Making Sense of Intelligence, both natural and artificial

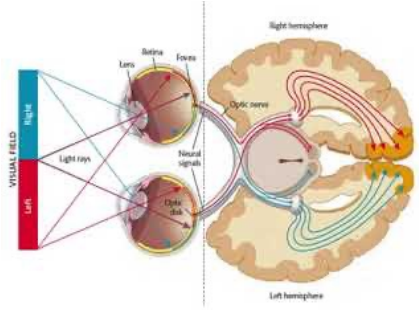
Mark Ho

New York University

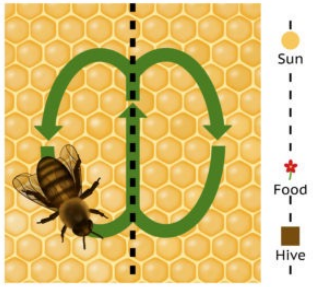
Department of Psychology

Making Sense of Intelligence

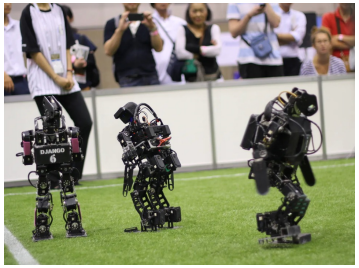
Psychology / Neuroscience

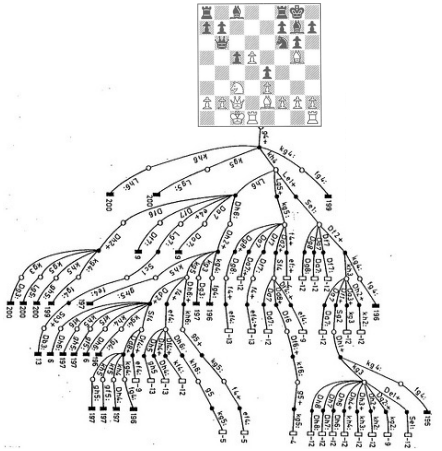


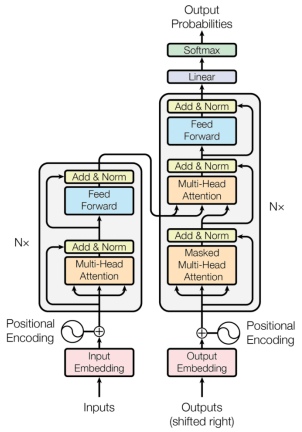
Purple Red Brown
Red Green Blue



Machine Learning / Robotics

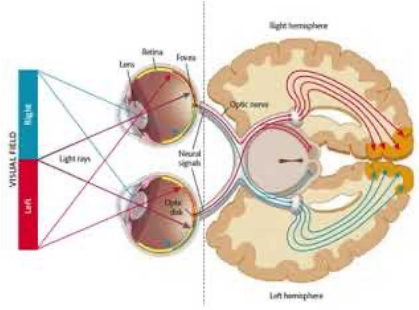




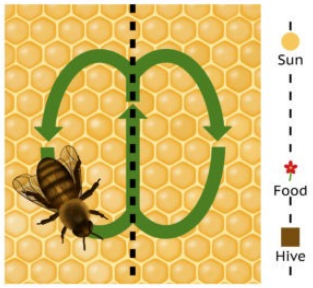


Making Sense of Intelligence

Psychology / Neuroscience

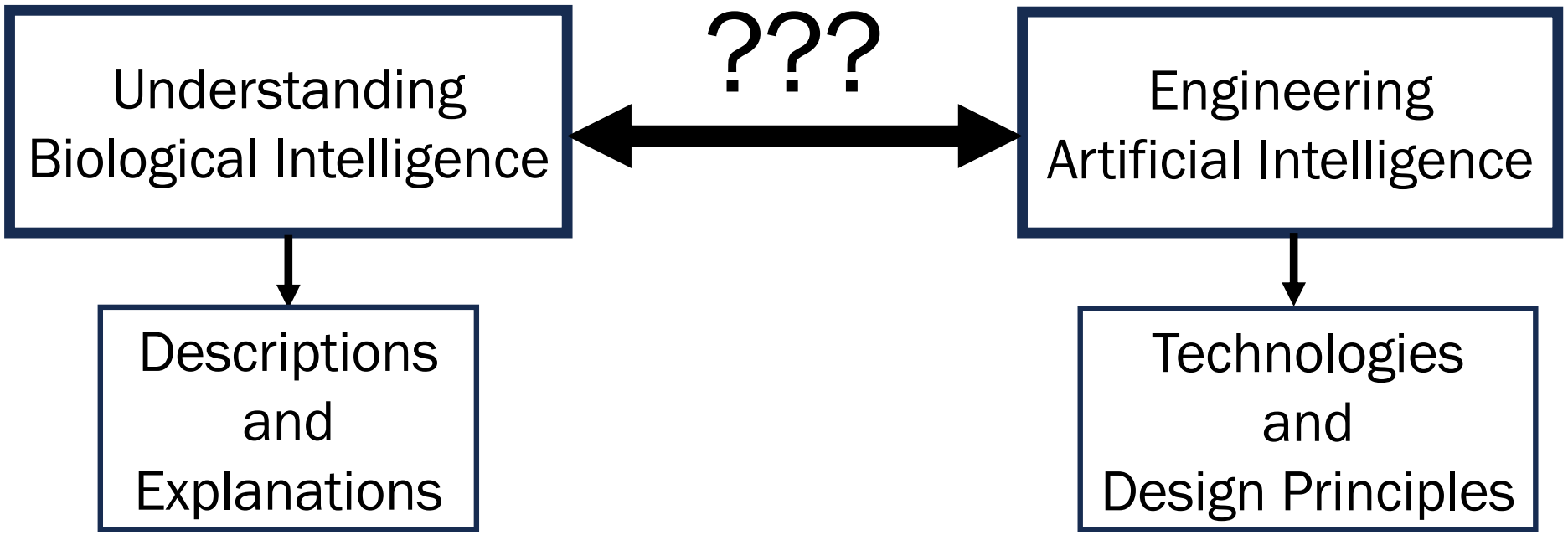


Purple Red Brown
Red Green Blue

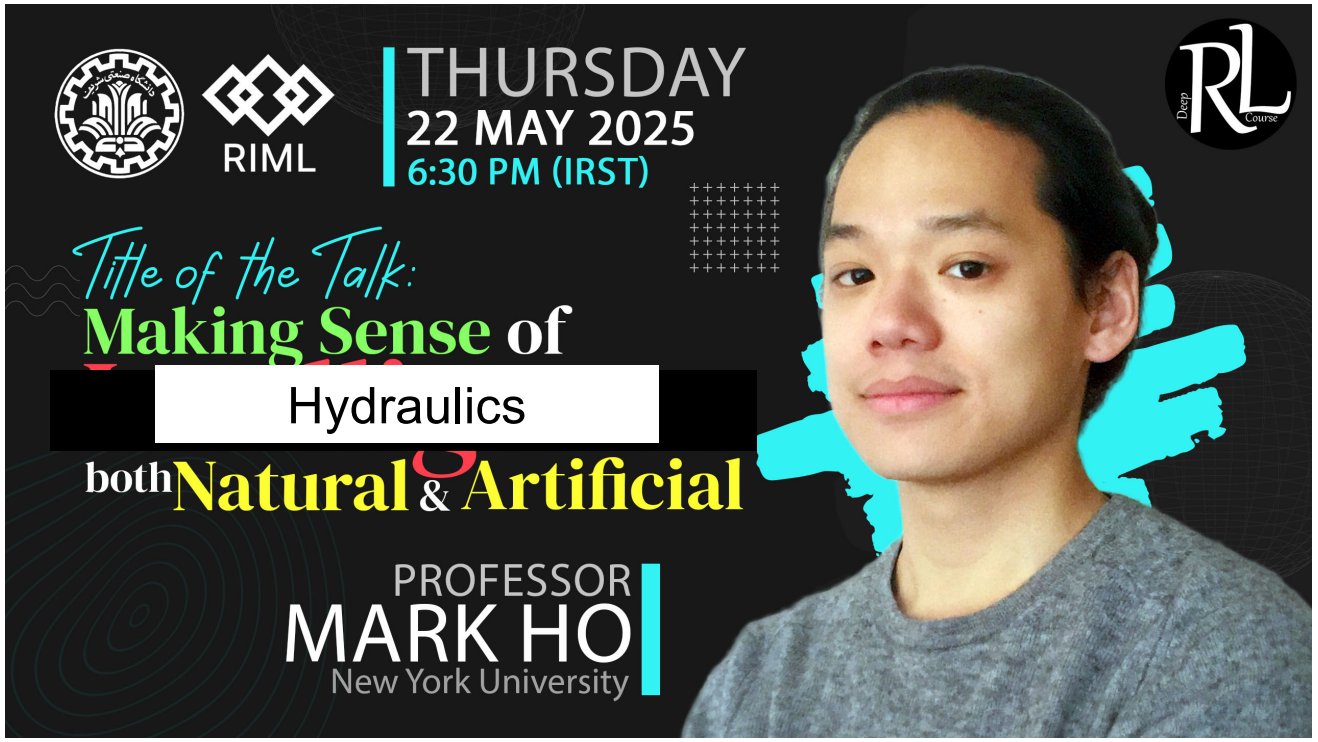


Engineering
Artificial Intelligence

Making Sense of Intelligence

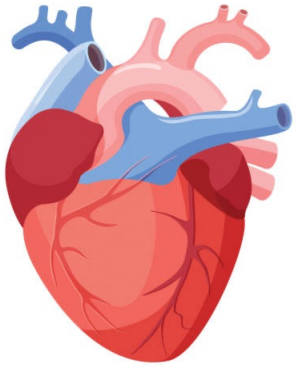


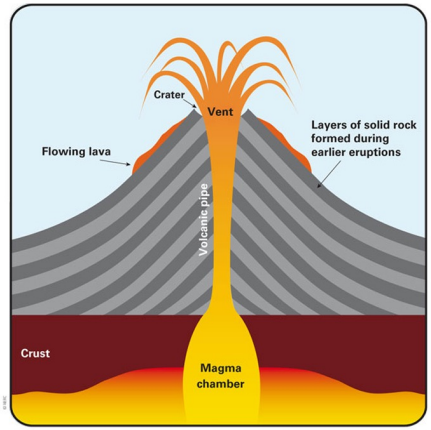
Making Sense of Intelligence



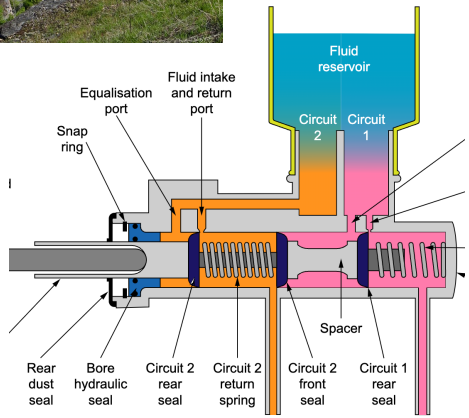
Making Sense of ~~Intelligence~~ ***Hydraulics?***

Natural Hydraulics



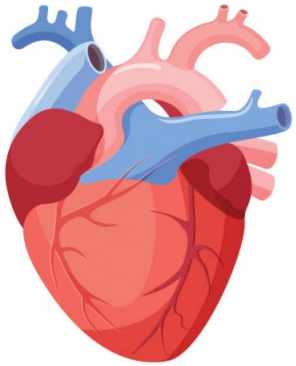


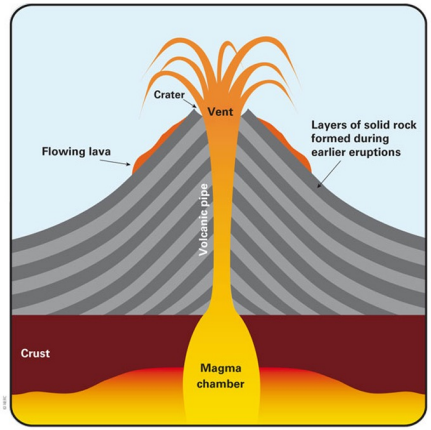
Artificial Hydraulics



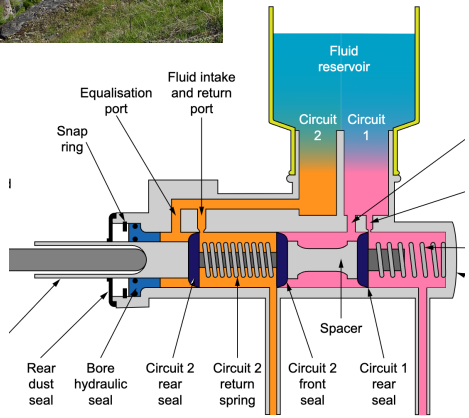
Making Sense of ~~Intelligence~~ ***Hydraulics?***

Natural Hydraulics





Artificial Hydraulics



Making Sense of ~~Intelligence~~ ***Hydraulics?***

Natural Hydraulics

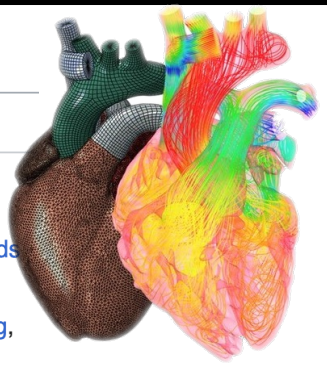
Artificial Hydraulics

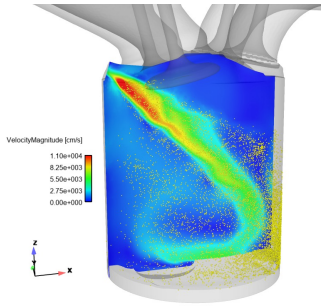
Fluid mechanics

Article Talk

From Wikipedia, the free encyclopedia

Fluid mechanics is the branch of physics concerned with the mechanics of fluids (liquids gases, and plasmas) and the forces on them.[1]:3 It has applications in a wide range of disciplines, including mechanical, aerospace, civil, chemical and biomedical engineering, geophysics, oceanography, meteorology, astrophysics, and biology.





$$\tau = -\mu \frac{\mathrm{d}u}{\mathrm{d}n} \qquad \tau_{ij} = \mu \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} - \frac{2}{3}\delta_{ij} \nabla \cdot \mathbf{v} \right)$$

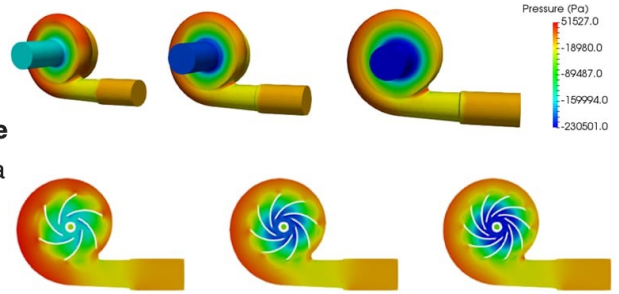
τ is the shear stress exerted by the fluid ("drag"),

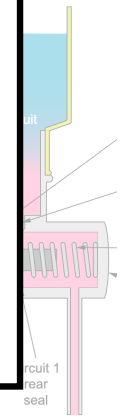
μ is the fluid viscosity—a constant of proportionalit

$\frac{\mathrm{d}u}{\mathrm{d}n}$ is the velocity gradient perpendicular to the di

The **Navier–Stokes e** the force balance at a are[13][14][15][16]

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u}$$





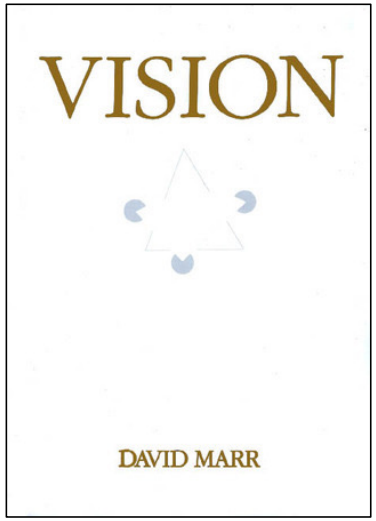
Two ways to make sense of intelligence: Rationality and computation

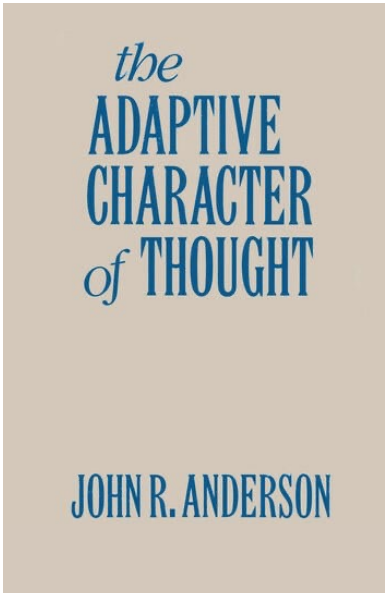
Rationales

- What problem is a system solving?
- *Functional* explanations
- Problem statements and utility functions

Computations

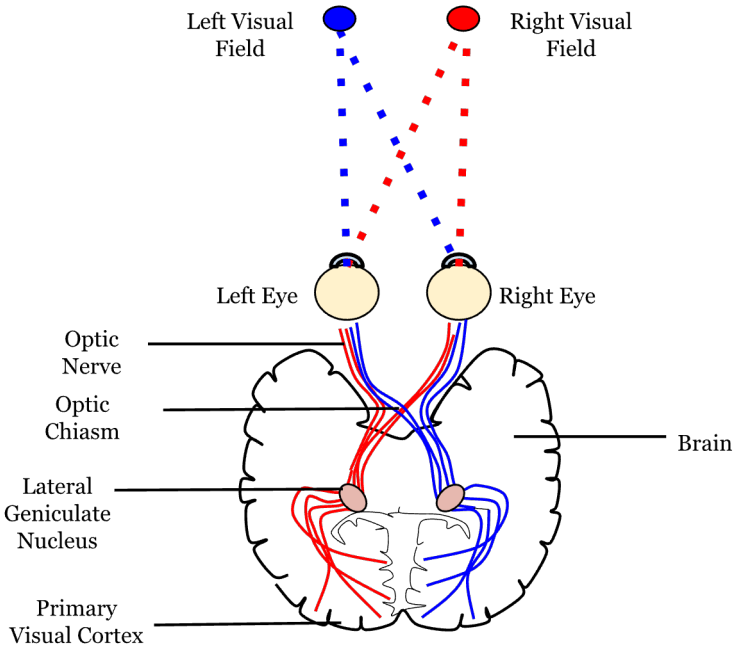
- How does a system find a solution? What does it use to find a solution?
- *Mechanistic* explanations
- Algorithms and search strategies

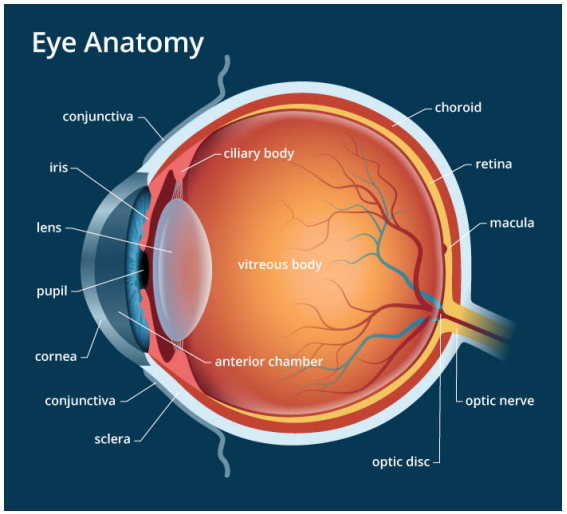




Also see Tinbergen, 1963 for a similar perspective in ethology

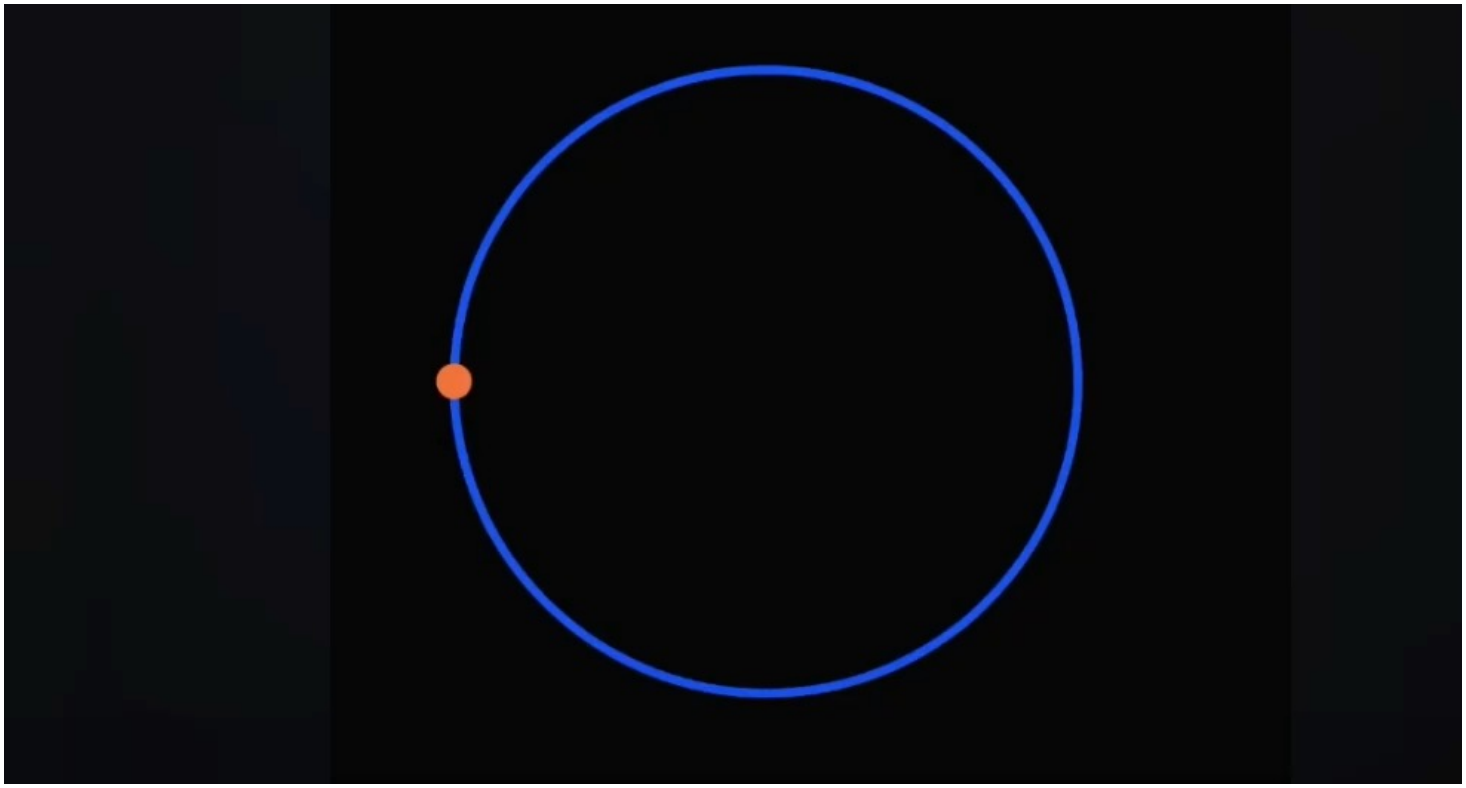
Mechanistic explanations of vision



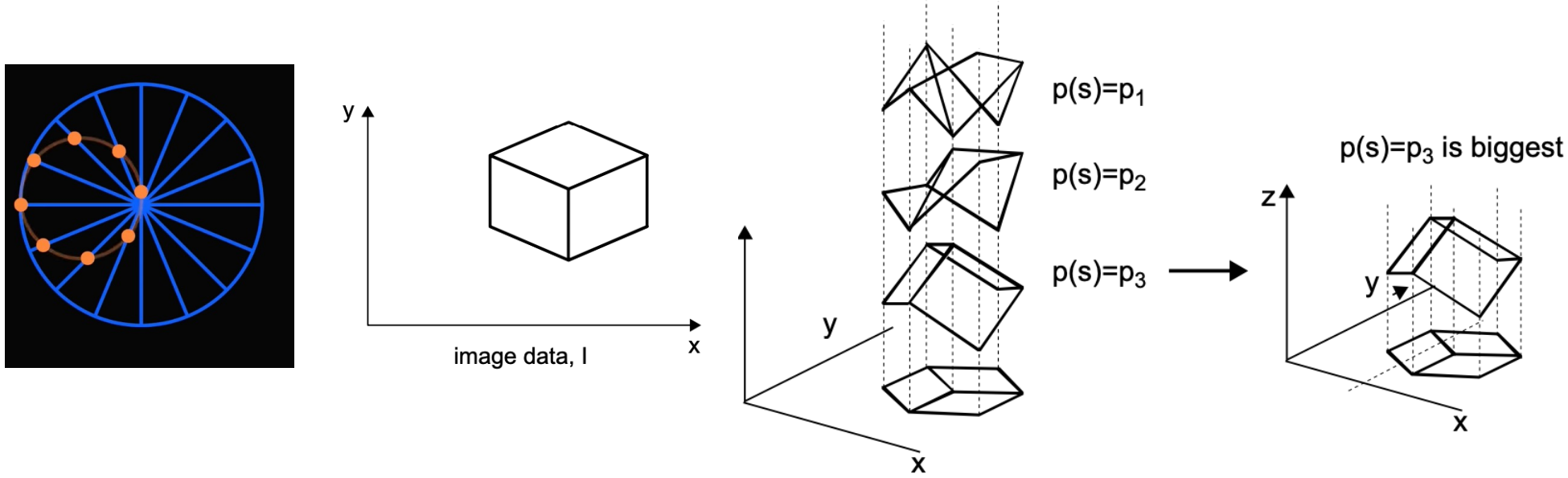


Mechanism of vision: light hits retina, causing neural firing, etc.

Functional explanations of vision



Functional explanations of vision



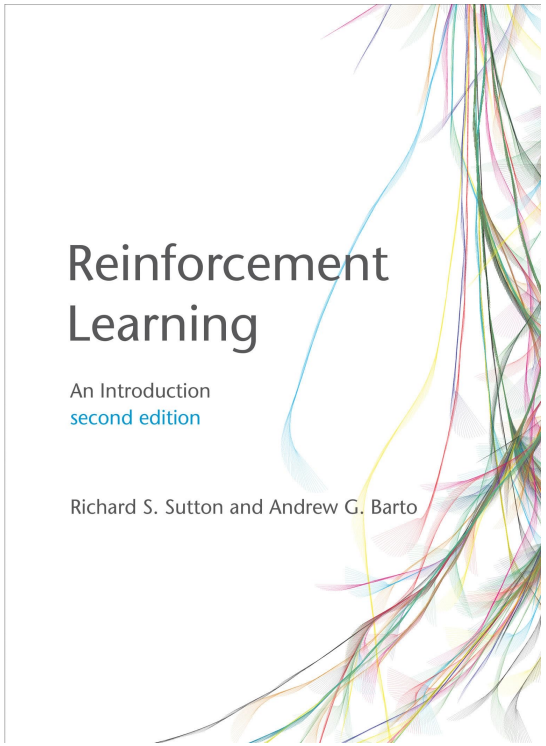
Kersten & Yuille, 2003

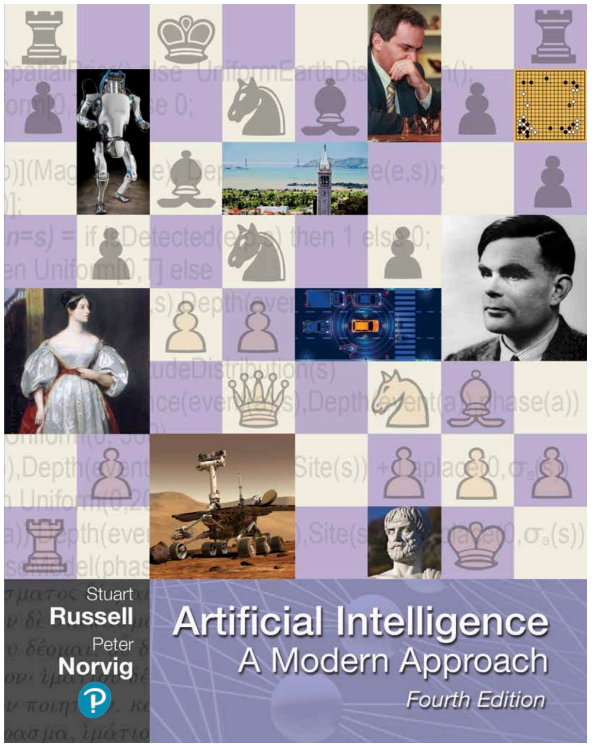
Function of vision: Identifying objects in environment

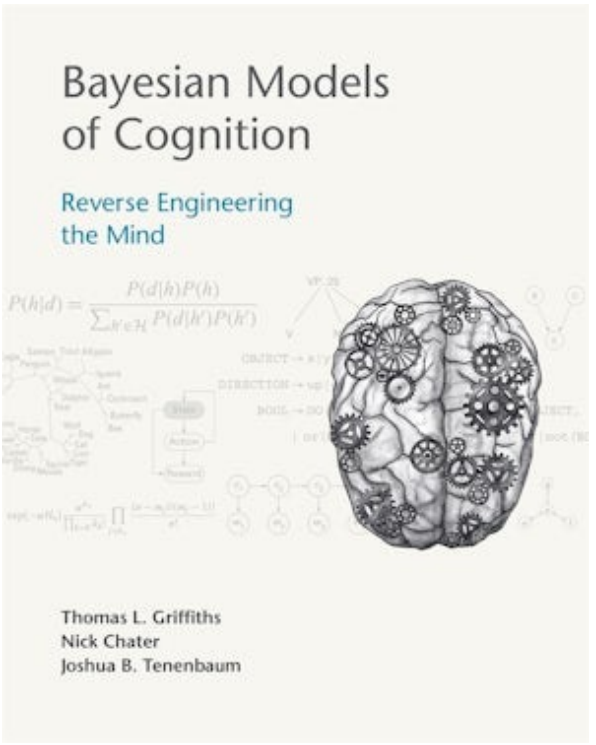
Functional accounts of higher-level and social cognition

- Natural and artificial intelligence have very different **mechanisms,** but they can share similar **functions**
  - Different mechanisms: neurons versus silicon
  - Same function: detecting objects
- Function(s) of perception are fairly clear-cut
  - Even so, computational theory is very useful!
- Function(s) of higher-level and social cognition are less obvious
  - Goals, intentionality, agency, learning, curiosity, adaptation, intelligence, habits, communication, pedagogy, norms, cooperation, morality …
  - Computational theory is extremely useful, maybe unavoidable

Functional accounts of higher-level and social cognition

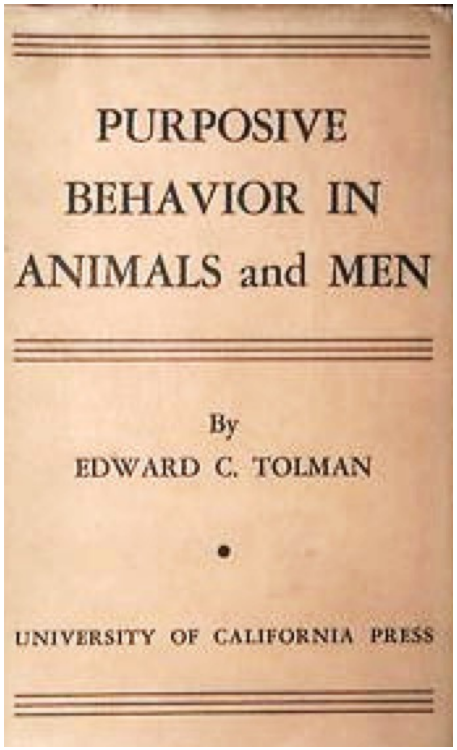




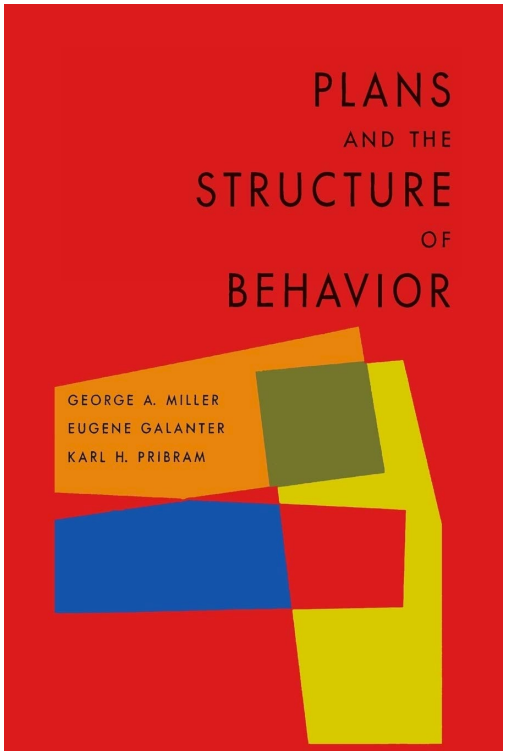


Human Problem Solving

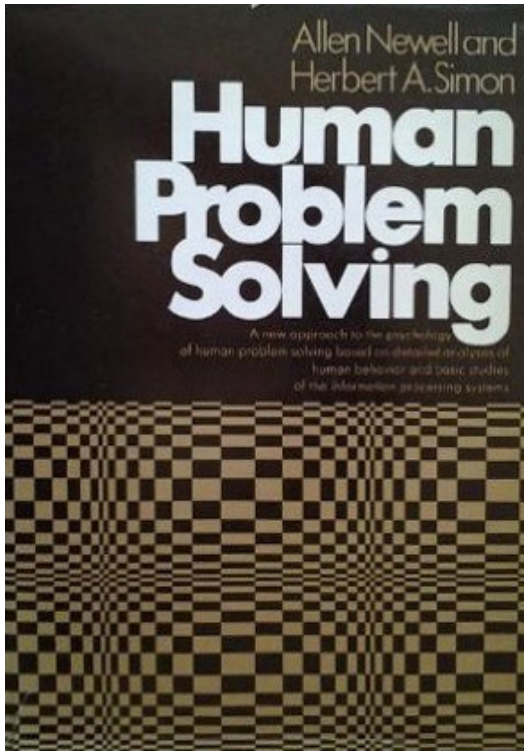
Human Problem Solving



Tolman (1932)



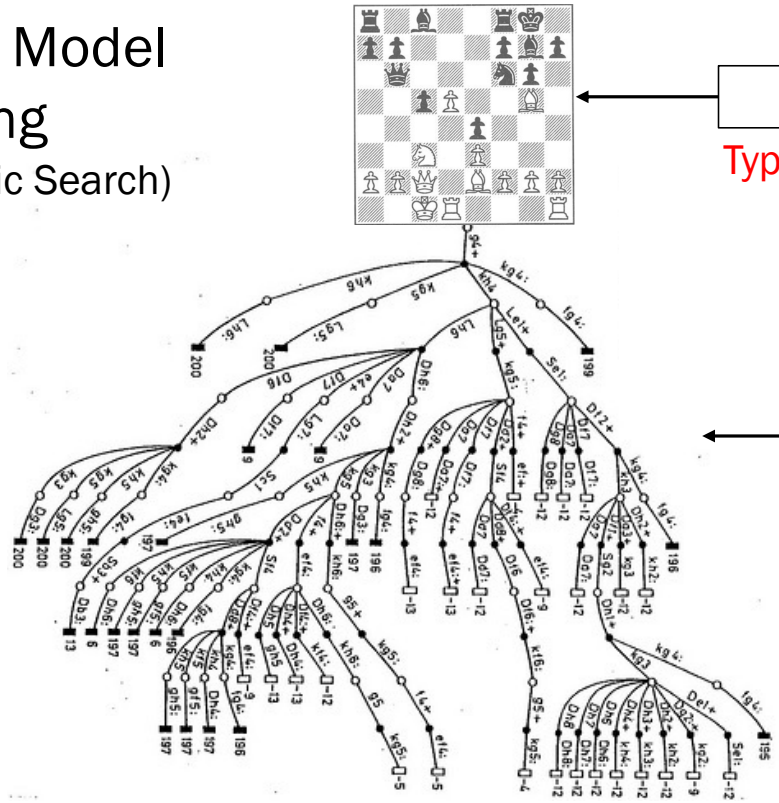
Miller, Galanter & Pribram (1960)



Newell & Simon (1972)

Planning

Standard Model
of Planning
(e.g., Heuristic Search)



Represent Task

Typically, this is ***fixed***

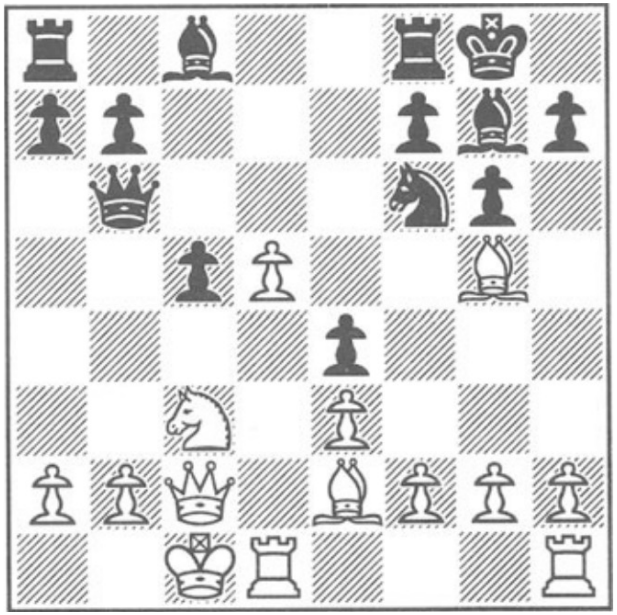
Insight? Analogy?
Expertise?

Search for
Actions

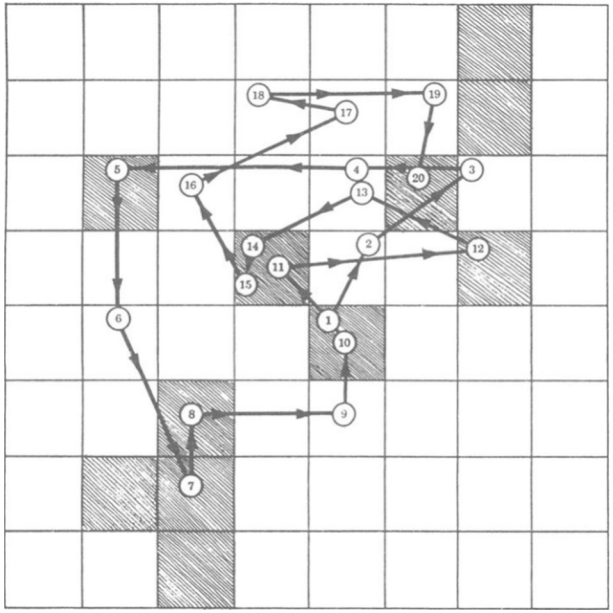
Newell & Simon, 1976; Puterman, 1994; Sutton & Barto, 1998; 2018

Planning

Full Representation



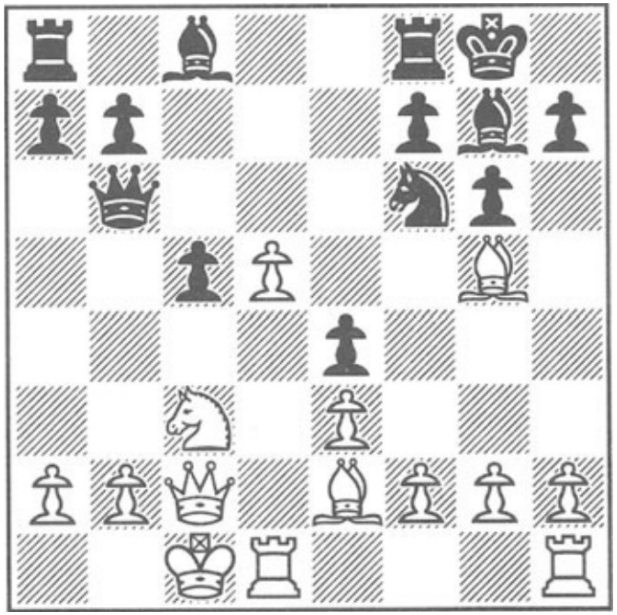
Expert Eye-Movements



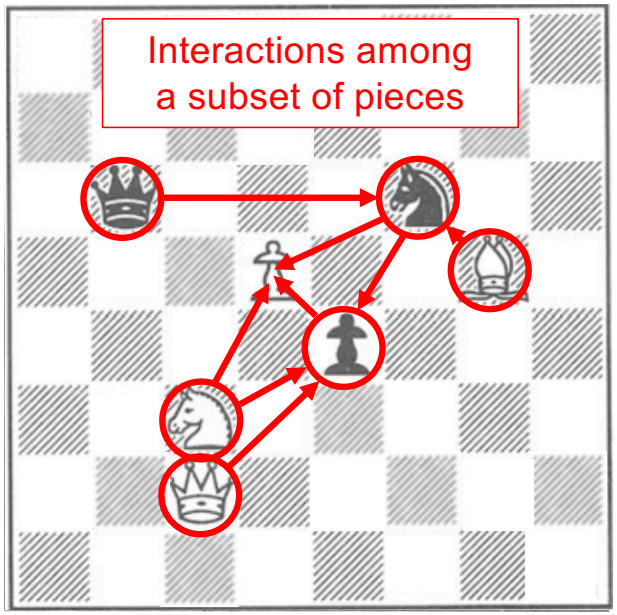
Tichomirov & Poznyanskaya (1966)

Planning

Full Representation

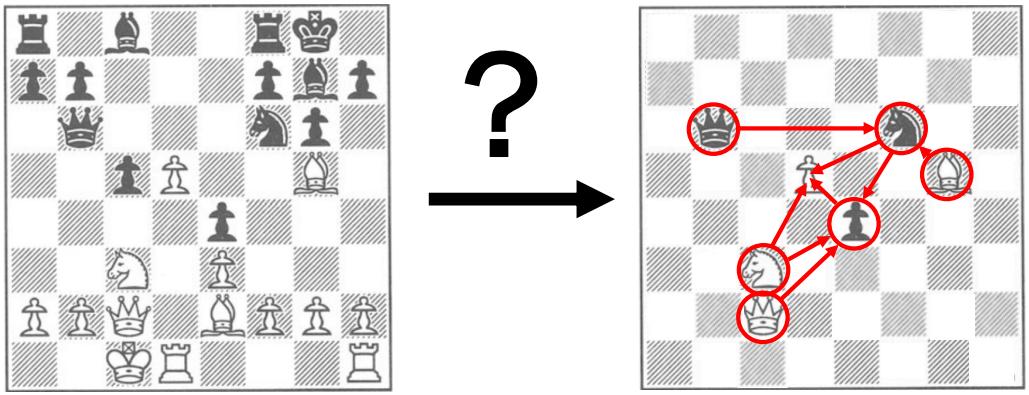


Expert Eye-Movements



Tichomirov & Poznyanskaya (1966)

Planning



How do humans represent tasks?

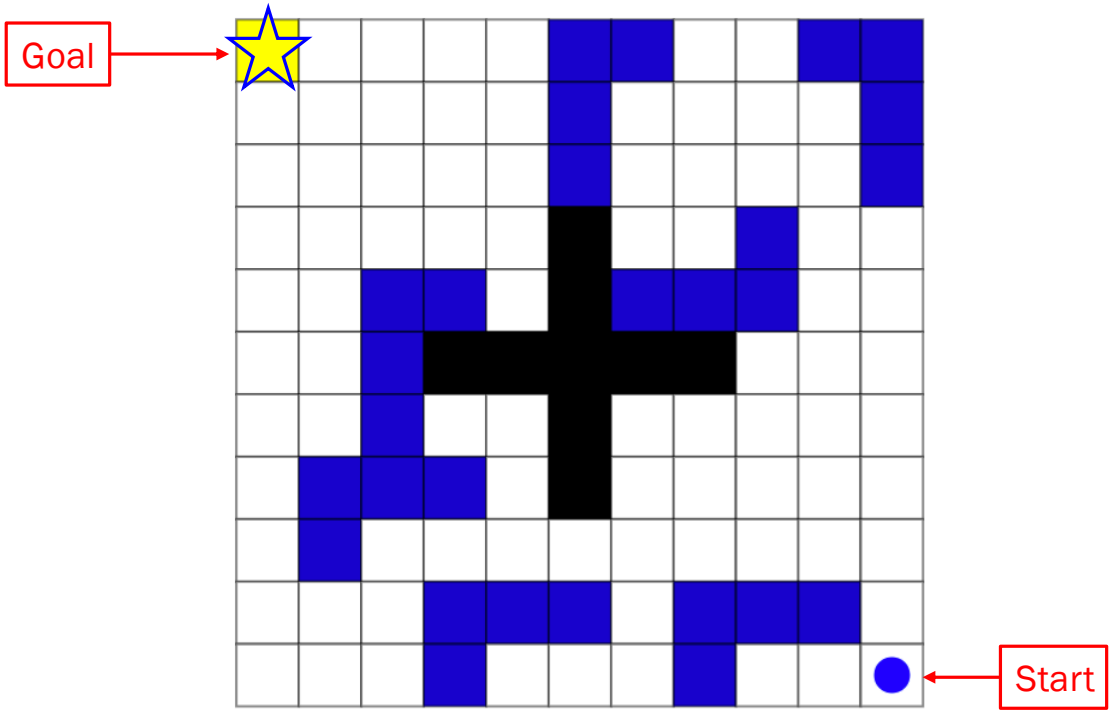
(Newell & Simon, 1972; Ohlsson, 2012)

People construct ***value-guided construals***

Ho, Abel, Correa, Littman, Cohen & Griffiths (2022). *Nature.*

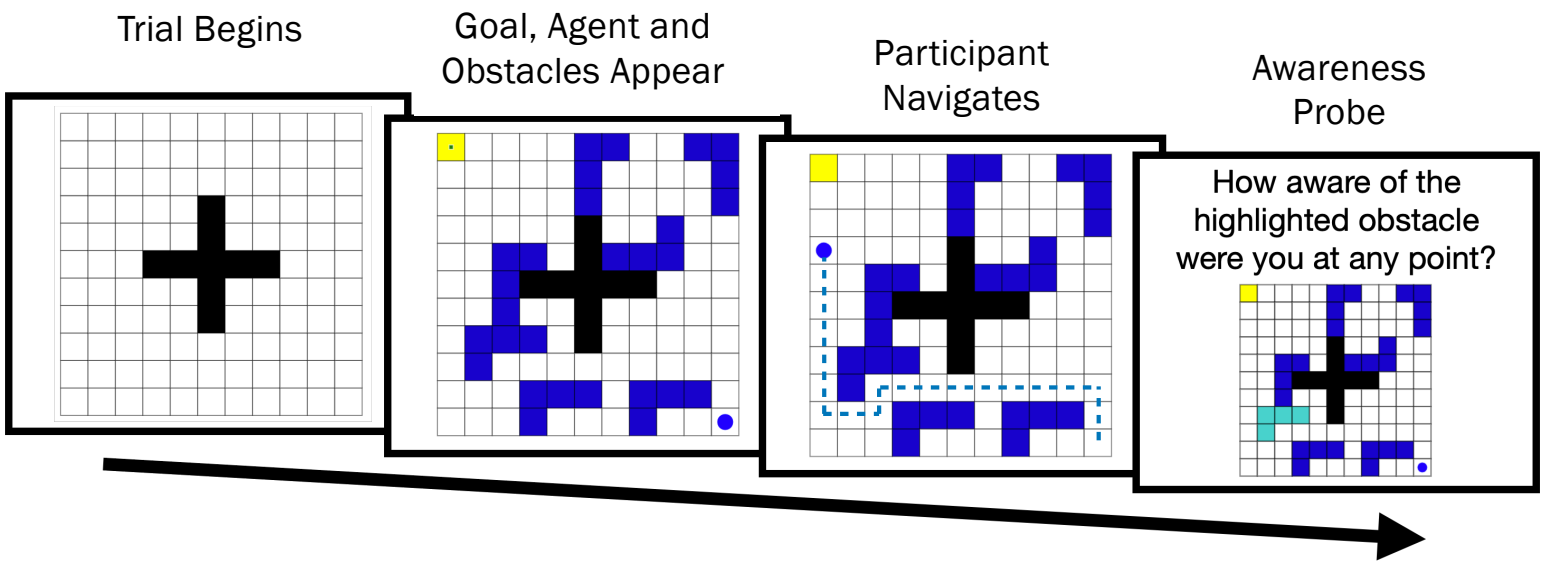
Maze Navigation Task

Mazes constructed out of fixed walls (black)
and changing obstacles (blue)

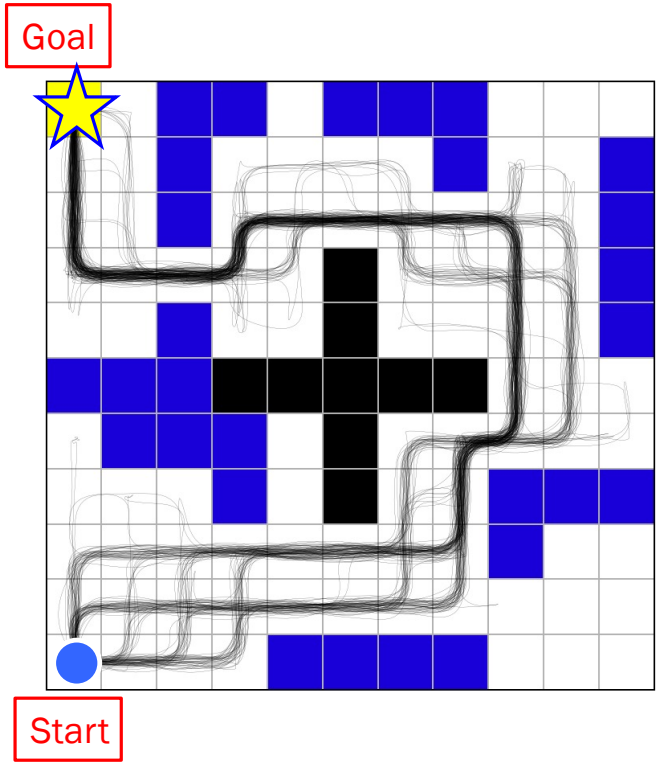


Ho, Abel, Correa, Littman, Cohen & Griffiths (2022). *Nature.*

Maze Navigation Task

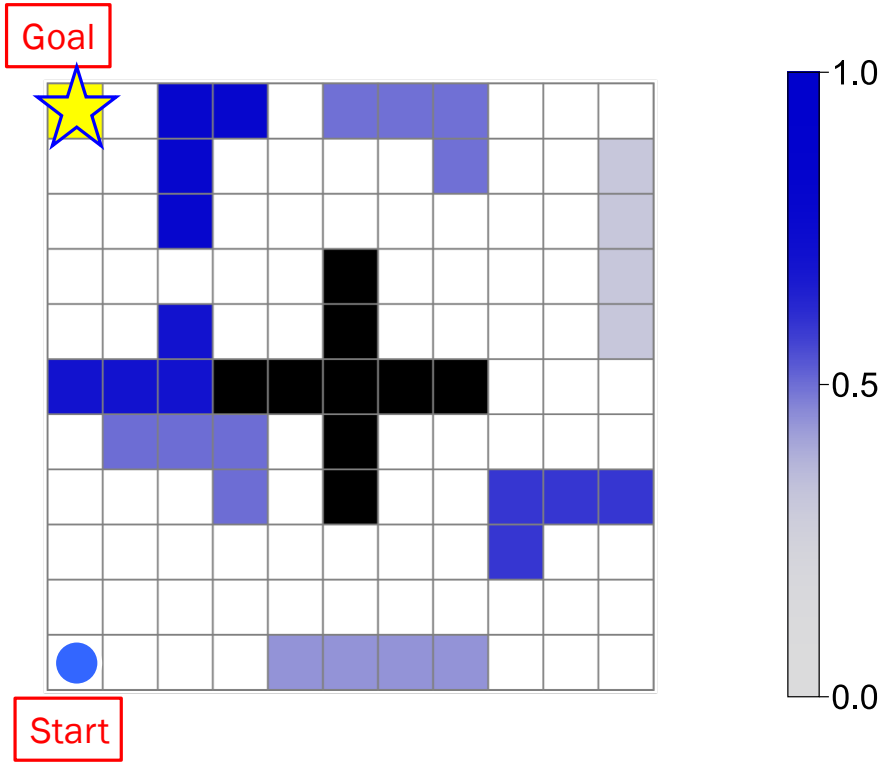


People plan and then act



N = 161

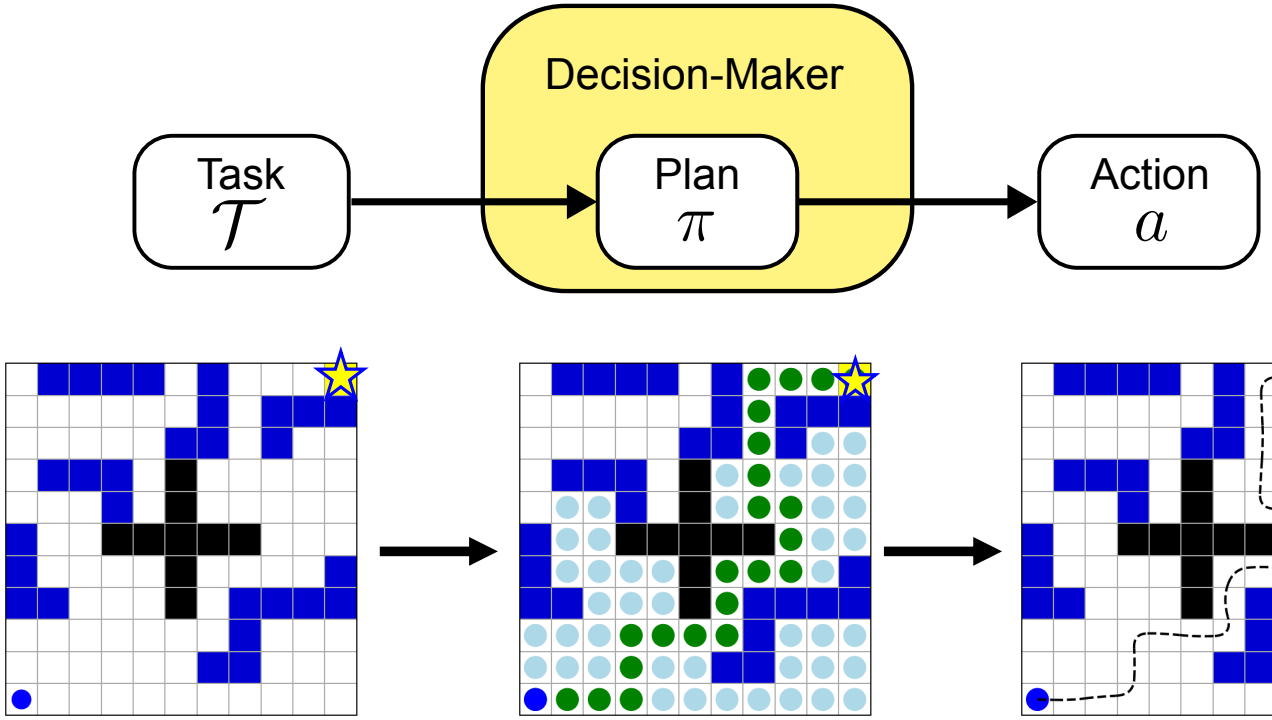
Mean Awareness Responses



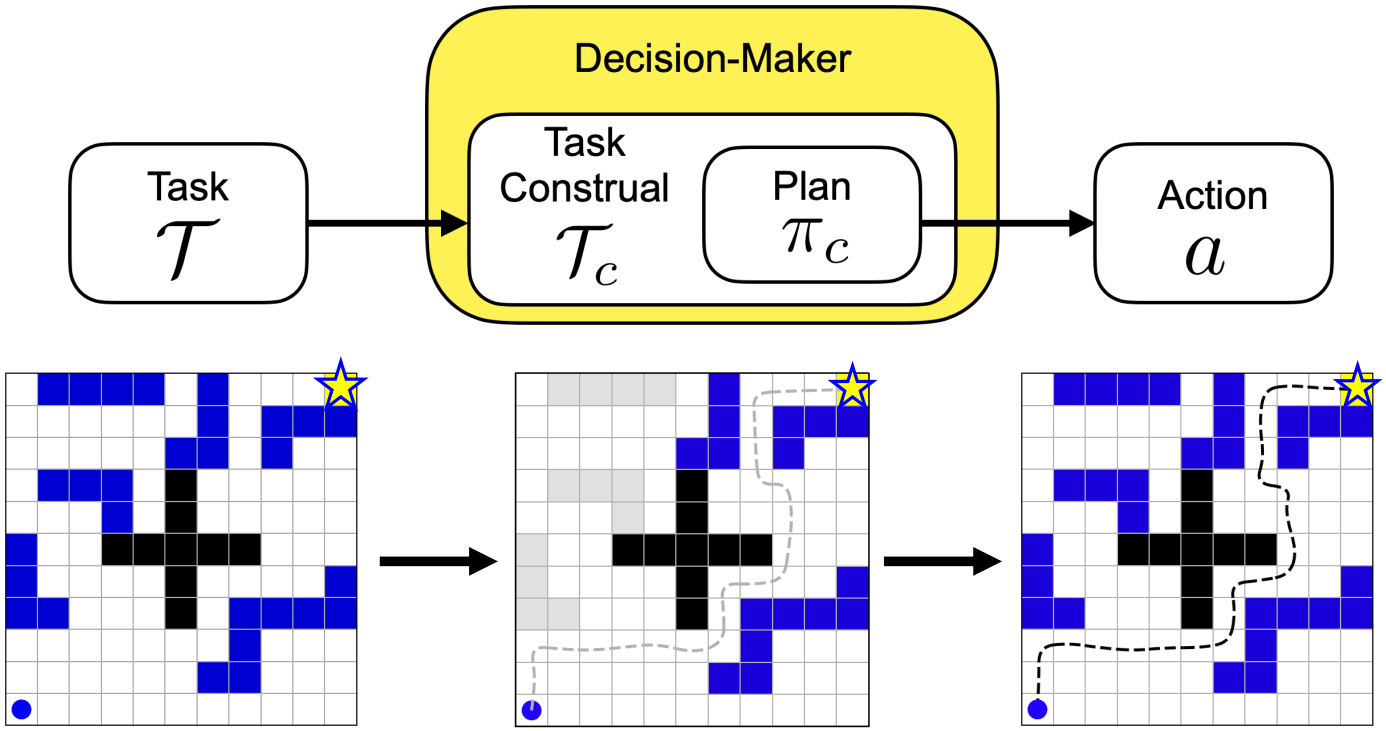
N = 161

Standard Model of Planning

(e.g., Heuristic Search)



Value-Guided Construal



Value-Guided Construal

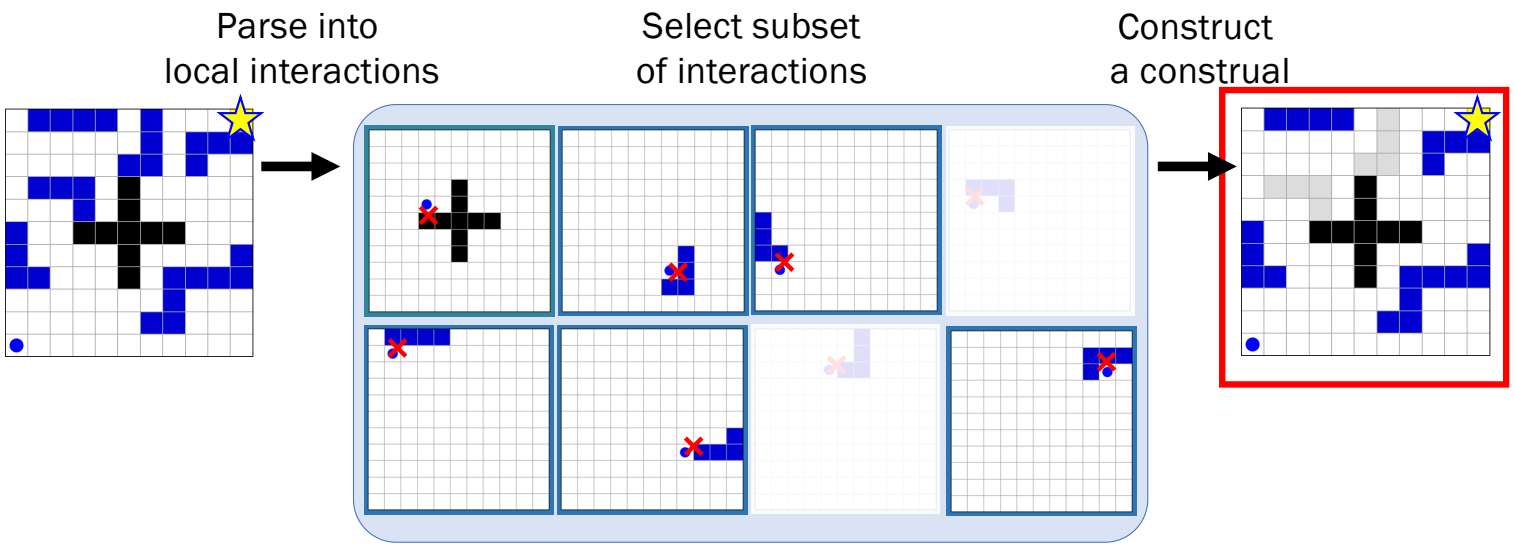
Value of Construal

Behavioral Utility

Cognitive Cost

$$V(c) = U(\pi_c) - C(c)$$

Construal

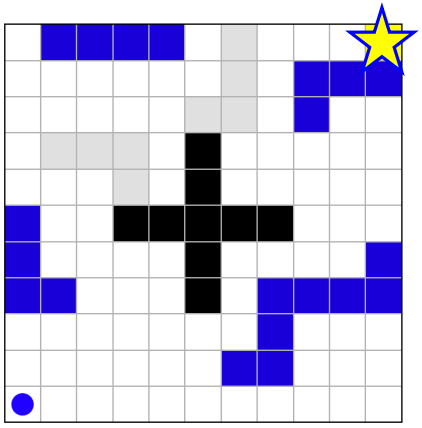


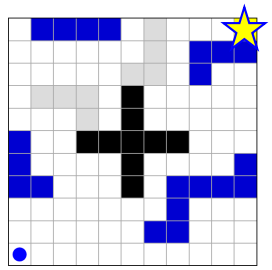
Value-Guided Construal

Value of Construal Behavioral Utility Cognitive Cost

$$V(c) = U(\pi_c) - C(c)$$

Construal



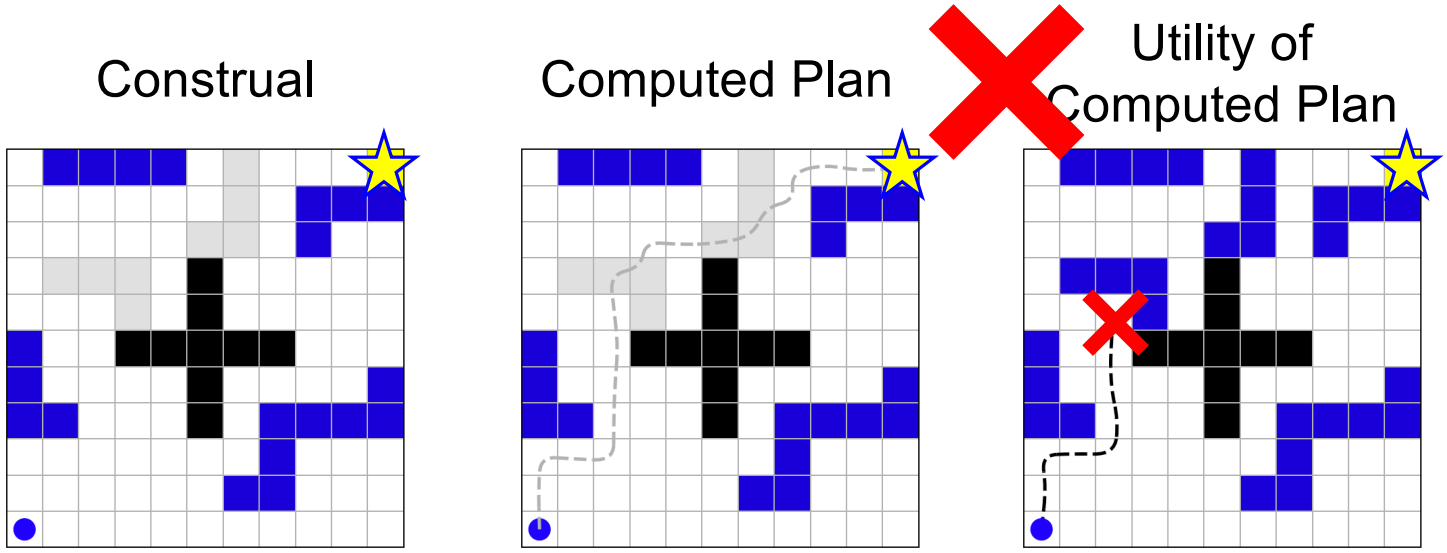


Value-Guided Construal

Value of Construal — Behavioral Utility — Cognitive Cost

$$V(c) = U(\pi_c) - C(c)$$

↑ Plan Computed with Construal

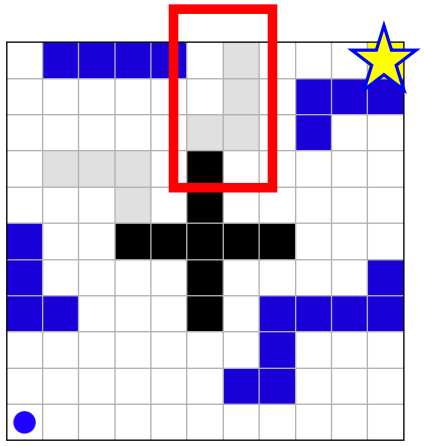


Value-Guided Construal

Value of Construal Behavioral Utility Cognitive Cost

$$V(c) = U(\pi_c) - C(c)$$

Construal

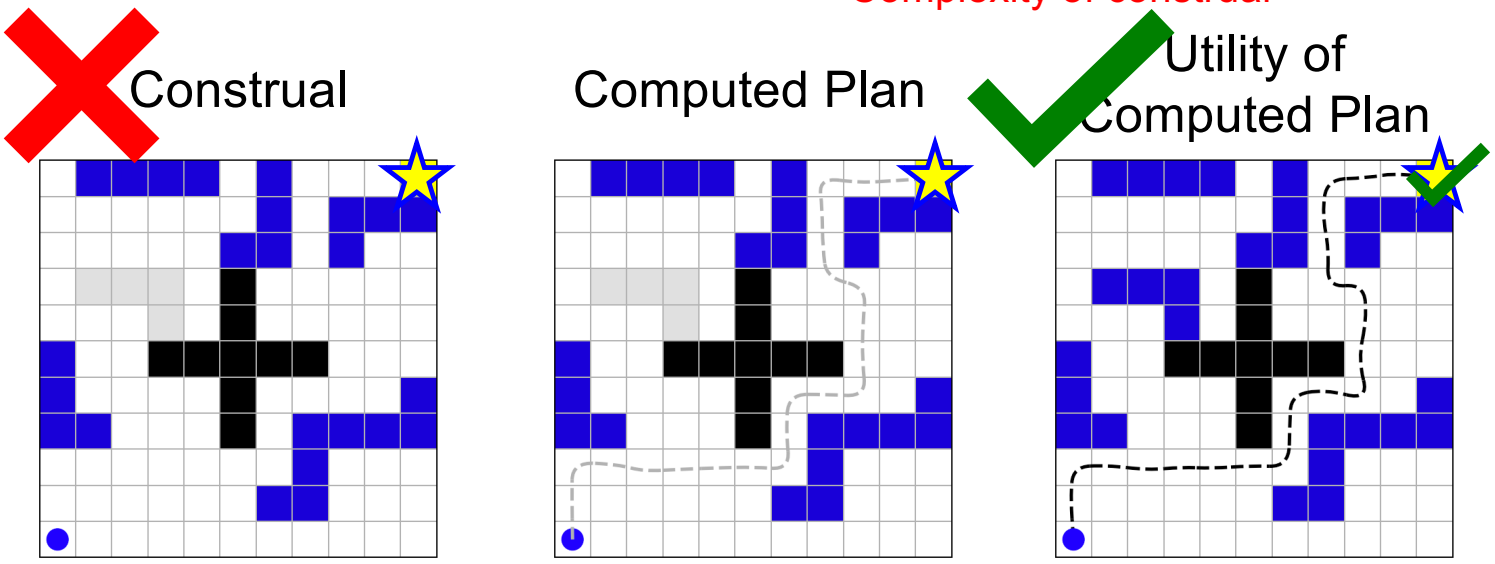


Value-Guided Construal

Value of Construal — Behavioral Utility — Cognitive Cost

$$V(c) = U(\pi_c) - C(c)$$

Complexity of construal

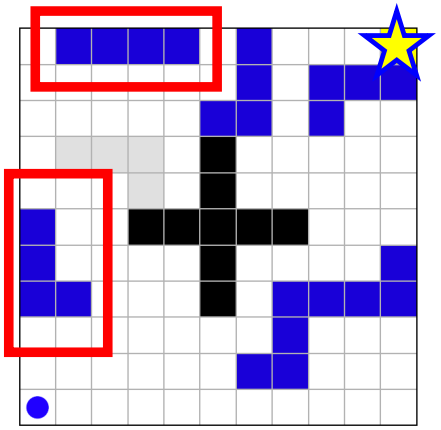


Value-Guided Construal

Value of Construal Behavioral Utility Cognitive Cost

$$V(c) = U(\pi_c) - C(c)$$

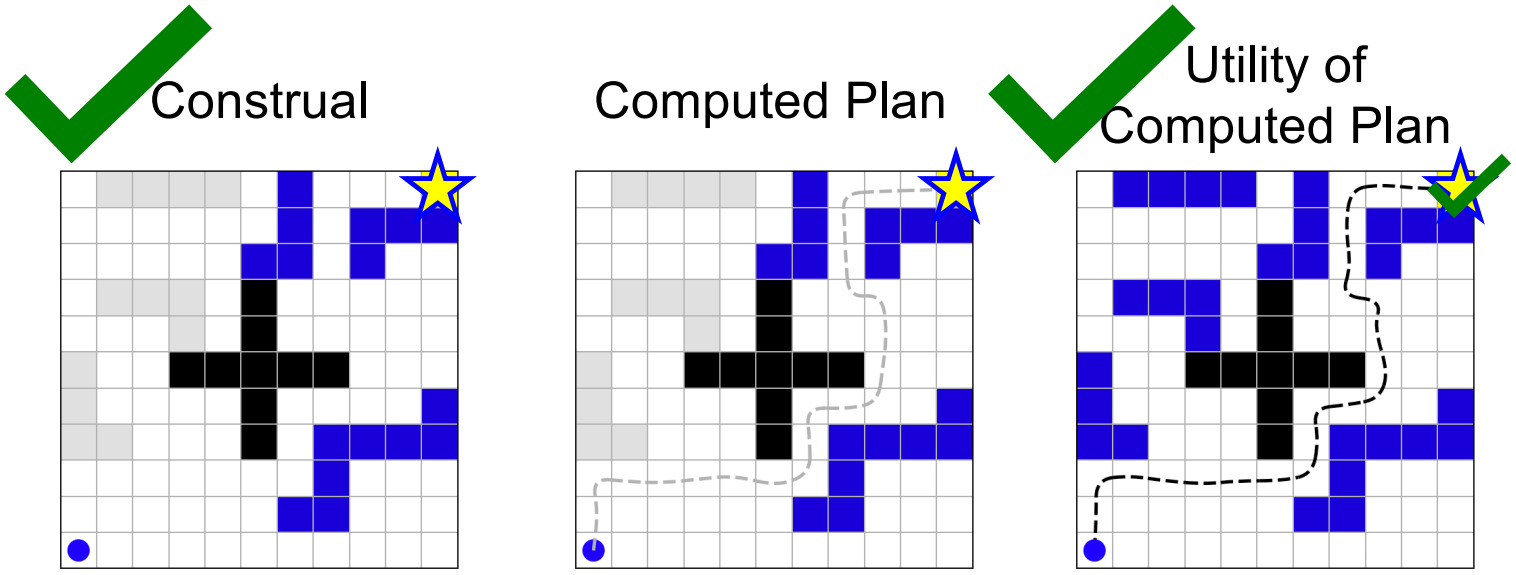
Construal



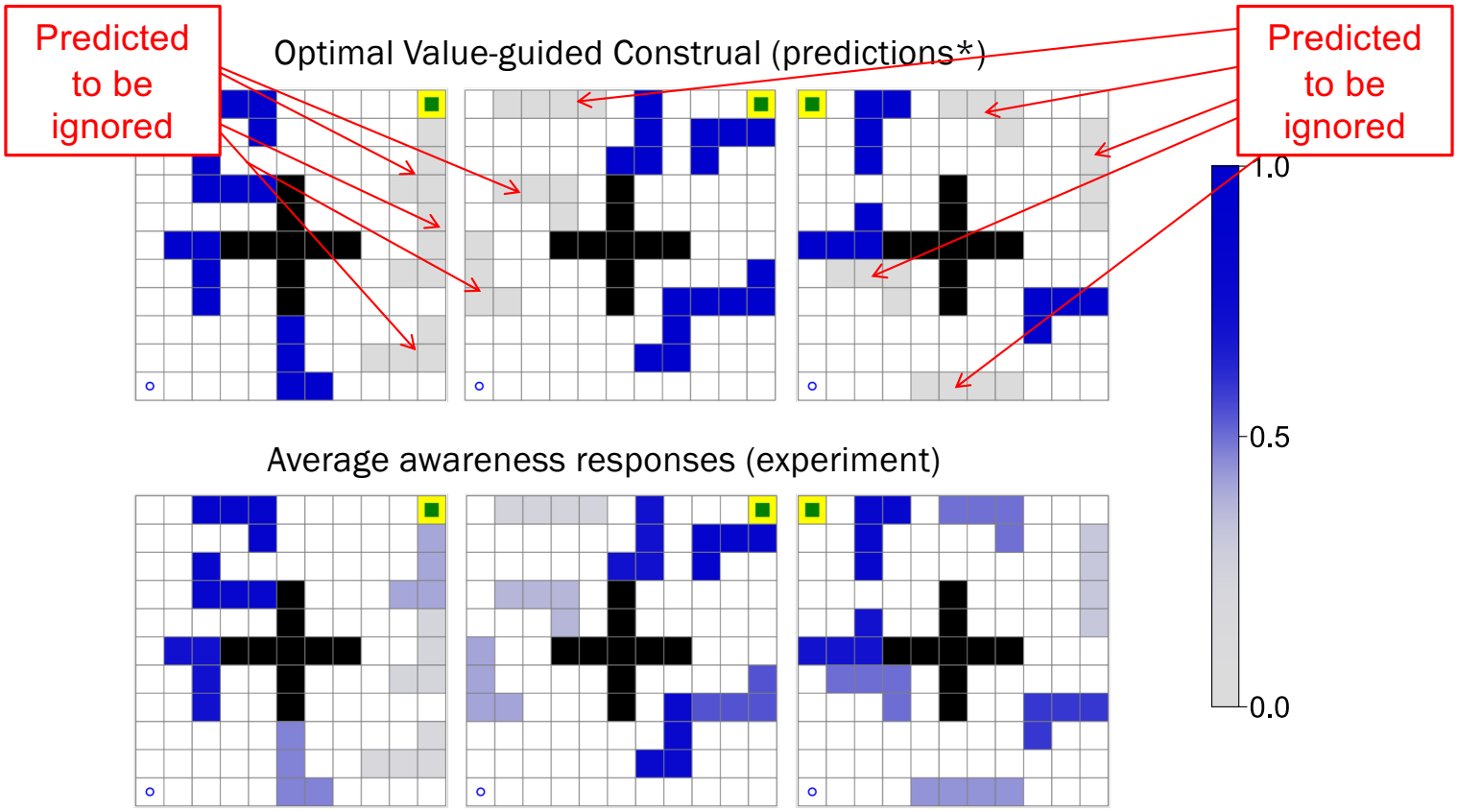
Value-Guided Construal

Value of Construal Behavioral Utility Cognitive Cost

$$V(c) = U(\pi_c) - C(c)$$

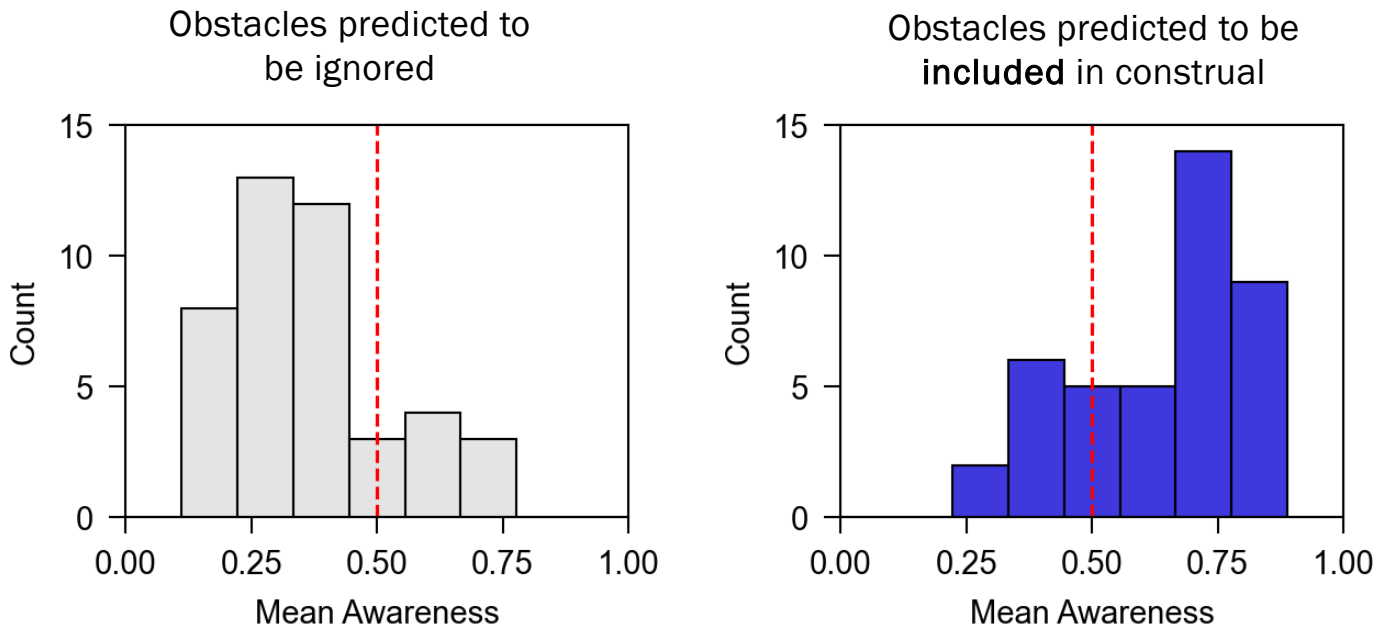


Do people form optimal construals?



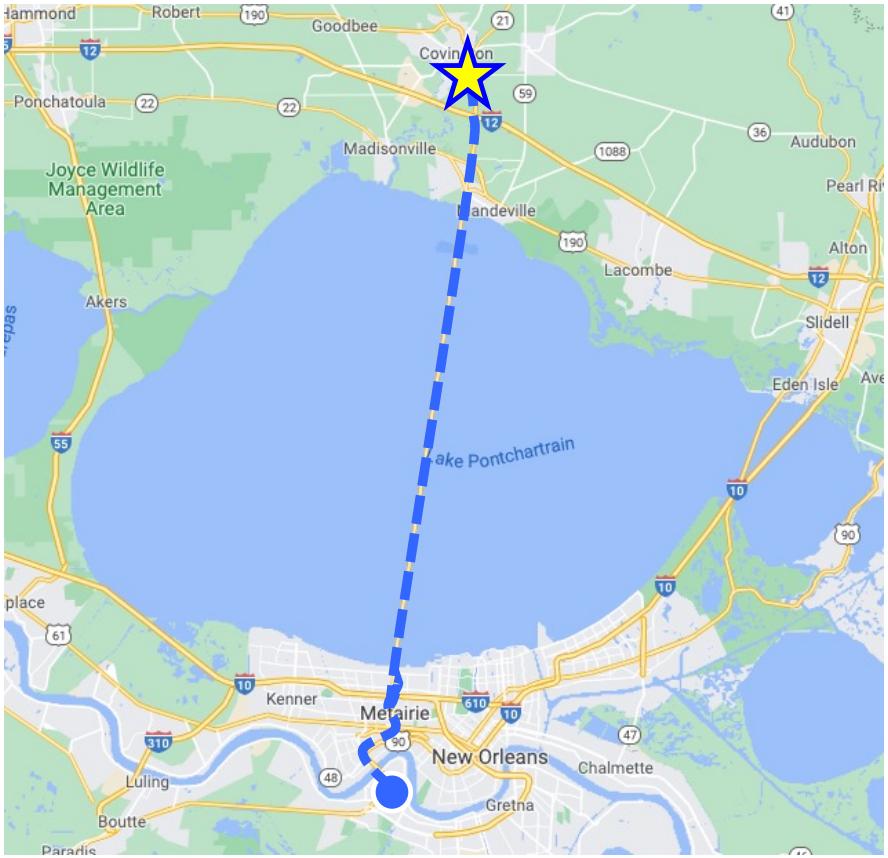
*Pre-registered model predictions: https://osf.io/zpq69

Do people form optimal construals?



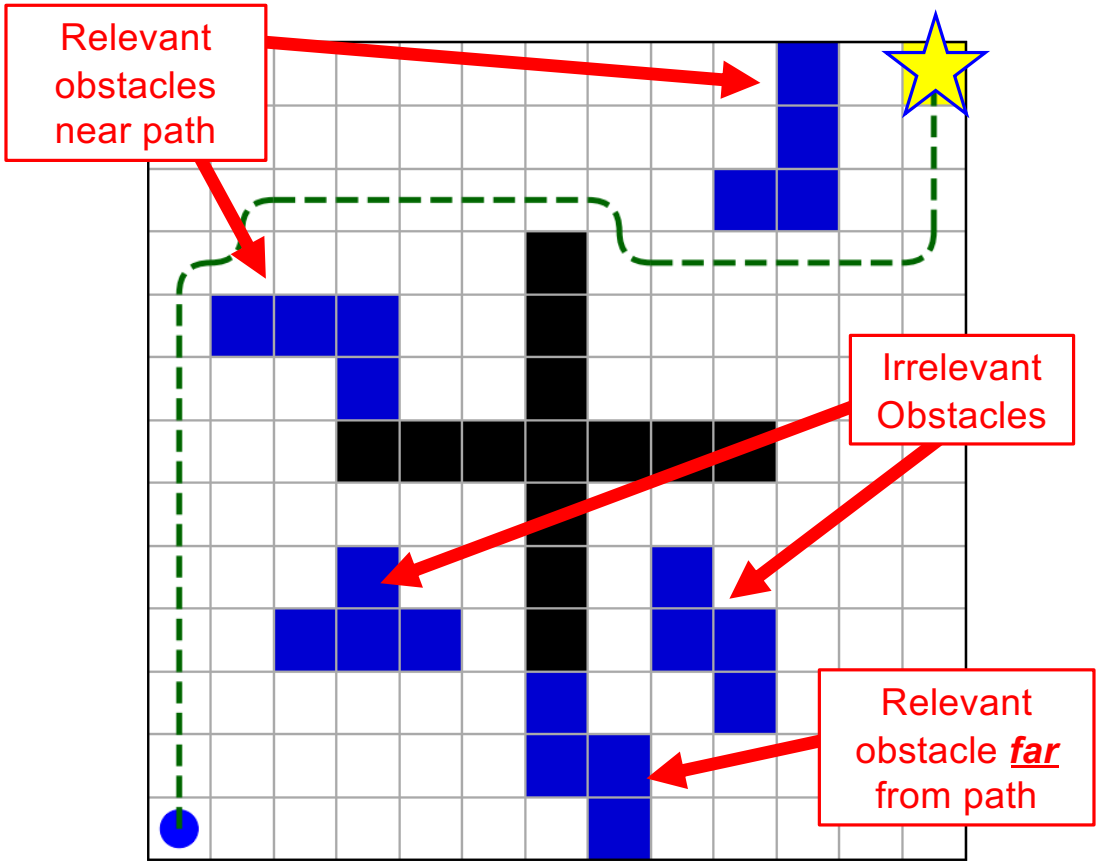
Obstacles split by 0.5 awareness: $\chi^2(1) = 23.03$, $p = 1.6 * 10^{-6}$, $w = 0.52$

Crossing a Bridge

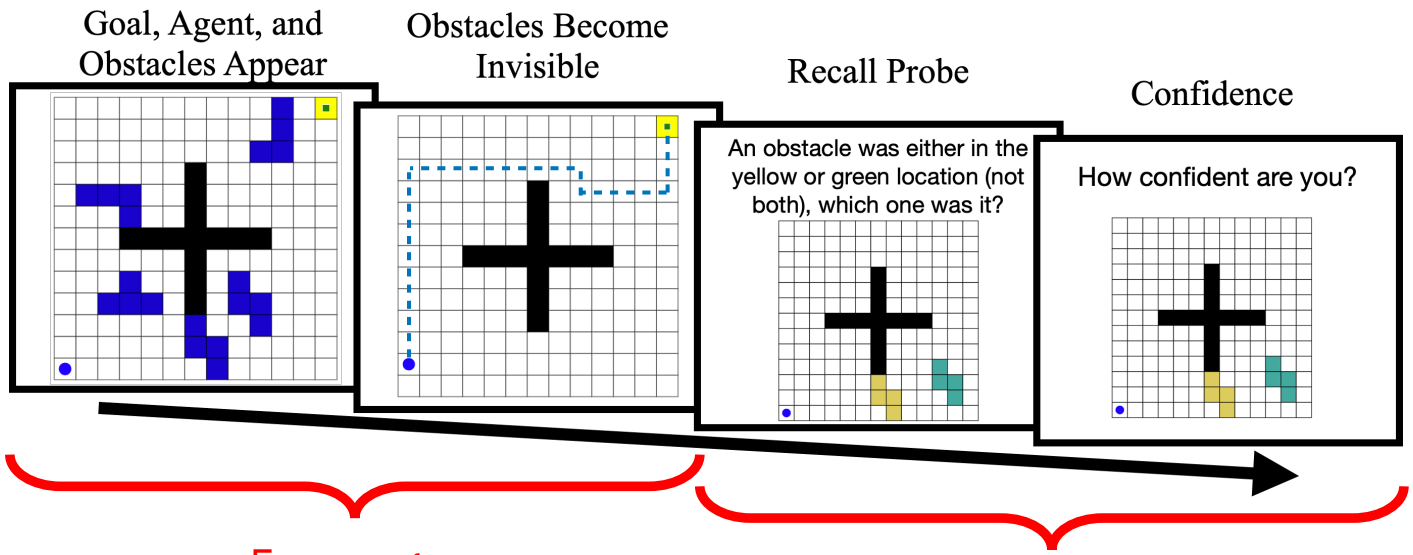


Crossing a Bridge

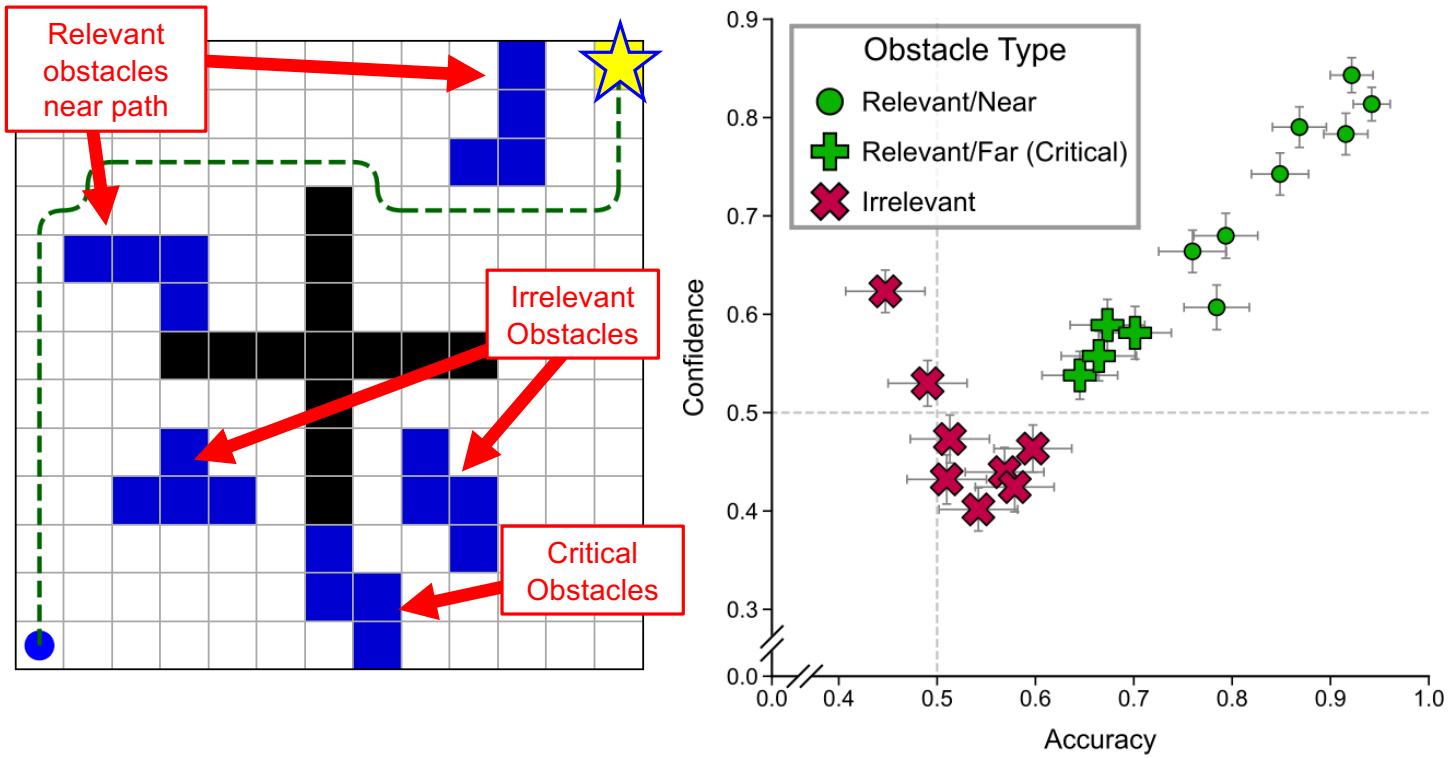
Critical Mazes and Obstacles



Obstacle Recall Probe



Critical Mazes and Obstacles



Error bars are SEM

Two concerns

1. Are patterns of construal a *side-effect* of other perceptual or cognitive mechanisms?

2. Is value-guided construal *computationally feasible*?

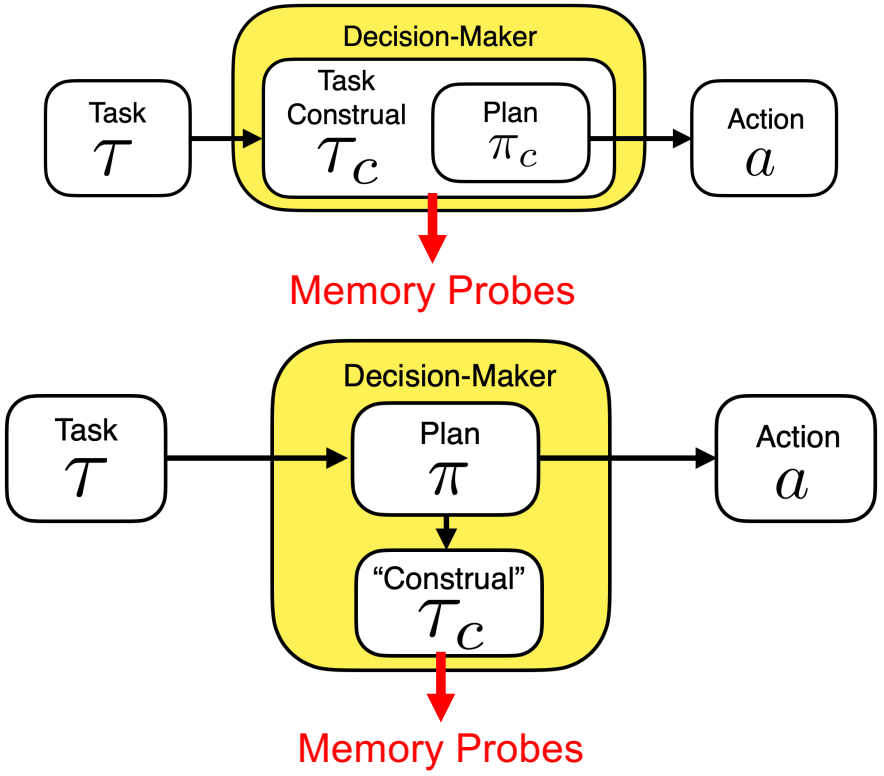
Is construal a side-effect?

Value-Guided Construal

Construal actively shapes planning

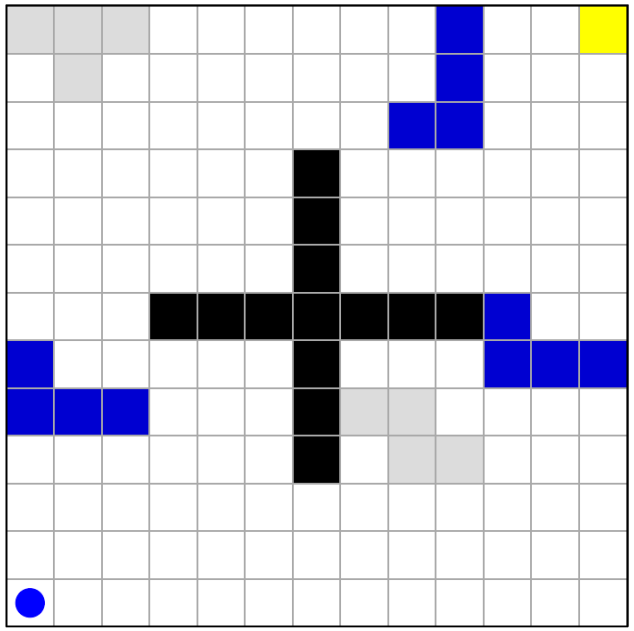
Planning ***without*** Construal

Construal is *side-effect* of other processes (e.g., low-level perceptual cues)



Is construal a side-effect of other processes?

Value-guided Construal



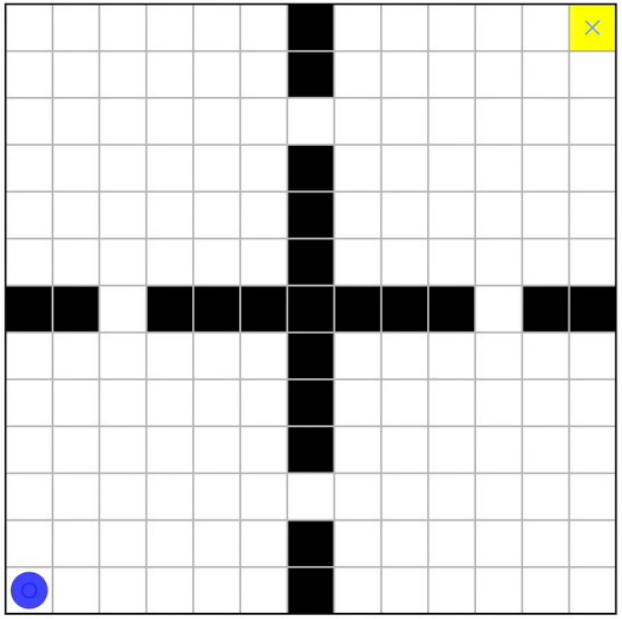
Is construal a side-effect of other processes?

Value-guided Construal

Trajectory-based Heuristic Search

Barto, Bradtke & Singh, 1995
Bonet & Geffner, 2003

Real-Time Dynamic Programming (RTDP) + Heuristic



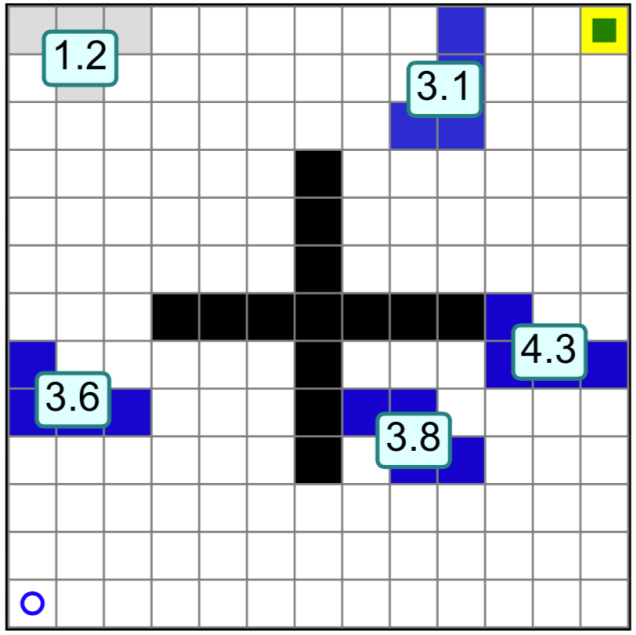
Is construal a side-effect of other processes?

Value-guided Construal

Trajectory-based Heuristic Search

Barto, Bradtke & Singh, 1995
Bonet & Geffner, 2003

Real-Time Dynamic Programming (RTDP) + Heuristic



Log “Hit Count”

Is construal a side-effect of other processes?

Value-guided Construal

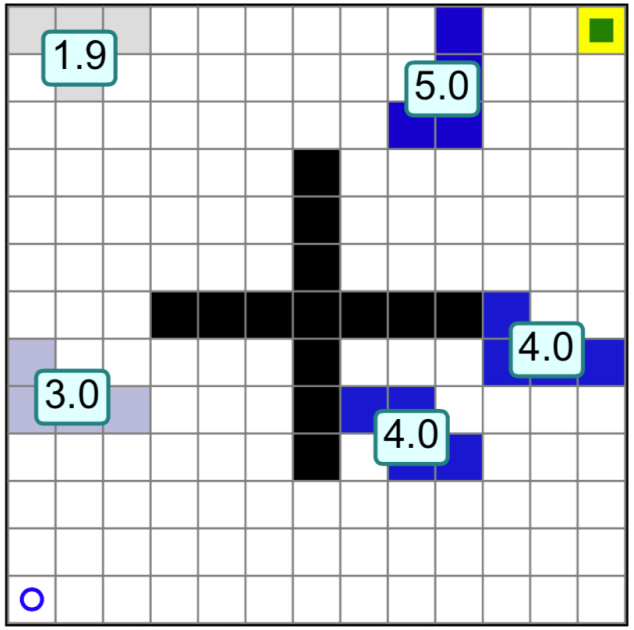
Trajectory-based Heuristic Search

Barto, Bradtke & Singh, 1995
Bonet & Geffner, 2003

Graph-based Heuristic Search

Hart, Nilsson & Raphael, 1968
Hansen & Zilberstein, 2001

LAO* + Heuristic
(A* for MDPs)



Hit Count

Is construal a side-effect of other processes?

Value-guided Construal

Trajectory-based Heuristic Search

Barto, Bradtke & Singh, 1995
Bonet & Geffner, 2003

Graph-based Heuristic Search

Hart, Nilsson & Raphael, 1968
Hansen & Zilberstein, 2001

Related to ***experience replay***
Mattar & Daw, 2018; Pfeiffer & Foster, 2013;
Diba & Buzsáki, 2007

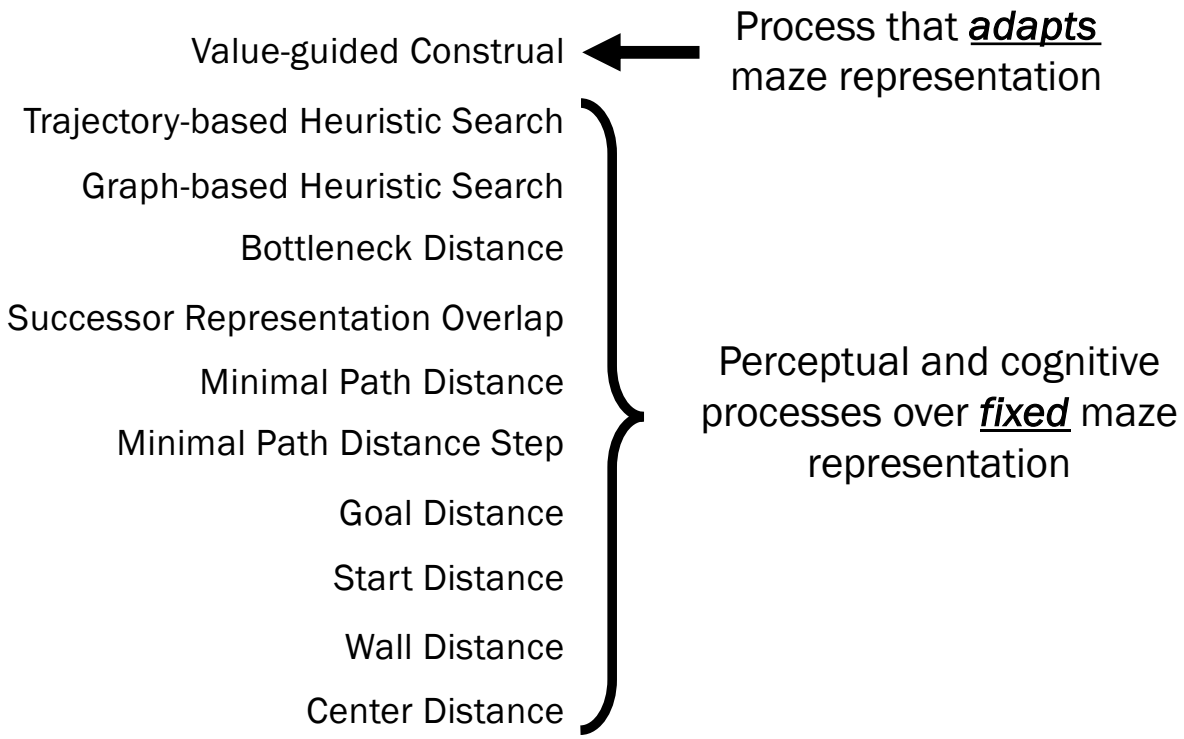
Is construal a side-effect of other processes?

Value-guided Construal

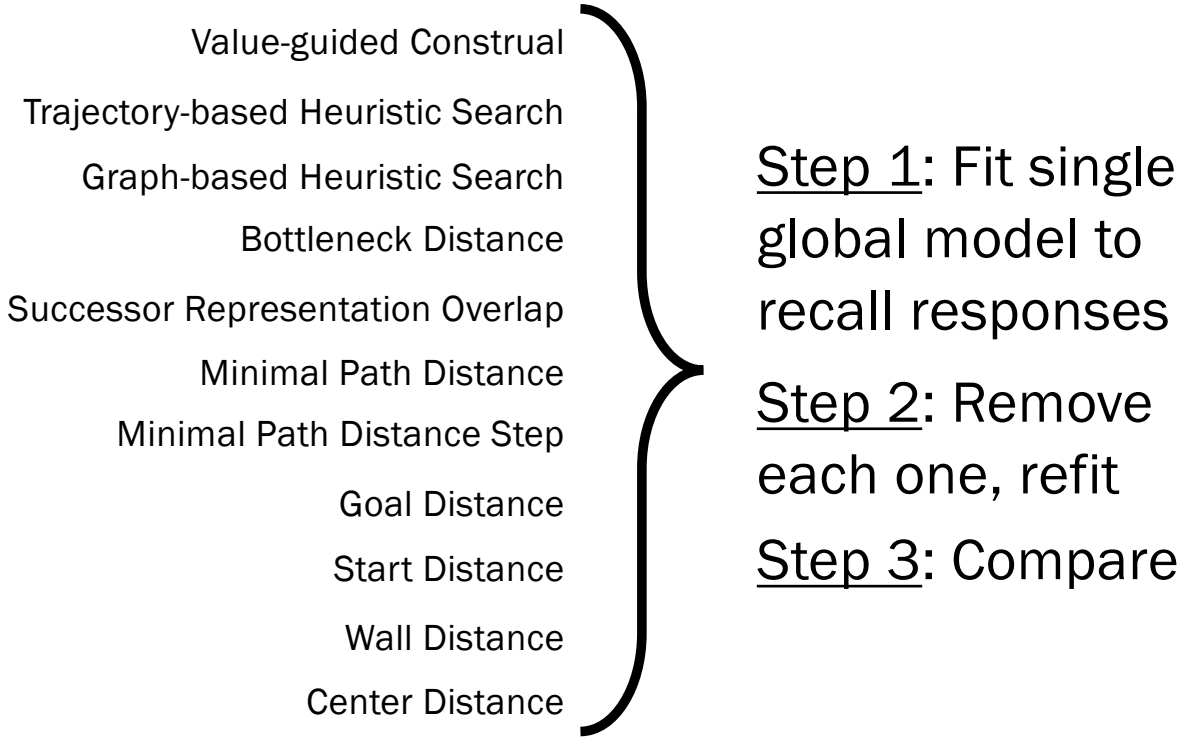
Trajectory-based Heuristic Search

Graph-based Heuristic Search

Is construal a side-effect?

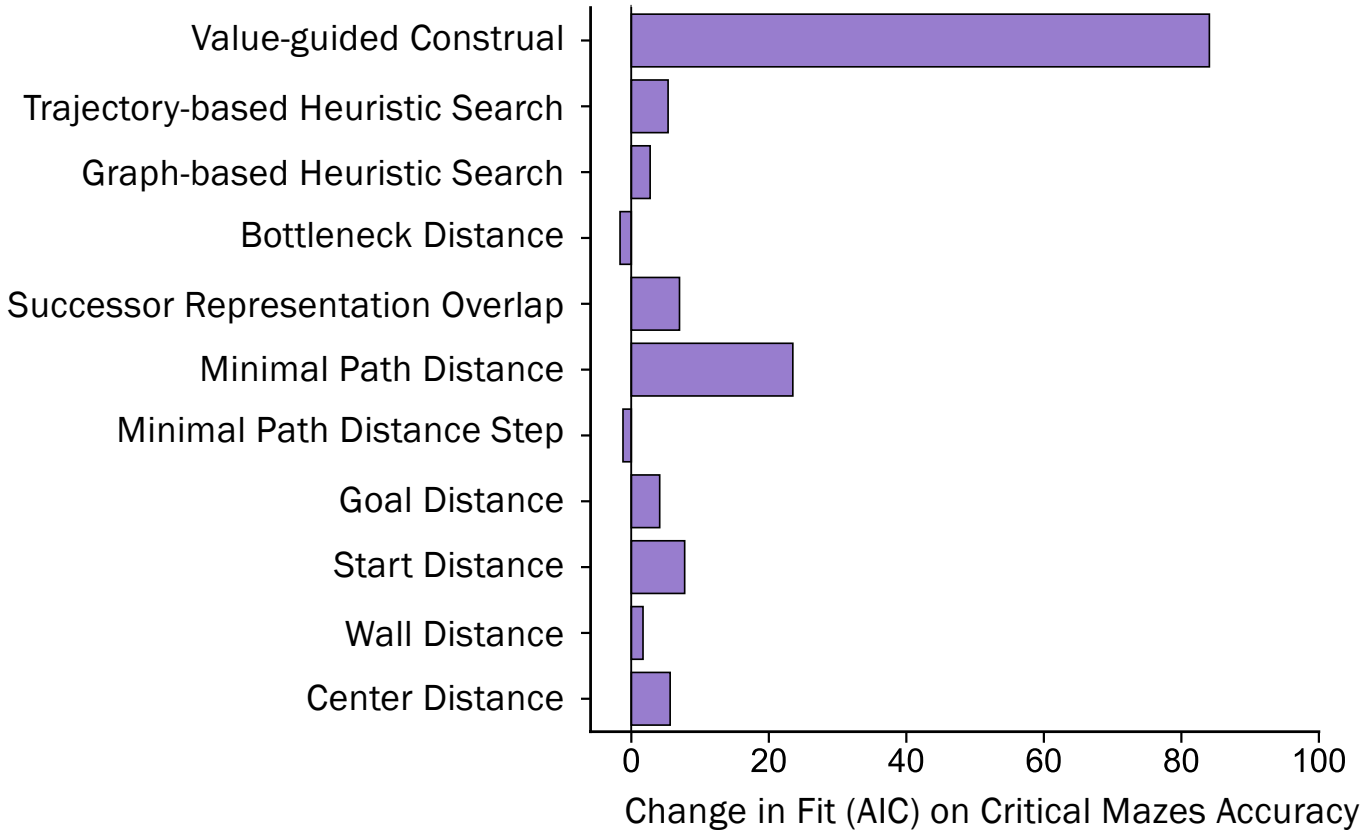


Is construal a side-effect?



Addressing concerns 1 and 2:

Is construal a side-effect?



Two concerns

1. Are patterns of construal a *side-effect* of other perceptual or cognitive mechanisms?

Even when accounting for alternative factors, value-guided construal explains responses

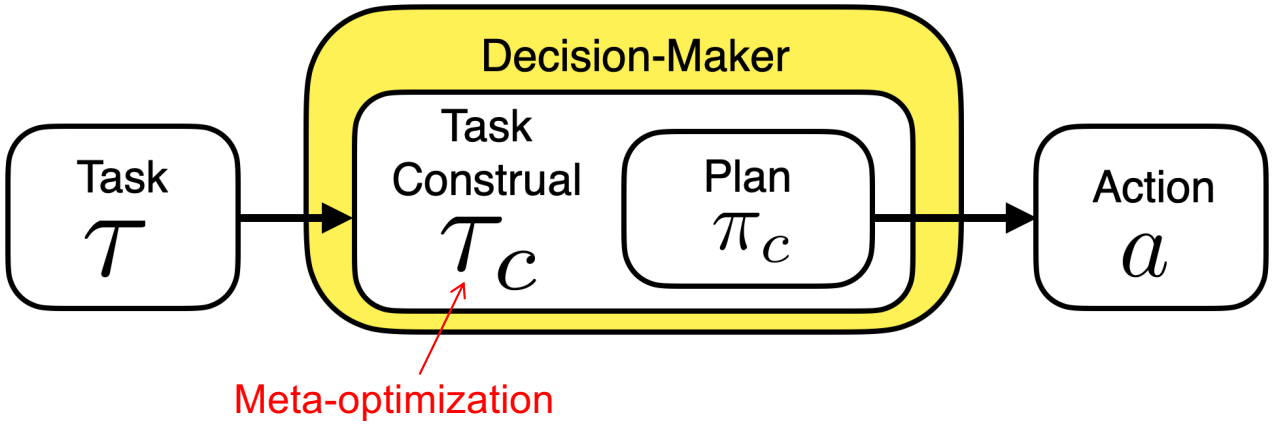
2. Is value-guided construal *computationally feasible*?

Is construal computationally feasible?

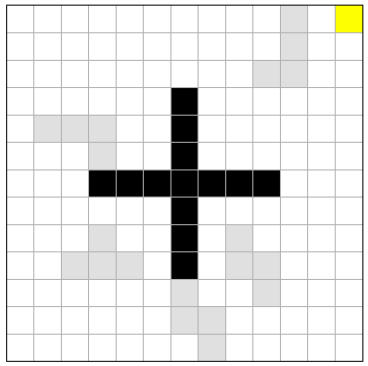
Value of Construal: $V(c)$ = Behavioral Utility: $U(\pi_c)$ − Cognitive Cost: $C(c)$

$$V(c) = U(\pi_c) - C(c)$$

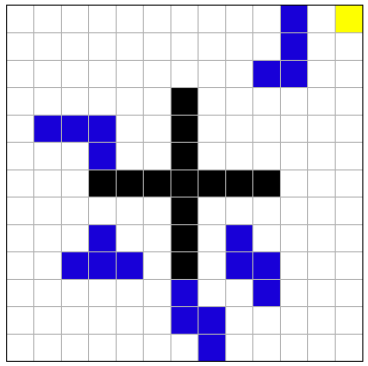
Can construals be efficiently optimized *in principle*?



One approach: Construal search

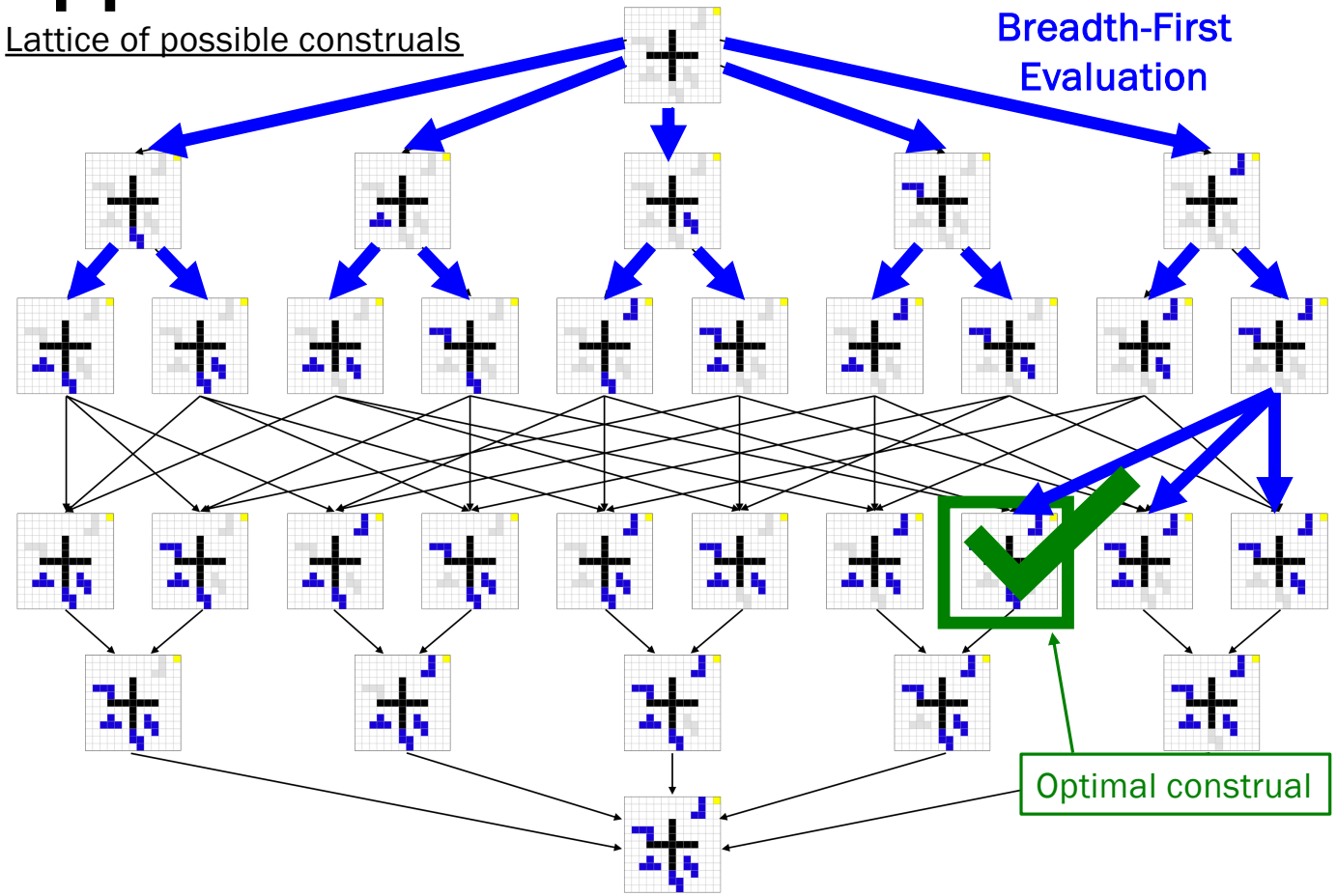


Empty construal



Full construal

One approach: Construal search



Two concerns

1. Are patterns of construal a *side-effect* of other perceptual or cognitive mechanisms?

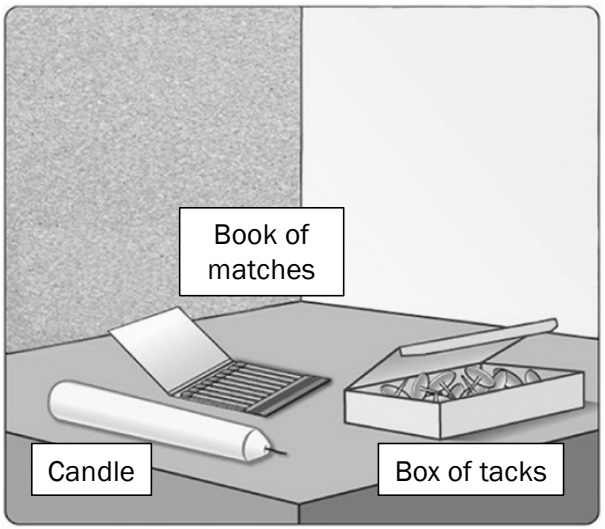
Even when accounting for alternative factors, value-guided construal explains responses

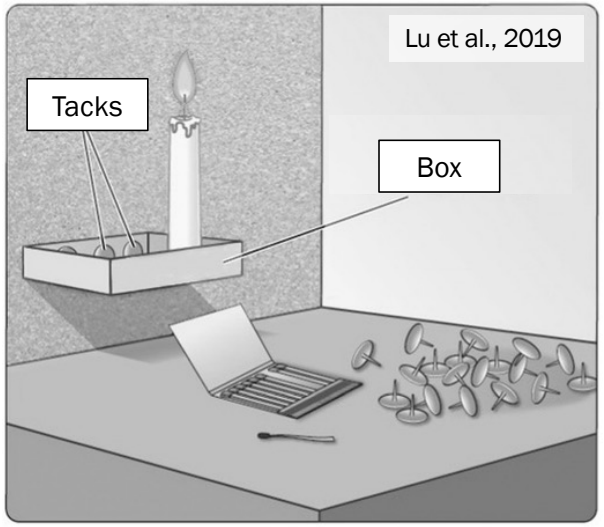
2. Is value-guided construal *computationally feasible*?

Optimal construal could be efficiently approximated using search-based methods.

Functional Fixedness

Goal: Mount the candle to the wall and light it



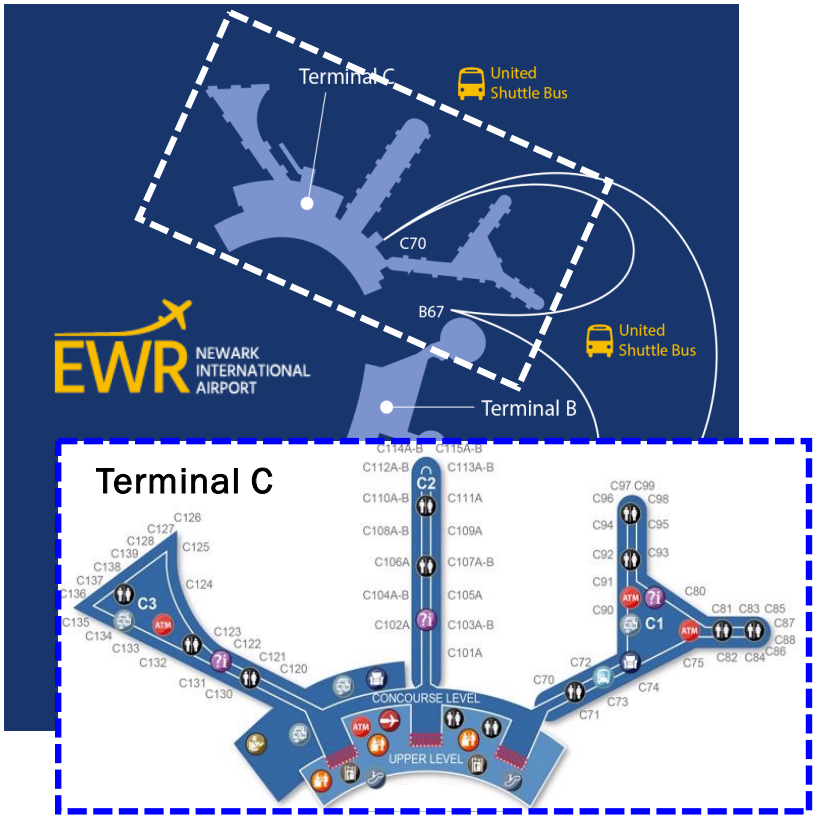


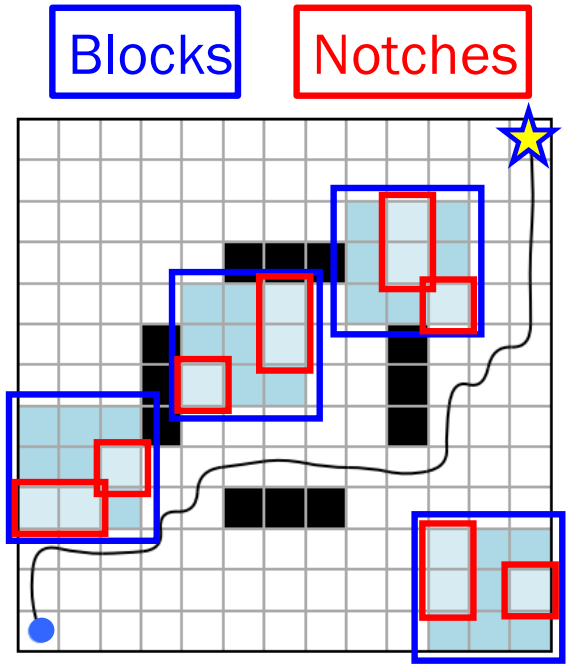
Participants used the wrong *task representation*

(e.g., construing the box as a *container* and not as a *support*)

Duncker, 1945

Functional Fixedness: A consequence of construal

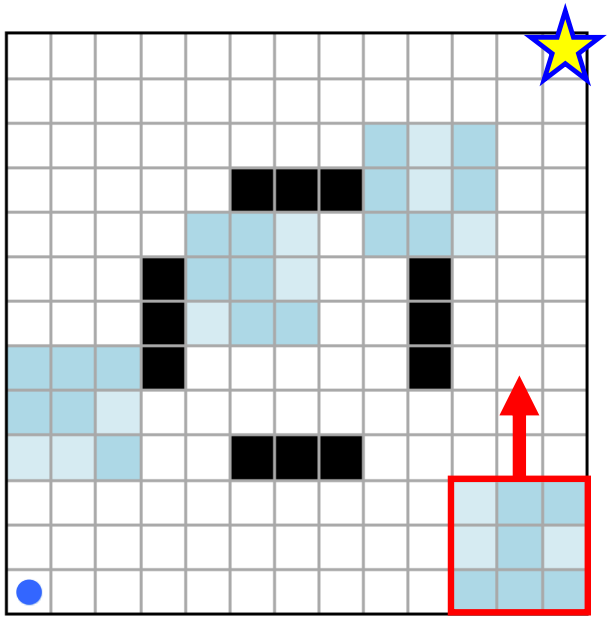




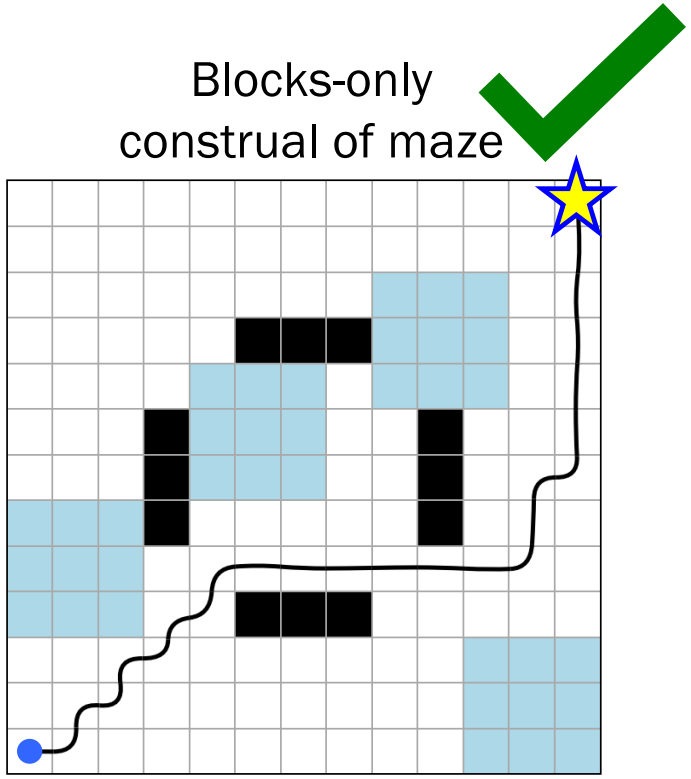
Ho, Cohen & Griffiths (in press) *Psychological Science*.

Blocks and Notches

True maze



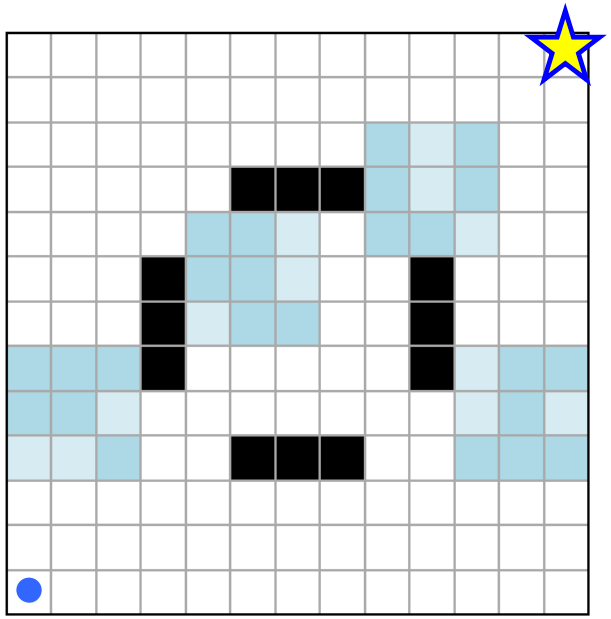
Blocks-only construal of maze



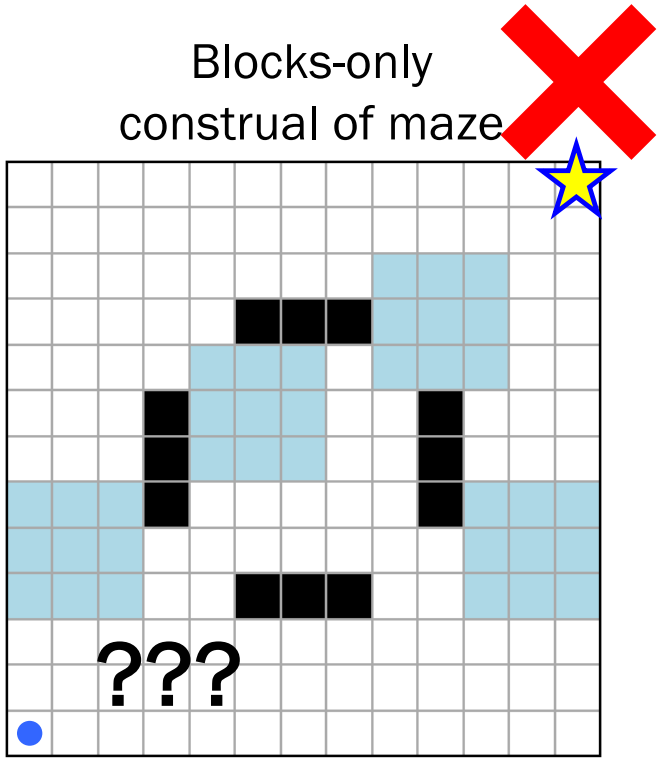
Ho, Cohen & Griffiths (in press) *Psychological Science*.

Blocks and Notches

True maze

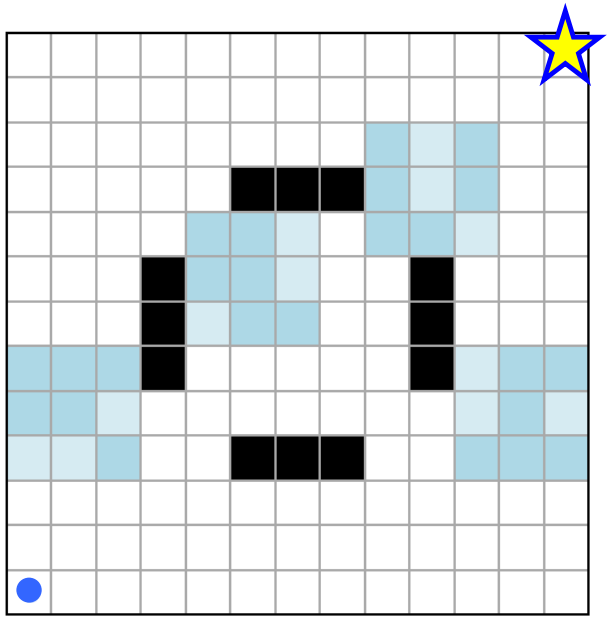


Blocks-only
construal of maze

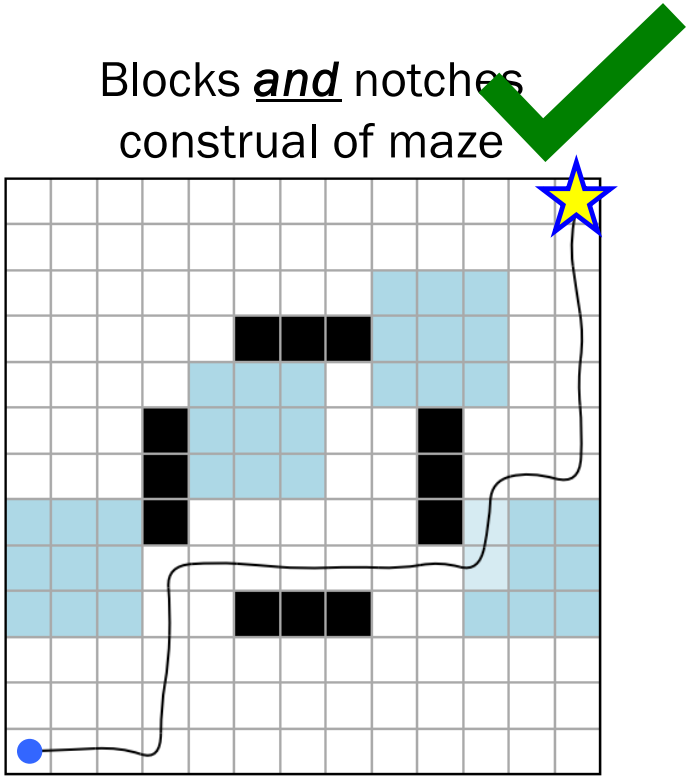


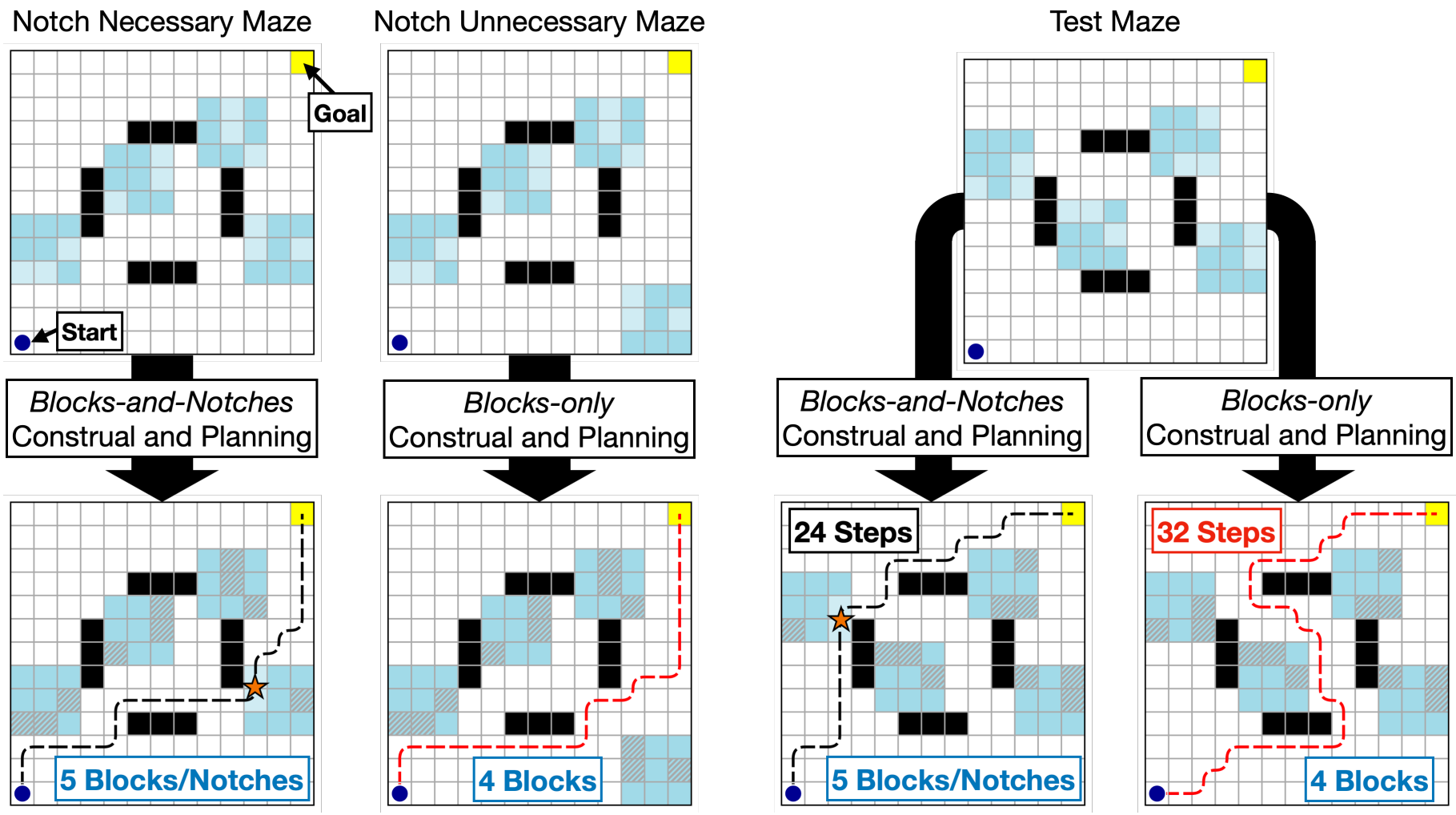
Blocks and Notches

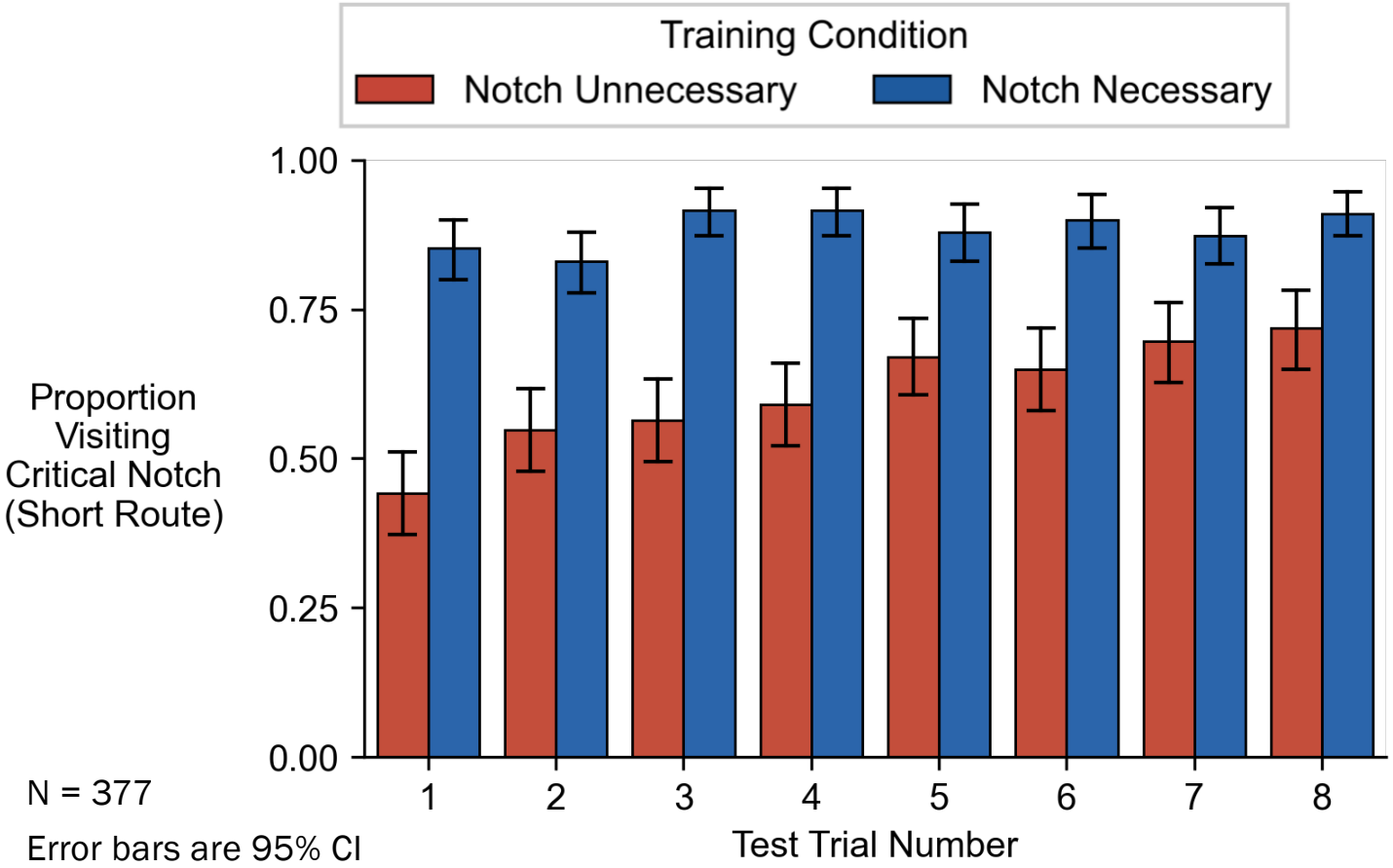
True maze



Blocks ***and*** notches construal of maze




Notch Necessary Maze
Notch Unnecessary Maze
Test Maze
Goal
Start
Blocks-and-Notches
Construal and Planning
Blocks-only
Construal and Planning
Blocks-and-Notches
Construal and Planning
Blocks-only
Construal and Planning
5 Blocks/Notches
4 Blocks
24 Steps
5 Blocks/Notches
32 Steps
4 Blocks



N = 377

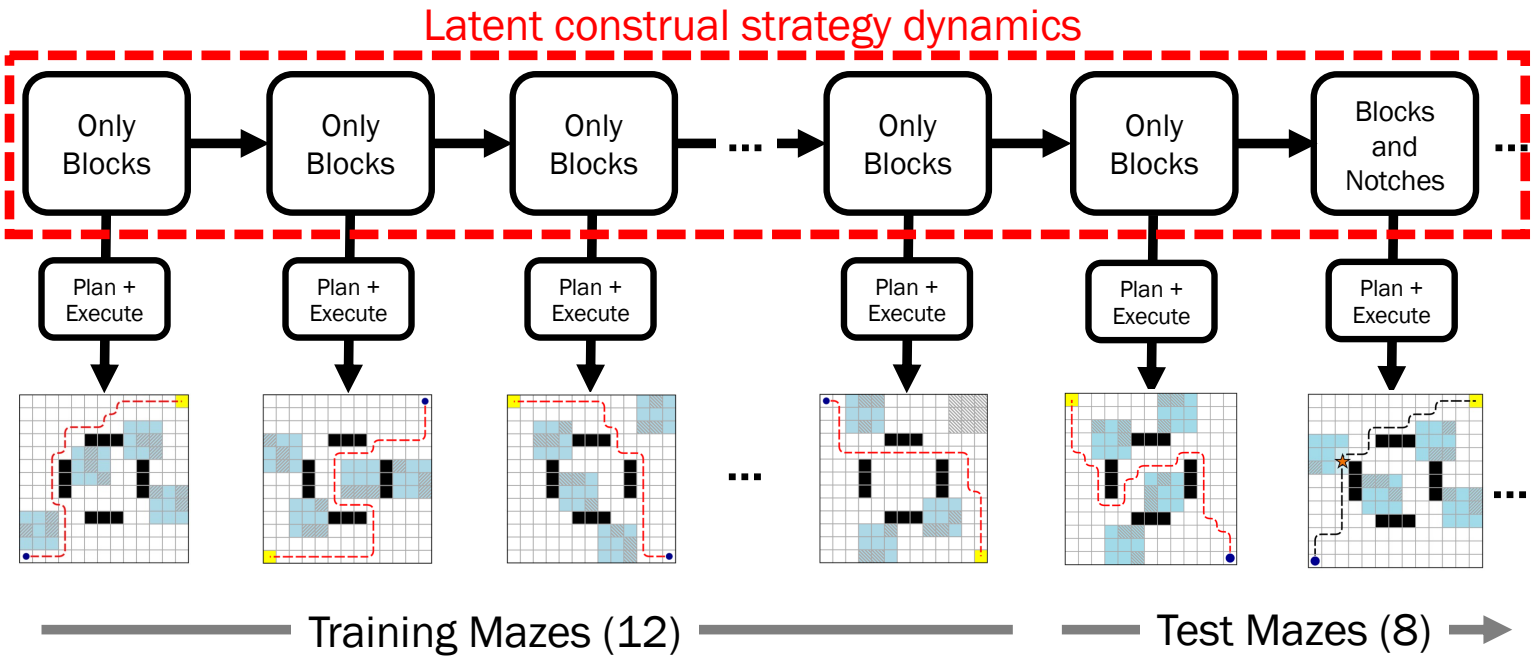
Error bars are 95% CI

Training Condition: β = -1.90, SE = 0.17, z = -11.22, $p < 2*10^{-16}$

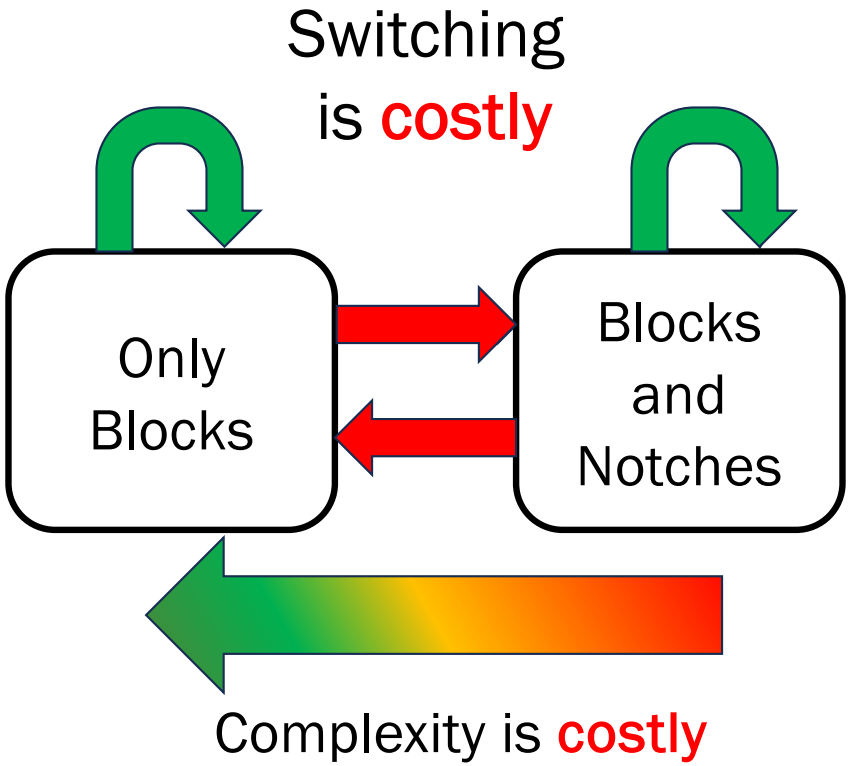
Trial Number: β = 0.11, SE = 0.02, z = 5.08, $p = 3.7*10^{-7}$

Interaction: β = 0.09, SE=0.04, z = 2.15, p = 0.03

Modeling Functional Fixedness



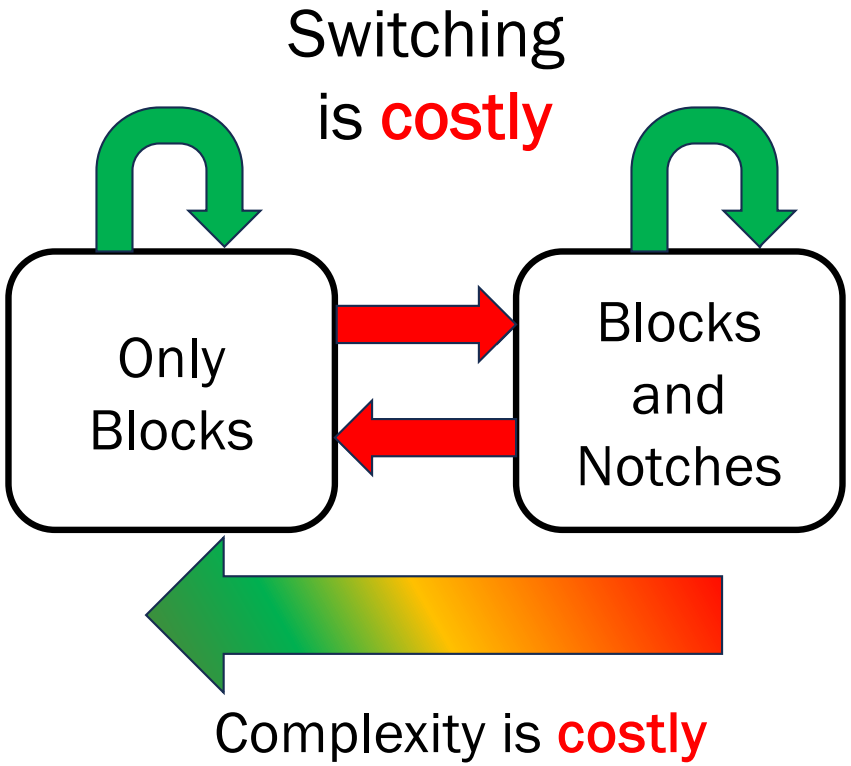
Modeling Functional Fixedness



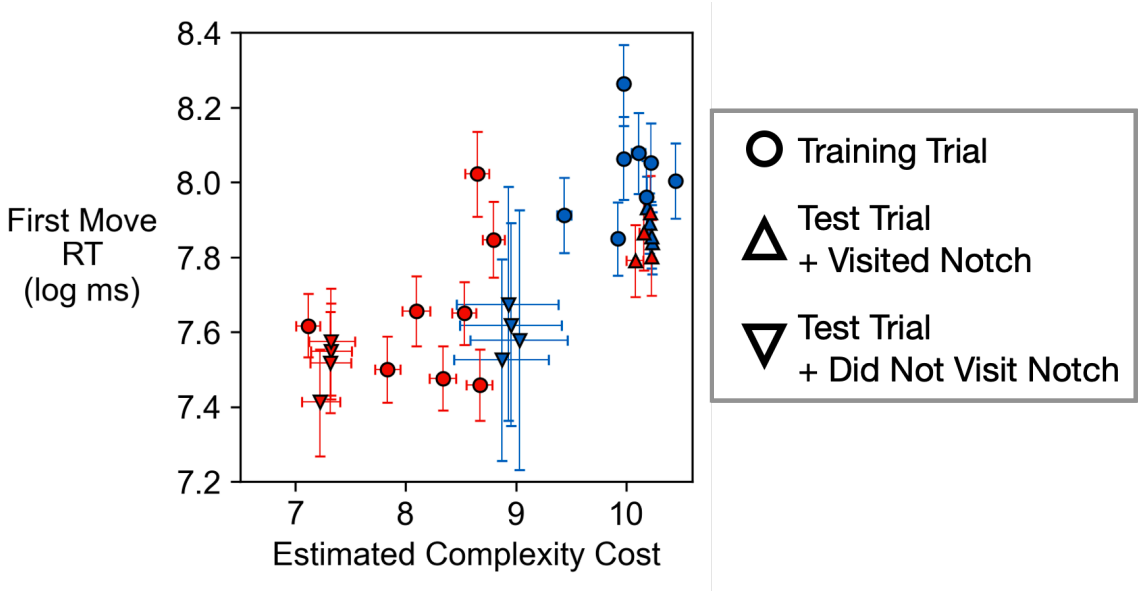
Complexity *and* switching best explain behavior (i.e. functional fixedness)

Model	df	AIC	ΔAIC
No complexity or switch cost	1	132094	2804
Only complexity cost	2	130535	1245
Only switch cost	2	130057	767
Both complexity and switch cost	3	129289	0

Modeling Functional Fixedness



Complexity costs predict reaction time

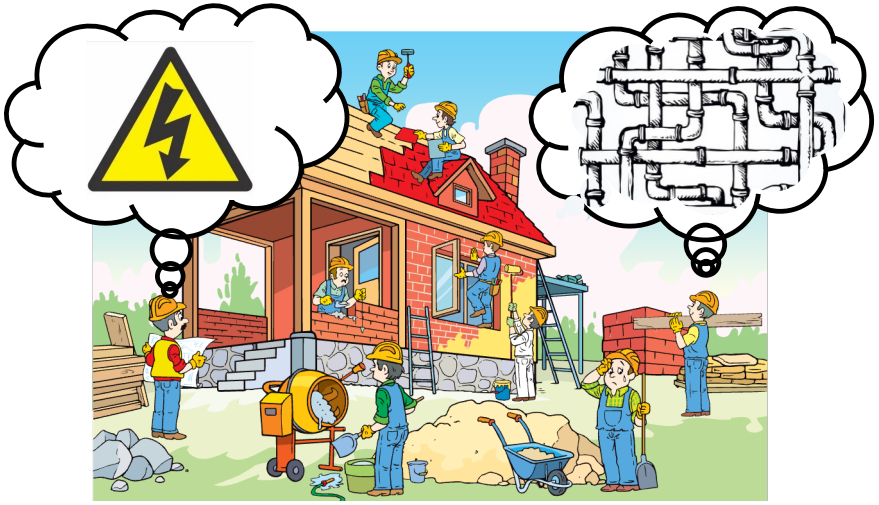


Making sense of social intelligence

Communicative Interactions

Ho et al. (2021) *JEP:G*;
Ho et al. (2018) *Cog Sci*;
Ho et al. (2016) *NeurIPS*

Coordination and Joint Planning



Ho et al. (2016) *CogSci*;
Carroll, et al. (2019) *NeurIPS*.

Natural stupidity and artificial intelligence

My colleagues, they study artificial intelligence; me, I study natural stupidity.
- Amos Tversky

How can we develop AI systems that complement, rather than enhance, our natural stupidity?

Closing thoughts

- Making sense of intelligence
  - Rational vs mechanistic accounts of natural and artificial intelligence
- Task representations in human problem solving
  - Value-guided construal
  - Functional Fixedness
- Making sense of social intelligence
- Using cognitive science to inform interactive ML

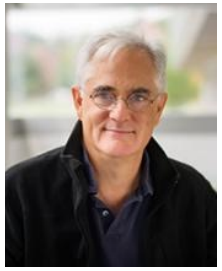
Collaborators

Carlos Correa
NYU

David Abel
DeepMind

Tom Griffiths
Princeton

Michael Littman
Brown



Jon Cohen
Princeton

Thanks for listening!


NYU


JOHN
TEMPLETON
FOUNDATION
Inspiring Awe & Wonder


NSF


Computation and
Decision-Making
Lab


TOYOTA
RESEARCH INSTITUTE

Questions?