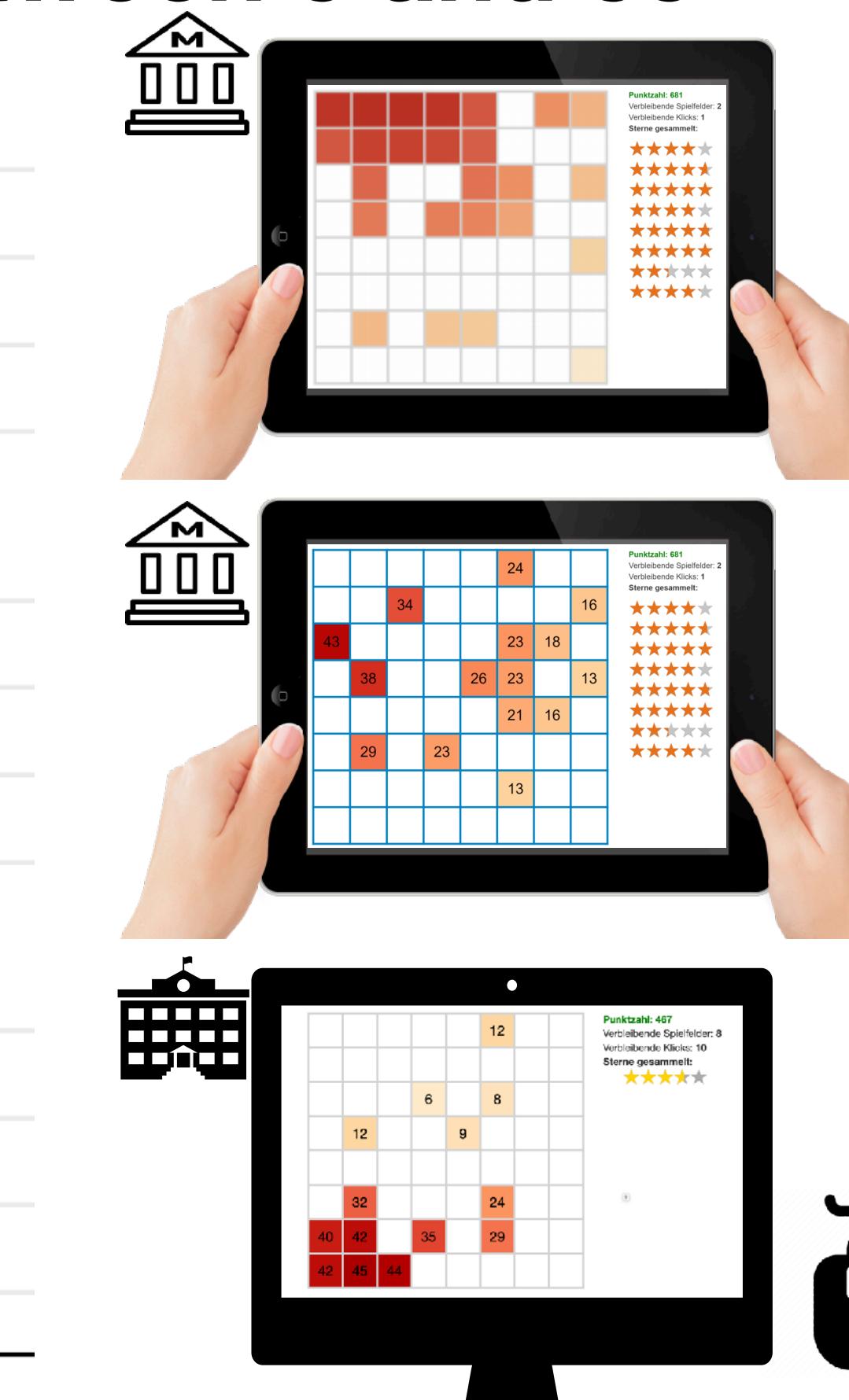
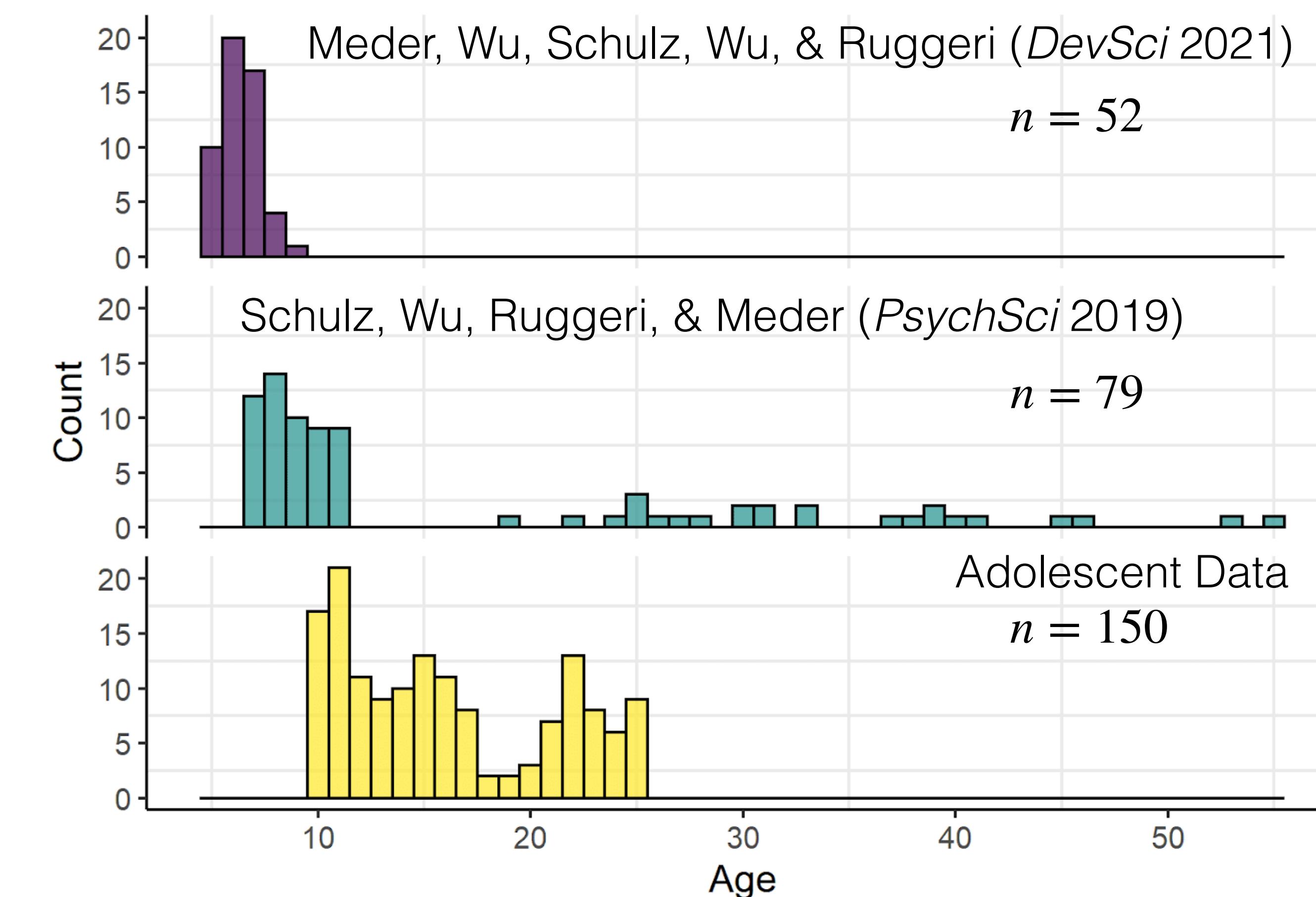


Click tiles on the grid to maximize rewards

Reward expectations are spatially correlated

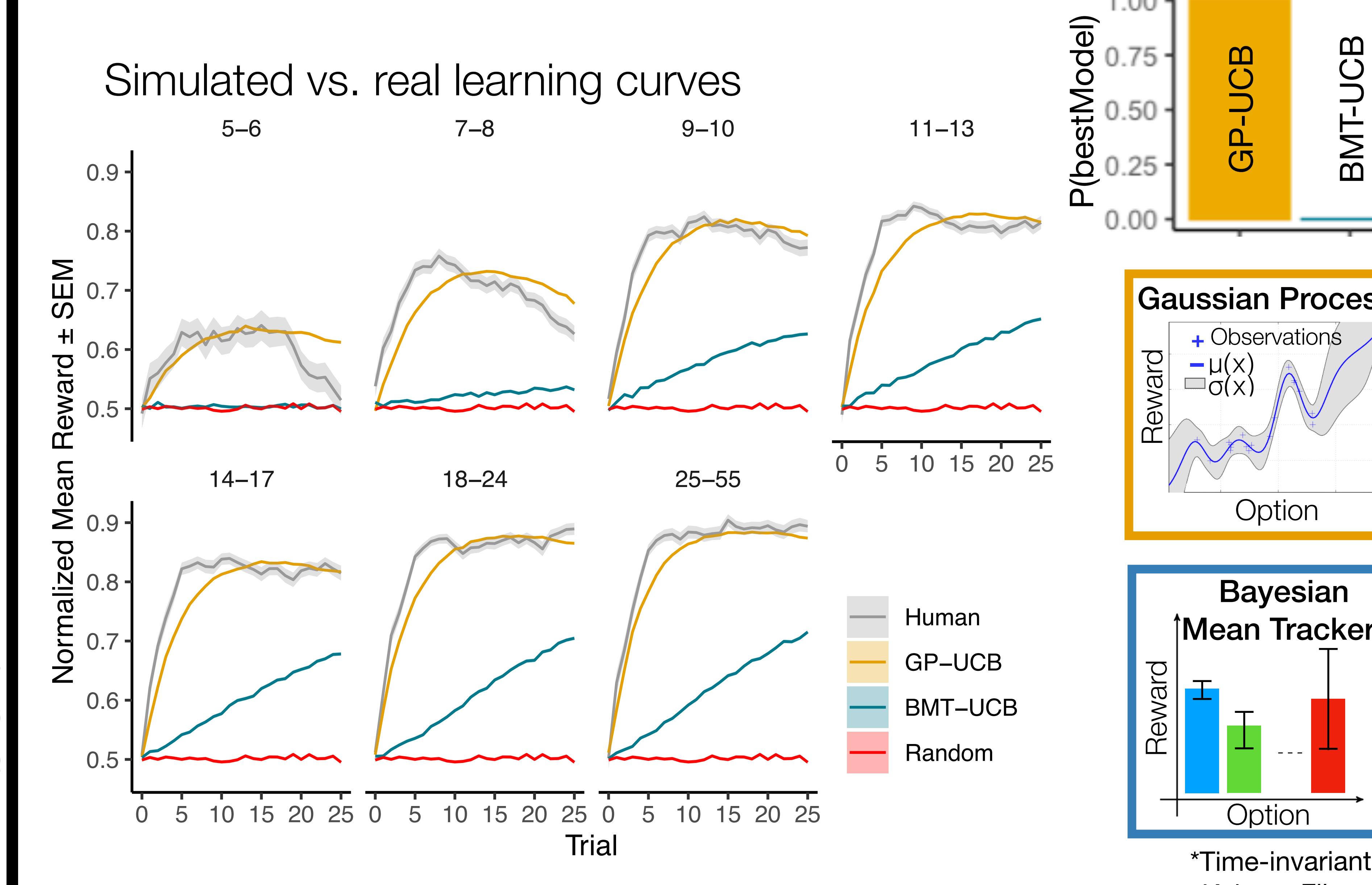
Limited search horizon (25 clicks) requires balancing explore-exploit

### Combined dataset with $n=281$ subjects between 5 and 55



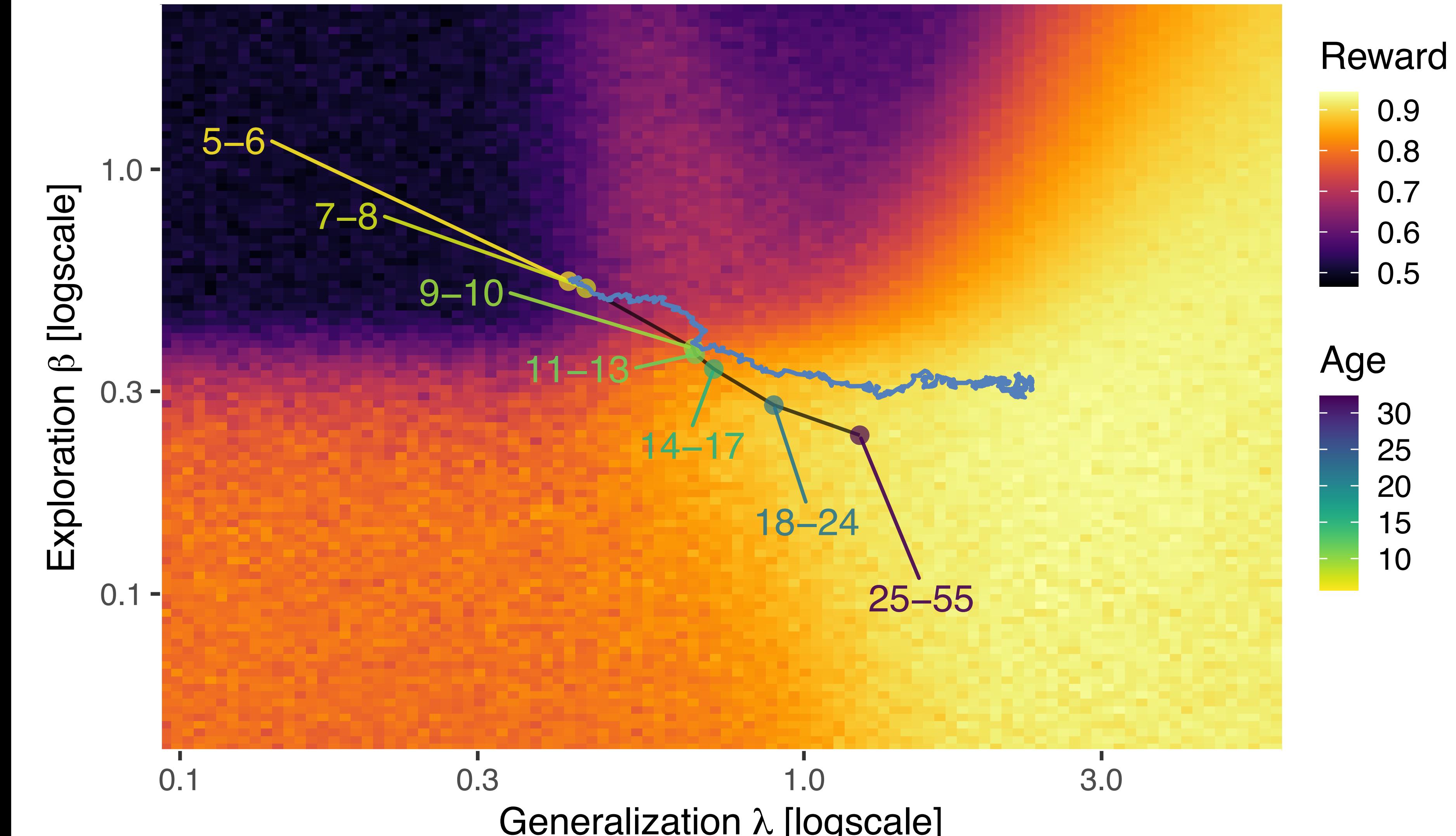
## Model Comparison

### Simulated vs. real learning curves



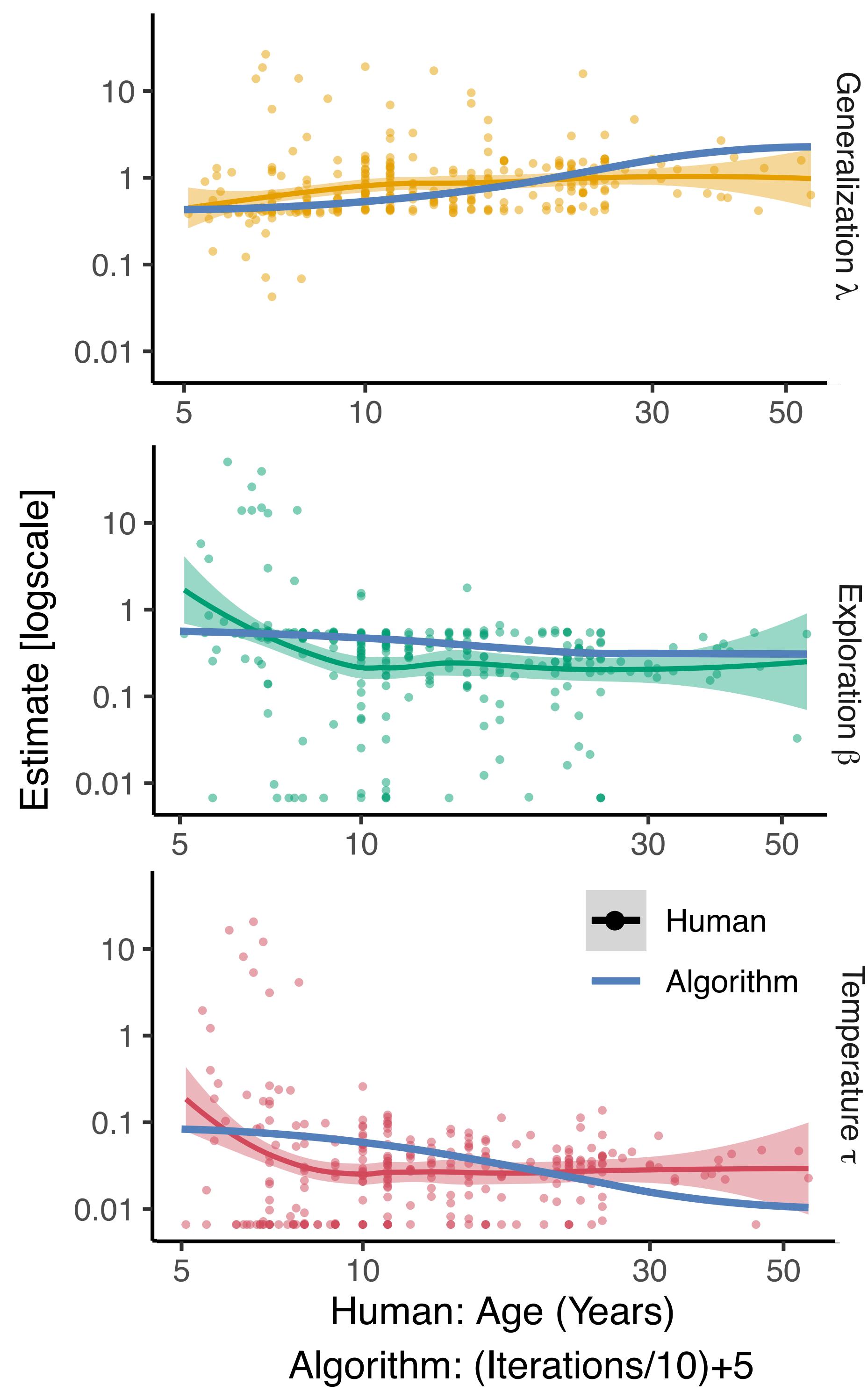
## Main Results

### Human vs. SGD trajectories



Using a fitness landscape computed across 1 million parameter combinations, we find that human development is highly similar to stochastic gradient descent (SGD), but starts to deviate during adolescence

- Human **generalization  $\lambda$**  and uncertainty-directed **exploration  $\beta$**  change much faster than SGD during childhood but peak at lower levels
- Human **temperature  $\tau$**  drops rapidly in childhood but stays much higher than SGD



**Conclusion:** Not just reduction in randomness, but close resemblance between development and stochastic optimization

- [1] Gopnik, A. et al. Changes in cognitive flexibility and hypothesis search across human life history from childhood to adolescence to adulthood. *Proc. Natl. Acad. Sci.* 114, 7892–7899 (2017)
- [2] Schulz, E., Wu, C. M., Ruggeri, A. & Meder, B. Searching for rewards like a child means less generalization and more directed exploration. *Psychol. Sci.* 30, 1561–1572 (2019).
- [3] Meder, B., Wu, C. M., Schulz, E. & Ruggeri, A. Development of directed and random exploration in children. *Dev. Sci.* e13095 (2021).
- [4] Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D. & Meder, B. Generalization guides human exploration in vast decision spaces. *Nat. Hum. Behav.* 2, 915–924 (2018).