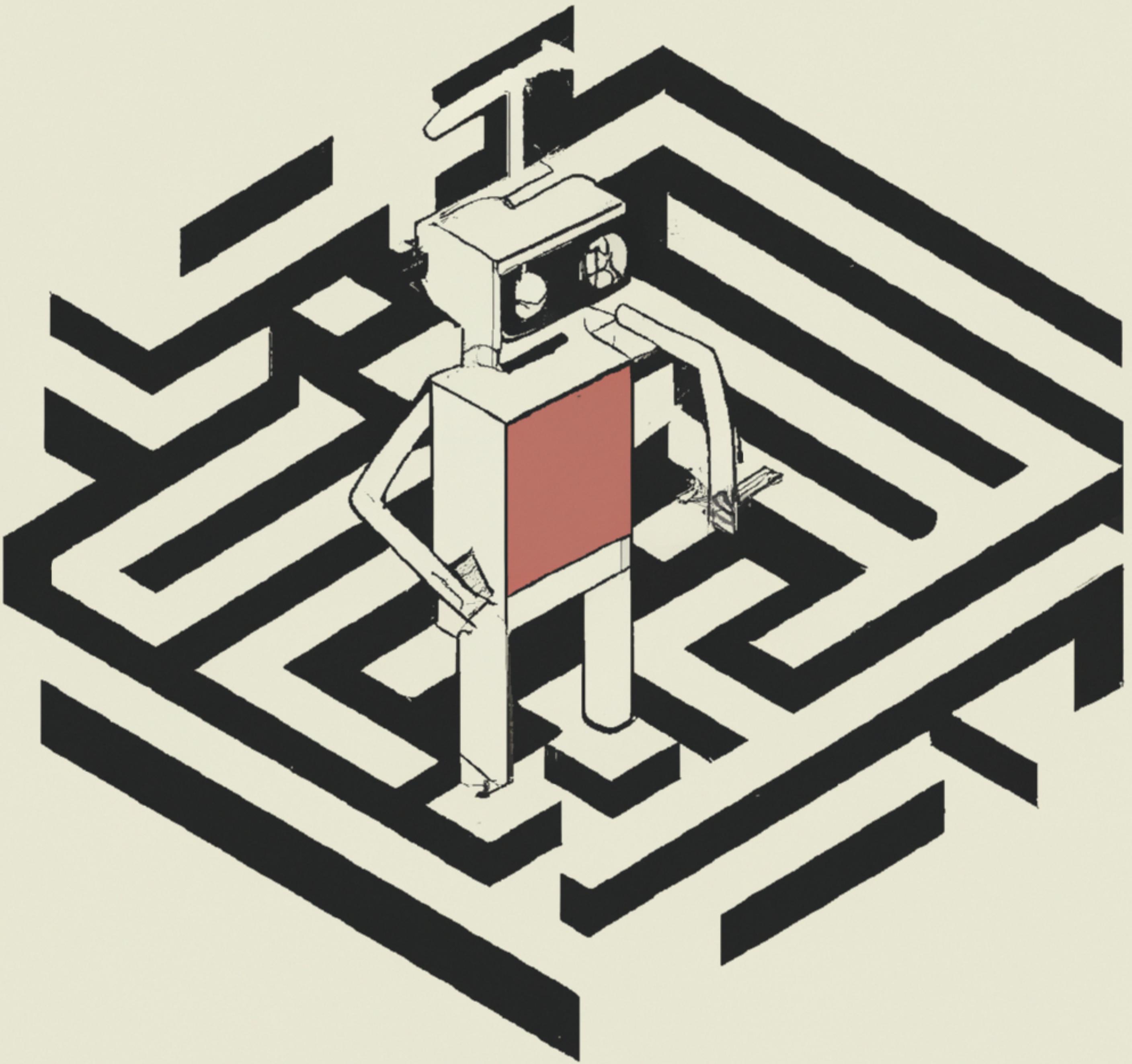


Seminar in Cognitive Modelling

Lecture 3 - Representation



Housekeeping

- The website is now updated to reflect presenters.
- The first week for portfolio reflections is *this week*
- Bring your Tuesday reflection, printed out, to class on Thursday if you didn't today!
- Reminder: You “Must attend” the time corresponding to the presentation you’re giving
 - And “May attend” the other time

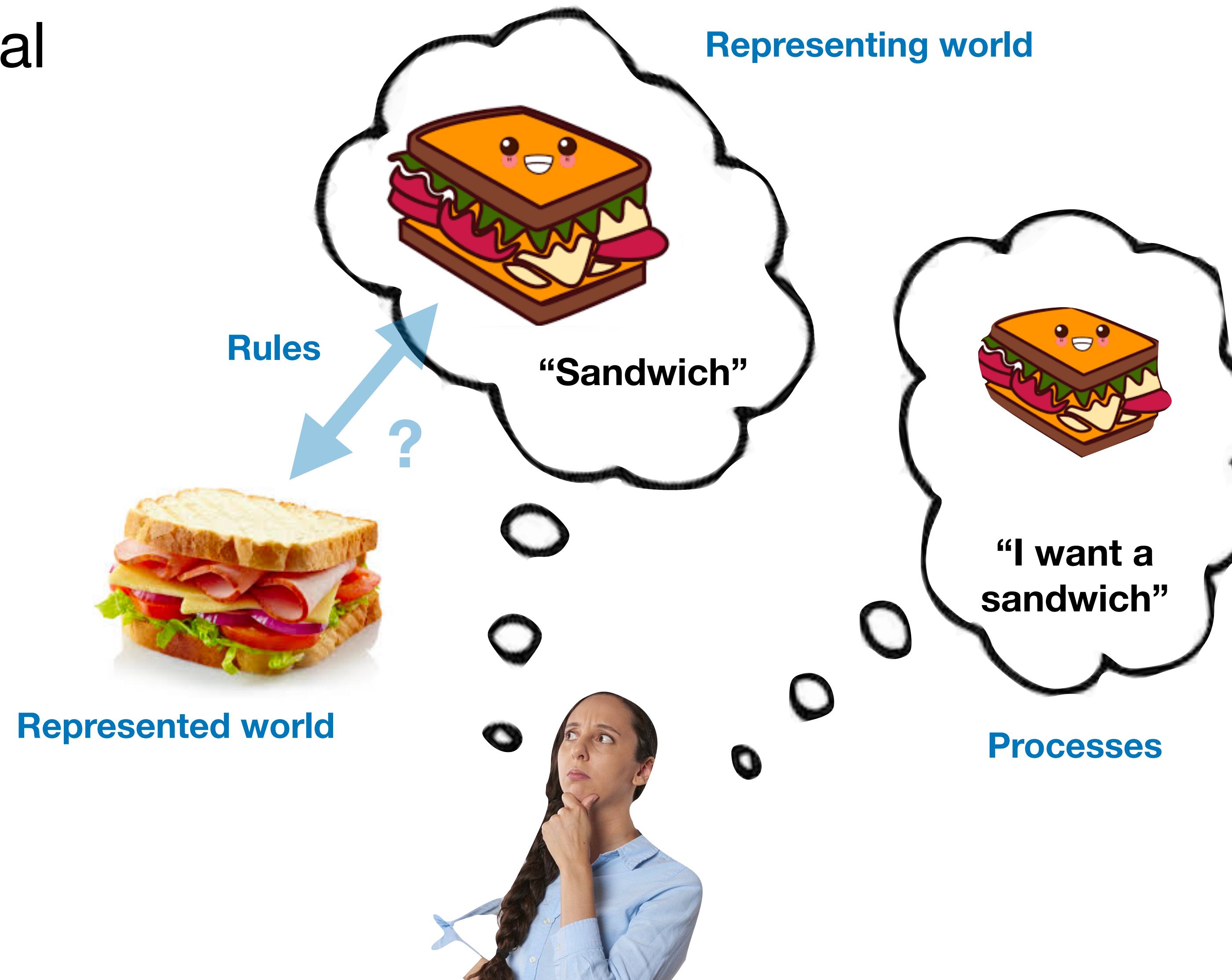
Outline

1. What's a representation?
2. How experiments probe at representation (and process)
3. Cognitive maps in rats and men

What is a representation?

Art Markman offers a general definition:

- Four ingredients:
 1. A represented world
 2. A representing world
 3. Representing rules
 4. Process that uses the representation



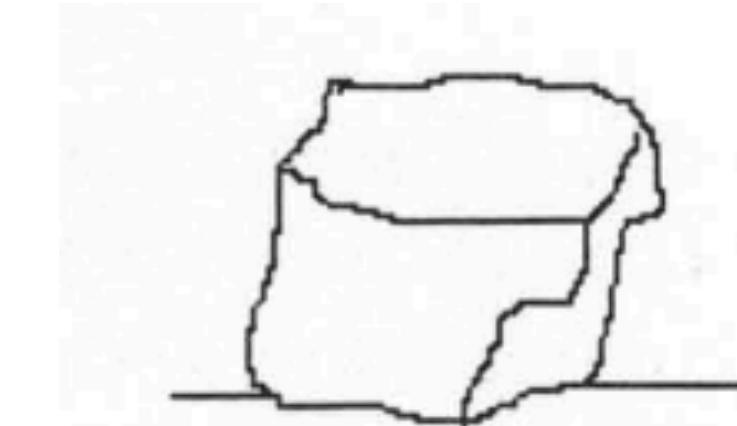
- There are many ways to represent things
- And world is rich with properties that we may or may not want to represent

E.g.

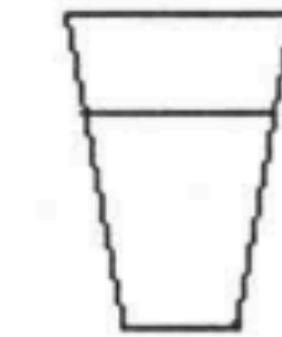
Fig 1.2 in Knowledge representation

Represented world

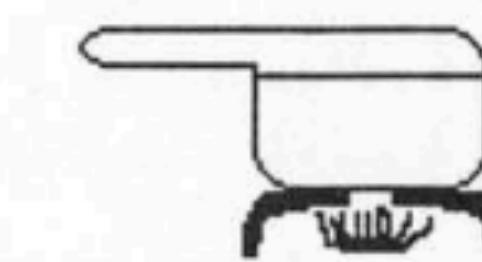
Block of ice



Glass of water



Pan of water over flame



These are also representations

Representing world

- Suppose I am only interested in representing one thing about these examples: **temperature**

Representation rule: ~~integer's share and (temperature)~~
Digital values in differences <1

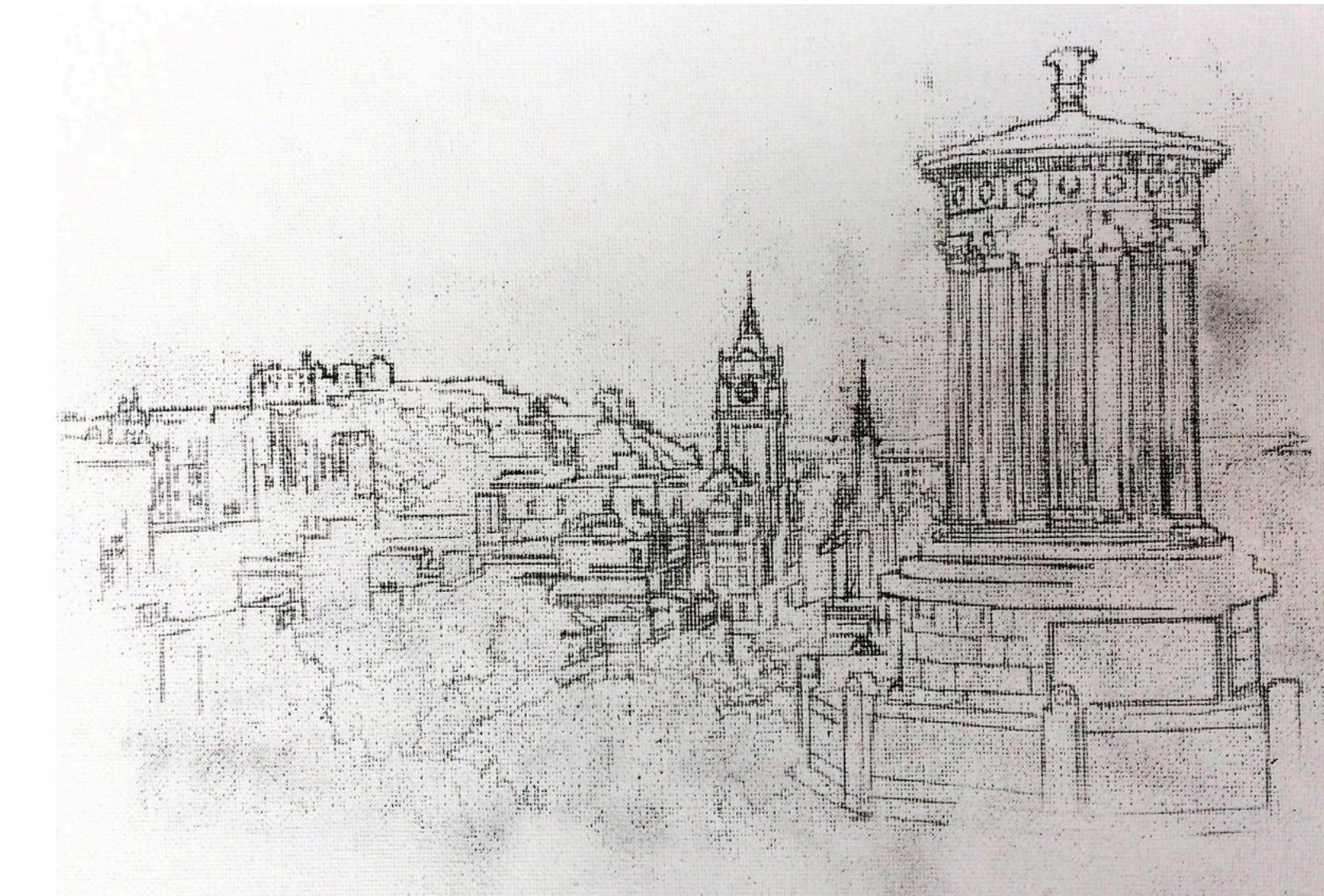
Why represent?



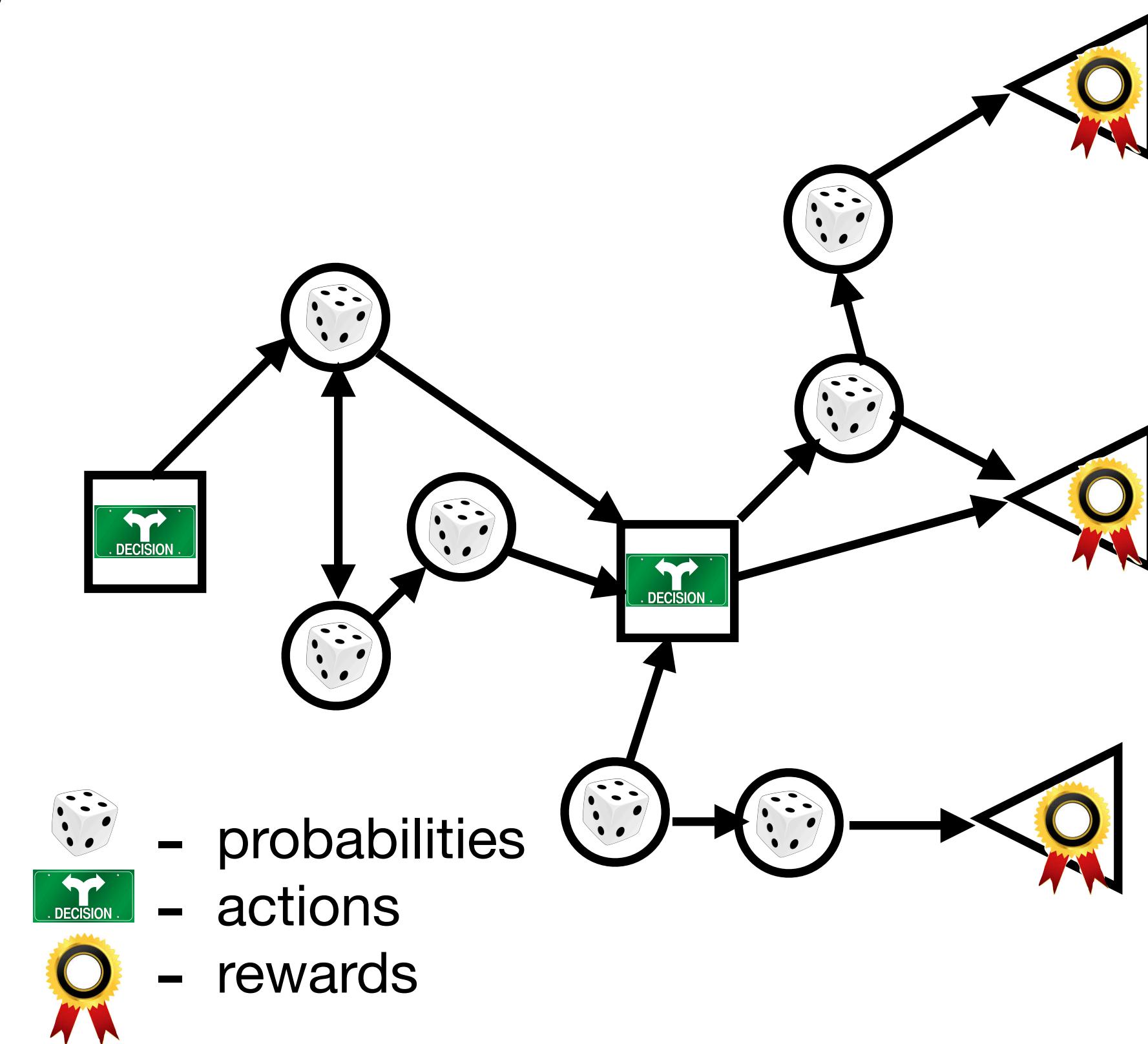
“If the organism carries a “small-scale model” of external reality and of its own possible actions within its head, it is able to **try out various alternatives**, conclude which is the best of them, **react to future situations** before they arise, **utilise the knowledge of past events** in dealing with the present and future, and in every way to **react in a much fuller, safer, and more competent manner** to the emergencies which face it.”

Kenneth Craik, *The Nature of Explanation*, 1943

- To what level of detail do we, *or should we*, represent the world in our heads?
- Photorealistically?
 - Presumably not
- More abstractly / efficiently /compactly?
 - Maybe, but how exactly?



- ...perhaps something like a map
- ...or an network / flow chart of personally relevant options & affordances
- Whatever best supports the processing we need to do...



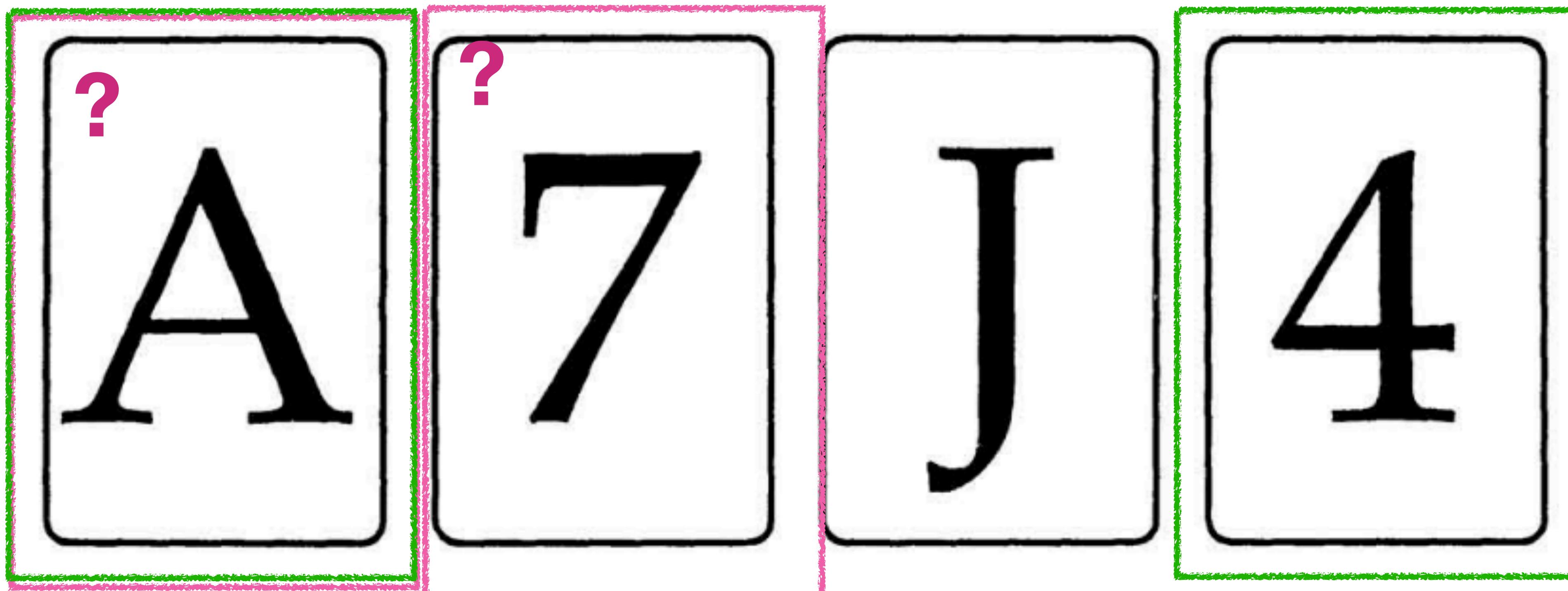
Interim summary

- So seems like part of what minds do is represent the world in ways that enable cognitive processing
- How can we probe how the mind represents things?
- Markman gives example of Wason card selection task

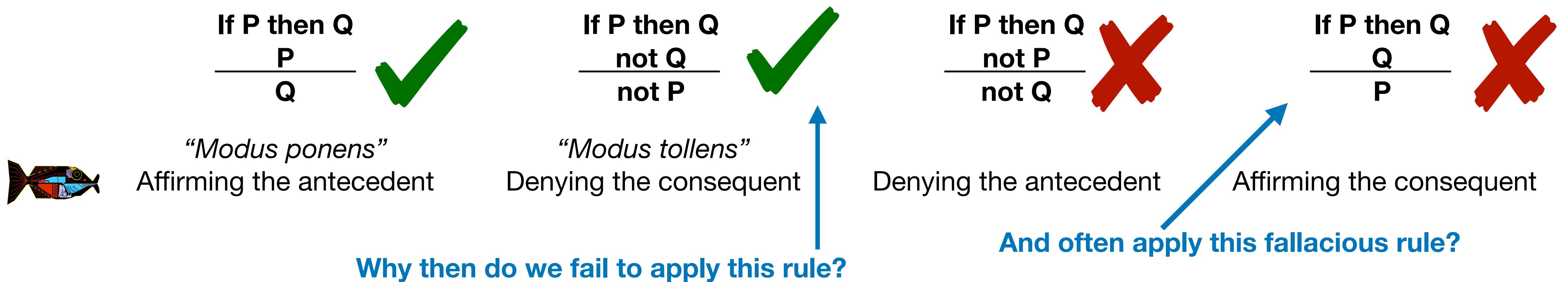
Let's play a game



- "*If there is a vowel on one side of the card, then there is an odd number on the other side of the card.*" (Wason & Johnson-Laird, 1972)
- Point to the card(s) you should turn over to test this rule



- People are weirdly bad at the Wason card selection task
- Prevailing view at the time was that we represent and reason about the world **logically** (e.g. Rips, 1994; Braine, Reiser, and Romain, 1984)
- E.g. Represent statements symbolically & apply rules of deductive logic:



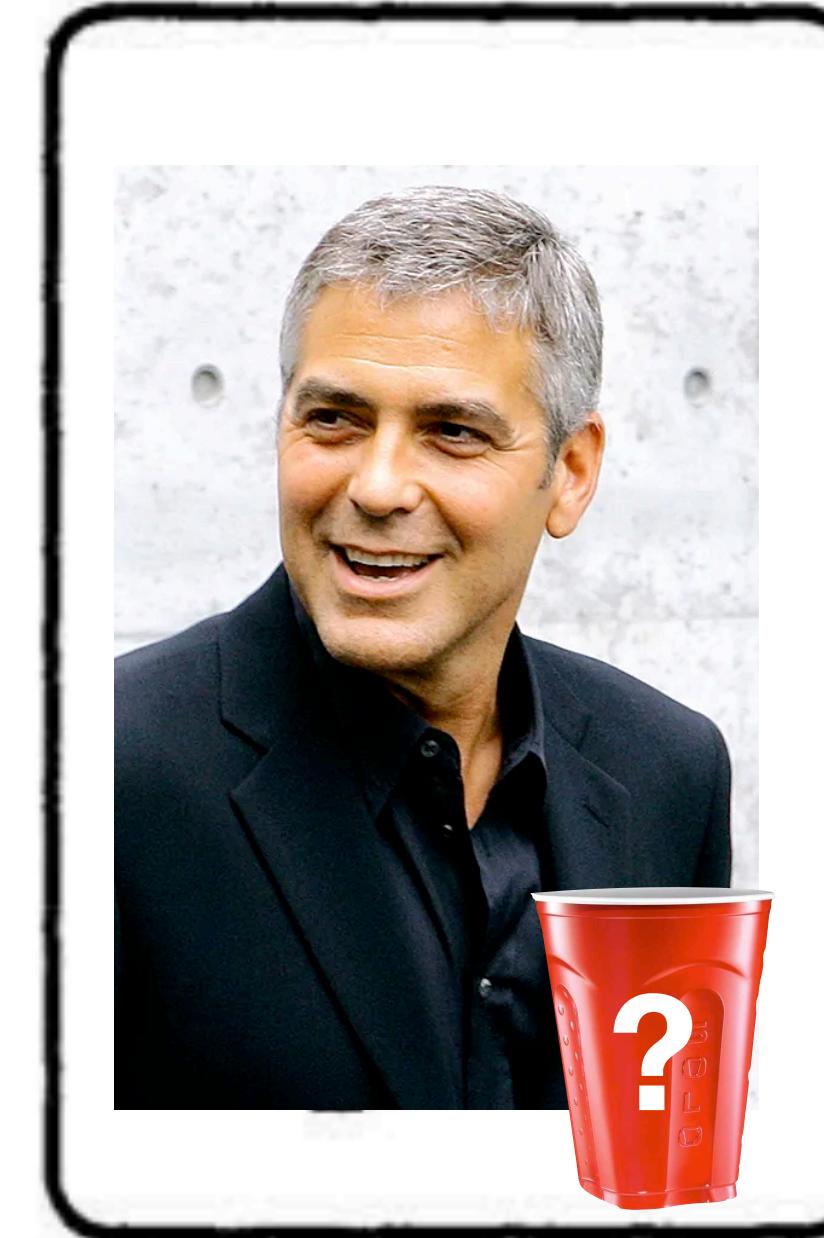
Something seems wrong with this explanation, perhaps this is not how we represent the problem...

Curiously people do far better when the scenario is presented in a less abstract context:

- “You are a bouncer enforcing the rule that if you are under 18 you should not drink alcohol in the club...”
- What do you need to check?

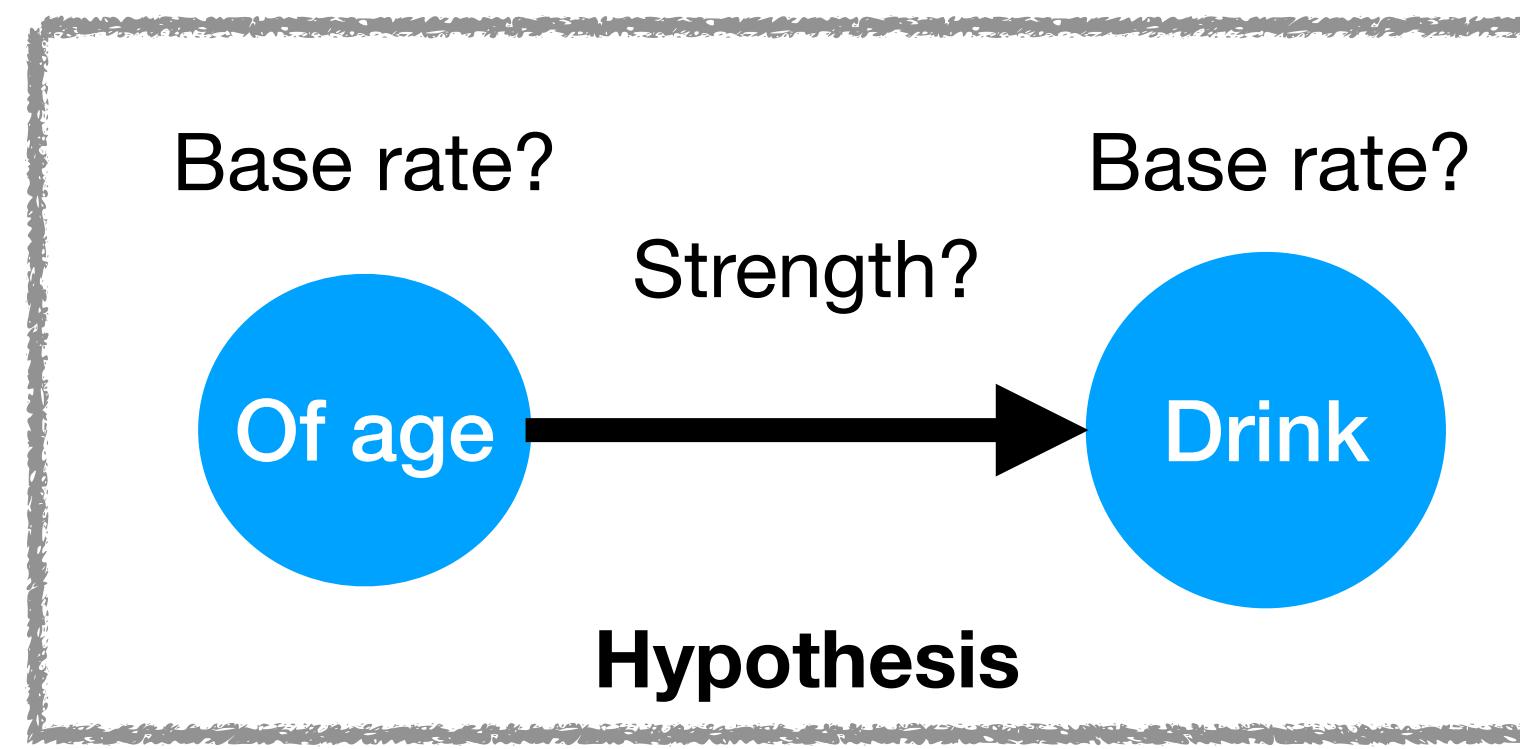
Clearly, you need to check child's cup, and beer drinker's ID

Clooney and the Oasis drinker are irrelevant

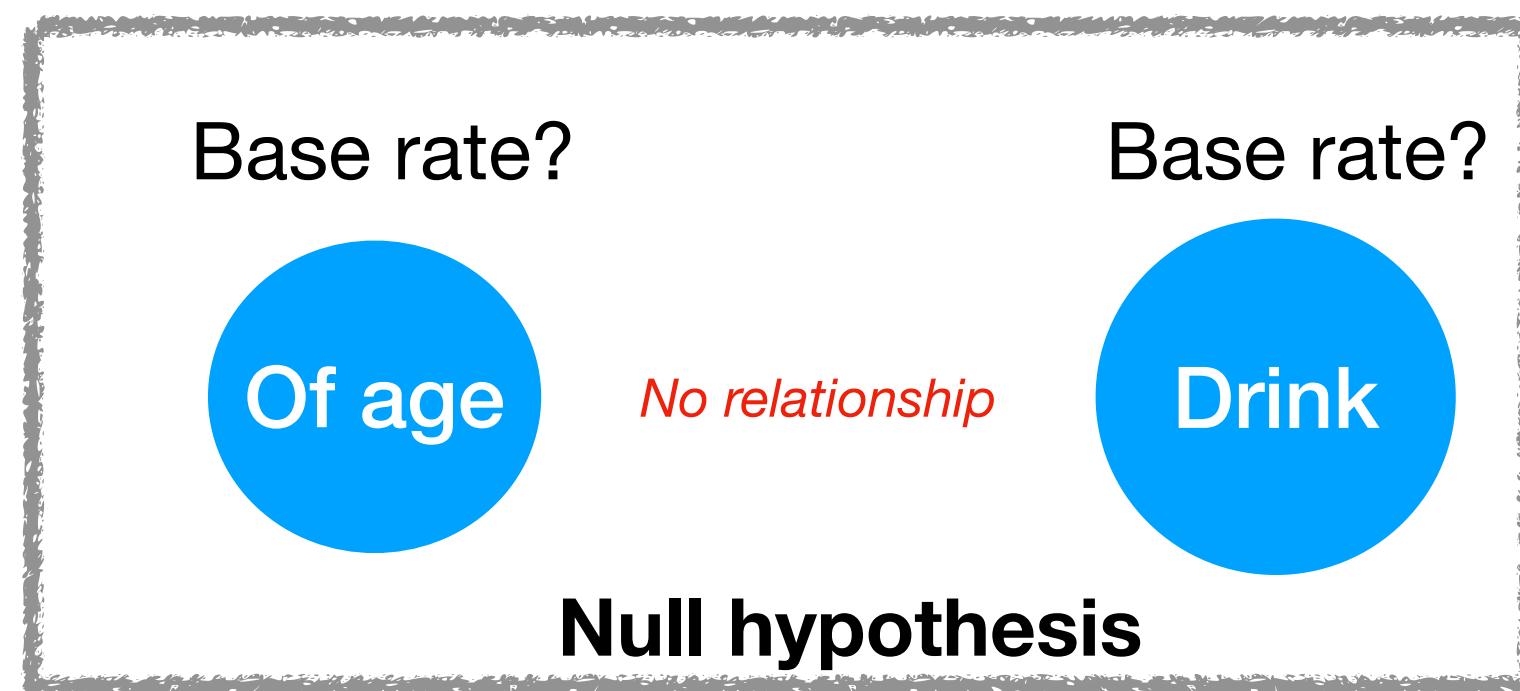


Different representations explain different behaviours

- Johnson-Laird (1983) Maybe people construct a schematic mental model of situation, and this is facilitated by familiarity with variables
- Oaksford & Chater (1994), maybe they represent the variables and relationship probabilistically, making all tests valuable for determining these probabilities
 - Maybe they bring in *priors* i.e. that vowels are rarer than consonants, bars contain more non-drinkers than underagers etc
 - Maybe they use similarity to past experiences and no model at all (Rips, 1994)



vs.



VOL. 55, NO. 4

JULY, 1948

THE PSYCHOLOGICAL REVIEW

COGNITIVE MAPS IN RATS AND MEN¹

BY EDWARD C. TOLMAN

University of California

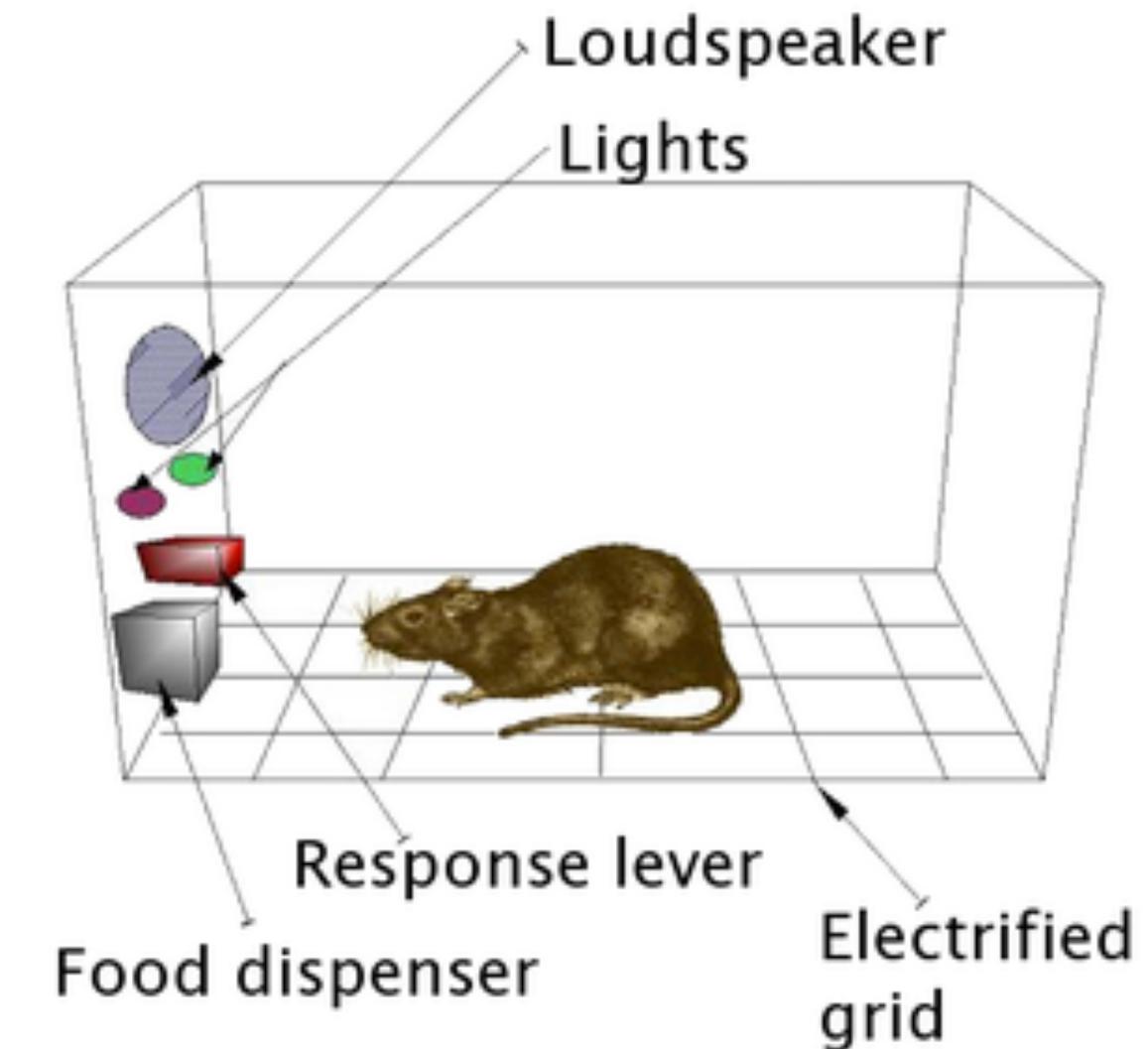
I shall devote the body of this paper to a description of experiments with rats. But I shall also attempt in a few words at the close to indicate the significance of these findings on rats for the clinical behavior of men. Most of the rat investigations, which I shall report, were carried out in the Berkeley laboratory. But I shall also include, occasionally, accounts of the behavior of non-Berkeley rats who obviously have misspent their lives in out-of-

the food box and eats. This is repeated (again in the typical experiment) one trial every 24 hours and the animal tends to make fewer and fewer errors (that is, blind-alley entrances) and to take less and less time between start and goal-box until finally he is entering no blinds at all and running in a very few seconds from start to goal. The results are usually presented in the form of average curves of blind-entrances or of seconds from start to

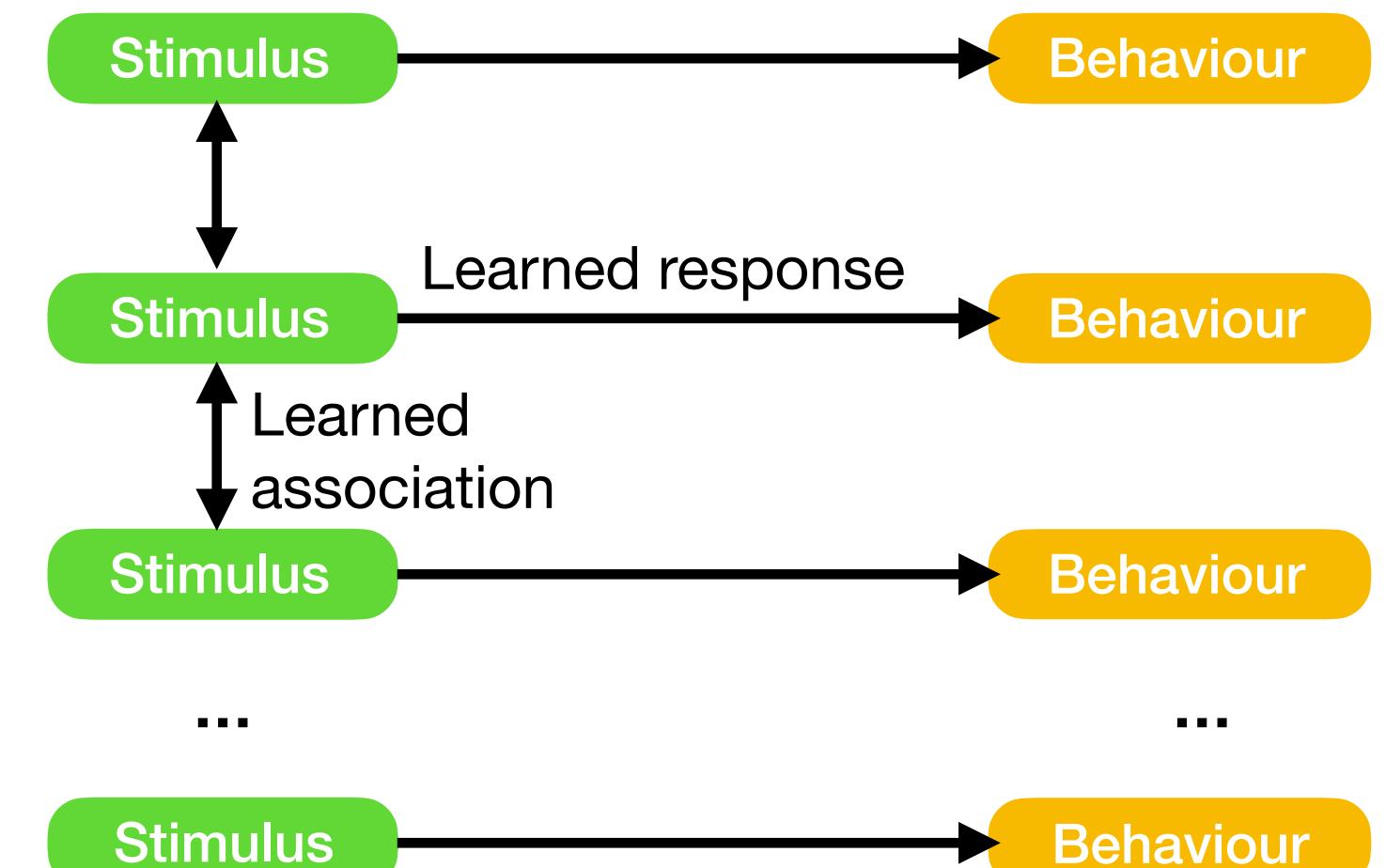
Context

- In the early—mid 20th century, psychology dominated by Behaviourism
 - Behaviours == learned responses to stimuli
 - Often studied animal learning in impoverished environments, documenting reliable associative learning from experiences of reward (e.g. food), or punishment (e.g. shocks) following stimuli (e.g. light flashes), or actions (e.g. lever presses)
 - Implicit assumption that the mind is representationally flat/ empty
 - e.g. no mental models, no mental maps, just lots of stim-stim and stim-action associations

A “Skinner box”, or operant conditioning chamber

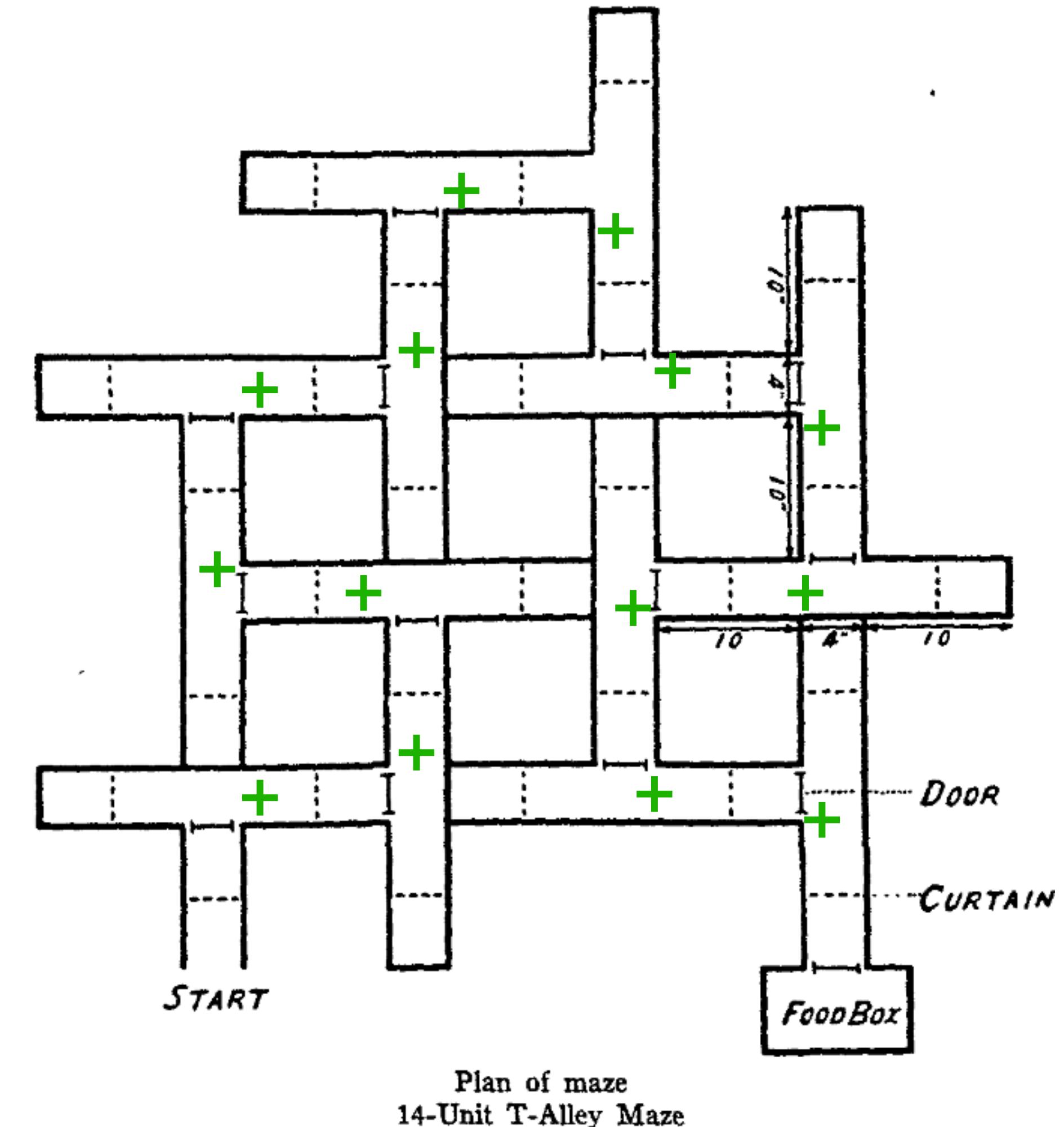


e.g. Edward Thorndike (🐱), Ivan Pavlov (🐶),
John Watson (👶), **B. F. Skinner** (🐦鼠)



Model-free learning

- On behaviourist view, rats trial-and-error way through mazes w/ what we now call “model free reinforcement learning” – forming associations between choices and subsequent rewards
- Over many encounters with a maze these form policy reflecting fastest route to food
- But offer no potential support for alternative goals or alterations to the maze...

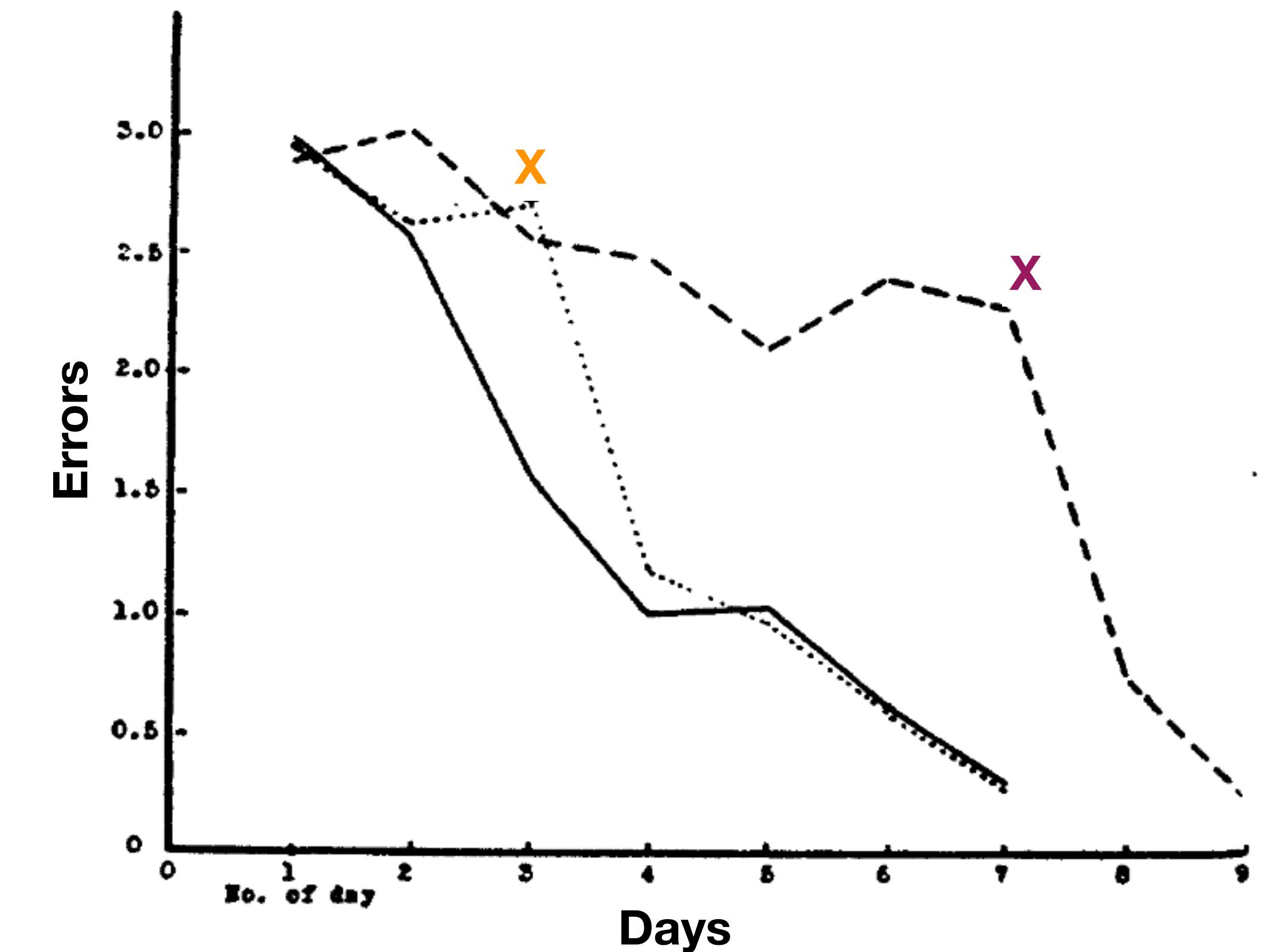
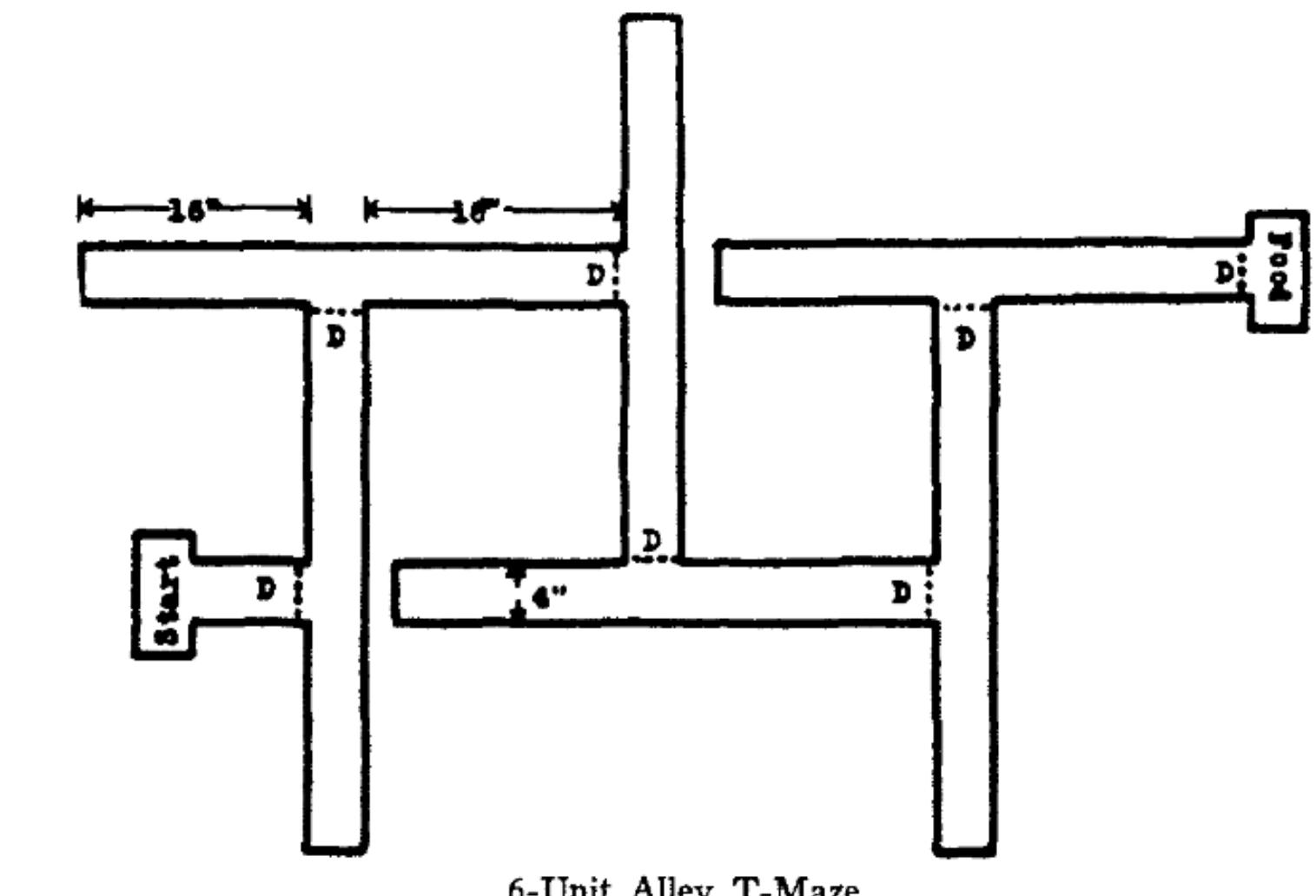


Cognitive maps

- Tolman entertains and tests the alternative idea
- That rats form a mental map-like representation of the maze
- Probes with 5 tasks:
 1. Latent learning
 2. Vicarious Trial and error
 3. Searching for stimulus
 4. Hypotheses
 5. Spatial orientation

1. Latent learning

- Three groups of rats allowed to run around in a maze over 9 days
 - Group 1 (—): There is food from the start. They gradually learn to go to it.
 - Group 2 (---): Food added on day 7 (X) They then learn to go to it quicker than Group 1
 - Group 3 (.....): Food added on day 3 (O) They learn to go to it quickly becoming indistinguishable from group 1 by day 9
- **Suggests they had learned structure of maze spontaneously & in absence of reward!**



2. Vicarious trial and error (aka looking before leaping)

- Contraption set up so rat must jump from platform to one of two doors, either reaching a food platform, or locked door & needing to return to initial platform
- Visible door marking associated with whether unlocked
- Rats turn their head and look at the doors before jumping, doing more looking ahead when the markings are similar
- **Suggests they are playing out the options in their minds before taking them**

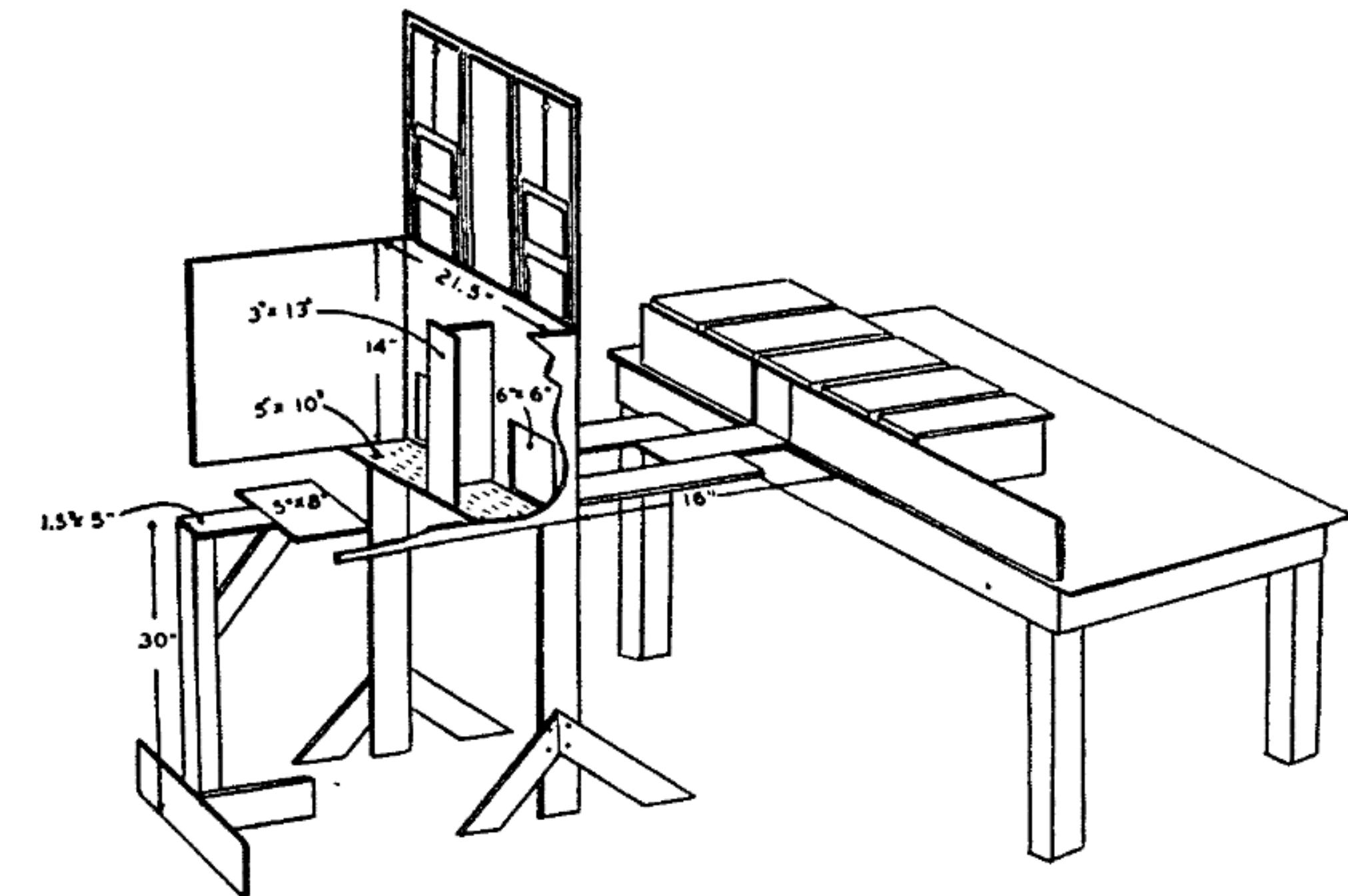


FIG. 9

3. Searching for the stimulus

- Rat put in skinner box with electrified food cup with distinctive pattern on it
- They go for the food, get shocked, and then avoid the pattern for weeks afterward
- But, noted that rats “appeared to search cage” after shock, potentially looking for explanation/cause & alighting on the unusual pattern and “blaming it” for the shock...
- Hypothesis tested by turning light off as they are shocked, and removing pattern. Found that they then did not form an aversion to the pattern
- **Interpretation: They do not just associate blind and serially, but actively and driven by surprise (i.e. need for model revision)**

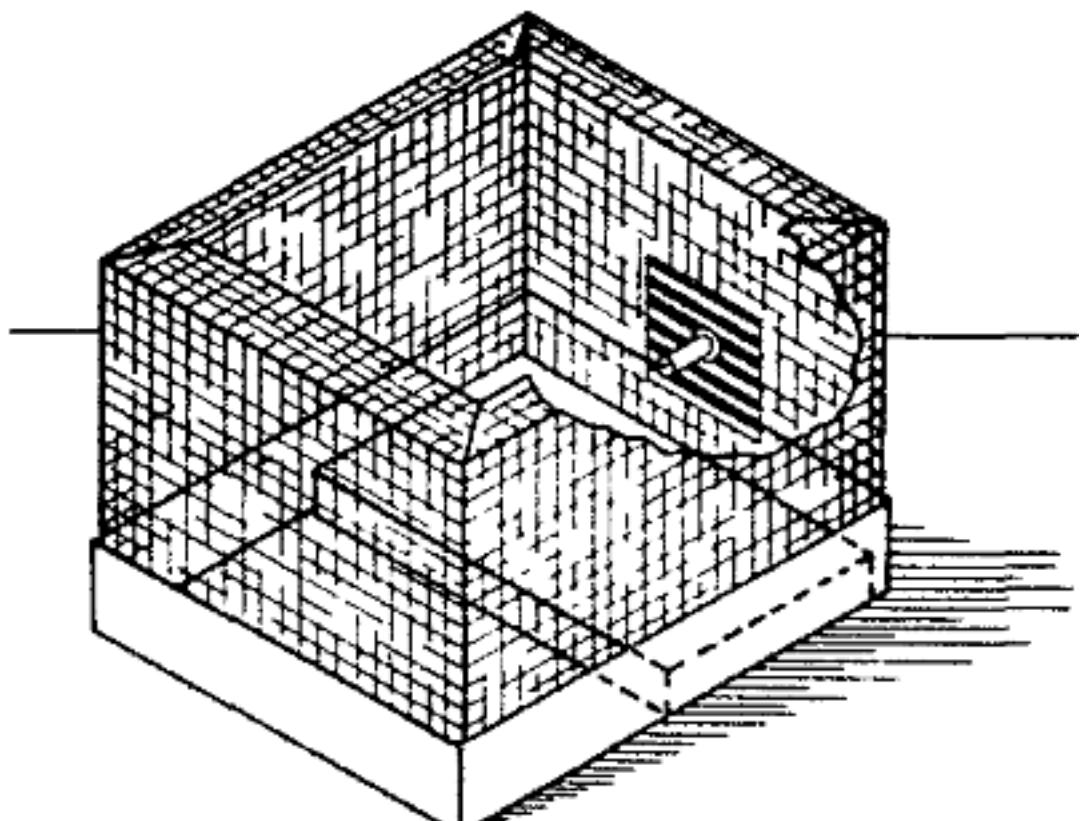
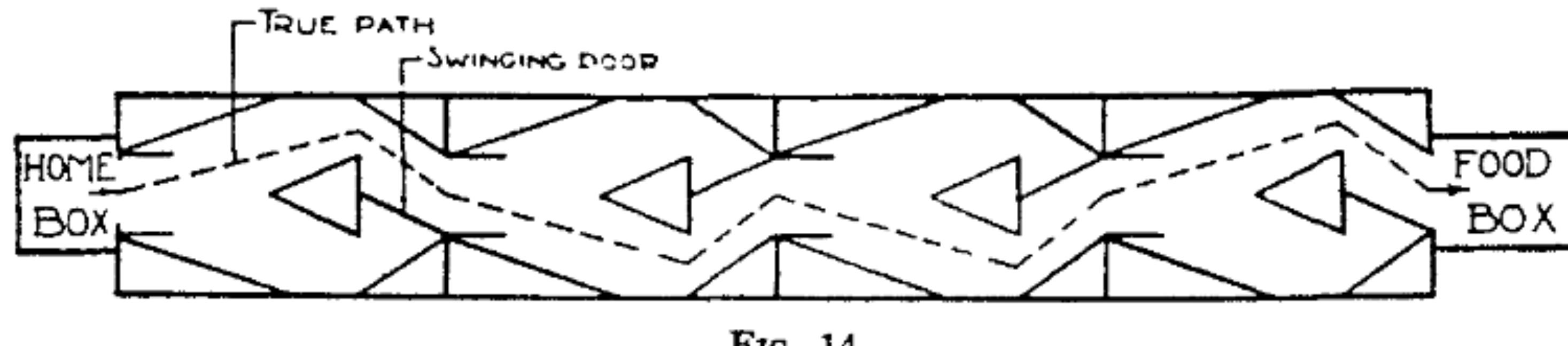


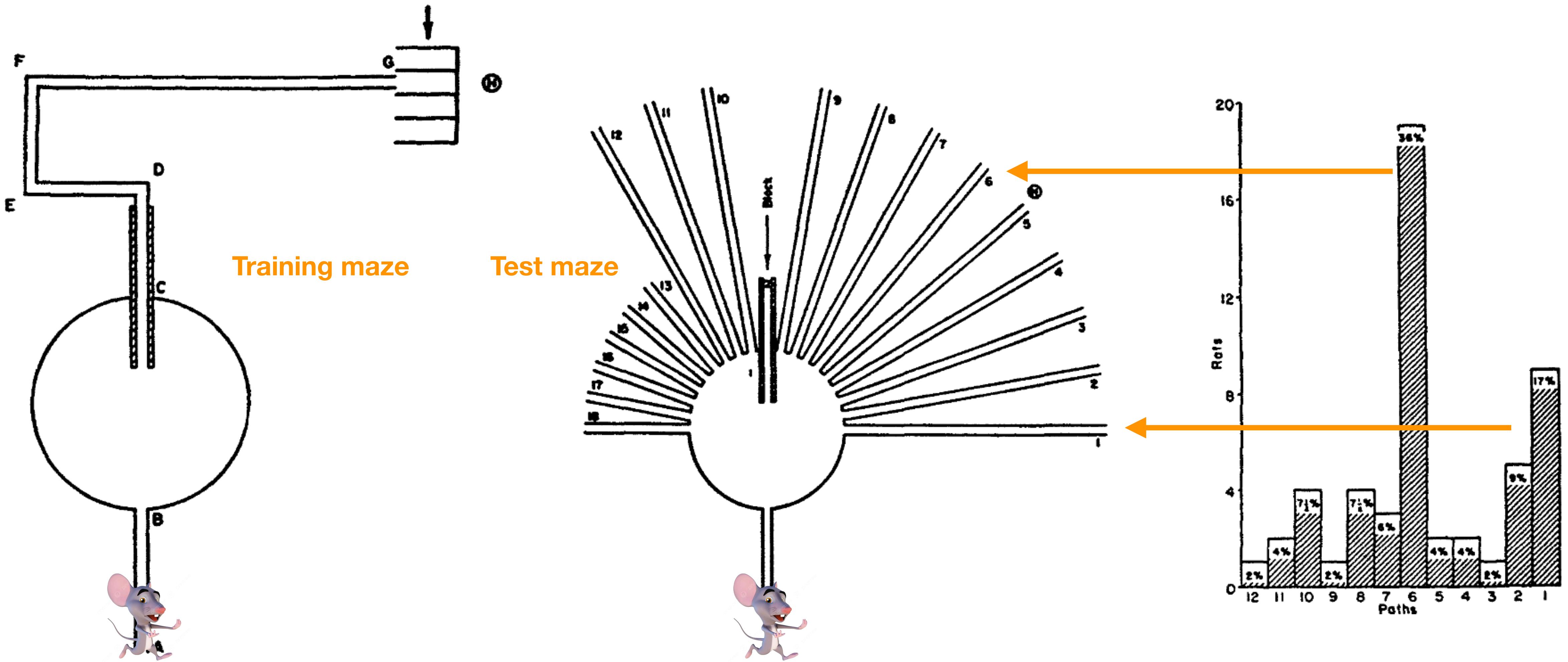
FIG 13

4. Hypothesis testing

- Rats put in 4-compartment discrimination box, i.e. must make sequence of 4 binary choices with two cues “left” vs “right” and “light” vs “dark” to get to end efficiently
- Can be set up to have a simple rule (i.e. stay left, or stay light) or complex rule, or random.
- Rat behaviour found to be statistically significantly “systematic”, doing things like, staying left, or following light, abandoning each behaviour if ineffective, **as if testing hypotheses**



5. Spatial orientation



Suggests the rat's mental maps represent space (allowing for calculation of the direct path)

Is Representation & Process even the right kind of description?

- Van Glender (1995) argues that we might do away with discussion of representations and processes
- & think of minds as dynamical systems
- Description of mechanism that enacts the required computation, seemingly without explicit representation/computation
- Something to think about...

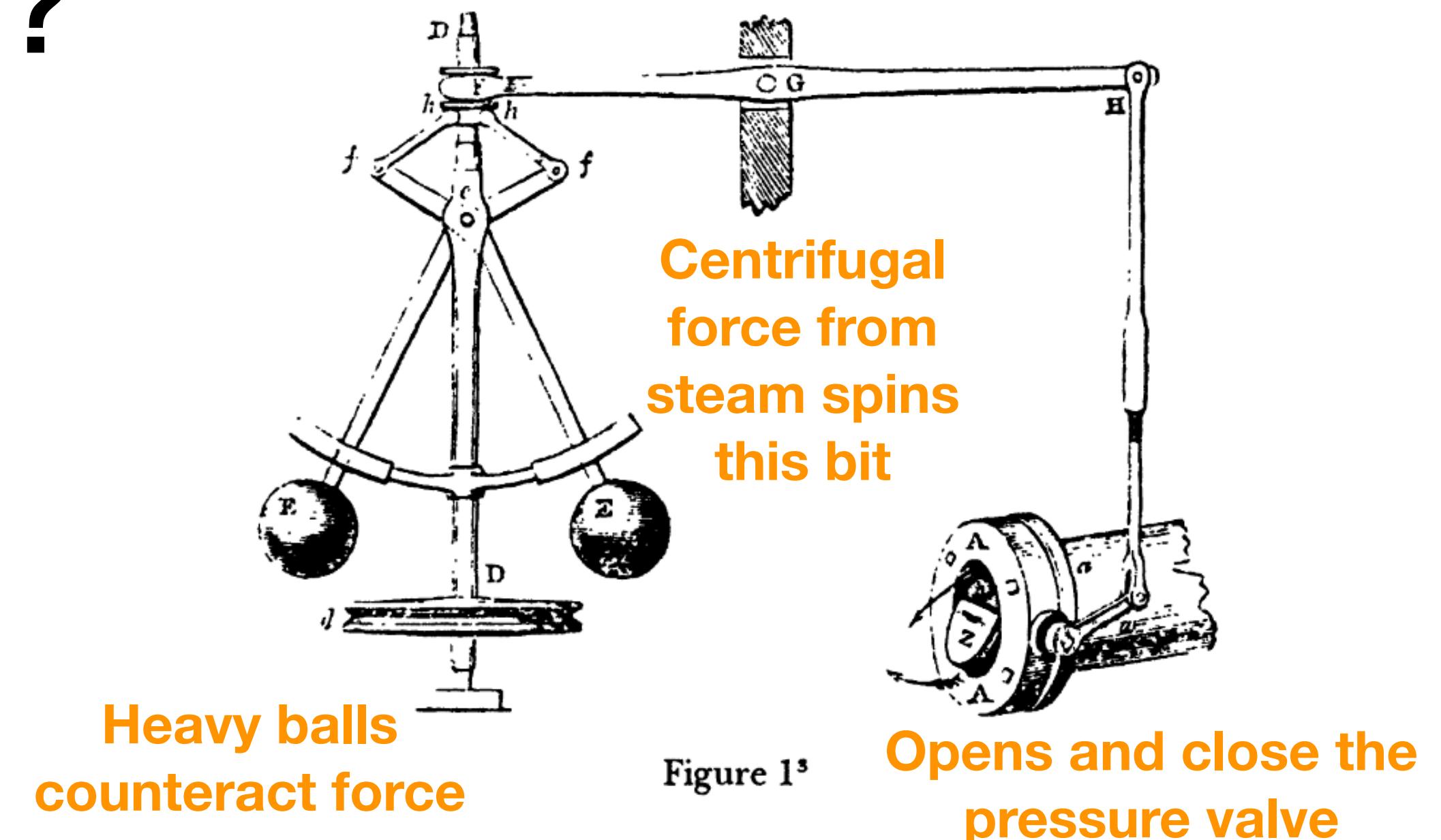
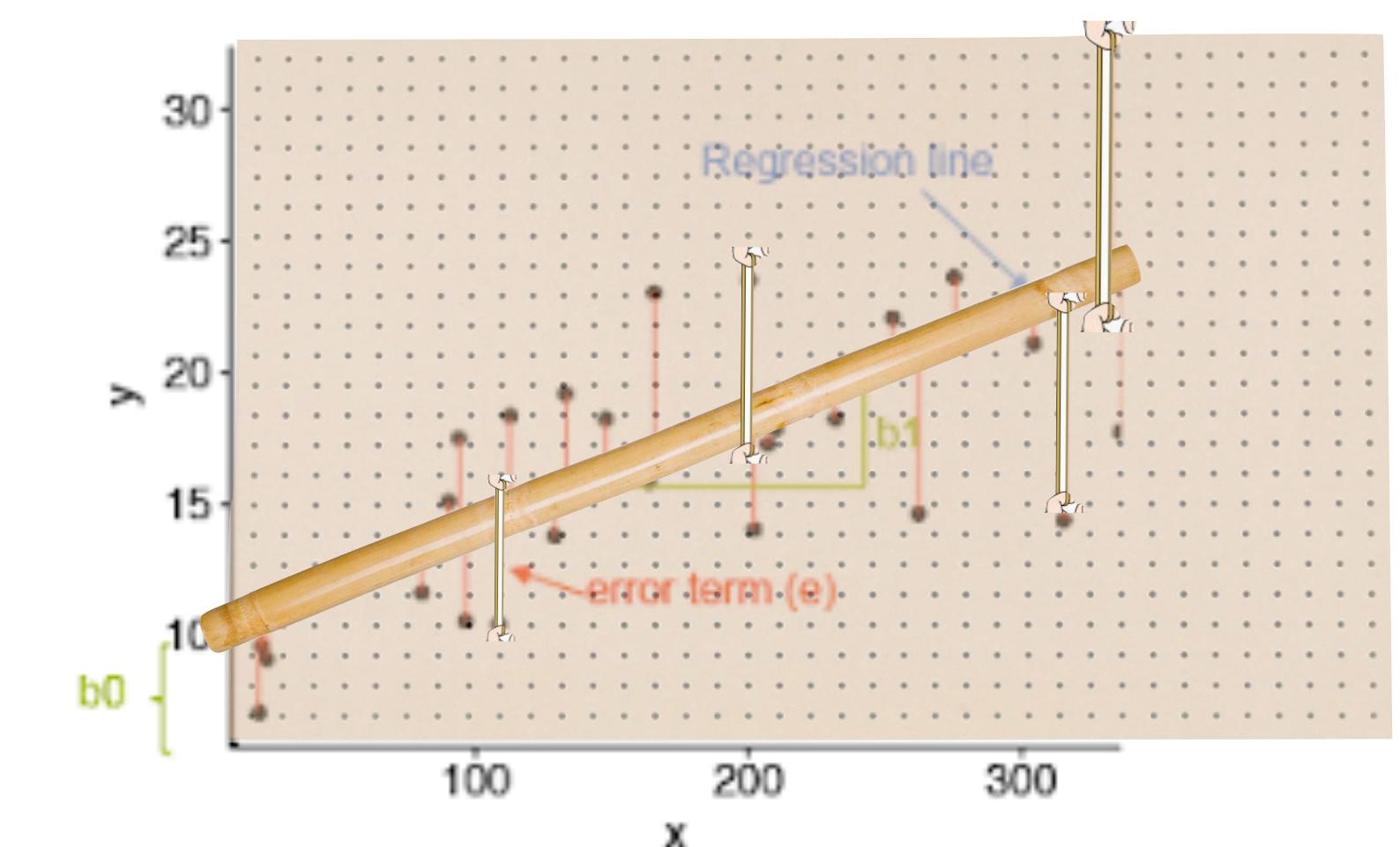


Figure 1³

³ The Watt centrifugal governor for controlling the speed of a steam engine—
from J. Farey, *A Treatise on the Steam Engine: Historical, Practical, and Descriptive*
(London: Longman, Rees, Orme, Brown, and Green, 1827).

$$\hat{\beta} = (X'X)^{-1}X'y$$

OR



Summing up

Knowledge representation chapter 1:

- Representations are “of” stuff in the world
- But abstract away details retaining just what is procedurally useful

Cognitive maps in rats and men:

- Synthesis of maze studies suggesting that even rats must be forming task representations (i.e. cognitive maps), that they do this actively, and that the maps retain spatial information
- This is an early example of cognitivist turn in psychology, toward taking the latent representations inside minds as real and serious objects of study

References

- Braine, M. D., Reiser, B. J., & Ruman, B. (1984). Some empirical justification for a theory of natural propositional logic. In *Psychology of Learning and Motivation*, 18, 313-371.
- Craik, K. (1943). *The Nature of Explanation*, Cambridge University Press.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Harvard University Press.
- Markman, A. B. (1998). *Knowledge Representation*. Psychology Press.
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 27-50.
- Larkin & Simon (1987). Why a Diagram is (Sometimes) Worth 10,000 Words. *Cognitive Science*, 11, 65-99.
- Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, 101(4), 608.
- Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. Mit Press.
- van Gelder, T. (1995) What Might Cognition Be, If Not Computation? *The Journal of Philosophy*, 92(7), 345-381.
- Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content* (Vol. 86). Harvard University Press.

Discussion...

- Thoughts about Wason card task?
- Comments on Tolman's experiments...
- ...and about representation idea generally:
 - What *are* representations?
 - What do ours look like?
 - Do we need them?
- Where have we got with these questions since Tolman paper?