Exploring Exploration: Comparing children with RL agents in unified environments

Eliza Kosoy, Jasmine Collins, David Chan, Jessica Hamrick, Sandy Huang, John Canny and Alison Gopnik

BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

Bridging AI and Cognitive Science (BAICS) - ICLR 2020

That baby video from every AI talk.....

Francis Vachon

Building agents inspired by children can be hard

Environments differ. Children act in the real world, lots of RL research takes place in grid-world settings or 2D Atari games

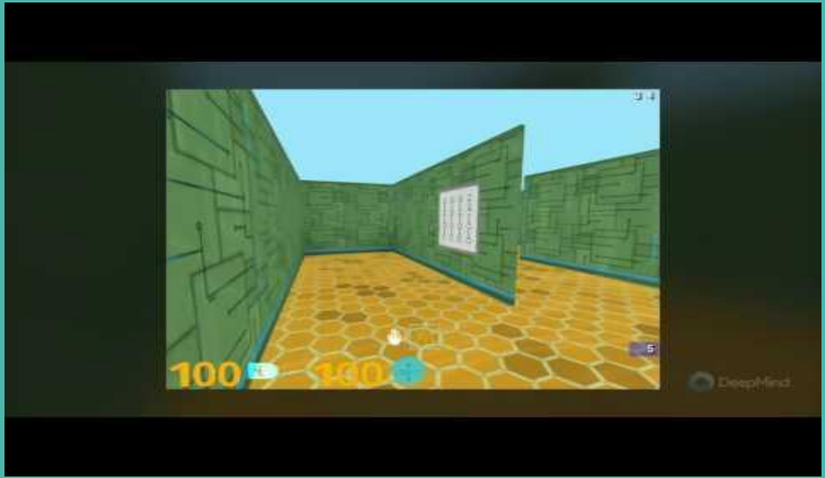
Comparisons are not controlled. Experiments on children engaged in free exploration, majority of research in AI is in goal-seeking domains

Can’t always ‘close the loop’. Ultimately want to learn something about human cognition from AI research

Studying child and agent behavior in the same controlled, rich 3-D environment may alleviate many of these problems!

Contribution:

A framework for directly comparing human children with artificial agents in DeepMind Lab.



Outline

How do children explore?

Comparing children with agents

Preliminary results

Outline

How do children explore?

Comparing children with agents

Preliminary results

What do we know about exploration of children?

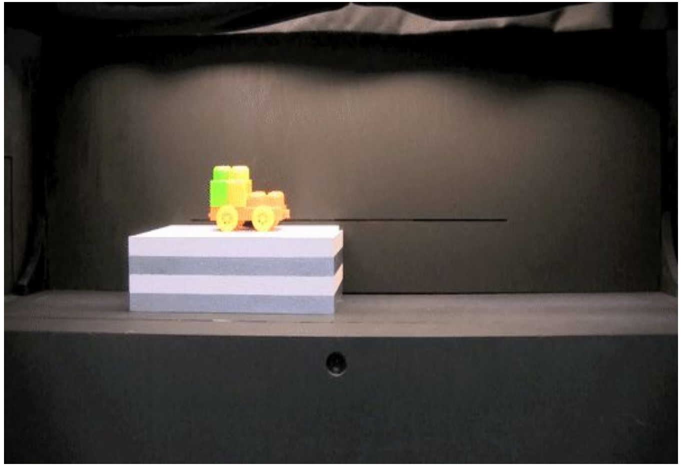
Preschoolers explore more when evidence does not distinguish between multiple hypotheses

(Bonawitz 2007)

What do we know about exploration of children?

Babies exhibit meaningful hypothesis testing behaviors

Example of violation



Baby investigating toy

(Feigenson 2015)

What do we know about exploration of children?

Children are more explorative in explore-exploit tasks

(Sumner 2019, Liquin 2019)

Outline

How do children explore?

Comparing children with agents

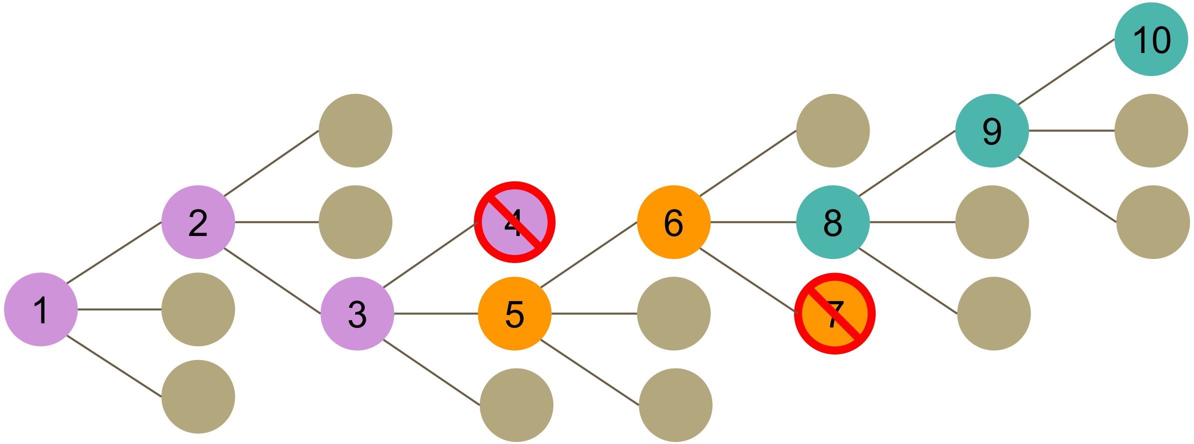
Preliminary results

Exploration in artificial agents

Goal-directed classical search methods

- **Depth-First Search (DFS)**: explore down a path until a dead-end, then backtrack and explore next unexplored path
- **Breadth-First Search (BFS)**: explore new states in the order they were observed
- **A* Search**: search to a goal guided by a heuristic

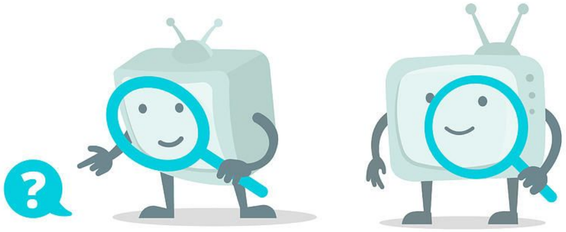
DFS:



Exploration in artificial agents

Reinforcement learning

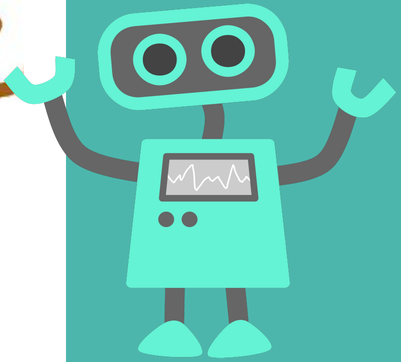
- **Random exploration:** take random actions occasionally (e.g., ε-greedy)
- **Intrinsic motivation (retrospective):** reward bonus for exploring "interesting" regions
  - count-based: bonus for states that are rarely visited
  - curiosity-based: bonus for regions that don't match the learned world model
- **Uncertainty-based methods (prospective):** explore regions with high uncertainty



Difficulties Comparing Children and Agents

Children

- Explore mostly at test time
- Most existing studies use real-world objects
- Given natural language instructions
- Easily switch between free & goal-directed exploration
- Lots of preexisting knowledge, e.g. “wall”, “move forward”, or “prize”
- Fine motor control is difficult

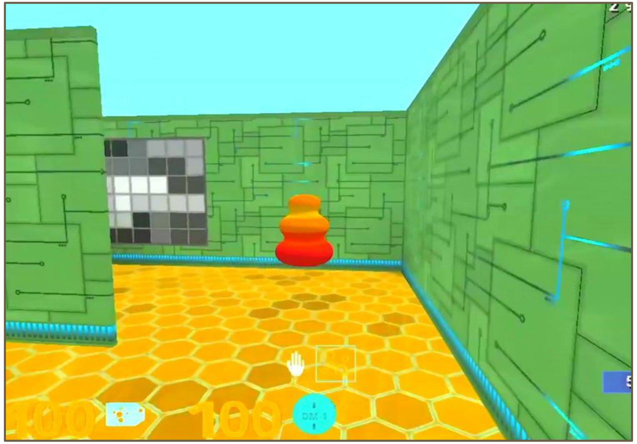


Agents

- Explore mostly during training/learning
- Hard to deploy in the real world
- Don’t understand natural language
- Most methods are for goal-directed tasks
- Almost no preexisting knowledge
- Excel at fine motor control

Leveling the Playing Field with DeepMind Lab

- Forces an exact comparison at test time
- Restricted action set shared by agents and children
- Can be used to generate lots of training data for agents



Outline

How do children explore?

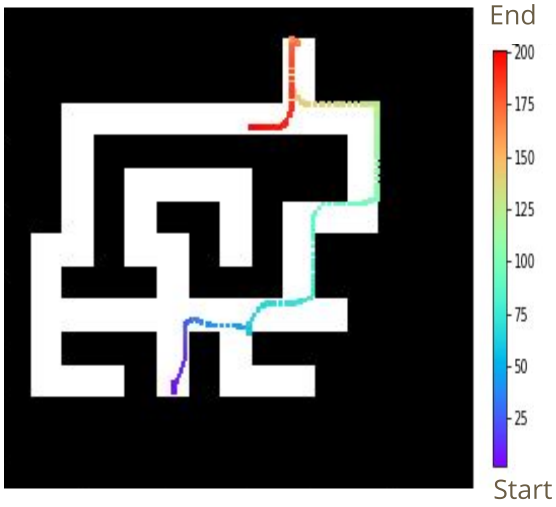
Comparing children with agents

Preliminary results

What does the setup look like?

View of child in task

What the child sees



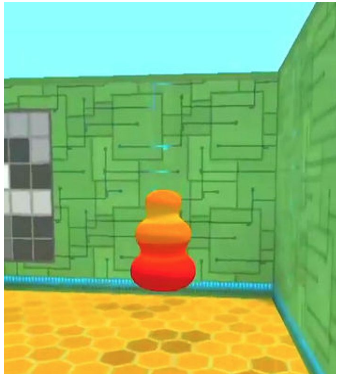
Child's trajectory through the maze

Children's exploration in DeepMind Lab

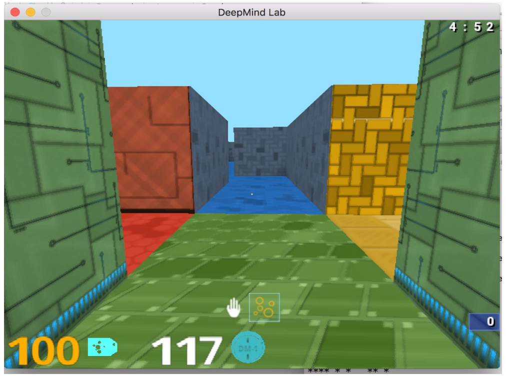
Question:

Are children (4-5 yrs. old, N=30) that are naturally more curious/exploratory in a random maze, more likely to succeed or find the goal in a smaller number of steps?

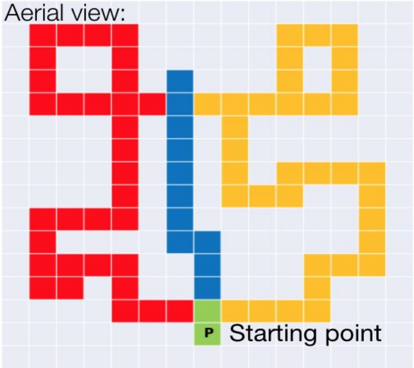
Goal (gummy):



First person P.O.V.

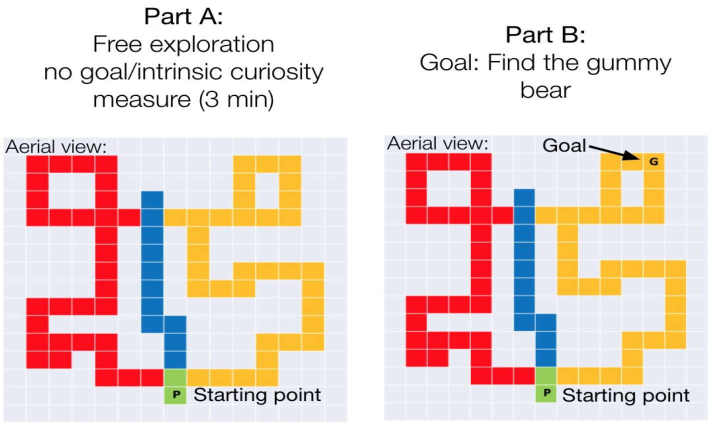


Part A:
Free exploration
no goal/intrinsic curiosity measure (3 min)



Part B:
Goal: Find the gummy bear

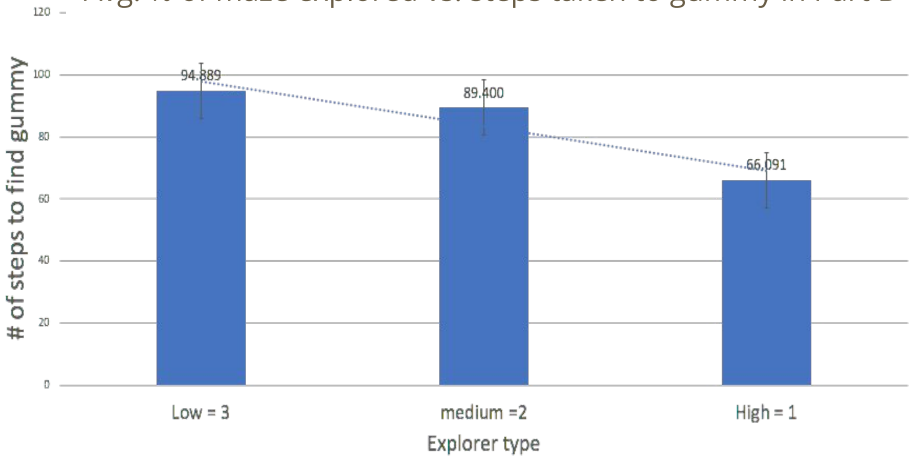
Results: Children's Exploration in DeepMind Lab



Results of K-means on
part A: Free exploration

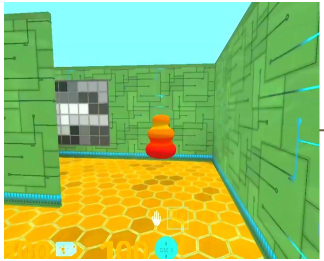
LOW – explored 22% (N=9)
MEDIUM – explored 44% (N=10)
HIGH - explored 71% (N=11)

Avg. % of maze explored vs. steps taken to gummy in Part B



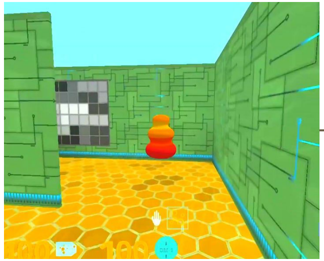
Choice-Consistency: Metric for Direct Comparison

State



Human Action

State



Possible Agent Actions

Is human action ∈ agent actions?

If yes:
Action choice **is** consistent
If no:
Action choice **is NOT** consistent

Measure percentage of states in the trajectory where the action choice is consistent.

Baseline: Do children “act” like Depth First Search?

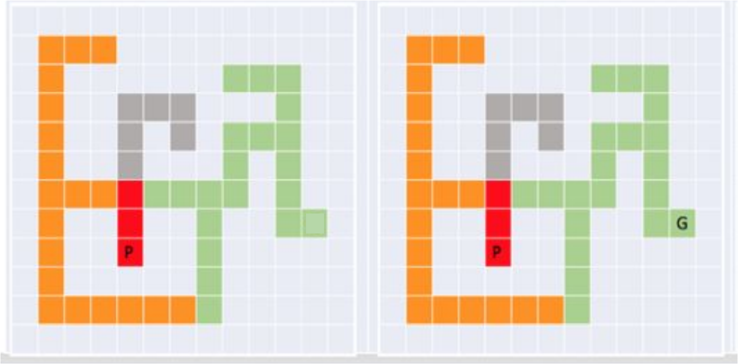
	Choice-Consistency w/ DFS Algorithm (p=0.0073)
No-Goal Exploration	89.61%
Goal-Oriented Exploration	96.04%

Beyond Depth First Search

Children have a lot of prior information about exploration - does this matter? When does this matter? **Can we quantify it?**

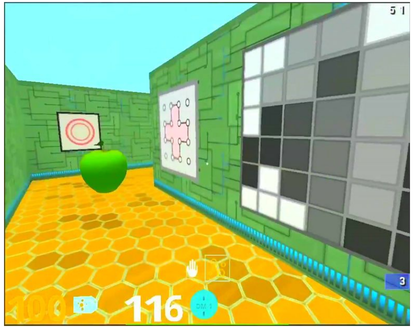
By comparing to other exploration methods from both in classical search (A*, UCB, etc.) and reinforcement learning (Curiosity, Uncertainty, count-based) **we can start to uncover these relationships.**

Free-Exploration Find Goal

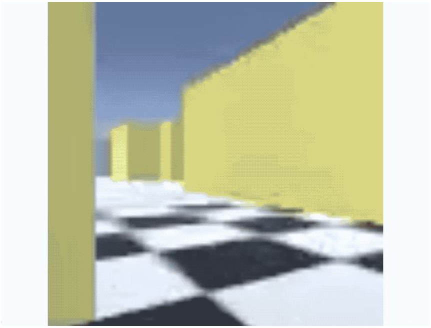


Future work:

- Sparse vs. dense rewards?
- Adding distractors (noisy TV)
- Testing memory / integration of information
- Positive vs. negative rewards



In asking these questions, we will be able to acquire a deeper understanding of the way that children and agents explore novel environments, and how to close the gap between them.



Takeaways:

Exploring Exploration:
Comparing children with RL agents in unified environments

Eliza Kosoy
Jasmine Collins
David Chan
Jessica Hamrick
Sandy Huang
John Canny
Alison Gopnik

With thanks to:
Jeff Liu
Deepak Pathak
Pulkit Agrawal

Bridging AI and Cognitive Science (BAICS) - ICLR 2020

BAIR
BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

CENTER FOR OPEN SCIENCE

DeepMind