A Task and procedure
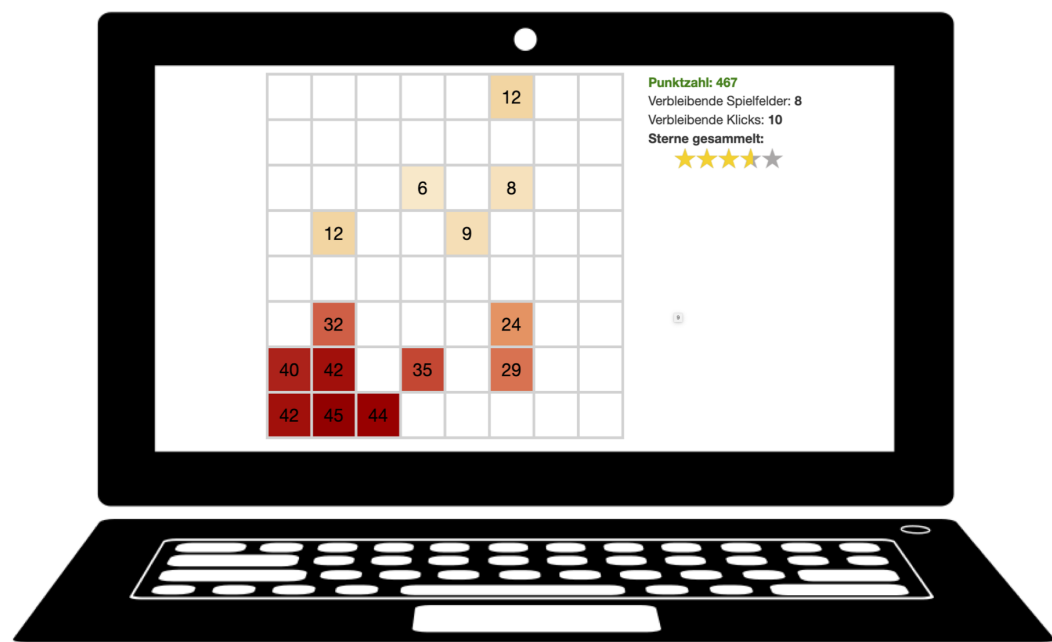
Task

Procedure

Tutorial round

8 rounds of search

25 clicks in each round
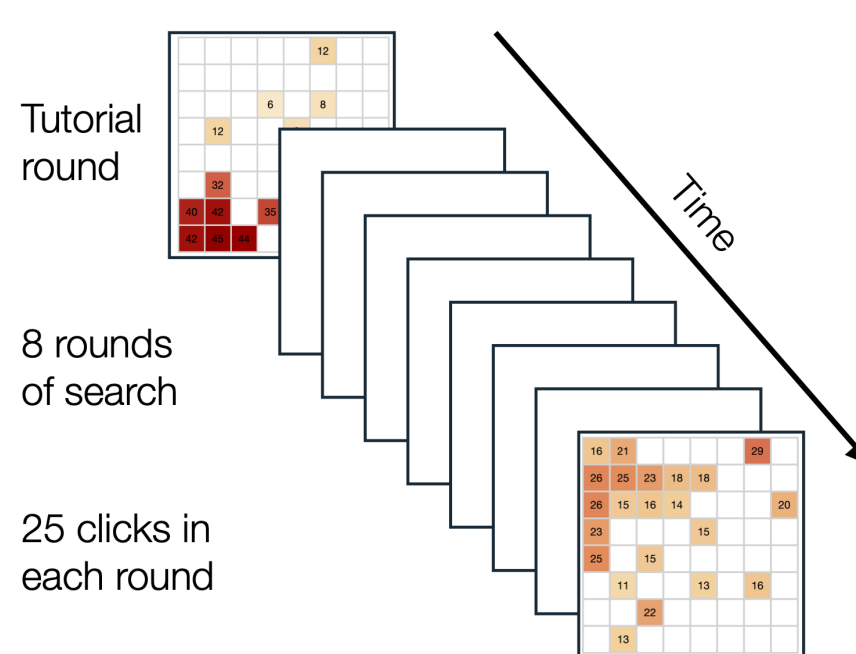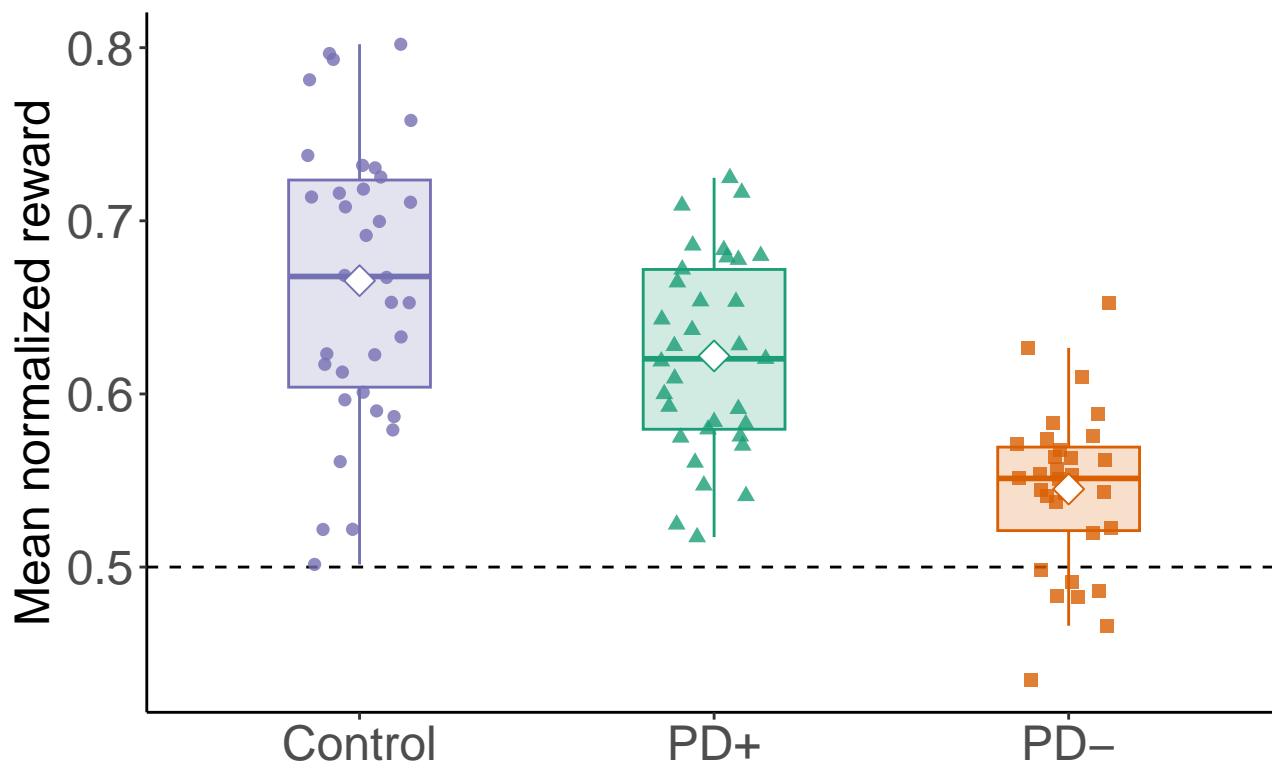
Time

Environments
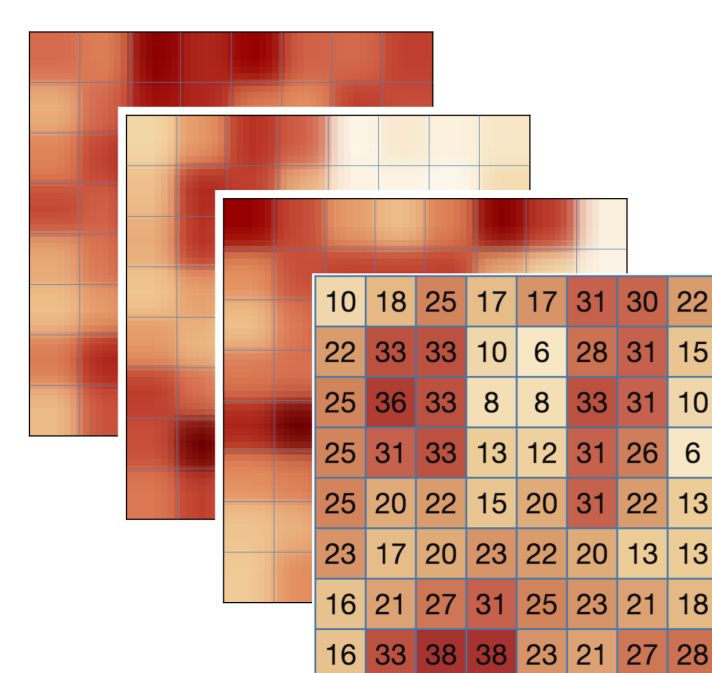
B Performance

C Learning curves

- Control
- PD+
- PD−
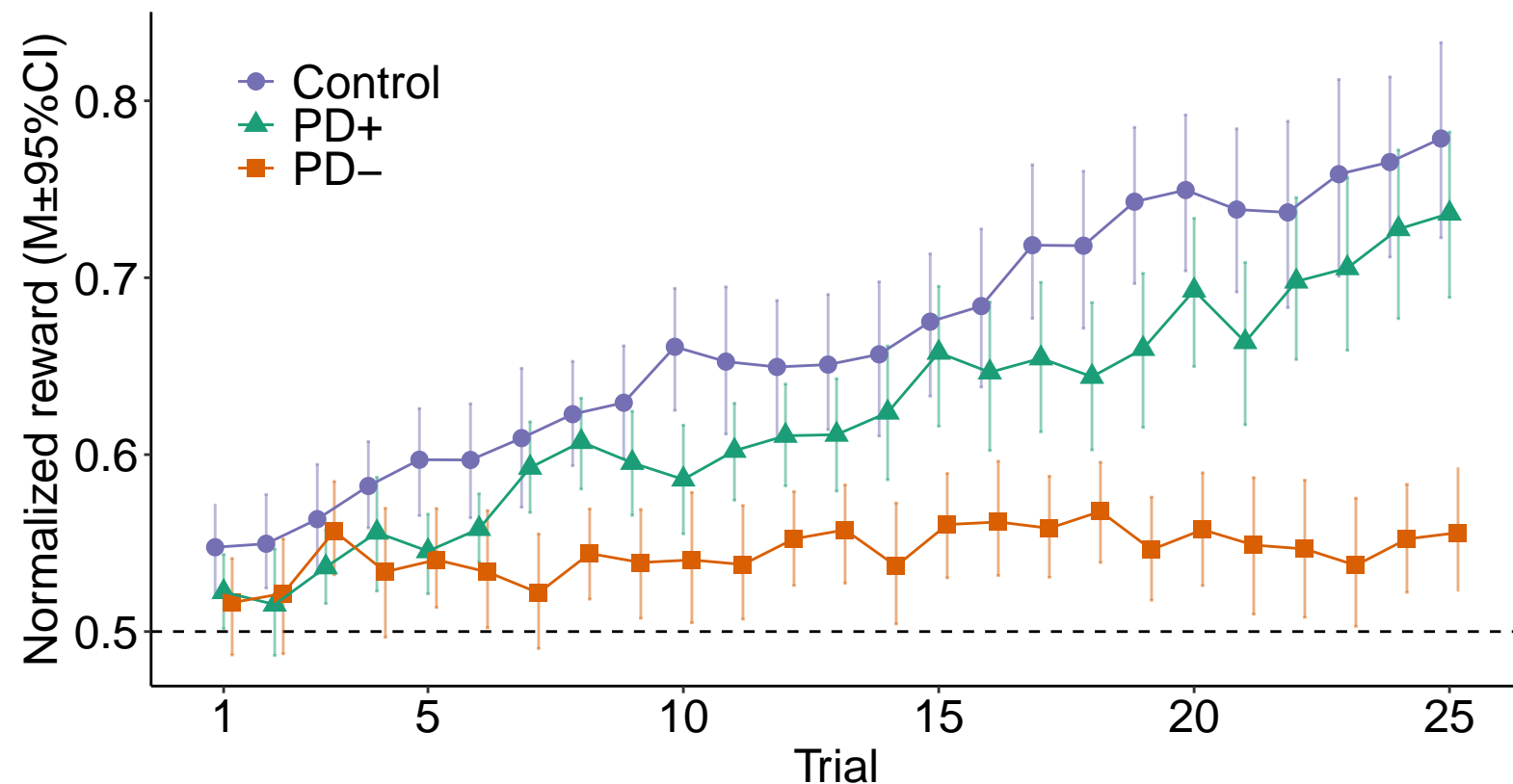
D Exploitation choices
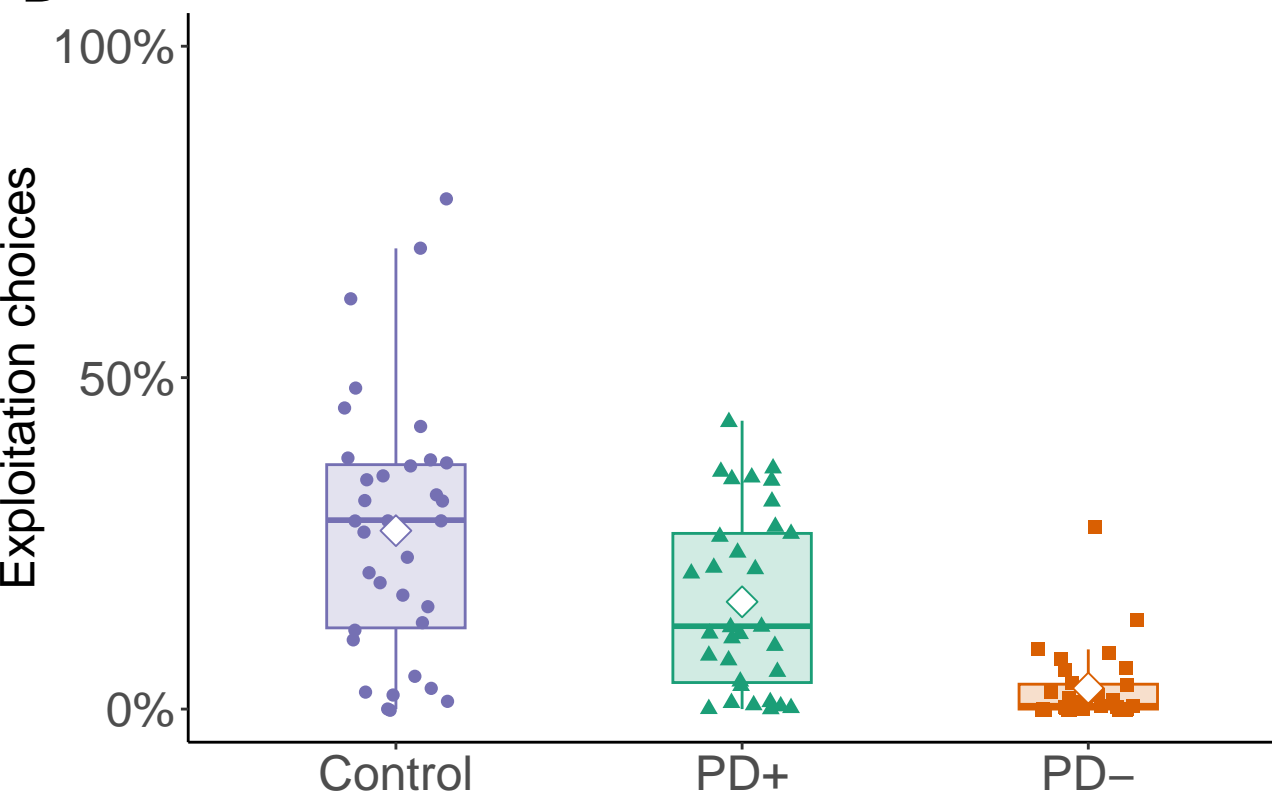
E Exploration and exploitation over time

- Control
- PD+
- PD−
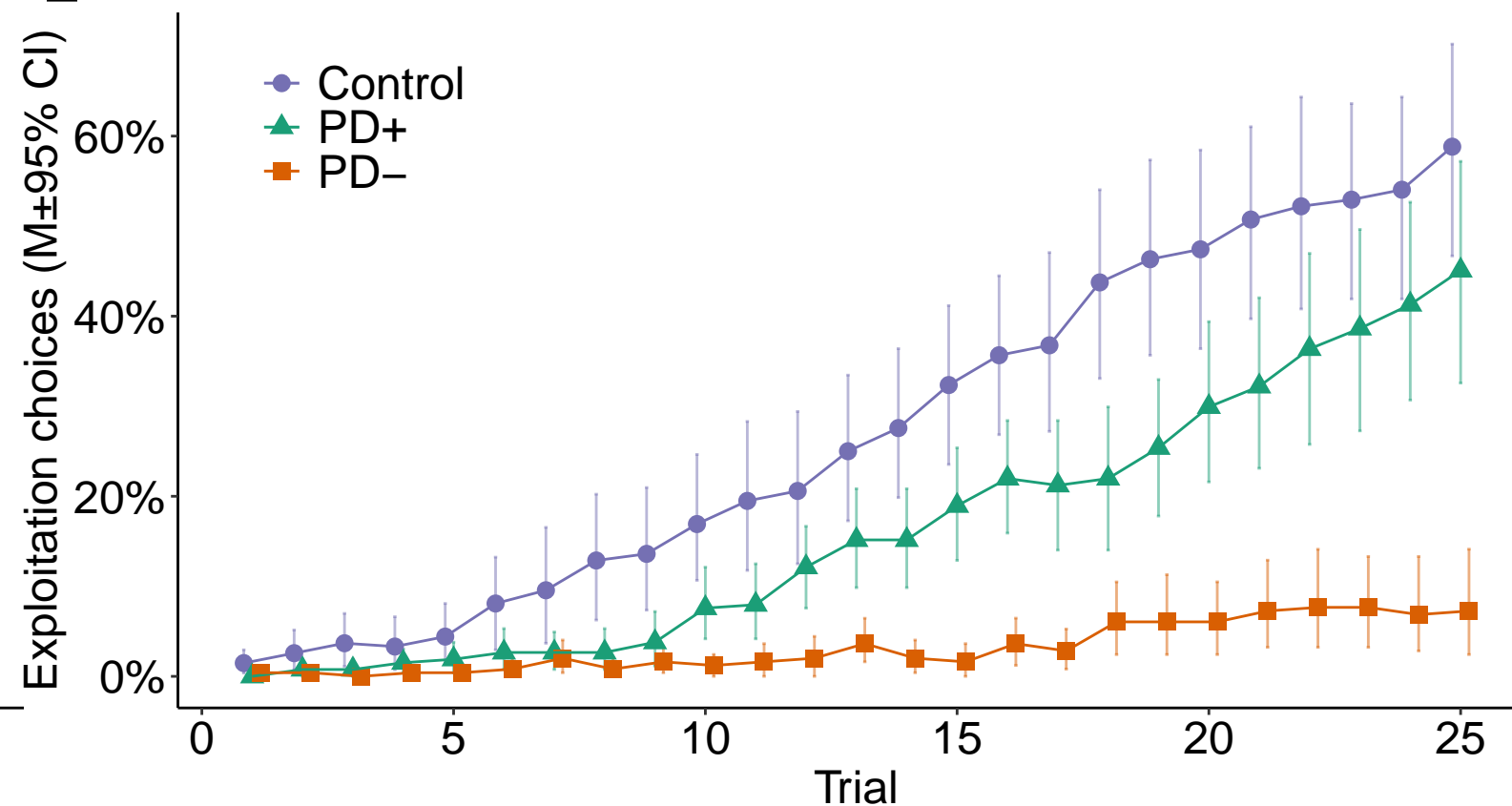
F Distance consecutive choices

Control Mean: 1.8
PD+ Mean: 2.1
PD− Mean: 2.1
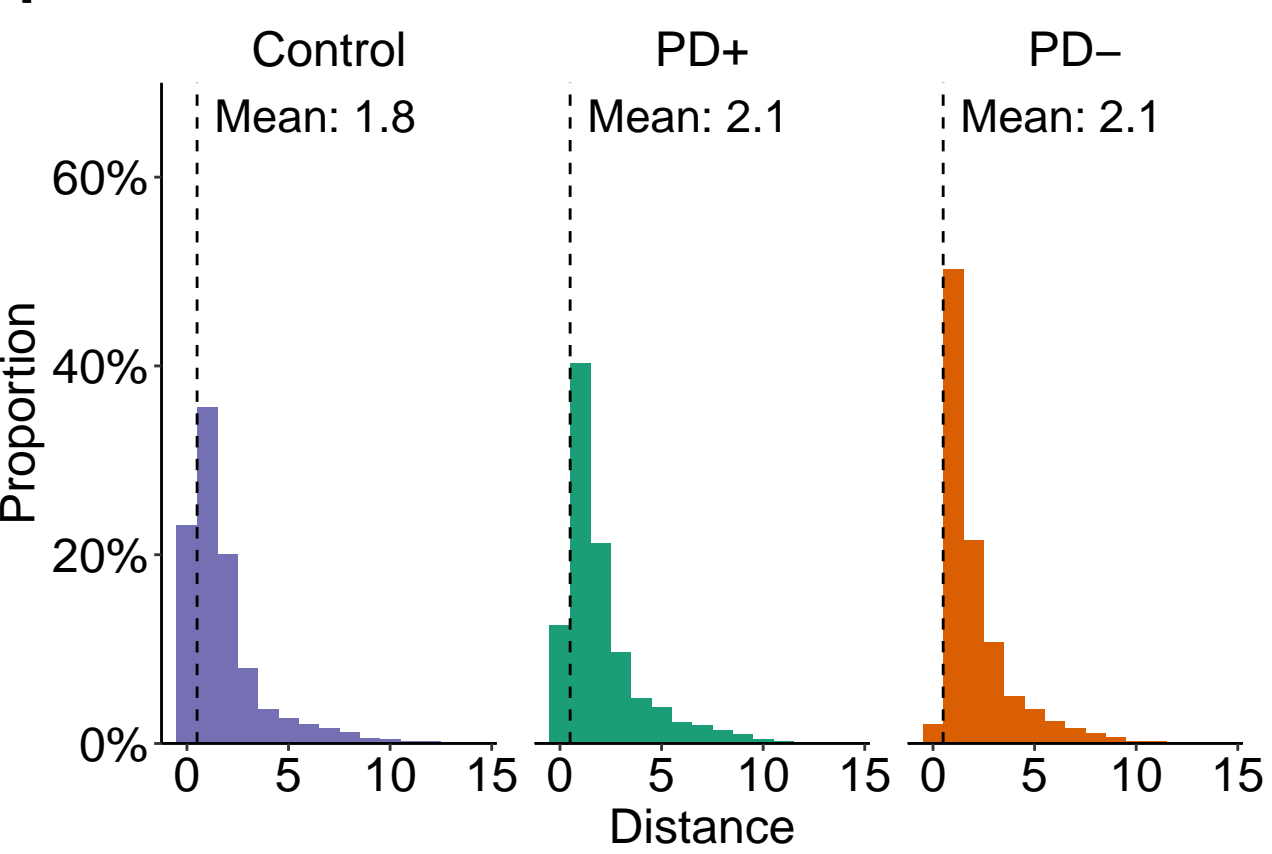
G Search Distance ~ Previous Reward

- Control
- PD+
- PD−