A Experiment Descriptions

Acquisition_ContinuousVsPartial [Wagner et al., 1967]: With repeated CS-US pairings, CS elicits CR that increases in magnitude and frequency with further reinforcement. Partial reinforcement leads to slower acquisition and a lower conditioning asymptote. Data: rat, fear conditioning, CS visual + US shock → auditory startle.

Extinction_ContinuousVsPartial [Wagner et al., 1967]: When CS-US pairings are followed by presentations of CS alone or unpaired CS and US, the CR decreases. Partial reinforcement leads to slower extinction and a higher conditioning asymptote. Data: rat, fear conditioning, CS visual, auditory + US shock → bar pressing.

Generalization_NovelVsInhibitor [Kutlu and Schmajuk, 2012a]: Adding a novel stimulus C to a trained stimulus A results in a smaller decrease in CR than does adding a conditioned inhibitor X to a trained stimulus A. Data: human, value prediction, CS visual + US value → value prediction.

Generalization_AddVsRemove [Brandon et al., 2000]: Adding a cue to a trained compound results in a smaller decrease in CR than does removing a cue from a trained compound. Data: rabbit, eyeblink conditioning, CS visual, auditory, tactile + US shock → eyeblink.

Discrimination_ReinforcedVsNonreinforced [Campolattaro et al., 2008]: A reinforced CS elicits significantly greater CR than a non-reinforced CS. Data: rat, eyeblink conditioning, CS visual, auditory + US shock → eyeblink.

Discrimination_PositivePatterning [Bellingham et al., 1985]: Reinforced AB+ intermixed with non-reinforced A- and B- results in responding to AB that is stronger than the sum of the individual responses to A and B. Data: rat, appetitive conditioning, CS visual, auditory + US water → drinking.

Discrimination_NegativePatterning [Bellingham et al., 1985]: Non-reinforced AB- intermixed with reinforced A+ and B+ results in responding to AB that is weaker than the sum of the individual responses to A and B. Data: rat, appetitive conditioning, CS visual, auditory + US water → drinking.

Discrimination_NegativePatterningCommonCue [Redhead and Pearce, 1998]: Adding a common cue C to negative patterning decreases discrimination. Data: pigeon, appetitive conditioning, CS visual + US food → feeding.

Discrimination_NegativePatterningThreeCues [Redhead and Pearce, 1995]: Non-reinforced ABC- intermixed with reinforced A+ and BC+ results in responding to ABC that is weaker than the sum of the individual responses to A and BC. Data: pigeon, appetitive conditioning, CS visual + US food → feeding.

Discrimination_Biconditional [Saavedra, 1975]: Biconditional discrimination between compounds (AC+/BD+ vs. AD-/BC-, where no single CS predicts reinforcement or non-reinforcement) is possible but harder than component discrimination between compounds (AC+/AD+ vs. BC-/BD-, where A and B predict reinforcement and non-reinforcement, respectively). Data: rabbit, eyeblink conditioning, CS visual, auditory + US shock → eyeblink.

Discrimination_FeaturePositive [Ross and Holland, 1981]: Reinforced BA+, alternated with non-reinforced A-, results in stronger responding to BA than A alone. In the simultaneous case (BA+), B gains an excitatory association with the US; in the serial case (B → A+), B does not gain an excitatory association with the US. Data: rat, appetitive conditioning, CS visual, auditory + US food → head jerk.

Discrimination_FeatureNegative [Holland, 1984]: Non-reinforced BA-, alternated with reinforced A+, results in weaker responding to BA than A alone. In the simultaneous case (BA-), B gains an inhibitory association with the US; in the serial case (B → A-), B does not gain an inhibitory association with the US. Data: rat, fear conditioning, CS visual, auditory + US shock → bar pressing.

Inhibition_InhibitorExtinction [Zimmer-Hart and Rescorla, 1974]: Inhibitory conditioning to X trained via A+ → AX- is extinguished by AX+ presentations. Data: rat, fear conditioning, CS visual, auditory + US shock → bar pressing.

`Competition_RelativeValidity` [Wagner et al., 1968]: Conditioning to X is weaker when training consists of pairing X with stimuli A/B that are correlated with reinforcement, than when training consists of pairing X with stimuli A/B that are not correlated. Data: rat, appetitive conditioning, CS visual, auditory + US food → bar pressing.

`Competition_OvershadowingAndForwardBlocking` [Holland and Fox, 2003]: Training AB+ results in weaker conditioning to A than training A+ alone (overshadowing). Training B+ → AB+ results in even weaker conditioning to A (forward blocking). Data: rat, appetitive conditioning, CS visual, auditory + US food → feeding.

`Competition_Unblocking` [Dickinson et al., 1976]: In forward blocking B → AB+, increasing or decreasing the US during AB presentation can increase responding to the blocked A. Data: rat, fear conditioning, CS visual, auditory + US shock → bar pressing.

`Competition_BackwardBlocking` [Miller and Matute, 1996]: Training AB+ → B+ results in weaker conditioning to A than training A+ alone (backward blocking). Data: rat, fear conditioning, CS auditory + US shock → drinking.

`Competition_Overexpectation` [Rescorla, 1970]: Training A+ → B+ → AB+ results in lower conditioning to A than without the AB+ compound. Data: rat, fear conditioning, CS visual, auditory + US shock → bar pressing.

`Competition_Superconditioning` [Rescorla, 1971]: Training B- → AB+ (superconditioning) results in higher conditioning to A than training AB+ only (overshadowing) and yet higher than training B+ → AB+ (forward blocking). Data: rat, fear conditioning, CS visual, auditory + US shock → bar pressing.

`PreExposure_LatentInhibitionVsPerceptualLearning` [Lubow et al., 1976]: Pre-exposure of CS A- before A+ pairings can result in reduced responding (latent inhibition) or increased responding (perceptual learning) depending if the context is the same or different, respectively. Data: rat, appetitive conditioning, CS olfactory + US food → feeding.

`PreExposure_USPreExposure` [Kamin, 1961]: Pre-exposure of US + before A+ pairings results in decreased responding. Data: rat, fear conditioning, CS auditory + US shock → bar pressing.

`Transfer_Reacquisition` [Ricker and Bouton, 1996]: Following acquistion A+ and extinction A-, A+ pairings can result in faster or slower reacquisition depending on the number of extinction trials. Data: rat, appetitive conditioning, CS visual, auditory + US food → feeding.

`Recovery_LatentInhibition` [Grahame et al., 1994]: Extensive exposure to the context after training results in reduction of latent inhibition. Data: rat, fear conditioning, CS auditory + US shock → drinking.

`Recovery_Overshadowing` [Matzel et al., 1985]: Extinction of B after overshadowing training AB+ results in increased responding to A. Data: rat, fear conditioning, CS visual, auditory + US shock → drinking.

`Recovery_ExternalDisinhibition` [Bottjer, 1982]: Presenting a novel stimulus immediately before a previously extinguished CS might produce renewed responding. Data: pigeon, appetitive conditioning, CS visual + US food → feeding.

`Recovery_SpontaneousRecovery` [Rescorla, 2004]: Presenting the CS some time after the subject has stopped responding might yield renewed responding. Data: rat, appetitive conditioning, CS visual, auditory + US food → feeding.

`Recovery_Renewal` [Harris et al., 2000]: After extinction, presentation of the CS in a novel context might yield renewed responding. Data: rat, fear conditioning, CS auditory + US shock → freezing.

`Recovery_Reinstatement` [Rescorla and Heth, 1975]: After extinction, presentation of the US in the context might yield renewed responding. Data: rat, fear conditioning, CS auditory + US shock → bar pressing.

`HigherOrder_SensoryPreconditioning` [Brogden, 1939]: When AB- pairings are followed by A+ pairings, presentation of B may generate a response. Data: dog, reflex conditioning, CS visual, auditory + US shock → flexion.

`HigherOrder_SecondOrderConditioning` [Yin et al., 1994]: When A+ pairings are followed by AB- pairings, presentation of B may generate a response. The number of BA- pairings determines whether second-order conditioning or conditioned inhibition is obtained. Data: rat, fear conditioning, CS auditory + US shock → drinking.

B Baseline Model Details

We implemented three standard classical conditioning models as baselines following Gershman [2015]. For all models, we used one-hot vector representations where the input is a conjunction of the CS and the context [Gershman, 2017]. We use value v_t as the response value, as conditioned responding to a stimulus is assumed to be proportional to the value for that stimulus.

For each experiment, we simulated 20 subjects for each model. All experiments were performed on a personal computer with 16 GB of RAM.

B.1 Rescorla-Wagner

The predicted reward (US) at time t, v_t, is a linear combination of the input stimuli, $\mathbf{x}_t$, weighted by their associative weights, $\mathbf{w}_t$:

$$v_t = \mathbf{w}_t^\top \mathbf{x}_t \tag{1}$$

Updating is governed by:

$$\mathbf{w}_{t+1} = \mathbf{w}_t + \alpha \mathbf{x}_t \delta_t \tag{2}$$

where the prediction error, δ_t, is the difference between the predicted and the actual reward:

$$\delta_t = r_t - v_t \tag{3}$$

Following Gershman [2015], we set the learning rate to $\alpha = 0.3$.

B.2 Kalman Filter

In a probabilistic interpretation of Rescorla-Wagner, the weights are assumed to be evolving according to a linear Gaussian dynamical system (LDS) [Gershman, 2015]:

$$\mathbf{w}_0 \sim \mathcal{N}(\mathbf{0}, \sigma_w^2 \mathbf{I}) \tag{4}$$

$$\mathbf{w}_n \sim \mathcal{N}(\mathbf{w}_{n-1}, \tau_w^2 \mathbf{I}) \tag{5}$$

$$r_n \sim \mathcal{N}(v_n, \sigma_r^2) \tag{6}$$

This induces a posterior over the weights which can be inferred using Kalman filtering:

$$\hat{\mathbf{w}}_{n+1} = \hat{\mathbf{w}}_n + \mathbf{k}_n \delta_n \tag{7}$$

$$\hat{\mathbf{\Sigma}}_{n+1} = \hat{\mathbf{\Sigma}}_n + \tau^2 \mathbf{I} - \mathbf{k}_n \mathbf{x}_n^\top (\hat{\mathbf{\Sigma}}_n + \tau^2 \mathbf{I}) \tag{8}$$

where the learning rate corresponds the Kalman gain:

$$\mathbf{k}_n = \frac{(\hat{\mathbf{\Sigma}}_n + \tau^2 \mathbf{I}) \mathbf{x}_n}{\mathbf{x}_n^\top (\hat{\mathbf{\Sigma}}_n + \tau^2 \mathbf{I}) \mathbf{x}_n + \sigma_r^2} \tag{9}$$

and the initial mean and covariance are $\hat{\mathbf{w}}_0 = \mathbf{0}$, $\hat{\mathbf{\Sigma}}_0 = \sigma_w^2 \mathbf{I}$.

Following Gershman [2015], we set the diffusion variance to $\tau^2 = 0.01$, the noise variance to $\sigma_r^2 = 1$ and the prior variance to $\sigma_w^2 = 1$.

B.3 Temporal Difference Learning

Temporal difference learning models seek to learn the expected discounted sum of future rewards:

$$V(\mathbf{x}_t) = \mathbb{E}\left[\sum_{k=0}^{\infty} r_{t+k}\right] \tag{10}$$

Assuming the value is approximated as a linear combination of the stimuli, similarly to Rescorla-Wagner, $V(\mathbf{x}_t) = \mathbf{w}_t^\top \mathbf{x}_t$, the associative weights are updated according to:

$$\hat{\mathbf{w}}_{t+1} = \hat{\mathbf{w}}_t + \alpha \mathbf{x}_t \delta_t \tag{11}$$

where the prediction error, δ_t, takes into account the future discounted rewards:

$$\delta_t = r_t + \gamma \hat{\mathbf{w}}_t^\top \mathbf{x}_{t+1} - \hat{\mathbf{w}}_t^\top \mathbf{x}_t \tag{12}$$

Following Gershman [2015], we set the discount factor to $\gamma = 0.98$ and the learning rate to $\alpha = 0.3$. In order to extend the stimulus representation across time, we used the complete serial compound representation [Sutton and Barto, 1990].