

Discrepant predictions from computational models of associative learning on the effect of contingency uncertainty*

Ji Hoon Jeong, Yang Seok Cho, June-Seek Choi[†]

School of Psychology, Korea University

The role of information uncertainty has wide implications ranging from emotional modulation to optimal decision making. Yet the concept has been employed as an ad-hoc explanation for various phenomena. One useful approach to the problem is to use a formal computational model to test different parameters extracted from animal and human studies on stimulus uncertainty.

We developed an integrated simulation environment written in Matlab (Korea University Conditioning Simulator: KUCS) which provides graphical user interface for several influential models of associative learning such as Rescorla-Wagner model, Mackintosh model, Pearce and Hall model, Schmajuk-Pearce-Hall model, Esber-Hasselgrove model, and Temporal Difference model. Using KUCS, We first demonstrated common predictions on basic conditioning phenomena: acquisition, extinction, blocking, conditioned inhibition, latent inhibition, and second-order conditioning to confirm the validity of the simulator and to find some novel limitations and predictions. We then generated a series of data under uncertainty and compared them with animal and human experiments to examine how the models' predictions on the associative strength and associability concur with the experimental data. The simulator program is available in <https://github.com/knowblesse/KUCS>

Keywords: modeling, associative learning, learning model, reward uncertainty

1차원고접수 21.07.30; 수정본접수: 21.10.31; 최종게재결정 21.10.31

심리학에서 가장 영향력이 큰 연구 중 하나인 Rescorla와 Wagner의 수리적 연합학습 모델이 1972년 발표된 이후 유사한 혹은 상이한 학습 원리를 적용한 다양한 모델들이 제안되었다. 예측 오차(prediction error)에 기반한 Rescorla-Wagner 모델(Rescorla & Wagner, 1972) 이외에도 주의(attention)에 기반한 Mackintosh 모델(Mackintosh, 1975), 불확실성에 기반한 Pearce-Hall 모델 및 이를 개선한 Schmajuk-Pearce-Hall 모델(Pearce & Hall, 1980; Schmajuk & Moore, 1985), 강화학습 이론에 기반한 Temporal Difference 모델(Sutton & Barto, 1987), 두 가지 모델을 결합한 하이브리드 모델(Esber & Haselgrove, 2011) 등이 대표적인 예이다. 이러한 연합학습 모델들은 신경과학

적 연구와 결합되어 뇌신경회로의 작동방식에 대한 통찰을 제공하기도 하였고(Moore & Choi, 1997; Schmajuk & DiCarlo, 1991), 인공지능 분야의 연구에서 핵심 알고리즘 중 하나인 강화학습의 개발에 기여하기도 하였다(Sutton & Barto, 2018).

한편, 위 모델들은 같은 학습 현상을 해석하는 방식이 다르고 사용하는 파라미터가 다양해서 모델 간 비교가 어렵다. 또한, 연합학습 모델들은 기존 모델의 문제점을 개선하거나, 특정한 학습 현상을 설명하는 것에 초점이 맞춰져 있어서 다양한 학습 계획에 대한 포괄적인 비교 분석이 진행된 적이 없다. 특히 무조건 자극(unconditioned stimulus: US)의 빈도와 강도가 변화하는 부분 강화계획(partial reinforcement)

* 본 연구는 한국연구재단의 지원을 받아 수행되었음(2020R1A2C2014830, 2021M3E5D2A01023887, 2017H1A2A1044665).

[†] 교신저자: 최준식, 고려대학교 심리학과, (02841) 서울시 성북구 안암로 145, E-mail: j-schoi@korea.ac.kr

이나 보상이 확률적으로 주어지는 보상 불확실성 (reward uncertainty) 상황에 대해서는 모델마다 예측하는 연합의 결과가 다르다. 이는 연합학습 모델마다 자극의 현저성 (saliency)의 변화나 연합의 강도를 예측하는 방식이 다르기 때문이다. 그러나 지금까지의 연합학습 모델에 관한 연구들은 개념적인 비교에 그치거나 연구자들의 개별 프로그램에 의해 진행되어 왔고, 이는 공개되지 않거나 사용 편의성이 떨어져 다양한 실험 패러다임을 구현하기가 어려웠다 (Mercier, 1996; Lipp, Stephens & Smith, 1999).

따라서 본 연구에서는 주요 학습 모델 6가지, Rescorla-Wagner (RW) 모델, Mackintosh (Mac) 모델, Pearce-Hall (PH) 모델, Schmajuk-Pearce-Hall (SPH) 모델, Esber-Haselgrove (EH) 모델, Temporal-Difference (TD) 모델들을 수식화 과정을 통해 정량화하고, 질적 비교 분석이 가능한 스크립트로 변환하였다. 더 나아가 다양한 학습 계획과 파라미터 분석을 위한 GUI 기반 시뮬레이터 프로그램도 제작했다. 연구 1에서 이 스크립트와 프로그램을 사용해 앞서 언급한 모델들이 다양한 학습 현상을 설명할 수 있는지 검증하였고, 한계점을 수리적으로 분석하였다.

이러한 시뮬레이터를 기반으로 연구 2에서는 모델들의 부분 강화 계획 패러다임 예측을 비교 분석했다. 이를 위해 무조건 자극을 지속 강화(continuous reinforcement) 조건과 부분 강화(partial reinforcement) 조건으로 제시했을 때의 연합 결과를 비교하였다. 또한, 동물과 인간 연구실험에서 공통으로 추출된 학습 계획을 사용해 불확실성이 존재하는 학습 상황에서 실험 결과를 예측할 수 있는지 시뮬레이션을 진행하고 한계점을 논의하였다.

연구 1. 기본적인 조건화 현상에 대한 예측

연구 1에서는 각 모델들이 다음의 다섯 가지의 학습 계획에서 어떻게 반응하는지를 확인하였다: 1) Acquisition/Extinction, 2) Conditioned inhibition, 3) Blocking, 4) Latent inhibition, 5) Second-order conditioning. 각 모델의 파라미터들은 모든 모델이 유사한 Acquisition/Extinction 커브를 그리도록 설정하여 사용하였고(부록 2), 여기서 변경한 값을 사용한 경우 따로 표기하였다.

방법

시뮬레이터. 본 연구에 사용된 모든 시뮬레이션은 Matlab © 스크립트를 통해 진행하였으며, R2020a 버전에서 테스트

하였다. (부록 2) 또한, 사용자 편의성을 위하여 작성한 GUI 버전의 소프트웨어 Korea University Conditioning Simulator (KUCS)는 같은 소프트웨어의 appdesigner 패키지를 이용하여 작성하였다.

학습 계획. 모델들의 복합적인 비교를 위하여 기본적인 다섯 가지의 학습 현상을 선택하였다. 이 다섯 가지의 학습 현상은 CS와 US 사이의 연합을 습득하고 소거하는 Acquisition/Extinction, 기존에 US와 연합이 형성된 CS를 새로운 CS와 함께 제시하면 새로운 CS가 US와 연합을 잘 형성하지 못하는 Blocking, US와의 연합 전에 CS를 반복적으로 노출시킨 경우 이후 US와의 연합 형성이 늦어지는 Latent inhibition, no-US와의 연합을 형성하는 Conditioned inhibition, 직접 적인 US 제시 없이 사전에 US와 연합된 CS를 함께 제시하여 US와의 연합이 전이되는 Second-order conditioning이다. 구체적인 학습 계획은 부록 1에 기재하였다.

핵심 변수의 결정. 사용한 모델마다 학습이 진행됨에 따라 변하는 변수가 다르기에 모델 간 정량적 비교를 위하여 최종적으로 관찰하는 변수를 CS의 연합의 강도와 연합가능성 (associability)으로 한정을 지었다. 먼저 연합의 강도로는 단일 V값을 사용하는 RW 모델, Mac 모델은 각각의 CS의 V값을 사용하였고, 부적 연합의 강도를 나타내는 \bar{V} 를 포함하는 PH 모델, SPH 모델, EH 모델은 정적 연합의 강도(V)와 부적 연합의 강도(\bar{V})의 차이 $\dot{V}(=V-\bar{V})$ 를 사용하였다. TD 모델의 경우는 V로 명시된 값이 없지만 의미상 w가 연합의 강도와 일치하기에 이를 택했다. 연합가능성 값은 RW 모델의 경우 고정된 학습률 β 를 사용하였고, Mac, PH, SPH 모델들에서는 α 를 사용하였다. EH 모델은 모델의 설계상 자극의 물리적 성질에 의해 결정되는 고정된 ϕ 와 학습에 의해 변하는 ϵ 의 합으로 연합가능성 α 를 사용하기에 이를 그대로 차용하였다. 연합가능성의 개념이 존재하지 않는 TD 모델에서는 α 를 계산하지 않았다

결과 및 논의

모든 모델에 대해 5 가지 학습 상황에 따른 시뮬레이션 결과를 비교하였다. 모든 모델이 acquisition과 extinction, 그리고 Kamin blocking을 설명할 수 있었지만, 연합가능성이 변하지 않는 RW 모델과 TD 모델은 latent inhibition을 설명할 수 없었다(Miller, 1995). 또한, 시간 요소가 개입된

Table 1. Summary of simulation results from basic conditioning phenomena

	Acquisition/ Extinction	Conditioned inhibition	Blocking	Latent inhibition	Second-order conditioning
RW Model	O	O	O	X	X
Mac Model	O	X	O	O	X
PH Model	O	O	O	△	X
SPH Model	O	O	O	O	X
EH Model	O	O	O	O	X
TD Model	O	O	O	X	△

O: Success; X: Failure; △: Partial success

TD 모델만이 second-order conditioning을 설명할 수 있는 점도 관찰하였다(Sutton & Barto, 1987). 흥미롭게도 US의 부재를 적극적으로 예측하는 conditioned inhibition 현상의 경우 Mac 모델은 설명하지 못하는 한계점을 발견하였다. 이는 기존의 이론적 논의에서 간과된 모델의 특성이라고 할 수 있다. 시뮬레이션 결과는 종합하여 Table 1에 요약하였다.

Conditioned inhibition (CI). US와 연합된 CS_A를 새로운 CS, CS_B와 US 없이 제시하면 이 CS는 “US의 부재”와 연합이 되며 이로 인해 통상 조건화에 의해 US와 연합된 흥분성 조건 자극(excitatory conditioned stimulus: ECS)과 대비되는 효과를 가지게 된다. 즉 CS_B는 억제적 조건 자극(inhibitory conditioned stimulus: ICS)이 된다. ICS의 효과는 다른 ECS와 함께 제시할 경우, ECS 단독으로 제시할 때보다 CR의 크기가 줄어드는 현상(summation test)이나, ICS를 다시 US와 정적 연합을 시킬 때 상대적으로 학습이 지연되는 현상(retardation test)을 통해 실제로 동물 학습에서 일어남이 확인되었다(Pavlov, 1927).

RW 모델과 TD 모델은 ICS와 ECS를 같은 차원에서의 일차원 벡터로 표현한다. 즉 V 혹은 w 값이 양수면 CS가 ECS이고 음수로 가면 ICS로 작동하게 된다. 한편 PH 모델, SPH 모델, EH 모델은 ICS와 ECS를 같은 차원에 두지 않고, CS의 부재(no-CS)를 예측하는 변수, \bar{V} 를 새로 만들어서 V값과 \bar{V} 값의 총합을 최종 반응에 반영하고 있다. 이들 모델도 CI나 extinction 등 ICS의 개념이 필요한 현상을 잘 설명한다.

흥미롭게도 Mac 모델은 conditioned inhibition을 설명하지 못하였다. 이는 기존 논문들에서 한 번도 지적되지 않은 Mac 모델의 한계를 발견한 것이다. RW 모델과 Mac 모델의 중요한 차이 중 하나는 예측 오차에 의한 V값의 변화

가 $(\lambda_t - \sum_{s' \in t} V_t^{s'})$ 에 기인하는지 아니면 $(\lambda - V_t^{CS_B})$ 에 기인하는지라고 할 수 있는데 전자는 다른 V^{CS}들과의 상호작용으로 인해 음수의 V값이 가능한 반면, 후자에서는 CS_B가 단독으로 예측 오차에 기여하기 때문에 V값이 음수가 될 수 있는 경우가 원천적으로 존재하지 않는다. 이를 보완하고자 다른 모델들이 \bar{V} 의 개념을 추가로 채택했으나, Mac 모델은 이러한 계산 장치를 언급하지 않고 있으므로 CI 현상을 재현하는 데 실패했다는 결론에 도달하였다.

Mac 모델의 CI 시뮬레이션 과정에서 발견한 또 한 가지 흥미로운 예측은 CS_A/CS_B- 시행에서 오히려 연합가능성 α 가 증가한다는 것이다. CS_A/CS_B- 시행은 RW 모델의 경우 CS_A+ 시행에서 양의 값을 가지게 된 CS_A로 인해 CS_B가 음수가 되는 시행인데 비해 Mac 모델의 경우 전술한 이유로 인해 V값이 음수가 되지 않을 뿐 아니라 연합가능성이 증가하는, 직관적인 예상과 일치하지 않는 예측을 하게 된다. 이는 Mac 모델의 연합가능성이 US를 얼마나 잘 예측하는가에 따라 증감이 되기 때문인데, CS_A/CS_B- 시행에서 US가 존재하리라 예측하는 CS_A보다 US가 없을 것으로 예측하는 CS_B가 예측력이 높기에 $(\lambda - V_t^{CS_B} < \lambda - ((\sum_{s' \in t} V_t^{s'}) - V_t^{CS_B}))$ CS_B의 예측력이 올라가 α 값이 증가한다. 따라서 Mac 모델은 CI 절차를 마친 CS_B가 새로운 조건화 학습에서 US와 결합하는 경우 acquisition 과정에서 중성적인 (naive) CS보다 오히려 빠른 학습을 보일 것으로 예측한다. 그러나 이러한 예측은 실제 동물이나 인간 대상 실험에서 증명된 바가 없다.

Latent inhibition. LI는 단독으로 CS에 대한 반복적 노출로 인해 이후 US와의 연합학습이 지연되는 현상으로 새로운 CS에 대한 학습이 US에 의존하는 RW 모델에서는 구현할

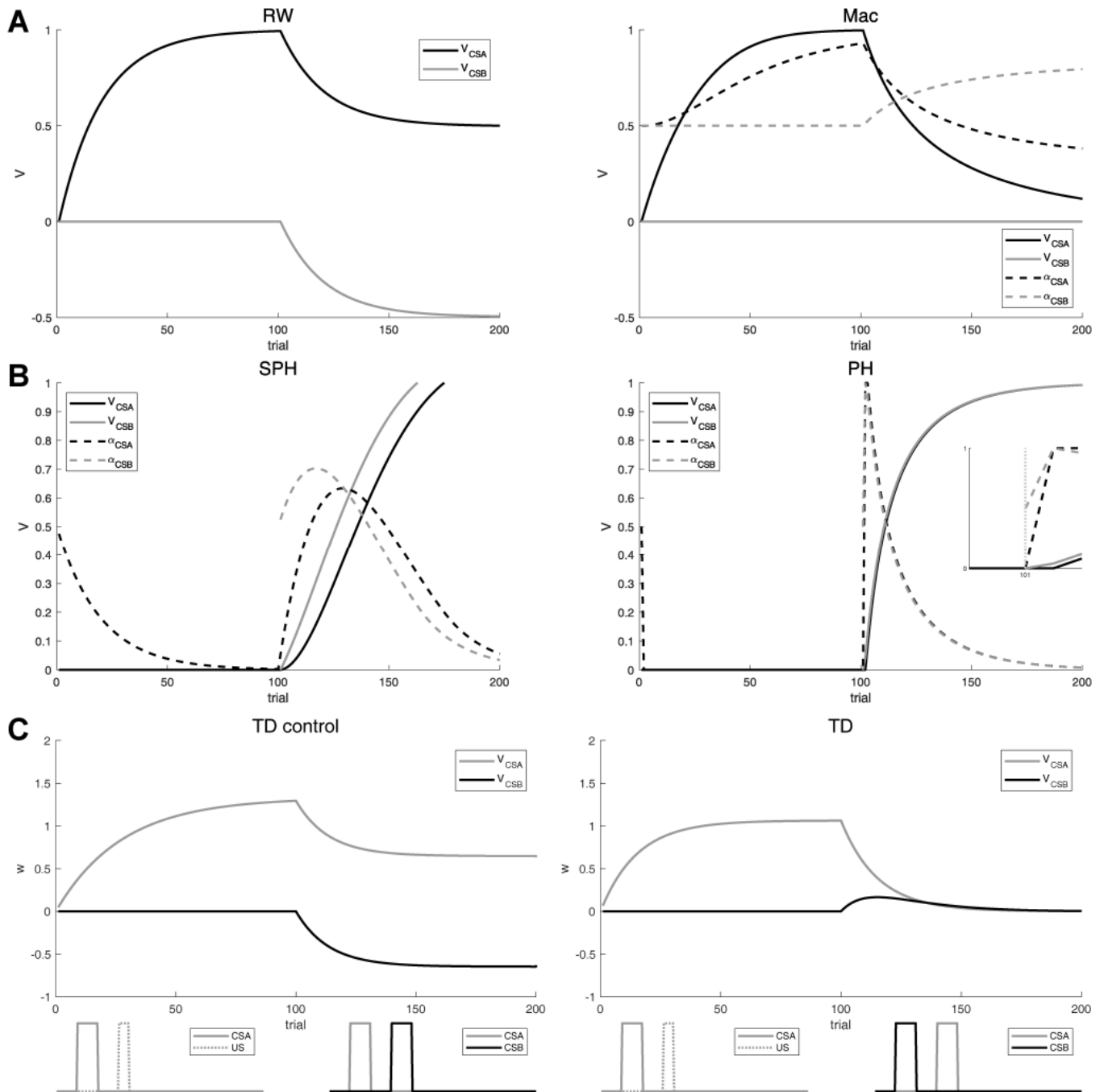


Figure 1. Example outputs from KUCS. Solid lines indicate the V value, and dashed lines indicate the associability of each CS. Key parameters for all the models were adjusted to produce similar scale in acquisition and extinction curves. Formal mathematical equations and parameters employed in each simulation are listed in the appendix. **A. Conditioned Inhibition:** RW model and Mac model. The first 100 trials were CS_{A+} and the second 100 trials were $CS_{A/B-}$. Mac model failed to produce conditioned inhibition as V for CS_B continue to remain at 0 (solid gray line) on trials 101–200. In addition, the associability of CS_B increased during the same period (dotted gray line), which is also counterintuitive. **B. Latent Inhibition:** SPH model and PH model. Each model simulated 100 trials of CS_{A-} and 100 trials of CS_{A+} . For the comparison, a new CS_B (gray line) is introduced from trial 101. SPH model shows retarded acquisition of the CS_A (solid black line) during the acquisition (trial 101–200) compared to the CS_B (solid gray line). This is due to the non-reinforced CS presentation in trial 1–100. PH model predicts latent inhibition only partially because the decreased associability during CS_{A-} trials only affects the first trial of following reinforced trials (trial 101–200; inset). SPH model used altered gamma value: $\gamma = 0.05$ to enlarge the discrepancy. **C.(top)** Simulation results of TD model with two different CS_{AB-} trials' stimulus time configurations. Only when previously conditioned CS_A is presented after the to-be-associated CS, the CS_B gains positive association with the US (**right**). **(bottom)** Stimulus time configurations of each experiment.

수 없다. TD 모델 역시 V값이 0에서 출발하는 새로운 CS가 단독으로 제시될 경우 이를 학습에 반영하는 수학적 장치가 없어서 LI를 구현하지 못한다. 반면의 연합가능성 α 가 가변하는 Mac 모델과 PH 모델, SPH모델, EH 모델은 이러한 학습 지연을 V값이 아닌 줄어든 α 에 의한 영향으로 설명한다.

PH 모델의 경우, α 값이 바로 전 시행에 의해서만 결정되므로 CS_{A-} 시행에서 CS pre-exposure를 진행하면 두 번째 시행부터 α 값이 0이 되어 이를 학습 현상에 대한 모사로 보기에는 어색한 측면이 있다. 유사하게 CS_{A+} 시행에서 US가 같이 제시되기 시작하면 첫 시행에서는 대조군보다 학습이 저조한 형태를 보이지만 그 이후부터 대조군과의 차이가 사라져서 (Fig. 1B inset) 거의 같은 형태의 학습곡선을 보인다. 반면 SPH 모델은 PH 모델의 급변하는 α 에 관성을 추가하여 변화를 완화하는 항목이 있으므로 PH 모델에 비해서 CS pre-exposure로 인한 $\alpha = 0$ 의 영향이 여러 시행에 영향을 미치며, 이는 대조군과 확실한 대비를 보인다.

Second-order conditioning. 이전에 US와 연합을 형성했던 CS_A의 연합력(association strength)이 새로운 CS_B로 전이되는 second-order conditioning은 conditioned inhibitor를 만드는 실험 계획과 시행 구성(trial configuration) 단위에서 동일하다(부록 1). 때문에, 시간 기반(time-based) 모델에서

만 이 학습 현상을 구현할 수 있으며, CS가 conditioned inhibitor가 될지, second-order conditioning에 의한 새로운 excitator가 될지는 세부적인 실험 절차에 의해 결정된다(Yin, Barnet, & Miller, 1994). 먼저 CS_{A+} 시행들이 CS_A/CS_{B-} 시행들보다 먼저 제시가 되어야 하고, CS_A/CS_{B-} 시행이 너무 많이 제시되지 않아야 한다. 특히 CS_A/CS_{B-} 시행에서 각 CS의 시간적인 배치도 중요한데, CS_B 제시 바로 직후에 CS_A가 제시되어야 second-order conditioning을 일으킬 수 있다(Kamin & Szakmary, 1977).

시뮬레이션에 사용한 유일한 시간 기반 모델인 TD 모델에서 second-order conditioning을 구현할 수 있었으나, 앞서 제시된 조건을 만족할 때만 CS_B가 양의 w 값을 가질 수 있었다. TD 모델도 RW 모델과 유사한 형태로 w 값을 변경하는데, 관측값 λ 과 예측값 $\sum V^s$ 의 차($\lambda - \sum V^s$)에 의해 학습이 진행되는 RW 모델과 달리 현재의 관측값과 예측값의 합($\lambda + w_t^T x_t$)과 이전 시간의 예측값($w_t^T x_{t-1}$)의 차 ($\lambda + \gamma w_t^T x_t - w_t^T x_{t-1}$)를 줄이는 방향으로 w 값의 변화가 일어난다. 여기서 현재의 예측값($w_t^T x_t$)이 US와 같은 역할을 해서 CS_A의 연합력이 CS_B로 전이될 수 있다. 이 때문에 양의 w 값을 가지는 CS_A가 CS_B이후에 나와야 US에 대한 정적 연합을 형성할 수 있다. (Fig. 1C)

Table 2. Partial reinforcement schedules adopted by Experiment 2

	Schedule Name	Schedule	
		Acquisition	Extinction(Testing)
Haselgrove et al. (2004) Partial Reinforcement Extinction Effect	Continuous	A+(0.5) 100 trials	A- 100 trials
	Partial	{A+(1.0), A-} 50 repetitions (100 trials)	A- 100 trials
Cho & Cho (2021) Uncertainty-Driven Attentional Capture	Exp1 : Uncertain	A+(1.0) 25% A- 75% 100 trials	A- 100 trials
	Exp1 : Certain	A+(0.25) 100% 100 trials	A- 100 trials
	Exp2 : Uncertain	A+(0.10) 25% A+(0.25) 25% A+(0.75) 25% A+(0.90) 25% 100 trials	A- 100 trials
	Exp2 : Certain	A+(0.5) 100% 100 trials	A- 100 trials

연구 2. 정보의 불확실성이 포함된 학습 상황에 대한 예측

결과 및 논의

연구 2에서는 연구 1에서 작동방식이 확인된 시물레이터를 활용하여 정보의 불확실성이 연합학습에 미치는 영향을 확인하였다. 구체적으로, 부분 강화 학습 계획을 사용하여 각 모델의 정량적인 예측을 확인하고 이를 실제 실험 결과와 비교했다. rat을 사용한 실험과 사람을 대상으로 한 연구를 사용하였는데, 동물 실험으로는 소리 자극에 대한 쥐의 먹이통 접근 반응을 사용해서 Partial Reinforcement Extinction Effect (PREE)를 측정하는 연구(Haselgrove, Aydin & Pearce, 2004)를 사용하였고, 사람 실험으로는 불확실한 자극에 대한 Value-Driven Attentional Capture (VDAC)를 측정하는 연구(Cho & Cho, 2021)를 참고하였다.

방법

시물레이션 및 모델 비교. 연구 1과 같은 방법 (KUCS)를 사용해서 시물레이션을 진행하였으며, 파라미터를 변경한 경우 따로 표기하였다.

학습 계획. 본 연구는 Table 3의 학습 계획을 사용했다. 두 연구 모두 한 가지 CS를 사용하였으며, 조건별로 CS와 함께 제시되는 US의 양과 확률이 달랐다. CS가 US가 함께 제시되는 경우 +, 제시되지 않는 경우 -로 표기하였으며 제시된 US의 상대적인 크기는 괄호 안에 표기하였다. 사람 대상 연구와 시물레이션 결과를 비교할 때, uncertain 그룹에서 제시되는 시행의 순서에 따라서 학습 결과가 다소 달라진다. 때문에, 서로 다른 시행 구성으로 100번의 시물레이션을 돌린 후, 평균값을 사용하여 비교를 진행했다.

Partial Reinforcement Extinction Effect in rats (Animal Study). 비교 연구에 사용한 동물 데이터는 PREE에 영향을 주는 요소를 찾기 위한 연구로, rat을 사용한 총 4개의 Pavlovian conditioning 실험으로 구성되어있다(Haselgrove et al., 2004). 이 중, 동일한 양의 US를 받았을 때 지속 강화와 부분 강화가 이후 extinction session에서 어떤 영향을 끼치는지를 확인한 첫 번째 실험 데이터를 모델 비교 분석을 위해 사용하였다. 연구에서는 소리와 food pellet을 사용한 욕구 조건화 패러다임을 사용하였고, 소리가 나올 때 먹이 공급기에 접근하는 횟수를 CR로 사용하였다. 동물들은 매 CS마다 1개의 food pellet이 제시되는 continuous 조건과 2개의 food pellet이 50% 확률로 제시되는 partial 조건으로 나누어 acquisition session을 진행한 뒤, 두 번의 extinction session을 통해 PREE를 측정하였다. 연구자들은 acquisition session에서는 두 그룹 간 차이가 없었으나, extinction session에서 partial 그룹에서 extinction이 지연되는 PREE를 관찰했다고 보고하였다(Figure 2A). 실험에서 사용된 학습 계획을 그대로 적용해 시물레이션을 진행하였고, 여기서 도출된 V 값과 동물 행동 결과를 비교하였다.

시물레이션 결과 TD 모델을 제외하고 모든 모델에서 partial 강화 조건에서 acquisition 후 V 값의 최대치가 continuous 강화 조건보다 높게 나왔다(Table 4, Figure 2B). 이는 두 조건에서 유사한 acquisition 곡선을 그리는 동물 행동 결과와는 다른 결과다. 또한, PH 모델의 경우 CS_{A+} 시행과 CS_{A-} 시행이 번갈아 제시되는 학습 계획에서는 정상적인 학습이 일어나지 못했다. 이는 PH 모델의 associability가 바로 직전 시행의 uncertainty에만 의존하여 변하도록 설계되어 있기 때문이다. 따라서 자극을 두 번의 A+, 두 번의

Table 3. Summary of simulation results from partial reinforcement schedules.

	Asymptotic V compared to continuous reinforcement	Time to reach the asymptotic V	PREE**
RW Model	Higher	Similar	X
Mac Model	Higher	Similar	O
PH Model*	Higher	Longer	X
SPH Model	Higher	Longer	X
EH Model	Higher	Similar	O
TD Model	Same	Same	X

O: Success; X: Failure

* The CS from PH model only forms association with the US when two + trials and two - trials are presented in alternating manner.

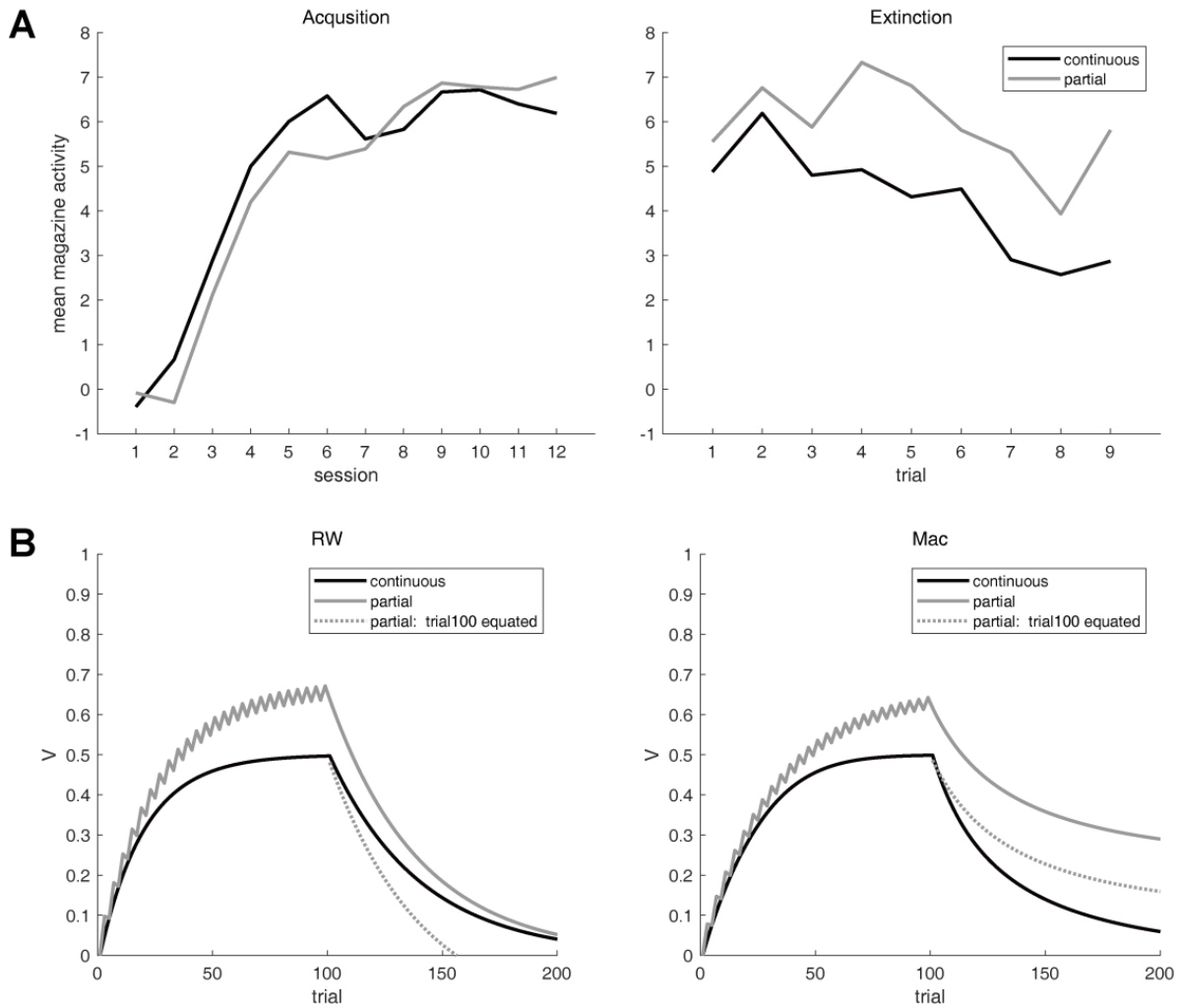


Figure 2. Partial reinforcement extinction effect (PREE) in experiments and simulations. **A. Results from animal experiments (modified from Haselgrove et al., 2004)** Animals went through appetitive Pavlovian conditioning with two different schedules: continuous and partial. Animals in the continuous schedule received one food pellet (US) after each CS when those in the partial group received double the amount (two pellets) on half of the trials and no pellet on the other half. The CR was measured by the approach toward the food magazine. During extinction, CS-only trials (no pellets) were presented. Although the asymptotic levels were similar for both continuous and partial reinforcement schedule at the end of the acquisition, the subsequent extinction session clearly shows PREE with partial group responding at a higher rate throughout the entire session. **B. Simulation results from KUCS. Trial 1–100: acquisition; Trial 101–200: extinction.** RW model and Mac model predicted a higher asymptotic associative value for the partial reinforcement condition due to greater lambda (0.5 vs. 1). After equating the level of associative value at the beginning of the extinction, PREE is only replicated by Mac model (dotted gray line).

A- 시행의 순서로 제시하는 학습 계획을 사용하여 비교하였다.

Mac 모델과 EH 모델의 경우 부분 강화된 자극의 소거가 느리게 일어나는 현상이 관찰되었다. 이러한 현상은 acquisition 이후 상대적으로 높은 부분 강화 조건의 V 값을 연속 강화 조건과 일치시켜도(Figure 2B 회색 점선) 유지되었다. 이러한 결과는 많은 동물 연구(Humphreys, 1939; Mowrer & Jones, 1945; Harris, Kwok & Gottlieb, 2019)에서 확인되었고 사용한 동물 실험 결과와도 일치하는 결과

이다. PH 모델과 SPH 모델도 extinction 중 상대적으로 높은 V 값을 보였으나, acquisition 후 두 조건 간 V 값을 일치시키자 그 효과가 사라졌다.

Effect of uncertainty in value-based attentional capture in human. Value-Driven Attentional Capture (VDAC)는 높은 가치와 연합된 자극에 불수의적으로 주의가 이동하는 현상이다(Anderson, Laurent & Yantis, 2011). 이러한 현상은 visual search task를 사용해서 측정할 수 있는데, 사전에 높

Table 4. Summary of simulation results from uncertainty-driven attention capture

	V value in uncertain condition compared to certain condition		Reversed V value	
	Exp1	Exp2	Exp1	Exp2
RW Model	Higher	Same	X	X
Mac Model	Same	Same	X	X
PH Model	Higher	Higher	X	X
SPH Model	Higher	Higher	X	X
EH Model	Higher	Higher	O	O
TD Model	Same	Same	X	X

O: Success; X: Failure

* The CS from PH model only forms association with the US when two + trials and two - trials are presented in alternating manner.

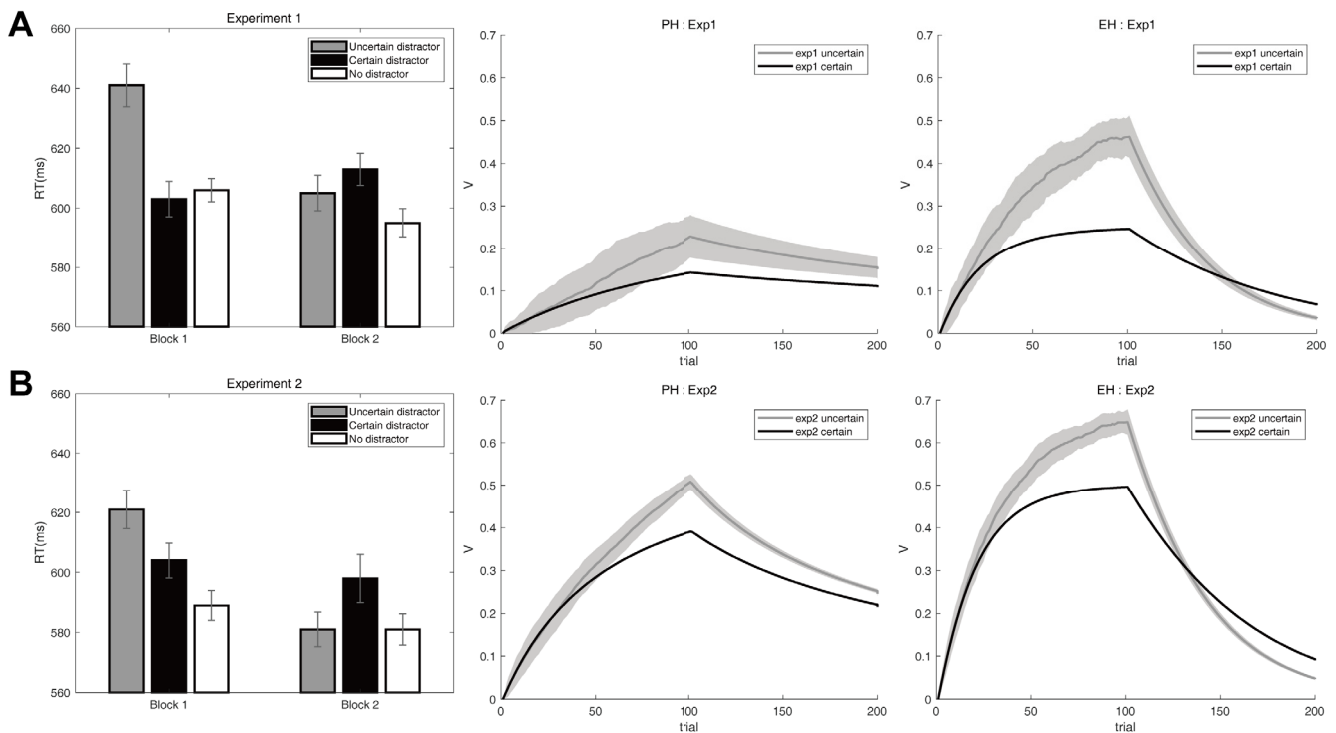


Figure 3. Comparison between experimental and simulation results based on uncertainty-driven attentional capture experiments in humans (Cho & Cho, 2021). During the training phase, participants were presented with two target colors that are different only by the associated reward variance: certain vs. uncertain. After the training, the degree of the Value-Driven Attentional Capture (VDAC) was measured through a visual search task with two previously paired colors as distractors. Experiment 1 modulated reward uncertainty with respect to the presence of the reward: 75% of no reward condition and 25% of reward condition, while maintaining the expected value same as the certain distractor condition. On the other hand, participants in Experiment 2 received reward in every uncertain condition, but the reward magnitude was varied. Experimental results were depicted on the left column of each figure. **A. (Left)** Increased RT in the uncertain distractor condition during Block 1 indicated the VDAC effect favored uncertain reward presentations although the effect faded in Block2. (Error bars ± 1 within-subject SEM) **(Right)** Simulation results of PH model and EH model. During the training period for the uncertain condition, all simulations used 100 pseudo-randomly mixed trial configurations and mean V value was used. (gray shaded area : 1sd) In both models, uncertain CS acquired higher V values during the acquisition (trials 1 - 100). This discrepancy was decreased during the extinction (trial 101 - 200) and even reversed (EH model) after trial 150. PH model and EH model used different set of parameters: PH : $S^A = 0.02$ EH : $\beta_{1_{acq}} = 0.06, \beta_{1_{ext}} = 0.01$ **B. (Left)** In experiment 2, the uncertainty distractor was paired with four different reward outcomes but had smaller variance compared to experiment 1. Interestingly, the VDAC effect was marginally reversed in Block 2, indicated by the higher RT in the certainty distractor condition. (Error bars ± 1 within-subject SEM) **(Right)** Although PH model showed a similar prediction as experiment 1, EH model simulated reversed VDAC effect during trial 141-200.

은 가치와 연합을 시킨 distractor가 함께 제시될 때 RT가 증가하는 것으로 확인할 수 있다. Cho와 Cho (2021)은 불확실성을 가진 보상과 연합된 distractor가 비슷한 현상을 일으키는지를 두 가지 실험을 통해서 확인하였다. 보상의 총량을 동일하게 유지하면서 보상의 크기를 변화시키는 방법으로 attentional capture를 측정했는데, testing session의 초반부에서만 RT가 높은 불확실성과 연합된 distractor에 의해 증가하는 것을 확인할 수 있었다(Figure 3A).

사람 실험과 같은 학습 계획을 6가지 모델로 시뮬레이션을 진행한 결과, 일부 모델에서 uncertain 조건에서 certain 조건보다 높아진 V 값으로 확인할 수 있는 VDAC를 관찰할 수 있었다(Table 5). 특히 PH 모델, SPH 모델, EH 모델에서 두 실험 모두 uncertain 조건에서 V 값이 높게 나타났다. RW 모델의 경우 US를 전혀 받지 않는 시행이 포함된 실험 1에서만 이러한 현상이 나타났고, 모든 시행에서 US를 받는 실험 2의 경우는 이 현상을 관찰할 수 없었다.

특히 EH 모델의 경우, extinction 세션을 진행하면서 높아진 α 값에 의해서 uncertain 조건에서 빠르게 extinction이 일어나는 것을 확인하였다. 일부 파라미터를 사용하여 모델을 시뮬레이션한 경우 초반의 급격한 α 값의 감소로 인해 uncertain 조건에서 130~140 시행부터 certain 조건의 V 값을 넘어 더 내려가는 현상을 보였다(Figure 3A 우측). 이 같은 현상은 실험 2의 학습 계획에서 두드러지게 나타나 extinction 세션의 후반부에선 certain 조건과의 격차가 크게 벌어졌다(Figure 3B 우측). EH 모델의 이러한 결과는 VDAC가 초반부에서만 관찰되는 실제 사람 실험 결과와 일치되는 현상이다. 특히 사람 실험 2에서는 오히려 certain reward와 연합된 distractor의 RT가 증가하는 현상이 관찰되었는데, EH 모델도 같은 결과를 예측한다는 점이 주목할만하다.

종합 논의

본 연구에서는 학습 심리학 분야에서 널리 알려진 여섯 가지의 모델을 수식화하고, 정량적 비교 분석이 쉽도록 시뮬레이터 프로그램을 개발하였다. 이러한 시도로 기존에 밝혀지지 않았던 학습 모델의 특성들을 확인할 수 있었고, 각각의 모델들의 강점과 한계점 또한 파악할 수 있었다. 특히 Mac 모델은 동물 실험 결과에서 잘 알려진 학습 현상인 CI를 설명하지 못한다는 점을 발견하였다. 그 원인은 Mac 모델이 RW 모델을 수정할 때, 여러 자극이 서로 경쟁해서 주의를 획득한다는 inverse hypothesis (Thomas, 1970)를 폐기하고

각 자극을 독립적으로 학습하도록 설계했기 때문이다. 또한, Mac 모델에서 ICS의 α 값이 증가하는 것으로 나타나 추후 Mac 모델의 수정에 반영되어야 할 이론적 시사점을 제공하였다.

또한, PH 모델의 경우는 US의 유무가 번갈아 반복되는 형태의 부분 강화 조건에서는 전혀 학습하지 못하였다. 이는 PH 모델의 연합가능성을 변경하는 방식이 다른 모델들과 다르기 때문이다. 연합가능성이 변하는 다른 모델들의 경우, 이전 시행의 값에 현재 시행의 결과에 의한 변화를 조금씩 가감하는 방식을 취하고 있다. 그러나 PH 모델은 바로 이전 시행에 의해서만 현재 시행의 연합가능성 값이 결정된다. 따라서 연속적으로 US가 제시되지 않으면 학습이 일어나지 않는다. 그러나 이를 개선한 SPH 모델에서는 이러한 문제가 발생하지 않았다.

연구 2에서는 부분 강화, 혹은 강화물에 불확실성이 존재하는 학습 계획에서 모델의 예측을 분석하였다. Mac 모델, EH 모델은 부분 강화 계획의 핵심 현상인 PREE를 예측하였다. 두 모델은 acquisition동안 부분 강화를 받으면서 연합가능성 α 가 감소하고 이로 인해 연이은 extinction 동안 연합강도 V 의 감소가 연속 강화를 받은 조건에 비해 느리게 일어나게 된다. 불확실성에 의해 연합가능성이 증가하는 모델인 PH 모델과 SPH 모델은 부분 강화로 acquisition을 진행한 경우 높아진 연합가능성으로 인해 더 빨리 extinction이 일어났다. 그러나 PREE를 예측한 두 모델도 acquisition 이후에 V 값이 부분 강화 계획보다 높은 동물 결과와는 다른 한계를 보여주었다. 한편으로, 실제로 동물 실험에서 부분 강화 조건과 연속 강화 조건에서 asymptotic 수준이 유사하게 나타나는 이유가 일종의 천정 효과 (ceiling effect)일 가능성이 존재한다. 또한, 부분 강화 조건에서 제시되는 US의 양을 맞추기 위해 강화물의 양을 2배로 증가시켜 진행했기에 asymptotic level이 증가하였는데 실제로 동물 연구에서도 부분 강화 조건의 강화물을 증가시키지 않으면 낮은 수준의 CR을 보인다는 증거가 존재하는 것으로 볼때 (Rescorla, 1999a; Rescorla, 1999b), asymptotic 수준의 차이는 근본적인 연합 강도의 차이라기보다는 강화물의 크기에 따른 실험 조건의 변동으로 보인다.

PH 모델, SPH 모델, EH 모델은 모두 uncertain 조건에서 높은 V 값을 보였고, 이는 uncertain reward와 연합된 distractor가 제시될 때 탐색 RT가 증가하는 사람 실험의 결과와 합치한다. 이는 이 모델들이 모두 불확실성의 정도에 따라 연합가능성이 증가하는 계산적 특성이 있기 때문이며 따라서 보상의 유무와 크기라는 두 가지 측면에서의 불확실

성이 존재하는 학습 계획들로부터 나온 결과를 잘 예측한다. 특히 EH 모델은 extinction 후반부로 갈수록 uncertain cue가 distractor로써의 힘을 잃고 오히려 certain distractor가 제시되는 조건에서 RT가 더 증가하는 현상을 설명할 수 있었다. 그러나 학습률 파라미터인 β 값들의 조합이 다를 경우에는 이러한 시뮬레이션 결과가 재현되지 않기 때문에 모델의 신뢰도를 높이기 위해 파라미터의 제한 조건이 추가로 필요할 것으로 보인다.

본 연구의 성과 중 하나는 학습 심리학 이론이나 프로그래밍에 능통하지 않는 연구자가 쉽게 연합 모델의 예측을 검사해 볼 수 있는 사용자 환경을 구현했다는 점이다. 즉, 각 연합학습 모델의 기본적인 작동 알고리즘만 알고 있다면 본 연구의 결과로 만들어진 공개 소프트웨어 (KUCS: Korea University Conditioning Simulator)를 이용해서 간편하게 결과를 산출해 볼 수 있다. KUCS에 포함된 6개의 연합학습 모델은 최대한 원래 모델의 핵심 아이디어를 반영하여 구현하고자 하였으나, 모델 자체가 일부 개념적 수준에서의 정의를 수학적 공식으로 제공하지 않는 경우 (예: Mac 모델)도 있었고 다른 모델들과 완벽하게 호환되는 결과함수를 제공하지 않는 경우 (예: EH 모델)도 있었다. 그런 경우 하나의 플랫폼 위에서 모든 모델을 정량적으로 비교 분석하기 위해서 모델의 일부만을 구현하거나, 변형하여 사용하였다. 예를 들어 EH 모델의 경우 학습의 영향으로 변화하는 연합가능성인 ϵ 값이 따로 있지만 이를 포함하는 ϕ 값을 비교에 사용하였다. 또한, latent inhibition을 설명하기 위해 도입된 ΔV^{pre-s} 값의 경우 저자들이 계산 공식을 의도적으로 명시하지 않았기에 RW 모델을 사용해서 구현하였다. 이러한 변경들은 모델의 핵심 아이디어를 위반하지 않는 선에서 진행되었기 때문에 모델에 예측 결과에 큰 영향을 미치지 않을 것으로 보인다. 그럼에도 이러한 모델들 간의 통합적 비교를 위해서는 더 많은 시뮬레이션 데이터와 각 모델을 제안한 연구자들과의 지속적인 논의가 더 필요할 것으로 예상된다.

또 다른 개선점은 본 연구에 포함된 6개의 연합학습 모델 외에도 다른 학습 모델들이 존재한다는 사실이다. 특히 최근에 등장한 다양한 기계학습 알고리즘들 또한 연합학습의 관점에서 검증될 경우 흥미로운 결과가 예상된다. 또 다른 모델들은 연합에 의한 “학습”과정 뿐 아니라 형성된 “기억”의 표상이 처리되는 과정을 구현하여 좀 더 포괄적인 학습과 기억 현상들을 설명하고자 하였다. 예를 들어 A. Wagner (1981)에 의해 제안된 Sometimes Opponent Process (SOP) 모델은 조건자극의 표상을 포함했으며 그 뒤로 등장한 좀 더 복잡한 다수의 표상을 가지는 모델들 (예: AESOP model,

Wagner & Brandon, 1989)들의 원형을 제공하였다. 추후 KUCS의 개발은 이러한 차세대 모델들의 작동방식을 포함하도록 확장되어야 할 것이다.

References

- Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proc Natl Acad Sci U S A*, 108(25), 10367-10371. doi:10.1073/pnas.1104047108
- Cho, S. A., & Cho, Y. S. (2021). Uncertainty modulates value-driven attentional capture. *Atten Percept Psychophys*, 83(1), 142-155. doi:10.3758/s13414-020-02171-3
- Esber, G. R., & Haselgrove, M. (2011). Reconciling the influence of predictiveness and uncertainty on stimulus salience: a model of attention in associative learning. *Proc Biol Sci*, 278(1718), 2553-2561. doi:10.1098/rspb.2011.0836
- Harris, J. A., Kwok, D. W. S., & Gottlieb, D. A. (2019). The partial reinforcement extinction effect depends on learning about nonreinforced trials rather than reinforcement rate. *J Exp Psychol Anim Learn Cogn*, 45(4), 485-501. doi:10.1037/xan0000220
- Haselgrove, M., Aydin, A., & Pearce, J. M. (2004). A partial reinforcement extinction effect despite equal rates of reinforcement during Pavlovian conditioning. *J Exp Psychol Anim Behav Process*, 30(3), 240-250. doi:10.1037/0097-7403.30.3.240
- Humphreys, L. G. (1939). The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. *Journal of Experimental Psychology*, 25(2), 141-158. doi:DOI 10.1037/h0058138
- Kamin, L. J., & Szakmary, G. A. (1977). Configural-like effects observed in the course of second-order conditioning in rats. *Learning and Motivation*, 8(1), 126-135.
- Lipp, O. V., Stephens, J., & Smith, T. A. (1999). RWMODEL II: computer simulation of the Rescorla-Wagner model of Pavlovian conditioning. *Behav Res Methods Instrum Comput*, 31(4), 735-736. doi:10.3758/bf03200756
- Mackintosh, N. J. (1975). Theory of Attention - Variations in

- Associability of Stimuli with Reinforcement. *Psychological Review*, 82(4), 276-298.
doi:DOI 10.1037/h0076778
- Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla-Wagner Model. *Psychological Bulletin*, 117(3), 363-386.
doi:Doi 10.1037/0033-2909.117.3.363
- Mercier, P. (1996). Computer simulations of the Rescorla-Wagner and Pearce-Hall models in conditioning and contingency judgment. *Behavior Research Methods Instruments & Computers*, 28(1), 55-60.
doi:Doi 10.3758/Bf03203636
- Moore, J. W., & Choi, J. S. (1997). Conditioned response timing and integration in the cerebellum. *Learn Mem*, 4(1), 116-129. doi:10.1101/lm.4.1.116
- Mowrer, O. H., & Jones, H. (1945). Habit Strength as a Function of the Pattern of Reinforcement. *Journal of Experimental Psychology*, 35(4), 293-311. doi:DOI 10.1037/h0056678
- Pavlov, I. P. (1927). *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex*. Oxford University Press, London. .
- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, 87(6), 532. doi:10.1037/0033-295X.87.6.532
- Rescorla, R. A. (1999a). Partial reinforcement reduces the associative change produced by nonreinforcement. *Journal of Experimental Psychology-Animal Behavioral Processes*, 25(4), 403-414. doi:10.1037/0097-7403.25.4.403
- Rescorla, R. A. (1999b). Within-subject partial reinforcement extinction effect in autoshaping. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology*, 52(1), 75-87.
doi:10.1080/713932693
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical conditioning: current research and theory*. New York: Appleton-Century-Crofts.
- Schmajuk, N. A., & DiCarlo, J. J. (1991). A neural network approach to hippocampal function in classical conditioning. *Behav Neurosci*, 105(1), 82-110.
doi:10.1037//0735-7044.105.1.82
- Schmajuk, N. A., & Moore, J. W. (1985). Real-Time Attentional Models for Classical-Conditioning and the Hippocampus. *Physiological Psychology*, 13(4), 278-290.
Retrieved from <Go to ISI>://WOS:A1985A611600010
- Sheffield, V. F. (1949). Extinction as a function of partial reinforcement and distribution of practice. *J Exp Psychol*, 39(4), 511-526. doi:10.1037/h0057242
- Sutton, R. S., & Barto, A. G. (1987). *A temporal-difference model of classical conditioning*. Paper presented at the Proceedings of the ninth annual conference of the cognitive science society.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2 ed.): MIT press.
- Thomas, D. (1970). Stimulus selection, attention, and related matters. *Current issues in animal learning*, 311-356.
- Wagner, A. (1981). SOP: A model of automatic processing in animal behavior. *Information Processing in Animals: Conditioned Inhibition*.
- Wagner, A. R., & Brandon, S. E. (1989). Evolution of a structured connectionist model of Pavlovian conditioning (AESOP). In *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory*. (pp. 149-189). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- Yin, H., Barnet, R. C., & Miller, R. R. (1994). Second-order conditioning and Pavlovian conditioned inhibition: operational similarities and differences. *J Exp Psychol Anim Behav Process*, 20(4), 419-428.
doi:10.1037/0097-7403.20.4.419

자극의 불확실성이 연합학습에 미치는 영향: 수리모델간 비교연구

정지훈, 조양석, 최준식

고려대학교 심리학부

보상 정보의 불확실성은 정서 조절부터 의사결정 최적화까지 다양한 영역에 영향을 끼친다. 그러나 불확실성은 정량적인 연구가 진행되기보다는 다양한 심리 현상들의 부수적인 설명으로 사용되어왔다. 이러한 문제를 해결하는 한 가지 방법은 계산 모델을 구현해, 이를 동물과 사람 실험 결과에 적용해서 불확실성이 포함된 조건화 상황에서 각 모델의 예측치를 비교하는 것이다. 본 연구에서는 Matlab 기반 연합학습 모델 시뮬레이션 통합환경(Korea University Conditioning Simulator: KUCS)을 구축하였다. 이 시뮬레이터에서는 GUI 기반으로 다양한 학습 계획을 적용해서 Rescorla-Wagner 모델, Mackintosh 모델, Pearce-Hall 모델, Schmajuk-Pearce-Hall 모델, Esber-Hasselgrove 모델, Temporal Difference 모델의 특성을 비교할 수 있다. 검증을 위하여 KUCS를 사용하여 기초적인 연합 조건화 현상인 acquisition, extinction, blocking, conditioned inhibition, latent inhibition, second-order conditioning을 각 모델에서 구현할 수 있는지 확인했고, 모델의 한계점과 예측에 관한 몇 가지 새로운 사실을 발견하였다. 또한, 보상의 불확실성이 존재하는 학습 계획을 사용해 모델의 연합력(association strength)과 연합가능성(associability) 값이 실제 동물과 사람 실험 데이터와 일치하는지를 확인하였다. 시뮬레이터 프로그램은 <https://github.com/knowblesse/KUCS>에 공개하였다.

주제어: 수리모델, 연합학습, 보상 불확실성

부록

부록 1. 연구 1의 학습 계획

Appendix Table 1. Experiment schedules

	Acquisition	Extinction (Testing)
Acquisition/ Extinction	A+ 100 trials	A- 100 trials
Conditioned inhibition	A+ 100 trials	AB- 100 trials
Blocking	A+ 20 trials	AB+ 100 trials
	rest	AB+ 100 trials
Latent inhibition	A- 100 trials	A+ 100 trials
	rest	B+ 100 trials
Second-order conditioning	A+ 100 trials	AB- 20 trials

부록 2. 시뮬레이션에 사용한 모델 수식

Appendix Table 2. Model equations

	Equation	Default parameters
RW	$V_{t+1}^s = V_t^s + \Delta V_t^s$ $\Delta V_t^s = \begin{cases} \alpha^s \beta_t (\lambda_t - \sum_{s' \in t} V_t^{s'}) & \text{if CS } s \in \text{ Trial } t \\ 0 & \text{otherwise} \end{cases}$ $\beta_t = \begin{cases} \beta_{acq} & \lambda_t > 0 \\ \beta_{ext} & \lambda_t = 0 \end{cases}$	$\beta_{acq} = 0.1$ $\beta_{ext} = 0.05$
	<p>V_t^s: Association strength of CS s at trial t</p> <p>λ_t: US intensity at trial t</p> <p>α^s: Saliency of CS s (Usually, the physical intensity of the stimulus)</p> <p>β_t: Learning rate at trial t</p> <p>$\sum_{s' \in t} V_t^{s'}$: Sum of all V values of CS s presented at trial t</p>	
Mac	$V_{t+1}^s = V_t^s + \Delta V_t^s$ $\Delta V_t^s = \begin{cases} \alpha_t^s \beta_t (\lambda_t - V_t^s) & \text{if CS } s \in \text{ Trial } t \\ 0 & \text{otherwise} \end{cases}$ $\beta_t = \begin{cases} \beta_{acq} & \lambda_t > 0 \\ \beta_{ext} & \lambda_t = 0 \end{cases}$ $\alpha_{t+1}^s = \alpha_t^s + \Delta \alpha_t^s$ $D_t^s = \lambda_t - V_t^s , \quad D_t^{X_s} = \lambda_t - ((\sum_{s' \in t} V_t^{s'}) - V_t^s) $ $\Delta \alpha_t^s = \begin{cases} D_t^s < D_t^{X_s} & k(1 - \alpha_t^s)(D_t^{X_s} - D_t^s)/2 \\ D_t^s = D_t^{X_s} & -k\epsilon \\ D_t^s > D_t^{X_s} & k\alpha \end{cases}$	$\beta_{acq} = 0.08$ $\beta_{ext} = 0.04$ $k = 0.05$ $\epsilon = 0.02$
	<p>D_t^s: Discrepancy between the real and predicted US intensity by CS s at trial t</p> <p>$D_t^{X_s}$: Discrepancy between the real and predicted US intensity by all other presented CS except CS s at trial t</p> <p>k: Proportional parameter ($0 \leq k$) (The degree to which predictability affects association changes)</p> <p>ϵ: Small value to make $\Delta \alpha_t^s$ negative when $D_t^s = D_t^{X_s}$</p>	
PH	$\dot{V}_t^s = V_t^s - \bar{V}_t^s$ $V_{t+1}^s = V_t^s + \Delta V_t^s \quad \Delta V_t^s = \begin{cases} S^s \alpha_t^s \lambda_t & \text{if CS } s \in \text{ Trial } t \\ 0 & \text{otherwise} \end{cases}$ $\bar{V}_{t+1}^s = \bar{V}_t^s + \Delta \bar{V}_t^s \quad \Delta \bar{V}_t^s = \begin{cases} S^s \alpha_t^s \bar{\lambda}_t & \text{if CS } s \in \text{ Trial } t \\ 0 & \text{otherwise} \end{cases}$ $\bar{\lambda}_t = (\sum_{s' \in t} V_t^{s'} - \sum_{s' \in t} \bar{V}_t^{s'}) - \lambda_t$ $\alpha_{t+1} = \lambda_t - (\sum_{s' \in t} V_t^{s'} - \sum_{s' \in t} \bar{V}_t^{s'}) $	$S^A = 0.04$ $S^B = 0.04$
	<p>\dot{V}_t^s: Net prediction of US</p> <p>\bar{V}_t^s: Associative strength between CS s and no US in trial t</p> <p>S^s: Saliency of CS s</p>	
SPH	$\dot{V}_t^s = V_t^s - \bar{V}_t^s$ $V_{t+1}^s = V_t^s + \Delta V_t^s \quad \Delta V_t^s = \begin{cases} S^s \alpha_t^s \beta_{ext} \lambda_t & \lambda_t - \sum_{s' \in t} \dot{V}_t^{s'} > 0 \\ 0 & \text{otherwise} \end{cases}$ $\bar{V}_{t+1}^s = \bar{V}_t^s + \Delta \bar{V}_t^s \quad \Delta \bar{V}_t^s = \begin{cases} S^s \alpha_t^s \beta_{inh} \bar{\lambda}_t & \lambda_t - \sum_{s' \in t} \dot{V}_t^{s'} \leq 0 \\ 0 & \text{otherwise} \end{cases}$ $\bar{\lambda}_t = (\sum_{s' \in t} \dot{V}_t^{s'}) - \lambda_t$ $\alpha_t^s = \left\{ \gamma \lambda_t - \sum_{s' \in t} \dot{V}_t^{s'} \right\}$	$S^A = 0.3$ $S^B = 0.3$ $\beta_{ext} = 0.1$ $\beta_{inh} = 0.09$ $\gamma = 0.2$
	<p>γ: Proportional parameter (The degree to which uncertainty affects association changes)</p>	

EH	$\dot{V}_t^s = V_t^s - \overline{V}_t^s$	
	$V_{t+1}^s = V_t^s + \Delta V_t^s$	
	$\Delta V_t^s = \begin{cases} \alpha_t^s \beta_1 (\lambda_t - (\sum_{s' \in t} V^{s'} - \sum_{s \in t} \overline{V}^{s'})) & \text{if CS } s \in \text{Trial } t \\ 0 & \text{otherwise} \end{cases}$	$\beta_{1_{\text{acq}}} = 0.05$ $\beta_{2_{\text{acq}}} = 0.03$
	$\overline{V}_{t+1}^s = \overline{V}_t^s + \Delta \overline{V}_t^s$	$\beta_{1_{\text{ext}}} = 0.04$ $\beta_{2_{\text{ext}}} = 0.02$
	$\Delta \overline{V}_t^s = \begin{cases} \alpha_t^s \beta_2 ((\sum_{s' \in t} V^{s'} - \sum_{s \in t} \overline{V}^{s'}) - \lambda_t) & \text{if CS } s \in \text{Trial } t \\ 0 & \text{otherwise} \end{cases}$	$k = 0.2$ $\beta^{\text{pre-}s} = 0.02$
	$V_{t+1}^{\text{pre-}s} = V_t^{\text{pre-}s} + \Delta V_t^{\text{pre-}s}$	
	Δ	

Φ : Unacquired associability value of CS s (Innate associability of the stimulus. Usually, the physical properties affects the value)

ϵ : Acquired associability value of CS s

$w_{t+1}^s = w_t^s + \Delta w_t^s$	$\beta = 0.875$
$\Delta w_t^s = c(\lambda_t + \gamma \max(\mathbf{w}_t^T \mathbf{x}_t, 0) - \max(\mathbf{w}_t^T \mathbf{x}_{t-1}, 0)) x_t^s$	$c = 0.08$
$\overline{x}_t^s = \beta x_{t-1}^s + (1 - \beta) x_{t-1}^s$	$\gamma = 0.95$

TD	\overline{x}_t^s : Eligibility trace of CS s at trial t
	w_t^s : Weight of CS s at trial t . Analogue of V_t^s
	$\mathbf{w}_t, \mathbf{x}_t$: Weight and CS vector of all CS at trial t
	c : Learning rate
	β : Eligibility trace parameter
	γ : Relative importance between presence and onset/offset of the CS