American Psychological Association

Journal of Experimental Psychology:
Animal Learning and Cognition


2329-8456/19/$12.00

2019, Vol. 45, No. 3, 280–289
http://dx.doi.org/10.1037/xan0000205

# Time-Scale-Invariant Information-Theoretic Contingencies in Discrimination Learning

Abigail Kalmbach
Columbia University and New York State Psychiatric Institute, New York, New York

Eileen Chun
New York State Psychiatric Institute, New York, New York

Kathleen Taylor
Columbia University and Barnard College at Columbia University

Charles R. Gallistel
Rutgers University

Peter D. Balsam
Columbia University; New York State Psychiatric Institute, New York, New York; and Barnard College at Columbia University

Animals optimize their behavior to maximize rewards by utilizing cues from the environment. In discrimination learning, cues signal when rewards can and cannot be earned by making a particular response. In our experiment, we trained male mice to press a lever to receive a reward on a random interval schedule. We then introduced a prolonged tone (20, 40, or 80 sec), during which no rewards could be earned. We sought to test our hypothesis that the duration of the tone and frequency of reward during the inter-tone-intervals affect the informativeness of cues and led to differences in discriminative behavior. Learning was expressed as an increase in lever pressing during the intertrial interval (ITI) and, when the informativeness of the cue was high, animals also reduced their lever pressing during the tone. Additionally, we found that the depth of discriminative learning was linearly related to the informativeness of the cues. Our results show that the time-scale invariant information-theoretic definition of contingency applied to excitatory conditioning can also be applied to inhibitory conditioning.

*Keywords:* conditioned inhibition, discrimination, information, operant behavior, contingency



Animals learn about the relations between stimuli, responses and outcomes. When positive or negative outcomes occur in the presence of some stimulus conditions but not in others, stimulus control over responding develops so the appropriate response is made under appropriate conditions. Stimulus control is manifest in the behavioral response following the onsets and offsets of stimuli. The onset of a positive conditioned stimulus ($S^+$) signals that an outcome has become more likely, while its offset signals that it has become less likely. Conversely, the onset of an $S^-$ signals that an outcome has become less likely and its offset signals that it has become more likely (Figure 1A). In both Pavlovian and operant conditioning such procedures are considered discrimination learning. The cues that come to control differential responding provide the most information about changes in the rate with which the outcomes occur (Balsam, 1988; Rescorla, 1969; Wagner, Logan, & Haberlandt, 1968; Wilkes & Gallistel, 2017). When a stimulus and reward are not correlated, that is they have zero contingency (Figure 1B), we do not expect the stimulus to gain control of behavior.

The potency of a cue to evoke responding as a function of its information value has been most extensively discussed in the Pavlovian conditioning literature. Rescorla brought the concept of contingency to prominence in his classic experiment in which he varied the rate of USs during the intertrial intervals, when the conditional stimulus (CS) (a tone) was absent, while holding constant their rate of reward during the CS (Rescorla, 1969). He found that when the USs were as frequent in the ITI as during the CS, a conditioned response to the CS did not develop, despite the repeated temporal pairing of the CS and the US. He concluded that it was the CS-US contingency that drove conditioning, not their

This article was published Online First April 25, 2019.

Abigail Kalmbach, Department of Psychiatry, Columbia University, and Department of Developmental Neuroscience, New York State Psychiatric Institute, New York, New York. Eileen Chun, Department of Developmental Neuroscience, New York State Psychiatric Institute. Kathleen Taylor, Department of Psychiatry, Columbia University, and Department of Psychology, Barnard College at Columbia University. Charles R. Gallistel, Rutgers Center for Cognitive Science, Rutgers University. Peter D. Balsam, Department of Psychiatry, Columbia University; Department of Developmental Neuroscience, New York State Psychiatric Institute; and Department of Psychology, Barnard College at Columbia University.

The authors thank Eleanor Simpson for excellent feedback and suggestions throughout this project.

Correspondence concerning this article should be addressed to Abigail Kalmbach, Department of Developmental Neuroscience, New York State Psychiatric Institute, Unit 40, 1051 Riverside Drive, New York, NY 10032. E-mail: ak3466@cumc.columbia.edu

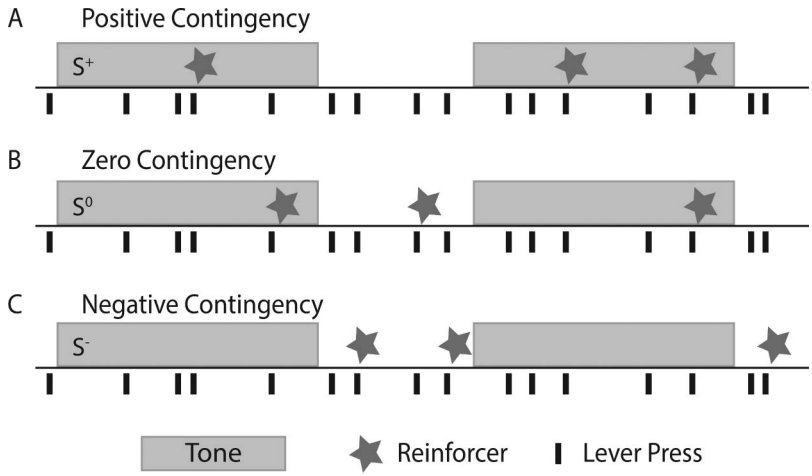


*Figure 1.* Contingencies of reward presentation with respect to stimuli. (A) Positive contingency occurs when reward availability is signaled by the conditioned stimulus ($S^+$). (B) Zero contingency occurs when the conditioned stimulus ($S^0$) is unrelated to reward availability. (C) Negative contingency occurs when reward availability is signaled by the absence of the conditioned stimulus ($S^-$). In our experiment, we tested the effects of varying the duration of $S^-$ (20s, 40s, and 80s) and the reward rate in the ITI (random interval of 20s and 40s).

temporal pairing. He did not, however, define contingency, and the well-known theory of associative learning subsequently developed made no reference to the concept (Rescorla & Wagner, 1972).

Contingency is now widely taken to be a critical variable in associative learning (Jackson, Horst, Pears, Robbins, & Roberts, 2016; Noonan, Chau, Rushworth, & Fellows, 2017; Schultz, 2015). It remains unclear, however, how to define it so as to make it a generally computable aspect of conditioning protocols. The most widely accepted definition of contingency is Allan's $\Delta p$, which is the difference between the probability of a reinforcer in the presence of a cue and its probability in the absence of that cue (Allan, 1993; Allan, 1980; Allan, Hannah, Crump, & Siegel, 2008; Allan & Jenkins, 1980; Allan, Siegel, & Tangen, 2005). This definition implicitly assumes that the continuous flow of experience in a conditioning protocol can be partitioned into trials on which associative connections (or action outcome values) are updated, depending on whether the cue was or was not present and on whether reinforcement did or did not occur. In this definition, the durations of intervals between events do not enter into the computation. Schultz defines contingency by reference to prediction errors, but again his definition is explicitly trial based; the errors are defined by the outcomes on trials, and the durations of those trials do not enter into the computation of the error (Schultz, 2015).

We have suggested that the conditioning process is to be understood not in terms of associative bonds but rather in the context of temporal learning (Balsam, Drew, & Gallistel, 2010; Balsam, Fairhurst, & Gallistel, 2006; Balsam & Gallistel, 2009; Balsam & Gibbon, 1981). In this view, the emergence of a conditioned response to a cue in excitatory Pavlovian conditioning results from subjects' learning that the expected wait to reinforcement is reduced by the onset of the cue. This view has led to a broadly applicable information-theoretic definition of contingency (Gallistel, Craig, & Shahan, 2014). The information-theoretic measure of contingency depends on the entropies of distributions defined over quantifiably or metrically represented intervals in the experimental protocol.

We hypothesize that animals form temporal maps of their experience and that these maps encoded metric information about the relationships between events (Balsam & Gallistel, 2009; Balsam et al., 2010). A metric representation is one to which the elementary arithmetic operations of addition, subtraction, multiplication and ordination may be applied. Intuitively, these quantities must be physically represented by physically realized symbols for numbers, as they are in a calculator or a computer. For example, when the number of occurrences of a reinforcing event is divided by the duration of the interval over which they were experienced to obtain the rate at which they were experienced, both the number of events and the duration of the interval must be represented in a way that makes it possible for a physically realized mechanism to divide the representation of the discrete quantity (the number of events) by the representation of the continuous quantity (the duration of the interval) to generate a representation of the third quantity (the rate).

There is a considerable literature on the representation of time in reinforcement learning (Daw, Courville, & Touretzky, 2006; Gershman, Moustafa, & Ludvig, 2014; Guilhardi, Yi, & Church, 2007; Kirkpatrick, 2014; Ludvig, Sutton, & Kehoe, 2008; Nakahara & Kaveri, 2010; Rivest, Kalaska, & Bengio, 2010; Suri & Schultz, 1999). Two representations appear repeatedly in this literature—in somewhat varying forms, and in combination with various rules for association formation (see Luzardo, Alonso, & Mondragón, 2017 for review): i) time is represented as a sequence of states (Schultz, Dayan, & Montague, 1997) or overlapping microstimuli (Gershman et al., 2014) or ii) a rate of accumulation or rate of directed drift is adjusted by reinforcement so that accumulation or net drift reaches a critical threshold value at the time of reinforcement (Luzardo et al., 2017; Simen, Balci, de Souza, Cohen, & Holmes, 2011; Simen, Rivest, Ludvig, & Killeen, 2013). It is

unclear how either form of temporal representation supports the basic arithmetic operations.

From a neural perspective—many of these models are more concerned with explaining the results from single neuron electrophysiological recordings than with explaining behavioral data—the microstimulus proposal is appealing, because there are many neurons that are temporal place cells: they fire at specific locations within elapsing intervals (Eichenbaum, 2014). Moreover, they have the properties posited by the microstimulus theory: the breadth of their temporal tuning curve increases with the duration of the interval to which they are tuned and the tuning curves overlap. There are also cells tuned to numerosity (Ditz & Nieder, 2016; Kutter, Bostroem, Elger, Mormann, & Nieder, 2018), with similar Weber-law properties. It is, however, unclear what neural mechanism could integrate numerocity-tuned and duration-tuned neurons to result in rate-tuned neurons or rate encoding neural activity. More importantly for present purposes, none of these proposals specifies how the brain might compute a temporal contingency. In the Discussion, we spell out the computation of information-theoretic temporal contingency, define the informativeness of our stimuli, and elaborate on its explanatory power.

In the present work, we extend the relevance of temporal contingency to inhibitory conditioning in an operant paradigm by studying protocols like the one shown in Figure 1C. Mice are taught to press a bar to earn food at unpredictable times only when a discriminative cue is absent. Lever presses made during the intervals between the onsets and the offsets of the cue are never reinforced, while presses made during the intervals between its offsets and its onsets are occasionally reinforced. In the operant literature, cues that modulate responding are called discriminative stimuli and are commonly denoted by S, whereas in the Pavlovian literature, these cues are called conditioned stimuli and denoted by CS. As this is an operant task with inhibitory conditioning, we will denote cues by $S^-$, and we will denote their onset and offset by $S\uparrow$ and $S\downarrow$, respectively. Because we are studying the role of a cue for nonreinforcement in operant discrimination, we use 'inhibition' somewhat loosely to refer to behavioral suppression. We do not know if the $S^-$ "truly" is a conditioned inhibitor because we have not done summation and resistance to reinforcement transfer tests. Our focus is on the magnitude of the difference between the rates of lever pressing in the presence and absence of the $S^-$, that is, on the extent to which it gains control over the subjects' operant behavior.

The atemporal conception of contingency partitions the protocol in Figure 1C into two sorts of trials, the trials when the cue is present and the trials when it is not. The first kind of trial is delimited by the onsets and offsets of the cue, while the second kind is delimited by the offsets and the onsets. The probability of reinforcement is 0 on the trials when the cue is present but very high on trials when it is not. Critically, on the atemporal definition of contingency, the relative durations of the two kinds of trials do not matter.

From a temporal learning perspective, however, the relative durations of the protocoled intervals are critical causal variables. In particular, the average delay to the next reinforcement in Figure 1C is the duration of the $S^-$ plus the average delay to reinforcement when it terminates. If this average delay controls discrimination then the longer the $S^-$, the greater should be the difference between responding in the presence and absence of the $S^-$. Consistent with this hypothesis, Andrzejewski et al. found that the speed with which behavior declined during an inhibitory discriminative stimulus when subjects learned an operant discrimination was directly related to the duration of the cue for nonreinforcement (Andrzejewski, Ryals, Higgins, Sulkowski, Doney, Kelley, & Bersh, 2007).

There is considerable evidence that subjects do in fact learn about time during both excitatory and inhibitory conditioning. Pavlov knew that if a US is presented a fixed delay after CS onset that the early parts of the CS may become inhibitory in what he termed inhibition of delay (Pavlov, 1927). Additionally, if subjects are trained to expect a US at a fixed time and then undergo conditioned inhibition training in which an added cue signals the omission of the US on that trial, subsequent transfer tests show that the inhibitory control of the added cue is greatest at the time the US would otherwise be expected. That is, the added cue is not uniformly inhibitory—it is best at inhibiting the expectation of a US that is learned to occur at the exact time that the added cue had previously signaled the omission of the US (Denniston, Blaisdell, & Miller, 2004; Molet & Miller, 2014). Here we investigate whether temporal knowledge affects the rapidity of acquisition and degree of response suppression by an $S^-$.

Another effect of lengthening the duration of a cue for nonreinforcement is to lower the overall rate of reinforcement in the experimental context, which we denote by $\lambda_C$. If the rate of reinforcement in the absence of the cue, which we denote by $\lambda_{iti}$, remains constant, then the offset of the cue ($S\downarrow$) signals a relatively greater increase in the rate of reinforcement, that is, a greater ratio, $\lambda_{iti}/\lambda_C$, between the rate in the absence of the inhibitory cue and the contextual rate. In Pavlovian procedures the ratio of the rate during the CS and the background rate ($\lambda_{CS}/\lambda_C$), determines the rapidity of excitatory conditioning (Balsam & Gallistel, 2009; Gallistel & Gibbon, 2000) and the strength of the conditioned response. Thus, this effect of the $S^-$ duration on the contextual rate of reinforcement is a second reason to anticipate that lengthening the duration of the $S^-$ while holding constant the rate of reward in its absence may increase the strength of conditioning to the inhibitory cue. Or conversely, holding the duration of the $S^-$ constant while increasing the rate of reward in its absence may increase the strength of conditioning. Please see the discussion for a more in-depth analysis.

The purpose of the experiment was therefore twofold: First, we investigated whether the extent of response suppression was affected by varying the duration of an inhibitory cue. Rewards were earned at unpredictable times following a lever press during the intervals when the inhibitory cue was absent ($S^+$). During the intervals when the auditory inhibitory cue was present ($S^-$), no rewards were available. The duration of the cues was fixed within a group but varied between groups from 20 to 80 seconds.

Second, we investigated whether the extent of response suppression during the ITI was affected by the rate of reinforcement during the ITI. To this end, subgroups of subjects with the same cue duration experienced different rates of reinforcement in the ITI. This allowed us to ask whether and under what assumptions the information-theoretic definition of contingency applied to the results of inhibitory conditioning in the same way as to the results of excitatory conditioning.

## Method

### Subjects

Sixty one male C57/bl6 mice were housed in groups of 4 or 5 in a colony room on a 12:12 hr light:dark cycle. The mice were fed an unlimited amount of chow in their home cages for one hour after experimental sessions, which occurred five days during the week. This food restriction resulted in body weights approximately 85% of ad lib feeding. On weekends the mice received an unlimited supply of food until approximately 18 hr before their Monday session. Water was available ad lib in the home cages at all times. Mice were approximately 90 days old at the start of the first training session and had been handled for one week prior to testing. All experiments and animal care protocols were in accordance with the New York State Psychiatric Institute Institutional Animal Care and Use Committees and Animal Welfare Regulations.

### Apparatus

Eight matching experimental chambers (Med-Asscociates, Inc., St. Albans, VT; model env-307w) equipped with liquid dippers were used in the experiment. Each chamber was located in a light- and sound-attenuating cabinet equipped with an exhaust fan, which provided 72dB background white noise inside the chamber. The internal dimensions of the experimental chamber were 22 × 18 × 13 cm and the floor consisted of metal rods placed 0.87 cm apart. A feeder trough was centered on one wall of the chamber. Inside the trough, an infrared photocell detector (4 mm from trough opening) was used to record head entries into the trough. A reward of one drop (0.01 cc) of evaporated milk could be provided by raising a dipper located inside the feeder trough. The start of each dipper presentation was signaled by a .5s illumination of a light inside the trough. Two retractable levers were mounted on the same wall as the feeder trough, 5 cm away on both sides. A houselight (Med Associates #1820) located at the top of the chamber was illuminated throughout all sessions. An audio speaker was positioned 8.5 cm from the floor on the wall opposite the feeder trough. The speaker delivered a tone (80db, 4kHz) to signal that no rewards were available. A computer running the Med-PC software controlled experimental events and recorded the time during which the events occurred.

### Procedure

All mice received 4 sessions of dipper training, during which 20 reinforcers were delivered with dippers over the course of 20 min. Next all mice were trained to press one of the two levers, half the subjects were trained to press the left lever and half the right. For four sessions every lever press resulted in a reward. Subjects were then trained for 5 sessions on a Random Interval 10s schedule in which the average time from the delivery of a reward until the next one could be earned was 10s. This was followed by 6 sessions on a RI20s schedule and then half of the animals were placed on a RI40s schedule for 9 days while the others continued on the RI20s schedule. Prior to commencement of the experiment, animals were divided into 6 groups so that the average press rate for each group was the same: S-20.RI20, S-40.RI20, S-80.RI20, S-40.RI40, S-80.RI40, and Random Control, $S^0$40.RI40 ($F_{(5,55)} = 0.049$, $p = .9986$). The number following the $S^-$ is the $S^-$ duration and the second part of the group designation is the mean of the random interval (RI) schedule that was in effect in the absence of the $S^-$.

The animals then proceeded to the experiment proper. In the experimental groups, it was presented for 20, 40 or 80 seconds in each of the different groups (as indicated by the group designations). In all 5 experimental groups the RI schedules (either RI20s or RI40s) were in effect in the absence of the tone and independent of the tone, but once the tone ($S^-$) was turned on no rewards could be earned and the timer for the reward schedule was paused until the tone turned off. In the random control group, the tone duration was 40s and the RI40 schedule was in effect throughout the session regardless of whether or not a tone was being presented. In all groups the times between tone presentations were selected from an exponential-like distribution with an average of 40s and ranges from 6s to 120s. In all experimental groups, sessions ended after 20 $S^-$ presentations. To control for potential number of earned rewards, in the random control group, sessions ended after 10 $S^-$ presentations.

Data were analyzed using Python and Matlab. Statistics were computed in Matlab and Prism, and partial omega squared ($\omega^2$) values and 90% confidence intervals for partial eta squared values were computed using a web-based calculator (https://effect-size-calculator.herokuapp.com/).

Data are freely available through Columbia University Academic Commons (https://dx.doi.org/10.7916/d8-vs3f-9f28).

## Results

Our measure of response suppression by the $S^-$ was the ratio of the lever press rate during the $S^-$ to the press rate during the ITI. We normalized this ratio to the first day of $S^-$ presentation to allow comparisons to be made between animals within and across groups. Figure 2 shows how the ratio changed across days for each group, and the means of each group are plotted in Figure 3A.

Increasing the duration of the $S^-$ increased its suppressive effect. The between group differences evident in Figure 2 were confirmed by a 2-way ANOVA ($F_{(5,55)} = 24.28$, $p < .0001$, $\omega^2 = 0.66$, [0.54, 0.74]), as was the effect of days ($F_{(18,990)} = 34.36$, $p < .0001$, $\omega^2 = 0.37$, [0.34, 0.41]), and the interaction between days and groups ($F_{(90,990)} = 3.249$, $p < .0001$, $\omega^2 = 0.16$, [0.12, 0.20]). Specific hypotheses are tested below but to assess whether the groups showed a lower rate of responding during the $S^-$ than during the intertrial intervals, all groups were compared to random group with a Bonferroni correction. We found all groups differed significantly from the random group ($p < .001$), except for the S-20.RI20group ($p = .078$).

The depth of the reduction in the ratio of the response rates was greatest when the $S^-$ was 80s as compared with $S^-$ durations of 40s and 20s (Figure 3B). A two-way repeated measures ANOVA revealed significant main effects of $S^-$ duration ($F_{(2,28)} = 35.09$, $p < .0001$, $\omega^2 = 0.69$, [0.52, 0.79]) and session number ($F_{(18,532)} = 27.62$, $p < .0001$, $\omega^2 = 0.47$, [0.42, 0.51]) and $S^-$ duration X session number interaction ($F_{(36, 504)} = 3.362$, $p < .0001$, $\omega^2 = 0.14$, [0.09, 0.19]). Post hoc Bonferroni test revealed significant pairwise comparisons across learning for each of 3 $S^-$ durations ($p < .005$). This result is not consistent with the atemporal (trial-based) definition of contingency, but it is consistent

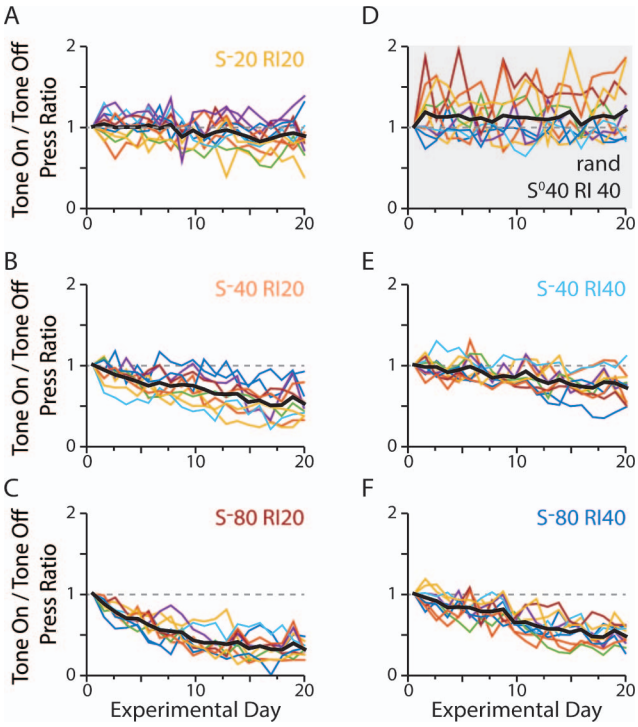


*Figure 2.* Discrimination learning curves for individual subjects. Learning curves for each subject (colors [thin gray lines]) and average (bold black) across subjects within a group are plotted as a function of the ratio of the average press rate during $S^-$ to the rate during the ITI and normalized to the first day for each subject. (A–C) Learning curves for 3 experimental groups when the RI was 20 seconds. (D) No change in behavior in the zero contingency group. (E, F) Learning curves for 2 experimental groups when RI was 40 seconds. See the online article for the color version of this figure.

with the temporal learning perspective on associative learning and with the information-theoretic definition of contingency.

Shortening the average interval between rewards during the ITI (by decreasing the RI) also deepened the suppressive effect of the $S^-$. We determined whether the RI mean contributed to behavioral change. We confirmed that decreasing the mean RI from 40 sec to 20 sec deepened the reduction in the response ratio whether the $S^-$ was 40 sec or 80 sec long. (Figure 3C,D). A two-way repeated measures ANOVA demonstrated a significant main effect of RI for S-40.RI20 versus S-40.RI40 ($F_{(1,\ 18)} = 7.581$, $p = .0131$, $\omega^2 = 0.25$, [0.04, 0.51]); and for S-80.RI20 versus S-80.RI40 ($F_{(1,\ 18)} = 20.93$, $p = .0002$, $\omega^2 = 0.50$, [0.23, 0.68]). Furthermore, this effect is not simply due to greater cumulative exposure to the $S^-$ as we compared the groups that earned rewards on the RI20s schedule at the point in training at which they had all received 4000s of cumulative exposure to the $S^-$. This occurred for Group S-80.RI20 at 5 days; for Group S-40.RI20 at 10 days and for Group S-20.RI20 at 20 days. We examined the last 40 trials of the cumulative exposure and found that indeed the ratio of responding in the $S^-$ to ITI declined with increasing $S^-$ duration using a 1-way ANOVA ($F_{(2,28)} = 4.31$, $p < .03$, $\omega^2 = 0.18$, [0.02, 0.40]).

This pattern of result is consistent with the temporal learning perspective and with the information-theoretic definition of contingency in conditioning. The information about the change in reward of reward rate at the offset of the $S^-$ ($I_{S\downarrow}$) depends on the ratio of the duration of the $S^-$ and the RI (see discussion below). Indeed, we found that equal $I_{S\downarrow}$ values produced equal degrees of suppression. We compared the tone on to tone off press ratios for three $I_{S\downarrow}$ values, 2 (S-20.RI20 and S-40.RI40) and 3 (S-40.RI20 and S-80.RI40) and 5 (S-80.RI20; Figure 3E). Using a 2-way repeated measures ANOVA, we found a significant difference among the three $I_{S\downarrow}$ values ($F_{(3,37)} = 11.4$, $p < .0001$, $\omega^2 = 0.43$, [0.24, 0.59]). To directly test whether keeping $I_{S\downarrow}$ constant would result in equal depths of learning, we could not use $p$ values to support the null hypothesis so we calculated the Bayes Factors associated with the S-RI ratios. We found that when the ratios are kept constant, the learning is similar ($I_{S\downarrow} = 2$: S-20.RI20 vs. S-40.RI40 BF = 1.8:1 in favor of the null hypothesis and $I_{S\downarrow} = 3$: S-40.RI20 vs. S-80.RI40 BF = 8.3:1 in favor of the null hypothesis). The equal effects of equal ratios is a manifestation of the time-scale invariance of the conditioning process (Gibbon & Balsam, 1981; Gallistel & Gibbon, 2000; Balsam & Gallistel, 2009). As expected with a time-scale invariant conditioning process, we observed a linear decrease in the depth of discriminative learning with $I_{S\downarrow}$, as plotted in Figure 3F ($r^2 = 0.815$).

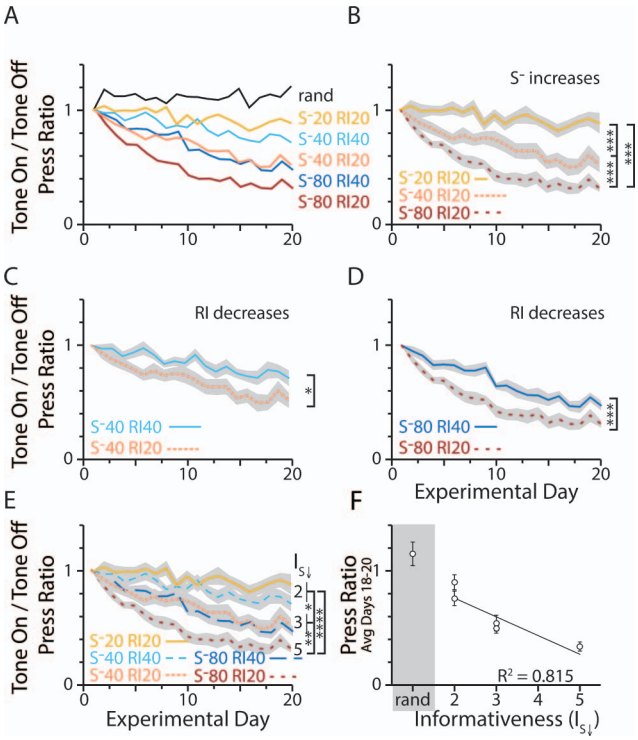


*Figure 3.* Depth of discrimination affected by $S^-$ duration and RI rate. (A) Average learning curves for each of the 6 experimental conditions. (B) As $S^-$ increases from 20 to 40 to 80s, depth of discrimination learning increases. (C, D) When the interval between rewards (RI) decreases from 40s to 20s but the $S^-$ duration remains constant, the depth of discrimination learning increases. (E) Informativeness of the $S^-$ and RI dictates the depth of discriminatory learning. $I_{S\downarrow}$(S-20.RI20) = 2; $I_{S\downarrow}$(S-40.RI40) = 2; $I_{S\downarrow}$(S-40.RI20) = 3; $I_{S\downarrow}$(S-80.RI40) = 3; $I_{S\downarrow}$(S-80.RI20) = 5; (F) Linear relationship between informativeness ($I_{S\downarrow}$) and depth of learning. Gray shading represents *SEM*. See the online article for the color version of this figure.

The behavioral metric that we have used thus far is a ratio of press rates. The decrease in this metric can occur if the numerator decreases (press rate during $S^-$ decreases), the denominator increases (press rate during ITI increases), or a combination of both occurs. To determine which of these scenarios is occurring in our experiment, we analyzed the mean press rates during the $S^-$ and ITI separately. The press rates during the ITI and $S^-$ periods were fit using linear regression and a slope that reflected change across days was obtained for each variable (see Figure 4). During the ITI period, all groups, except random, showed an increase in press rate that was positive and significantly different than zero ($R^2 \geq 0.43$, Figure 4A, C). Interestingly, during the $S^-$ period, only groups with an $I_{S\downarrow}$ of 3 or greater showed a decrease in press rate that was significantly different than zero (Figure 4B, D, $R^2 \geq 0.25$; all other $R^2 \leq 0.06$). To determine whether there was a systematic change in rates of learning across $I_{S\downarrow}$, we plotted the slopes for the $S^-$ and ITI periods for each $I_{S\downarrow}$. We found a linear increase in the slopes of the ITI response rates across training days ($R^2 = 0.36$) and a decrease in the slopes of the $S^-$ response rates across training ($R^2 = 0.89$; Figure 4E) as a function of the $I_{S\downarrow}$.

Finally, we looked at whether the distribution of lever pressing during the $S^-$ changed during learning. We determined the distribution of lever pressing by creating a histogram of all lever presses during $S^-$ across all $S^-$ presentations for Days 1 and 20 of training, normalized for each animal and day and grouped in 5 sec bins (see Figure 5). We found that the greater the $I_{S\downarrow}$, the greater the change in press dynamics during the $S^-$. Two-way repeated measures ANOVAs within each of the experimental groups revealed significant differences in the press count distributions between Days 1 and 20 in all groups except S-20.RI20 (S-20.RI20: $F_{(1, 20)} = 2.0$, $p = .1723$, $\omega^2 = 0.04$, [0, 0.30]; S-40.RI20: $F_{(1, 18)} = 10.66$, $p = .0043$, $\omega^2 = 0.33$, [0.08, 0.56]; S-40.RI40: $F_{(1, 18)} = 30.38$, $p < .0001$, $\omega^2 = 0.59$, [0.34, 0.75]; S-80.RI20: $F_{(1, 18)} = 70.02$, $p < .0001$, $\omega^2 = 0.78$, [0.60, 0.86]; S-80.RI40:$F_{(1, 18)} = 45.94$, $p < .0001$, $\omega^2 = 0.69$, [0.47, 0.81]). There were also significant interactions between days and time within $S^-$ for these four groups (S-40.RI20: $F_{(7, 126)} = 8.168$, $p < .0001$, $\omega^2 = 0.27$, [0.17, 0.38]; S-40.RI40:$F_{(7, 126)} = 3.366$, $p = .0025$, $\omega^2 = 0.11$, [0.035, 0.21]; S-80.RI20: $F_{(15, 270)} = 11.18$, $p < .0001$, $\omega^2 = 0.35$, [0.28, 0.42]; S-80.RI40: $F_{(15, 270)} = 7.911$, $p < .0001$, $\omega^2 = 0.27$, [0.19, 0.34]). To determine the time bin within the $S^-$ for which the response rates were different early versus late in training for each of these groups, we used Sidak's multiple comparisons test for the time bins indicated in Figure 5B, C, E, F (S-40.RI20: bins 3 ($p = .009$), bin 4 ($p < .0001$) and bins 6 and 7 ($p < .02$); S-40.RI40: bin 2 ($p = .0244$), bin 4 ($p = .0001$), and bin 5 ($p = .006$); S-80.RI20: bin 3 ($p = .0011$) and bins 4–16 ($p < .0001$); S-80.RI40: bins 2–8 ($p < .0001$), bin 9 ($p = .0002$), and bin 10 ($p = .001$)).

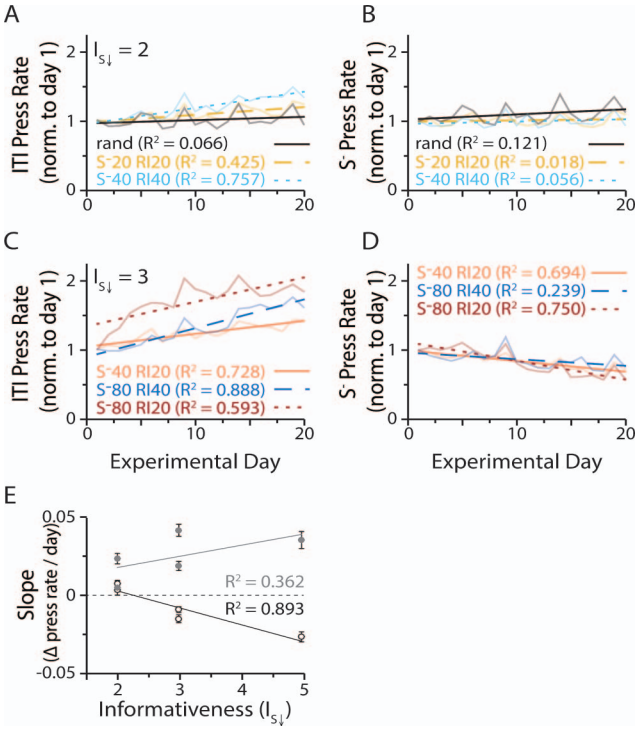


*Figure 4.* Discrimination learning driven by increase in responding during ITI. (A, B) Press rate for the random group and experimental groups with $I_{S\downarrow} = 2$ during ITI (A) and $S^-$ (B) across experimental days fit by linear curve. (C, D) Press rate for experimental groups with $I_{S\downarrow} > 3$ during ITI (C) and $S^-$ (D) across experimental days fit by linear curve. (E) Slopes of linear fits from A–D plotted according to of experimental group. See the online article for the color version of this figure.

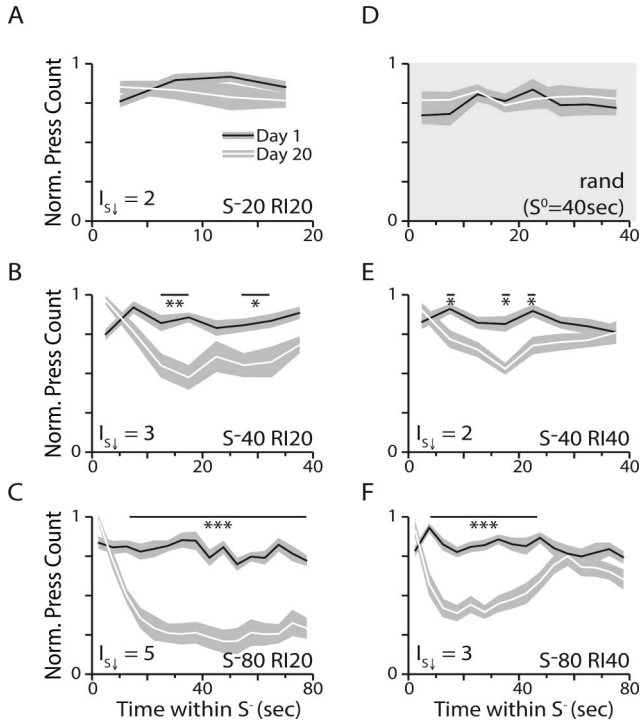


*Figure 5.* Distribution of presses in 5 sec bins on day 1 (black) and day 20 (white) in training for each experimental group. (A) No significant change in normalized press count during $S^-$ observed in S-20.RI20 group. (B, C) Distribution of presses during $S^-$ becomes U shaped in the course of training when RI = 20 sec and $S^-$ is 40 or 80 sec. (D) No change in normalized press count in the zero contingency group. (E, F) Decrease in distribution of presses observed at beginning of $S^-$ when RI = 40 sec and $S^-$ is 40 or 80 sec. Gray shading represents *SEM*.

## Discussion

We hypothesized that increasing the duration of the $S^-$ would result in a greater depth of discrimination. Indeed, we found that when the duration of the $S^-$ increased, the difference between responding in the presence and absence of the $S^-$ was greater, and this was not simply do to different cumulative exposure to the $S^-$. This result is consistent with Andrzejewski et al. who found more rapid discrimination learning with longer $S^-$ durations (Andrzejewski et al., 2007). Taken together the empirical results are consistent with the hypothesis that information about the timing of reward underlies discrimination learning.

The delay to the next reward is dependent on both the duration of the $S^-$ and the rate of reward in the ITI. Therefore, we determined whether the reward rate during the ITI would have an impact on the depth of discrimination. We found that indeed increasing the rate of reward during the ITI enhanced discrimination learning. This is consistent with the idea that it is the contrast between the rate of reward during the ITI ($\lambda_{S\+}$) and the rate of reward in the experimental context ($\lambda_C$) that determines the rapidity and depth of conditioning. This ratio has been described as the informativeness statistic, that is the ratio of the two rates of reinforcement (Balsam et al., 2006; Balsam & Gallistel, 2009).

We defined the depth of discrimination as a ratio of the rate of responding during the $S^-$ to the rate of responding during the ITI. Therefore, if either the rate of responding during the $S^-$ decreases relative to the first day of training or the rate of responding during the ITI increases relative to the first day of training or a combination of the two, the ratio will decrease. We determined that the rate of responding during the ITI increased across sessions in all experimental conditions which is consistent with the idea that the ITI became more excitatory over the course of discrimination learning. However, only in experimental conditions in which the ratios of the $S^-$ duration to RI duration were greater than 1 was there a decrease in pressing during the $S^-$ across sessions. Thus, in groups with ratios equal to 1, animals did not reduce their responses in $S^-$ as might be expected from an inhibitory learning point of view. These results are slightly different than the previous study by Andrzejewski et al. in which they observed a decrease in responding during the CS when the CS:RI ratio was 1 or 0.5 but, similar to this study, they observed a dependence on both the CS duration and the reward rate on the depth of discrimination (Andrzejewski et al., 2007).

In addition to looking at the average responding across sessions, we examined the pattern of responding during $S^-$ presentations. We found that in all groups with an $S^-$ of 40 seconds or longer, the distribution of pressing during the $S^-$ changed throughout training. The flat distribution that was observed early in training often developed into a U-shaped distribution by the end of training, indicating that the subjects anticipated the end of the low-reinforcement state (the fixed duration inhibitory $S^-$). In Group S-80.RI20, approximately half of the animals showed some anticipatory pressing toward the end of the cue but, in the average, this was absent. In subsequent experiments, we have observed variable press distributions during the cue, with some animals having strong anticipatory responses and others not.

In all the groups that showed discrimination learning, the minimum amount of pressing occurred at approximately 20 seconds following onset of $S^-$ indicating that it takes the animals about 20 seconds on average to inhibit their pressing. Interestingly, even in Group S-40.RI40 where the average press rate (presses per second) across the whole $S^-$ did not change across sessions (see Figure 4), the distribution of presses during $S^-$ did change across training (see Figure 5) with a minimum at about 20s by the end of training. Perhaps, in the group with $S^-$ of 20 seconds, and RI of 20 seconds the slow decline in responding may have been offset by the slow increase of responding in anticipation of the end of the $S^-$. Under this hypothesis, a flat distribution of lever pressing throughout training is expected and this is what we observed.

### Informativeness and Information-Theoretic Contingency in Excitatory and Inhibitory Conditioning

Balsam, et al. defined the informativeness of an excitatory Pavlovian CS as the ratio of rate of reinforcement predicted by CS onset ($\lambda_{CS\uparrow}$) to the contextual rate of reinforcement ($\lambda_C$) (Balsam et al., 2006; Balsam & Gallistel, 2009). The rate of reinforcement is the reciprocal of the average wait for reinforcement. Thus, this rate ratio is the factor by which the onset of the CS reduces the expected wait to the next reinforcement.

The natural extension of this definition to inhibitory case is that the informativeness of an inhibitory $S^-$, which we denote by $I_{S\downarrow}$, is the ratio of the rate predicted by $S^-$ offset ($\lambda_{S^+} = 1/RI$) to the contextual rate of reinforcement, which is $1/(RI + d_S)$, where $d_S$ denotes the duration of the $S^-$. After algebraic rearrangement, we have $I_{S\downarrow} = 1 + d_S/RI$. This is in accord with our experimental data results in that there was greater information value due to greater informativeness of the cue when either: i) the duration of the $S^-$ increased ($d_S$) or ii) the reward interval decreased (RI). Furthermore, when $I_{S\downarrow}$, was kept constant, the degree of learning was constant. In other words, our results suggest that inhibitory conditioning, like excitatory conditioning is time-scale invariant (Balsam & Gallistel, 2009; Balsam & Gibbon, 1981; Gallistel & Gibbon, 2000). It is time-scale invariant because it depends on the ratios of the protocol's temporal parameters.

We have previously suggested that contingency in excitatory Pavlovian and operant conditioning be measured by information-theoretic contingency, which is the normalized mutual information (Gallistel & Balsam, 2014; Gallistel et al., 2014). The temporal information conveyed by a cue for reinforcement is the amount of uncertainty cue events dispel about when reinforcement will next occur. The contextual uncertainty about when reinforcement will next occur is—$\log \lambda_C = -\log(n_R/T_C)$, where $\lambda_C$ denotes the contextual rate of reinforcement, $n_R$ denotes the number of reinforcements in that context and $T_C$ denotes the cumulative exposure to that context. (The time intervals, such as cumulative exposure, must be measured in a time unit much shorter than the average interval between reinforcements for this to be true, and the expected wait for reinforcement is, of course, the inverse of the rate. The cue-conditional uncertainty is—$\log \lambda_{CS} = -\log(n_{CS}/T_{CS})$, where $n_{CS}$ denotes the number of reinforcements that have occurred in the state signaled by the CS event and $T_{CS}$ denotes the cumulative duration of that state. In excitatory conditioning, the reinforcement-richer state is signaled by the *onset* of the CS. In our inhibitory conditioning protocol, the reinforcement-richer state is signaled by the *offset* of the CS. We denote the rate of reinforcement following CS offset, that is, the rate during the intertrial intervals, by $\lambda_{ITI}$. The uncertainty about when to expect the next

reinforcement during those intervals is $-\log\lambda_{ITI} = -\log(n_{ITI}/T_{ITI})$, where $n_{ITI}$ denotes the cumulative number of reinforcements during those intervals and $T_{ITI}$, denotes their cumulative duration.

We denote the mutual information between CS offset and reinforcement by $H_{CS\downarrow,R}$. The mutual information between CS offsets (here denoted by CS $\downarrow$) and the reinforcements (here denoted by R) is the contextual or *unconditional* uncertainty, (in this case, –log $\lambda_C$) minus the *conditional* uncertainty (denoted $-\log\lambda_{iti}$). Thus, we have:

$$\begin{aligned} H_{CS\downarrow,R} &= -\log\lambda_C - (-\log\lambda_{CS\downarrow}) \\ &= \log\lambda_{CS\downarrow} - \log\lambda_C \\ &= \log(\lambda_{CS\downarrow}/\lambda_C) \\ &= \log(\lambda_{ITI}/\lambda_C), \end{aligned}$$

from which we see that the mutual information in our inhibitory conditioning protocol is the log of the informativeness of the CS offset.

Contingency is the mutual information normalized by (that is divided by) the available information. The available information is another term for the contextual or unconditional uncertainty, because the greater this uncertainty is, the more there is to be learned from a cue that reduces it. Thus, the information-theoretic measure of temporal contingency is the fraction by which cue events could reduce a subject's uncertainty about the wait for reinforcement. Note that this temporal contingency is an objective quantitative fact about the protocol. It is computed purely from the values of the protocol parameters without any assumptions about processes in a subject's brain.

Ever since Rescorla's classic experiment (Rescorla, 1968), contingency has been recognized as a causal factor in associative learning (e.g., Schultz, 2015), but, as we noted in the introduction, it has resisted objective definition; hence, measurability. By adopting the information-theoretic measure of contingency we make it a measurable property of both Pavlovian and operant conditioning protocols (Gallistel et al., 2014).

## The Explanatory Power of Contingency

From their inception (Gibbon, 1977) to their most recent instantiations, timing theories have assumed that brains form metric (Guilhardi et al., 2007; Wilkes & Gallistel, 2016, 2017)—or at least ordinal (Arcediano, Escobar, & Miller, 2003; Savastano & Miller, 1998)—temporal maps of the distribution of events in time (Balsam & Gallistel, 2009). They assume that brains store interval durations in a form to which arithmetic operations apply (at the very least, ordination). Adding to this assumption, the assumption that temporal contingency drives conditioning extends timing theory in the directions called for by Luzardo et al. (2017). They stress the desirability of integrating theories that explain three different aspects of conditioning: the timing of the conditioned response, the acquisition and extinction of the conditioned response, and cue competition (e.g., blocking).

As just explained, temporal contingency is normalized mutual information. In Pavlovian conditioning, the mutual information between a CS event (onset or offset) is the log the informativeness, which is the ratio of the rate during the state cued by the event to the contextual rate. A striking quantitative property of acquisition, whose theoretical importance has long been recognized (Gibbon & Balsam, 1981; Rescorla, 1988), is that trials to acquisition is proportional to the ratio between the CS-US interval and the US-US interval. This fact implies a fully metric representation. It also suggests a simple, parameter free quantitative acquisition law: Acquisition depends on the product of the informativeness of the CS event (whether onset or offset) and the cumulative number of reinforcements in the state signaled by that event. When this product exceeds a subject-specific decision criterion, the conditioned response abruptly appears (Gallistel, Fairhurst, & Balsam, 2004).

Finally, the hypothesis that temporal contingency drives conditioning can be applied to other phenomena such as partial reinforcement, extinction, and cue competition phenomena (Balsam & Gallistel, 2009; Gallistel et al., 2014; Gallistel, 2012). With respect to contingency learning, when, the rate following CS onset is the same as the contextual rate, as in Rescorla's (1968) seminal experiment, the ratio of the two rates is 1 and the log of that ratio, which is the mutual information between CS onset and reinforcement, is 0, so the contingency is 0. Similarly, in blocking experiments, (e.g., Kamin, 1967, 1969) one cue is conditioned and, then superimposed on a second cue. In that protocol, the already conditioned cue is the context and the superimposed cue does not alter the rate of reinforcement in that context. Therefore, reinforcement is not contingent on the second cue, which is why a conditioned response to that cue does not develop.

We conclude that Rescorla was correct when he concluded that contingency drives conditioning. Here we show how an objective, measurable temporally based calculation of contingency can be applied to inhibitory conditioning. This frames a challenge for neutrally oriented theories of conditioning and reinforcement learning. They must provide mechanisms for the neural encoding and representation of duration and number that make it possible for brains to compute temporal contingency, the experiential variable that drives conditioning.

## References

Allan, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society, 15,* 147–149. http://dx.doi.org/10.3758/BF03334492

Allan, L. G. (1993). Human contingency judgments: Rule based or associative? *Psychological Bulletin, 114,* 435–448. http://dx.doi.org/10.1037/0033-2909.114.3.435

Allan, L. G., Hannah, S. D., Crump, M. J., & Siegel, S. (2008). The psychophysics of contingency assessment. *Journal of Experimental Psychology: General, 137,* 226–243. http://dx.doi.org/10.1037/0096-3445.137.2.226

Allan, L. G., & Jenkins, H. M. (1980). The judgment of contingency and the nature of the response alternatives. *Canadian Journal of Psychology/Revue. Canadian Psychology, 34,* 1–11. http://dx.doi.org/10.1037/h0081013

Allan, L. G., Siegel, S., & Tangen, J. M. (2005). A signal detection analysis of contingency data. *Learning & Behavior, 33,* 250–263. http://dx.doi.org/10.3758/BF03196067

Andrzejewski, M. E., Ryals, C. D., Higgins, S., Sulkowski, J., Doney, J., Kelley, A. E., & Bersh, P. J. (2007). Is extinction the hallmark of operant discrimination? Reinforcement and S(Delta) effects. *Behavioural Processes, 74,* 49–63. http://dx.doi.org/10.1016/j.beproc.2006.09.010

Arcediano, F., Escobar, M., & Miller, R. R. (2003). Temporal integration and temporal backward associations in human and nonhuman subjects.

*Learning & Behavior, 31,* 242–256. http://dx.doi.org/10.3758/BF03195986

Balsam, P. D. (1988). Selection, representation and equivalence of controlling stimuli. In R. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Handbook of experimental psychology* (Vol. 2, pp. 111–166). New York, NY: Wiley.

Balsam, P. D., Drew, M. R., & Gallistel, C. R. (2010). Time and Associative Learning. *Comparative Cognition & Behavior Reviews, 5,* 1–22. http://dx.doi.org/10.3819/ccbr.2010.50001

Balsam, P. D., Fairhurst, S., & Gallistel, C. R. (2006). Pavlovian contingencies and temporal information. *Journal of Experimental Psychology: Animal Behavior Processes, 32,* 284–294. http://dx.doi.org/10.1037/0097-7403.32.3.284

Balsam, P. D., & Gallistel, C. R. (2009). Temporal maps and informativeness in associative learning. *Trends in Neurosciences, 32,* 73–78. http://dx.doi.org/10.1016/j.tins.2008.10.004

Balsam, P. D., & Gibbon, J. (1981). Spreading associations in time. In C. M. Locurto, H. S. Terrace, & J. Gibbon (Eds.), *Autoshaping and conditioning theory* (pp. 219–253). New York, NY: Academic.

Daw, N. D., Courville, A. C., & Touretzky, D. S. (2006). Representation and timing in theories of the dopamine system. *Neural Computation, 18,* 1637–1677. http://dx.doi.org/10.1162/neco.2006.18.7.1637

Denniston, J. C., Blaisdell, A. P., & Miller, R. R. (2004). Temporal coding in conditioned inhibition: Analysis of associative structure of inhibition. *Journal of Experimental Psychology: Animal Behavior Processes, 30,* 190–202. http://dx.doi.org/10.1037/0097-7403.30.3.190

Ditz, H. M., & Nieder, A. (2016). Sensory and working memory representations of small and large numerosities in the crow endbrain. *The Journal of Neuroscience, 36,* 12044–12052. http://dx.doi.org/10.1523/JNEUROSCI.1521-16.2016

Eichenbaum, H. (2014). Time cells in the hippocampus: A new dimension for mapping memories. *Nature Reviews Neuroscience, 15,* 732–744. http://dx.doi.org/10.1038/nrn3827

Gallistel, C. R. (2012). Extinction from a rationalist perspective. *Behavioural Processes, 90,* 66–80. http://dx.doi.org/10.1016/j.beproc.2012.02.008

Gallistel, C. R., & Balsam, P. D. (2014). Time to rethink the neural mechanisms of learning and memory. *Neurobiology of Learning and Memory, 108,* 136–144. http://dx.doi.org/10.1016/j.nlm.2013.11.019

Gallistel, C. R., Craig, A. R., & Shahan, T. A. (2014). Temporal contingency. *Behavioural Processes, 101,* 89–96. http://dx.doi.org/10.1016/j.beproc.2013.08.012

Gallistel, C. R., Fairhurst, S., & Balsam, P. (2004). The learning curve: Implications of a quantitative analysis. *Proceedings of the National Academy of Sciences of the United States of America, 101,* 13124–13131. http://dx.doi.org/10.1073/pnas.0404965101

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review, 107,* 289–344. http://dx.doi.org/10.1037/0033-295X.107.2.289

Gershman, S. J., Moustafa, A. A., & Ludvig, E. A. (2014). Time representation in reinforcement learning models of the basal ganglia. *Frontiers in Computational Neuroscience, 7,* 194. http://dx.doi.org/10.3389/fncom.2013.00194

Gibbon, J. (1977). Scalar expectancy theory and Weber's Law in animal timing. *Psychological Review, 84,* 279–325. http://dx.doi.org/10.1037/0033-295X.84.3.279

Gibbon, J., & Balsam, P. D. (1981). Spreading associations in time. In C. M. Locurto, H. S. Terrace, & J. Gibbon (Eds.), *Autoshaping and conditioning theory* (pp. 219–253). New York, NY: Academic.

Guilhardi, P., Yi, L., & Church, R. M. (2007). A modular theory of learning and performance. *Psychonomic Bulletin & Review, 14,* 543–559. http://dx.doi.org/10.3758/BF03196805

Jackson, S. A. W., Horst, N. K., Pears, A., Robbins, T. W., & Roberts, A. C. (2016). Role of the perigenual anterior cingulate and orbitofrontal cortex in contingency learning in the marmoset. *Cerebral Cortex, 26,* 3273–3284. http://dx.doi.org/10.1093/cercor/bhw067

Kamin, L. J. (1967). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: Aversive stimulation* (pp. 9–33). Miami, FL: University Miami Press.

Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 276–296). New York, NY: Appleton-Century-Crofts.

Kirkpatrick, K. (2014). Interactions of timing and prediction error learning. *Behav Processes, 101,* 135–145. http://dx.doi.org/10.1016/j.beproc.2013.08.005

Kutter, E. F., Bostroem, J., Elger, C. E., Mormann, F., & Nieder, A. (2018). Single neurons in the human brain encode numbers. *Neuron, 100,* 753–761.e4. http://dx.doi.org/10.1016/j.neuron.2018.08.036

Ludvig, E. A., Sutton, R. S., & Kehoe, E. J. (2008). Stimulus representation and the timing of reward-prediction errors in models of the dopamine system. *Neural Computation, 20,* 3034–3054. http://dx.doi.org/10.1162/neco.2008.11-07-654

Luzardo, A., Alonso, E., & Mondragón, E. (2017). A Rescorla-Wagner drift-diffusion model of conditioning and timing. *PLoS Computational Biology, 13*(11), e1005796. http://dx.doi.org/10.1371/journal.pcbi.1005796

Molet, M., & Miller, R. R. (2014). Timing: An attribute of associative learning. *Behavioural Processes, 101,* 4–14. http://dx.doi.org/10.1016/j.beproc.2013.05.015

Nakahara, H., & Kaveri, S. (2010). Internal-time temporal difference model for neural value-based decision making. *Neural Computation, 22,* 3062–3106. http://dx.doi.org/10.1162/NECO_a_00049

Noonan, M. P., Chau, B. K. H., Rushworth, M. F. S., & Fellows, L. K. (2017). Contrasting Effects of Medial and Lateral Orbitofrontal Cortex Lesions on Credit Assignment and Decision-Making in Humans. *The Journal of Neuroscience, 37,* 7023–7035. http://dx.doi.org/10.1523/JNEUROSCI.0692-17.2017

Pavlov, I. P. (1927). *Conditioned reflexes*. London, England: Oxford University Press.

Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology, 66,* 1–5. http://dx.doi.org/10.1037/h0025984

Rescorla, R. A. (1969). Conditioned inhibition of fear resulting from negative CS-U.S. contingencies. *Journal of Comparative and Physiological Psychology, 67,* 504–509. http://dx.doi.org/10.1037/h0027313

Rescorla, R. A. (1988). Pavlovian conditioning. It's not what you think it is. *American Psychologist, 43,* 151–160. http://dx.doi.org/10.1037/0003-066X.43.3.151

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, NY: Appleton-Century-Crofts.

Rivest, F., Kalaska, J. F., & Bengio, Y. (2010). Alternative time representation in dopamine models. *Journal of Computational Neuroscience, 28,* 107–130. http://dx.doi.org/10.1007/s10827-009-0191-1

Savastano, H. I., & Miller, R. R. (1998). Time as content in Pavlovian conditioning. *Behavioural Processes, 44,* 147–162. http://dx.doi.org/10.1016/S0376-6357(98)00046-1

Schultz, W. (2015). Neuronal reward and decision signals: From theories to data. *Physiological Reviews, 95,* 853–951. http://dx.doi.org/10.1152/physrev.00023.2014

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275,* 1593–1599. http://dx.doi.org/10.1126/science.275.5306.1593

Simen, P., Balci, F., de Souza, L., Cohen, J. D., & Holmes, P. (2011). A model of interval timing by neural integration. *The Journal of Neuro-*

*science, 31,* 9238–9253. http://dx.doi.org/10.1523/JNEUROSCI.3121-10.2011

Simen, P., Rivest, F., Ludvig, E. A., Balci, F., & Killeen, P. (2013). Timescale Invariance in the Pacemaker-Accumulator Family of Timing Models. *Timing & Time Perception, 1,* 159–188. http://dx.doi.org/10.1163/22134468-00002018

Suri, R. E., & Schultz, W. (1999). A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. *Neuroscience, 91,* 871–890. http://dx.doi.org/10.1016/S0306-4522(98)00697-6

Wagner, A. R., Logan, F. A., & Haberlandt, K. (1968). Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology, 76,* 171–180. http://dx.doi.org/10.1037/h0025414

Wilkes, J. T., & Gallistel, C. R. (2016). *Information theory, memory, prediction, and timing in associative learning* (Unpublished manuscript). Retrieved from https://osf.io/y9sg2

Wilkes, J. T., & Gallistel, C. R. (2017). Information theory, memory, prediction, and timing in associative learning. In A. Moustafa (Ed.), *Computational Models of Brain and Behavior* (pp. 481–492). New York, NY: Wiley/Blackwell. http://dx.doi.org/10.1002/9781119159193.ch35

Received November 13, 2018
Revision received February 5, 2019
Accepted February 5, 2019 ■