Title

Inhibitory S-R Associations Influence the Partial Reinforcement Extinction Effect

Permalink

https://escholarship.org/uc/item/3h34z46z

Journal

International Journal of Comparative Psychology, 36(1)

ISSN

0889-3675

Authors

Seitz, Benjamin M

Blaisdell, Aaron P

Publication Date

2024-04-03

DOI

10.46867/ijcp.2024.36.5639

Supplemental Material

https://escholarship.org/uc/item/3h34z46z#supplemental

Copyright Information

Peer reviewed

Inhibitory Stimulus–Response Associations Influence the Partial Reinforcement Extinction Effect

Benjamin M. Seitz[1], Alexandra Stolyarova[1], & Aaron P. Blaisdell[1]

[1] University of California, Los Angeles

The Partial Reinforcement Extinction Effect (PREE) is the finding of more rapid extinction to a continuously reinforced cue than to a partially reinforced cue. Despite its well-documented existence, it has challenged many theoretical accounts of learning that assume the likelihood of responding is correlated with the strength of the learned Stimulus (S) – Response (R) association. While a number of explanations (notably by Amsel (1967) and Capaldi (1966)) of the PREE have been proposed, Rescorla (1999) reported a PREE using an entirely within-subjects design—a finding that runs counter to these proposed explanations. We describe how the PREE might be explained by the formation of inhibitory S-R associations that develop during the acquisition of partial reinforcement, but not continuous reinforcement. To test this proposal, we propose a mathematical model of S-R formation and simulate the described result. We then trained pigeons to peck a response key in the presence of one partially reinforced stimulus but also omitted the response key during the training of an alternative partially reinforced stimulus—thus preventing the development of inhibitory S-Rs to that latter stimulus. We found evidence for the PREE only for the partially reinforced stimulus where inhibitory S-Rs were capable of being formed, thus supporting our proposal. Alternative explanations of our procedure and its relevance to broader computational rules of learning is discussed.

抑制性の刺激—反応連合が部分強化—消去効果に影響を及ぼす

部分的強化—消去効果（PREE：Partial Reinforcement Extinction Effect）とは、連続的に強化された手がかりよりも、部分的に強化された手がかりの方が、より速やかに消去されるという知見である。その存在は十分に立証されているにもかかわらず、反応が起こる確率が、学習した刺激(S)と反応(R)の結合の強さに相関していると仮定する学習に関して多くの理論的説明が挑んできた。PREE に関する多くの説明（特に Amsel (1967)と Capaldi (1966)による）が提唱されているが、Rescorla (1999)は、完全に被験者内デザインによる PREE（この発見は、それらの提唱された説明とは反するものであった）を報告した。我々は、PREE が、連続強化ではなく、部分強化の獲得時に発達する抑制性 S-R 連合の形成によってどのように説明されるかを述べる。この提案を検証するために、我々は S-R 形成の数学的モデルを考案し、前述の結果をシミュレーションした。そこでハトに対して、部分強化されたある刺激の下では反応キーをつつくように訓練し、部分強化された別の刺激の訓練下では反応キーを省略した（このように、後者の刺激に対する抑制性 S-R の発達を妨げた。）我々は、抑制性 S-R が形成可能な部分強化刺激においてのみ PREE の証拠を発見し、このことは我々の提言を支持した。本手法の代替的説明と、より広範な計算論的学習規則との関連性について議論する。

Asociaciones Estímulo-Respuesta Inhibitorias Influyen en el Efecto en la Extinción del Reforzamiento Parcial

El efecto en la extinción del reforzamiento parcial (EERP) es el hallazgo de una extinción más rápida desde una señal continuamente reforzada que desde una señal parcialmente reforzada. A pesar de su bien documentada existencia, el efecto ha cuestionado muchas explicaciones teóricas del aprendizaje que asumen que la probabilidad de responder está correlacionada con la fuerza de la aprendida asociación estímulo (S) – respuesta (R). Aunque se han propuesto varias explicaciones (en particular las de Amsel (1967) y Capaldi (1966)) del EERP, Rescorla (1999) reportó un EERP utilizando un diseño completamente intrasujeto, hallazgo que va en contra de estas explicaciones propuestas. Describimos cómo el EERP podría explicarse por la formación de asociaciones S-R inhibidoras que se desarrollan durante la adquisición de refuerzo parcial, pero no de refuerzo continuo. Para evaluar esta propuesta, proponemos un modelo matemático de formación S-R y simulamos el resultado descrito. Posteriormente, entrenamos palomas para responder a una clave de respuesta en presencia de un estímulo parcialmente reforzado, pero también omitimos la clave de respuesta durante el entrenamiento de un estímulo alternativo parcialmente reforzado, evitando así el desarrollo de S-R inhibidores de ese último estímulo. Encontramos evidencia para el PREE solo para el estímulo parcialmente reforzado donde los S-R inhibidores eran capaces de formarse, lo que soporta nuestra propuesta. Se discuten explicaciones alternativas de nuestro procedimiento y su relevancia para reglas computacionales más amplias del aprendizaje.

Inhibitory S-R Associations Influence the Partial Reinforcement Extinction Effect

The partial reinforcement extinction effect (PREE) is the finding of more rapid extinction of an instrumental or Pavlovian response that had previously been reinforced on a continuous schedule compared to a response that had only been partially reinforced (Humphreys, 1939). The PREE has been a challenge for many learning theories (e.g., Rescorla & Wagner, 1972) that assume the strength of conditioned responding to be a product of the associative strength between stimulus (S) and response (R), or cue and outcome. Because a continuously reinforced cue or response should have a stronger associative strength than a partially reinforced cue or response, it seems paradoxical that partially reinforced cues or responses would be more resistant to extinction.

The two most popular accounts of the PREE have been Amsel's (1967) frustration theory and Capaldi's (1966) sequential hypothesis. According to Amsel's theory, frustration is elicited by nonreinforcement on nonreinforced trials involving a partially reinforced cue or response. This frustration carries over to the following trial. When the subsequent trial sometimes is a rewarded trial, an association develops between the internal state of frustration and reward. Thus, frustration itself becomes a positive stimulus (S+) that motivates the conditional response. A S+ is always reinforced, and thus frustration never develops to it. Thus, when extinction treatment begins, and every trial evokes frustration due to nonreward, responding continues to be motivated to a partially reinforced cue or response, but not to a continuously reinforced cue or response. This extended responding to the partially reinforced cue but not the continuously reinforced cue results in the PREE.

Capaldi's (1966) sequential hypothesis is quite similar, except with memory for nonreward becoming a S+ to motivate conditional responding. For example, if during a partial-reinforcement schedule, a nonrewarded trial is followed by a rewarded trial, and the subject has a memory of nonreward (N) from the previous trial when the reward is delivered on the subsequent trial, N can acquire properties of a S+. Since all continuously reinforced trials are rewarded, the subject never experiences N, nor develops any association between N and reward in the continuous reward condition. Thus, during extinction, N continues to motivate responding in the partially reinforced condition, but not in the continuously reinforced condition, leading to the PREE.

While both theories have garnered much theoretical support, a clever experiment by Rescorla (1999) provides evidence that PREE might not be explained by either account. Rescorla used a completely within-subject procedure in which trials with separate partially reinforced and continuously reinforced cues were interleaved within a session. Thus, some continuously reinforced cues would have been experienced in a state of frustration or following N. During extinction, the partially reinforced cue extinguished more slowly than did the continuously reinforced cue, thereby demonstrating the PREE. This finding challenges both theoretical accounts of the PREE discussed above because both frustration and memory for nonreward should motivate responding on all trials during extinction, both to the partially reinforced cue or response and the continuously reinforced cue or response. Thus, according to Rescorla, frustration (Amsel, 1967) or memory for nonreward (Capaldi, 1966) should be elicited on each extinction trial, which in turn should contribute motivation to respond equally to continuously reinforced and partially reinforced cues on extinction trials.

The evidence of a within-subject PREE has now been replicated several times (Chan & Harris, 2017; Harris & Bouton, 2020; Harris et al., 2019) which further challenges the frustration and sequential hypothesis. However, each of these replications was conducted using rats whereas Rescorla's (1999) original demonstration was in pigeons. We sought to replicate Rescorla's findings of within-subject PREE in pigeons while also testing a plausible mechanism of the effect: the acquisition of an inhibitory S-R association during partial reinforcement.

Can Preventing Inhibitory S-Rs During Acquisition Affect the PREE?

While earlier models of learning such as the Law of Effect (Thorndike, 1911) posited that responses followed by unsatisfying outcomes (e.g., extinction) would simply weaken an existing S-R association, the last century of empirical work strongly supports the view that extinction results in the formation of a separate inhibitory association that interferes with or inhibits excitatory behavior (Bouton, 2004; Colwill, 1991; Rescorla, 1993; Todd et al., 2014). Thus, training on a partial reinforcement schedule should result in the learning of separate, strong excitatory S-Rs and inhibitory S-Rs to the same stimulus. Training on a continuously reinforced schedule, on the other hand, should only result in a strong excitatory S-R but no inhibitory S-R.

These mechanisms might explain why extinction is more rapid to a continuously reinforced cue. The change in associative strength has been shown to follow an error-reduction process. When a stimulus is followed by reward, the increment in excitatory associative strength is largest on the first rewarded trial, and progressively decreases as the association approaches asymptotic levels supported by the reward. Likewise, the increment in inhibitory strength is largest on the first nonrewarded trial, and progressively decreases as the inhibitory association approaches asymptotic levels supported by nonreward. These acquisition functions for excitatory and inhibitory associations have been best documented by Rescorla (2000; 2001). Also note that the rate at which excitatory and inhibitory associations reach asymptote may differ on a number of contextual factors (e.g., magnitude of reward, prior expectation of reward, generalization/discrimination) with the general pattern being that excitation develops quicker than inhibition (Rescorla, 1969). In extinction, the continuously reinforced cue should have an inhibitory S-R association that starts at zero strength and increases rapidly. The partially reinforced cue, however, has previously received a history of nonreward during acquisition, and thus should have a relatively well-established inhibitory S-R association. As a result, on each nonrewarded trial during extinction, its inhibitory S-R association should undergo much smaller increments in strength. Thus, while both continuously reinforced and partially reinforced cues undergo increments in inhibitory S-R strength during extinction, these negatively accelerating functions start at different levels of preexisting S-R inhibition, and thus change at different rates during extinction. Assuming adequate response acquisition to continuously reinforced and partially reinforced cues, the more rapid increase in inhibitory S-R learning to the continuously reinforced than partially reinforced cue during extinction predicts more rapid extinction to the continuously reinforced than partially reinforced cue—that is, the PREE.

Testing the Role of Inhibitory S-Rs in the PREE

If the basis for the PREE is the different levels of S-R inhibition at the start of extinction treatment, then blocking the development of inhibitory S-R associations during acquisition of the partially reinforced cue should prevent the PREE. That is, blocking the development of an inhibitory S-R association to the partially reinforced cue should result in both the partially reinforced and continuously reinforced cues starting extinction with the same zero strength inhibitory S-R association. As a result, they should both undergo similar rates of acquisition of the inhibitory S-R association during extinction, and thus the response should extinguish at the same rate to both cues.

This prediction can be tested by preventing responses from being made to a partially reinforced cue during nonreinforced trials in the acquisition phase of training. We achieved this in the following simulation and experiment by omitting presentation of the response key on nonreinforced trials (see Figure 1), which should prevent the formation or strengthening of an inhibitory S-R association to the partially reinforced cue on nonreinforced trials. Amsel (1967) and Capaldi's (1966) theories still predict the PREE, even when subjects are prevented from making the instrumental response on nonreinforced trials, because the partially reinforced cue is still partially reinforced in a Pavlovian manner. That is, on nonreinforced trials, the partially reinforced cue is present and nonreinforced. According to Amsel's (1967) theory, nonreinforcement would elicit frustration. Thus, on subsequent reinforced trials the frustration that remains from the prior trial now has the opportunity to become associated with, or serve as, a discriminative cue for reinforcement on the next reinforced trial. As a result, during extinction, frustration should slow down extinction of both partially reinforced and continuously reinforced cues using a within-subject procedure. An analogous prediction is made by Capaldi's (1966) sequential hypothesis in terms of memories of nonreward slowing down the rate of extinction on all trial types during extinction.

Figure 1

Example of trial screen displays used in Experiment 1
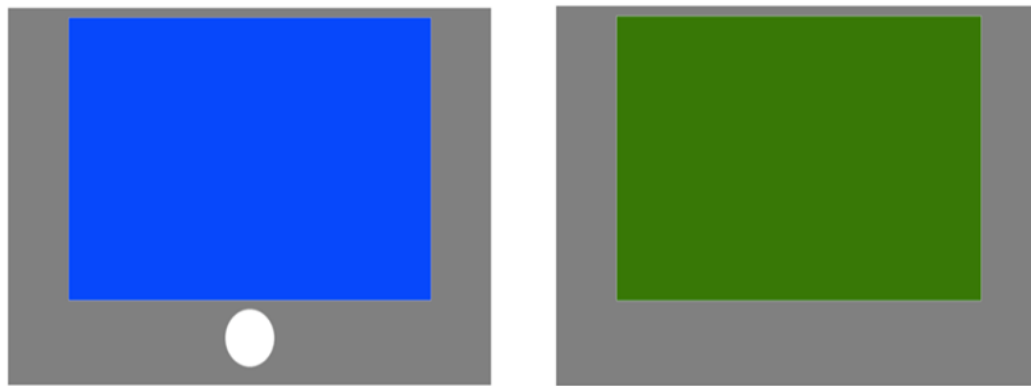
Note. Colored rectangles served as discriminative stimuli (cues) and a white circle served as the response key. On left, an example of a trial with one cue and the response key present. On right, an example of an acquisition-phase trial with partially-reinforced Cue Y and the response key absent.

S-R Learning Model and Simulated Experiment

The task implemented in the simulation was the same as the task used for pigeons as described for the experiment below, with the same numbers of sessions and trials, and the same reinforcement contingencies. In short, the model simulated response frequencies on a task for pigeons that involved key responses during the separate presentations of four different discriminative stimuli, A, B, X, and Y, which were colored rectangles presented on a computer monitor. Each trial involved the presentation of one of the discriminative stimuli, with an equal number of each trial type per session. During acquisition, key pecks during A and X were continuously reinforced, while key pecks during B and Y were reinforced on 25% of trials. During extinction, A continued to be reinforced at 100%, whereas X, B, and Y were nonreinforced on every trial (see Table 1). A critical manipulation is that there was no response key available on nonreinforced trials with Cue Y during acquisition, though the response key was available on Cue Y trials during extinction and on reinforced trials during acquisition (see Figure 2). This manipulation prevents the formation of an inhibitory S-R association to Cue Y during acquisition, but allows it to form and strengthen during extinction.

Table 1

*Reinforcement schedule for the experiment and simulations*

Cue	Acquisition	Extinction
A	100%	100%
B	25%	0%
X	100%	0%
Y	25% *key absent on nonreinforced trials	0%

Note. The experiment design was entirely within subject and each bird was exposed to 25 trials of each cue. Trials lasted a total of 20s. The first presentation of each cue was always reinforced, and the remaining 24 trials were either reinforced 100% of the time or 25% of the time (i.e. 6 reinforced trials). During nonreinforced Y trials, the key was not presented. During extinction, each cue was presented 24 times and the key was always present for all cues.

Figure 2

Simulated key pecks averaged across the 1000 agents following the training (upper panel) and extinction (lower panel) schedule described in Table 1
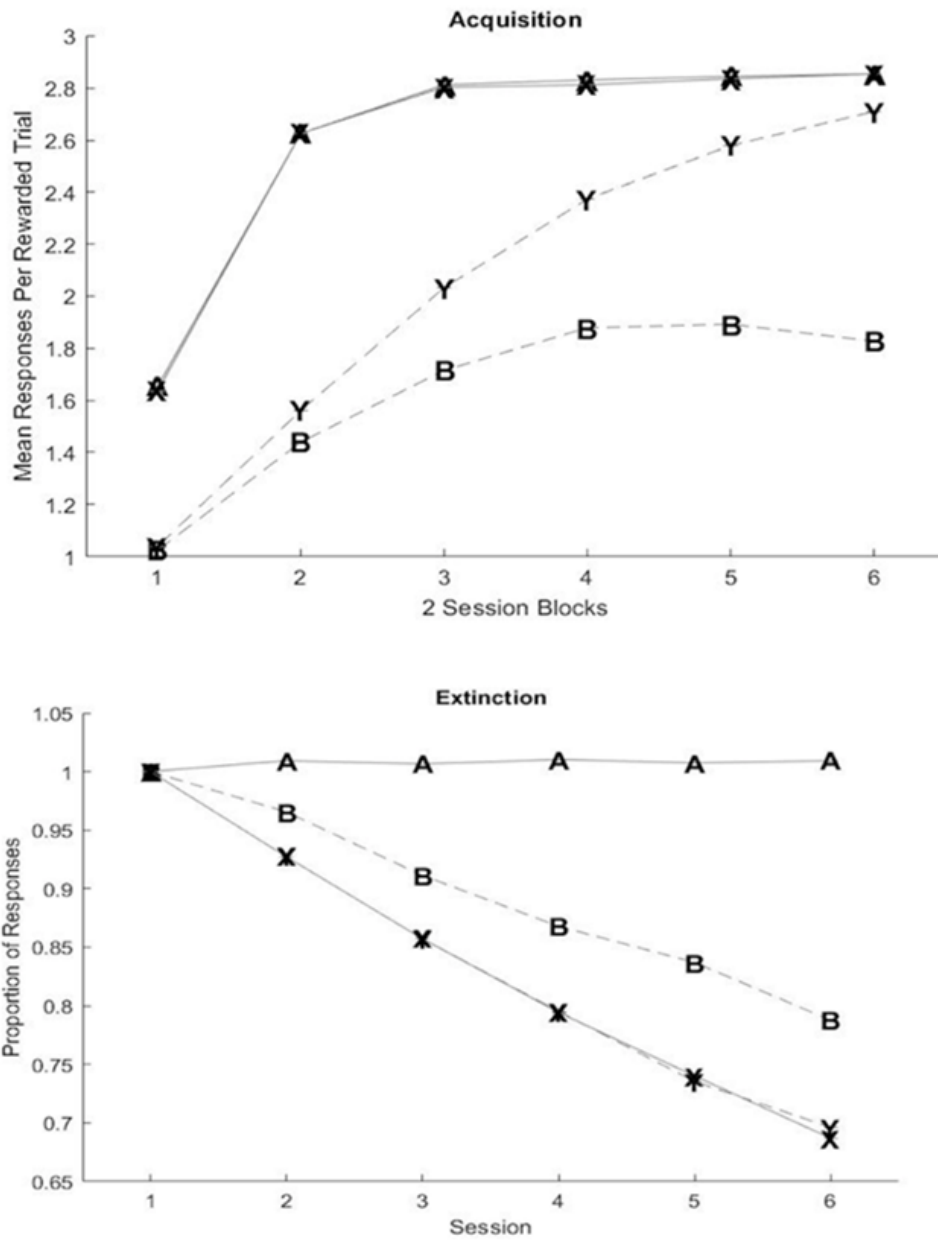
Note. Cues A and X were continuously reinforced and Cues B and Y were reinforced on 25% of trials but the key was not present on nonreinforced Y trials. Blocks represent two daily sessions in acquisition and one session in extinction. Upper) Mean number of key pecks on rewarded trials during the presentation of each cue throughout acquisition. Lower) Proportion of key pecks during the presentation of each cue as a function of the number of pecks made on the first day of extinction.

We simulated a total of 1000 agents performing the task (following Blaisdell et al., 2016). Each agent formed a set of excitatory and inhibitory S-R associations for the four stimuli used in the study. One of the responses contingent on each stimulus was the target action (meant to represent pecking the key). The other response represented all alternative actions; these actions were modeled as a single response because reinforcement contingencies for all behaviors other than pecking the key are the same (i.e., no reward delivery). At the beginning of the simulation each response had the same excitatory strength, $Q(R_{i,ex}|S_j)$, and the same inhibitory strength $Q(R_{i,in}|S_j)$. Here R_i represents response i and S_j represents stimulus j. All responses also had mutual inhibitory connections that were proportional to the net strength of the S-R associations:

$$\omega_{m\_in} \cdot (Q(R_{i,ex}|S_j) - Q(R_{i,in}|S_j)), \text{ where } \omega_{m\_in} \text{ is the strength of mutual inhibition.}$$

Therefore, the total strength of response i when presented with stimulus j, $Q(R_i|S_j)$, is calculated as:

$$Q(R_i|S_j) = Q(R_{i,ex}|S_j) - Q(R_{i,in}|S_j) - \omega_{m_{in}} \cdot (Q(R_{\sim i,ex}|S_j) - Q(R_{\sim i,in}|S_j)), .$$

Because each trial lasted 20 seconds for pigeons (*Columba livia*) thereby allowing for multiple actions, for the agents each trial was also broken down into several bins (set to 4) and a response was selected in each bin. The probability of the target response (key peck) in each bin was equal to $Q(R_i|S_j)$: a single uniformly distributed random number in the interval (0,1) was drawn in each bin and if $Q(R_i|S_j)$ exceeded that number the response was made.

The strength of excitatory and inhibitory associations was updated after the response was made and an outcome was observed. A response followed by a reward strengthens the excitatory S-R association. In contrast, a response followed by no reward leads to strengthening of a separate inhibitory S-R association. The rules for updating the excitatory and inhibitory associations were as follows:

1. When a response is made and a reward is delivered, the strength of the excitatory association for response i when presented with stimulus j, $Q(R_{i,ex}|S_j)$, is increased according to:

$$Q(R_{i,ex}|S_j) \leftarrow Q(R_{i,ex}|S_j) + \alpha_r \cdot (1 - Q(R_{i,ex}|S_j)),$$

where α_r is the updated rate for excitatory associations following reinforcement. The maximum possible value of $Q(R_{i,ex}|S_j)$ is 1 (because the rewarding event is represented as 1) and the change in the strength of the excitatory association is greater early during conditioning. As conditioning progresses $Q(R_{i,ex}|S_j)$ approaches its maximum value and the difference, $1 - Q(R_{i,ex}|S_j)$, that drives learning grows progressively smaller.

2. When a response is made and a reward is *not* delivered, the strength of the inhibitory association of response i when presented with stimulus j, $Q(R_{i,in}|S_j)$, is increased according to:

$$Q(R_{i,in}|S_j) \leftarrow Q(R_{i,in}|S_j) + \alpha_{nr} \cdot (1 - Q(R_{i,in}|S_j)).$$

Here, α_{nr} is the update rate for inhibitory associations following nonreinforcement.

The response data from this simulation were summarized in the same way as the pigeon behavioral data and the results are presented in Figure 2, separately for acquisition and extinction treatments. The following parameter values were used in the simulation presented in Figure 2: $\omega_{m\_in} = 0.2, \alpha_r = 0.02, \alpha_{nr} = 0.0015$. The strengths of excitatory and inhibitory associations at the beginning of the simulation for each agent were set to 0. The higher update rate for the strength of the excitatory, compared to inhibitory, S-R associations ($\alpha_r\ vs\ \alpha_{nr}$) is needed to compensate for the lower frequency of rewarded, compared to non-rewarded, trials. If the update rates are set to the same value, inhibition quickly overpowers excitation with a 75% non-reinforced schedule (because there are 5 bins in each trial, non-reinforcement can occur 15 times more frequently than reinforcement). The difference between these parameter selection rates of excitatory versus inhibitory learning are also consistent with the huge empirical literature showing that excitatory responding develops faster than inhibitory responding (both for Pavlovian conditioned inhibition and for discriminative operant responding), and for the ability to establish excitatory responding on lean schedules of reinforcement, such as 25% (or lower).

The simulation was implemented in MATLAB (MathWorks R2017a). Figure 2 shows the results of the simulation for both acquisition treatment (upper panel) and extinction treatment (lower panel). Continuously reinforced Cues A and X approach asymptotic levels of responding sooner (around Blocks 2 or 3) than partially reinforced Cue Y (around Blocks 5 or 6), for which the target response was not available on nonreinforced trials. Partially reinforced Cue Y did eventually reach a similar asymptotic level of responding as did the continuously reinforced cues. The slower rate of acquisition to Cue Y is due to having fewer reinforced trials per session. Partially reinforced Cue B also exhibited a slower rate of acquisition (around Blocks 5 or 6), but unlike Cue Y, Cue B also reached a lower asymptotic level of responding. The lower asymptote is due to the acquisition of an inhibitory S-R association that competes with the excitatory S-R association to produce a net level of responding lower than observed to the continuously reinforced cues for which no inhibitory S-R association was acquired. In theory, because the MLOE treats the excitatory and inhibitory associations as separate entities, with enough training, responding should eventually decrease back to 0 for a partially reinforced cue. This, of course, does not appear to be supported by empirical studies of partial reinforcement, and so future iterations of this model—or in fact any model that treats excitation and inhibition separately—should consider implementing rules by which excitation and inhibition interact with each other and when one might dominate the other.

The lower panel of Figure 2 depicts the proportion of simulated key pecks during extinction for the presentation of each cue as a function of the number of pecks made on the first day of extinction. This mirrors the normalization of extinction data in the experiments below. Cue A continues to be reinforced during extinction, and thus maintains 100% of responding as observed on Day 1 of extinction treatment. Continuously reinforced Cue X shows a more rapid rate of extinction than does partially reinforced Cue B (i.e., the PREE). Importantly, responding to partially reinforced Cue Y extinguishes as rapidly as that to continuously reinforced Cue X, and faster than to partially reinforced Cue B. This demonstrates the prediction that preventing key responses on nonreinforced trials of Cue Y during acquisition, should prevent the formation of an inhibitory S-R association for Cue Y, thus preventing the PREE to Cue Y. This is due to both Cues X and Y entering into the extinction treatment with a zero-strength inhibitory S-R association for the target response, and thus, the inhibitory S-R association should undergo rapid and large increases at equal rates for each cue.

Now that we have formally simulated how this model explains PREE and the prevention of PREE through the omission of the target response on nonreinforced trials with Cue Y, we turn to the empirical investigation of these predictions.

Experimental Test of the Role of S-R inhibitory Associations in the PREE

In the following experiment, we tested the rate of extinction to two partially reinforced cues—one for which the instrumental key peck response was prevented on nonreinforced trials during acquisition, and the other for which the instrumental key peck response was available on all trials during acquisition. We predicted faster extinction for the former cue over the latter.

To test the influence of S-R inhibitory associations in facilitating the PREE, pigeons were trained to peck a response key during the separate presentations of four different discriminative stimuli, A, B, X, and Y, which were colored rectangles presented on a computer monitor. Key pecks during A and X were continuously reinforced, while key pecks during B and Y were reinforced on 25% of trials (see Table 1). The critical element of our procedure was the omission of the response key on nonreinforced trials with Y. Omitting the response key should prevent the formation of an inhibitory S-R association between Cue Y and the key peck response. As a result, compared to extinction of continuously reinforced Cue X, we predicted similar, rapid extinction to Y, but slower extinction to B.

Methods

Subjects

Two male and two female homing pigeons (*Columba livia*) from Double T Farms (Glendwood, IA) served as subjects. Subjects were individually housed in steel home cages with metal wire mesh floors in a vivarium and they were maintained at 80% of their free-feeding weight. Testing occurred at approximately the midpoint of the light portion of the 12-hour light-dark cycle. Pigeons were between 2 and 16 years of age at the time of experimental treatment, and had prior experience in behavioral studies involving instrumental screen pecking and autoshaping, but were naïve with respect to the stimuli used in this study. All experimental procedures were conducted in accordance with the University of California, Los Angeles (UCLA) Institutional Animal Care and Use Committee.

Materials

Testing was conducted in a flat-black Plexiglas chamber (38 × 36 × 38 cm). Stimuli were generated on a monitor (NEC MultiSync LCD1550M). Pecks to the monitor were detected by an infrared touchscreen (Carroll Touch, Elotouch Systems, Fremont, CA) mounted on the monitor. A customized food hopper was located in the center of the front panel, its access hole was flush with the chamber floor and could deliver mixed grain. All experimental procedures were programmed using PsychoPy Software (Peirce et al., 2019). Four colored (red, blue, green, yellow) rectangles (20 × 15 cm) served as discriminative stimuli (DS) and are subsequently referred to as A, B, X, and Y. Color assignments for A, B, X, and Y were counterbalanced across birds. Below the DS, a 2.5 cm diameter white circle centered on the lower fifth of the screen served as the response key. See Figure 1 for an example screenshot of a trial display.

Procedure

A summary of the full reinforcement schedule for each cue is provided in *Table 1*.

Acquisition

A mixed autoshaping/operant shaping procedure to the key began by displaying the key below one of four colored rectangles (DS). Trials lasted 20 s and multiple rewards could be obtained on each rewarded trial. Each 60-min session consisted of the presentation of 25 trials of each stimulus (A, B, X, and Y). The first presentation of each stimulus was always rewarded. For the remaining 24 trials, Cues A and X were reinforced on each trial (continuously reinforced), while Cues B and Y were only reinforced on 25% of trials (partially reinforced). All reinforced trials ended with a non-contingent delivery of the reward (i.e., autoshaping). Stimulus presentations were followed by a variable inter-trial interval (ITI) with a mean time of 10 s, where the screen was gray and the key was present. On rewarded trials, each time the key was pecked, the subject received a reward, which consisted of a 3 s delivery of grain from the hopper. While the key was present on all trials with Cues A, X, B, and during the ITI, the key was not present on nonreinforced trials with Cue Y (see Figure 1). Nevertheless, the key was present for the 25% of Y trials that were reinforced, thus allowing excitatory but not inhibitory S-R associations to form. We also recorded pecks on nonreinforced Y trials from the screen location where the response key was otherwise present on other trials, which allowed us to continue to document the subject's behavior during these trials. Sessions were conducted five days per week and acquisition lasted a total of 12 daily sessions (one session per day).

Extinction

      After twelve acquisition sessions, peck rates to A, X, B, and Y had reached asymptotic levels (visually assessed, see Figure 3a), and birds were placed on an extinction schedule, in which key pecks to Cues X, B, and Y were never reinforced. Each cue was presented 24 times per session during extinction. Key pecks during presentations of Cue A continued to be reinforced on 100% of trials like what had occurred during acquisition. The key was present on all trials during extinction, including trials with Cue Y. The study concluded after 6 days of extinction training as peck rates had reached asymptotically low levels (visually assessed, see Figure 3b).

Figure 3

Key pecks averaged across the 4 pigeons
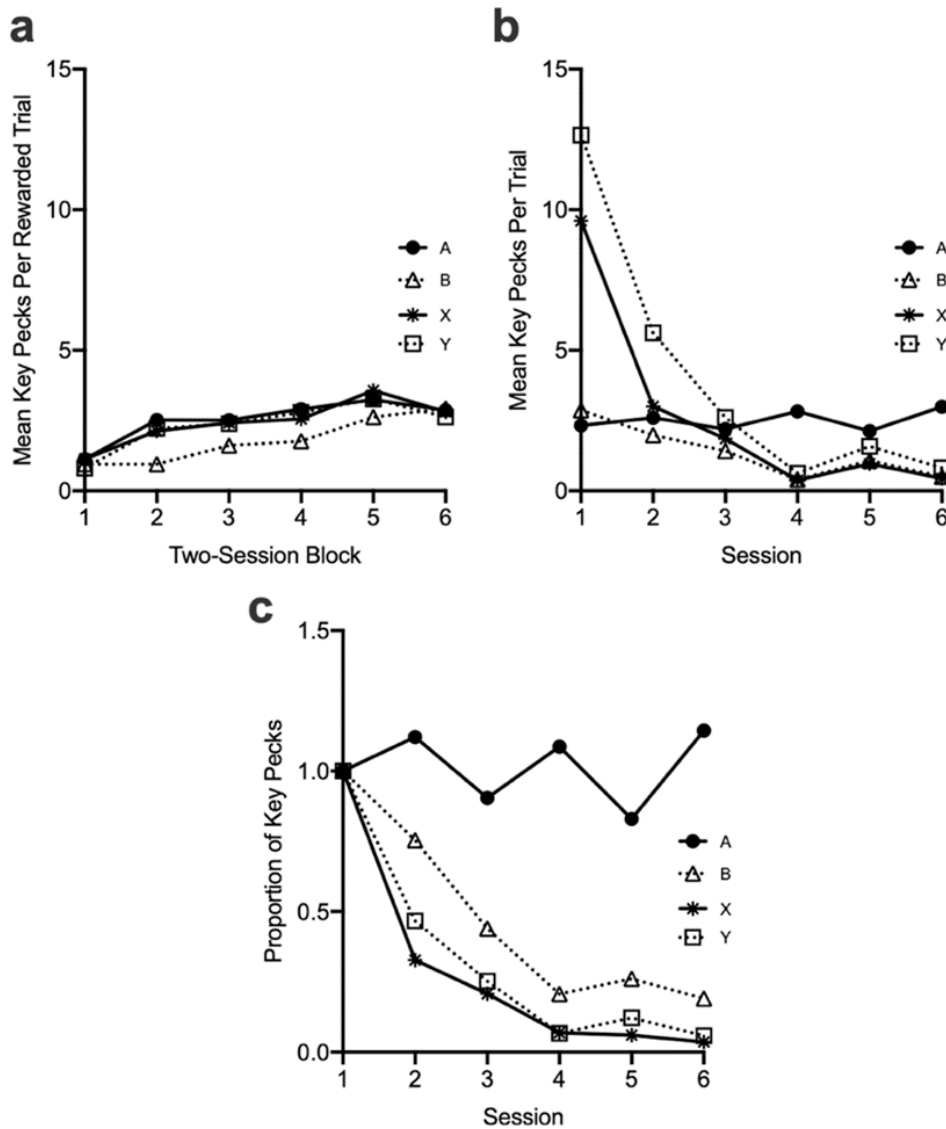
Note. Cues A and X were continuously reinforced and cues B and Y were reinforced on 25% of trials but the key was not present on nonreinforced Y trials. Blocks represent two daily sessions. a) Mean number of pecks made to the key on rewarded trials during presentation of each cue during acquisition. b) Mean number of pecks made to the key during the presentation of each cue during extinction. c) Proportion of pecks made to the key during the presentation of each cue as a function of the number of pecks made on the first day of extinction.

Results

The mean number of pecks made to the key on rewarded trials per acquisition session was averaged across all 4 birds and presented in 2-session blocks (Figure 3a, see figure S1 for responding on nonreinforced trials). Given the small sample, a Friedman's test was used in lieu of a repeated measures ANOVA, with Cue and Block used as grouping variables. This revealed a significant effect of Cue $\chi^2(3) = 29.95$, $p < .001$, Kendall's W = 0.695, and a significant effect of Block $\chi^2(5) = 28.41$, $p < .001$, Kendall's W = 0.523. Birds increased their responding from a mean of 0.5 pecks per 20 s trial in the beginning of training to a mean of 3 pecks per trial at the end of training. Birds appeared to learn to peck the key during the presentation of each cue at similar rates though key pecks were highest during nonreinforced B trials, (see S1) which should result in the acquisition of a strong inhibitory S-R association and an initial lag in responses for Cue Y.

The extinction phase lasted 6 days and the extinction data are presented in Figure 3b. While it is typical to directly compare rates of responding from acquisition to extinction, such comparison is complicated by the details of our procedure. Specifically, because the cue duration was 20 s, pigeons could receive multiple rewards on rewarded trials. This ultimately lowered responding relative to nonreinforced trials (see Figure S1 where responding is highest on nonreinforced B trials). Thus, on extinction trials, many more pecks could be made, because pigeons were not slowed down by receiving a reward. Interestingly, on the first day of extinction, responding to partially reinforced Cue B remained similar to responding on rewarded trials at the end of the acquisition phase. On the other hand, continuously reinforced Cue X and partially reinforced Cue Y elicited large extinction bursts on the first day of extinction (note continuously reinforced Cue A remained reinforced to motivate responding during the extinction sessions). Extinction bursts are not uncommon when transitioning from continuous reinforcement to an extinction schedule (Katz, 2018), but one must ask why an extinction burst was observed to partially reinforced Cue Y and not partially reinforced cue B. A frustration account is unsatisfying, because frustration putatively would have been experienced during the 75% of Y trials that were nonreinforced during training. By contrast, the lack of inhibitory S-R associations for Cues X and Y should result in an extinction burst observed to each cue at the beginning of extinction treatment.

Because of the extinction burst to Cues X and Y and the direct comparison of acquisition and extinction data, we then analyzed the proportion of pecks made to each cue during sessions 2-6 of extinction relative to the number of pecks made to each cue on the first day of extinction (Figure 3c). This approach mirrors that of Rescorla (1999, Exp 3) who examined performance during extinction as a percentage of the final asymptotic level for the same reason—that differences in absolute levels of response complicate the interpretation of extinction results. A Friedman's test with Block and Cue as grouping variables was used to analyze the normalized data, eliminating the first day due to no variability between cues. This revealed a significant effect of Cue, $\chi^2(3) = 35.91$, $p < .001$, Kendall's W = 0.834 and significant effect of Block, $\chi^2(4) = 29.95$, $p = .007$, Kendall's W = 0.495. A planned nonparametric comparison between partially reinforced Cue B and continuously reinforced Cue X revealed a significant difference, $\chi^2(1) = 5.32$, $p = .02$, Kendall's W = 0.761. This faster rate of extinction to continuously reinforced Cue X than to partially reinforced Cue B demonstrates the PREE. In line with the prediction that inhibitory S-Rs mediate the PREE, a planned nonparametric comparison between continuously reinforced Cue X and partially reinforced Cue Y revealed no difference, $\chi^2(1) = 0.62$, $p = .43$, Kendall's W = 0.213. Thus, partially reinforced Cue Y did not demonstrate the resistance to extinction that was observed for partially reinforced Cue B and that is typically observed in the PREE. This suggests acquiring an inhibitory S-R association is necessary for the PREE, and that preventing a key response on nonreinforced trials prevented the formation of this association.

Discussion

The PREE is one of the most robust and well-documented phenomena in associative learning. Nevertheless, the PREE appears paradoxical to most accounts of learning that presume responding to reflect the acquired associative strength between stimulus and response (or outcome). A number of theoretical accounts have been proposed to explain this effect, most notably Amsel's (1967) frustration account and Capaldi's (1966) sequential hypothesis. Nevertheless, both accounts struggle to explain how the PREE can be observed using a within-subjects paradigm (Rescorla, 1999).

We hypothesized that during the acquisition of a partially reinforced cue, the learning of inhibitory S-Rs might play a causal role in producing the PREE. The development of parallel excitatory and inhibitory associations fits within a more contemporary understanding of associative processes and has also been adopted in more recent models of learning. Indeed, the Modified Law of Effect (MLOE), proposed by Blaisdell et al. (2016), describes a S-R relationship with separate excitatory and inhibitory associations, as well as formalizing the role of response competition via mutual inhibition between possible actions. According to this theory for which our proposal is a test of, during extinction the rapid increase in the inhibitory S-R association for the continuously reinforced cue results in faster extinction compared to the partially reinforced cue which undergoes a comparatively slow increase in S-R inhibition due to already having acquired a strong inhibitory S-R association during acquisition (cf. Rescorla, 2001). Thus, the PREE can be diminished by preventing the participant from making the instrumental response on nonreinforced trials during acquisition. This is precisely what we found by omitting the response key on nonreinforced trials for a specific discriminative stimulus (Cue Y) that was only reinforced on 25% of trials. Both Amsel's (1967) and Capaldi's (1966) theories of the PREE do not predict that omitting the response key on nonreinforced trials will attenuate the PREE, because the cue remains partially reinforced which should cause frustration or a memory trace of nonreinforcement followed by reinforcement, and these processes should exert their effects on all cue types during extinction. Finally, we found no evidence of resistance to extinction to partially reinforced Cue Y which supports the role of inhibitory S-Rs influence on the PREE.

Two alternative explanations of our data should be considered. The first is that during acquisition, pigeons simply learned not to respond during Cue Y, thus making this an easy pattern to continue during extinction. We find this alternative explanation unsatisfactory. First, on reinforced trials when the key was present, responding to Cue Y was equally as robust as responding to the key while other cues were presented. Also, by the end of acquisition, some birds made pecks on nonreinforced Y trials to the location where the key would have been on reinforced trials (see Y_null in S1), which suggests they did not learn a response rule of not making responses during the nonreinforced Y trials.

The other alternative explanation is that during acquisition, nonreinforced trials of partially reinforced Cue Y may have been perceived differently than reinforced trials of this cue. This could serve as a context change and these contextual changes could be driving our effects. That said, Boughner and Papini (2006) report that the PREE persists even with changes in context from acquisition to extinction, suggesting this phenomenon might be somewhat immune to context effects. Moreover, the purpose of including the key during the ITI was to reduce the likelihood that the key would modulate instrumental behavior, and instead only serve as the target of instrumental behavior. Additionally, in several unpublished studies, our lab has consistently shown pigeons are capable of learning about discriminative stimuli in the presence/absence of various other response keys or physical stimuli in the operant chamber. As an example, we recently had pigeons complete a Pavlovian-Instrumental-Transfer procedure in which they first learned that a colored screen was a CS+ and then that a response key yielded the same reward as the CS+. During the transfer phase, pigeons were presented with the response key on top of the colored screen for the very first time. According to context-specific account, pigeons would perceive this transfer trial as entirely separate from that on which they had been previously trained. Contrary to this account, presentation of the CS+ at test elevated instrumental key pecking significantly above baseline (i.e., the key by itself).

Finally, it is worth calling attention to the fact that the PREE we observed was largely due to an extinction burst experienced to the continuously reinforced cue and partially reinforced cue whose response key was hidden during nonreinforcement. Extinction bursts are not always seen in demonstrations of the PREE but the fact that both these cues demonstrated an extinction burst while the traditionally trained partially reinforced cue did not, suggests inhibitory S-Rs might contribute to extinction bursts in addition to, and can potentially explain the PREE. Future experiments with procedures that minimize extinction bursts would be fruitful to tease apart this distinction. Discovering a similar lack of PREE to a partially reinforced cue and no extinction burst would be stronger evidence in support of the role inhibitory S-Rs moderating the PREE.

These results indicate pigeons are capable of generalizing prior learning to situations where novel stimuli are present or familiar stimuli are absent. Lastly, while a context-specific account might provide an alternative account of these data, this is true of many phenomena in associative learning. The possibility of virtually any environmental or internal state change to be viewed as a distinct context makes testing the context-specific account a protean endeavor. At the very least, the S-R inhibitory account provides the most parsimony and *a priori* predictive power, whereas context-modulation accounts are typically more easily constructed post-hoc.

It is worth considering one further discrepancy between our demonstration of PREE in a within subject procedure and others who have failed to demonstrate a PREE under some conditions using a within-subject procedure (reviewed in Nevin, 1988). Perhaps the most comparable experiment to ours was an experiment by Nevin (1985; reported in Nevin, 1988) in which pigeons received repeated cycles of training followed by extinction on a concurrent schedule involving two keys, one where pecks were reinforced on a continuous reinforcement (CRF) schedule and the other reinforced on a variable interval (VI) 30 s schedule (data presented in Figure 4 of Nevin, 1988). During Days 2 through 5 of daily acquisition/extinction cycles, response rates where much higher to the VI schedule than the CRF schedule, whereas by Days 75 through 125, response rates had dramatically diminished with little difference between schedules. More importantly, when analyzing the data as the proportion of number of responses emitted during the first 16 min of the total extinction session, he reported a small PREE on Days 2-5 (especially in the second quarter of extinction), but a much more robust effect of faster extinction to the CRF than VI schedule on Days 75-125. This latter effect is consistent with Nevin's Behavioral Momentum Theory, while the former is not, and perhaps reflects some other transient process that decays after sufficient training.

While the results form Days 75-125 appear to contradict our demonstration of PREE, perhaps they are actually consistent with it and are anticipated by the MLOE. Remember that the MLOE explains the PREE as the initially larger trial-wise increases in the inhibitory S-R to the CRF Cue X (and, in our study to Cue Y due to the omission of the response key on nonreinforced trials during acquisition). This should only be the case, however, for the FIRST session of extinction. During the first session of extinction, a strong inhibitory S-R association should be acquired to Cues X and Y. Thus, if pigeons received a round of reacquisition, followed by a second round of extinction, the large previously acquired inhibitory S-R associations for Cues X and Y should result in much smaller further increments in these inhibitory S-R associations. Thus, MLOE only predicts the PREE on the first round of extinction, but not on subsequent rounds of extinction if repeated rounds of acquisition followed by extinction are given.

Does the MLOE actually predict a reversal of the PREE following repeated rounds of acquisition/extinction such that the CRF cues should show faster extinction than PRF cues, as documented in Nevin (1988)? In our simulations of MLOE, we set the learning rate term for changes in excitatory associations to be greater than for changes in inhibitory associations. Indeed, there is strong empirical support that excitatory associations are acquired more rapidly than inhibitory associations which appear to more slowly (Harris et al., 2014; 2016; Rescorla, 2002). Thus, by incorporating this assumption into the updated rule for excitatory and inhibitory associative changes into the MLOE, it appears to be fully consistent with the results of the study by Nevin (1988) and behavioral momentum theory more generally. Perhaps MLOE provides a mechanistic instantiation of behavioral momentum theory.

In sum, we have proposed a novel explanation for the PREE and provide empirical evidence in support of this explanation and its predictions. Further investigations on the potential role of inhibitory S-Rs in mediating behavior may grant insight into the computational rules central to habit formation, addiction-like behaviors, general reinforcement learning, and the neural circuitry that govern these processes.

Acknowledgments

References

Amsel, A. (1967). Partial reinforcement effects on vigor and persistence: Advances in frustration theory derived from a variety of within-subjects experiments. *Psychology of Learning and Motivation*, *1*, 1–65. https://doi.org/10.1016/S0079-7421(08)60512-5

Blaisdell, A. P., Stolyarova, A., & Stahlman, W. D. (2016b). The law of expect or a modified law of effect? Outcome expectation and variation in learned behavior 1, 2. *Conductual*, *4*(2), 61–90. https://doi.org/ISSN: 2340-0242

Boughner, R. L., & Papini, M. R. (2006). Survival of the partial reinforcement extinction effect after contextual shifts. *Learning and Motivation*, *37*(4), 304–323. https://doi.org/10.1016/j.lmot.2005.09.001

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory*, *11*(5), 485–494. https://doi.org/10.1101/LM.78804

Capaldi, E. J. (1966). Partial reinforcement: A hypothesis of sequential effects. Psychological Review, 73(5), 459–477. https://doi.org/10.1037/h0023684

Chan, C. K. J., & Harris, J. A. (2017). Extinction of Pavlovian conditioning: The influence of trial number and reinforcement history. *Behavioural Processes*, *141*(Part 1), 19–25. https://doi.org/10.1016/J.BEPROC.2017.04.017

Colwill, R. M. (1991). Negative discriminative stimuli provide information about the identity of omitted response contingent outcomes. *Animal Learning & Behavior*, *19*(4), 326–336. https://doi.org/10.3758/BF03197893

Harris, J. A., Kwok, D. W. S., & Andrew, B. J. (2014). Conditioned inhibition and reinforcement rate. *Journal of Experimental Psychology: Animal Learning and Cognition, 40*, 335–354. http://dx.doi.org/10.1037/xan0000023

Harris, J. A., Patterson, A. E., Andrew, B. J., Kwok, D. W. S., & Loy, I. (2016). Asymmetry between excitatory and inhibitory learning. *Journal of Experimental Psychology: Animal Learning and Cognition, 42*, 297–312. http://dx.doi.org/10.1037/xan0000110

Harris, J. A., & Bouton, M. E. (2020). Pavlovian conditioning under partial reinforcement: The effects of nonreinforced trials versus cumulative conditioned stimulus duration. *Journal of Experimental Psychology: Animal Learning and Cognition*, *46*, 256–272. https://doi.org/10.1037/xan0000242

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. *Journal of Experimental Psychology: Animal Learning and Cognition*. https://doi.org/10.1037/XAN0000220

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. https://doi.org/10.1037/h0058138

Nevin J. A. (1985) Behavioral momentum and the partial reinforcement extinction effect. *Paper read at Psychonomic Society meeting, Boston, MA*.

Nevin, J. A. (1988). Behavioral momentum and the partial reinforcement effect. *Psychological Bulletin, 103*(1), 44–56. https://doi.org/10.1037/0033-2909.103.1.44

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H, Katsman, E., Lindeløv, J. K. (2019).

PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Rescorla, R A. (2001). Are associative changes in acquisition and extinction negatively accelerated? *Journal of Experimental Psychology. Animal Behavior Processes*, *27*(4), 307–315. http://www.ncbi.nlm.nih.gov/pubmed/11676082

Rescorla, Robert A. (1969). Pavlovian conditioned inhibition. *Psychological Bulletin*, *72*(2), 77–94. https://doi.org/10.1037/h0027760

Rescorla, R. A. (1993). Inhibitory associations between S and R in extinction. *Animal Learning & Behavior*, *21*(4), 327–336. https://doi.org/10.3758/BF03197998

Rescorla, R. A. (2000). Associative changes in excitors and inhibitors differ when they are conditioned in compound. *Journal of Experimental Psychology: Animal Behavior Processes*, *26*(4), 428–438. https://doi.org/10.1037//0097-7403.26.4.428

Rescorla, R. A. (1999). Within-subject partial reinforcement extinction effect in autoshaping. *Quarterly Journal of Experimental Psychology*, *52b*(1), 75–87. https://journals.sagepub.com/doi/pdf/10.1080/713932693

Recorla, R. A., & Wagner, A. R. (1972). A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement. In A. H. Black, & W. F. Prokasy (Eds.), Classical Conditioning II: Current Research and Theory (pp. 64-99). New York: Appleton- Century-Crofts. https://pdfs.semanticscholar.org/afaf/65883ff75cc19926f61f181a687927789ad1.pdf

Thorndike, E. (1911). *Animal intelligence : Experimental studies*. https://www.worldcat.org/title/animalintelligence-experimental-studies/oclc/844297

Todd, T. P., Vurbic, D., & Bouton, M. E. (2014). Behavioral and neurobiological mechanisms of extinction in Pavlovian and instrumental learning. *Neurobiology of Learning and Memory*, *108*, 52–64. https://doi.org/10.1016/J.NLM.2013.08.012