Navigating uncertainty through information search

Charley M. Wu[1] (cwu@mpib-berlin.mpg.de), Björn Meder[2], & Jonathan D. Nelson[2,3]

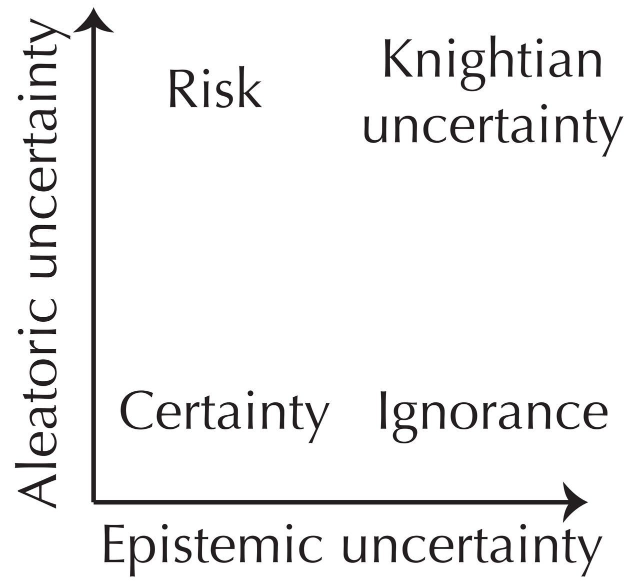
[1] Center for Adaptive Rationality, Max Planck Institute for Human Development; [2] MPRG iSearch, Max Planck Institute for Human Development; [3] School of Psychology, University of Surrey, Guilford

Max-Planck-Institut für Bildungsforschung
Max Planck Institute for Human Development

Introduction. Two definitions of learning

i) Reducing *"Aleatoric uncertainty"* by querying the environment

- Information Search paradigm
- Learning as conditioning beliefs on new evidence acquired from performing a medical test, conducting an experiment, or asking a question

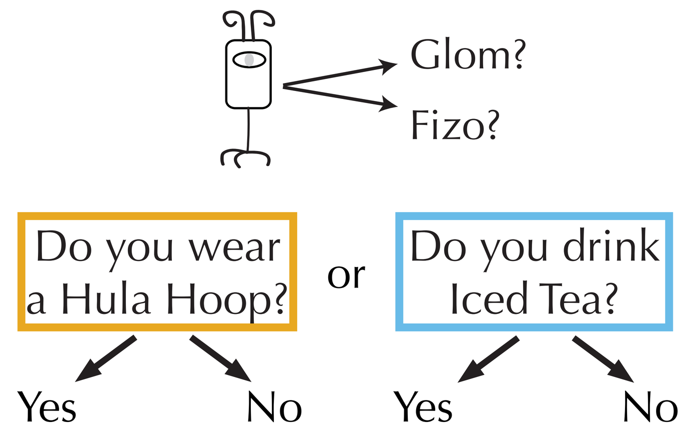


ii) Reducing *"Epistemic uncertainty"* by observing more samples

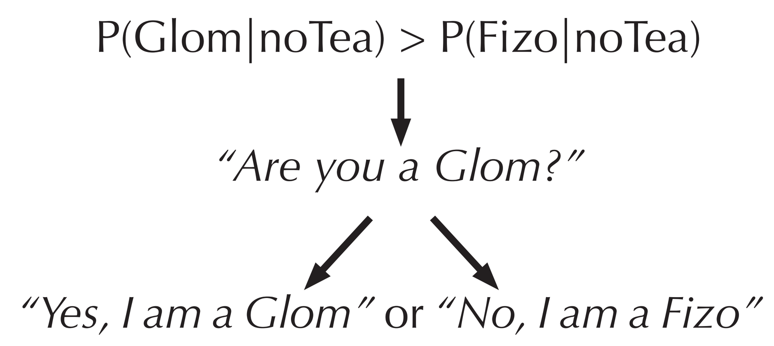
- Reinforcement Learning (RL) paradigm
- Learning as acquiring more observations of a stochastic event (e.g., coin flips) in order to estimate the true probabilities

Task. Informational Bandit Problem

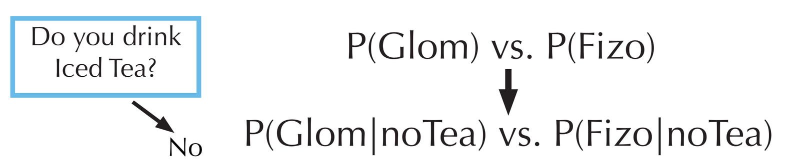
1. Select a query to aid the classification of an alien:



3. Classify alien based on most likely class:

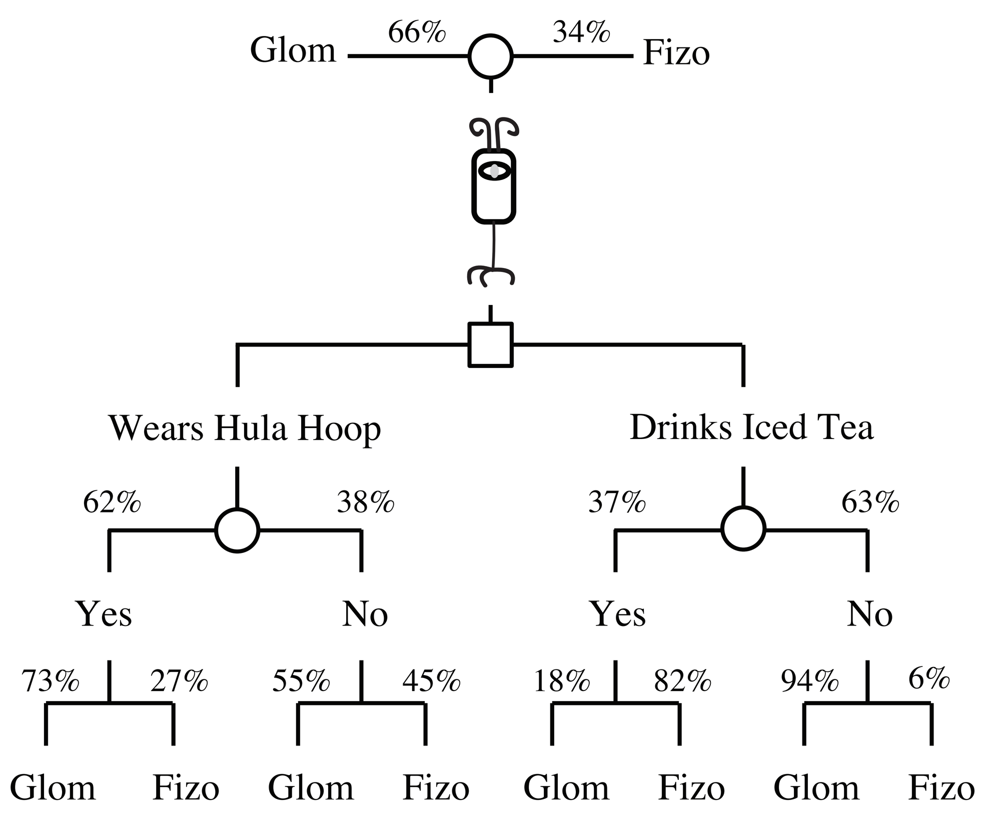


2. Observe outcome & condition beliefs:

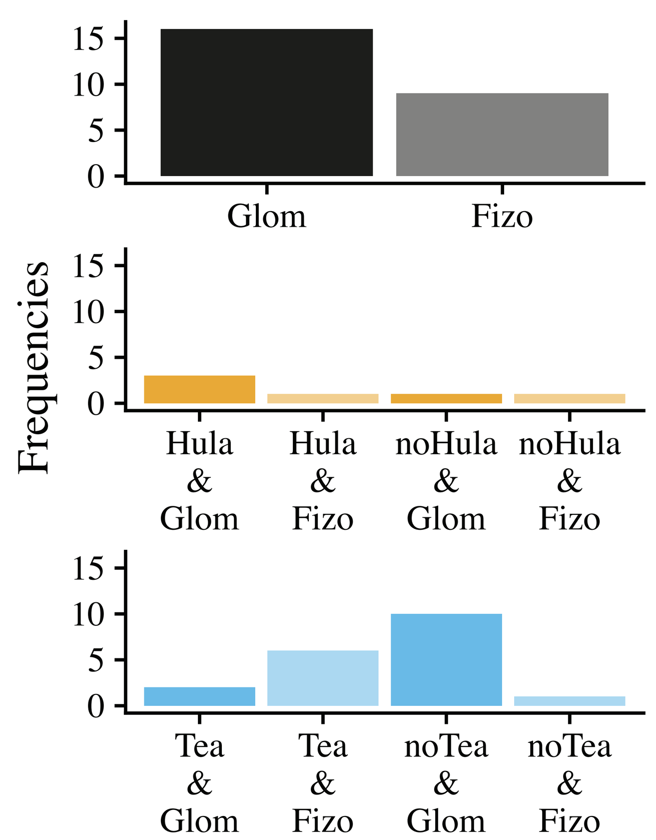


4. Update memory/beliefs:

True Environmental Structure



Observations at t=25



Models of Query Selection.

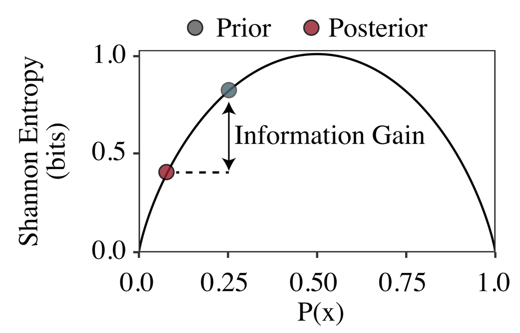
Optimal Experimental Design (OED) models:

- Choose query with highest expected usefulness EU(Q) based on observed probabilities
- No account of epistemic uncertainty from small samples

$$EU(Q) = \sum_{q \in Q} P(q)U(q)$$

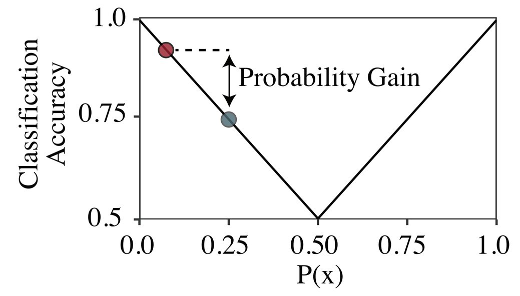
Information Gain

$$U_{InfoGain}(q) = \underbrace{H\left[P(C)\right]}_{\text{Prior entropy}} - \underbrace{H\left[P(C|q)\right]}_{\text{Posterior entropy}}$$



Probability Gain

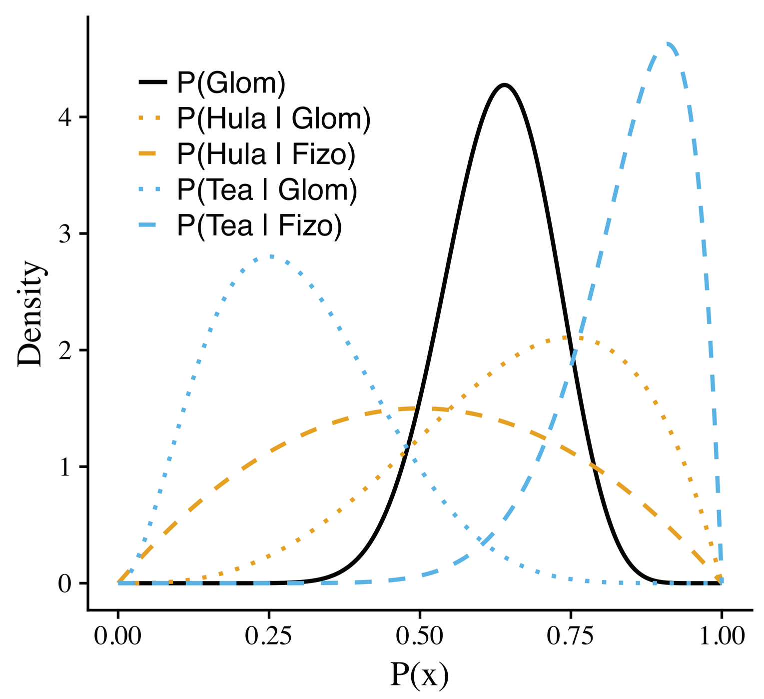
$$U_{ProbGain}(q) = \underbrace{\max P(c|q)}_{\text{P(correct)}_{posterior}} - \underbrace{\max P(c)}_{\text{P(correct)}_{prior}}$$



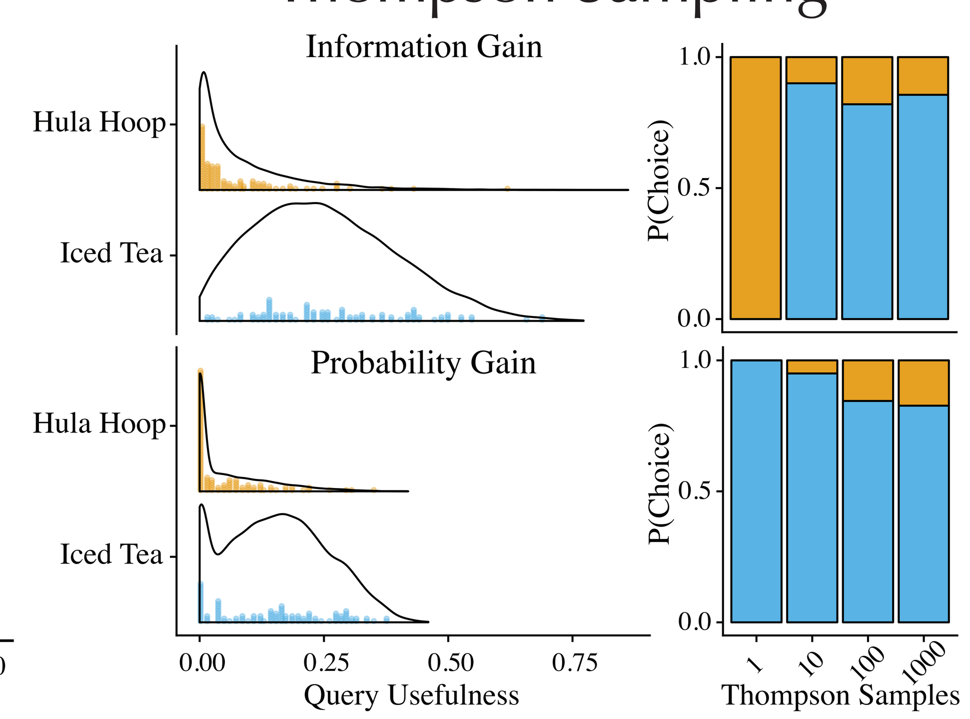
OED with Bayesian Sampling:

- Model probabilities with beta distributions based on observed frequencies, where variance corresponse to epistemic uncertainty
- Thompson sampling for selecting queries (balancing exploration-exploitation)

Beta Distributions



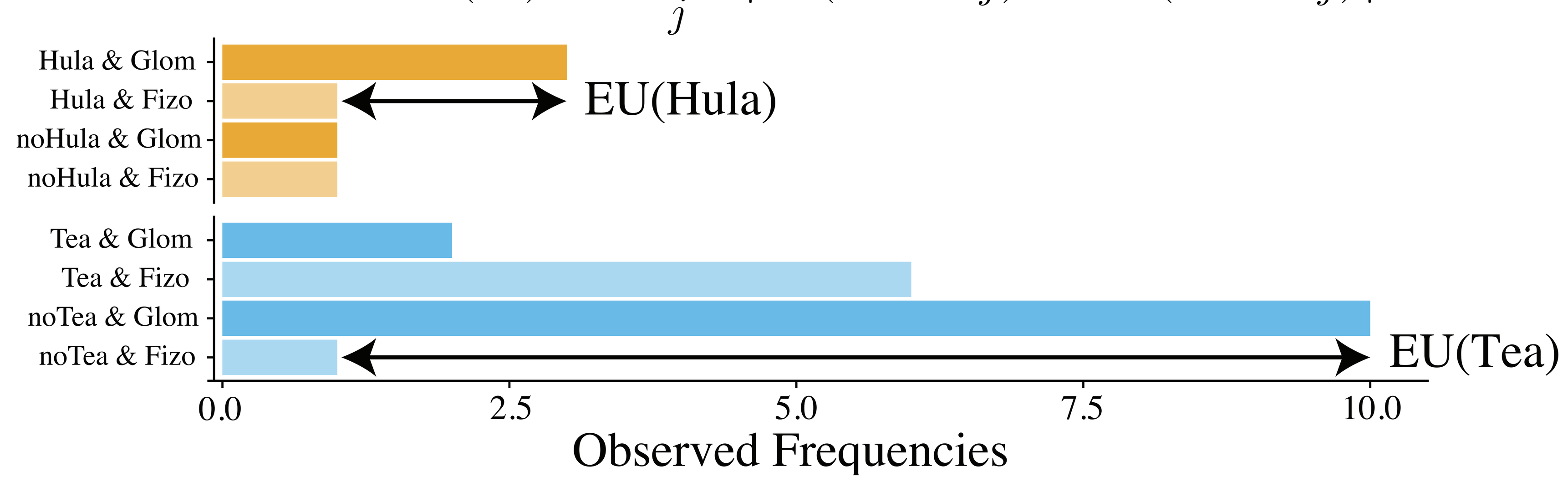
Thompson Sampling



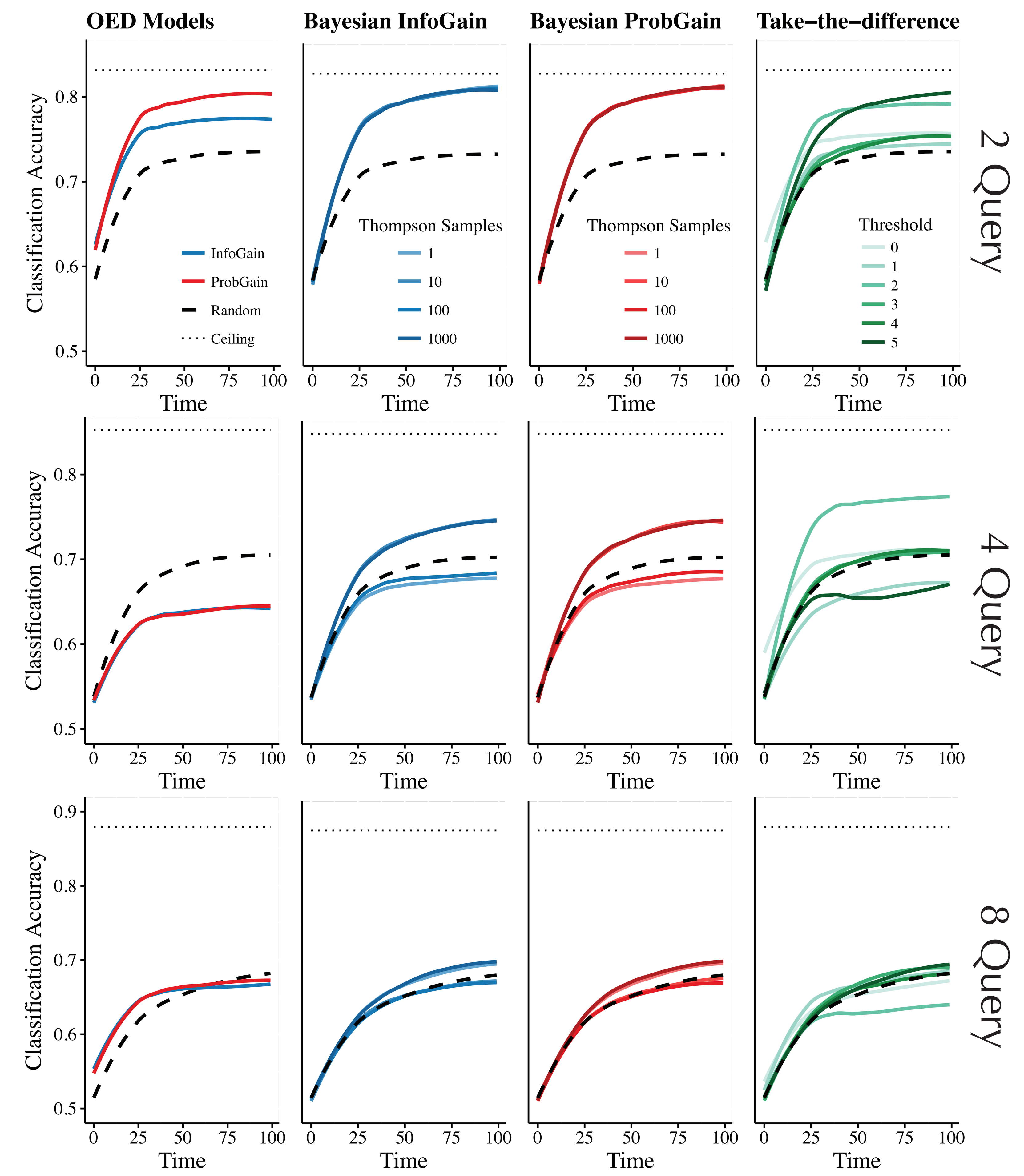
Take-the-difference (TTD) Heuristic:

- Represents knowledge about the world using natural frequencies rather than probabilities
- In the limit of epistemic certainty, never disagrees with probability gain (Wu, Meder, Filimon & Nelson, *JEP:LMC* 2017)

$$EU_{TTD}(Q) = \max_j \left|N(c_1 \wedge q_j) - N(c_2 \wedge q_j)\right|$$



Simulation Results.



*Each n-query simulation consisted of 40k randomly generated environments. In some environments, query selection has very little impact on classification accuracy, so here we present results from the quartile of environments where the difference between tests are largest.

Conclusions.

1. Standard OED models fail to adequately explore, and sometimes perform worse than random query selection

2. Bayesian sampling models almost always outperform standard OED models by representing epistemic uncertainty and balancing the explore-exploit dilemma

3. TTD is incredibly efficient in early stages of learning, and can sometimes even outperform the more complex Bayesian models in the late stages of the task