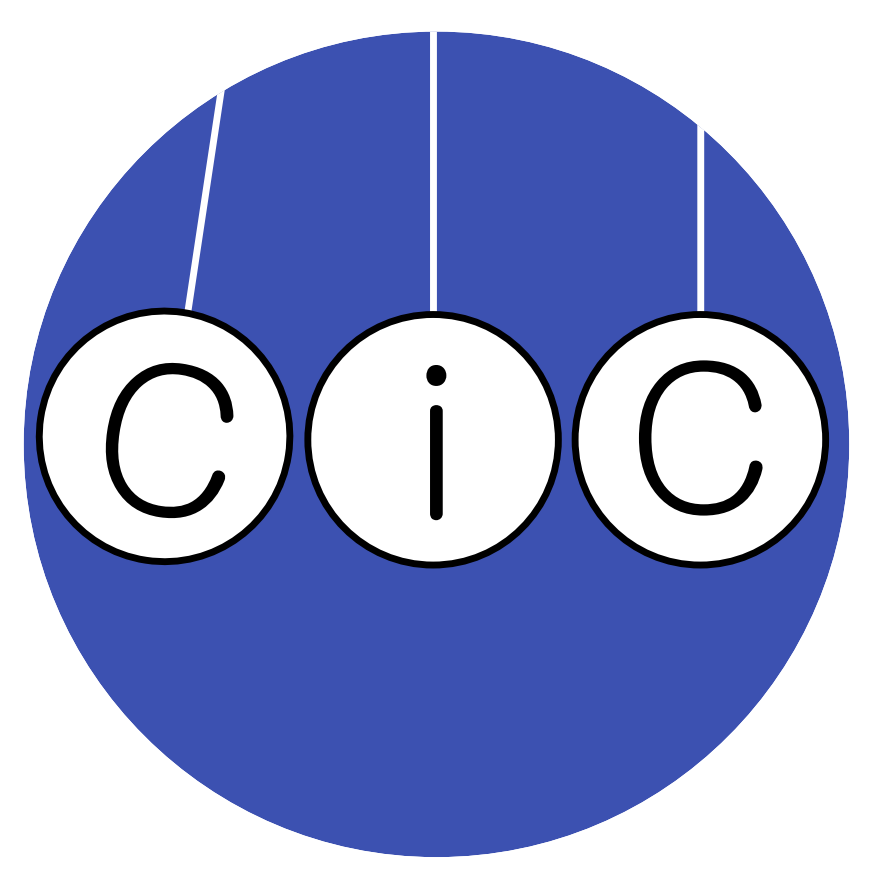




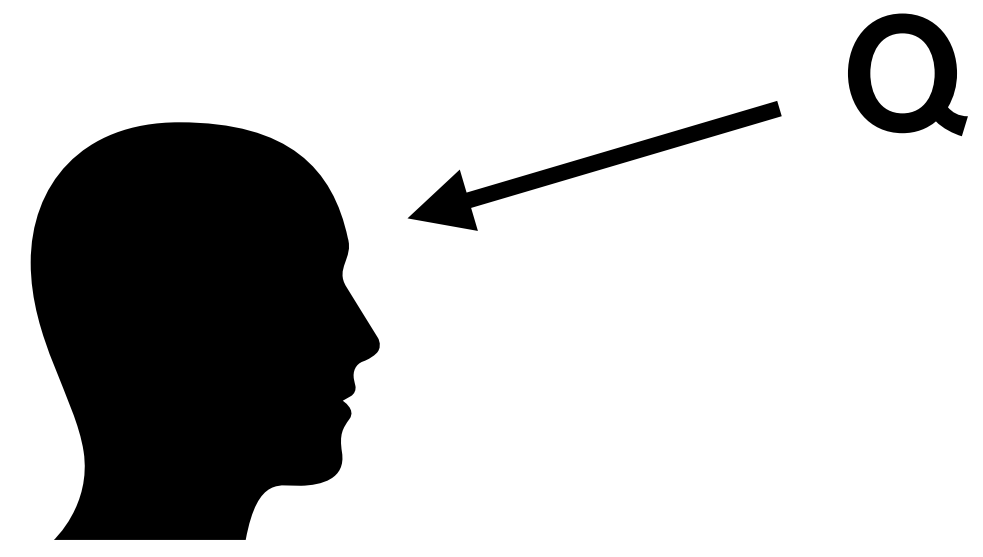
Learning to Plan from Actual and Counterfactual Experiences

Justin Yang and Tobias Gerstenberg



What can we learn from simulating past experience?

Background



Learning is usually thought of as acquiring new information from the **outside world**



But we may also learn things by **simulating alternatives** to what we've already seen



world model



action policy

Such counterfactuals might help us better understand the world and take better actions

hypothesis

People can use counterfactual simulation to learn a better **world model**

Experiment

We manipulated whether participants explicitly engaged in counterfactual simulation. Participants learned a model of the environment through experience and simulation, before being tested in an exam.

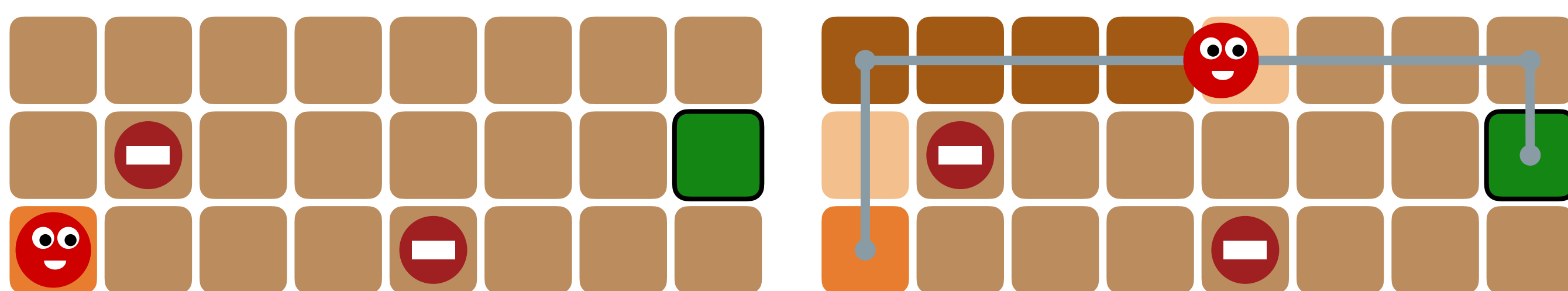
learning phase

experience

Across multiple trials participants navigate a desert while trying to avoid **quicksand**, likely to spawn on **unsafe** tiles

Plan the quickest path to the goal!

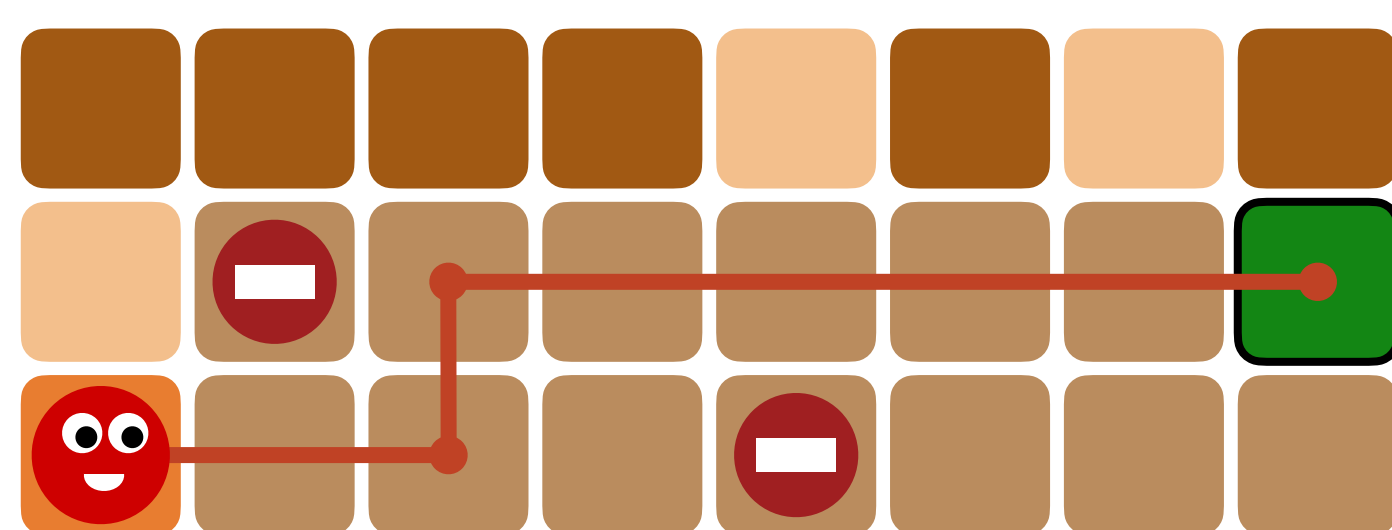
Watch the agent follow your path



counterfactual

Some participants made counterfactual paths after each trial

Given how the path turned out, how would you have reached the goal from here?



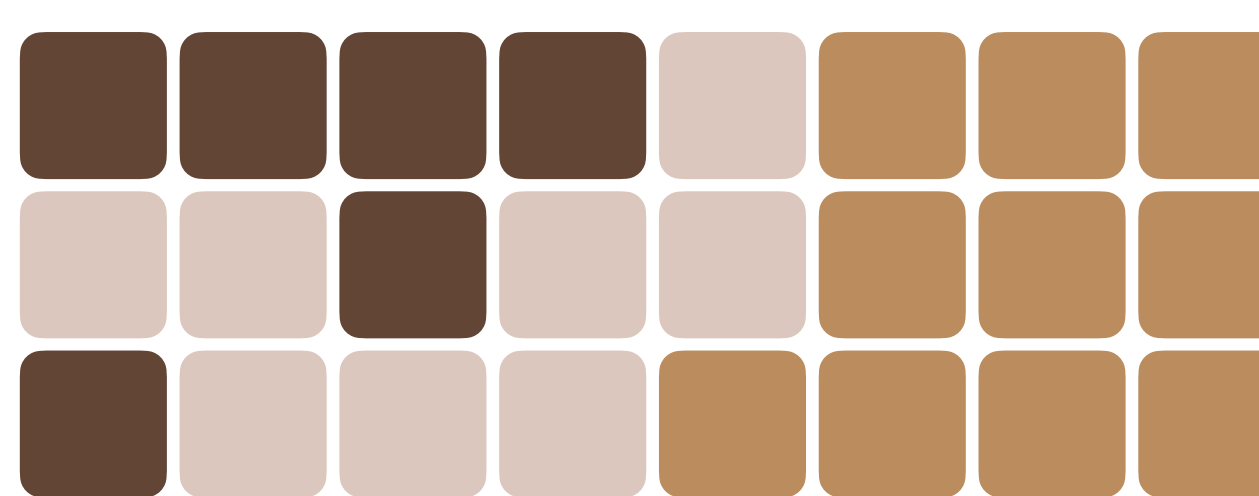
start goal sand quicksand road block

sand is easy to walk through
quicksand slows the agent down

exam trial

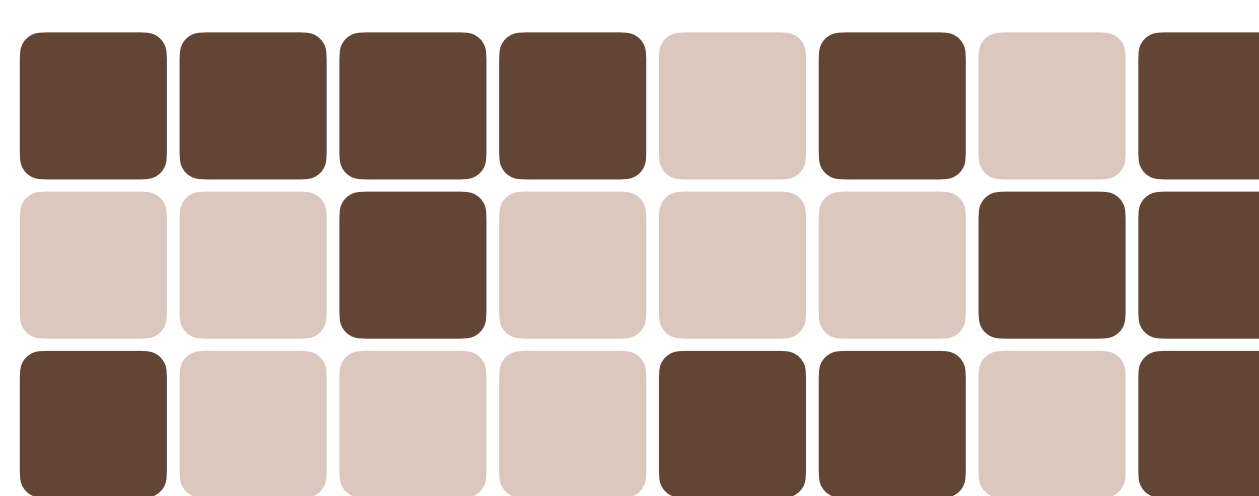
Afterwards they were tested on their understanding of the environment

How safe is each tile?
Click once for safe, and twice for unsafe



Their goal was to recover which tiles were **safe** or **unsafe**

Ground Truth

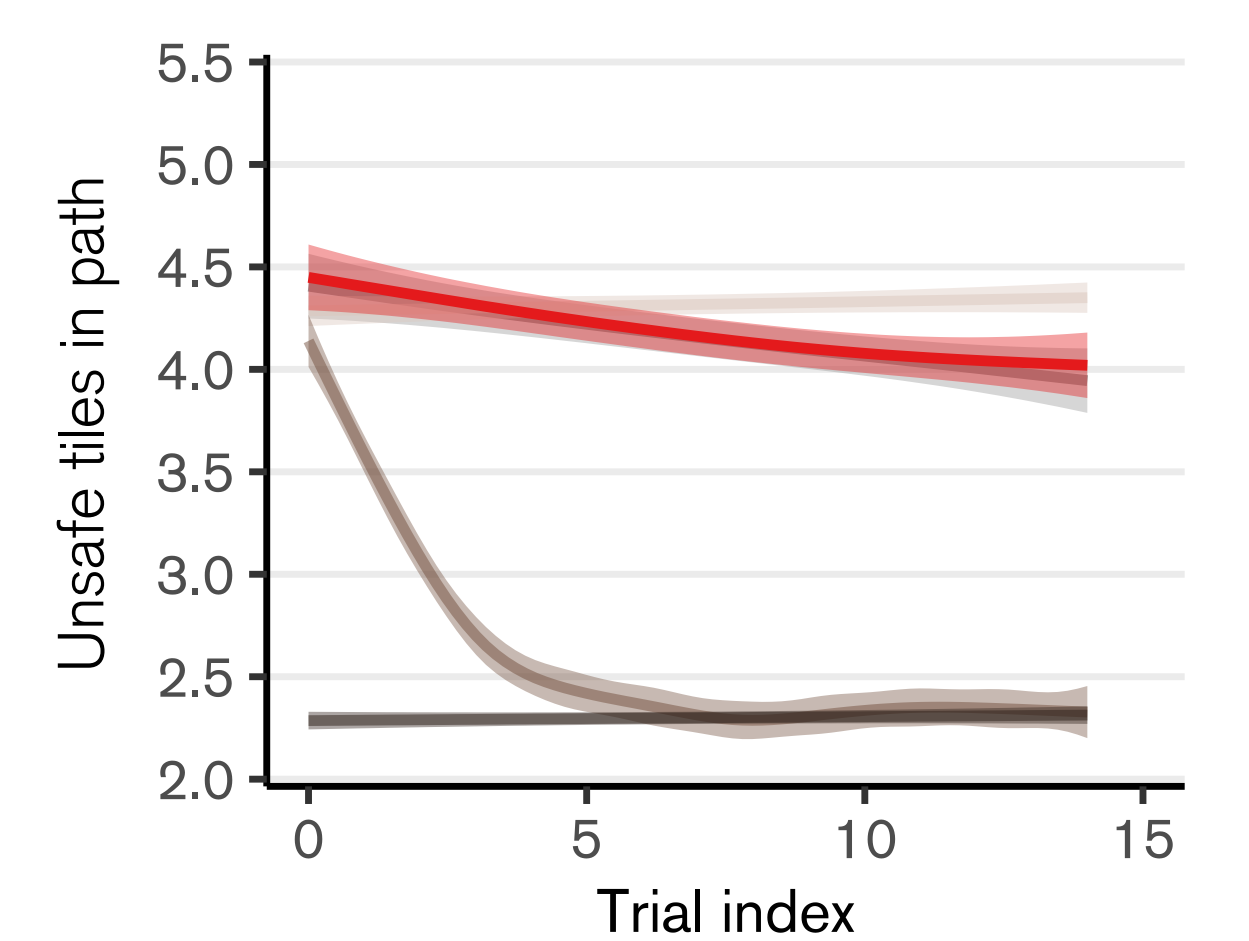


safe unsafe

safe tiles are never quicksand
unsafe tiles are **quicksand** 80% of the time

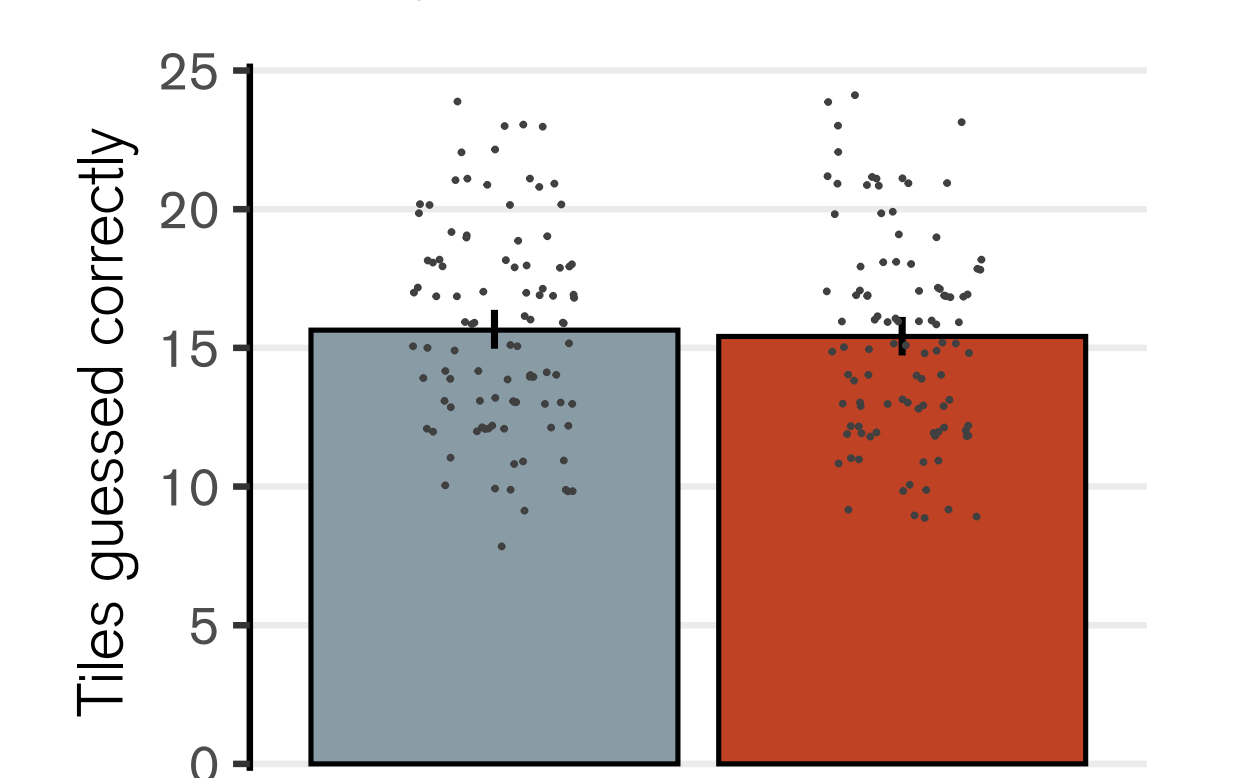
Results

No evidence for improved performance from counterfactual simulation, despite overall improvement from experience



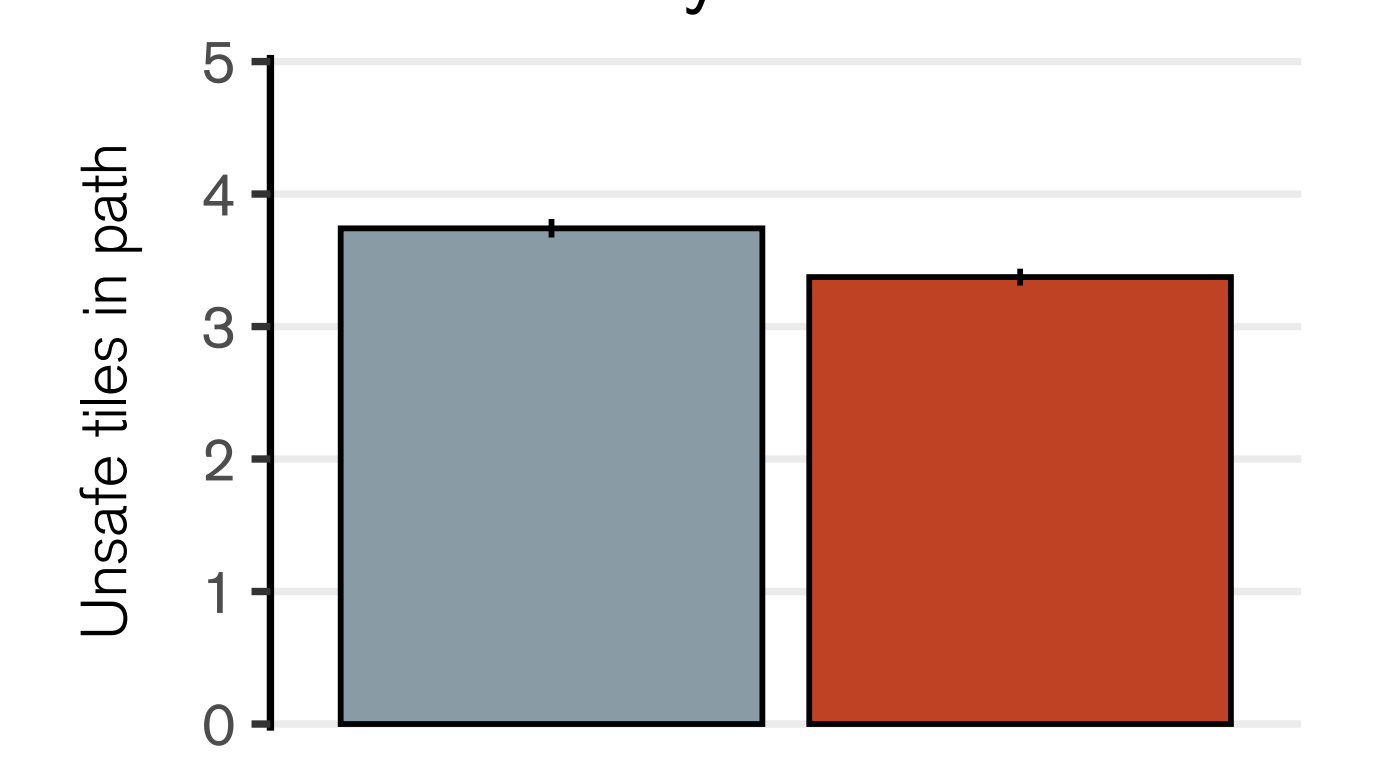
Condition — experience-only — exp + counterfactual
Baseline model — random prior — optimal learner — oracle

No evidence that engaging in counterfactuals yields a better world model



Condition — experience — counterfactual

Counterfactual paths would have contained fewer unsafe tiles than paths actually taken



Path type — experience — counterfactual

Next steps

People did not learn a better world model when simulating counterfactuals, despite generating safer counterfactual paths
Future work: exploring the extent to which people can use counterfactual simulation to learn better **action policies**

