Summary
How do we solve this explore-exploit dilemma in the domains of economic gains vs. losses?

Two strategies used equally for gains and losses:
- Directed exploration – choose unknown options
- Random exploration – choose randomly

A bias in the domain of losses:
- Uncertainty seeking – choose unknown option independent of exploration

Horizon task with losses

- First four trials are “forced” to give unbiased information about each bandit
- Four conditions:
  - gains: horizon 1 vs. horizon 6
  - losses: horizon 1 vs. horizon 6

Performance

Results

Choice curves

- Unequal information [1 3]
- Equal information [2 2]

- Difference in means between more and less informative options
- Difference in means between right and left options

Basic model: parameter fits

- Information bonus [points]
- Decision noise [points]

Uncertain seeking with losses

- Sum of information bonus for horizon 6 & 1
- Difference in information bonus between horizon 6 & 1

Cognitive model

- Uncertainty preference [points]
- Difference in uncertainty preference between horizon 6 & 1

Model

Basic model

- \( Q_a = R_a + \alpha I_a + Bs_a \) value of option \( a \)

Cognitive model

- \( Q_a = R_a + \alpha I_a + \beta U_a + Bs_a \)

Values

- \( R_a \) mean reward
- \( I_a \) information (-1, 0, or +1)
- \( s_a \) spatial location (-1 or +1)
- \( \alpha \) information bonus
- \( B \) spatial bias
- \( \beta \) uncertainty preference

Choice probabilities

- \( p_a = \frac{1}{1 + \exp\left(\frac{Q_d - Q_a}{\sigma_d}\right)} \)