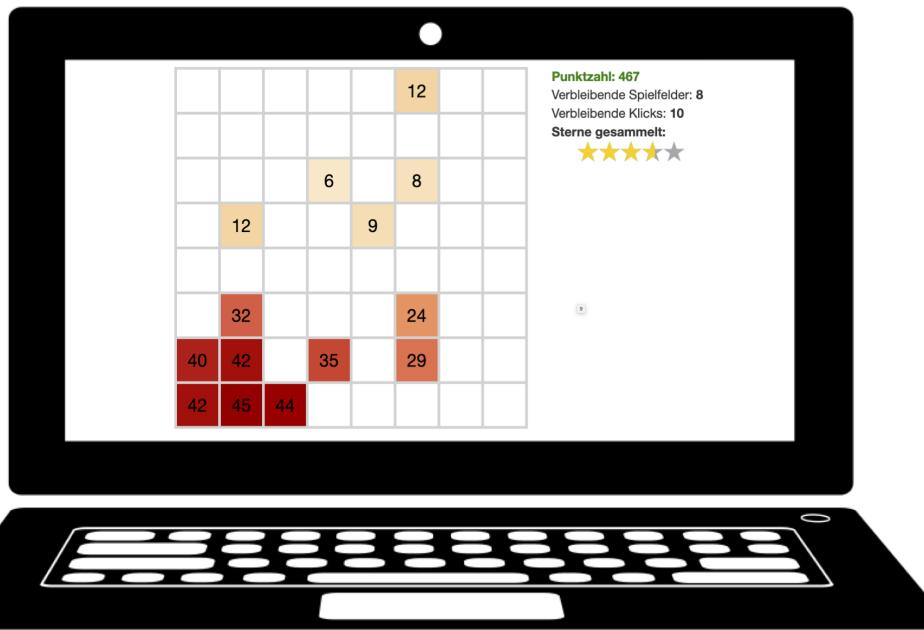
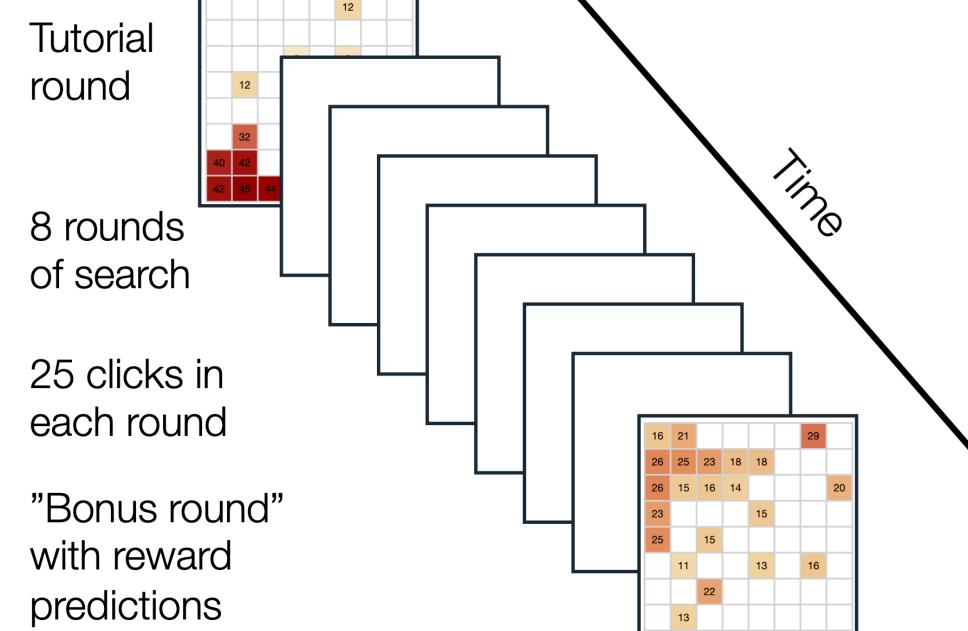


a Experiment and design

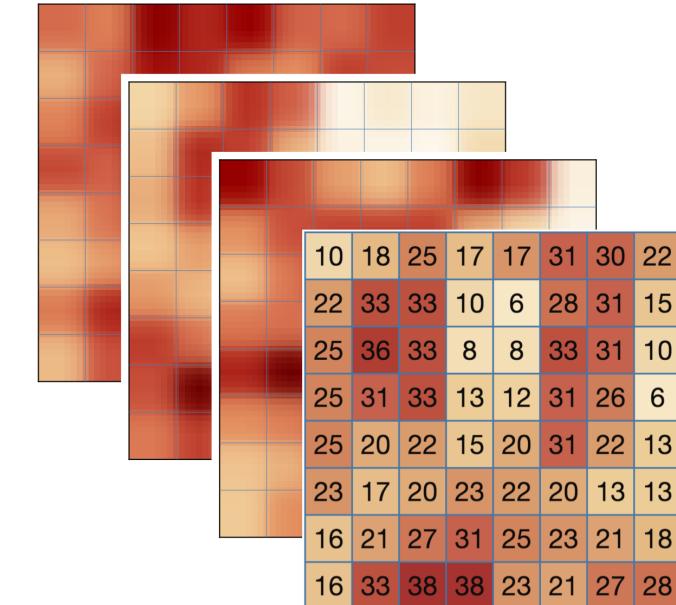
Task



Procedure

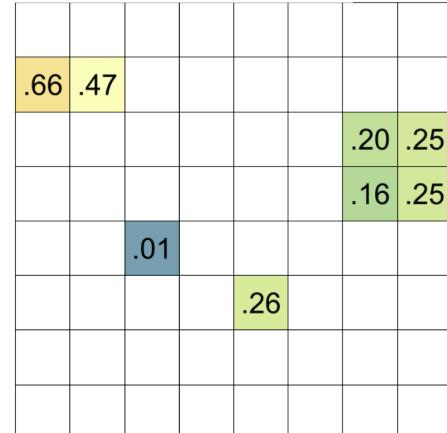


Environments

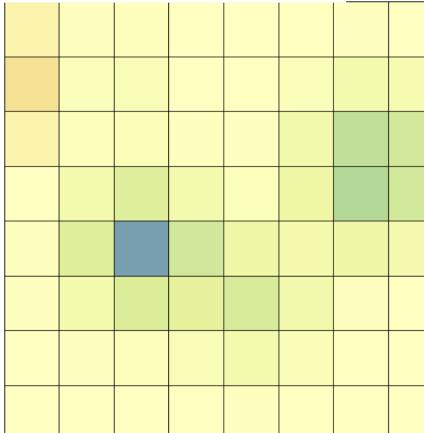


b Computational GP–UCB model

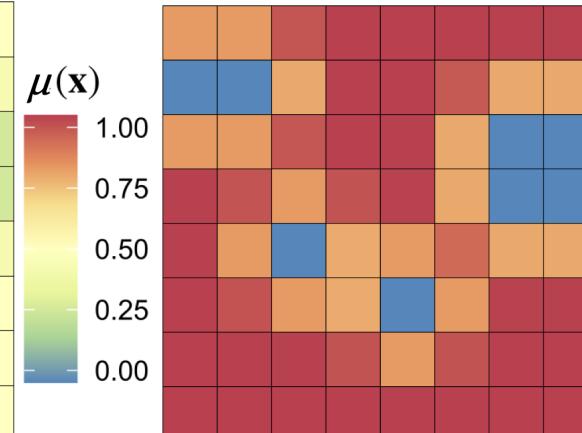
Observations ($t=8$)



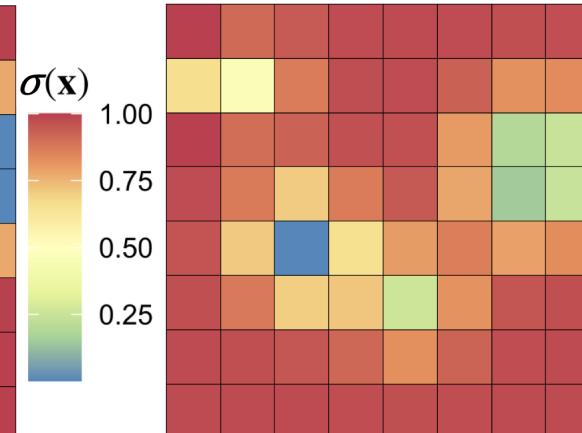
Expected rewards



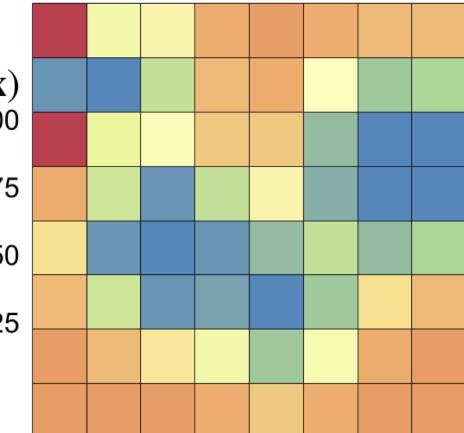
Uncertainty



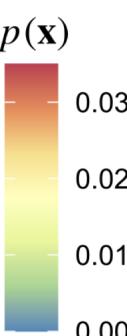
Upper Confidence Bound (UCB)



Softmax predictions



$p(x)$



A vertical color scale for softmax predictions, ranging from 0.00 (blue) to 1.00 (red).

Generalization λ

Exploration bonus β

Random temperature τ