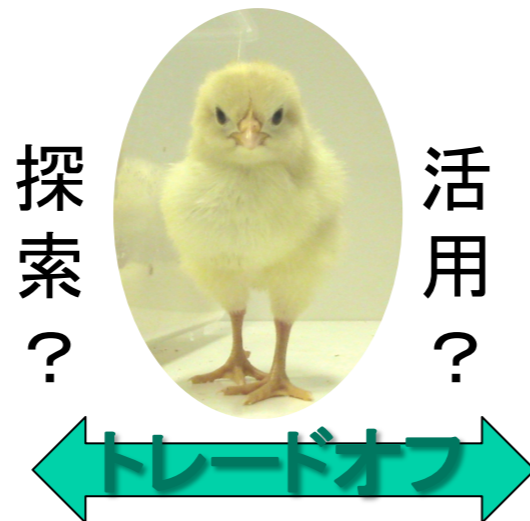


# ニワトリ雛の強化学習モデル: 報酬が不確実な場合の選択行動

川森 愛 モデリング研究系 特任研究員

## 背景

環境中の餌に関して完全な情報を持っている場合、採餌効率を最大化する行動ルールは、「利潤率(=餌量/遅延)の高い餌を選ぶ」ことであることが分かっている(Charnov 1976)。しかしながら、自然状態において餌に関する完全な情報を手にいれることは難しい。動物個体は環境を「探索」し、餌に関する情報を逐次学習していかなくてはならない。その際用いる学習ルールとして、強化学習モデルが提案されている。さて、探索するには環境中を満遍なく巡る必要があるが、中には質の悪い餌場もある。平均採餌効率をあげるには、ある程度のところで探索を切り上げる方がよい。学習を行わない場合の行動ルールとして、習慣が強く影響することが知られている。つまり、その時点までに経験した回数が多い選択肢ほどより高い確率で選択するのである。その際餌場の報酬は無視される。この行動により質の良い餌場を効率よく「活用」できる。しかし、いずれ環境が変動してより良い餌場が出現するかもしれないので、いつまでも同じ餌場を利用し続けることも良い戦略ではない。探索と活用の間にはトレードオフが存在する。採餌の成功のためには、両方の戦術を駆使しなくてはならないのである。本研究は、様々な複雑さの課題に対しニワトリ雛がどのように戦術を使い分けるかを調べた。



## 採餌戦術モデル

ニワトリ雛の行動ルールを探るため、以下の3つのモデルを検証した。

### 1. 強化学習モデル

(選択ルール)

$$choice_c \sim \text{Bernoulli}(P_{rl})$$

$$P_{rl} = 1/(1+\exp(-\beta(V_c - V_v)))$$

(価値アップデート)

$$V_{new} = V_{old} + \alpha \left( \frac{R}{1+D/\tau} - V_{old} \right)$$

R: 報酬の量, D: 遅延時間

$\alpha$ : 学習パラメータ  
 $\beta$ : 感度パラメータ  
 $\tau$ : 遅延割引スケールパラメータ

### 2. 習慣行動モデル

(選択ルール)

$$choice_c \sim \text{Bernoulli}(P_{habit})$$

$$P_{habit} = 1/(1+\exp(-\beta(\log N_{oc} - \log N_{ov})))$$

$N_{oc}/N_{ov}$ : これまでにその餌場を経験した回数

$\beta$ : 感度パラメータ

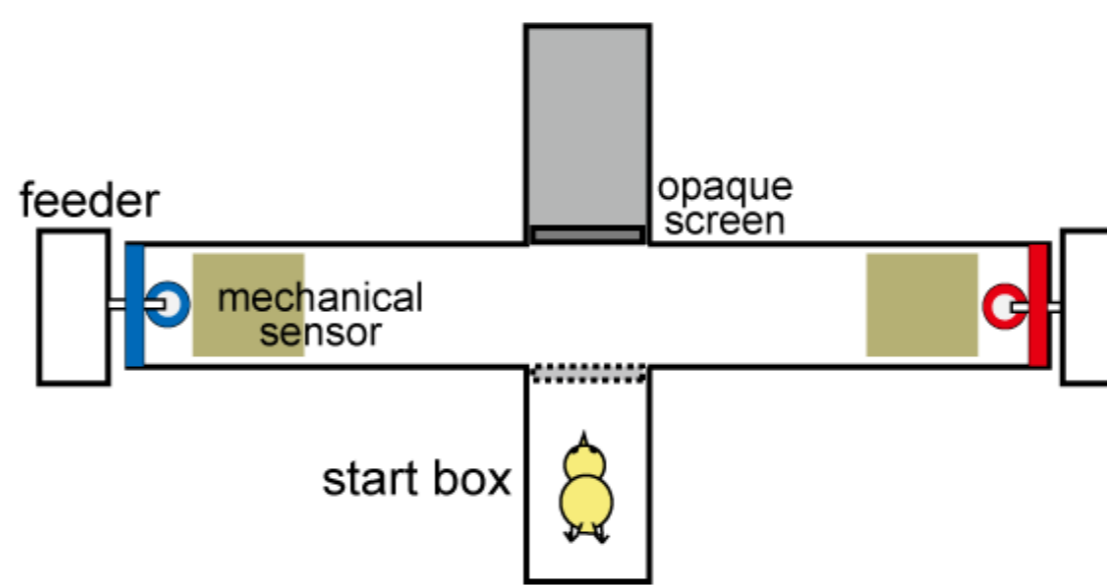
### 3. ランダム行動モデル

(選択ルール)

$$choice_c \sim \text{Bernoulli}(P_{random})$$

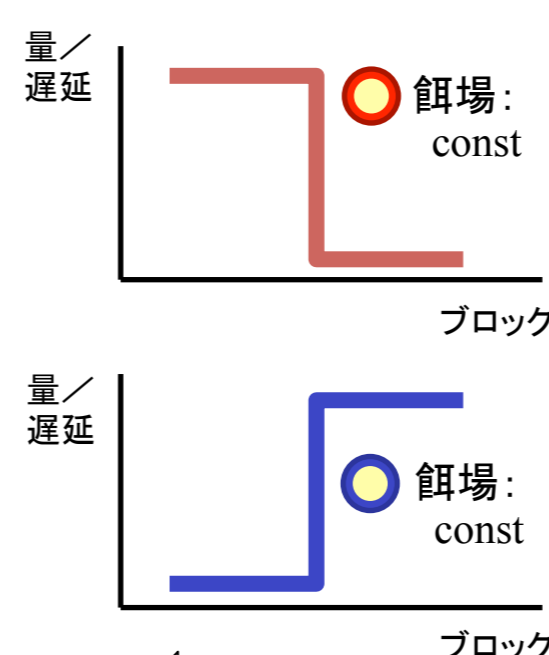
$P$ : 選択確率パラメータ

## 行動データ

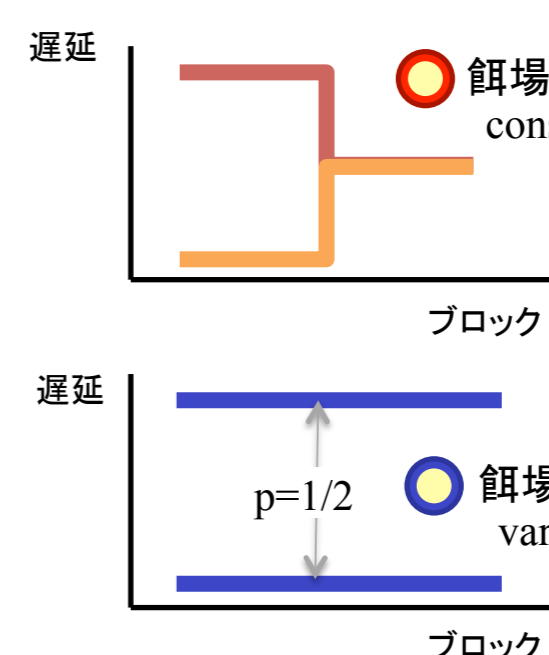


左のような十字迷路を用いて実験を行った。長腕の端には色の異なる2つの餌場がある。ニワトリ雛は短腕からスタートし、T字路でどちらか一方の餌場を選ぶ。餌場前の機械センサーに到達すると、決められた遅延時間の後、決められた量の餌が給仕される。実験者は餌場の餌量、または遅延時間を操作した。餌場として、結果(量または遅延)が一定であるため予測可能な餌場(const)と、確率的に変動するため予測できない餌場(var)の2種類を用意した。ニワトリ雛に対し、以下の3つの課題を課した。

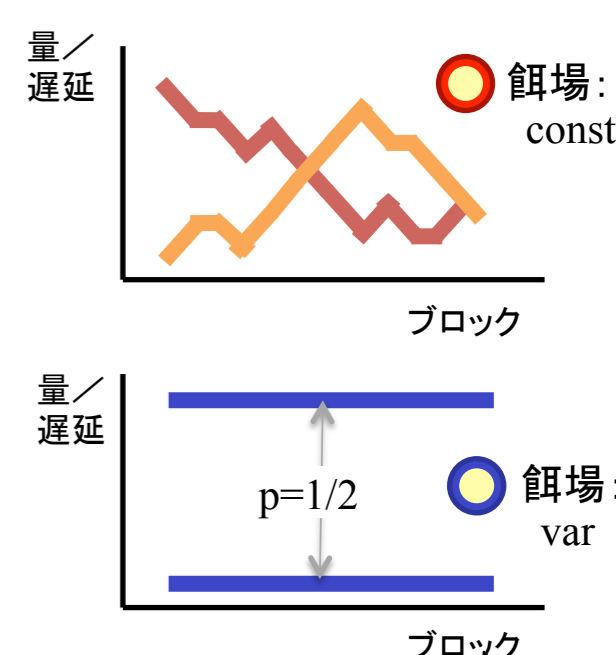
#### 反転学習課題



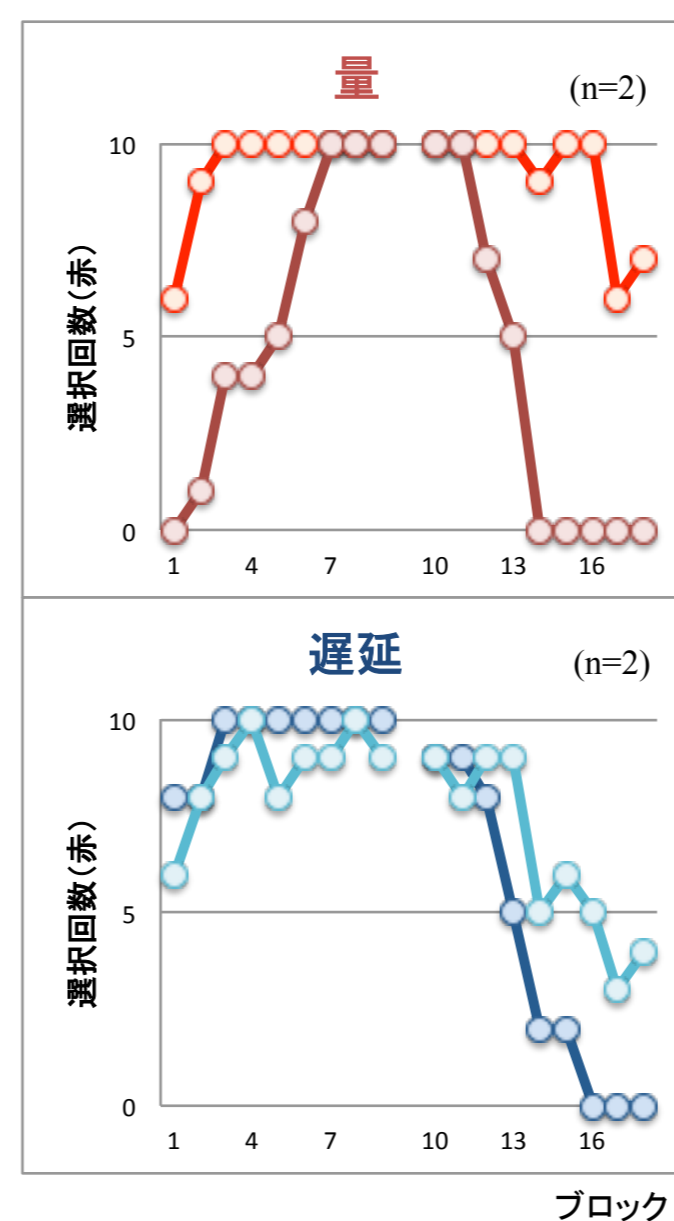
#### スイッチ学習課題



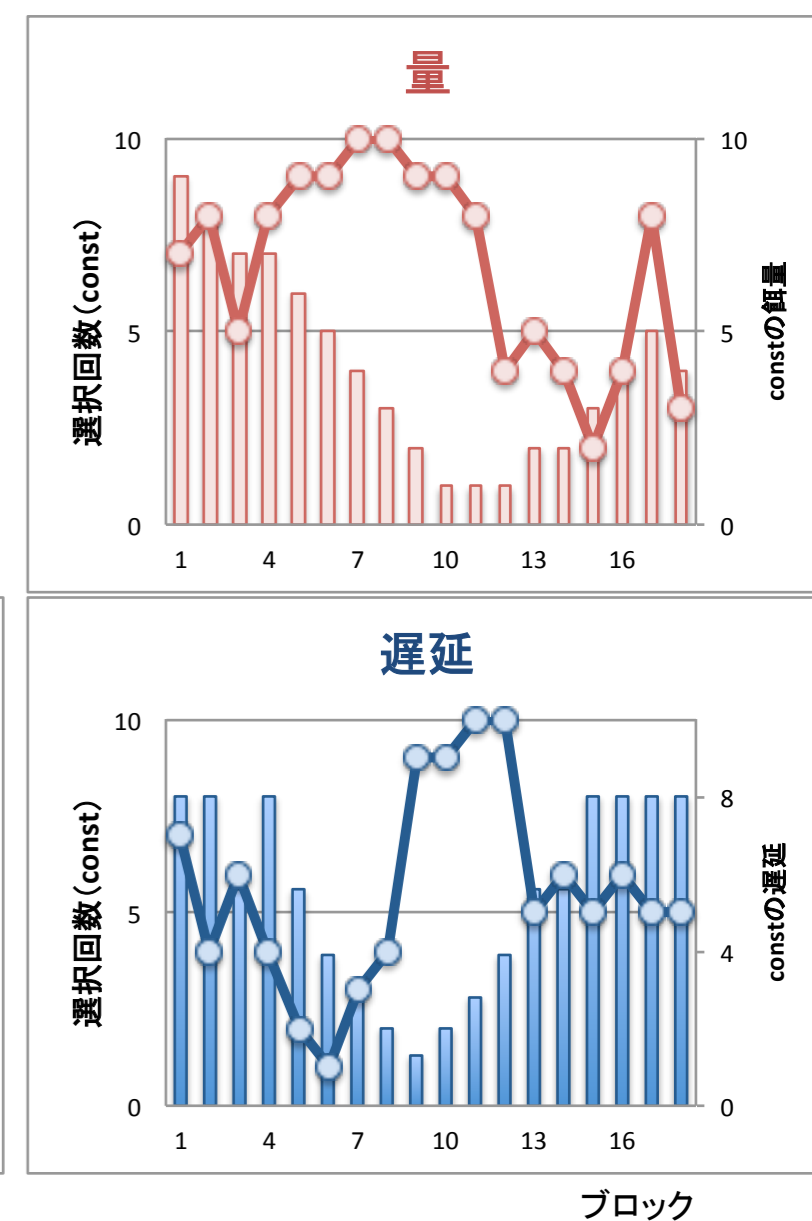
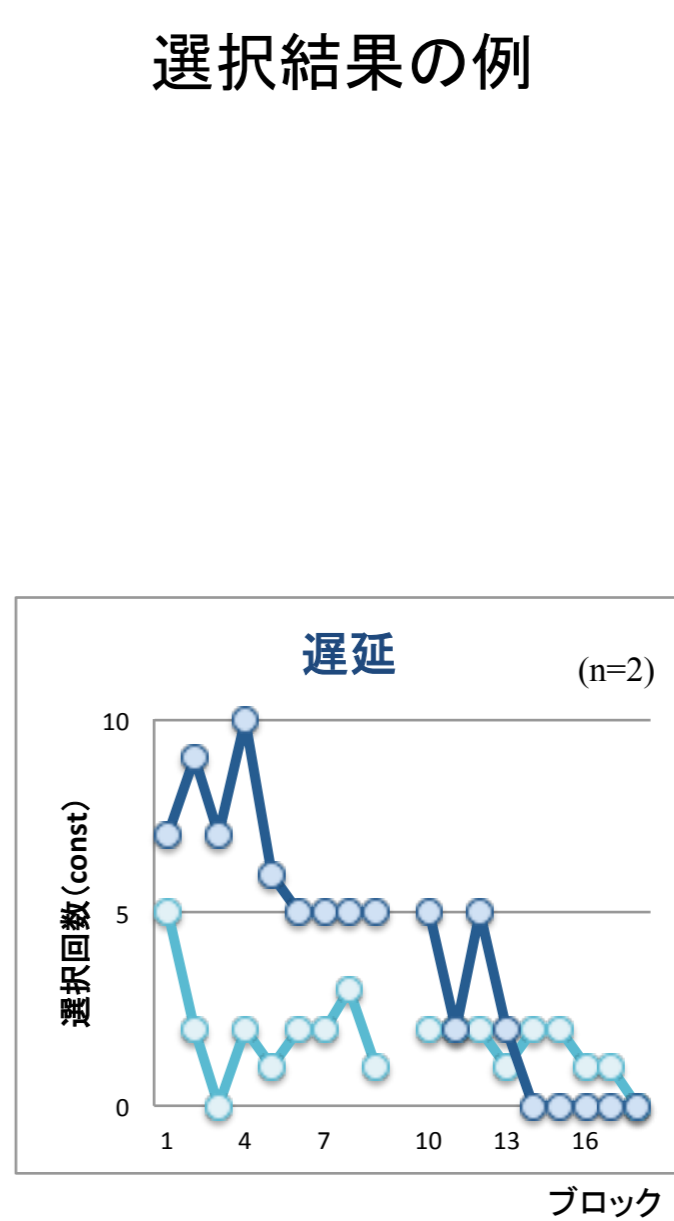
#### 行動滴定課題



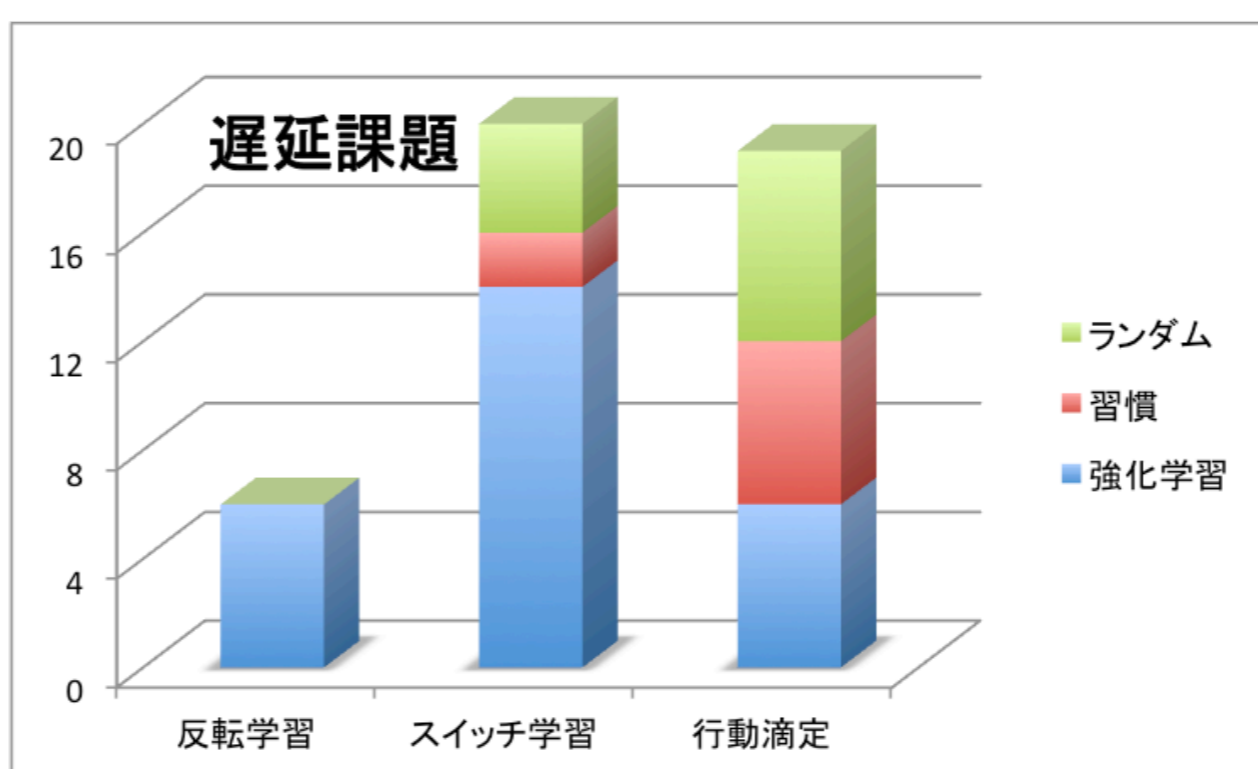
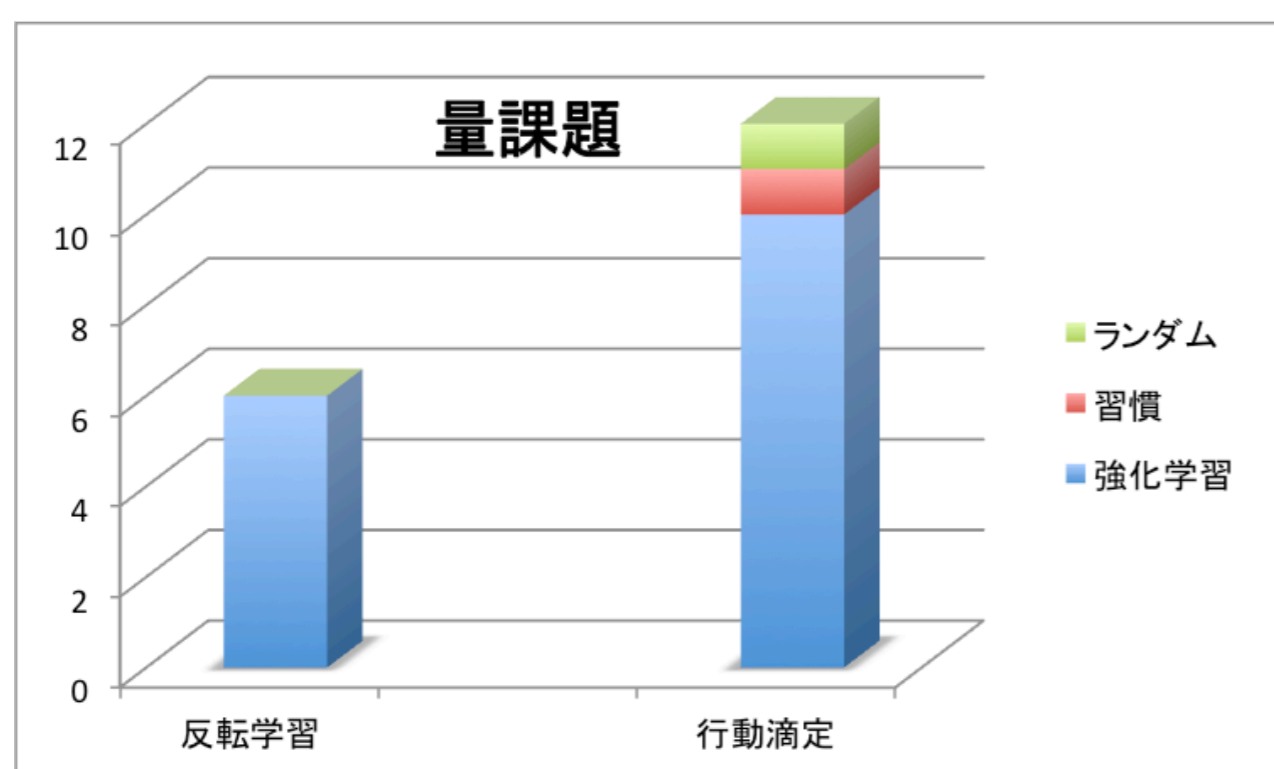
← 単純 → 複雑



#### 選択結果の例



## 課題による採餌戦術の違い



● 量課題ではほとんどの個体が強化学習を行っていた

● 遅延課題では課題が複雑になるにつれて、強化学習を行わない個体が増加した

● 強化学習を行う個体は課題が複雑になるにつれて減少した

本研究の餌条件はニワトリ雛がこれまで経験したことのない環境であり、学習しないことが適応的であるとは考えにくい。遅延に関し、課題が複雑になるほど強化学習を行わない個体が増加した理由として、2つの可能性が考えられる。

1. できなかった(能力がない)

ニワトリ雛が生育する環境において、多少の遅延の差異が生存を分けるような状況は少ないと考えられる。従って、わずかな遅延の差異を検出し、それに反応するような採餌能力は進化的に生じなかったのかもしれない。

2. できるがやらなかった(モチベーションがない)

遅延条件ではどの選択肢を選んでも餌量一定であり、収益の違いはない。ニワトリ雛は差し迫った問題に直面しない限り、複雑な問題に対しては強化学習を行わないのかもしれない。

