In *Journal of Experimental Psychology: Animal Learning & Cognition*
2014, vol. 40 (3), pp 335-354
http://www.apa.org/pubs/journals/xan/index.aspx

DOI: 10.1037/ xan0000023

Conditioned inhibition and reinforcement rate

Justin A. Harris, Dorothy W. S. Kwok, & Benjamin J. Andrew

School of Psychology, University of Sydney, Australia

Abstract

We investigated conditioned inhibition in a magazine approach paradigm. Rats were trained on a feature negative discrimination between an auditory conditioned stimulus (CS) reinforced at one rate versus a compound of that CS and a visual stimulus (L) reinforced at a lower rate. This training established L as a conditioned inhibitor. We then tested the inhibitory strength of L by presenting it in compound with other auditory CSs. L reduced responding when tested with a CS that had been reinforced at a high rate, but had less or even no inhibitory effect when tested with a CS that had been reinforced at a low rate. The inhibitory strength of L was greater if it signaled a decrease in reinforcement from an already low rate than if it signaled an equivalent decrease in reinforcement from a high rate. We conclude that the strength of inhibition is not a linear function of the change in reinforcement that it signals. We discuss the implications of this finding for models of learning (e.g. Rescorla & Wagner, 1972) that identify inhibition with a difference (subtraction) rule.

Address correspondence to:

Justin Harris
School of Psychology
University of Sydney
Sydney 2006
Australia
Tel: +61 2 9351 2864
Email: justin.harris@sydney.edu.au

Key words: Rat; delay conditioning; magazine approach; variable interval; conditioned inhibition.

Author note: The work was supported by grants DP1092695 and FT100100091 from the Australian Research Council.

Pavlov (1927) described the manner in which a stimulus could come to exert inhibitory control over conditioned responding if it was repeatedly presented with an excitatory conditioned stimulus (CS+) and if the unconditioned stimulus (US) was omitted on these trials. The ability of the inhibitory stimulus (CS-) to reduce responding was not confined to the response elicited by the original CS+, but could readily transfer to another CS+ that had been conditioned with the same or similar US. Conditioned inhibition has become a touchstone phenomenon in conditioning research, and is of central importance for theoretical accounts of learning. The reduction in responding is usually understood as resulting from a decrease in net conditioning strength when the inhibitory value of a CS- is summed with the excitatory value of a CS+ (Rescorla, 1969a). Indeed, it is commonplace to treat inhibition in simple arithmetic terms when explaining not only the effect of a CS- on responding but also the way that inhibition is acquired (Rescorla & Wagner, 1972).

One of the most significant theoretical contributions of the Rescorla-Wagner model was in providing a mechanism by which a stimulus would acquire net inhibitory value when consistently associated with a negative discrepancy between the expectancy of a US (based on the aggregate associative strength of all CSs present) and the asymptotic conditioning strength (λ) that can be sustained by the absent US on that trial (Rescorla & Wagner, 1972; Wagner & Rescorla, 1972). Thus, if a stimulus, X, is repeatedly presented in the presence of a CS+ with associative strength V_i, but on these trials the US is omitted (and thus $\lambda = 0$), the amount of inhibition that X ultimately acquires will equal $-V_i$. This difference rule specifies the net effect of X on responding to any CS+ because X will reduce by $-V_i$ the associative strength of any compound that contains X. These simple arithmetic principles are clearly identified with the Rescorla-Wagner model, and the many other models of conditioning that incorporate its learning rule. However, similar principles have also been incorporated in other quite different models of conditioning. In their Rate Estimation Theory, Gallistel and Gibbon (2000) identify the content of learning about a CS- with the difference between the reinforcement rate signaled by the CS+ and the reinforcement rate observed when that CS+ is presented with the CS-. The Comparator Hypothesis (Stout & Miller, 2007), which eschews notions of inhibitory associative strength, nevertheless identifies the effect of a CS- with a subtractive influence on responding that is based on a second-order link between the CS- and its original CS+. In effect, the excitatory strength of the original CS+ is subtracted from the excitatory strength of any new compound containing the CS-.

As just described, contemporary accounts of inhibition provide a description of both the circumstances that lead to inhibition and the amount of inhibition that is acquired. However, rigorous empirical test of the quantitative aspects of these accounts is out of reach without some detailed knowledge of how variations in learning map onto levels of responding. Recent experiments from our laboratory have gone some way to uncovering this relationship, and have revealed a very orderly relationship between the rats' response rates to a CS and the rate that the CS was reinforced (Harris & Carpenter, 2011). Andrew and Harris (2011), exploited this relationship when investigating the summation of responding when two separately conditioned stimuli are presented simultaneously as a compound. Their experiments consistently showed that rats responded to a compound of two CSs at

exactly the same rate as they responded to a third CS that had been reinforced at a rate equal to the sum of reinforcement rates of the two CSs in the compound. That is, the rats treated the compound as having a reinforcement rate equal to the sum of the rates at which its component CSs had been reinforced. A similar principle was at work in determining what rats learned about each of two CSs that were conditioned together as a compound: what the rats learned about one CS was equal to the reinforcement rate of the compound minus the reinforcement rate of the other CS (Harris, Andrew, & Livesey, 2012).

The experiments just described demonstrate that rats, in effect, learn about reinforcement rates, inasmuch as they combine reinforcement rates according to a simple linear summation rule when learning about or responding to CSs presented in compound. The present series of experiments sought to investigate whether these principles apply to the acquisition and expression of conditioned inhibition. As already stated, theoretical accounts of inhibition describe it as a simple subtractive computation: the CS- acquires negative associative strength (Rescorla & Wagner, 1972), is attributed with a negative reinforcement rate (Gallistel & Gibbon, 2000), or is compared negatively with an excitatory CS+ (Stout & Miller, 2007). In each case, the effect of the CS- is equal in magnitude to the positive strength of the CS+. In the present experiments, rats were trained with a number of auditory CS+s, each signaling food with a different rate of reinforcement. Concurrent with this training, a light was established as a CS- in a feature negative design by being presented with one of the auditory CS+s and this compound was reinforced at a lower rate than was the auditory CS+ on its own. The theoretical accounts described above predict that the strength of inhibition acquired by the CS- should equal the difference in the rate at which the auditory CS+ was reinforced alone and the rate of reinforcement of the compound of the CS+ and CS-. Therefore we should be able to predict precisely the effect of the CS- on responding to each of the other auditory CS+s. This aim meant that, in all experiments, we used summation tests of inhibitory strength rather than a retardation test (Rescorla, 1969b) given that the latter is not amenable to such quantitatively precise predictions.

Experiments 1a, 1b and 1c

The first three experiments trained rats with three auditory CS+s, A, B, and C, in a delay conditioning procedure. The reinforcement rate of each CS+ differed as a function of the mean CS duration (the mean CS-US interval, see Harris, Gharaei, & Pincham, 2011): A was reinforced after an average interval of 20 s (this rate is equivalent to, on average, three pellets per minute), B after an average of 30 s (two pellets per minute), and C after an average of 60 s (one pellet per minute). These values were chosen because they represent equal sized steps in mean reinforcement rate: C is 1/60-s above zero, B is 1/60-s above C, and A is 1/60-s above B. In each experiment a light, L, was established as a CS- with a nominal inhibitory strength of -1/60. In Experiment 1a L was combined with A and the compound was reinforced after a mean interval of 30 s; in Experiment 1b L was combined with B and the compound was reinforced after a mean interval of 60 s; in Experiment 1c L was combined with C and the compound was never reinforced. Thus in each experiment, L was associated with a decrease in reinforcement rate of 1/60-s compared to the reinforcement rate of the CS+ against which it was conditioned. After extended training, the inhibitory strength of L was assessed by occasional probe trials in which L was presented in compound with each of the

other two auditory CS+s. If L consistently acquired an inhibitory strength of -1/60-s, it should have a predictable effect on responding to the auditory CS+s. In Experiment 1a, the rats should respond to the compound of L and B at the same rate as they respond to C because both should have a net reinforcement rate of 1/60-s; their responses to the compound of L and C should not exceed baseline (pre-CS) levels because the compound should signal a reinforcement rate of zero. In Experiment 1b, the rats should respond to the compound of L and A at the same rate as they respond to B; their response rates to the compound of L and C should equal baseline response rates in the pre-CS period. In Experiment 1c, the rats should respond to the compound of L and A at the same rate as they respond to B; their response rates to the compound of L and B should match response rates to C. The design of the three experiments, and the predicted results, are summarized in Table 1.

Table 1. Summary of designs of Experiments 1a, 1b and 1c.

Experiment	Training	Probe trials (*+ predicted result*)
1a	A[20s+] B[30s+] C[60s+] **LA**[30s+]	LB (= *C*) and LC (= *pre-CS*)
1b	A[20s+] B[30s+] C[60s+] **LB**[60s+]	LA (= *B*) and LC (= *pre-CS*)
1c	A[20s+] B[30s+] C[60s+] **LC**[60s–]	LA (= *B*) and LB (= *C*)

Note: A, B, and C were auditory CSs; L was a light. CS and compound durations were on average 20, 30, or 60 s long, and ended with food reinforcement (+) or no reinforcement (–).

Methods

Subjects

In each experiment, there were 24 experimentally naive male Hooded Wistar rats (Rattus norvegicus; 8 to 10 weeks of age at the start of the experiment). They were obtained from the Laboratory Animal Services breeding unit at The University of Adelaide, South Australia. During the experiment, they were housed in groups of 8 in large white plastic tubs, measuring 26 x 59 x 37cm (height x length x depth), located in the animal colony maintained by the School of Psychology at the University of Sydney. They had unrestricted access to water in the home tubs. Three days prior to commencement of the experiment, they were placed on a restricted food schedule. Each day, half an hour after the end of the daily training session, each tub of rats received a ration of their regular dry chow (3.4 kcal/g) equal to 5% of the total weight of all rats in the tub. This amount is approximately equal to their required daily energy intake (Rogers, 1979), and took at least 2 h to be eaten (but was usually finished within 3 h). This meant that all rats in the tub had access to food for an extended period, which should reduce differences between rats in their levels of hunger.

Apparatus

Rats were trained and tested in 16 Med Associates conditioning chambers measuring 28.5 x 30 x 25 cm (height x length x depth). The end walls of each chamber were made of aluminum; the sidewalls and ceiling were Plexiglas™. The floor of the chamber consisted of stainless-steel rods, 0.5 cm in diameter, spaced 1.5 cm apart. Each chamber had a recessed food magazine in the center of one end wall, with an infra-red LED and sensor

located just inside the magazine to record entries by the rat. A small metal cup measuring 3.5 cm in diameter and 0.5 cm deep was fixed on the floor of each food magazine. Attached to the food magazine was a dispenser delivering 45 mg food pellets (Noyes Formula P; Research Diets Inc., New Brunswick, NJ). Each chamber was enclosed in a sound- and light-resistant wooden shell. Throughout all sessions, fans located in the rear wall provided ventilation; the operation of these created a background noise level measuring 70dB. Experimental events were controlled and recorded automatically by computers and relays located in the same room. White noise (78dB) was presented from a speaker mounted on the wall of each conditioning chamber above and to the left of the food magazine. A tone (2.9 kHz) was produced from a piezo buzzer positioned on the floor of the sound-attenuating shell behind each conditioning chamber. A clicker was delivered from a module (Med Associates, product ENV 135M) located on the wall of the conditioning chamber above and to the right of the magazine. A steady light (30cd/m^2) was produced by an incandescent bulb mounted high on the back wall of the sound-attenuating shell. The allocation of the three auditory stimuli to different CS+ types was counterbalanced evenly across squads of four rats using a Latin square.

Procedure

Prior to the start of conditioning, rats received a single 20-min magazine training session during which 20 food pellets were presented on a VT 1-min schedule, with no stimulus presentations. Rats that ate fewer than half of the pellets were given a second session of magazine training the following day. After all rats completed magazine training, they commenced daily conditioning sessions 5 days per week for a total of 35 days (Experiment 1a) or 32 days (Experiments 1b and 1c). Each session contained 48 trials, 12 trials of each of the three auditory CSs and 12 trials of the compound containing L. CS durations varied randomly from trial to trial, sampled from a uniform distribution of durations, so as to produce a stable response rate across the length of the CS (Harris et al., 2011). The durations of the auditory CSs were between 2 and 38 s (mean = 20 s) for A, between 2 and 58 s (mean = 30 s) for B, and between 2 and 118 s (mean = 60 s) for C. The durations of the compounds were between 2 and 58 s (mean = 30 s) for LA in Experiment 1a, and between 2 and 118 s (mean = 60 s) for LB and LC in Experiments 1b and 1c. All trials with an auditory CS alone included delivery of a single food pellet at the termination of the CS. Compound presentations of LA and LB (Experiments 1a and 1b) also terminated with a food pellet, but LC (Experiment 1c) was never followed by food. Trials of each type were randomly intermixed with the constraint that each quarter of the session included equal numbers of each trial type. Photo-beam interruptions by head entry into the magazine were recorded during each CS and each 20-s pre-CS period. Sessions lasted approximately 2 h.

The final 3 days (Experiment 1a) or 4 days (Experiments 1b and 1c) included “probe” trials in which L was presented with each of the other two auditory CS+s (those with which it had not been presented in compound during training). Thus LB and LC were presented as probe trials in Experiment 1a, and LA and LC in Experiment 1b, and LA and LB in Experiment 1c. These probe trials had a fixed duration equal to the mean duration of the auditory CS+ (i.e., 20 s for LA, 30 s for LB, and 60 s for LC), and were never reinforced with food. Each of the probe compounds was presented once during each session, at Trial 19 or Trial 30, and the order of the two probes was counterbalanced across squads of

rats, and was reversed across consecutive days.

The data across the final test sessions were analyzed as follows. Response rates on non-probe trials were first analyzed by ANOVA, with Greenhouse-Geisser corrections to the degrees of freedom whenever the data failed the test of sphericity, to establish whether the rats were indeed discriminating between the different stimuli as intended. The key analyses, however, involved the probe trials. To test the inhibitory effect of L, we compared responding during each probe compound with responding to the auditory CS+ that formed part of that compound (e.g., LB was compared with B). These comparisons were made using paired *t*-tests, and to control the type-1 error rate we divided alpha by 2. Bayesian analyses were performed to test the predictions that responding to each compound would equal responding to the CS+ that was reinforced at a rate matching the net rate signaled by the compound (e.g., that LB would equal C), as shown in Table 1. We used the tool described by Rouder, Speckman, Sun, Morey and Iverson (2009) to calculate a Bayes Factor based on the "JZS" prior, available on the website http://pcl.missouri.edu/bayesfactor. This allowed us to estimate the evidence in favor of the null hypothesis (that responding to the compound and CS+ were equal) versus the alternative hypothesis (that they were different).

Results

For all three experiments, the mean response rates for each trial type on each session are shown in the plots in the top half of Figure 1. The mean response rates, averaged over the final sessions that included probe compound trials, are shown in the bar graphs in the bottom of Figure 1.

Experiment 1a. Across the first 20 days, response rates during each auditory CS+ increased and then remained stable for the remaining sessions. An ANOVA, conducted on the data from the last 3 days (but excluding probe trials), confirmed that the rats did discriminate effectively between the different trial types on which they had been trained, $F(2, 47) = 24.37$, $p < .001$, $\eta^2_p = 0.51$ (95% confidence interval on η^2_p: 0.37 and 0.67, Hentschke & Stüttgen, 2011). The key analyses involved comparisons between the probe compounds and the auditory CS+s, averaged across the final three sessions. Response rates during LB did not differ significantly from those during B, $t(23) = 1.92$, $p = .067$, Cohen's d = 0.39 (95% confidence interval: -0.03 and 0.80), and response rates during LC were not different from those during C, $t(23) < 1$, Cohen's d = -0.17 (-0.58 and 0.23). The Bayes Factor for the comparison between LB and C favored the alternative hypothesis, that LB and C were different, by 9 to 1. The Bayes Factor for the comparison between LC and pre-CS response rates also favored the alternative hypothesis, that LC and pre-CS were different, by more than 12000 to 1.

Experiment 1b. Across the first 20 days, response rates during each auditory CS+ increased and then remained stable for the remaining sessions. An ANOVA, with Greenhouse-Geisser correction to the degrees of freedom, on the data from the last 4 days (but excluding probe trials) confirmed that the rats did discriminate effectively between the different trial types on which they had been trained, $F(1.9, 44) = 22.16$, $p < .001$, $\eta^2_p = 0.49$ (95% confidence interval: 0.33 and 0.65). Averaged across the final 4 sessions, the difference in response rates during probe trials with LA versus A fell short of the adjusted level of significance, $t(23) = 2.28$, $p = .037$, Cohen's d = 0.45 (0.03 and 0.87). Response rates during LC were significantly less than those during C, $t(23) = 3.21$, $p = .006$,

Cohen's d = 0.62 (0.18 and 1.05). The Bayes Factor for the comparison between LA and B favored the null hypothesis (that LA = B) by 4 to 1. However, for the comparison between LC and pre-CS response rates, the Bayes Factor favored the hypothesis that these rates were different by more than 12000 to 1.

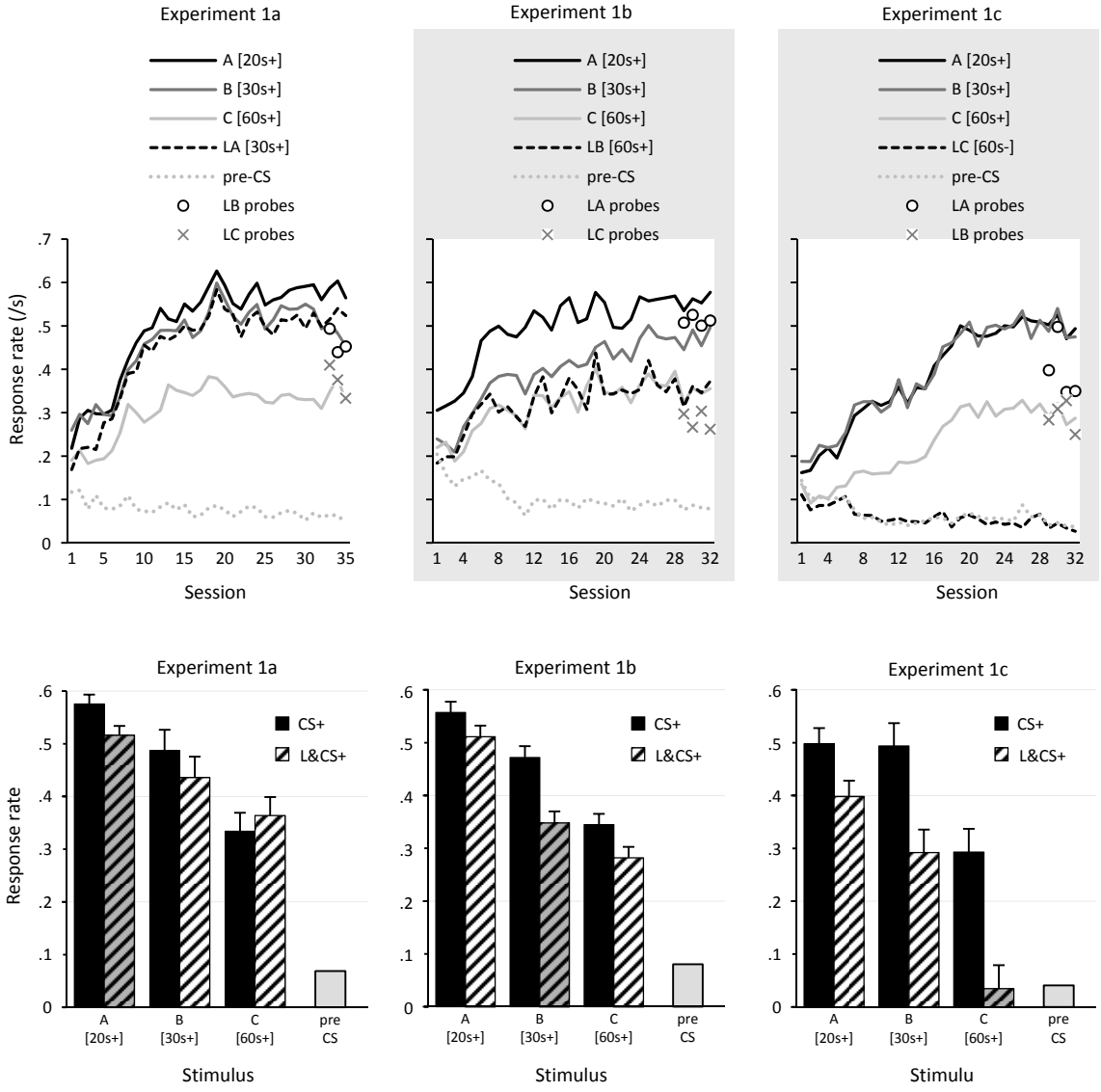

Figure 1. Results of Experiments 1a (left), 1b (center), and 1c (right). The line plots in the top half of the figure show average daily response rates to each of three auditory CS+s, A, B, and C, that were reinforced (+) on average after 20 s, 30 s, and 60 s, respectively. The plots also show responding to the compound of a light (L) and A that was reinforced after 30 s (Experiment 1a), the compound of L and B that was reinforced after 60 s (Experiment 1b), or the compound of L and C that was never reinforced (-; Experiment 1c). The last three or four sessions included non-reinforced probe trials in which L was presented in compound with each of the other two CS+s: as LB and LC in Experiment 1a; LA and LC in Experiment 1b; LA and LB in Experiment 1c. The bar graphs in the lower half of the figure show the response rates averaged over these final sessions. Black bars show responses to each auditory CS+, striped bars show response rates to compounds: the compound used in feature negative training (black and gray) and the probe compounds (black and white). Vertical bars show the standard error of the mean difference between response rates to each CS+ and L&CS+ compound.

Experiment 1c. Across the first 20 days, the rats' response rates during each auditory CS+ increased and then remained stable for the remaining 15 days. An ANOVA, with Greenhouse-Geisser correction, on the data from the last 4 days (excluding probe trials) confirmed that the rats did discriminate effectively between the different trial types on which they had been trained, $F(1.7, 40) = 43.04$, $p < .001$, $\eta^2_p = 0.65$ (95% confidence interval: 0.53 and 0.78). Averaged across the final 4 sessions, response rates on probe trials with LA were significantly below those to A, $t(23) = 3.35$, $p = .003$, Cohen's d = 0.68 (0.23 and 1.12). Response rates during LB were significantly less than those during B, $t(23) = 4.66$, $p < .001$, Cohen's d = 0.95 (0.46 and 1.43). The Bayes Factor for the comparison between LA and B gives even odds for the null (LA = B) and alternative (LA ≠ B) hypotheses, but favored the null hypothesis by 6 to 1 for the comparison between LB and C.

The effectiveness of L differed markedly between the three experiments. The difference was most pronounced when comparing Experiment 1a, in which L failed to significantly reduce responding in either test compound, with Experiment 1c, in which L significantly reduced responding in both test compounds. To test this difference directly, we performed an across-experiment statistical analysis comparing the effect of L on responding to CS B (this CS+ was chosen because both experiments included probe trials with the compound LB). We ran a 2x(2) ANOVA with Experiment as the between-subject factor, and LB versus B as the within-subject factor. There was no main effect difference between the two experiments, $F < 1$, $\eta^2_p = 0.02$ (95% confidence interval: 0.00 and 0.16), but there was a main effect for the comparison between LB and B, $F(1, 46) = 24.74$, $p < .001$, $\eta^2_p = 0.35$ (0.23 and 0.50). Of most importance, there was a significant interaction, $F(1, 46) = 8.78$, $p = .005$, $\eta^2_p = 0.16$ (0.03 and 0.35), confirming that L was more effective at reducing responding to B in Experiment 1c than in Experiment 1a.

Discussion

The experiments described above investigated the inhibitory effect of a light, L, that signaled a fixed decrease in reinforcement rate (-1/60-s). In Experiment 1c, L was trained with C, and signaled a reduction in reinforcement rate from 1/60-s to zero. As a result of this training, L was effective at reducing responding to A and B that signaled higher reinforcement rates (1/20-s and 1/30-s, respectively). Moreover, there was some evidence to confirm the specific prediction that the rats learned that L had a reinforcement rate of -1/60-s. The clearest evidence came from the tests with L and B: the rats responded to this compound at almost the identical rate that they responded to C, consistent with the prediction that L reduced the expected reinforcement rate of B by -1/60-s. A Bayesian analysis showed that the likelihood that LB and C were equal was six times greater than the likelihood that they were different. The response rate to LA was also where one might expect if the net reinforcement rate of the compound was calculated to be 1/30-s (= 1/20 – 1/60). However, this prediction could not be confirmed in the present data because the rats' response rates did not differ between A and B.

In Experiment 1b, L was effective as an inhibitor when tested against A and C. However, response rates to LC remained well above the pre-CS rate. Thus the impact of L in this compound fell short of the prediction that L would reduce the signaled rate of reinforcement to zero. Finally, in Experiment 1a, L was completely ineffective as an inhibitor. When L was tested in compound

with either B or C, it failed to produce a significant reduction in responding. Indeed, responding to LC was slightly higher than responding to C alone.

Together these experiments show that the effectiveness of a CS- is not simply determined by the size of decrease in reinforcement rate that it signals. Rather, a CS- that signals a decrease in reinforcement rate at the low end of the reinforcement spectrum is effective at inhibiting responses to a CS+ that signals a higher rate of reinforcement, whereas a CS- that signals the same size of decrement in reinforcement rate but at the higher end of the reinforcement spectrum is less effective at reducing responding to a CS+ that signals a low rate of reinforcement. This conclusion was subjected to further analysis that compared the effect of L on responding to B between Experiments 1a and 1c. While one needs to remain cautious in placing undue weight on such post-doc analyses across experiments, the analysis was consistent with the conclusion that the inhibitory strength of the CS- was weaker when it was trained against a strong CS+ than when trained against a weak CS+.

Experiments 2a and 2b

Experiments 1a, 1b and 1c showed that the effectiveness of a CS-, L, is a function of the reinforcement rates of the CS+ against which it is trained and the CS+ against which it is tested. In those experiments, L signaled a decrease in reinforcement rate of 1/60-s. When this decrease was established against a CS+ with a low reinforcement rate (e.g., 1/60-s), L was effective in reducing responding to other CS+s with higher reinforcement rates (1/20-s and 1/30-s). However, when L was established against a CS with high reinforcement (e.g., 1/20-s), L was less effective at reducing responding to other CSs with lower reinforcement rates (e.g., 1/60-s). The asymmetry in effectiveness of L implies that the acquisition and expression of inhibition is not a simple linear function of change in reinforcement rate. One potential source of this asymmetry stems from the fact that reinforcement rate in the previous three experiments was manipulated by varying the mean CS-US interval. Even though L signaled the same magnitude of change in rate of reinforcement in each experiment, the time window across which L was presented during training and tests differed. For example, in Experiment 1c, inhibitory training with L involved presentations of LC that had a mean duration of 60 s, whereas probe test trials with LA and LB were only 20 and 30-s long. Conversely, in Experiment 1a, inhibitory training with L involved presentations of LA that had a mean duration of 30 s, and probe trials with LB and LC were 30-s and 60-s long. Therefore, it is possible that L was more effective as an inhibitor in Experiment 1c because the time window across which it had been trained was longer than the time window across which it was tested.

To control for the possibility just described, Experiments 2a and 2b used the same range of CS-US intervals for all trials, and varied reinforcement rate by varying the proportion of trials that ended with food. In both experiments, three auditory CS+s were reinforced on 100% (A), 66% (B), or 33% (C) of trials. A light, L, was established as an inhibitor by pairing it with one of these CSs and reducing the reinforcement rate by 33%. The most informative design is that in which L signals a reduction from 66% to 33%, because this allows us to assess the inhibitory effect of L against a CS+ with a higher reinforcement rate (100%) and against a CS+ with a lower reinforcement rate (33%). Replication of our previous findings would be obtained if L was more effective at reducing responding to the 100% CS+ than to the 33% CS+. This design

was used in Experiment 2a. Experiment 2b used the design in which L was paired in training with C (33%) and reduced the reinforcement rate to zero, before L was tested against A and B. This design was used to test our initial hypothesis that the impact of an effective CS- on responding could be accurately predicted by the amount it reduced the reinforcement rate of the trained CS+. In other words, for Experiment 2b, we wished to test whether L would reduce the net reinforcement rate of each test CS+ by 33%, and therefore the rats should respond to LA at the same level as they respond to B, and they should respond to LB at the same level as they respond to C. The design of the experiments are summarised in Table 2.

Table 2. Summary of designs of Experiments 2a and 2b.

Experiment	Training	Probe trials
2a	A[100%] B[66%] C[33%] **LB[33%]**	LA and LC
2b	A[100%] B[66%] C[33%] **LC[0%]**	LA and LB

Note: A, B, and C were auditory CSs reinforced on 100%, 66%, and 33% of trials. L was a light that reduced the rate of reinforcement from 66% to 33%, or from 33% to 0%.

Methods

Subjects and apparatus

Each experiment used 16 rats of the same strain and source, and housed in the same manner, as described for Experiments 1a, b, and c. The same apparatus as described above was used in these experiments.

Procedure

Rats first received magazine training as described previously, before commencing daily conditioning sessions 5 days per week for a total of 35 days. Each session contained 48 trials, 12 trials of each of the three auditory CS+s and 12 trials of the compound containing L. All trial types had a mean duration of 30 s. At the end of each CS+ presentation, a food pellet was delivered on 100% of trials with A, 66% of trials with B, and 33% of trials with C. In Experiment 2a, food was delivered on 33% of trials with the compound LB; in Experiment 2b, food was delivered on 0% of trials with LC. Trials of each type were randomly intermixed with the constraint that each quarter of the session included equal numbers of each trial type. Photo-beam interruptions by head entry into the magazine were recorded during each CS and each 20-s pre-CS period. Sessions lasted approximately 2 h.

The final 5 days included "probe" trials in which L was presented with each of the other two auditory CSs (those with which it had not been presented in compound during training). Thus LA and LC were presented as probe trials in Experiment 2a, and LA and LB in Experiment 2b. These probe trials had a fixed duration of 30 s, and were never reinforced with food. Each of the probe compounds was presented once during each session, at Trial 19 or Trial 30, and the order of the two probes was counterbalanced across squads of rats, and was reversed across consecutive days.

Results

For both experiments, the mean response rates for each trial type on each session are shown in the plots in the top half of Figure 2. The mean response rates, averaged over the final sessions that included probe compound trials, are shown in the bar graphs in the bottom of Figure 2.

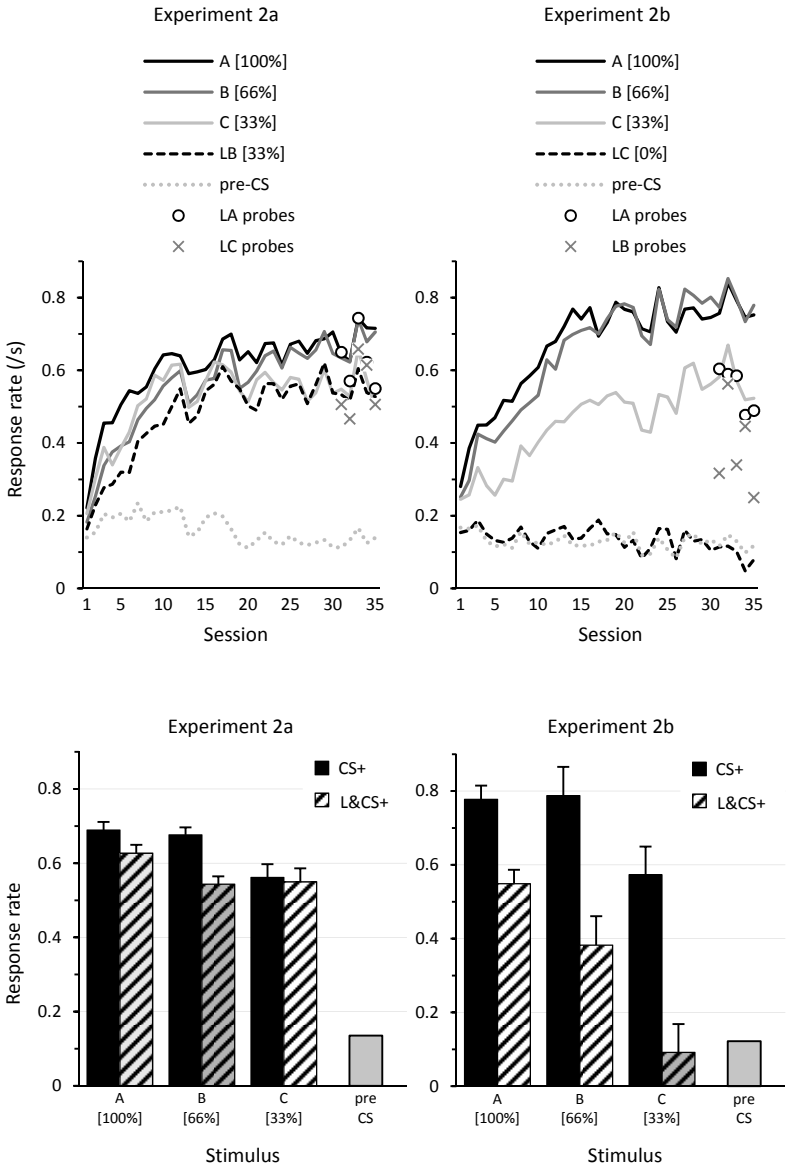

Figure 2. Results of Experiments 2a (left) and 2b (right). The line plots in the top part of the figure show average daily response rates to each of three auditory CSs, A, B, and C, that were reinforced on 100%, 66%, and 33% of trials, as well as to the compound LB (Experiment 2a) reinforced on 33% of trials, or the compound LC (Experiment 2b) reinforced on 0% of trials. The last five sessions included probe trials in which L was presented with A and C (Experiment 2a) or with A and B (Experiment 2b). The bar graphs at the lower part of the figure show the response rates averaged over the final five sessions. Black bars show responses to each auditory CS+, striped bars show response rates to compounds: the compound used in feature negative training (black and gray) and the probe compounds (black and white). Vertical bars show the standard error of the mean difference between response rates to each CS+ and L&CS+ compound.

***Experiment 2a*.** Responding increased steadily over the first 20 days and remained stable thereafter. Response rates, averaged over the last 5 sessions, are shown in the bar graph on the bottom left of Figure 2. The rats' response rates did not differ between A and B. However the response rates to these CSs were higher than to either C or the compound LB, both of which were reinforced at 33%. An ANOVA, with Greenhouse-Geisser correction, confirmed that there were significant differences among the CSs and compound, $F(1.9, 28) = 6.33$, $p = .006$, $\eta^2_p = 0.30$ (95% confidence interval: 0.17 and 0.58). The key analyses compared responding to the two probe compounds, LA and LC with their respective CSs, A and C. Paired *t*-tests, with appropriate correction to alpha, confirmed that response rates to LA were significantly less than to A, $t(15) = 2.82$, $p = .014$, Cohen's d = 0.70 (0.14 and 1.24), whereas response rates to LC were not different from C, $t < 1$, Cohen's d = 0.08 (-0.41 and 0.57). Indeed, the mean response rate to C and LC were virtually identical (0.56 and 0.55, respectively). The Bayes Factor, calculated on the *t*-statistic (Rouder et al., 2009), indicated that these data provide five-times stronger support for the hypothesis that C and LC were equal than for the hypothesis that they were different.

***Experiment 2b*.** Responding increased over the first 20 days and remained stable for the next 15 days. Response rates, averaged over the last 5 sessions, are shown in the bar graph on the bottom right of Figure 2. Once again, the rats' response rates did not differ between A and B. Nonetheless, the response rates to these CSs were higher than to C and to the compound LC, reinforced at 0%. An ANOVA, with Greenhouse-Geisser correction, confirmed that there were significant differences among the CSs and compound, $F(2.7, 40) = 44.39$, $p < .001$, $\eta^2_p = 0.75$ (95% confidence interval: 0.68 and 0.83). This was followed by two pairwise comparisons to establish whether L was effective at reducing responding in each test compound. Thus, paired *t*-tests with appropriate correction to alpha compared LA with A, and compared LB with B. These confirmed that response rates to LA were significantly less than to A, $t(15) = 6.14$, $p < .001$, Cohen's d = 1.53 (0.79 and 2.25), and that response rates to LB were significantly less than to B, $t(15) = 5.18$, $p < .001$, Cohen's d = 1.30 (0.61 and 1.96). The Bayes Factors showed that, according to these data, LA and B were 49 times more likely to be different than to be equal, but that the odds were approximately even for the comparison between LB and C (the null hypothesis was 1.66 times more likely than the alternative).

The final analysis compared the strength of L's inhibitory effect on CS A across the two experiments. A 2x(2) ANOVA compared A and LA as the within-subject factor, and Experiment 2a versus Experiment 2b as the between-subject factor. There was no main effect difference between the two experiments, $F < 1$, $\eta^2_p < 0.01$ (95% confidence interval: 0.00 and 0.16), but there was a main effect for the comparison between LA and A, $F(1, 30) = 44.91$, $p < .001$, $\eta^2_p = 0.60$ (0.45 and 0.75). Of most importance, there was a significant interaction, $F(1, 30) = 14.90$, $p = .001$, $\eta^2_p = 0.33$ (0.13 and 0.56), confirming that L was more effective at reducing responding to A in Experiment 2b than in Experiment 2a.

Discussion

These experiments investigated the inhibitory effect of a light, L, that signaled a decrease in reinforcement rate. In Experiment 2a, L signaled a change from 66% to 33% reinforcement. On transfer tests with new CS+s, L was effective at reducing responding to the CS+ (A) that had been reinforced on 100% of trials, but was completely ineffective

when tested against a CS+ (C) that had been reinforced on 33% of trials. In Experiment 2b, L signaled a change in reinforcement rate from 33% to 0%. Here, L was effective at reducing responding on transfer tests against the 100% CS+ (A) and against a CS+ (B) reinforced on 66% of trials. Even though L was effective at reducing responding to A in both experiments, an across-experiment ANOVA confirmed the observation that the effect of L was greater in Experiment 2b than in Experiment 2a. These findings are consistent with the results of Experiments 1a, b and c. They show that a stimulus acquires inhibitory strength if it signals a decrease in reinforcement rate when presented in compound with a CS+. However, they also show that the amount of inhibition acquired by the CS-, and its effectiveness at reducing responding to a CS+, are determined by the absolute reinforcement rate of the training and test CS+.

Experiment 2b also failed to provide evidence to confirm the prediction that on probe tests L would reduce the expected reinforcement rate of the compound by -33%. In that experiment, there was much stronger evidence that LA and B were different than that they were the same. This contrasts with the prediction that the expected reinforcement rate of LA should have been equal to that of B, based on the assumption that L reduced the expected reinforcement rate by 33%.

Experiment 3

Across-experiment comparisons among the previous experiments have indicated that, when a CS- signals a fixed decrease in reinforcement rate (-1/60-s or -33%), the amount of inhibition it acquires is greater when the absolute reinforcement rate is low than when it is high. Experiment 3 was designed to confirm this observation within a single experiment. It also sought to investigate these effects when the CS- signaled a larger decrease in reinforcement rate. Two groups of rats were trained on a feature negative discrimination between a CS+, A, and a compound LA that was reinforced at a rate 50% lower than A. For one group ("100-50"), A was reinforced on 100% of trials and LA was reinforced on 50% of trials; for the other group ("50-0"), A was reinforced on 50% of trials and LA was reinforced on 0% of trials. Concurrent with this discrimination, the rats were trained with two other CS+s, B and C, used for transfer tests to assess the inhibitory strength of L. B was reinforced on 100% of trials, and C was reinforced on 50% of trials. After 28 sessions of training, the last 5 sessions included probe trials in which the compounds LB and LC were presented. If L acquired the same inhibitory strength in both groups, based on the fact that it signaled a reduction in reinforcement rate of 50% in each case, then L should reduce responding to B and C by the same amount for both groups. However, if, as suggested by the evidence so far, L acquires greater inhibitory strength when signaling a reduction from 50% to zero than when signaling a reduction from 100% to 50%, then L should produce a bigger decrease in responding to B and C in the 50-0 group than in the 100-50 group. The design of the experiment is summarized in Table 3.

Methods

Subjects and apparatus

The experiment used 32 rats of the same strain and source, and housed in the same manner, as described for Experiments 1a, b, and c. The same apparatus as described above was used in these experiments.

Table 3. Summary of design of Experiment 3.

Group	Training	Probe trials
100-50	A[100%] B[100%] C[50%] **LA[50%]**	LB and LC
50-0	A[50%] B[100%] C[50%] **LA[0%]**	LB and LC

Note: A, B, and C were auditory CSs reinforced on 100% or 50% of trials. L was a light that, when paired with A, reduced the reinforcement rate from 100% to 50% in Group 100-50, and reduced the reinforcement rate from 50% to 0% in Group 50-0.

Procedure

Rats first received magazine training as described previously, before they were split into two groups (n=16) and commenced daily conditioning sessions 5 days per week for a total of 33 days. Each session contained 48 trials, 12 trials of each of the three auditory CS+s, A, B, and C, and 12 trials of the compound LA. All trial types had a mean duration of 20 s. At the end of each CS+ presentation, a food pellet was delivered on 100% of trials with B, and 50% of trials with C. Presentations of A were reinforced on 100% of trials in Group 100-50, but were reinforced on only 50% of trials in Group 50-0. Presentations of LA were reinforced on 50% of trials for Group 100-50, but were never reinforced for Group 50-0. Trials of each type were randomly intermixed with the constraint that each quarter of the session included equal numbers of each trial type. Photo-beam interruptions by head entry into the magazine were recorded during each CS and each 20-s pre-CS period. Sessions lasted approximately 2 h. The final 5 days included "probe" trials in which L was presented in compound with B or C. These probe trials had a fixed duration of 30 s, and were never reinforced with food. LB and LC probes were presented once during each session, at Trial 19 or Trial 30, and the order of the two probes was counterbalanced across squads of rats, and was reversed across consecutive days.

Results

A defect in the infra-red detector measuring nose-pokes in one conditioning chamber meant that, for rat 9, data from Days 29, 32, and 33 had to be excluded from all analyses. For both experiments, the mean response rates, averaged over the final sessions that included probe compound trials, are shown in the bar graphs in the bottom of Figure 3.

As previously, responding increased steadily over the first 20 days and remained stable thereafter. In order to save space, these pre-test data are not shown for this and the remaining experiments. Response rates to each trial type, averaged over the last five sessions, are shown in Figure 3. An ANOVA, with Greenhouse-Geisser correction, was run on the response rates to the three CS+s and the training compound (LA). This analysis, which excluded the probe trials, confirmed that there were significant differences among the training trial types, $F(2.6, 79) = 56.04$, $p < .001$, $\eta^2_p = 0.65$ (95% confidence interval: 0.60 and 0.73), and there was a significant interaction between the within-group factor of trial-type and the between group factor, $F(2.6, 79) = 37.01$, $p < .001$, $\eta^2_p = 0.55$ (0.46 and 0.68). This interaction reflects the fact that CS A and LA were reinforced at different rates between the two groups. Despite this, there was no overall difference in response rate between the two groups, $F < 1$, $\eta^2_p < 0.01$ (0.00 and 0.17).

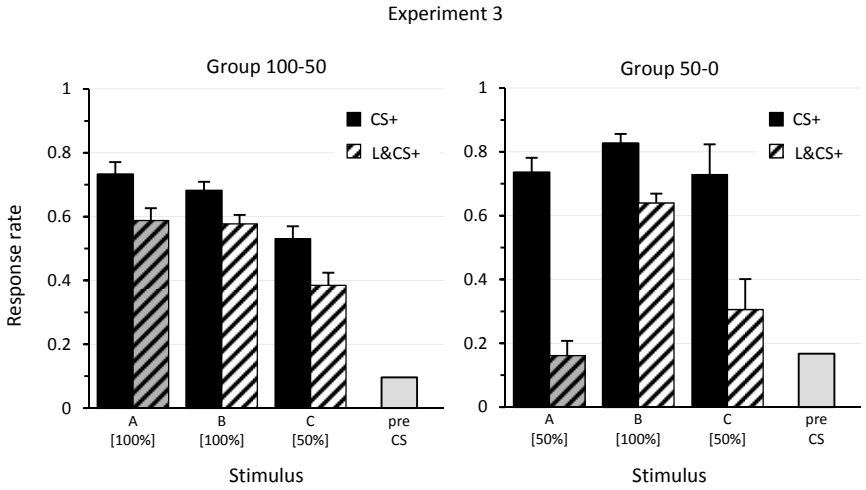

Figure 3. Average response rates over the last five sessions of Experiment 3. Two groups of rats were trained for 33 days with three auditory CS+s, A, B, and C, and a compound of a light, L, with A. For Group 100-50, A was reinforced on 100% of trials, and LA was reinforced on 50% of trials. For Group 50-0, A was reinforced on 50% of trials, and LA was reinforced on 0%. For both groups, B and C were reinforced on 100% and 50% of trials, respectively. The final five session included non-reinforced probe trials in which L was presented in compound with B or C. Black bars show responses to each auditory CS+, striped bars show response rates to compounds: the compound used in feature negative training (black and gray) and the probe compounds (black and white). Vertical bars show the standard error of the mean difference between response rates to each CS+ and L&CS+ compound.

The key analyses compared responding to the two probe compounds, LB and LC with their respective CS+s, B and C. We ran a 2x(2)x(2) ANOVA that compared the two groups as the between-subjects factor, compared trials with B versus trials with C as one within-subjects factor, and compared compound trials (LB and LC) versus CS+ alone trials (B and C) as the second within-subjects factor. Analyses of main effects showed that there was no overall difference between the two groups, $F < 1$, $\eta^2_p = 0.03$ (95% confidence intervals: 0 and 0.24), but there was a significant overall difference between trials with B and trials with C, $F(1, 30) = 45.78$, $p < .001$, $\eta^2_p = 0.60$ (0.47 and 0.75), consistent with the difference in their reinforcement rates. There was a significant main effect for the presence versus absence of L, $F(1, 30) = 53.79$, $p < .001$, $\eta^2_p = 0.64$ (0.52 and 0.78), confirming that L acted as an inhibitor to reduce response rates. There was a significant interaction between these two within-subjects factors, $F(1, 30) = 7.21$, $p = .012$, $\eta^2_p = 0.19$ (0.02 and 0.51), indicating that L had a bigger effect on responding to C than to B. While the group factor did not interact with the difference between B and C, $F < 1$, $\eta^2_p = 0.02$ (0 and 0.21), it did interact significantly with the effect of L, $F(1, 30) = 9.47$, $p = .004$, $\eta^2_p = 0.24$ (0.05 and 0.51). This interaction confirms that the effect of L was significantly greater in Group 50-0 than in Group 100-50. The triple interaction fell just short of a conventional level of significance, $F(1, 30) = 3.53$, $p = .070$, $\eta^2_p = 0.11$ (0 and 0.24). We conducted separate 2x(2) ANOVAs for trials involving B versus C in order to establish whether the difference in effectiveness of L between the two groups was reliable for both CS+s. These analyses confirmed that there was a significant interaction between group and the comparison between LB and B, $F(1, 30) = 4.25$, $p = .048$, $\eta^2_p = 0.12$ (0 and 0.38); there was

also a significant interaction between group and the comparison between LC and C, $F(1, 30) = 7.28$, $p = .011$, $\eta^2_p = 0.20$ (0.02 and 0.47). Thus L had significantly greater inhibitory effect in Group 50-0 than in Group 100-50, both for tests with B and tests with C. Follow-up paired *t*-tests confirmed that response rates to LB were significantly less than to B, both for Group 100-50, $t(15) = 3.81$, $p = .002$, Cohen's d = 0.94 (0.34 and 1.52), and for Group 50-0, $t(15) = 6.36$, $p < .001$, Cohen's d = 1.59 (0.83 and 2.33). Response rates to LC were also significantly less than to C, both for Group 100-50, $t(15) = 3.80$, $p = .002$, Cohen's d = 0.93 (0.33 and 1.52), and for Group 50-0, $t(15) = 4.44$, $p < .001$, Cohen's d = 1.11 (0.47 and 1.73).

Discussion

This experiment confirmed that the inhibitory strength of a CS- is greater when it is trained against a weak CS+ than when trained against a strong CS+ if the CS- signals the same decrease in reinforcement rate in each case. Specifically, a CS- that signaled a decrease in reinforcement rate of 50% was more effective when it signaled a decrease from 50% to 0% during training than when it signaled a decrease from 100% to 50%. This difference in effectiveness was evidence both when the CS- was tested against a CS+ with high reinforcement rate (100%) and when tested against a CS+ with low reinforcement rate (50%).

Experiment 4

In the experiments thus far, a CS- that signaled a decrease in reinforcement from a high rate to a lower rate had limited inhibitory effect when tested against a CS+ reinforced at a low rate. In most of these experiments, the reinforcement rate of the test CS+ was the same as that of the CS- compound during feature-negative training. However in one experiment (Experiment 1a), the CS- was tested against a CS+ that had been reinforced at an even lower rate than the training compound. In that case, the CS- not only failed to reduce responding to the weak CS+, it slightly increased the rats' response rates. While this increase was not statistically significant, it warrants further attention because similar findings have been reported by others (Mackintosh & Cotton, 1985; Matzel, Gladstein, & Miller, 1988; Nelson, 1987). Mackintosh and Cotton described an experiment in which rats could press a lever to obtain sucrose in the presence of a signal. The rats were trained on a discrimination in which a tone signaled the availability of a strong (20%) sucrose solution, whereas a light-tone compound signaled availability of a weaker (8%) solution. Under these conditions, the light acquired inhibitory strength as demonstrated by a reduction in lever pressing when the light was presented in compound with another cue that signaled 20% sucrose. However, the same light produced a paradoxical increase in responding when presented in compound with a cue that signaled a very weak (2%) sucrose solution. A comparable result was reported by Nelson (1987). She trained rats to press a lever for food in the presence of a tone that signaled availability of food on 100% of trials. The rats additionally learned that a compound of the tone and a light signaled a reduced availability (50%). The inhibitory strength of the light was assessed in a summation test against a clicker that signaled either a high (90%) or very low (15%) availability of food. The light reduced responding to the clicker that signaled high availability of food, but increased responding to the clicker that signaled low availability of food. Finally, Matzel et al. (1988) reported that a CS- that signaled a decrease in the rate of shock from 67% to 33% could act as an inhibitor in a summation test with a strong CS+ or in a retardation test, but showed

evidence of excitatory conditioning when presented on its own. All of these authors concluded that a CS- acquires inhibitory strength by virtue of signaling a reduction in strength or probability of reinforcement, but can simultaneously acquire excitatory strength based on its own (albeit weak) association with the reinforcer, and this weak excitation can be revealed when the nominal CS- is presented on its own or in compound with an even weaker CS+.

Experiment 4 was designed to replicate this observation in the present Pavlovian conditioning paradigm (cf. Harris, Andrew, & Kwok, 2013). A summary of the design is shown in Table 4. Rats were trained with a feature-negative discrimination between an auditory CS+, A, that was reinforced at a high rate (1/10-s) and a compound, LA, that was reinforced at a lower rate (1/30-s). They were concurrently trained with two other auditory CS+s, one (B) was reinforced at a high rate (1/10-s) and the other (C) was reinforced at a very low rate (1/120-s). The rats were eventually tested for responding to probe presentations of the compounds LB and LC. We expected that L would reduce responding to B. Based on findings reported by Mackintosh and Cotton (1985), Nelson (1987), and Matzel et al. (1988), we predicted that L would increase responding to C. The results were consistent with these predictions. Therefore, we extended the experiment for a further 8 days during which we included presentations of L on its own to see whether rats would respond to the CS- by itself. We observed a modest but significant level of responding to L that gradually extinguished across the 8 days. Over this period, we maintained the probe presentations of LB and LC to assess how L continued to influence responding to B and C.

Table 4. Summary of designs of Experiments 4 and 5.

Experiment	Training	Probe Test	Extinction	Probe Test
4	A10s+ B10s+ C120s+ **LA30s+**	LB & LC	L-	LB & LC
5	A100% B100% C13% **LA33%**	LB & LC	A-	LB & LC

Note: A, B, and C are auditory CSs, L is a light. In Experiment 4, each CS or compound was reinforced (+) after 10 s, 30 s, or 120 s. In Experiment 5, each CS or compound was reinforced on 100%, 33% or 13% of trials.

Methods

Subjects and apparatus

The experiment used 16 rats of the same strain and source, and housed in the same manner, as described previously. The apparatus described earlier was used in this experiment.

Procedure

Rats first received magazine training before commencing daily conditioning sessions for 34 days. Each session contained 48 trials, 12 trials of each of the three auditory CSs and 12 trials of the compound LA. CS and compound durations varied randomly between 2 and 18 s (mean = 10 s) for A and B, between 2 and 238 s (mean = 120 s) for C, and between 2 and 58 s (mean = 30 s) for LA. A single food pellet was delivered at the termination of the CS or

compound on every trial. The final 6 of these sessions (Day 29 to 34) included "probe" trials in which L was presented once with B and once with C. These probe trials had a fixed duration of 30 s, and were never reinforced with food. The experiment was then extended for a further 8 days to assess responding to L on its own. In each of these eight sessions, presentations of A, B, and C continued as before, as well as the probe trials with LB and LC. Presentations of LA were reduced in number to only one per day (fixed duration 30 s), and instead there were 12 presentations of L on its own (variable duration with a mean of 30 s).

Results

Response rates for each trial type (CS, compound or probe trial) from Days 29 to 42 are shown in Figure 4. The initial focus of interest was on the responses during the first six of these days (Day 29 to 34), when probe trials with LB and LC were first included in the training schedule. Over this period, the rats were responding differentially to the auditory CSs and the compound LA according to their reinforcement rates. To analyze these data we ran an ANOVA to compare response rates across the four training trial types (A, B, C, and LA) averaged over the six days. This showed a significant main effect of trial type, $F(1.8, 27) = 82.05$, $p < .001$, $\eta^2_p = 0.85$ (95% confidence interval: 0.79 and 0.91). The data of primary interest are the response rates on the probe trials with the compounds LB and LC. As is evident in Figure 4, response rates to LB were lower than to B, but response rates to LC were actually higher than to C. To test these effects, we conducted a 2x2 ANOVA on the data averaged over the six days. One factor compared the two training CSs (B and C) versus the two probe compounds (LB and LC); the other factor compared trials containing B (B and LB trials) with trials containing C (C and LC trials). There was no significant main effect for the difference between the two training CSs and the two probe compounds, $F < 1$, $\eta^2_p < 0.01$ (0 and 0.30), but there was a significant main effect for the difference between trials containing B (B and LB) and trials containing C (C and LC), $F(1, 15) = 69.98$, $p < .001$, $\eta^2_p = 0.82$ (0.72 and 0.92). There was a significant interaction between these main effects, $F(1, 15) = 98.83$, $p < .001$, $\eta^2_p = 0.87$ (0.82 and 0.93). This interaction was investigated using paired *t*-tests which showed that responses to LB were significantly less than to B, $t(15) = 5.76$, $p < .001$, Cohen's d = 1.43 (0.72 and 2.14), and response rates to LC were significantly greater than to C, $t(15) = 3.97$, $p = .001$, Cohen's d = 0.99 (0.38 and 1.59).

From Day 35 on, the experiment was extended for 8 days to test for responding to L alone. To analyze these data we ran a 2x8 ANOVA comparing responding during L versus the pre-CS period across the 8 days. There was a significant difference overall between L and pre-CS response rates, $F(1, 15) = 8.05$, $p = .012$, $\eta^2_p = 0.35$ (95% confidence interval: 0.03 and 0.78), as well as a significant effect across days, $F(4.3, 65) = 6.03$, $p < .001$, $\eta^2_p = 0.29$ (0.17 and 0.48), and a significant interaction between the two main effects, $F(3.8, 57) = 7.66$, $p < .001$, $\eta^2_p = 0.34$ (0.24 and 0.53). Follow up *t*-tests showed that responding during L was significantly higher than during the pre-CS period on each of Days 35 to 37, smallest $t(15) = 3.12$, largest $p = .007$, Cohen's d ≥ 0.78 (0.21 and 1.33), but there was no difference between L and pre-CS rates on any of the subsequent five days, all *t*s < 1, Cohen's d ≤ 0.21 (-0.27 and 0.71). Thus the rats exhibited conditioned responding to L, but these responses extinguished across the repeated non-reinforced presentations of L each day.

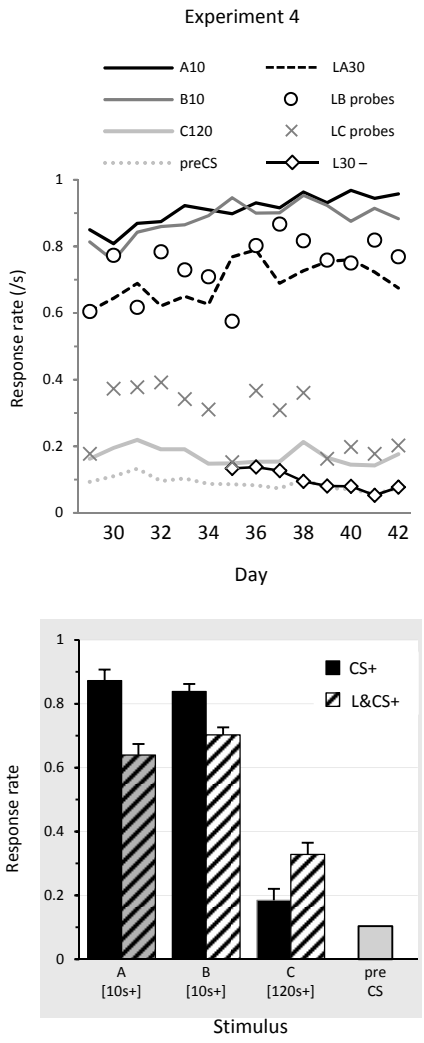

Figure 4. Top: Mean response rates for each trial type on each day from Days 29 to 42 in Experiment 4. The training schedule from Days 1 to 34 involved 12 reinforced presentations of each of three auditory CSs, A (mean duration 10 s), B (10 s), and C (120 s), as well as the compound LA (mean duration 30 s). Each session from Days 29 to 34 also included two non-reinforced probe trials, one with the compound LB and one with LC. Sessions from Days 35 to 42 continued this schedule except that LA trials were reduced in number from 12 to just one, and 12 trials were added in which L was presented on its own (mean duration 30 s). The lower left graph shows mean response rates averaged over the six sessions from Days 29 to 34. Black bars show responses to each auditory CS+, striped bars show response rates to compound trials with L. Vertical bars show the standard error of the mean difference between response rates to each CS+ and L&CS+ compound.

Given the change in responding to L across the 8 days, we conducted two 2x8 ANOVAs to assess the inhibitory effect of L against B and against C. The first of these ANOVAs showed that responding to LB was significantly less than to B, $F(1, 15) = 21.70$, $p < .001$, $\eta^2_p = 0.59$ (95% confidence interval: 0.30 and 0.84), that there was no main effect of day, $F(3.1, 47) = 1.42$, $p = .248$, $\eta^2_p = 0.07$ (0.03 and 0.31), but there was a significant interaction $F(3.2, 47) = 2.89$, $p = .043$, $\eta^2_p = 0.16$ (0.08 and 0.38). Despite this significant interaction, follow-up *t*-tests did not identify any clear trend for a change in the difference between LB and B across days – the difference between LB and B was significant on Days 35 ($t(15) = 3.50$, $p = .003$, Cohen's d = 0.87 [0.28 and 1.44]), 38 ($t(15) = 2.73$, $p = .016$, Cohen's d = 0.68 [0.13 and 1.22]), 39 ($t(15) = 3.60$, $p = .003$, Cohen's d = 0.90 [0.30 and 1.47]), and 40 ($t(15) = 2.38$, $p = .031$, Cohen's d = .59 [0.05 and 1.12]), but not on Days 36 ($t(15) = 1.67$, $p = .116$, Cohen's d = 0.42 [-0.10 and 0.92]), 37 ($t < 1$, Cohen's d = 0.15 [-0.34 and 0.64]), 41 ($t(15) = 2.02$, $p = .062$, Cohen's d = 0.50 [-0.03 and 1.02]) and 42 ($t(15) = 1.79$, $p = .094$, Cohen's d = 0.45 [-0.08 and 0.95]). The second ANOVA showed that responding to LC was significantly greater than to C overall, $F(1, 15) = 27.14$, $p < .001$, $\eta^2_p = 0.64$ (0.32 and 0.89), that there was a significant main effect of day, $F(4.3, 64) = 3.05$, $p = .021$, $\eta^2_p = 0.17$ (0.12 and 0.34), and there was a significant interaction between these main effects, $F(4.3, 64) = 3.70$, $p = .008$, $\eta^2_p = 0.20$ (0.12 and 0.39). This interaction did correspond to a systematic change in the difference between LC and C, and paired *t*-tests confirmed that responding to LC was significantly greater than to C on Days 36 to 38, smallest $t(15) = 3.42$, largest $p = .004$, Cohen's ds > 0.84 (0.27 and 1.42), whereas there was no difference between LC and C on any of the subsequent four days, largest $t(15) = 1.25$, smallest $p = .231$, Cohen's ds < 0.32 (-0.2 and 0.81).

Discussion

This experiment has confirmed that a CS- that signals a decrease from a high to a lower reinforcement rate is effective on transfer test at reducing responding to another CS+ with a high reinforcement rate, but is ineffective at reducing responding to a CS+ with very low reinforcement rate. Indeed, the nominal inhibitor significantly increased responding to the low CS+, and could elicit magazine responses even when presented on its own. The effect of L on responding to the low CS+ was obtained despite the fact that the inhibitor signaled a much larger decrease in reinforcement rate (from 1/10 s to 1/30 s = -1/15 s) than the increase in reinforcement rate signaled by the low CS+ (1/120 s). The paradoxical excitatory effect of L replicates earlier findings that a light that signaled a reduction in strength or probability of a reinforcer would increase responding in a summation test with a CS+ that signaled a very weak or very improbable reinforcer (Mackintosh & Cotton, 1985; Matzel et al., 1988; Nelson, 1987). Finally, in the present experiment, the paradoxical excitatory effect of L was extinguished when it was repeatedly presented alone without reinforcement; this treatment did not appear to affect the inhibitory strength of L when tested against the highly reinforced CS+, and did not uncover any inhibitory effect of L against the low CS+.

Experiment 5

The paradoxical excitatory effect of L that was observed in Experiment 4 was investigated further in Experiment 5. Rats were trained with a CS, A, that was reinforced on 100% of trials, and with a compound LA that was reinforced on 33% of trials. Thus L signaled a decrease in reinforcement rate of 67%. The inhibitory effect of L was then assessed on summation tests in which L was presented in compound with each of two CS+s: B, reinforced on 100% of trials, and C reinforced on 13% of trials. If the results of Experiment 4 were replicated here, rats would respond less to LB than to B, but would respond more to LC than to C. One potential source for the apparent excitatory effect of L could involve a second order association via A (Rescorla, 1982; Williams, Travis, & Overmier, 1986). That is, L might increase responding when tested with a very weak CS+ (C) not because L itself has a direct association with food, but because L is associated with A that has a strong association with food. The present experiment was designed to test this possibility by including a phase in which responding to A was extinguished between two test periods that included probe trials of compounds LB and LC. The design of the experiment is summarized in Table 4. If the apparent excitatory effect of L is due to a second-order association via A, then extinction of A should remove that excitation, and thus responding to LC should not be greater than C.

Methods

Subjects and apparatus

The experiment used 16 naive rats of the same strain and source, and housed in the same manner, as described previously. The apparatus described earlier was used in these experiments.

Procedure

Rats first received magazine training before commencing 34 daily conditioning sessions that contained 60 trials, 15 trials of each of the three auditory CSs and 15 trials of the compound LA. All CS and compound presentations varied between 2 and 58 s (mean = 30 s); the probability of food delivery at the end of the presentation was 100% for A and B, 13% (2 out of 15 trials) for C, and 33% for LA. The final five sessions of these 34 days

included "probe" trials in which L was presented once with B and once with C. The probe trials had a fixed duration of 30 s, and were never reinforced with food. The experiment was then extended for a further 16 days (until Day 50). In each session, trials with B and C continued exactly as before, as did the probe trials of LB and LC, but trials with A were now presented without food to extinguish responding to A. During these 16 days, there were no trials with the compound LA.

Results

The defect in the infra-red detector of one conditioning chamber referred to in Experiment 3 meant that data from one rat had to be excluded from all analyses of the present experiment as well. The plot in the top half of Figure 5 shows the response rates of the remaining 15 rats to each trial type on each of the final 21 days of the experiment, covering the period when LB and LC probe trials were included in each session. The two charts in the lower part of the figure show the response rates averaged over Days 30 to 34 (left panel) before responding to A was extinguished, and averaged over Days 46 to 50 (right panel), after responding to A had been extinguished. An overall ANOVA confirmed that there were significant differences in the rats' response rates to the trained trial types (A, B, C and LA for the first test period; A, B, and C for the second test period): $F(1.5, 22) = 33.05$, $p < .001$, $\eta^2_p = 0.70$ (95% confidence interval: 0.61 and 0.83), across Days 30 to 34; $F(1.1, 15) = 49.45$, $p < .001$, $\eta^2_p = 0.78$ (0.70 and 0.89), across Days 46 to 50.

To analyze the effect of L on responding to B and C, we conducted a 2x2x2 ANOVA; the three factors of this analysis were: test phase (1 versus 2), presence versus absence of L, and test CS+ (B versus C). All three main effects were significant, smallest $F(1, 14) = 9.53$, $p = .008$, $\eta^2_p = 0.40$ (95% confidence interval: 0.12 and 0.71), showing that responding during the second test was lower overall than during the first, that responding during C was less than B, and that responding during probe trials of LB and LC was lower than during B and C trials without L. The presence of L interacted significantly with CS+, $F(1, 14) = 78.47$, $p < .001$, $\eta^2_p = 0.85$ (0.78 and 0.92), showing that L's inhibitory effect was greater when tested with B than with C. L also interacted significantly with test, $F(1, 14) = 27.38$, $p < .001$, $\eta^2_p = 0.66$ (0.54 and 0.91), indicating that L's effectiveness changed from the first test phase to the second test phase. Finally, the three-way interaction fell just short of being significant, $F(1, 14) = 4.00$, $p = .065$, $\eta^2_p = 0.22$ (0 and 0.64), which suggests that the change in L's effect over time differed for its test trials with B versus C. We followed up this large ANOVA with two smaller 2x2 ANOVAs. These compared the presence versus absence of L (Factor 1), and the first versus second test phases (Factor 2), analysed separately for trials with B and trials with C. There was a significant main effect of L when tested with B, $F(1, 14) = 40.31$, $p < .001$, $\eta^2_p = 0.74$ (0.62 and 0.88), but not when tested with C, $F(1, 14) = 2.74$, $p = .120$, $\eta^2_p = 0.16$ (0 and 0.48). There was a significant main effect of test phase for C, $F(1, 14) = 8.14$, $p = .013$, $\eta^2_p = 0.37$ (0.03 and 0.81), but not for B, $F(1, 14) = 3.09$, $p = .101$, $\eta^2_p = 0.18$ (0 and 0.59). Of most importance, there were significant interactions between the test phases and the effect of L, both for tests with B, $F(1, 14) = 20.73$, $p < .001$, $\eta^2_p = 0.60$ (0.46 and 0.82), and for tests with C, $F(1, 14) = 9.94$, $p = .007$, $\eta^2_p = 0.42$ (0.17 and 0.66). These interactions establish that the extinction of A between test phases 1 and 2 increased the inhibitory effect of L on responding to B, and decreased evidence for L's paradoxical excitatory effect on responding to C. The latter conclusion, with respect to responding

to C, was confirmed by follow-up paired *t*-tests. These showed that responding to LC was significantly above responding to C during the first test phase, $t(14) = 2.76$, $p = .015$, Cohen's d = 0.71 (0.13 and 1.27), but there was no longer any difference between C and LC by the second test phase, $t < 1$, Cohen's d = 0.08 (-0.43 and 0.58). Indeed, in the second phase response rates during C and LC were almost identical (mean = .17 for C; mean = .16 for LC). The Bayes factor calculated on the *t* statistic suggested that the likelihood that LC and C were equal was five times greater than the likelihood that they were different.

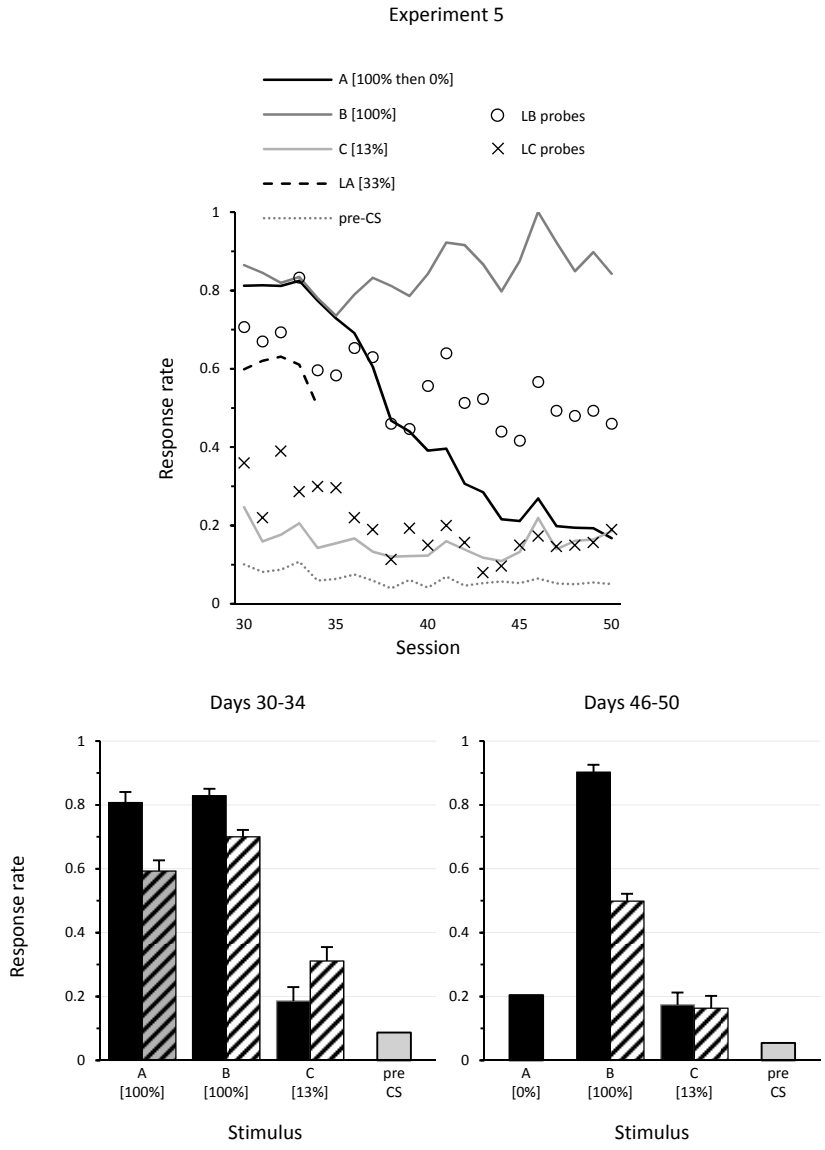

Figure 5. Results of Experiment 5. The line plots in the top part of the figure show average daily response rates, from Days 30 to 50, to each of three auditory CSs, A, B, and C, that were reinforced on 100% or 13% of trials, as well as to the compound LA reinforced on 33% of trials. Each of the sessions shown here included probe trials in which L was presented once with B and once with C. From Day 35 on, presentations of A were no longer reinforced, and all trials with LA were omitted. The bar chart on the lower left shows the response rates averaged over the first five sessions that included the probe trials (Days 30 to 34), before extinction of A. The chart on the lower right shows response rates averaged over the final five sessions (Days 46 to 50), after responses to A had been extinguished. Black bars show responses to each auditory CS+, striped bars show response rates to compound trials with L. Vertical bars show the standard error of the mean difference between response rates to each CS+ and L&CS+ compound.

Discussion

Experiment 5 confirmed that a stimulus, L, that signals a large decrease in reinforcement rate can reduce responding when presented in compound with a CS+ that signals a high rate of reinforcement but fails to reduce responding when tested with a CS+ that signals a low rate of reinforcement. The failure of L to reducing responding to the weak CS+ is especially noteworthy because, across training, L had signaled a 66% decrease in reinforcement rate that was five times larger than the increase in reinforcement rate signaled by the weak CS+ (13%). If L had acquired a simple inhibitory associative strength across training, one would expect this inhibition to be sufficient to reduce if not eliminate responding to the weak CS+.

When the effect of L on responding to a weak CS+ was initially tested, L significantly increased responding, as it had done in Experiment 4. However, after responding to L's training CS+, A, had been extinguished, L lost this excitatory influence on responding. This suggests that L's excitatory effect when tested with the weak CS+ was due to a second-order association between L and A, rather than a direct association between L and food.

Experiment 6

Experiments 4 and 5 have shown that, when a CS- signals a decrease in reinforcement from a high to an intermediate rate, it can carry excitatory strength via a second order association with the training CS+, as well as acquiring inhibitory strength. While this second order excitatory strength has only been evident when the CS- was presented alone or in compound with a very weak CS+, the existence of this second order association may influence the impact of the CS- at other times by countering its inhibitory strength. This could explain our observation in Experiments 1 to 3 that a CS- that signals a decrease in reinforcement from a high rate to an intermediate rate acquires less inhibitory strength than a CS- that signals the same magnitude decrease in reinforcement but from an intermediate rate to a low (or zero) rate. It is possible that both of these CS-s acquire the same amount of inhibition, as predicted by models that equate learning with a difference between expected and actual reinforcement, but the expression of that inhibition differs between the CS-s due to differences in the strength of their second order excitatory associations with the US. If this explanation is correct, then both CS-s should show equivalent inhibitory strength if their second order association with the US is extinguished prior to testing. This was the primary objective of Experiment 6. Two groups of rats were trained on a discrimination between A and LA. For one group, "100-50-extA", A was reinforced on 100% of trials, and LA on 50% of trials. For the other group, "50-0-extA", A was reinforced on 50% of trials, and LA on 0% of trials. For both groups, the inhibitory strength of L was assessed in transfer tests when L was presented in compound with either B, reinforced on 100% of trials, or C, reinforced on 17% of trials. Before these tests were conducted, responding to A was extinguished for both groups, thus minimizing any second order association between L and the US via A.

Experiment 6 had a second objective related to our demonstration in Experiment 5 that L's excitation was lost after extinction of the training CS+. Experiment 6 sought to confirm this finding, while including a control group for which the training CS+, A, was not extinguished but another CS+ was extinguished instead. This group thus controls for the effects of extinction *per se*, and therefore to establish whether the effect

observed in the previous experiment was indeed due to extinction of the second order connection between L and food via A. Therefore, in addition to the two groups described in the preceding paragraph, Experiment 6 included a third group, "100-50-extD", that was trained with the discrimination between A, reinforced on 100% of trials, and LA, reinforced on 50% of trials. These rats then received extinction of a different CS+, D, that had been reinforced on 100% of trials, before they were tested with L in compound with B and C. A summary of the training schedule for all three groups is shown in Table 5.

Table 5. Summary of design of Experiment 6.

Group	Training	Extinction	Probe trials
100-50-extD	A[100%] B[100%] C[17%] D[100%] **LA[50%]**	D-	LB and LC
100-50-extA	A[100%] B[100%] C[17%] D[100%] **LA[50%]**	A-	LB and LC
50-0-extA	A[50%] B[100%] C[17%] D[100%] **LA[0%]**	A-	LB and LC

Note: A, B, C and D were auditory CSs reinforced on 100%, 50% or 17% (one in six) of trials. L was a light.

Methods

Subjects and apparatus

The experiment used 48 female rats, aged 8 weeks at the start of the experiment. They were of the same strain and source, and housed in the same manner, as described previously. The apparatus described earlier was used in these experiments.

Procedure

Rats initially received 20 daily conditioning sessions with 12 trials each of A, LA, B, C, and D. For Groups 100-50-extA and 100-50-extD, A was reinforced on 100% of trials, LA on 50%; for Group 50-0-extA, A was reinforced on 50% of trials, LA on 0%. For all three groups, B was reinforced on 100% of trials, C on 17% (two of 12), and D on 100%. All single and compound trials had variable length with a mean of 20 s.

For the next 15 days, trials with B and C continued to be reinforced as before. Trials with A were not reinforced for Groups 100-50-extA and 50-0-extA, but trials with D continued to be reinforced as before. For Group 100-50-extD, trials with D were not reinforced, but trials with A continued to be reinforced as before. There were no LA trials for any groups. In all groups, the CS+ that was undergoing extinction (A or D) increased in number from 12 to 24 trials per session, in order to facilitate extinction and to make up for the omission of the LA trials. The final 5 days of the experiment (Days 31 t 35) included probe trials with the compounds LB and LC. Each session included two LB and two LC probe trials; each had a fixed duration of 20 s and was not reinforced with food. They occurred at trials 11, 22, 33, and 44, in interleaved fashion and their order was counterbalanced across rats and across days.

Results

Response rates averaged over the final five sessions (Days 31-35) are shown for all three groups in Figure 6. The data of primary interest are response rates to CSs B and C, and during probe presentations to the compounds LB and LC. These data were analyzed as two separate 2x(2)x(2) ANOVAs. The first analysis compared Group 100-50-extA with Group 100-50-extD to assess how the extinction of A (rather than D) affected the inhibitory strength of L. The second analysis compared Group 100-50-extA with

Group 50-0-extA, to assess whether the inhibitory strength of L differed between these groups (as it had in Experiment 3) after extinction of A. Both ANOVAs tested the same within-subjects factors: they compared trials with L versus trials without L, and compared trials with B versus trials with C. Because these ANOVAs were only partially independent (they both included Group 100-50-extA), we divided alpha by 2 to control the familywise error rate.

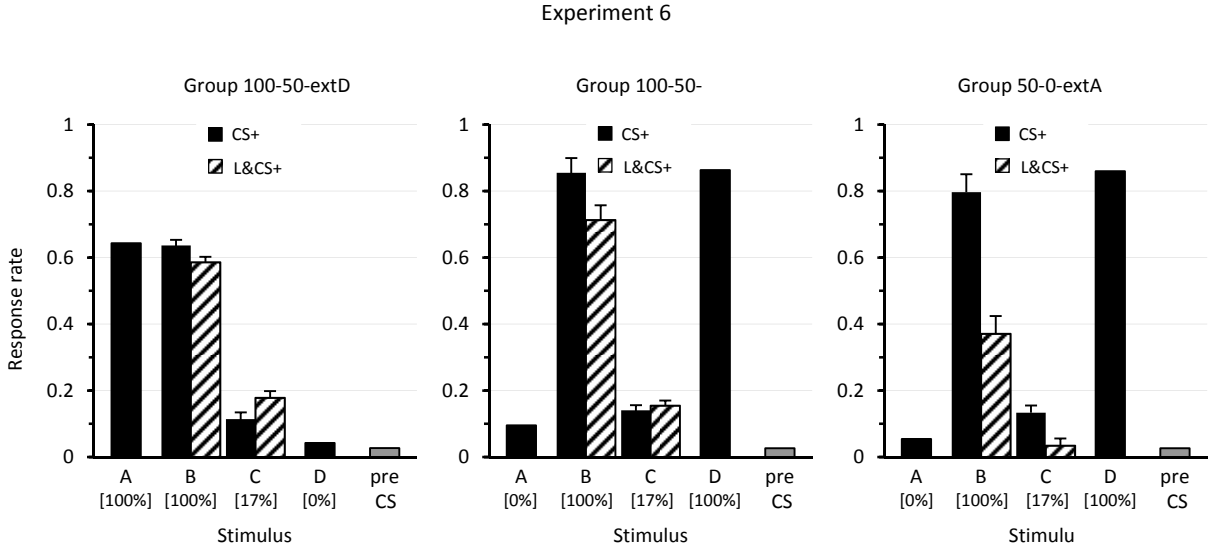

Figure 6. Average response rates over the last five sessions of Experiment 6. Three groups of rats were trained for 20 days with four auditory CS+s, A, B, C and D, and a compound of a light, L, with A. For Groups 100-50-extD and 100-50-extA, A was reinforced on 100% of trials, and LA was reinforced on 50% of trials. For Group 50-0-extA, A was reinforced on 50% of trials, and LA was reinforced on 0%. For all three groups, B and C were reinforced on 100% and 17% of trials, respectively, and D was reinforced on 100% of trials. After these 20 days, all groups received 15 further days of training in which either A (for Groups –extA) or D (for Group –extD) was extinguished by repeated presentations without any reinforcement. During these 15 days, LA trials were omitted from the schedule, and trials with the other 3 CS+s continued as on the first 20 days. The final five sessions included non-reinforced probe trials in which L was presented in compound with B or C. Black bars show responses to each auditory CS+, striped bars show response rates to LB and LC compounds. Vertical bars show the standard error of the mean difference between response rates to each CS+ (B and C) and its probe compound (LB and LC).

The left panel of Figure 6 shows the response rates in Group 100-50-extD. In this group, responding to LB was slightly lower than to B, and responding to LC was higher than to C. A different pattern of responding was seen in Group 100-50-extA (middle panel of Figure 6); these rats showed a larger difference between LB and B and a smaller difference between LC and C. This observation was confirmed by the 2x(2)x(2) ANOVA. There was a significant main effect of CS+ (trials with B versus trials with C), $F(1, 15) = 71.16$, $p < .001$, $\eta^2_p = 0.70$ (95% confidence intervals: 0.62 and 0.81). The main effect for L did not reach the adjusted level of significance, $F(1, 15) = 4.16$, $p = .050$, $\eta^2_p = 0.12$ (0.00 and 0.34), and there was no significant main effect for Group, $F(1, 15) = 1.00$, $p = .329$, $\eta^2_p = 0.03$ (0.00 and 0.23). There was a significant two-way interaction between L and Group, $F(1, 15) = 6.28$, $p = .018$, $\eta^2_p = 0.17$ (0.02 and 0.40). This interaction is important because it

establishes that the inhibitory effect of L was greater in Group 100-50-extA than in Group 100-50-extD. There was no two-way interaction between Group and CS+, $F(1, 15) = 1.72$, $p = .199$, $\eta_p^2 = 0.05$ (0.00 and 0.28), but there was a two-way interaction between L and CS+, $F(1, 15) = 28.29$, $p < .001$, $\eta_p^2 = 0.49$ (0.33 and 0.67). There was no triple interaction, $F(1, 15) = 0.66$, $p = .422$, $\eta_p^2 = 0.02$ (0.00 and 0.22).

The right panel of Figure 6 shows the response rates in Group 50-0-extA. These rats responded much less to LB than to B, and also responded less to LC than to C, a pattern that was quite clearly different from that of Group 100-50-extA. The ANOVA that compared these groups revealed a significant main effect of CS+ (trials with B versus trials with C), $F(1, 15) = 80.34$, $p < .001$, $\eta_p^2 = 0.73$ (95% confidence intervals: 0.66 and 0.82), and a significant main effect for L, $F(1, 15) = 71.68$, $p < .001$, $\eta_p^2 = 0.70$ (0.60 and 0.82), but no significant main effect for Group, $F(1, 15) = 2.65$, $p = .114$, $\eta_p^2 = 0.08$ (0.00 and 0.30). There was a significant two-way interaction between L and Group, $F(1, 15) = 26.58$, $p < .001$, $\eta_p^2 = 0.47$ (0.25 and 0.70), such that the inhibitory effect of L was greater in Group 50-0-extA than in Group 100-50-extA. There was no two-way interaction between Group and CS+, $F(1, 15) = 1.16$, $p = .289$, $\eta_p^2 = 0.04$ (0.00 and 0.25), but there was a two-way interaction between L and CS+, $F(1, 15) = 47.71$, $p < .001$, $\eta_p^2 = 0.61$ (0.49 and 0.75). There was a significant triple interaction, $F(1, 15) = 5.99$, $p = .020$, $\eta_p^2 = 0.17$ (0.01 and 0.44).

Simple effects analyses tested the effect of L in all three groups. These paired comparisons confirmed that responding to LB was significantly less than to B for all three groups: Group 100-50-extD, $t(15) = 3.04$, $p = .008$, Cohen's d = 0.76 (0.19 and 1.31); Group 100-50-extA, $t(15) = 3.21$, $p = .006$, Cohen's d = .80 (0.23 and 1.36); and Group 50-0-extA, $t(15) = 8.07$, $p < .001$, Cohen's d = 2.02 (1.14 and 2.87). Responding to LC was significantly greater than responding to C in Group 100-50-extD, $t(15) = 3.16$, $p = .007$, Cohen's d = 0.79 (0.21 and 1.34), but LC and C did not differ in Group 100-50-extA, $t(15) = 0.92$, $p = .374$, Cohen's d = 0.23 (-0.27 and 0.72). Responding to LC was significantly less than to C in Group 50-0-extA, $t(15) = 4.72$, $p < .001$, Cohen's d = 1.18 (0.52 and 1.81).

Discussion

Experiment 6 has shown that a CS-, L, that signals a decrease in reinforcement from a high rate (100%) to an intermediate rate (50%) has greater inhibitory strength after its training CS+, A, has been extinguished than after a different CS+, in this case D, has been extinguished. Thus, this experiment confirms the observation in Experiment 5 that extinction of A removes any evidence for excitatory strength of L, and shows this excitatory strength was dependent specifically on the excitatory strength of A. Therefore, we can conclude that the excitatory strength of the CS- was indeed derived from a second order association with the US via the training CS+.

Experiment 6 additionally showed that when L signals a decrease in reinforcement rate from 50% to 0%, it acquires greater inhibitory strength than when it signals a decrease from 100% to 50% reinforcement. This replicates the findings obtained in Experiment 3, but goes beyond that result in demonstrating that the difference in inhibitory strength exists even after extinction of the training CS+, A. Therefore, the difference in effectiveness of L is not due to differences in second order excitation via A, but must reflect a true difference in inhibition acquired by L.

General Discussion

The present experiments investigated the relationship between inhibition and reinforcement rate in conditioned magazine approach with rats. Our previous studies had shown that what rats learn about an excitatory CS (CS+) is linearly related to the rate at which the CS+ is reinforced, inasmuch as their learning about, or responding to, a compound of two CS+s is the sum of what is learned about each individual CS+ (Andrew & Harris, 2011; Harris et al., 2012). On the basis of this evidence, we hypothesized that learning about a conditioned inhibitor (CS-) during a feature negative discrimination would be linearly related to the size of the decrease in reinforcement rate that it signaled. Moreover, when the CS- is combined with a CS+ in a summation test, its effect on responding could be estimated by simply subtracting the change in reinforcement rate that was signaled by the CS- in the feature negative discrimination from the reinforcement rate signaled by the CS+ in the summation test.

Neither of the hypotheses described above were confirmed by the present experiments. First, learning about a CS- was not a linear function of the size of the decrease in reinforcement rate that it signaled during the feature negative training. Rather, the strength of the CS- depended on the absolute value of the reinforcement rate during training. The clearest evidence for this was obtained in Experiments 3 and 6. Two groups of rats were given feature negative training, for one group the CS- signaled a decrease from 100% to 50% reinforcement, for the other group the CS- signaled a decrease from 50% to 0% reinforcement. Despite the fact that the CS- signaled an equivalent change in reinforcement rate (-50%) for the two groups, the inhibitory strength of the CS- was greater in the latter group than in the former group, both when the CS- was tested against a CS+ that signaled a high rate of reinforcement and when it was tested against a CS+ that signaled a lower rate of reinforcement.

The current experiments also disconfirmed our second hypothesis that the strength of a CS- could be estimated by simply subtracting the change in reinforcement that it signaled during the feature negative training from the reinforcement rate of the CS+ used in the summation test. Instead, we consistently observed that the effectiveness of a CS- varied depending on the reinforcement rate of the CS+ used in the summation test. If, during feature negative training, the CS- signaled a decrease in reinforcement from a high rate to an intermediate rate (greater than zero), then that CS- was effective at reducing responding when tested against a CS+ that signaled a high rate of reinforcement, but the same CS- was less effective, or even ineffective, when tested against a CS+ that signaled a low reinforcement rate.

Mackintosh and Cotton (1985) and Nelson (1987) reported a similar fate for a CS- that signaled a decrease in, but not omission of, reinforcement (also Matzel et al., 1988). They trained rats on a feature negative discrimination in which a CS- signaled a decrease in strength or rate of reinforcement from a high to an intermediate level. On subsequent test, this CS- was ineffective against a CS+ that had been conditioned with a weak US or very low rate of reinforcement. Indeed, they reported that the CS- could elicit a paradoxical increase in responding under some circumstances. That paradoxical result was also observed here in Experiments 4, 5 and 6. When a CS- signaled a decrease in reinforcement from a high rate to an intermediate rate (e.g., 100% to 33%), it produced a significant increase in responding when tested against a CS+ reinforced at a very low rate (13%). The authors of those earlier

papers concluded that, during feature negative training, the CS- acquired inhibitory strength because it signaled a fall in rate or strength of reinforcement, but it simultaneously acquired some excitatory associative strength because it was still associated with reinforcement. Experiments 5 and 6 in the present series have identified a different mechanism for this paradoxical excitatory effect of the CS-. In these experiments, all evidence for the excitatory effect of the CS- was eliminated by extinction of the CS+ with which it had been presented during feature negative training. Therefore, the ability of the CS- to increase responding when tested with a very weak CS+ depended on a second order association between the CS- and the training CS+, rather than any direct association between the CS- and the US.

In the present experiments, extinction of the training CS+ did not weaken, but rather increased, the inhibitory strength of the CS-. This observation is consistent with some earlier reports that extinction of the training CS+ can increase the inhibitory strength of a CS-, arguably by reducing the second order association between the CS- and US via the training CS+ (Rescorla, 1982; Williams et al., 1986). However, there are also existing studies in which extinction of the training CS+ was found to weaken the inhibitory strength of a CS- (Hallam, Matzel, Sloat, & Miller, 1990; Lysle & Fowler, 1985). In the experiments by Lysle and Fowler, evidence for inhibition was lost only if all excitors (including the conditioning context) were extinguished prior to a retardation test with the CS-. Inhibition could be restored to the CS- if a new excitatory CS were conditioned between extinction of the training CS+ and the retardation test of the CS-. Since our experiments all used summation tests to assay the inhibitory strength of the CS-, this would have fulfilled the conditions under which, according to Lysle and Fowler, inhibition would be preserved. But our findings cannot be so easily reconciled with those of Hallam et al., who found that extinction of the training CS+ (a light) reduced the inhibitory strength of a CS- (noise) even when that inhibition was assayed in a summation test against another CS+ (clicker). There are many procedural differences between our experiments and those of Hallam et al., the most obvious being related to the conditioning paradigm: we measured conditioned magazine approach to a CS signaling food, Hallam et al. measured fear-induced suppression of drinking by a CS signaling shock. However, it is unlikely that the differences in paradigm can explain the difference in results given the report by Williams et al., who obtained results comparable to ours but measured conditioned suppression to a CS paired with a shock US.

The results reported by Hallam et al. (1990) are uniquely consistent with the predictions of the Comparator Hypothesis (Stout & Miller, 2007), and therefore that model is contradicted by our finding that extinction of the training CS+ increases the inhibitory strength of the CS-. In the Comparator Hypothesis, the mechanism by which a CS- influences responding is very different from that of other associative models, in that the CS- itself has no inhibitory strength (and may even have some excitatory strength). Rather, evidence for inhibition depends on the excitatory strength of the training CS+ which is subtracted from the excitatory strength of any new compound containing the CS-. This description of inhibition is clearly incompatible with our observation that extinction of the training CS+ increases, rather than removes, the inhibitory strength of a CS-. Nonetheless, there are aspects of the

Comparator Hypothesis that do speak specifically to some of our data, since the model allows a CS- to have both excitatory strength of its own, if it has been paired with the US, and to have an inhibitory influence on responding via the excitatory strength of its comparator (the training CS+). This characterization of a CS- does resonate with the seemingly ambivalent effects of a CS- that has signaled a decrease in reinforcement rate but to a non-zero level.

There is a further implication of our conclusion that the inhibitory strength of a CS- can be undermined by a second order association between the CS- and US (via the training CS+). This conclusion uncovers a potential explanation for our observation that inhibition is relatively weak when a CS- signals a decrease from a high to an intermediate rate of reinforcement and is relatively strong when the CS- signals a decrease from intermediate to zero reinforcement. This difference could arise because of differences in the strength of the second order association with the US. That is, a CS- that signals a decrease from high to intermediate reinforcement is also associated with a strongly reinforced CS+, whereas a CS- that signals a decrease from intermediate to no reinforcement is associated with a weaker CS+. Thus the inhibitory strength of the former CS- could be masked by its second order excitatory association with the US. However, the results of Experiment 6 disconfirmed that hypothesis. That experiment compared the inhibitory strength of a light when it signaled a decrease from 100% to 50% reinforcement versus a decrease from 50% to 0%. In each case, the excitatory strength of the training CS+ was extinguished before testing the light's inhibitory strength, in order to minimize the impact of any second order association between the light and the US. Despite this treatment, the light acquired much stronger inhibition when it signaled a decrease from 50% to 0% reinforcement than from 100% to 50%, just as we had observed in experiments that did not include extinction of the training CS+.

As reviewed in the Introduction, theories of conditioning that provide a mechanism for inhibition typically predict that its acquisition and expression are determined by a simple subtraction. The current findings have important implications for these theories. They are particularly problematic for any theory, such as Rate Estimation Theory (RET, Gallistel & Gibbon, 2000), that assumes that the content of conditioning is directly (linearly) related to reinforcement rate. These theories are faulted by the finding that a CS- signaling a decrease in reinforcement from a high to an intermediate rate has only a weak inhibitory effect, or even no effect, against a CS+ that has been reinforced at a low rate. They are similarly troubled by the demonstration that a CS- will acquire more inhibitory strength if the base rate of reinforcement is lower even though the magnitude of decrease in reinforcement rate signaled by the CS- is held constant.

Our findings also have important implications for associative models, such as the Rescorla-Wagner model. According to that model, the relationship between the rate at which a CS is reinforced and its terminal associative strength, Vt, after extended training, depends on particular assumptions that must be made about learning rates. If learning during reinforcement and non-reinforcement proceeds at the same rate (specified by the parameter □ in the Rescorla-Wagner model), then Vt is linearly related to reinforcement rate (see Harris et al., 2012, for a derivation of this statement). As such, the predictions for inhibition are, like RET's predictions, contradicted by the present findings. However, if learning during non-

reinforcement is slower than during reinforcement (i.e., if □- < □+), as originally proposed by Rescorla and Wagner (1972) (see also Rescorla, 2002), then Vt is related to reinforcement rate by a hyperbolic function[1] of the sort shown in Figure 7 (see, Andrew & Harris, 2011, for derivation). Because the relationship is not linear under these assumptions, the inhibitory strength that develops to a CS- is not a constant proportion of the size of the decrease in reinforcement rate that is signaled by the CS-. If two inhibitors signal the same size of decrement in reinforcement rate but differ in the absolute level of reinforcement, then those inhibitors will acquire different inhibitory strengths. Consistent with the results obtained in the present experiments, the CS- that has been trained at a low rate of reinforcement will acquire greater inhibitory strength than the CS- trained at a higher rate of reinforcement.

As noted above, the present results are broadly consistent with a hyperbolic relationship between learning and reinforcement rate. This relationship correctly predicts that a given decrease in reinforcement rate will produce weak inhibition if the CS- is trained in the context of a high base rate of reinforcement. However, this account says nothing about our other observation that the evidence for inhibition is poor when the CS- is pitted in a summation test against a CS+ with low rate of reinforcement. This observation was most striking in Experiments 4, 5 and 6. Despite the fact that the CS- signaled a large reduction in reinforcement during feature negative training (e.g., from 100% to 33%), it had no detectable inhibitory effect on responding to a weak CS+ (e.g., reinforced at 13%). On the face of it, this observation could suggest a positively accelerating relationship between expectancy of reinforcement and responding, such that a given change in expectancy produces a larger difference in responding when the absolute expectancy is higher rather than lower. But this conclusion is unlikely to be correct for at least two reasons. First, we have previously shown there is decelerating non-linear relationship between reinforcement rate and response rate, similar to the hyperbolic function shown in Figure 7 (Harris & Carpenter, 2011). This would imply that the impact of a CS- on responding should be greater when tested against a lower base rate of reinforcement. The second problem for this conclusion is that all CS-s should show the same sensitivity to the reinforcement rate of the test CS+. Thus, a CS- that signals a reduction in reinforcement from a low rate to zero should have a larger effect when tested against a strongly reinforced CS+ than when tested against a weakly reinforced CS+. But in Experiments 1c, 2b, and 3, when such comparisons were made the reverse difference was observed – the CS- had a smaller suppressive effect on responding when tested against a more strongly reinforced CS+. These observations show that the evidence for inhibition is sensitive to the reinforcement rate of the training CS+, the reinforcement rate of the test CS+, and the relationship between the two. Clearly the amount of information acquired about a CS- during feature negative training is more complex than can be captured by a one dimensional variable such as V.

[1] This analysis of the Rescorla-Wagner model assumes that the absolute asymptote of learning, λ, is constant across all CSs because the reinforcer itself is never varied in our experiments (the reinforcer is always a single food pellet). The model can, nonetheless, account for differences in terminal associative strength, Vt, as a function of differences in reinforcement rate. As we explain here, the relationship between Vt and reinforcement rate depends on assumptions about the rate of learning during reinforcement and non-reinforcement, specified by the parameter β.

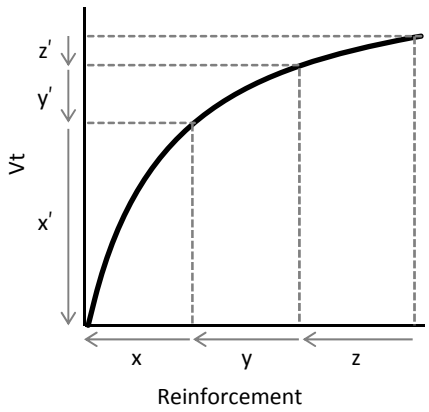

Figure 7. A hyperbolic function representing the relationship between reinforcement rate and terminal associative strength (Vt), as derived from the Rescorla-Wagner model (Rescorla & Wagner, 1972) when the rate of change is larger during reinforcement than non-reinforcement (i.e., b+ > b-). X, y and z, show equal sized changes in reinforcement rate from different base rates. Despite the fact that x = y = z, they produce different amounts of change in associative strength, with x' > y' > z'.

The simple conceptualization of inhibition as a single negative value has faced other long-standing problems that arise whenever the net strength of inhibition and excitation is less than zero. A key example is when an inhibitor is presented on its own, in the absence of any CS+ (or excitatory context); in such cases the net associative strength would be negative. One problem is to specify how a net negative strength could translate into an observable behavioral response. The second problem is whether a net negative strength in itself would produce new learning. According to the Rescorla-Wagner model (Rescorla & Wagner, 1972), when an inhibitor is presented on its own without a US, there would be a positive discrepancy between □ (zero) and the negative net associative strength. As a result, there should be an increase in associative strength. This account predicts extinction of inhibition, or if a neutral stimulus were presented with an inhibitor, that stimulus would acquire excitatory associative strength even if the compound were not reinforced. Zimmer-Hart and Rescorla (1974) showed that this prediction was incorrect (see also Rescorla, 1982; Williams et al., 1986; Witcher & Ayres, 1984).

As a solution to this problem, Rescorla (1979) invoked Konorski's (1948) conceptualization of inhibition as raising the threshold for activation of a US center. According to this account, an inhibitor shifts to the right the function relating strength of excitatory input to activation of the US center (see Figure 8A). Our findings suggest that this shift only operates on a specific part of the function. That is, whereas Rescorla (1979) argued that the lower bound on inhibition was zero net strength, our findings suggest that the lower bound on inhibition is set by the reinforcement schedule used during training. More specifically, if a CS- signals a decrease from high to low reinforcement, the CS- will only shift that part of the US activation function that lies above the lower level of reinforcement (see Figure 8B). It is admittedly difficult to understand how such an operation could be instantiated computationally, and as such the proposal may appear as little more than a re-description of the data. Nonetheless, linking the current results to that conceptual framework offers a potentially useful insight. It suggests that the failure to extinguish inhibition by non-reinforced presentations of the CS- is not just because □V cannot take a net negative value. Rather, the failure to extinguish inhibition reflects the more general constraint that inhibition only operates at or above values of excitation as specified during training. By extension, just as non-reinforced presentations of a CS- do not extinguish its inhibitory strength, we might expect that weak reinforcement of a CS- would also fail to affect its inhibitory strength if training of that

CS- involved a decrease from high to intermediate reinforcement.

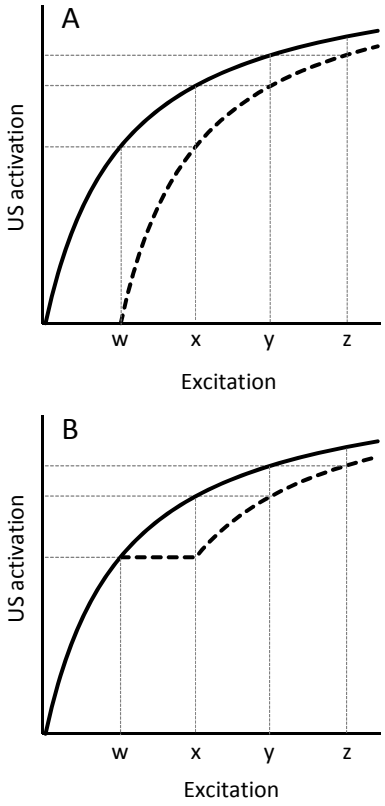

Figure 8. In both A and B, the solid curve shows a function relating excitation from a CS+ to activation of the US center. Under the influence of inhibition, this function is shifted to the right (dashed lines), effectively reducing the response to a given level of excitatory input. In **A**, a CS- that signals the total absence of the US during training (e.g., reinforcement decreases from w to zero) shifts the entire function (as proposed by Konorski, 1948; see also Rescorla, 1979). We propose that a CS- that signals a decrease from high to intermediate reinforcement (e.g., from x to w) only shifts a section of the function, as shown in **B**. Under such conditions, the CS- would still reduce responding to any CS+ with high excitatory strength (e.g., at y or z), but there would be no change in the response to weak excitatory input (below w).

References

Andrew, B. J., & Harris, J. A. (2011). Summation of reinforcement rates when conditioned stimuli are presented in compound. *Journal of Experimental Psychology: Animal Behavior Processes, 37*, 385-393.

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review, 107*, 289-344.

Hallam, S. C., Matzel, L. D., Sloat, J. S., & Miller, R. R. (1990). Excitation and inhibition as a function of posttraining extinction of the excitatory cue used in Pavlovian inhibition training. *Learning & Motivation, 21*, 59-84.

Harris, J. A., Andrew, B. J., & Kwok, D. W. S. (2013). Magazine approach during a signal for food depends on Pavlovian, not instrumental, conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 39*, 107-116.

Harris, J. A., Andrew, B. J., & Livesey, E. J. (2012). The content of compound conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 38*, 157-166.

Harris, J. A., & Carpenter, J. S. (2011). Response rate and reinforcement rate in Pavlovian conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 37*, 375-384.

Harris, J. A., Gharaei, S., & Pincham, H. L. (2011). Response rates track the history of reinforcement times. *Journal of Experimental Psychology: Animal Behavior Processes, 37*, 277-286.

Hentschke, H., & Stüttgen, M. C. (2011). Computation of measures of effect size for neuroscience data sets. *European Journal of Neuroscience, 34*, 1887-1894.

Konorski, J. (1948). *Conditioned reflexes and neuron organization*. Cambridge: Cambridge University Press.

Lysle, D. T., & Fowler, H. (1985). Inhibition as a "slave" process: Deactivatino of conditioned inhibition through extinction

of conditioned excitation. *Journal of Experimental Psychology: Animal Behavior Processes, 11*, 71-94.

Mackintosh, N. J., & Cotton, M. M. (1985). Conditioned inhibition from reinforcement reduction. In R. R. Miller & N. E. Spear (Eds.), *Information processing in animals: conditioned inhibition* (pp. 89-111). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Matzel, L. D., Gladstein, L., & Miller, R. R. (1988). Conditioned excitation and conditioned inhibition are not mutually exclusive. *Learning and Motivation, 19*, 99-121.

Nelson, K. J. (1987). Conditioned inhibition from incomplete reductions in the probability of reinforcement. *Quarterly Journal of Experimental Psychology, 39B*, 365-391.

Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex.* (G. V. Anrep, Trans.). New York: Dover.

Rescorla, R. A. (1969a). Conditioned inhibition of fear resulting from negative CS-US contingencies. *Journal of Comparative & Physiological Psychology, 67*(4), 504-509.

Rescorla, R. A. (1969b). Pavlovian conditioned inhibition. *Psychological Bulletin, 72*, 77-94.

Rescorla, R. A. (1979). Conditioned inhibition and extinction. In A. Dickinson & R. A. Boakes (Eds.), *Mechanisms of learning and motivation: A memorial volume to Jerzy Konorski* (pp. 83-110). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Rescorla, R. A. (1982). Some consequences of associations between the excitor and the inhibitor in a conditioned inhibition paradigm. *Journal of Experimental Psychology: Animal Behavior Processes, 8*, 288-298.

Rescorla, R. A. (2002). Comparison of the rates of associative change during acquisition and extinction. *Journal of Experimental Psychology: Animal Behavior Processes, 28*, 406-415. doi: 10.1037/0097-7403.28.4.406

Rescorla, 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.

Rogers, A. E. (1979). Nutrition. In H. J. Baker, J. R. Lindsey & S. H. Weisbroth (Eds.), *The laboratory rat* (Vol. 1, pp. 123-152). New York: Academic Press.

Rouder, J. N., Speckman, P. L., Sun, D., & Morey, R. D. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*, 225-237.

Stout, S. C., & Miller, R. R. (2007). Sometimes competing retrieval (SOCR): A formalization of the comparator hypothesis. *Psychological Review, 114*, 759-783.

Wagner, A. R., & Rescorla, R. A. (1972). Inhibition in Pavlovian conditioning: application of a theory. In M. S. Halliday & R. A. Boakes (Eds.), *Inhibition and learning* (pp. 301-336). San Diego, CA: Academic Press.

Williams, D. A., Travis, G. M., & Overmier, J. B. (1986). Within-compound associations modulate the relative effectiveness of differential and Pavlovian conditioned inihbition procedures. *Journal of Experimental Psychology: Animal Behavior Processes, 12*, 351-362.

Witcher, E. S., & Ayres, J. J. (1984). A test of two methods for extinguishing Pavlovian conditioned inhibition. *Animal Learning and Behavior, 12*, 149-156.

Zimmer-Hart, C. L., & Rescorla, R. A. (1974). Extinction of Pavlovian conditioned inhibition. *Journal of Comparative & Physiological Psychology, 86*(5), 837-845.