Feedback Promotes Learning and Knowledge of the Distribution of Values Hinders Exploration in an Optimal Stopping Task

Erin H. Bugbee and Cleotilde Gonzalez
{ebugbee, coty}@cmu.edu
Department of Social and Decision Sciences, Carnegie Mellon University

Summary

In many naturalistic situations such as deciding on an apartment to rent or selecting a life partner, people explore options before making a selection. We investigate the following:

How do **feedback** and **knowledge of the distribution of options values** affect **learning** in sequential search?

How do people **deviate from optimal** based on these factors, and can this be modeled with a **cognitive model** of decisions from experience?

Introduction

- Previous work indicates that people are suboptimal at stopping exploration and often stop earlier than optimal [1, 5]
- Recent work shows that people can learn to stop at the optimal time with experience [3]
- We investigate factors that may influence learning in stopping decisions and extend our previous modeling work [2] to this task

Methods

Human Participants

- *Participants*: 226 from Amazon MTurk
- *Design*: Between-subjects
- *Conditions*: 2 (Knowledge of the Distribution: Known or Unknown) x 3 (Feedback: Outcome, Detailed, or No Feedback)

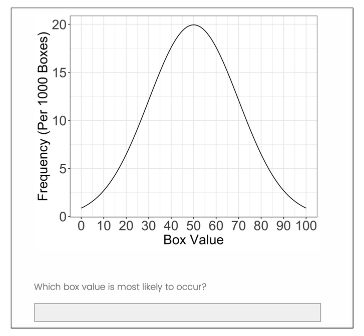

Figure 1: Known Distribution

Model Agents

- Simulate 256 IBL model and optimal agents

Optimal Stopping Task

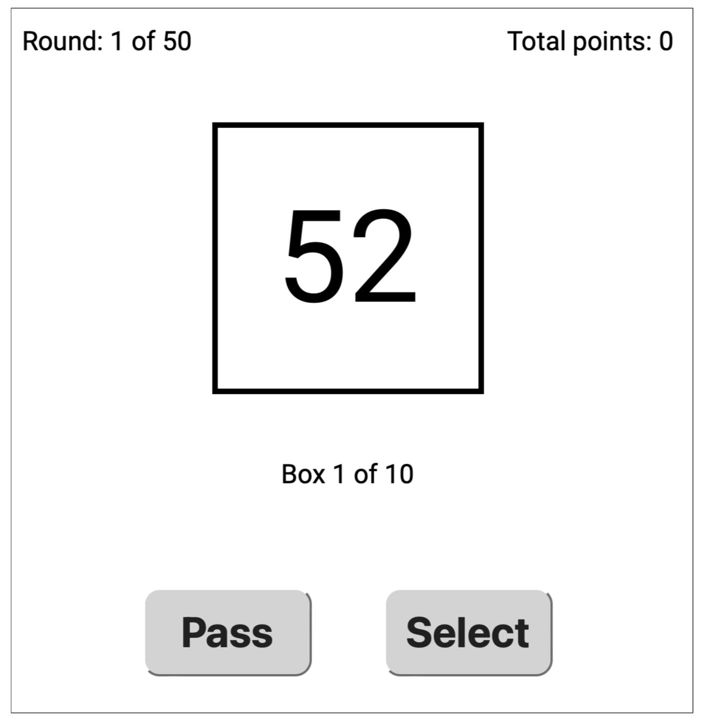

Figure 2: Choice

Outcome Feedback:
Your choice was [Correct/Wrong]. You chose **Box** [**Number**] with a value of [**Value**].
Detailed Feedback:
Your choice was [Correct/Wrong]. You chose **Box** [**Number**] with a value of [**Value**]. The maximum value box was **Box** [**Number**] with a value of [**Value**].
No Feedback:
You chose **Box** [**Number**] with a value of [**Value**].

Cognitive Model: Instance-Based Learning Model

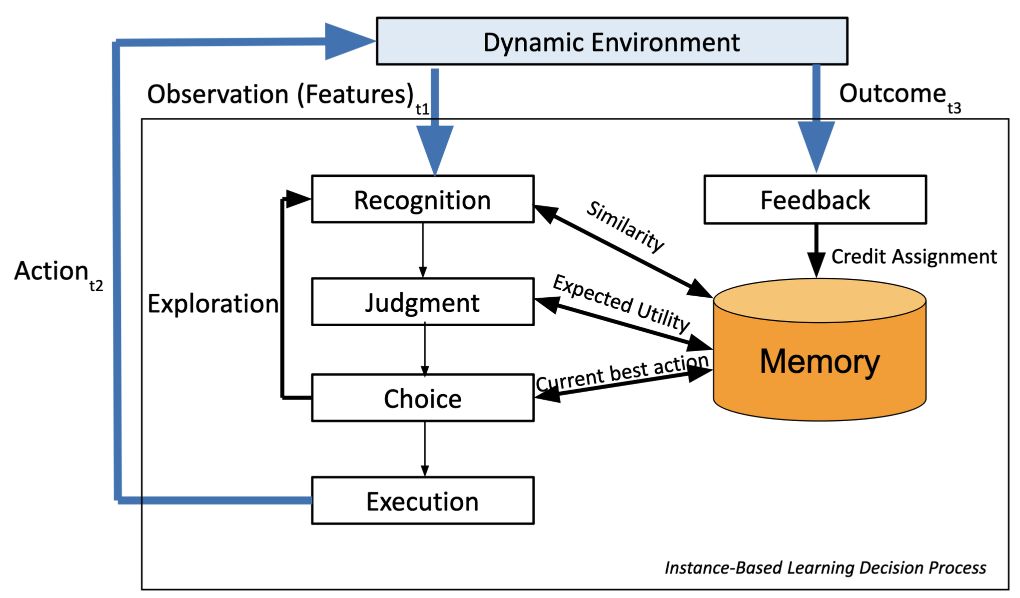

Figure 3: Instance-Based Learning Theory [4]

- Learning occurs through the accumulation of memory units called *instances*
- Past instances are *retrieved* based on *similarity* to the current situation, frequency, and recency
- A *blended value* (BV) is calculated based on the *utility* of the retrieved instances
- The agent *chooses* option with the highest BV

State		Action	Utility
Value	Boxes Remaining	{Select, Pass}	{0, 1}

Table 1: Instance Structure

Results

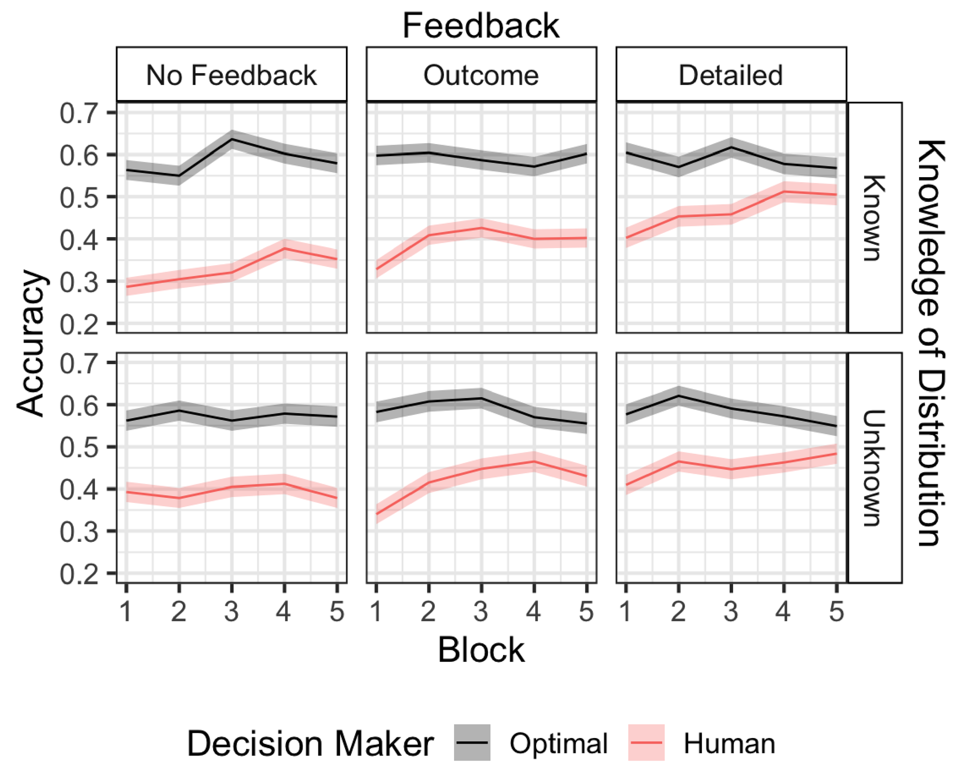

Figure 4: Average Reward for Human Participants and Optimal Agents over Blocks of 10 Problems

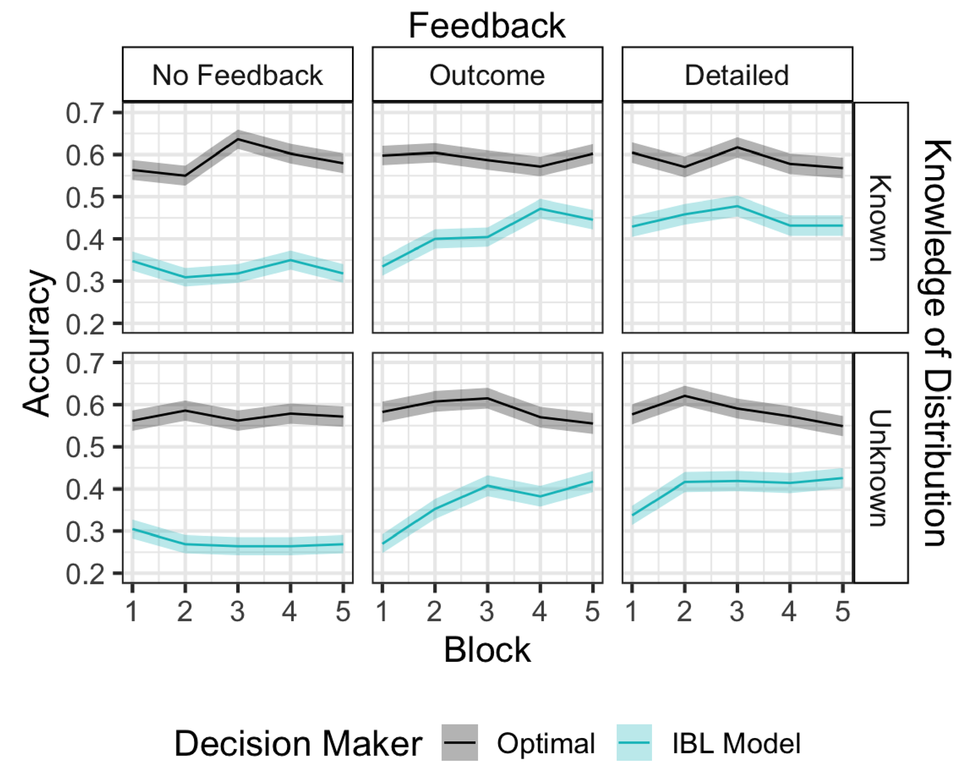

Figure 5: Average Reward for IBL and Optimal Agents over Blocks of 10 Problems

Conclusion

- We develop a **novel optimal stopping task** to investigate the impact of various factors on stopping decisions in sequential search
- Humans and IBL model agents have **lower thresholds (stop earlier) than optimal**
- People **learn more when provided feedback** relative to no feedback
- **Knowledge** of the distribution **hinders exploration**
- ***Future Work***: Investigate additional factors (variability of sequence length, crowd decisions)

References

[1] Baumann, C., Singmann, H., Gershman, S. J., and von Helversen, B. A linear threshold model for optimal stopping behavior. *Proceedings of the National Academy of Sciences 117*, 23 (June 2020), 12750–12755.

[2] Bugbee, E., and Gonzalez, C. Deciding When to Stop: Cognitive Models of Sequential Decisions in Optimal Stopping Tasks. Publisher: Carnegie Mellon University. Under review.

[3] Goldstein, D. G., McAfee, R. P., Suri, S., and Wright, J. R. Learning When to Stop Searching. *Management Science 66*, 3 (Mar. 2020), 1375–1394.

[4] Gonzalez, C., Lerch, J. F., and Lebiere, C. Instance-based learning in dynamic decision making. *Cognitive Science 27* (2003).

[5] Guan, M., Stokes, R., Vandekerckhove, J., and Lee, M. D. A cognitive modeling analysis of risk in sequential choice tasks. *Judgment and Decision Making 15*, 5 (Sept. 2020), 823–850.

Contact Information

- **Email:** ebugbee@cmu.edu
- **Lab Website:** cmu.edu/ddmlab
- **Personal Website:** erinbugbee.com