

Number and time in acquisition, extinction and recovery

C. R. Gallistel and E. B. Papachristos

Rutgers University

We measured rate of acquisition, trials to extinction, cumulative responses in extinction, and the spontaneous recovery of anticipatory hopper poking in a Pavlovian protocol with mouse subjects. We varied by factors of 4 number of sessions, trials per session, intersession interval, and span of training (number of days over which training extended). We find that different variables affect each measure: Rate of acquisition [$1/(\text{trials to acquisition})$] is faster when there are fewer trials per session. Terminal rate of responding is faster when there are more total training trials. Trials to extinction and amount of responding during extinction are unaffected by these variables. The number of training trials has no effect on recovery in a 4-trial probe session 21 days after extinction. However, recovery is greater when the span of training is greater, regardless of how many sessions there are within that span. Our results and those of others suggest that the numbers and durations and spacings of longer-duration “episodes” in a conditioning protocol (sessions and the spans in days of training and extinction) are important variables and that different variables affect different aspects of subjects’ behavior. We discuss the theoretical and clinical implications of these and related findings and conclusions—for theories of conditioning and for neuroscience.

Most accounts of spontaneous recovery view the time after extinction as a period during which changes in the strengths of excitatory and inhibitory associations occur. McConnell and Miller (2014) point out that the treatment of extinction in major associative theories of learning fall into two categories, those that posit the degradation of the excitatory association and those that posit the development of inhibitory associations. Perhaps the most common assumption regarding spontaneous recovery is that the inhibitory associations that develop during extinction fade faster than the excitatory associations (Hull, 1943; Pavlov, 1927; Rescorla, 1979, 1993; Wagner, 1981).

Other theories focus not on the postulated effects of the passage of time on associations

but rather on its effect on the performance function that determines the behavioral expression of associative strength. For example, the activation threshold for the extinction memory rises with time after extinction, making it harder to retrieve (Kraemer & Spear, 1993); or, the temporal context of extinction changes, which in turn results in the renewal of the acquisition memory at the time of testing (Bouton, 1991, 1993); or, the conditioned and extinguished CS elements are redistributed, so that the extinguished elements from extinction are outnumbered by the still conditioned elements formed from the lengthier acquisition training (Estes, 1955; Estes & Burke, 1953).

These explanations for recovery have two things in common. First, time and number are treated as the *media* for alterations in associative processes or their expression, not as crucial parts of the *content* of what is learned (Savastano & Miller, 1998). In these theories, the subject does not remember the durations and numerosities—the intertrial intervals, the intersession intervals, the span of days within which training occurs, the numbers of trials in a session, the numbers of reinforced trials in a session, the numbers of sessions, etc. Second, the effects of time and number are assumed to be mediated by the intrinsic time course of processes of association formation, decay and expression. The associations do not encode experiential facts (trial durations, session

The work reported here comprised a portion of a dissertation submitted to Rutgers University in partial fulfillment of the requirements for the doctoral degree. EBP designed and did the experiments, did the initial data analyses, and wrote a draft, with some advice from CRG. CRG did several further analyses and wrote the final draft.

The research was supported by National Institutes of Health Grant R21 MH 63866 to Charles R. Gallistel, and Fulbright, Alexander S. Onassis Public Benefit Foundation, and Gerondelis Foundation scholarships to Efstathios B. Papachristos. The authors gratefully acknowledge very helpful suggestions and references to relevant previous work made by the Editor and two anonymous referees.

Address correspondence to: C.R. Gallistel, 252 7th Ave. 10D, New York, NY 10001. Email: galliste@rucss.rutgers.edu

doi: 10.1002/jeab.571

durations, numbers of trials, numbers of reinforced trials, etc.). Associations, both historically in philosophy and psychology, and currently in their neuroscientific interpretation, are conductive pathways whose strength (a.k.a. conductance) varies depending on the temporal pairing of events. Because associative and synaptic-plasticity accounts of associative learning do not assume that associations encode facts, they do not specify a code (Gallistel, 2017).

Experimental work over the last several decades has led to two conclusions that are at variance with these assumptions: First, in associative conditioning, time and number are *messages*, not *media*. The numerosities of the events in a protocol and the intervals between them are among the contents of memory. Second, many aspects of the conditioned behavior depend on derived quantities, such as rate (number/duration) and probability (number/number) and the C/T ratio (the ratio between the average US–US interval and the average CS–US interval in Pavlovian delay conditioning). We term these *derived* quantities because they must be computed from directly and separately measured counts and durations. Brains appear to do arithmetic computations on encoded abstract quantities like number and duration, or, at least, subjects behave as if their brains did such computations.

Time in Conditioning

Subjects in associative learning experiments make a temporal map of the conditioning experience (Balsam & Gallistel, 2009; Honig, 1981; Taylora, Joseph, Zhaoc, & Balsam, 2014). The temporal information in the map affects every aspect of Pavlovian and operantly conditioned behavior (Arcediano & Miller, 2002; Barnet, Grahame, & Miller, 1993; Barnet & Miller, 1996; Blaisdell, Denniston, & Miller, 1998; Burger, Denniston, & Miller, 2001; Cole, Barnet, & Miller, 1995; Cunningham & Shahan, 2018; Denniston, Blaisdell, & Miller, 1998, 2004; Gallistel, Craig, & Shahan, 2019; Gür, Duyan, & Balci, 2017; Shahan & Cunningham, 2015; Theunissen & Miller, 1995; Thraillkill & Shahan, 2014).

That subjects learn the distribution of wait times in conditioning experiments is revealed by the fact that their behavior during the

conditioned stimulus (CS for short) varies depending on whether the delay of reinforcement during the CS is fixed or exponential. When it is fixed, the onset of responding is abrupt on any given trial. The distribution of onsets is centered approximately half way through a trial, regardless of CS duration, that is, the mean of the onset distribution scales with CS duration and the right edge is close to the anticipated reinforcement time (Church, Meck, & Gibbon, 1994; Gallistel, King, & McDonald, 2004; Gibbon, 1977). Moreover, in the peak procedure, where “probe” trials are not reinforced and the CS continues for three or four times the anticipated reinforcement latency, subjects stop abruptly soon after that moment has passed. The distribution of the stops is almost perfectly normal, and it is narrow (coefficient of variation is approximately 0.16, Balci, Allen, et al., 2009; Balci et al., 2011; Church et al., 1994; Gallistel et al., 2004). On the other hand, when the distribution of wait times during CSs is exponential, response rate during the CS reflects the flat hazard function unique to the exponential (Libby & Church, 1974, 1975).

Subjects appear to do computations with the wait times they have learned (Gür et al., 2017): For example, the rate of learning [$1/(\text{trials to acquisition})$] is a scalar function of the C/T ratio, which, as already noted, is the ratio of the basal average wait time for reinforcement in the training context (C) and the average wait time (T) for reinforcement following the onset of the conditioned stimulus (CS for short). The larger the C/T ratio is, the fewer the number of reinforced trials required for the appearance of a conditioned response to the CS (Gallistel & Gibbon, 2000; Gibbon & Balsam, 1981; Gottlieb, 2008; Jenkins, Barnes, & Barrera, 1981; Sunsay & Bouton, 2008). The claim that this implies an arithmetic computation over encoded abstract quantities is, of course, both strong and controversial. In the Discussion, we argue that the ball is in the court of those who are uncomfortable with it.

Numerosity in Conditioning

There is also an extensive literature showing that the subjects in conditioning experiments learn the numerosities (Anobile, Cicchini, &

Burr, 2015; Davison & Cowie, 2019; Gallistel & Gelman, 1990; Gallistel, 1990; Geary, Berch, & Koepke, 2015; Kutter et al., 2018). Evidence that subjects compute with the numerosities they have learned comes from studies of the partial reinforcement extinction effect. The probability of reinforcement during training has a scalar effect on trials to extinction. If *on average* only 1 in n trials is reinforced during training, it takes nk trials to reach any given extinction criterion, where k is a constant that depends on the criterion (Chan & Harris, 2019; Gibbon, Farrell, Locurto, Duncan, & Terrace, 1980).

The probability of reinforcement is the ratio of the number of reinforced trials to the total number of trials. The trial types whose relative numerosity determines the probability of reinforcement are widely spaced episodes (Crystal & Smith, 2014; Kheifets, Freestone, & Gallistel, 2017). Moreover, reinforcement events have been shown to be but one among the types of events that can distinguish between trials/episodes in determining behaviorally important probabilities. In the switch protocol, trials are typed by their durations (Balci, Freestone, & Gallistel, 2009; Fetterman & Killeen, 1995; Kheifets et al., 2017; Kheifets & Gallistel, 2012). On some fraction of the trials, reinforcement is obtained after a short interval by poking into the “short” hopper, while on the complementary fraction, it is obtained by poking into the “long” hopper. At the beginning of each trial, subjects do not know which type of trial the computer has chosen. They learn to do the sensible thing: They go to the short hopper at the beginning of every trial and they switch their poking to the long hopper on those trials when poking in the short hopper fails to deliver reinforcement at the anticipated short latency specific to that hopper. When they depart too soon from the short hopper on a short trial type, they do not get a reinforcer; when they depart too late on a long trial type (after the long reinforcement latency has elapsed), they also do not get reinforcement.

When there is a 2- or 3-fold ratio between the reinforcement latencies, subjects position the distribution of their departure times approximately optimally between the two temporal goal posts (between the short and long reinforcement latencies). Therefore, they get reinforcement on almost every trial. The position of the distribution of their departure

latencies is, however, sensitive to the complementary probabilities of the short and long trial types. When the short probability is high and the long low, they shift their distribution away from the short latency hopper. When the reverse probabilities obtain, they shift their distribution in the opposite direction. When the probabilities abruptly change, subjects abruptly change their distribution. On a substantial percentage of the change occasions, subjects shift their departure distribution before they have missed a single reinforcement (Kheifets & Gallistel, 2012).

Content-based Theories of Learning

The discovery of the rich mnemonic contents produced by conditioning protocols and the long-standing evidence for the importance of derived ratio variables (rates, probabilities, the C/T ratio) has stimulated the development of nonassociative content-based theories of learning (Gallistel et al., 2019; Gallistel & Wilkes, 2016; Gallistel, 1990, 2012; Gallistel & Gibbon, 2000; Gibbon, 1977; Wilkes & Gallistel, 2017). In these theories, learning has two components, the second of which presupposes the first. First, there is the encoding into memory of the sensory properties (e.g., texture and color) and first-order nonsensory properties (e.g., duration and numerosity) of hierarchically structured episodes (Gallistel, 2017). Second comes the computation of stochastic models (Gallistel & Wilkes, 2016) based on these raw data. The stochastic models, which are themselves stored in memory, have two functions: 1) they enable more efficient coding of the data on which they are based; 2) they enable the predictions underlying the anticipation of future episodes.

It seems likely that some of the second component—the computation of stochastic models—occurs off line. Consolidation and reconsolidation phenomena are plausibly considered manifestations of off-line stochastic model computation, because stochastic model development may lead to recoding memories so as to reduce the amount of memory required to preserve the same data (Dudai, 2012; Wang & Morris, 2010). The computation of stochastic models may also be the computational explanation for the replay of episodes during sleep and quiet wakefulness (Foster & Wilson, 2006; Jafarpour, Fuentemilla, Horner, Penny, & Duze, 2014; Mattar & Daw, 2018;

Ólafsdóttir, Bush, & Barry, 2018; Panoz-Brown et al., 2018; Zentall, 2019).

As suggested by the preceding brief and very incomplete review, most of the literature that has shown that subjects in conditioning experiments learn durations and numerosities has focused on the effects of trial parameters: trial duration, intertrial interval, number of trials and number of reinforced trials. In the experiments we now report, we looked for effects at higher levels of the hierarchically structured episodes that conditioning protocols present to subjects. Trials are episodes. Embedded within them are events such as the onset and offset of the CS and the reinforcements. Sessions are episodes. Embedded within them are the trials. The days over which training occurs constitute a lengthy “episode” (one might prefer the term *epoch*). Embedded within the training epoch are the sessions during which reinforced trials occurred. Embedded within an extinction session (or epoch) are trials on which reinforcements did not occur. In some extinction protocols, several extinction sessions are embedded within the extinction epoch. In our experiments, we ask whether the numerical and durational parameters of these higher-level chunks of experience affect acquisition, extinction and recovery.

We examined the effect on the post-extinction recovery of conditioned nose-poking in the mouse of four numerical and temporal training variables: total number of training trials, number of training sessions,

trials per session, and the span in days over which the sessions were spread. The variations in these parameters of acquisition training are shown in Table 1.

The intertrial intervals in all four experiments were drawn from an exponential distribution with a mean of 180 s, to which a 10-s interval was added, so that there was no intertrial interval shorter than 10 s. Thus, the average intertrial interval was 190 s, and it did not vary between experimental groups.

As may be seen in Table 1, the number of trials in a session varied from as few as two to as many as 40. Because the average intertrial interval was the same for every group, session duration scaled with trials-per-session.

We trained with a Pavlovian protocol in which reinforcement was the delivery of a food pellet at the termination of the 10-s white noise CS. Our index of conditioned responding was the *elevation score*, the difference in the number of pokes during the 10-s CS and the number during the 10-s interval immediately preceding the onset of the CS.

Extinction occurred during a single session the day after the last session of acquisition training. Our first measure of performance during extinction was trials to extinction; five CS presentations without a response terminated the session after five additional trials, thereby yielding the trials to extinction measure. Our second measure was the cumulative elevation score during the extinction session. If, as sometimes happened, the mouse made more pokes during the 10-s pre-CS intervals than during the CS intervals, this measure could be negative.

Table 1
Training parameters

Group	# Sessions	Trials/Session	Span (days)	Total Trials
Experiment 1				
Group 1.1 (n = 6)	28	2 or 3, mean = 2.5	28	70
Group 1.2 (n = 6)	7	40	7	280
Experiment 2				
Group 2.1 (n = 6)	7	40	7	280
Group 2.2 (n = 6)	28	10	28	280
Group 2.3 (n = 6)	7	40	28	280
Experiment 3				
Group 3.1 (n = 6)	24	10	24	240
Group 3.2 (n = 6)	6	40	6	240
Experiment 4				
Group 4.1 (n = 6)	8	10	8	80
Group 4.2 (n = 6)	8	40	8	320

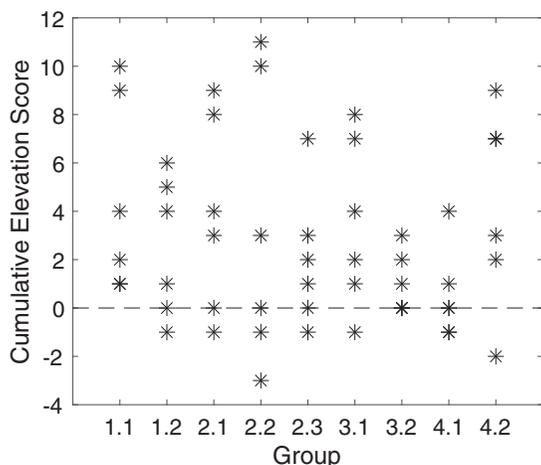


Fig. 1. Cumulative elevation scores by group from the first probe for recovery (Day 7 postextinction, except for 1.1 and 1.2, where it was Day 3).

A pilot experiment showed negligible recovery of the elevation score in a four-trial session on the 7th day postextinction, but substantial recovery in a second four-trial session on the 21st day postextinction. Therefore, each of the groups had a four-trial recovery session with no reinforcement at a “short” postextinction lapse and again at a “long” lapse. For the groups in Experiment 1, the probe days were Day 3 and Day 18. For the seven groups in the last three experiments, the short probe for recovery was on Day 7 postextinction and the long probe on Day 21. The short probes always yielded little or no recovery (Fig. 1), replicating the pilot result. Therefore, we focus our analysis on the results from the long probes, which always yielded significant recovery.

Method

Subjects

The subjects were male C57Bl/6 mice obtained from Harlan (Indianapolis, IN). They were about 9-11 weeks old and weighed between 16.3 and 20.9 g when the experiments started. They were housed individually in plastic tubs, and maintained on a 12:12 hr photoperiod, with lights on at 22:00 hr. Behavioral testing occurred during the dark phase of the photoperiod. Water was available ad lib in both the home cage and the experimental chambers, while food was restricted to keep body weight at approximately 85% of free-feeding weight. Standard rodent chow was given at the end of each

session. Mice remained on their deprivation schedule until the first spontaneous recovery test, after which they received unrestricted food until 4 days prior to the second test, when we returned them to the deprivation schedule.

Apparatus

Experimental sessions took place in modular operant chambers (Med Associates, Georgia, VT, model # ENV307W) measuring 21.6 cm x 17.8 cm x 12.7 cm, housed in individual ventilated, sound-attenuating boxes. Each chamber was equipped with a pellet dispenser connected to a feeding station on the center of one side. The station was a cubic hopper, 24 mm on a side, equipped with an infrared (IR) beam that detected nose pokes and a 5-watt light that illuminated the hopper when turned on. Mounted on the opposite wall were a clicker generator (80 dB, 10 Hz), a white noise generator (80 dB, flat 10-25,000 Hz), and a house light (28 V DC, 100 mA). At the end of the feeding latency (10 s) a 20-mg precision pellet (TestDiet, 5TUM 1811143) was delivered in the feeding station. The experiment was controlled by computer software (Med-PC IV, Med Associates) that also logged and time-stamped the events—the onsets and offsets of interruptions of the IR beams in the station, the onsets and offsets of white noise, and the delivery of food pellets. Event times were recorded with a resolution of 20 ms.

Procedure

Body weights were recorded right before the start of each session. The house light remained illuminated throughout the experiment.

Acquisition. Sessions started with an inter-trial interval (ITI) drawn from an exponential distribution with a mean of 180 s. This ITI was followed by a fixed, unsignaled 10-s interval (pre-CS period), at the end of which a trial started (10 s white noise terminating in pellet release). A 100-ms clicker signaled pellet delivery.

Extinction. The day after their last acquisition session all mice received a single extinction session. There were no pellets or clickers delivered at the end of the white noise. In every other aspect, the extinction session was

identical to an acquisition session. After the first 20 trials, an extinction criterion was employed. A mouse should make no responses during the CS for five consecutive trials. The session ended five trials after this criterion was met.

Spontaneous recovery tests. All mice were tested for spontaneous recovery 1 and 3 weeks postextinction, except in Experiment 1, where they were tested at 3 and 18 days post extinction. Each test included four presentations of the white noise in the absence of a reward.

Statistical Analyses

Estimating the trial at which consistent anticipatory responding begins. This was taken to be the point in the cumulative record (the cumsum of the elevation scores) at which the cumulative record permanently exceeded 20 presses above its minimum.

Statistical comparisons. We did two tests for each comparison, a two-tailed t test followed by a Bayesian alternative. In the Bayesian test, the normalized likelihood function for the mean of the “control” group is the null prior. It represents the hypothesis that the mean of the “experimental” does not differ from the mean of the “control” group. A second prior distribution, called the alternative prior, represents the hypothesis that the mean of the “experimental” group may differ from that of the “control” group. It specifies the range of deviations that might reasonably be expected. In the Bayesian equivalent to the *two-tailed t* test, the alternative prior distribution spreads out beyond both tails of the null prior.

The more widely the alternative prior spreads, the more the resulting BF will favor the null hypothesis. Thus, there is always the question what a reasonable spread is. If the data are numerous (say, more than 20) on both sides of a comparison and the distributions do not overlap, there is no point in doing a statistical test. Regardless of what test one uses, the probability that the two distributions have the same mean will be infinitesimal. Therefore, the fact that one thinks it appropriate to do a statistical test implies that the data on at least one side of a comparison are not numerous (e.g., $n = 6$) and/or the distributions overlap. In the light of these considerations, we chose as the prior for our alternative to the null a flat distribution that

extended to 2σ to either side of the mean of the null prior, where the value for sigma was the pooled standard deviation. Figure 2 shows the graphs of the null prior, alternative prior and likelihood function produced by the BF2 command applied to three of the comparisons. The question of which group was the “control” and which the “experimental” was moot in these experiments, but the results from BF2.m are the same regardless of which group is assigned which role (Gallistel, 2009).

The raw data and the Matlab™ code that analyzed them are in a publicly accessible repository (<https://github.com/CRGallistel/TimeNumberAcqExtRecov>). Included in the *DataFiles* folder in the repository are files *LongTable.mat* and *LongTable.csv*. They contain a long table with the subject-by-subject summary statistics that entered into the plots and statistical comparisons in this report. The Matlab™ code that generated the figures and the statistical comparisons, using data in that long table, is in the file *CRGcode.m* in the *MatlabCode* folder in that repository. The Matlab™ commands for the statistical comparisons and the results they produced may be accessed in the *LT.Properties.UserData* of the *LT* table variable in the *LongTable.mat* file.

Results

Acquisition

Figure 3 plots the cumulative records (cumsum) of the trial-by-trial elevation scores (number of responses during the CS minus number during 10 s pre-CS interval) throughout the procedure (acquisition, extinction, and recovery) for the 12 subjects in Experiment 1. The top six panels are for the subjects that received either two or three (on average 2.5) trials per session for 28 daily sessions of acquisition training (Group 1.1 in Table 1). The bottom six panels are for the subjects that received 40 trials per session for seven daily sessions (Group 1.2). Thus, between the two groups there was a four-fold difference in the number of training sessions and in the number of training trials, but in opposite directions; the group with the 28 training sessions had four-fold fewer total training trials than the group with only seven sessions. The relatively few training trials for the top group explain why their cumulative records attain

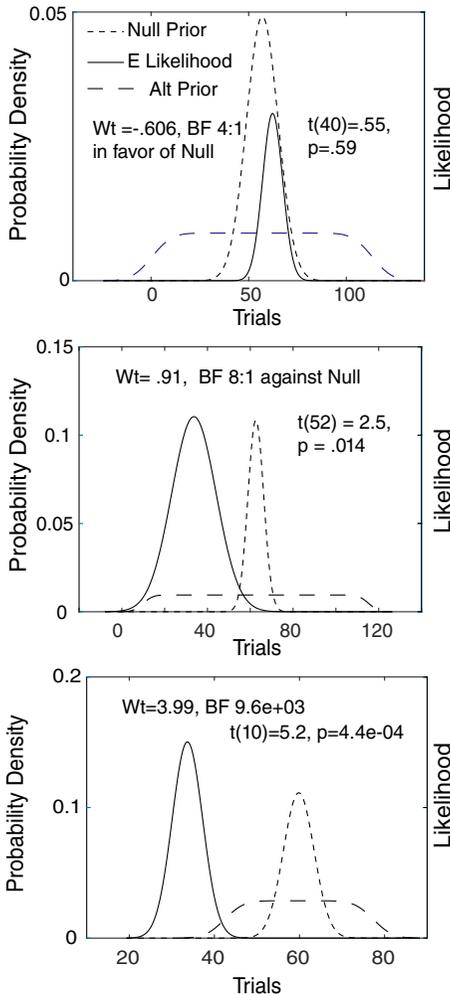


Fig. 2. The computation of the Bayes Factor asks which prior hypothesis better predicts the likelihood function. The prior hypotheses are represented by prior distributions, one for the null hypothesis that the two means do not differ (the null prior, finely dashed curve) and one for the alternative to the null (coarsely dashed). These prior distributions are plotted against the left axis (probability density). The question is which hypothesis puts more probability under the likelihood function (solid curve), which describes the likelihood of different possible values for the mean of the “experimental” group given the data from that group. The likelihood function is plotted against the right axis (likelihood). These three examples come from the comparisons made in the course of this data analysis. The width of the alternative prior was determined by the results of the *t* test. The alternative prior was the mean of the null prior $\pm 2\sigma$, where σ is the pooled estimate of the standard deviation of the distributions. The *t* test assumes the two distributions have equal variance. The *t* test was 2-tailed, meaning that there was no prior hypothesis about the direction of the possible difference in the means. Notice that the Bayes alternative prior extends to either side of the mean of the null prior. The null prior is the normalized likelihood function given the data from the “control” experiment.

much lower asymptotic cumulative differences (y axis) and why these cumulative records terminate at fewer than 150 trials (x axis).

Inspection of the inserts in Figure 3 suggests a conclusion that we confirm later: a four-fold increase in the number of training trials—from 70 or 80 total trials to 240 or 280—has no effect on the cumulative elevation score in the recovery sessions. What increases recovery is the span of training, not the number of trials in that training.

The thin solid vertical lines in Figure 3 mark trials to acquisition (the onset of a reliable response to the CS). The distributions of the loci of the vertical lines in the two sets of panels do not overlap; every subject in the top group (1.1) began responding reliably to the CS in fewer than 50 trials; every subject in the bottom group (1.2) began after more than 50 trials. For this comparison, we have a 2-tailed $t(10) = 5.16$, with $p < .001$ and a bi-directional Bayes Factor (BF) of almost 1,000:1 against the null hypothesis that the means of these two distributions do not differ. This result is further confirmation of the well-established and robust effect of trials per session on the rate of acquisition, when measured by either trials or reinforcements to acquisition: Fewer trials per session increase rate of acquisition (Kehoe & Macrae, 1994; Papini & Dudley, 1993; Papini & Overmier, 1985).

A different measure of the efficiency of reinforced trials in promoting the appearance of the conditioned response may also be considered: cumulative training time—the amount of time the subject has been in the training context experiencing reinforced trials when it begins to respond. Cumulative context experience and cumulative CS experience, together with the cumulative counts of reinforcements occurring in the context and during CS, are the determinative quantities in Rate Estimation Theory, a parameter-free theory of cue competition (aka assignment of credit) in Pavlovian protocols (Gallistel, 1990; Gallistel & Gibbon, 2000). For Group 1.1 (two or three trials/session), the mean cumulative time in the training context when a consistent conditioned response to the CS appeared was 116 min; for Group 1.2 (40 trials/session), it was 198 min ($t_2(10) = 4.63$, $p < .001$, Cohen’s $d = 2.7$; BF $\gg 100:1$ against the null). Thus, by either measure of trial efficiency (trials to

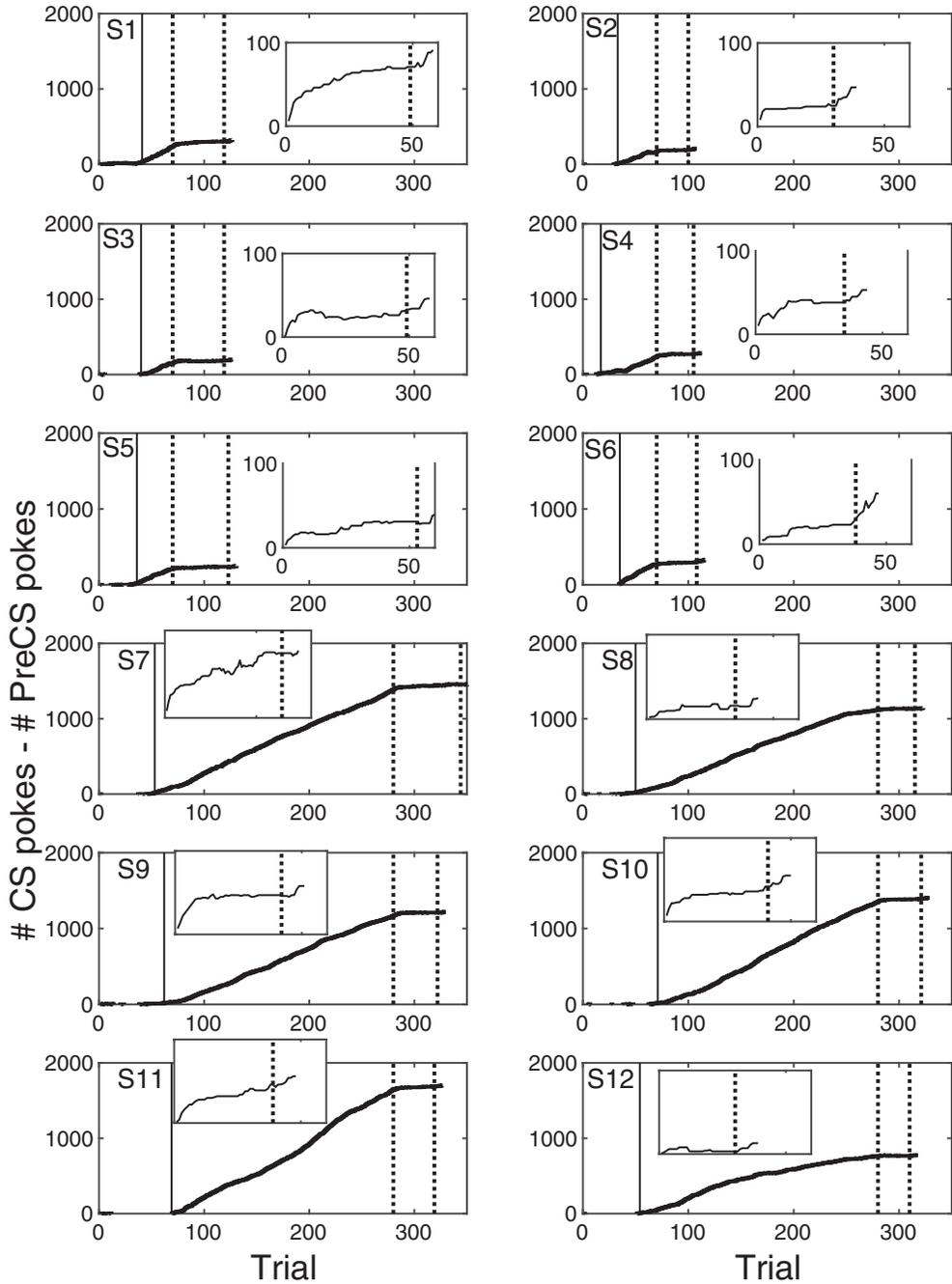


Fig. 3. Cumulative records of the elevation scores for subjects in Experiment 1 (cumulative elevation score versus cumulative number of trials). The top six panels are the records for the subjects in Group 1.1 (28 sessions with only 2 or 3 trials/session); the bottom six are for the subjects in Group 1.2 (7 daily sessions with 40 trials/session). Thin solid vertical lines mark the trial on which consistently elevated nose poking during the CS is estimated to have first appeared. The dotted vertical lines mark the end of training (end of the acquisition sessions) and the end of the extinction session. The insets show only the extinction and recovery portions of the cumulative records. The dotted vertical lines in the insets mark the end of the extinction session. The axes' scales for the insets are 0 to 60 trials on the x axis and 0 to 100 for the cumulative elevation score on the y axis (with one exception, S7, where the x-axis scale is 0 to 80).

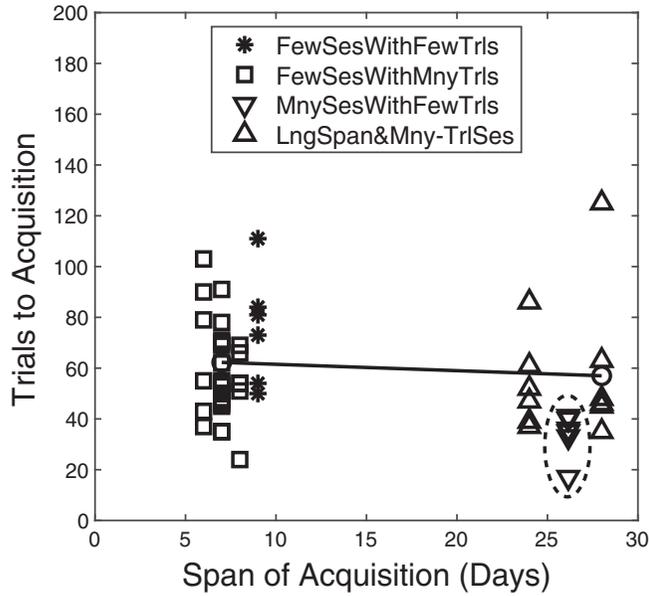


Fig. 4. Trials to acquisition for all 54 subjects broken down by the four possible combinations of few versus many trials per session and few versus many sessions. The data from Group 1.1 with 28 two- or three-trial sessions (the subjects in the top six panels of Fig. 2), are the inverted triangles within the dashed oval. The data for the other six subjects in that experiment (Group 1.2) are among those in the cluster of squares on the left (short span, many trials). The thin line with circles at either end connects the mean of all the subjects with fewer than nine daily sessions (Groups X.1, Y.2, Z.1, Z.2 & X.2) to the mean of all the subjects whose training spanned more than 23 days with 10- or 40-trial sessions (Groups X.2, 3 & Y.1, upright triangles). The inverted triangles (inside the dashed oval) have been offset to an earlier day, to separate them from the upright triangles at 28 days.

acquisition or training time to acquisition), conditioning progresses more rapidly when sessions are short with very few trials per session. In the Discussion, we consider why this might be.

When sessions have only one or a very few trials, there will generally be more sessions before the conditioned response appears, in which case elevated responding appears only after a longer *span* of training (number of days over which sessions are distributed). This suggests that the span of training may itself be an important variable.

Figure 4 plots trials to acquisition as a function of the span of training. One sees that the inverted triangles (inside the dashed oval) are outliers. They are the data from the upper group in Figure 1 (Group 1.1). If we exclude this group, there is no effect on trials to acquisition of the other combinations of span and number of trials per session. The thin almost horizontal line connects the mean of all the points to the left of a span of 9 days to the mean of all the points to the right of a span of 23 days, excluding only the inverted triangles.

The difference between these two means does not approach statistical significance whether assessed by a two-tailed *t* test or by a Bayes Factor. The comparisons that justify pooling the data that went into the mean for the left cluster and pooling the data that went into the mean of the right cluster did not yield *p* values or BFs that approached conventional alpha levels (.05 for *p*'s, 3 for BFs).

Besides the training of Group 1.1, the training of three other groups in the cluster at right in Figure 4 spanned more than 23 sessions: Group 2.2 had 28 daily 10-trial sessions. Group 3.1 had 24 daily 10-trial sessions. Group 2.3 had seven 40-trial sessions at intersession intervals that ranged between 1 and 7 days, with an average intersession interval of 4 days. Our intent in including this group was to determine whether a long span of training induced stronger recovery even when that span contained relatively few sessions.

A one-way ANOVA for all the Groups in Figure 4 except Group 1.1 (inverted triangles within dashed oval) yielded an *F* that did not approach conventional levels of significance.

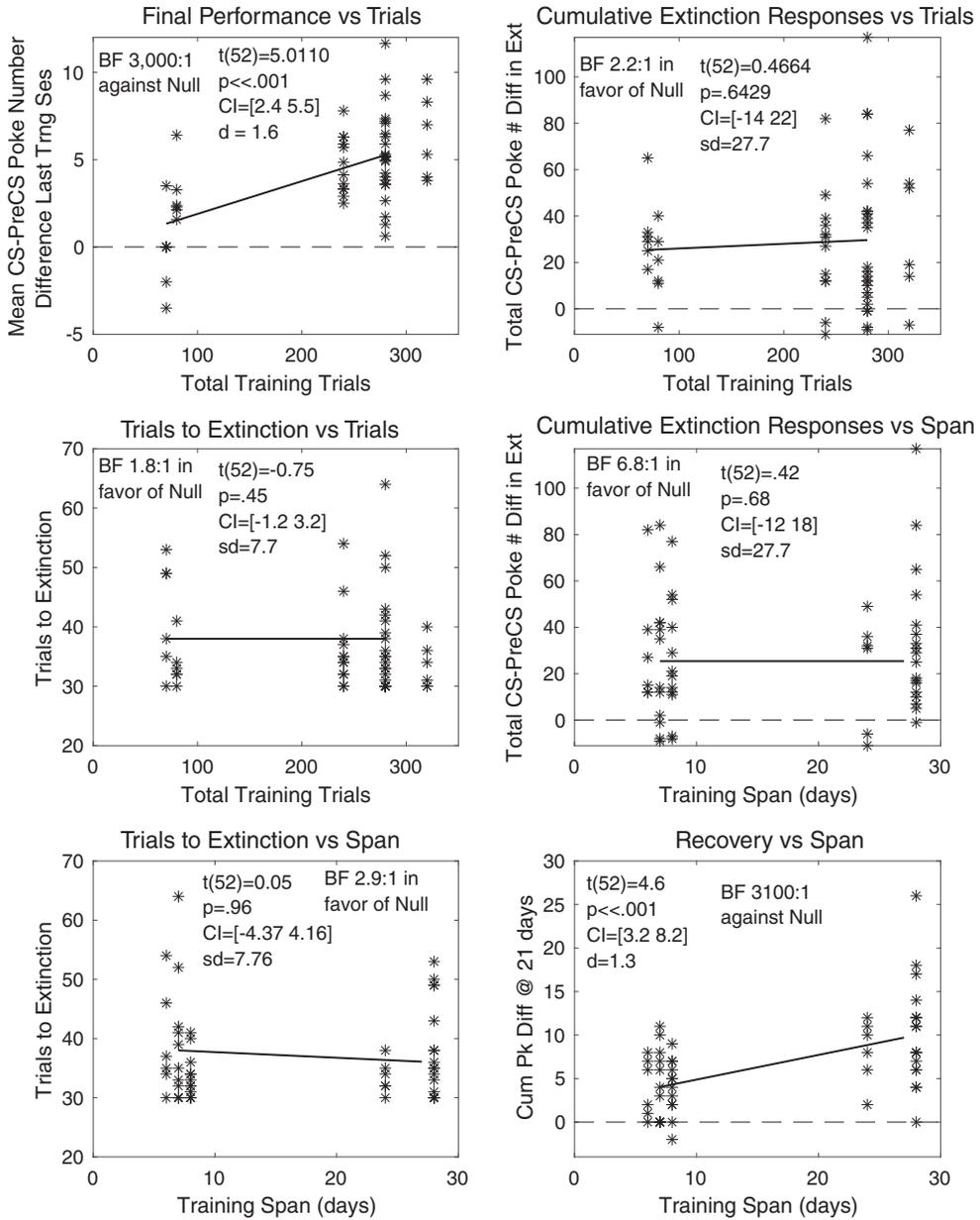


Fig. 5. Results from the Extinction and Recovery phases. The thin lines connect the means of the clusters at the left and right of each panel. Also shown are the results from a 2-tailed t test comparing the two clusters and from a bidirectional Bayes Factor computation. In the two cases where there is a significant effect, Cohen's d ($\Delta\mu/\sigma$) is given beneath the t stats; in the other cases, σ (the pooled standard deviation) is given beneath the t stats.

That and the BFs favoring the null from all the pairwise comparisons between these groups (all groups except 1.1) provide statistical support for the conclusion that when sessions have more than a very few trials, the

efficacy of those trials in promoting the appearance of conditioned responding is reduced. The groups with twenty-eight 10-trial sessions or seven 40-trial sessions fall in with all the groups for which the span of training was

less than 9 days with daily sessions containing 10 to 40 trials. Thus, we conclude that “many” trials means not more than a few (two or three); 10 trials in a session is already “many.” This conclusion is consistent with previous results on the “trials per session effect” (Kehoe & Macrae, 1994; Papini & Dudley, 1993; Papini & Overmier, 1985).

Extinction and Recovery

Figure 5 plots the results from the extinction and recovery phases of the experiment. (See insets in Fig. 3 for examples of the cumulative records of extinction and recovery.) The different panels contrast the effects and lack of effects of widely differing numbers of training trials and widely differing training spans on final performance, trials to extinction, total responses in extinction, and recovery from extinction.

The number of training trials has a strong effect ($d = 1.6$) on the elevation score at the end of conditioning (top left panel in Fig. 5). Rate of responding is a common measure of associative strength. However, the number of training trials has no effect on trials to extinction (middle left panel) nor on the cumulative elevation score during extinction session. Therefore, if rate of responding at the end of training measures associative strength, then associative strength at the conclusion of training has no effect during extinction.

The span of days over which training occurs has no effect on trials to extinction (bottom left panel), nor on the amount of responding observed during extinction (right middle). The span of training does, however, have a strong effect ($d = 1.3$) on the cumulative responses during a four-trial probe for recovery 21 days postextinction (bottom right panel). That this is an effect of the span rather than the number of sessions is shown by comparing the recovery in the three groups in Experiment 2, the experiment that varied session spacing and session number while holding total trials constant (Fig. 6). Three weeks postextinction, there was robust and very similar spontaneous recovery for the two groups with a long training span, despite the fact that one group had only seven sessions (with an average of 4 days between them), while the other had 28 daily sessions. A mixed models

ANOVA revealed significant main effects of Test ($F(1,15) = 48.4, p < .001$), Group ($F(2,15) = 11.62, p = .001$), and their interaction ($F(2,15) = 11.03, p = .001$). The source of this interaction was the absence of spontaneous recovery in the short-span group with the seven daily sessions ($t(5) = 0.27, p = .40$, one-tailed), as opposed to the two groups with a 28-day training span.

Discussion

Theoretical Implications

Implications for content-based theories of Pavlovian conditioning. We take our results and those of others (Bratch et al., 2016; Crystal & Smith, 2014; Panoz-Brown et al., 2016; Panoz-Brown et al., 2018; Wilson, Mattell, & Crystal, 2015; Zhou, Hohmann, & Crystal, 2012) to suggest that the memory contents that drive classically conditioned behavior come from a hierarchically structured record of the episodes in the conditioning experience. The levels of the hierarchy are dictated by the different time spans at which the temporally structured conditioning experience unfolds—from point events (CS onsets and offsets and reinforcement deliveries), to the trial events in which these point events are embedded, to the sessions in which the trial events are embedded, to the epochs over which training may extend, in which the session events are embedded.

The construal of experience. A hierarchically structured representation of the conditioning experience enables subjects to construe it in different ways for different purposes. Different construals permit different policies and mixtures of policies. Consider for an example a protocol in which a subject must press one lever, the “prepare” lever, a fixed number of times in order to arm a “deliver” lever. Pressing the deliver lever after it is armed delivers a pellet. In a sequential counting task such as this, the count and the time elapsed in making it are highly correlated. Therefore, two strategies are possible, based on two different construals of the target during the preparatory phase: 1) Press the “prepare” lever *a target number of times*. 2) Press the “prepare” lever steadily for a *target amount of time*. These are not mutually exclusive strategies. If a subject sometimes loses the count of

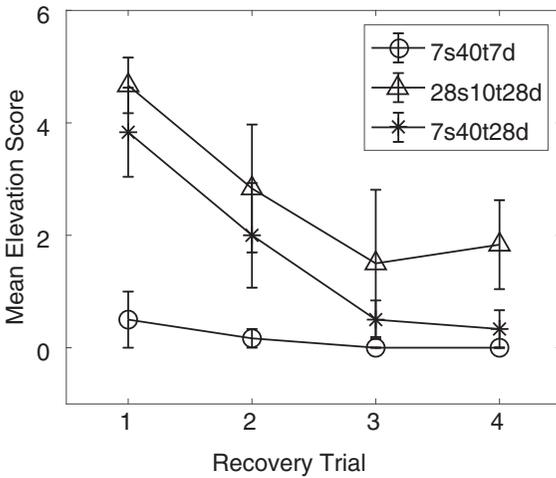


Fig. 6. Recovery at 21 days postextinction for the three groups in Experiment 2. The legend gives for each group the number of sessions (s), trials per session (t), and span of days (d) over which the sessions were distributed. Total trials ($s \times t$) was the same in all groups (280). The two groups with a long training span showed greater (and very similar) recovery, while the group with the short training span showed almost no recovery.

the number of presses so far made while retaining a measure of the time elapsed since it began pressing the “prepare” lever, it may switch from a count target to a time target. Light et al. (2019) devised an analysis of the sequence of presses on the prepare lever that enabled them to detect any of three possible strategies, a count strategy, a timing strategy or a mixed strategy. In a group of eight mice subjects, one mouse relied almost entirely on a count strategy, one relied almost entirely on a timing strategy, and the other six relied on a mixed strategy. A representation of the training experience that holds both possible targets in memory makes this possible.

The minimalist construals allowed in model-free reinforcement learning theories allow only a few policies (Dayan, 2002; Dayan & Berridge, 2014). A construal that takes account only of the number of training trials in our protocol would not enable a strategy that takes session-level properties (trial spacing, number of trials per session, session duration, the probability of reinforcement) into account. A construal that takes account only of session-level statistics would allow neither trial-level statistics (e.g., trial duration, the distribution of CS–US intervals) nor span-level statistics (e.g., the duration of the span, the

number of sessions within the span, the distribution of interspan intervals) to affect the behavior.

Our results and others imply that rodents and birds construe their experience of a Pavlovian conditioning protocol at different levels in different phases of the experiment, because the different policies appropriate to different phases depend on different construals. The policy that determines how rapidly they respond during a CS takes the reinforced trials so far accumulated as an input (Fig. 5, top left). The policy that determines how many responses they will generate before not responding at all for five successive trials does not take reinforcements so far accumulated into account (Fig. 5, top right and middle left). The strategy pigeons and rodents implement during extinction takes into account the probability that a trial during the training epoch was reinforced (Bouton, Woods, & Todd, 2014; Drew, Walsh, & Balsam, 2017; Drew, Yang, Ohyama, & Balsam, 2004; Harris, 2019; Harris & Andrew, 2017). The policy that determines behavior during extinction construes trials as reinforced or not; it is, however, indifferent to how often reinforcement occurred within a reinforced trial (Harris, Kwok, & Gottlieb, 2019) and it is indifferent to the C/T interval, the variable that has a scalar effect on trials and reinforcements to acquisition (Gibbon et al., 1980).

The just-cited finding that increasing the number of reinforcements within the reinforced trials in a partial-reinforcement protocol does not affect trials to extinction (Harris et al., 2019) comports with our analogous finding at the session level: What matters when it comes to the strength of recovery is not how many reinforced trials there were during training nor how many such sessions there were. What matters is whether there were reinforced trials, however few, and the span of training over which sessions containing reinforced trials were spread, not how many such sessions there were within that span.

There may also be an analogy between the effect of spacing trials within sessions and the effect of spacing sessions within the span of training. Because of the effect of trial spacing on acquisition, the number of trials does not affect the progress of conditioning (Gallistel, 2009; Gottlieb, 2008). Similarly, because of the effect of session spacing on recovery, the

number of sessions did not matter in our experiments. What matters for acquisition (the appearance of a conditioned response) is the CS-US interval and cumulative training time, not the cumulative number of trials within that time (Gallistel, 2009; Gottlieb, 2008). Similarly, in our results, what mattered for the robustness of recovery at 21 days post-extinction was the span of days covered by the training sessions, not the number of sessions within that span (Figs. 5 and 6). Testing the generality of this conclusion is a task for the future.

When rats adjust to unpredictable changes in the relative rates of reinforcement in a choice protocol, their strategy for timing the durations of their hopper visits takes into account the frequency of the changes in the rates of reinforcement at the two locations. If there have been no changes for many sessions, rats detect the change quickly, but they adjust to it slowly. Moreover, their pre-change pattern of visit durations recovers for a while at the beginnings of the next two or three sessions after this unexpected change (Gallistel, 2012; Gallistel, Mark, King, & Latham, 2001). When, however, the changes in the relative rates of reinforcement have recently been frequent, rats adjust completely and abruptly shortly after each change—and they show no recovery.

Together, these trial-level and session-level results suggest that subjects count and time episodes defined at very different time scales (from seconds to days). They further suggest that in making the counts and timing the durations of lower level events (trials and hopper visits), subjects distinguish between these different units of experience (episodes) on the basis of different events or different measures embedded within those episodes—whether the trials were reinforced or not and how long they lasted.

The policy a subject adopts when confronted with a change in the stochastic properties of its multilevel representation of its experience depends on the numerical and temporal properties it has encoded. The policy adopted during extinction following training with partially reinforced trials depends on the probability of a reinforced trial, which is to say on counts of reinforced trials and total trials. On the other hand, the policy a subject adopts when confronted with context

extinction—placement in a context where reinforcements were previously provided gratis at random times, but where now no more reinforcements occur—must be based on elapsed time without reinforcement, because the only objectively observable episode in such a protocol is the session. Mustaca, Gabelli, Papine, & Balsam (1991, their Fig. 4) show context conditioning extinguishing during a single session, followed by diminishing spontaneous recovery in subsequent sessions, each recovery followed by within-session reextinction. Elapsed time without reinforcement must have driven the within-session progress of the extinction of context conditioning in their protocol. Thus, the policy in force during extinction may take either the count of unreinforced trials or the time elapsed without reinforcement, depending on the circumstances.

Our intuitive account of our finding that the span of training affects the strength of recovery is that the longer the span of time in which the CS had predictive power, the longer and more vigorously subjects will explore the possibility that its predictive power has returned. The memory of an earlier state in which the CS had predictive power also leads the subject to rapidly resume anticipatory responding when given renewed evidence of CS's predictive power (Napier, Macrae, & Kehoe, 1992; Ricker & Bouton, 1996). In our view, the phenomena of recovery, renewal, reinstatement and rapid reacquisition are all manifestations of a policy for dealing with manifest non-stationarity. Recovery is probing for whether what was once true may be again true. Reinstatement is this same probing elicited by an unexpected reinforcement. Renewal is this same probing when elicited by a change in context. Rapid reacquisition reflects the same remembered fact that drives recovery, renewal and reinstatement, namely, that there was an epoch during which the CS predicted reinforcement. The increasing rate of extinction in protocols with repeated acquisition and extinction (Clark, 1964; Craig, Sweeney, & Shahan, 2019; Davenport, 1969) reflects the memory for previous extinction episodes and how long they lasted.

So far as we know, there is little quantitative work on the effect of training span on recovery. The experiments here reported were exploratory. To get some idea of what variables mattered and what did not, given the

large parameter space, we had to have many groups. For practical reasons, they had to have small n s. The data in hand do not support the development of a computational model of recovery because we have almost no data on the time course of recovery when the span and the session frequency during acquisition are both varied. The relevant experiments will require a great many experimental groups because the different variables that have been shown to be relevant covary and interact in poorly understood ways (Papini & Overmier, 1985). Serious quantitative modeling will require data on trade-offs between these variables, because trade-off functions are much more powerful revealers of underlying processes than are psychometric functions (Gallistel, Shizgal, & Yeomans, 1981).

A prediction. The current state of our knowledge does, however, suggest predictions to guide further research. The trials required before the appearance of a conditioned response in appetitive Pavlovian conditioning (and in eyeblink conditioning) generally number close to half a hundred and often much more (see Fig. 3, for example). It has long been known, however, that this number depends very strongly on the ratio of the average intertrial interval to the average wait for reinforcement following CS onset (Gallistel & Gibbon, 2000; Gibbon & Balsam, 1981; Gottlieb, 2008). Once one has chosen an average duration for the CS, the longer one makes the average interval between trials, the more informative CS onset becomes. The more informative it is, the fewer the trials required for the conditioned anticipatory response to appear (Balsam, Fairhurst, & Gallistel, 2006; Gottlieb, 2008; Jenkins et al., 1981; Ward, Gallistel, & Balsam, 2013; Ward et al., 2012).

Given our current results and those just cited, we assume that the policy that determines the appearance of consistent anticipatory responding in a Pavlovian delay conditioning protocol takes into account both the informativeness of the trials and the evidence that there is within-session and session-to-session (day-to-day) stability in the predictive value of the CS. In any given session, one or a very few reinforced trials establish for the subject that the predictive power of the CS remains in force. That is why adding more reinforced trials to the sessions does not affect the amount of responding seen in probes for

recovery. By contrast, adding more (few-trial) sessions does affect recovery. We assume that this is an effect of the increased span of training, because the same effect is produced by adding a few (trial-rich) sessions spread out over a comparable span.

These conclusions led us to a prediction: One should be able to get appetitive conditioned behavior in a Pavlovian delay protocol after a very few trials, provided one uses highly informative CSs (short with very long average intervals between them) and provided one has more than one session, thereby spreading training over a span of 2 or more days. We are grateful to an anonymous reviewer for calling to our attention experiments that turn our prediction into a successful retrodiction. In a long sequence of fascinating and important experiments on acquisition in autoshaping, Jenkins et al. (1981) had more than one condition in which there were daily 15-min-long (900 s) sessions with a single 8 s conditioning trial—thus, with a highly informative C/T ratio of 60:1. The median trials to acquisition in these conditions was 2.5. This is dramatically fewer trials to acquisition than is commonly observed in autoshaping experiments with many trials per session (see Fig. 9 in Gallistel & Gibbon, 2000 and other conditions in Jenkins et al., 1981; Kehoe & Macrae, 1994; Papini & Dudley, 1993; Papini & Overmier, 1985).

The Time Scale of Recovery

The only theory of spontaneous recovery which treats *encoded* time as a critical determinant is Devenport's Temporal Weighting Rule (TWR, Devenport, 1998; Devenport, Hill, Wilson, & Ogden, 1997). According to TWR, spontaneous recovery reflects the animal's decision to reinvest in a CS because on average, over the longer run, it has produced more than it has failed to produce. Specifically, the animal computes an estimate of the value of a CS as a signal of reward. This estimate is a weighted average (V_w) of all experiences with the CS, with each experience (Q_i) being weighted by its recency (i.e., inverse of time, T_i , since that experience). Equation 1 shows the mathematical formulation of the model:

$$V_w = \sum_{i=1}^n \left(Q_i \times \frac{1}{T_i} \right), \sum_{i=1}^n \left(\frac{1}{T_i} \right) \quad (1)$$

More recent experiences are weighted more heavily but their privileged weight is discounted in a hyperbolic fashion with the passage of time. Thus, soon after extinction, responding is still depressed because the extinction experience (the quality of which is 0) carries a considerably heavier weight. After a longer delay, however, the extinction weight becomes more similar to the acquisition weight(s) (time makes the extinction experience look less and less recent relative to acquisition¹) and the internal estimate regresses to the true (unweighted) mean, thereby causing the appearance of spontaneous recovery.

TWR's treatment of time as an encoded part of the learning episode allows it to explain an interesting finding by Rescorla (2004). He showed that not just the extinction-test interval but also the training-test interval affects spontaneous recovery. He trained rats in a magazine approach procedure with two stimuli that differed only in the interval between their acquisition and extinction training. For one of the stimuli the two training regimes were separated by 8 days, while for the other that interval lasted only 1 day. Both stimuli were tested 2 days after extinction. Despite identical extinction curves, the stimulus with the shorter interval between acquisition and extinction showed greater spontaneous recovery during testing. Thus, the acquisition-extinction interval inversely affected spontaneous recovery.

Devenport's Temporal Weighting Rule anticipates the finding that the group given many trial-poor sessions shows more recovery than the group given a few trial-rich daily sessions, only if the session, as opposed to the trial, is the unit of experience that enters into the weighted averaging process. However, TWR wrongly predicts that spacing the same number of sessions should, if anything, reduce spontaneous recovery, since it moves the acquisition experiences farther into the past, thus diminishing their positive influence on the weighted average at the time of testing. Our data indicate that spacing the sessions

enhances spontaneous recovery, in a similar way that spacing trials enhances conditioning.

Regardless, we think that Devenport's TWR is the right sort of theory for two reasons: First, it posits a rational strategy for coping with the non-stationarity of predictive relations between events. Second, it assumes that the temporal properties of those events are stored in an accessible memory, from which they may be retrieved to serve as the construal on which a strategy (aka policy) is based. However, as just noted, some of our results are inconsistent with TWR. It seems likely that spontaneous recovery has a time course and that this time course depends on both the span of training and the span of extinction in ways that must be elucidated by further experiment.

A longer span of extinction has been shown to suppress spontaneous recovery (Tapias-Espinosa, Kádár, & Segura-Torres, 2018). However, as in most studies of recovery, recovery was probed at only one postextinction delay. In our exploratory experiments, we probed at two delays (1 or 3 days, and 18 or 21 days). Within-subject probes at multiple delays raise a delicate methodological issue, because each probe is another extinction session. It is plausible in the light of the arguments we have already made that what matters in these probes is the spacing of the sessions more than the number of trials in them. When one probes with even a very few unreinforced at 21 days, one informs the subjects that the predictive power of the CS has not returned after that span of time. This may well affect the strength of recovery observed a few days later.

For future work, we suggest a method that takes into account what has so far been learned: Train with a short CS, because the shorter the CS is, the shorter will be the cumulative session time required for the appearance of the conditioned response. Train with only one trial per session, with session times very much longer than the CS duration; a C/T ratio in the range of 50 to 100 appears desirable. The times of occurrence of the CS within a session should be randomly chosen from a uniform distribution. Under these conditions, one can reasonably expect the conditioned response to appear after only two or three sessions. The days spanned by those training sessions should be varied, because our results imply that it is important. The interval between the end of training and the onset of

¹The effect of hyperbolic discounting will be better exemplified by considering the following example. Suppose an animal receives acquisition training during day 1, followed by extinction training on day 2. On day 3 the extinction experience is twice more recent ($1/1d$) than the acquisition experience ($1/2d$). However, 11 days after extinction both experiences are almost equally recent (extinction recency = $1/11d$, acquisition recency = $1/12d$).

extinction sessions should also be varied (Rescorla, 2004). The span of extinction should also be varied (Tapias-Espinosa et al., 2018). Fortunately, this can be done in such a way as to also vary the intervals between the end of extinction and the probe for recovery: As we have just noted, probes for recovery are further extinction sessions. Thus, in varying the interval between extinction sessions one is varying the intervals at which one is probing for recovery.

Clinical Implications

If our conclusions hold up under further experimental tests, they suggest an explanation for the difficulty of permanently extinguishing maladaptive learned behavior. They suggest that the longer the underlying construal of the situation that elicits the behavior has lasted, the more difficult it will be for extinction experiences to persuade the brain that there is a vanishingly small probability of that construal becoming again worth entertaining at some point in the perhaps distant future. The only way to forestall this would be to repeat the extinction experience at least briefly at ever increasing intervals.

The Challenge for Computational Neuroscience

Formal modeling in computational behavioral and cognitive neuroscience is understandably concerned with neural plausibility. This concern, however, confronts modelers with a dilemma: On the one hand, memory has factual content, which is retrieved on demand in order to inform current behavior. That human brains contain a great many retrievable quantitative facts about their past experiences is evident from introspection. A vast range of well documented facts about learned behavior in nonhuman subjects (Gallistel, 1990), including even insects (Menzel et al., 2005; Menzel et al., 2011; Menzel et al., 2012) implies that nonhuman subjects also remember abstract quantitative facts, such as times of day, durations, distances and directions. On the other hand, the neuroscientific theory of memory—Hebbian synapses, also known as plastic synapses—does not attempt to explain how factual content is encoded in memory (Gallistel, 2017; Gallistel &

Matzel, 2013). Indeed, the problem of encoding factual content is rarely if ever discussed in neurobiological reviews on the status of the search for the engram (Poo et al., 2016). It is as if the genetic code were not discussed in reviews of the material basis of heredity.

This dilemma is manifest in the kinds of models modelers tend to prefer. They prefer models in which behavior is based on one or a very few sufficient statistics that 1) can be computed event by event as running averages, and 2) can be regarded as being in some sense associative strengths, hence encodable in plastic synapses. (This latter more implicit part of the agenda often requires reading between the lines.) We resist giving citations for two reasons. First, a list of appropriate citations would occupy several pages. Second, no matter how many citations we give, some will protest that there exist exceptions, which we do not deny. Despite the exceptions, it is a historical fact that: 1) associative bonds have always been conceived of as conductors of activation, not as symbols that encode the facts about the world revealed by a subject's experiences; and 2) this associative conception of memory has determined almost all efforts to discover the physical basis of memory, whether by experiment or by neurobiologically oriented formal modeling of learned behavior. The focus, both experimentally and in modeling, is on trials, because trials are the episodes within which occur the temporal pairing of events that is assumed to drive the associative process (Poo et al., 2016; Schultz, 2015).

We believe our results and other long-established results that we have repeatedly cited pose a strong challenge to the associative conception of learning and memory. One indication of the seriousness of the challenge is that the results now to be summarized are robust and large effects that are well established experimentally and have been for decades, yet there are very few attempts to deal with them within formalized associative theories of associative learning.

The Fundamental Importance of Computationally Derived Ratios

Probability of reinforcement. The probability of reinforcement during training is known to have a scalar effect on trials to acquisition

and on trials to extinction (Chan & Harris, 2019; Gibbon et al., 1980). Because these effects on trials to acquisition and trials to extinction are scalar, the probability of reinforcement has no effect on reinforcements to acquisition and omitted reinforcements to extinction. The difficulties that this partial reinforcement extinction effect: PREE) poses for associative theories were discussed at length in decades-old influential reviews (“The most critical problem facing any theory of extinction is to explain the effect of partial reinforcement. And, for inhibition theory, the difficulties are particularly great,” Kimble, 1961, p. 286). That quote is 58 years old. Nonetheless, the authors of a recent ambitious modeling effort, who stress the importance of a model that accounts for a wide range of results (Luzardo, Alonso, & Mondragón, 2017), do not deal with the scalar PREE. One wonders where physics would be today if mathematically inclined physicists in the 17th century had ignored the variables that have scalar effects.

Why is the dramatic effect of probability of reinforcement on the most basic variable in any associative theory—the number of trials—such a challenge for associative theorists? Because the probability of reinforcement is the ratio of two counts, a count of the number of reinforced trials and a count of the total number of trials (or, if one thinks the psychologically relevant statistic is the odds ratio rather than the probability, then a count of the nonreinforced trials). The events that must be counted to derive the ratio that drives the behavior are widely and variably spaced in time, and often they do not co-occur. A counter of sequential events must contain a memory mechanism that retains the current count. Whatever the physical variable is that encodes the current count, it must not decay with time. A probability or an odds ratio is one count divided by another count. Because both counts must be retained in a memory of some kind, their ratio can only be derived by a mechanism that implements division and has access to the counts in memory. In short, the counts of different events widely separated in time cannot be conceptualized as *stimuli* that can activate neurons that become connected to other neurons by means of an associative process. The effects of the ratios between

counts seem to require both counters, a memory capable of storing counts, and a mechanism that can access stored counts to compute their ratio. Neuroscience has nothing to say about what such a mechanism might look like. That’s the problem.

The C/T ratio. Again, we have a ratio of two quantities—two durations—and, again, the ratio of these two quantities is known to have a scalar effect on the most basic variable in associative theories of learning, the rate of acquisition (the reciprocal of trials to acquisition). Decades ago, Jenkins et al. (1981, p. 255) wrote, “The effect of trial spacing is so large that no theory of [Pavlovian conditioning] can be considered adequate unless it provides an account of how spacing exerts its effects.” In that same volume, Gibbon and Balsam (1981) showed that the effect of the spacing was scalar: Increasing the cycle duration (the US–US interval) by a factor f reduces reinforced trials to acquisition by $1/f$. The effect of this increase is itself scaled by the CS duration: Trials to acquisition is a function of the ratio between the trial spacing (the US–US interval) and the wait for reinforcement in the presence of the CS (Balsam & Gallistel, 2009; Ward et al., 2012).

More recently, Gottlieb confirmed an extremely counterintuitive consequence of this scalar effect. He showed that reducing the number of reinforced trials in a Pavlovian protocol by a factor of 8 while maintaining the spacing of the remaining reinforced trials (thereby increasing the cycle duration by a factor of 8) reduced trials to acquisition by a factor of 8 (Gottlieb, 2008). This result is a mathematical consequence of what Gibbon and Balsam (1981) had shown more than a quarter century earlier. Nonetheless, Gottlieb’s result was so counterintuitive that a reviewer of his manuscript wrote “Only a few crazies in the Gallistel lab could believe that the number of trials does not matter.” In short, this is another long-established, obviously important quantitative fact that most authors of associative models—even the authors of ambitious models (Luzardo et al., 2017)—make no attempt to deal with and that some reviewers think can only be treated with disbelief. Why is this such a problem?

Like numbers, durations cannot be treated as stimuli in any neurobiologically meaningful

sense. There are neurons that are tuned to the proportion of a learned interval that has so far elapsed (Eichenbaum, 2013, 2014; MacDonald, Lepage, Eden, & Eichenbaum, 2011; Mau et al., 2018). But the existence of these neurons does not justify treating different elapsed times as stimuli in the neurobiological sense (Gershman & Uchida, 2019; Gershman, Moustafa, & Ludvig, 2014). Their existence could have been inferred from the long-established fact that conditioned behavior is appropriately timed: If the moment at which reinforcement occurs can be anticipated, then conditioned responding peaks at that moment, and it subsides soon afterwards if reinforcement fails to occur (Church et al., 1994). The discovery of these neurons is a step forward, but it does not address the essential mechanistic question, which is, What do we have to assume exists in a machine in order to explain behavior driven by the ratio of two very different durations obtained by timing different event types? Time does not act on sensors in either the engineer's or the neurobiologist's understanding of what a sensor is. From the engineer's perspective at least, the machine whose behavior depends on temporal ratios must contain timers. And, given that its behavior is driven by the ratio of the averages of two very different variable durations, then the machine must also possess a memory mechanism capable of retaining at the very least two averages. It must also contain a mechanism that can access those averages and output their ratio. Neuroscience has nothing to say about what such a mechanism might look like. That's the problem.

Different Policies are Based on Different Protocol Quantities

A third problem has been the focus of this report—conditioned behavior has many different aspects: the behavior during acquisition, the pattern of behavior once the conditioned response has appeared, the pattern of behavior during extinction, and the pattern of behavior after extinction (recovery, reinstatement, resurgence, renewal, and rapid relearning). Subjects employ different policies during these different phases. Those different policies depend on different quantitative aspects of the training and extinction protocols. They depend on what we have called

different construals of the subjects' remembered experiences in the experimental context. These different quantities are often temporal and numerical, which means that they cannot be treated as stimuli in the neurobiological sense of the term. And, finally, these abstract quantities come from very different levels of structure in their experience.

The scalar effect of the C/T ratio depends only on the average US–US interval and the average CS–US interval. Its effect on trials and reinforcements to acquisition does not depend on the distributions of those intervals. The pattern of responding during the CS, however, depends on the distribution of the CS–US intervals. It differs dramatically depending on whether those intervals are drawn from an exponential distribution or from a delta distribution (an unvarying interval), even when the means are the same (Libby & Church, 1974; 1975). The C/T ratio has a scalar effect on trials to acquisition but no effect on trials to extinction (Gibbon, Baldock, Locurto, Gold, & Terrace, 1977). The probability of reinforcement has a scalar effect on both (Gibbon et al., 1980). The cumulative number of reinforced trials has a strong effect ($d = 1.6$) on the response rate at the conclusion of training, but no effect on trials to acquisition, nor on extinction, nor on the magnitude of recovery at 21 days postextinction (Fig. 5). Trials per session has a strong effect on reinforcements to acquisition and on elapsed training time to acquisition (Kehoe & Macrae, 1994; Papini & Dudley, 1993; Papini & Overmier, 1985 and Fig. 5) but no effect on extinction or recovery (Figs. 3 and 4). The span of training has a strong effect ($d = 1.3$) on recovery (Figs. 5 and 6) but no effect on extinction (Fig. 5).

Subjects appear to count and time and compute the averages of the wait times for reinforcements, both in the context and in the presence of the CS. They appear to be timing and counting the durations of the individual trials. They appear to be counting and timing the sessions. And they appear to be timing the spans of the different phases of their experience, which may involve counting the days. In most associative theories, the different effects of these different variables are to be explained by a single quantity, variously called associative strength or value, because the construal on which policies (or performance functions)

depend is just the strength of an association or the value of an option.

Perhaps it is time for modelers to accept the fact that memory has factual content and for experimentalists to search for a memory mechanism capable of encoding an abstract quantitative fact, like a numerosity or a duration or a distance or a direction. The behavioral facts are not going to go away. Sooner or later neurobiologists and computational neuroscientists must face them.

References

- Anobile, G., Cicchini, G. M., & Burr, D. C. (2015). Number as a primary preceptual attributed: A review. *Perception*, *45* (1-2), 5-31. doi:https://doi.org/10.1177/0301006615602599
- Arcediano, F., & Miller, R. R. (2002). Some constraints for models of timing: A temporal coding hypothesis perspective. *Learning and Motivation*, *33*, 105-123. https://doi.org/10.1006/lmot.2001.1102
- Balci, F., Allen, B. D., Frank, K., Gibson, J., Gallistel, C. R., & Brunner, D. (2009a). Acquisition of timed responses in the peak procedure. *Behavioural Processes*, *80*, 67-75. doi:https://doi.org/10.1016/j.beproc.2008.09.010
- Balci, F., Freestone, D., & Gallistel, C. R. (2009b). Risk assessment in man and mouse. *Proceedings of the National Academy of Science U. S. A.*, *106*(7), 2459-2463. doi: https://doi.org/10.1073/pnas.0812709106
- Balci, F., Freestone, D., Simen, P., deSouza, L., Cohen, J. D., & Holmes, P. (2011). Optimal temporal risk assessment. *Frontiers in Integrative Neuroscience*, *5*, 56. doi:https://doi.org/10.3389/fnint.2011.00056
- Balsam, P. D., Fairhurst, S., & Gallistel, C. R. (2006). Pavlovian contingencies and temporal information. *Journal of Experimental Psychology: Animal Behavior Processes*, *32*, 284-294. DOI: https://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*(2), 73-78. doi:https://doi.org/10.1016/j.tins.2008.10.004
- Barnet, R. C., Grahame, N. J., & Miller, R. R. (1993). Temporal encoding as a determinant of blocking. *Journal of Experimental Psychology: Animal Behavior Processes*, *19*, 327-341. https://doi.org/10.1037/0097-7403.19.4.327
- Barnet, R. C., & Miller, R. R. (1996). Temporal encoding as a determinant of inhibitory control. *Learning and Motivation*, *27*, 73-91. https://doi.org/10.1006/lmot.1996.0005
- Blaisdell, A. P., Denniston, J. C., & Miller, R. R. (1998). Temporal encoding as a determinant of overshadowing. *Journal of Experimental Psychology: Animal Behavior Processes*, *24*(1), 72-83. https://doi.org/10.1037/0097-7403.24.1.72
- Bouton, M. E. (1991). Context and retrieval in extinction and in other examples of interference in simple associative learning. In L. Dachowski & C. R. Flaherty (Eds.), *Current topics in animal learning* (pp. 25-53). Hillsdale, NJ: Lawrence Erlbaum.
- Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, *114*, 80-99. https://doi.org/10.1037/0033-2909.114.1.80
- Bouton, M. E., Woods, A. M., & Todd, T. P. (2014). Separation of time-based and trial-based accounts of the partial reinforcement extinction effect. *Behavioural Processes*, *101*(1), 22-31. doi: https://doi.org/10.1016/j.beproc.2013.08.006
- Bratch, A., Kann, S., Cain, J. A., Wu, J.-E., Rivera-Reyes, N., Dalecki, S., ... Crystal, J. D. (2016). Working memory systems in the rat. *Current Biology*, *26*, 351-355. doi: https://doi.org/10.1016/j.cub.2015.11.068
- Burger, D. C., Denniston, J. C., & Miller, R. R. (2001). Temporal coding in conditioned inhibition: Retardation tests. *Animal Learning & Behavior*, *29*(3), 281-290. https://doi.org/10.3758/BF03192893
- Chan, C. K. J., & Harris, J. A. (2019). The partial reinforcement extinction effect: The proportion of trials reinforced during conditioning predicts the number of trials to extinction. *Journal of Experimental Psychology: Animal Learning and Cognition*, *45*(1), 43-58. doi: https://doi.org/10.1037/xan0000190
- Church, R. M., Meck, W. H., & Gibbon, J. (1994). Application of scalar timing theory to individual trials. *Journal of Experimental Psychology: Animal Behavior Processes*, *20* (2), 135-155.
- Clark, F. C. (1964). Effects of repeated VI reinforcement and extinction upon operant behavior. *Psychological Reports*, *15*, 943-955. https://doi.org/10.1037/0097-7403.20.2.135
- Cole, R. P., Barnet, R. C., & Miller, R. R. (1995). Temporal encoding in trace conditioning. *Animal Learning and Behavior*, *23*(2), 144-153. https://doi.org/10.3758/BF03199929
- Craig, A. R., Sweeney, M. M., & Shahan, T. A. (2019). Behavioral momentum and resistance to extinction across repeated extinction tests. *Journal of the Experimental Analysis of Behavior*, *112*(3), 290-309. doi: https://doi.org/10.1002/jeab.557
- Crystal, J. D., & Smith, A. E. (2014). Binding of episodic memories in the rat. *Current Biology*, *24*, 2957-2961. doi:https://doi.org/10.1016/j.cub.2014.10.074
- Cunningham, P. J., & Shahan, T. A. (2018). Suboptimal choice, reward-predictive signals, and temporal information. *Journal of Experimental Psychology: Animal Learning and Cognition*, *44*(1), 1-22. https://doi.org/10.1037/xan0000160
- Davenport, J. W. (1969). Successive acquisitions and extinctions of discrete bar-pressing in monkeys and rats. *Psychonomic Science*, *16*, 242-244. https://doi.org/10.3758/BF03332665
- Davison, M., & Cowie, S. (2019). Timing or counting? Control by contingency reversals at fixed times or numbers of responses. *Journal of Experimental Psychology: Animal Learning and Cognition*, *45*(2), 222-241. doi: https://doi.org/10.1037/xan0000201
- Dayan, P. (2002). Reinforcement learning. In H. Pashler & R. Gallistel (Eds.), *Steven's handbook of experimental psychology* (3rd ed. 103-129). New York, NY: John Wiley & Sons.
- Dayan, P., & Berridge, K. C. (2014). Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation. *Cognitive, Affective, &*

- Behavioral Neuroscience*, 14(2), 473-492. <https://doi.org/10.3758/s13415-014-0277-8>
- Denniston, J. C., Blaisdell, A. P., & Miller, R. R. (1998). Temporal coding affects transfer of serial and simultaneous inhibitors. *Animal Learning and Behavior*, 26(3), 336-350. <https://doi.org/10.3758/BF03199226>
- 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. <https://doi.org/10.1037/0097-7403.30.3.190>
- Devenport, L. D. (1998). Spontaneous recovery without interference: Why remembering is adaptive. *Animal Learning and Behavior*, 26, 172-181. <https://doi.org/10.3758/BF03199210>
- Devenport, L. D., Hill, T., Wilson, M., & Ogden, E. (1997). Tracking and averaging in variable environments: A transition rule. *Journal of Experimental Psychology: Animal Behavior Processes*, 23(4), 450-460. <https://doi.org/10.1037/0097-7403.23.4.450>
- Drew, M. R., Walsh, C., & Balsam, P.D. (2017). Rescaling of temporal expectations during extinction. *Journal of Experimental Psychology: Animal Behavior Learning and Cognition*, 43(1), 1-14. <https://doi.org/10.1037/xan0000127>
- Drew, M. R., Yang, C., Ohyama, T., & Balsam, P. D. (2004). Temporal specificity of extinction in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes*, 30(3), 163-176. <https://doi.org/10.1037/0097-7403.30.3.163>
- Dudai, Y. (2012). The restless engram: consolidations never end. *Annual Review of Neuroscience*, 35, 227-247. <https://doi.org/10.1146/annurev-neuro-062111-150500>
- Eichenbaum, H. (2013). The hippocampus, time and memory across scales. *Journal of Experimental Psychology: General*, 142(4), 1211-1230. doi: <https://doi.org/10.1037/a0033621>.
- Eichenbaum, H. (2014). Time cells in the hippocampus: a new dimension for mapping memories. *Nature Neuroscience* 15, 732-744. doi:<https://doi.org/10.1038/nrn3827>
- Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review*, 62, 145-154. <https://doi.org/10.1037/h0048509>
- Estes, W. K., & Burke, C. J. (1953). A theory of stimulus variability in learning. *Psychological Review*, 60, 276-286. <https://doi.org/10.1037/h0055775>
- Fetterman, J. G., & Killeen, P. R. (1995). Categorical scaling of time: Implications for clock-counter models. *Journal of Experimental Psychology: Animal Behavior Processes*, 21, 43-63. <https://doi.org/10.1037/0097-7403.21.1.43>
- Foster, D. J., & Wilson, M. A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, 440(7084), 680-684. <https://doi.org/10.1038/nature04587>
- Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: Bradford Books/MIT Press.
- Gallistel, C. R. (2009). The importance of proving the null. *Psychological Review*, 116(2), 439-453. <https://doi.org/10.1037/a0015251>
- Gallistel, C. R. (2012). Extinction from a rationalist perspective. *Behavioural Processes*, 90, 66-88. doi:<https://doi.org/10.1016/j.beproc.2012.02.008>
- Gallistel, C. R. (2017). The coding question. *Trends in Cognitive Science*, 21(7), 498-508. doi:<https://doi.org/10.1016/j.tics.2017.04.012>
- Gallistel, C. R., Craig, A., Shahan, T. A. (2019). Contingency, contiguity and causality in conditioning: Applying information theory and Weber's Law to the assignment of credit problem. *Psychological Review*, 126(5), 761-773. doi:<https://doi.org/10.1037/rev0000163>
- Gallistel, C. R., & Gelman, R. (1990). The what and how of counting. *Cognition*, 34(2), 197-199. doi:[https://doi.org/10.1016/0010-0277\(90\)90043-j](https://doi.org/10.1016/0010-0277(90)90043-j)
- Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review*, 107(2), 289-344. <https://doi.org/10.1037/0033-295X.107.2.289>
- Gallistel, C. R., King, A., & McDonald, R. J. (2004). Sources of variability and systematic error in mouse timing behavior. *Journal of Experimental Psychology: Animal Behavior Processes*, 30(1), 3-16. <https://doi.org/10.1037/0097-7403.30.1.3>
- Gallistel, C. R., Mark, T. A., King, A., & Latham, P. E. (2001). The rat approximates an ideal detector of changes in rates of reward: Implications for the law of effect. *Journal of Experimental Psychology: Animal Behavior Processes*, 27, 354-372. <https://doi.org/10.1037/0097-7403.27.4.354>
- Gallistel, C. R., & Matzel, L. D. (2013). The neuroscience of learning: Beyond the Hebbian synapse. *Annual Review of Psychology*, 64, 169-200. <https://doi.org/10.1146/annurev-psych-113011-143807>
- Gallistel, C. R., Shizgal, P., & Yeomans, J. S. (1981). A portrait of the substrate for self-stimulation. *Psychological Review*, 88(3), 228-273. <https://doi.org/10.1037/0033-295X.88.3.228>
- Gallistel, C. R., & Wilkes, J. T. (2016). Minimum description length model selection in associative learning. *Current Opinion in Behavioral Science*, 11, 8-13. doi: <https://doi.org/10.1016/j.cobeha.2016.02.025>
- Geary, D. C., Berch, D.B., & Koepke, K. M. (Eds.) (2015). *Evolutionary origins and early development of number processing*. New York: Elsevier/Academic Press.
- 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, 8. doi:<https://doi.org/10.3389/fncom.2013.00194>.
- Gershman, S. J., & Uchida, N. (2019). Believing in dopamine. *Nature Reviews Neuroscience*, 20, 703-714. doi: <https://doi.org/10.1038/s41583-019-0220-7>
- Gibbon, J. (1977). Scalar expectancy theory and Weber's Law in animal timing. *Psychological Review*, 84, 279-335. <https://doi.org/10.1037/0033-295X.84.3.279>
- Gibbon, J., Baldock, M. D., Locurto, C., Gold, L., & Terrace, H. S. (1977). Trial and intertrial durations in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes*, 3, 264-284. <https://doi.org/10.1037/0097-7403.3.3.264>
- 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: Academic.
- Gibbon, J., Farrell, L., Locurto, C. M., Duncan, H. J., & Terrace, H. S. (1980). Partial reinforcement in autoshaping with pigeons. *Animal Learning and Behavior*, 8, 45-59. <https://doi.org/10.3758/BF03209729>

- Gottlieb, D. A. (2008). Is the number of trials a primary determinant of conditioned responding? *Journal of Experimental Psychology: Animal Behavior Processes*, *34*(2), 185–201. <https://doi.org/10.1037/0097-7403.34.2.185>
- Gür, E., Duyan, Y. D., & Balci, F. (2017). Spontaneous integration of temporal information: implications for representational/computational capacity of animals. *Animal Cognition*, *21*(2). doi:<https://doi.org/10.1007/s10071-017-1137-z>
- Harris, J. A. (2019). The importance of trials. *Journal of Experimental Psychology: Animal Learning and Cognition*, *45*(4), 390–404. <https://doi.org/10.1037/xan0000223>
- Harris, J. A., & Andrew, B. J. (2017). Time, trials and extinction. *Journal of Experimental Psychology: Animal Learning and Cognition*, *43*(1), 15–29. <https://doi.org/10.1037/xan0000125>
- Harris, J. A., Kwok, D.W.S., & Gottlieb, D. A. (2019). The partial reinforcement extinction effect depends on learning about nonreinforced trials rather than reinforcement rate. *Journal of Experimental Psychology: Animal Learning and Cognition*, *45*(4). doi:<https://doi.org/10.1037/xan0000220>
- Honig, W. K. (1981). Working memory and the temporal map. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 167–197). Hillsdale, NJ: Erlbaum.
- Hull, C. L. (1943). *Principles of behavior*. New York: Appleton-Century-Crofts.
- Jafarpour, A., Fuentemilla, L., Horner, A. J., Penny, W., & Duzé, E. (2014). Replay of very early encoding representations during recollection. *The Journal of Neuroscience*, *34*(1), 242–248. doi:<https://doi.org/10.1523/JNEUROSCI.1865-13.2014>
- Jenkins, H. M., Barnes, R. A., & Barrera, F. J. (1981). Why autoshaping depends on trial spacing. In C. M. Locurto, H. S. Terrace, & J. Gibbon (Eds.), *Autoshaping and conditioning theory* (pp. 255–284). New York: Academic.
- Kehoe, E. J., & Macrae, M. (1994). Classical conditioning of the rabbit nictitating membrane response can be fast or slow: Implications for Lennartz and Weinberger's (1992) two-factor theory. *Psychobiology*, *22*, 1–4.
- Kheifets, A., Freestone, D., & Gallistel, C. R. (2017). Theoretical implications of quantitative properties of interval timing and probability estimation in mouse and rat. *Journal of the Experimental Analysis of Behavior*, *108*(1), 39–72. <https://doi.org/10.1002/jeab.261>
- Kheifets, A., & Gallistel, C. R. (2012). Mice take calculated risks. *Proceedings of the National Academy of Science*, *109*, 8776–8779. <https://doi.org/10.1073/pnas.1205131109>
- Kimble, G. A. (1961). *Hilgard and Marquis' conditioning and learning*. NY: Appleton-Century-Crofts.
- Kraemer, P., & Spear, N. E. (1993). Retrieval processes and conditioning. In T. R. Zentall (Ed.), *Animal cognition* (pp. 87–107). Hillsdale, NJ: Erlbaum.
- Kutter, E. F., Bostroem, J., Elger, C. E., Christian, E. E., Mormann, F., & Nieder, A. (2018). Single neurons in the human brain encode numbers. *Neuron*, *100*(3), 753–761. <https://doi.org/10.1016/j.neuron.2018.08.036>
- Libby, M. E., & Church, R. M. (1974). Timing of avoidance responses by rats. *Journal of the Experimental Analysis of Behavior*, *22*, 513–517.
- Libby, M. E., & Church, R. M. (1975). Fear gradients as a function of the temporal interval between signal and aversive event in the rat. *Journal of Comparative and Physiological Psychology*, *88*(2), 911–916. <https://doi.org/10.1037/h0076420>
- Light, K. R., Cotten, B., Malekan, T., Dewil, S., Bailey, M. R., Gallistel, C. R., & Balsam, P. D. (2019). Evidence for a mixed timing and counting strategy in mice performing a Mechner counting task. *Frontiers in Behavioral Neuroscience*, *13*(109). doi:<https://doi.org/10.3389/fnbeh.2019.00109>
- Luzardo, A., Alonso, E., & Mondragón, E. (2017). A Rescorla-Wagner drift-diffusion model of conditioning and timing. *PLoS Computational Biology*, *13*(11), e1005796. doi:<https://doi.org/10.1371/journal.pcbi.1005796>
- MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). Hippocampal “time cells” bridge the gap in memory for discontinuous events. *Neuron*, *71*(4), 737–749. doi:<https://doi.org/10.1016/j.neuron.2011.07.012>
- Mattar, M. G., & Daw, N. D. (2018). Prioritized memory access explains planning and hippocampal replay. *Nature Neuroscience*, *21*, 1609–1617. doi:<https://doi.org/10.1038/s41593-018-0232-z>
- Mau, W., Sullivan, D. W., Kinsky, N. R., Hasselmo, M. E., Howard, M. W., & Eichenbaum, H. (2018). The same hippocampal CA1 population simultaneously codes temporal information over multiple timescales. *Current Biology*, *28*, 1499–1508. <https://doi.org/10.1016/j.cub.2018.03.051>
- McConnell, B. L., & Miller, R. R. (2014). Associative accounts of recovery-from-extinction effects. *Learning and Motivation*, *46*, 1–15. doi:<https://doi.org/10.1016/j.lmot.2014.01.003>
- Menzel, R., Greggers, U., Smith, A., Berger, S., Brandt, R., Brunke, S., ... Watz, S. (2005). Honey bees navigate according to a map-like spatial memory. *Proceedings of the National Academy of Sciences*, *102*(8), 3040–3045. <https://doi.org/10.1073/pnas.0408550102>
- Menzel, R., Kirbach, A., Hass, W-D., Fischer, B., Fuchs, J., Koblöfsky, M., ... Greggers, U. (2011). A common frame of reference for learned and communicated vectors in honeybee navigation. *Current Biology*, *21*, 645–650. <https://doi.org/10.1016/j.cub.2011.02.039>
- Menzel, R., Lehmann, K., Manz, G., Fuchs, J., Koblöfsky, M., & Greggers, U. (2012). Vector integration and novel shortcutting in honeybee navigation. *Apidologie*, *43*(5), 229–243. doi:<https://doi.org/10.1007/s13592-012-0127-z>
- Mustaca, A. E., Gabelli, F., Papine, M. R., & Balsam, P. D. (1991). The effects of varying the interreinforcement interval on appetitive contextual conditioning. *Animal Learning and Behavior*, *19*, 125–138. <https://doi.org/10.3758/BF03197868>
- Napier, R. M., Macrae, M., & Kehoe, E. J. (1992). Rapid reacquisition in conditioning of the rabbit's nictitating membrane response. *Journal of Experimental Psychology: Animal Behavior Processes*, *18*, 182–192. <https://doi.org/10.1037/0097-7403.18.2.182>
- Ólafsdóttir, H. F., Bush, D., & Barry, C. (2018). The role of hippocampal replay in memory and planning. *Current Biology*, *28*, R37–R50. doi:<https://doi.org/10.1016/j.cub.2017.10.073>

- Panoz-Brown, D., Corbin, H. E., Dalecki, S. J., Sluk, C. M., Wu, J.-E., & Crystal, J. D. (2016). Rats remember items in context using episodic memory. *Current Biology*, *26*(20), 2821-2826. doi:https://doi.org/10.1016/j.cub.2016.08.023
- Panoz-Brown, D., Iyer, V., Carey, L. M., Sluka, C. M., Rajic, G., Kestenman, J., ... Crystal, J. D. (2018). Replay of episodic memories in the rat. *Current Biology*, *28*(10), 1628-1634.e1627. doi:https://doi.org/10.1016/j.cub.2018.04.006
- Papini, M. R., & Dudley, R. T. (1993). Effects of the number of trials per session on autoshaping in rats. *Learning and Motivation*, *24*(2), 175-193. https://doi.org/10.1006/lmot.1993.1011
- Papini, M. R., & Overmier, J. B. (1985). Partial reinforcement and autoshaping of the pigeon's key-peck behavior. *Learning and Motivation*, *16*(1), 109-123. https://doi.org/10.1016/0023-9690(85)90007-4
- Pavlov, I. (1927). *Conditioned reflexes* (G. V. Anrep, Trans.). New York: Dover.
- Poo, M., Pignatelli, M., Ryan, T. J., Tonegawa, S., Bonhoeffer, T., Martin, K. C., ... Stevens, C. (2016). What is memory? The present state of the engram. *BMC Biology*, *14*, 40. doi:https://doi.org/10.1186/s12915-016-0261-6
- 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). New York: Wiley.
- Rescorla, R. A. (1993). Inhibitory associations between S and R in extinction. *Animal Learning & Behavior*, *21*, 327-336. https://doi.org/10.3758/BF03197998
- Rescorla, R. A. (2004). Spontaneous recovery varies inversely with the training-extinction interval. *Learning & Behavior*, *32*, 401-408. https://doi.org/10.3758/BF03196037
- Ricker, S. T., & Bouton, M. E. (1996). Reacquisition following extinction in appetitive conditioning. *Animal Learning and Behavior*, *24*(4), 423-436. https://doi.org/10.3758/BF03199014
- Savastano, H. I., & Miller, R. R. (1998). Time as content in Pavlovian conditioning. *Behavioural Processes*, *44*(2), 147-162. https://doi.org/10.1016/S0376-6357(98)00046-1
- Schultz, W. (2015). Neuronal reward and decision signals: From theories to data. *Physiological Reviews*, *95*(3), 853-951. doi:https://doi.org/10.1152/physrev.00023.2014
- Shahan, T. A., & Cunningham, P. (2015). Conditioned reinforcement and information theory reconsidered. *Journal of the Experimental Analysis of Behavior*, *103*, 405-418. https://doi.org/10.1002/jeab.142
- Sunsay, C., & Bouton, M. E. (2008). Analysis of trial-spacing effect with relatively long intertrial intervals. *Learning and Behavior*, *36*, 92-103. https://doi.org/10.3758/LB.36.2.104
- Tapias-Espinosa, C., Kádár, E., & Segura-Torres, P. (2018). Spaced sessions of avoidance extinction reduce spontaneous recovery and promote infralimbic cortex activation. *Behavioural Brain Research*, *336*, 59-66. doi:https://doi.org/10.1016/j.bbr.2017.08.025
- Taylor, K. M., Joseph, V., Zhao, A. S., & Balsam, P. D. (2014). Temporal maps in appetitive Pavlovian conditioning. *Behavioural Processes*, *101*, 15-22. https://doi.org/10.1016/j.beproc.2013.08.015
- Theunissen, F., & Miller, J. P. (1995). Temporal encoding in nervous systems: A rigorous definition. *Journal of Computational Neuroscience*, *2*(2), 149-162. https://doi.org/10.1007/BF00961885
- Thraillkill, E. A., & Shahan, T. A. (2014). Temporal integration and instrumental conditioned reinforcement. *Learning & Behavior*, *42*, 201-208. doi.org/10.3758/s13420-014-0138-x
- Wagner, A. R. (1981). SOP: a model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Information Processing in Animals: Memory Mechanisms* (pp. 5-47). Hillsdale, NJ: Erlbaum.
- Wang, S.-H., & Morris, R. G. M. (2010). Hippocampal-neocortical interactions in memory formation, consolidation and reconsolidation. *Annual Review of Psychology*, *61*, 49-79. https://doi.org/10.1146/annurev.psych.093008.100523
- Ward, R. D., Gallistel, C. R., & Balsam, P. D. (2013). It's the information! *Behavioural Processes*, *95*, 3-7. https://doi.org/10.1016/j.beproc.2013.01.005
- Ward, R. D., Gallistel, C. R., Jensen, G., Richards, V. L., Fairhurst, S., & Balsam, P. D. (2012). Conditional stimulus informativeness governs conditioned stimulus-unconditioned stimulus associability. *Journal of Experimental Psychology: Animal Behavior Processes*, *38*(1), 217-232. doi:https://doi.org/10.1037/a0027621
- 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: Wiley/Blackwell.
- Wilson, G., Mattell, M. S., & Crystal, J. D. (2015). The influence of multiple temporal memories in the peak-interval procedure. *Learning & Behavior*, *43*(2), 163-178. doi.org/10.3758/s13420-015-0169-y
- Zentall, T. R. (2019). Rats can replay episodic memories of past odors. *Learning & Behavior*, *47*(1), 5-6. doi:https://doi.org/10.3758/s13420-018-0340-3
- Zhou, W., Hohmann, A. G., & Crystal, J. D. (2012). Rats answer an unexpected question after incidental encoding. *Current Biology*, *22*, 1149-1153. https://doi.org/10.1016/j.cub.2012.04.040

Received: October 28, 2019

Final Acceptance: November 21, 2019

Editor in Chief: Timothy Shahan

Associate Editor: Timothy Shahan