The Predictability of Task-unrelated Action Outcome Modulates Reinforcement Learning

Takumi Tanaka and Hideaki Kawabata

Keio university, Japan.

Background

Goal-directed action is driven by anticipations of the (intended) action outcomes

(e.g., rewards or other incentives that are worth the effort)

This depends on “bidirectional association” between actions and outcomes

(cf., Effects of stimulus–response compatibility)

Recent studies have shown that

* The predictability of action outcome improved the motivation and performance even when these outcomes were not related to the goal of the task

Goal

Why do the task-unrelated action outcomes modulate action?

Previous studies tell us two possibility

* Human implicitly judge whether an action outcome is expected or not.
* The accuracy of action outcome prediction modulates processing of reward associated with specific action or stimulus eliciting it.

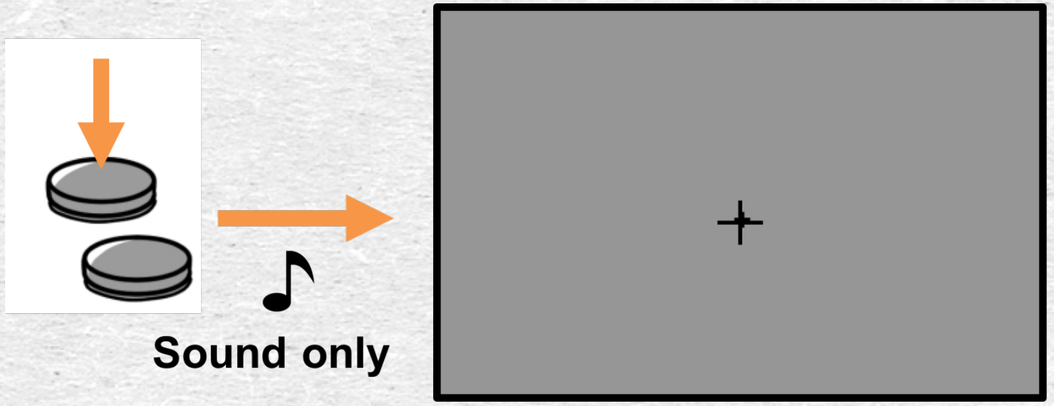
✔ We investigated the effect of predictability of task-unrelated action outcome on reward processing more directly using reinforcement learning.

Experiment (N=24, 10 female, mean age = 20.12 ± 1.13)

Action-Outcome Association Task

Right/Left key press caused a **Low/High** tone

The specific action-outcome relations were fixed throughout all 200 trials

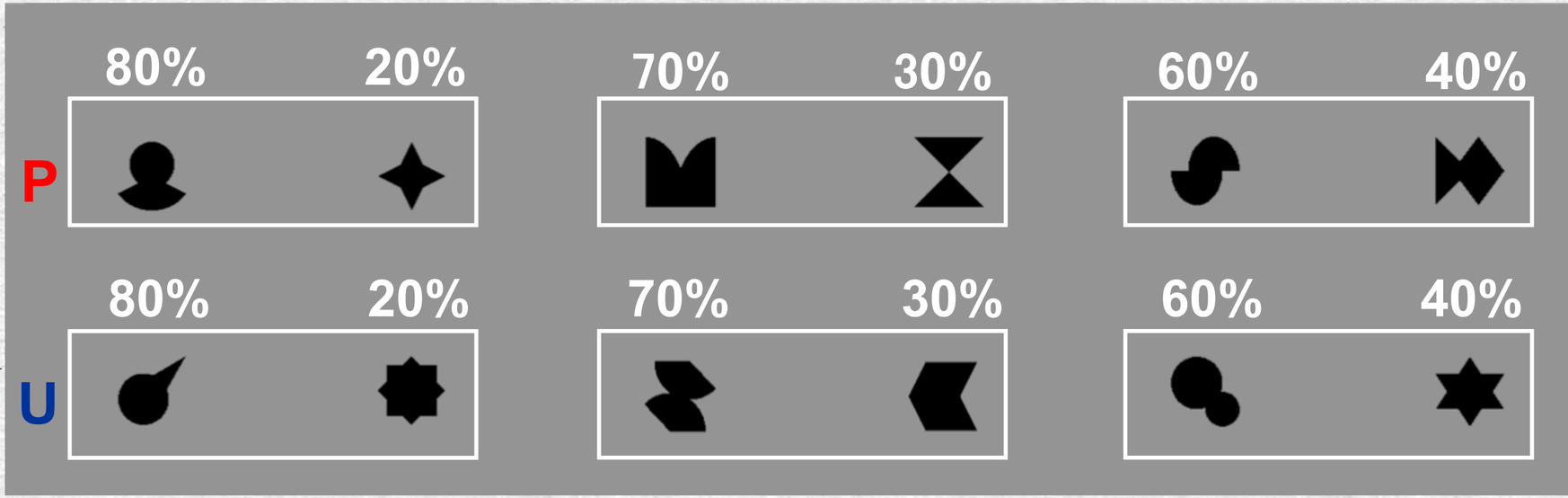


Reinforcement Learning

6 steps of **reward** probability (20-80%)

×

2 conditions of **predictability** (**P**redictable vs. **U**npredictable)



Reinforcement Learning -Learning phase

One out of six figure pairs was presented

Participants Chose one by the same keys as those in association task

➡□ ***Visual feedback***

Reward / Punishment = Green / Red circle

+

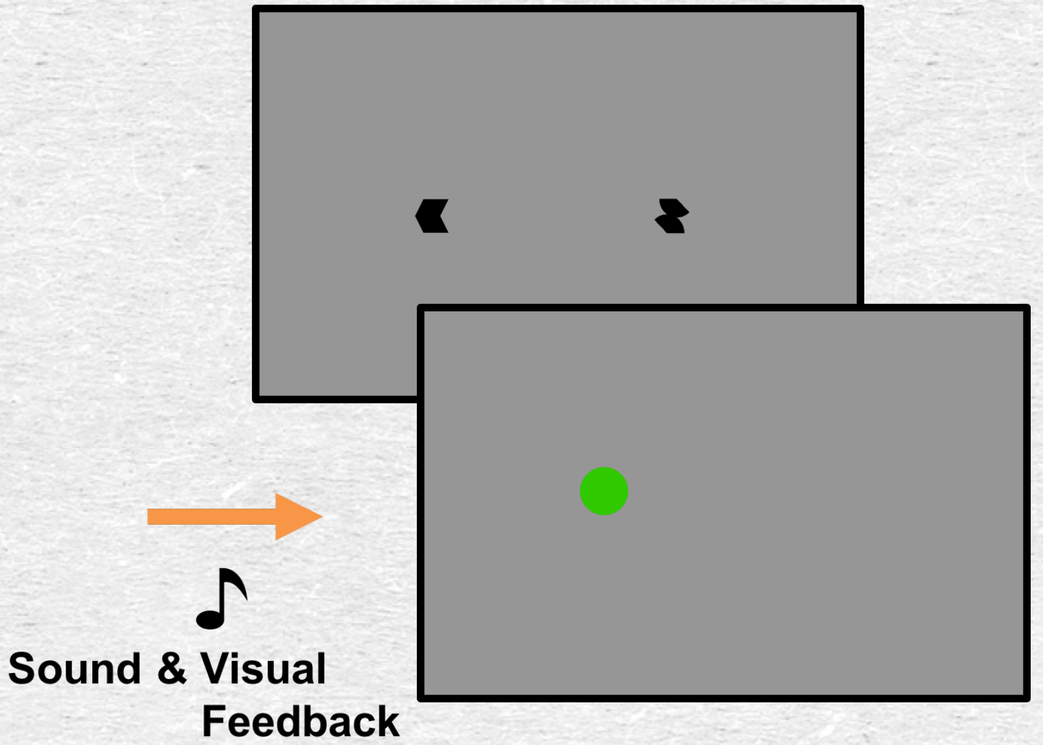
Auditory feedback

- In the choice from **Predictable Pairs**

  Associated (Low/High) tone in every trial (**100%**)

- In the choice from **Unpredictable Pairs**

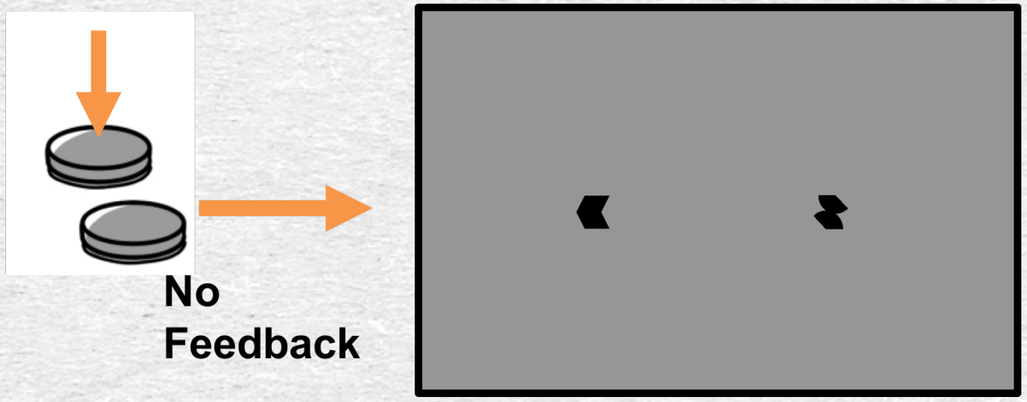
  Associated (**65%**) or ***NOT*** associated tone (= the tone associated with another key; **35%**)



Reinforcement Learning -Test phase

All possible option pairings ($_{12}C_2 = 66$) × **2**

No Visual & Auditory feedback



The combinations of the key and associated tone

of the colors and reward/punishment

Reward probability and predictability of each stimulus

] Counterbalanced

The number of trials in learning phase varied (240-360) by each participant's performance

Results

The analysis of **paired comparison with Bradley-Terry method** for choice in the test phase

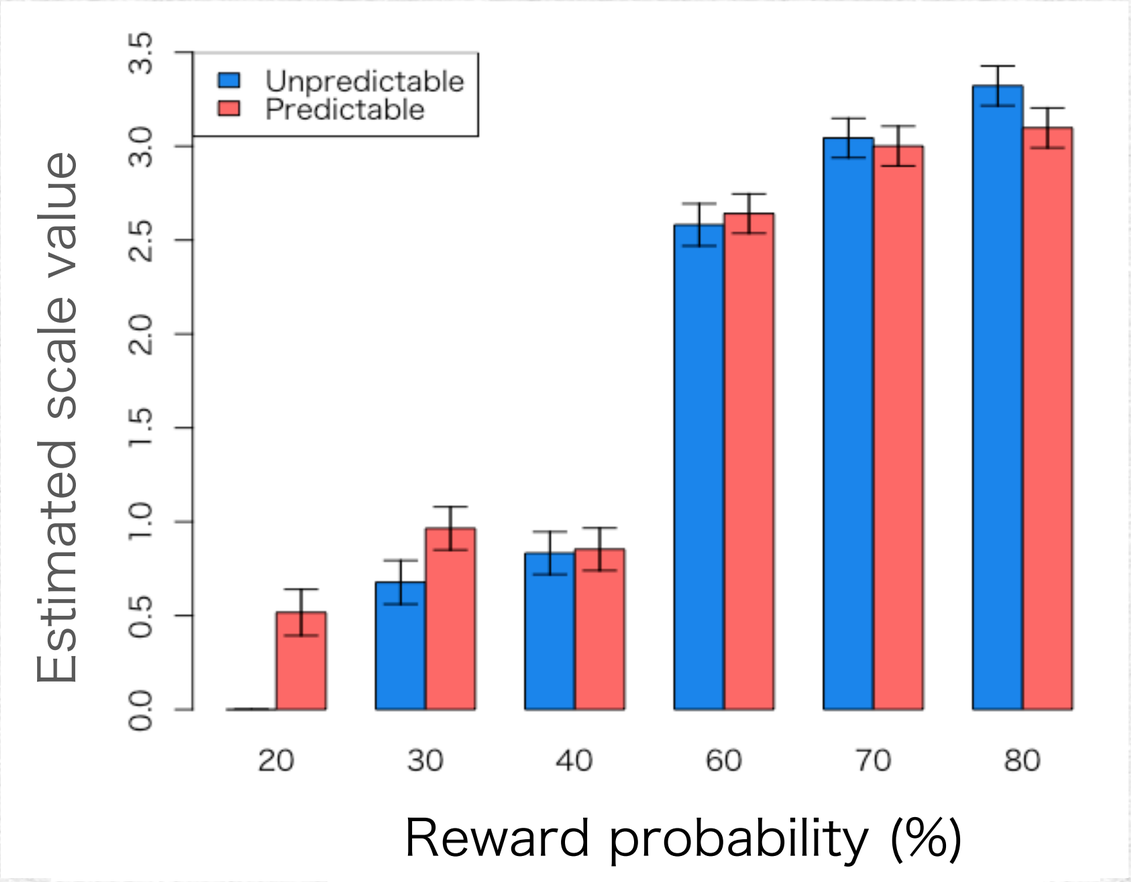
＊ In choice of **the lowest reward probability**

Participants were *unlikely* to choose stimuli accompanied with the unpredictable tone compared to that with the predictable one

(Unpredictable20% < Predictable20%: $z = 4.46$, $p < .001$)

⇔ In choice of **the highest reward probability**

(Unpredictable80% > Predictable80%: $z = 2.50$, $p < .05$)



Results

Fitting of generalized linear mixed model (GLMM) using logit link function

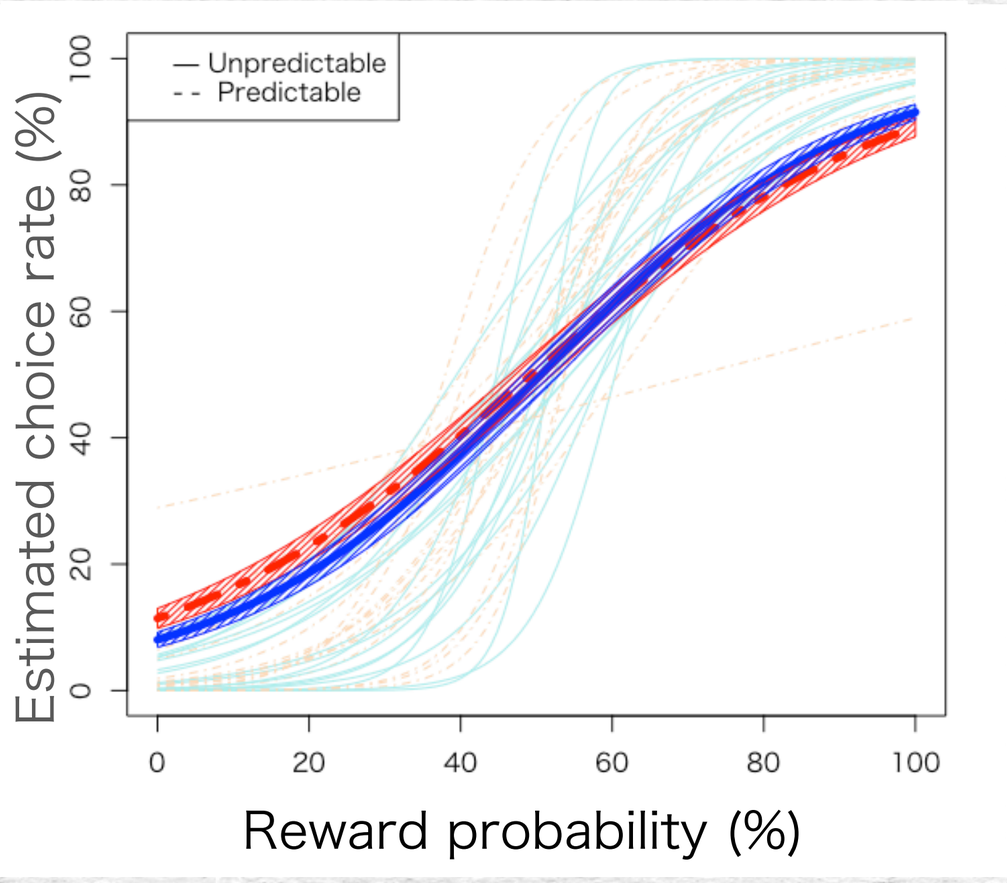
Random intercept: the factor of participants and presentation position of stimuli (left/right)

＊ An interaction effect confirmed **Unpredictable** tone

Decreased choice of stimuli with **lower** reward probability

Increased choice of those with **higher** reward probability

Factor	Coefficient	Wald's test
Intercept	-1.78	$p < 0.001$
Predictability	-0.38	$p = 0.596$
Reward	0.51	$p < 0.001$
Predictability × Reward	0.10	$p < 0.001$



Discussion

Q 1. Is the accuracy of action outcome prediction calculated implicitly?

Q 2. Can it influence the processing of reward?

A 1. YES.

Even when an action outcome is unrelated with goal of task,
man could judge whether action outcome is expected or not.

A 2. (At least partly) YES.

The predictability of action outcome modulated choice in reinforcement learning task,
particularly in the choice of specific worth.

Discussion

∗ How did the action outcome prediction modulate choice?

On the whole, the unpredictable outcomes promoted **appropriate** choice

(= Less choice of worthless stimuli + More choice of valuable stimuli)

✔ The agreement between predicted and observed action outcome might be processed as an implicit cue of the appropriateness of an action choice

※ There are some potential mechanism underlying this

The detection of unpredicted outcomes might drive …

more careful processing of other information of outcome, attention, or motivation to task…?

Thank you for your kind attention.

Any Questions?