Information, certainty, and learning

Justin A. Harris

The University of Sydney

C. R. Gallistel

Rutgers University

Address for correspondence:

Justin Harris
School of Psychology
The University of Sydney,
Sydney 2006
Australia

Email: Justin.Harris@sydney.edu.au

Running head: Information, certainty, and learning

Author Note. This research was supported by funding from the Australian Research Council, grant DP210102343. The data and analysis code presented in this article are available for download at https://osf.io/vmwzr/. The work was not pre-registered.

ABSTRACT

More than four decades ago, Gibbon and Balsam (1981) showed that the acquisition of Pavlovian conditioning in pigeons is directly related to the informativeness of the conditioning stimulus (CS) about the unconditioned stimulus (US), where informativeness is defined as the ratio of the US-US interval (*C*) to the CS-US interval (*T*). However, the evidence for this relationship in other species has been equivocal. Here, we describe an experiment that measured the acquisition of appetitive Pavlovian conditioning in 14 groups of rats trained with different *C*/*T* ratios (ranging from 1.5 to 300) to establish how learning is related to informativeness. We show that the number of trials required for rats to start responding to the CS is determined by the *C*/*T* ratio, and the specific scalar relationship between the rate of learning and informativeness is similar to that previously obtained with pigeons. We also found that the response rate after extended conditioning is strongly related to *T*, with the terminal CS response rate being a scalar function of the CS reinforcement rate (1/*T*). Moreover, this same scalar relationship extended to the rats' overall response rates (during the CS and inter-trial interval) which was directly proportional to the overall rate of reinforcement in the context (1/*C*). The findings establish that animals encode rates of reinforcement, and that conditioning is directly related to how much information the CS provides about the US. The consistency of these observations across species, captured by a simple regression function, suggests a universal model of conditioning.

More than a century of laboratory-based research has been devoted to investigating how animals learn about simple relationships between events, such as learning to respond to a conditioned stimulus (CS) that is followed by an unconditioned stimulus (US) or learning to perform a specific action that is reinforced by a rewarding US. Much of that research has focussed on identifying what properties about the CS-US or response-US relationship are most important for learning. There is widespread consensus about the importance of three particular properties. One is the *temporal contiguity* between the events: conditioning emerges sooner when the US follows the CS or response closely in time. Another is the *spacing* of the learning trials: conditioning takes fewer trials when there is a long time-interval between each CS-US or response-US pairing. The third property is the *contingency* between the events: conditioning is more successful when the US occurs reliably in the presence of the CS or response, and does not occur in their absence.

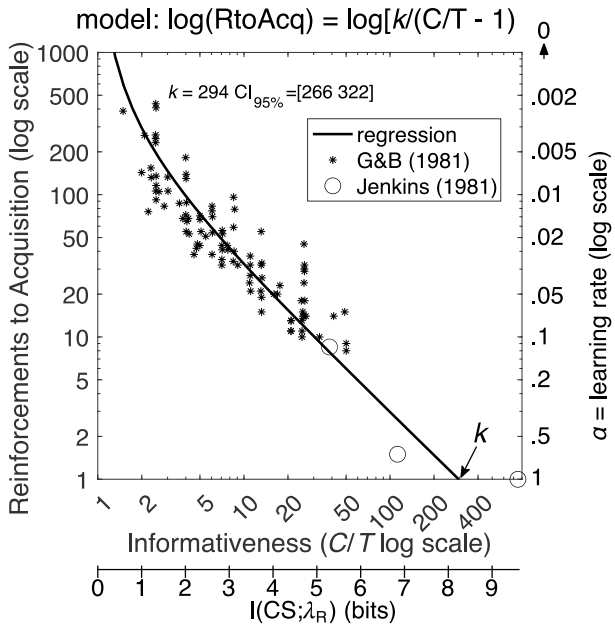
Acquisition of Conditioning, C-over-T, and Informativeness

A landmark study, conducted more than 40 years ago, demonstrated that the first two of these properties—contiguity and trial spacing—are interdependent and subserved by a single principle that encompasses all three properties. In two large-scale experiments, Gibbon and colleagues (1977) measured the number of trials required for pigeons to start pecking at a key-light (the CS) as they learned it was followed by food (the US). Different groups of pigeons were trained with different trial durations (i.e., the interval between onset of the CS and delivery of the US; henceforth *T*). *T* ranged from 1 s to 64 s across 41 groups in two experiments. The inter-trial interval (ITI) also varied between groups, ranging from 6 s to 768 s. Gibbon *et al*. recorded the number of training trials required for each pigeon to start responding reliably during the CS (responding on 3 out of 4 consecutive trials). The birds required more trials as *T* increased, confirming the effect of CS-US contiguity. They also required fewer trials as the ITI increased, confirming the trial-spacing effect. More importantly, however, these two effects were complementary; an increase in *T* had no effect if it was accompanied by a proportional increase in the ITI. Thus, the rate at which the pigeons acquired the conditioned response systematically increased as the ratio of ITI to *T* increased, but there was no separate effect of varying ITI or *T* when their ratio was held constant.

The extent of the relationship between contiguity and trial-spacing was established by a meta-analysis (Gibbon & Balsam, 1981) that combined data from all 41 groups tested by

Gibbon *et al*. (1977) with data from 11 other experiments investigating the acquisition of key-peck responses in pigeons. This analysis compared the number of trials to criterion against the ratio of *C*/*T* (where *C* is the interval between USs, equal to *T* + ITI). The data on pigeon autoshaping are well described by a regression model whose only parameter is the x-intercept, the *C*/*T* value that produces acquisition after a single trial in the median pigeon (*k* in Figure 1). When *C*/*T* > 4, as it is in most Pavlovian protocols, the slope of the regression model on loglog coordinates is approximately -1. Thus, over most of the *C*/*T* range, the number of reinforcements at acquisition is inversely proportional to the *C*/*T* ratio; doubling the ratio reduces required reinforcements by a factor of 2.

Figure 1. **Median reinforcements to acquisition plotted against the *C*/*T* ratio**



Note. *Asterisks show the data from Gibbon and Balsam's (1981) metanalysis of acquisition in pigeons. The two open circles are from Jenkins, Barnes and Barrera (1981) who used even bigger C/T ratios in some groups. The log of the informativeness is the mutual information, I, between the CS and the expected wait for reinforcement (λ_R; lower x axis). The learning rate (right axis, reversed) is the reciprocal of reinforcements to acquisition (thus the axis increases towards zero). The regression model was fit only to the asterisks, but it also predicts the Jenkins et al. data. The value k is the x intercept, the informativeness that produces one-trial learning. The regression model curves upward to infinity as informativeness goes to 1 and mutual information to 0 because, when the CS rate does not differ from the contextual rate of reinforcement, differential responding to the CS does not emerge (Rescorla, 1967, 1968).*

When reinforcements are delivered only during the CS, the *C*/*T* ratio is the ratio of the rate of reinforcement during CSs to the *contextual rate*, the rate subjects expect when in the experimental chamber, without regard to whether the CS is or is not present. Balsam and Gallistel (2009) used the term *informativeness* to refer to the ratio of the CS rate of reinforcement to the contextual rate, because the log of informativeness is the *mutual information* that CS onsets transmit to a subject about the expected wait for the next reinforcement (lower x-axis in Figure 1). The mutual information transmitted by a CS is the upper limit on the extent to which the CS can reduce the subject's uncertainty about the wait for the next reinforcement. When the CS rate equals the contextual rate, no information is transmitted because the informativeness ratio is 1, and log(1) = 0.

Contingency is mutual information divided by available information (Gallistel & Latham, 2023). *Available information* is the amount that reduces a subject's uncertainty to 0. Because temporal measurement error scales with duration measured (Weber's Law, Gibbon, 1977b), contingency = 1 only when two events coincide in time (Gallistel & Latham, 2023).

The analysis presented thus far identifies a fundamental principle of associative learning; what animals learn about events is defined in terms of measurable properties of their temporal distributions. To claim this as a general principle of learning, we must seek evidence for the importance of *C*/*T* in conditioning paradigms with other species. Several studies have sought evidence that conditioning is related to *C*/*T* using an appetitive Pavlovian conditioning paradigm in which rats or mice learn to anticipate the arrival of a food reward (indexed by monitoring their activity at the food cup or licking at a spout) during a CS (Bouton & Sunsay, 2003; Burke et al., 2024; Holland, 2000; Kirkpatrick & Church, 2000; Lattal, 1999; Thrailkill, Todd, & Bouton, 2020; Ward et al., 2012). However, as summarised below, the evidence from these studies is mixed.

The first evidence for an effect of *C*/*T* on the acquisition of responding in rats was reported by Lattal (1999) and Holland (2000), who observed effects that persisted even when all rats were tested with an identical ITI (different groups had been trained with different itis but were shifted to a common interval between trials when tested for responding to the CS). More recently, Ward *et al*. (2012) reported that the log of number of trials for mice to acquire responding scaled with log(*C*/*T*), and this relationship was similar to that described for pigeons (Gibbon et al., 1977; Gibbon & Balsam, 1981). Bouton and Sunsay (2003) also provided

evidence for the importance of *C*/*T* on Pavlovian conditioning in rats. They showed that conditioning was negatively affected when *T* was increased 3-fold by inserting two CS-alone presentations between each CS-US trial (a manipulation that affects *T* but does not affect *C* and, therefore, reduces *C*/*T*) but that conditioning was not affected by a 3-fold increase in *T* brought about by omitting the US from two out of three CS-US trials (a manipulation that increases both *T* and *C* equally and, therefore, does not change *C*/*T*). Most recently, Burke *et al.* (2024) have shown that removing 9 out of 10 CS-US trials from a Pavlovian conditioning schedule, thus increasing *C* 10-fold without changing *T*, reduced the number of trials required for learning by a factor of 10.

In addition to finding an effect of *C*/*T* on conditioning, both Lattal (1999) and Holland (2000) also found an effect of *T* that was independent of the *C*/*T* ratio. They observed that an increase in *T* resulted in a decrease in responding even when the ITI was also increased to keep the *C*/*T* ratio constant. Further evidence against an influence of *C*/*T* ratio comes from a study by Kirkpatrick and Church (2000) in which differences in *C*/*T*, ranging from 1.5 to 12, did not produce differences in the acquisition of conditioned responding in rats. Most recently, Thrailkill *et al*. (2020) observed clear effects of *T*, but not *C*/*T*, on conditioning. Their rats responded at much higher rates to a 10-s CS than to a 60-s CS, even though the groups had identical *C*/*T* ratios. When comparing groups on how quickly responding was acquired, they found no systematic effect of *C*/*T* on the number of 4-trial conditioning blocks required to reach a response criterion.

In sum, in contrast with the impressive evidence from experiments with pigeons, studies of appetitive conditioning in rats and mice have provided inconsistent evidence for the importance of *C*/*T* to conditioning. At the same time, these studies have shown that *T* has an effect on responding that is independent of *C*/*T*. This latter observation is in fact consistent with the evidence from experiments with pigeons. In the original study that established the importance of *C*/*T* on trials to acquisition, Gibbon *et al*. (1977) also observed that the rate at which the pigeons responded to the CS after extended training was negatively related to *T* and not related to *C*/*T*. In other words, while *C*/*T* affected how quickly conditioned responding emerged, *T*, rather than *C*/*T*, determined the level of responding that was ultimately acquired. This distinction may go some way to explaining the inconsistency in the evidence for the effect of *C*/*T* in rats and mice. The inconsistencies may be due to differences in how the point of

acquisition was identified in different studies, and, in particular, whether differences in the amount of responding affected the measure of when responding was first acquired. This concern is not relevant to the results from pigeon experiments because the appearance of their key-peck response is unambiguous thanks to the fact that the baseline rate of that response is effectively zero. In contrast, rats and mice show activity at the food-cup in the ITI, and therefore researchers using this appetitive paradigm must decide how to take account of the baseline response rate when quantifying conditioned responding during the CS (Lattal, 1999).

The current experiment

The experiment described here attempts to elucidate the role of *C*/*T* and *T* in an appetitive Pavlovian conditioning paradigm with rats by distinguishing their impact on the emergence of responding from their effect on the level of responding subsequently acquired after extended conditioning. Fourteen groups of rats were trained for 42 sessions with a single CS that was followed on every trial by delivery of food (the US). Each session contained either 10 CS-US trials (Groups 1-11) or 3 CS-US trials (Groups 12-14). Both *T* and *C* varied between the groups in an uncorrelated fashion (r = -0.19, p = .519) so that effects of *T* and *C* could be assessed independently (summarised in Table 1).

When the location of the CS is spatially separated from the US, as in most appetitive conditioning experiments with pigeons or rodents, the two types of response are mutually incompatible, which can impact on the measurement of conditioned responding. For example, any factor that increases food-cup activity, such as reducing the ITI, may reduce evidence for conditioning that is indexed by CS-directed responses like key-pecks. Conversely, evidence for conditioning indexed by food-cup activity may be reduced to the extent that animals also acquire CS-directed responses. These problems can be mitigated if the CS and US are co-located. In the present experiment, the CS was illumination of a small LED inside the magazine. Conditioned responses were measured using an infra-red photobeam across the opening of the magazine that should detect both food-cup activity and approach responses to the CS. CS-US intervals varied randomly from trial to trial (around a mean equal to *T*) so that response rates remained constant across the length of each trial (see Figure S1 of Supplementary Materials; and Harris & Carpenter, 2011; Harris, Gharaei, & Pincham, 2011).

Table 1. Summary of groups.

Group	*T*	*T* range	*C*	ITI [range]	*C*/*T*
1	42	2-82	63	21 [15-27]	1.5
2	48	2–94	144	96 [15–177]	3
3	21	2-40	94.5	73.5 [15-132]	4.5
4	12	2–22	72	60 [15–105]	6
5	6	2–10	54	48 [15–81]	9
6	62	2–122	930	868 [15–1721]	15
7	30	2-58	600	570 [15–1125]	20
8	18	2–34	486	468 [15–921]	27
9	36	2–70	1296	1260 [15–2505]	36
10	26	2–50	1404	1378 [15–2741]	54
11	8	2–14	576	568 [15–1121]	72
12	16	2-30	1760	1744 [15-3473]	110
13	10	2-18	1800	1790 [15-3565]	180
14	14	2-26	4200	4186 [15-8357]	300

Note. *All values for T (CS-US interval), C (US-US interval), and the inter-trial interval (ITI) are in seconds.*

RESULTS

Rats in all groups eventually poked at higher rates during the CS than during the ITI. (See Supplementary Materials for mean response rates of each group as well as CS and ITI response rates of each individual rat.) Except for one rat (#104) that registered very few responses across the whole experiment, the difference between CS and ITI response rates over the final 5 sessions was significant for every individual rat ($p < .05$, one-tailed Wilcoxon signed rank test) and decisively so in 96% of rats ($p < .001$). Trial-by-trial response data for every rat are available from https://osf.io/vmwzr/.

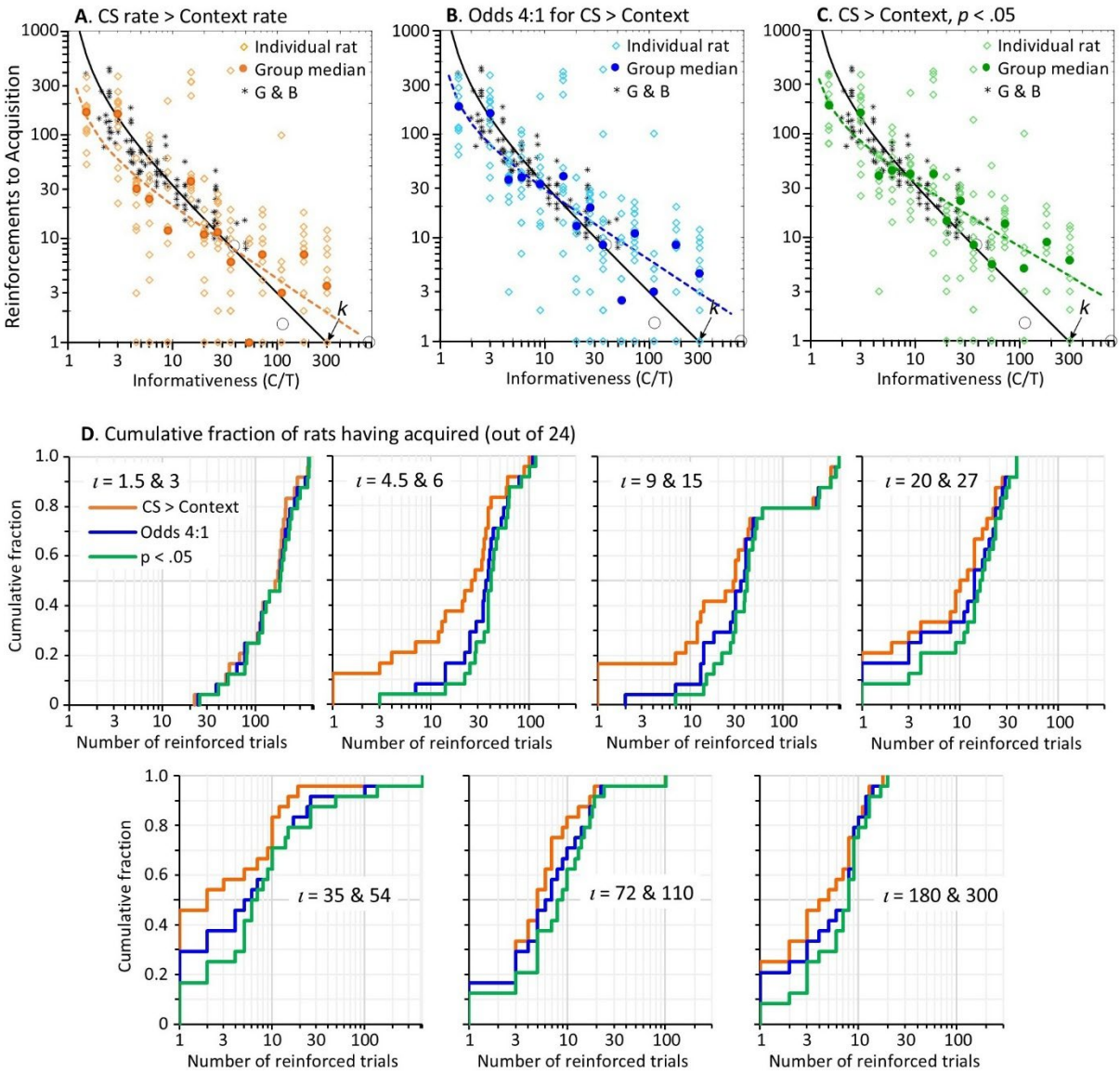
Learning Rate (Trials to Acquisition)

Given that the evidence for response to the CS was decisive in all but one rat, the question of interest is not *whether* there was a significant effect of CS reinforcement on responding, but *when* it first appeared. Our operational definition of learning rate is the reciprocal of reinforcements prior to the appearance of the conditioned response. In our protocol, the conditioned response has appeared when the response rate during the CS is greater than the overall response rate in the experimental context (i.e., the rate across both the ITI and CS intervals, which should not differ before the CS has acquired control over responding). To estimate the point of acquisition in each rat, we calculated the difference between the cumulative response rate during the CS at each trial (cumulative number of responses during the CS divided by the cumulative duration of the CS) and the cumulative contextual response rate (across the CS and ITI). We took the trial when this difference became permanently positive as the first trial on which the rat had acquired responding to the CS (i.e., when the cumulative CS response rate was permanently greater than the cumulative contextual rate). Therefore, the trial before this was our first estimate of the number of reinforcements to learn. Figure 2A plots the estimated reinforcements to acquisition for each individual rat and the median for each group. It shows a steady decrease in the number of reinforced trials prior to acquisition as the *C*/*T* ratio increased, thus confirming the scalar relationship between learning rate and informativeness. The data are provided in Acquisition_Table.xlsx available at https://osf.io/vmwzr/ .

The data in Figure 2A put a number on trials to acquisition for each rat, but these estimates come without an index of the strength of evidence, such as an odds ratio or *p* value. We have used an information-theoretic statistic, the $n\mathrm{D_{KL}}$ (Gallistel & Latham, 2023), to evaluate the strength of evidence at each trial that the rat's CS response rate exceeds its overall response rate. The $n\mathrm{D_{KL}}$ estimates the cumulative coding inefficiency of a model that treats the response rate during the CS as indistinguishable from the contextual response rate. An important advantage of the $n\mathrm{D_{KL}}$ is that, unlike a Fisherian *p* value, its interpretation does not require specification of sample size in advance because the effective sample size is the *n* in the $n\mathrm{D_{KL}}$. This means that the test can be applied iteratively with the data from each successive trial, where *n* is the number of trials up to each iteration. Another useful property of the $n\mathrm{D_{KL}}$ is that its value, measured in nats (natural-log units), can easily be converted into

more familiar odds ratios or *p* values. A detailed explanation of the $n\mathrm{D_{KL}}$ and justification for its use are provided in the Methods section at the end of this paper.

Figure 2. Number of trials for acquisition of responding to the CS, plotted against CS informativeness (*C/T* ratio).



Note. A, B *and* ***C*** *plot the number of trials to reach three different acquisition criteria for CS response rates exceeding the contextual rate. Each plot shows the median number of trials to criterion for each group (filled coloured circles) and the number of trials for each rat (unfilled coloured diamonds) plotted against informativeness of the CS. These data are superimposed onto the data (black asterisks) from Gibbon & Balsam (1981) and the black regression line shown in Figure 1. The dashed coloured lines show the regression model, log[RtoAcq] = log[k·(C/T-1)^s], fitted to the group medians. The slope, s, and R^2 of each regression line is: in* ***A****, s = -0.68 [95% confidence intervals = -0.94, -0.42], R^2 = 0.73; in* ***B****, s = -0.66 [-0.86, -0.46], R^2 = 0.81; in* ***C****, s = -0.6 [-0.76, -0.45], R^2 = 0.86. The seven plots in* ***D*** *show the cumulative fraction of rats that had acquired responding to the CS after* n *reinforced trials (x-axis) according to*

the same 3 acquisition criteria in A to C. In each plot, the data from two groups with similar informativeness (i, iota) are combined to increase the sample size.

We iteratively computed the $n\mathrm{D_{KL}}$ for the difference between CS responding and context responding across trials. Because divergence is a scalar measure of distance, rather than a measure of difference, the $n\mathrm{D_{KL}}$ is positive even when the rate of responding during the CS is less than the overall response rate. In our experiment, some rats did initially respond less during the CS, which we put down to a temporary disruption to ongoing magazine activity by the unfamiliar magazine light. Therefore, because we are only interested in evidence for a positive difference between CS and context response rates, we have added a negative sign to the $n\mathrm{D_{KL}}$ on trials where the cumulative CS rate is less than the overall rate. As a result, the trial on which the $n\mathrm{D_{KL}}$ becomes permanently positive is the trial when the cumulative CS rate permanently exceeds the overall rate (as plotted in Figure 2A). The trial-by-trial values of the $n\mathrm{D_{KL}}$ for every rat are plotted in the figures for individual rats in the Supplementary Materials and are included in the data file nHT.mat available at https://osf.io/vmwzr/. For rats that responded less during the CS such that their $n\mathrm{D_{KL}}$ was initially negative (e.g., Rat #10), the $n\mathrm{D_{KL}}$ began to increase from its lowest point when responding during the CS began to increase, yet it could take many trials before the $n\mathrm{D_{KL}}$ became positive. In these cases, a more accurate marker of acquisition of responding to the CS would be the minimum of the $n\mathrm{D_{KL}}$ rather than the trial on which it became permanently positive. However, the minimum $n\mathrm{D_{KL}}$ is a questionable localiser of acquisition in other cases where it does not mark the start of a systematic increase in the $n\mathrm{D_{KL}}$ (e.g., Rats 1 and 7). Nonetheless, we have included the minimum $n\mathrm{D_{KL}}$ as an index of acquisition in the data plots for each rat in the Supplementary Materials and in Acquisition_Table.xlsx.

We have used the $n\mathrm{D_{KL}}$ to identify in each rat when the evidence for acquisition, based on a positive difference between its response during the CS and its overall response rate, had exceeded one of two statistical thresholds. The more lenient threshold was for odds of at least 4:1 that CS responding was higher than the overall response rate; this threshold is met when the $n\mathrm{D_{KL}}$ is permanently greater than 0.82 (Gallistel & Latham, 2023). While odds of 4:1 would normally be regarded as only weak evidence for an effect, it is important to remember that the evidence for acquisition was eventually decisive in every subject. Our

second, more stringent, criterion was set at $p < .05$ (odds 19:1) for the probability that the CS response rate was not more than the overall rate; this threshold is met when $n\mathrm{D_{KL}} > 1.92$. For both of these statistical criteria, we calculated the $n\mathrm{D_{KL}}$ starting from the trial on which the cumulative CS response rate was permanently greater than the contextual response rate. We started from this trial, rather than from Trial 1, because response rate data from trials prior to the point of acquisition would dilute the evidence for a statistically significant difference in responding once it had emerged, and thereby increase the number of trials required to observe significant responding to the CS. The data from Rat 1 illustrates this point. The CS response rate of Rat 1 permanently exceeded its overall response rate on Trial 52 (when the $n\mathrm{D_{KL}}$ became permanently positive). The $n\mathrm{D_{KL}}$, calculated from that trial onwards, surpassed 0.82 (odds 4:1) after a further 11 trials (on Trial 63) and reached 1.92 ($p < .05$) on Trial 81. By contrast, the $n\mathrm{D_{KL}}$ for this rat, calculated from Trial 1, did not permanently exceed 0.82 until Trial 83 and did not exceed 1.92 until Trial 93, adding 10 or 20 trials to the point of acquisition.

Figure 2B plots number of trials for each rat to reach odds greater than 4:1 of a difference between CS responding and overall responding. Figure 2C plots the number of trials to reach $p < .05$ against that difference in response rates. (Other trials to criterion using statistical thresholds $p < .01$ and .001 are included in Acquisition_Table.xlsx.) As expected, the trials to acquisition in Figures 2B and 2C are higher than those shown in Figure 2A simply because it requires additional trials to accumulate statistical evidence of the difference between CS responding and overall responding. Nonetheless, in all three plots there is clear evidence that the log of trials to acquisition scales with the log of CS informativeness, a fact confirmed by regression lines plotted to each set of data based on the regression model presented in Figure 1. Indeed, the data in Figure 2 are superimposed on Figure 1, revealing a striking correspondence between the data from Gibbon and Balsam's (1981) meta-analysis of learning rates in pigeons and our data with rats, especially in Figures 2B and 2C where statistical thresholds have been applied.

To assess how specific the relationship is between trials to acquisition and informativeness, we performed correlational analyses for the relationship between the median log of trials to acquisition and log(C/T), log(C), and log(T). We set α at .017 to correct for multiple tests. Log(C/T) was the strongest predictor of trials to criterion (r = -0.84, -0.89, and -0.92; p's < .001; for the 3 criteria shown in Figures 2A, B, and C, respectively). Log(C) was

also significantly correlated with the log of trials to criterion (r = -0.71, -0.81, and -0.83; p's < .006) but log(T) was not significantly correlated with trials to any criterion (r = 0.43, 0.35, 0.35; p = .121, .221 and .225). Given that log(C) and log(C/T) were strongly correlated with one another (r = 0.90), we calculated their partial correlations to test whether each made independent contributions to the rate of acquisition. After partialling out the effect of log(C), correlations with log(C/T) remained significant (r's = -0.65, -0.64, and -0.69; $p \leq .017$). In contrast, after partialling out the effect of log(C/T), each correlation with log(C) was not significant (r's = 0.19, -0.03, and -0.07; p's = .543, .916, and .816). In sum, the number of trials required for conditioned responding to emerge was strongly affected by the C/T ratio, whereas neither C nor T alone had any independent effect.

From the plots in Figures 2A, B, and C, it can be hard to appreciate the distribution of trials to acquisition when the data for individual subjects are superimposed. This is particularly the case for subjects learning after their first reinforcement, where individual data pile up on the x-axis. To show the distribution more clearly, Figure 2D contains 7 plots of the cumulative fraction of animals that had acquired responding to the CS by trial n (x-axis) based on the 3 criteria shown in Figures 2A to C. To increase the sample size, each plot combines data from two groups with similar informativeness ratios (ι, iota). Three things become particularly evident in these plots. One is that the curves migrate leftwards as informativeness increases, showing that the whole distribution of subjects takes fewer trials to reach each criterion for learning. The second observation is that, apart from the two groups with the lowest informativeness (ι = 1.5 & 3), many rats (33 out of 144, 23%) acquired responding in just 1 trial, as determined by the criterion that the CS response rate permanently exceeded the contextual rate (orange lines). Of these 33 rats, the evidence for 1-trial learning exceeded 4:1 odds in 60% of cases and was significant at p <.05 in 30% of cases. The third observation is that, while the number of rats showing 1-trial learning increased as informativeness increased from low to intermediate levels, 1-trial learning was less common among rats trained with the highest levels of informativeness. This is surprising because 1-trial learning should be most frequent in rats trained with the highest informativeness. We will return to this issue shortly.

The regression lines in Figures 2A, 2B and 2C confirm Gibbon and Balsam's (1981) observation of timescale invariance in the rate of learning, in that trials-to-acquisition scales

with the ratio of *C* over *T*. Indeed, the rates of learning in the first 10 groups of rats (*C*/*T* ratios <72) agree quite closely with those reported by Gibbon and Balsam across a comparable range of *C/T* ratios, particularly when a statistical threshold was applied to the evidence for responding (panels B and C). However, the slopes of the regressions fitted to the data from all 14 groups are less negative than the slope (-1) of the black regression line modelled on Gibbon and Balsam's data. The primary reason for the shallower gradients in our regressions appears to be the higher-than-expected number of trials to acquisition for the 4 groups trained with the highest *C*/*T* ratios (≥ 72). Indeed, if we fit regressions to our data but excluding those 4 groups, thereby restricting the analysis to the range of informativeness values used by Gibbon and Balsam, we obtain slopes closer to -1: for Figure 2A, slope = -0.93 [95% confidence intervals = -1.33, -0.54]; for Figure 2B, slope = -0.82 [-1.11, -0.52]; for Figure 2C, slope = -0.75 [-0.97, -0.50]. This reiterates the surprising observation made above that the number of rats showing 1-trial learning was smaller in these four groups than in groups trained with intermediate informativeness. It seems implausible that a rat should learn more from one exposure to a moderately informative CS than to a CS that is 6 times more informative about the US.

The fact that groups trained with very high informativeness acquired more slowly than expected, and were less likely to show acquisition after 1-trial, might imply that the scaling between learning rate and informativeness changes at high levels of informativeness. However, the apparent limit on learning rate at high informativeness is likely to be, at least in part, an artefact of our training protocol and, in particular, to the fact that our rats were not given magazine training—sessions in which food was delivered without the CS—before conditioning with the CS. Magazine training ensures that the animals are familiar with the arrival of food into the magazine and, therefore, can promptly detect it during subsequent conditioning trials with the CS. Without magazine training, it is likely that many rats did not find the food immediately after its delivery on the first few conditioning trials. We have confirmed this in a recent unpublished experiment that delivered food pellets at widely spaced intervals (on average every 45 min) and measured the latency for rats to enter the magazine after each pellet was delivered. The median latency for the 32 rats to enter the magazine was 39 s on Trial 1, which decreased to 2 s after 5 trials. Such a delay in finding the food US would substantially increase the effective CS-US interval, measured from CS onset to

discovery of the food pellet by the rat, making the CS much less informative over those initial trials. Crucially, this would affect groups trained with high informativeness much more than those with low informativeness, first because those initial few trials are more critical for establishing learning when informativeness is high, and second because the effective increase in CS-US interval will reduce high levels of informativeness more than low informativeness. To illustrate this second point, consider the impact of a 39-s delay in finding the food pellet for rats in Group 1, with C = 63 and T = 42, versus rats in Group 14, with C = 4200 and T = 14. For Group 1, the 39-s delay would reduce C/T by just 16%, from 1.5 to 1.26 (=102/81); For Group 14, the same 39-s delay would reduce C/T by 73%, from 300 to 80 (=4239/53). Future experiments could reduce the likelihood of large delays in how quickly the rats find the food by giving them magazine training with the same US-US interval to be used during conditioning, thereby retaining the contextual rate of reinforcement at $1/C$.

Another factor that might have concealed evidence for 1-trial learning in the groups trained with extremely informative protocols is that their initial response rates were so low that, in many cases, the $n\mathrm{D_{KL}}$ could not be computed for several trials. Take, for example, Rat 176, trained with the highest level of informativeness (300). This rat had been in the chamber for over an hour when the CS first came on for just 4 s. It was in the chamber for another hour before the CS came on again for 15 s. Because the rat did not register a response during either CS presentation, nor during either 30-s pre-CS interval, the $n\mathrm{D_{KL}}$ could not be computed until Trial 3 when the rat made its first response during the 23-s CS (still failing to register any response in the pre-CS period). At this point, the evidence of responding to the CS was already significant at $p < .05$ ($n\mathrm{D_{KL}}$ = 2.35). The complete absence of responses during the first 2 trials makes it difficult to know whether the rat had in fact learned about the CS on Trial 1 but it failed to register a response during the 15-s presentation of the CS on Trial 2 because its response rate was still very low.

We have developed a principled solution to the problem caused by the lack of responses during the first few short CSs in groups with extreme values of informativeness. This solution uses the $n\mathrm{D_{KL}}$ to parse the trial-by-trial response rates into segments of one or more trials based on inter-poke intervals. Response rates for trials within a segment are treated as the same but the boundary between segments marks a significant change in rate. These parsed response rates during the CS and during the ITI are included in the individual

data plots for each rat in the Supplementary Materials and are in ParseTable.xlsx available at https://osf.io/vmwzr/. When there is a long wait to the first response, followed by a sequence of much shorter inter-response intervals, the parsing algorithm (see Methods) finds a change in rate at that first response, ascribing a response rate of zero before the first response and a higher poke rate in the segment that includes the first response. It does this if the long interval before the first response is unlikely to have come from the distribution of inter-response intervals observed thereafter. However, when the duration of the interval before the first response plausibly belongs to the distribution of subsequent inter-response intervals, the algorithm includes that initial interval within the segment of subsequent responses, such that the onset of the estimated response rate extends back to the start of observation. This can be seen in the parsed data for Rat 176 (see Supplementary Materials) which ascribes a single response rate to the CS and another, lower, rate to the ITI across all of the first 10 trials. Acquisition_Table.xlsx includes estimates, based on the $n\mathrm{D}_{\mathrm{KL}}$, of the number of trials before the parsed CS response rates exceeded the parsed context response rates using increasingly stringent statistical thresholds (odds of 4, 10, 20, 100 & 1000:1 against the null hypothesis that CS response rate equals the context response rate). It shows that, using this method, evidence for 1-trial acquisition with odds > 4:1 is obtained in the majority of rats (56%) that were trained with informativeness ratios of 72 and above, and 44% of rats show 1-trial acquisition with odds > 10:1.

Rate of Responding as Function of Reinforcement Rate (1/T)

After the initial point at which each rat had begun responding to the CS, response rates typically increased as conditioning continued, but there were large differences in how much responding increased. By the end of the experiment, response rates varied by more than three orders of magnitude, from as low as once every several minutes to as high as several times per second. From our analyses (see Supplementary Materials) we realized that the conventional way to compute response rates seriously underestimates the high poke rates observed during some CSs. The conventional calculation of response rate simply divides the poke count by the time interval across which pokes are counted. This implicitly assumes that the durations of the poke themselves—during which another response cannot be initiated—constitute a negligible fraction of the duration in the denominator. However, this is far from true when there is more than one poke per second. Our analysis showed that the mean

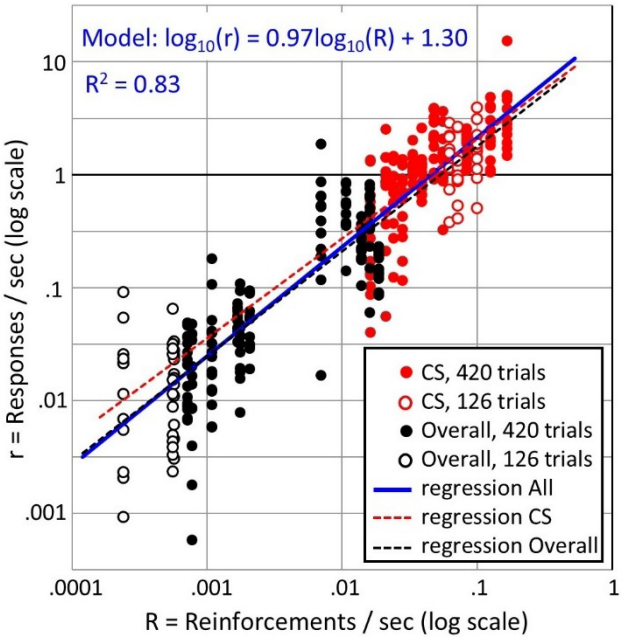
duration of a nose poke is 0.5 s, which will curtail the number of responses that can be produced within a fixed interval. Therefore, in our further analyses of response rates, we have corrected the duration in the denominator by subtracting cumulative poke duration from cumulative time to obtain the cumulative interval over which it was possible for the rat to initiate a poke. This is equivalent to computing rate as the reciprocal of the inter-response interval, measured from the end of one response to the start of the next, except that it can be computed even when there are fewer than two responses in a trial (unlike the inter-response interval).

We have adopted a further correction to the estimate of response rates during the CS. This correction was warranted because the distribution of latencies to the first poke after CS onset is different from the distribution from the intervals separating subsequent pokes (see Supplementary Materials). This was particularly true during CSs that elicited high response rates, where the inter-poke intervals were often measured in milliseconds whereas the latencies of the first pokes were usually around 2 s. This latency may reflect a lower limit on how quickly the rats could reach the magazine after CS onset or occur because 2 s was the lower limit on reinforcement latency after CS onset. Regardless of the reason, in our further analyses of CS response rates we did not include the first pokes in the poke count, and we subtracted the latency of each first poke from the CS durations.

Our analysis of response rates across the last 5 sessions uncovered several important findings. (Very similar results are obtained when using data from the final 10 sessions.) First, the response rate during the CS scaled with the reinforcement rate of the CS. This is shown in Figure 3, where each red dot shows one rat's response rate during the CS plotted against the CS reinforcement rate ($1/T$) for that rat. To extend the evidence for this relationship between response rate and reinforcement rate, the black dots on Figure 3 show the overall response rate of each rat (the estimate of the total responses made across both the ITI and CS divided by the total time out of the magazine) plotted against the overall reinforcement rate for that rat ($1/C$). This reveals that the overall response rate scales with the overall reinforcement rate and, strikingly, the scaling is almost identical to that observed for the CS response rate and reinforcement rate. Thus, there is very little difference between the two regression lines plotted to the CS data and the overall data (red and black dashed lines in Figure 3). Indeed, a single regression line (blue in Figure 3) plotted through all the data accounts for 83% of the

variance in log response rates. To compare these regression models, we computed Bayesian Information Criterion (BIC) scores for the 2-parameter model that fits the single (blue) regression line to all the data and for the 4-parameter model that fits separate regression lines for to CS data and overall response rate data. The comparison of these scores strongly favoured the single regression model (BIC = -712) over the separate regressions model (BIC = -696), with the estimated Bayes Factor = 2,981 (=e^8).

Figure 3. Terminal response rates (final 5 sessions) of individual rats as a function of the reinforcement rate on double-logarithmic coordinates.



Note. *Red circles show response rates during the CS (number of pokes after the first poke divided by the remaining CS time when the rat's head was out of the magazine) plotted against the CS reinforcement rate (1/T). Black circles show overall response rates across the inter-trial interval and CS (calculated as total number of pokes divided by total time with head out of the magazine) plotted against the baseline reinforcement rate (1/C). Filled circles show data of rats given 420 reinforced trials (Groups with C/T ratios ≤ 72) and open circles show data of rats given 126 reinforced trials (Groups with C/T ratios ≥ 110). The Model gives the equation for the solid blue regression line (R^2 = 0.83) fitted to all the data (CS and overall response rates). The slope of this line is virtually indistinguishable from 1 (R^2 = 0.83, when slope fixed at 1). The dashed red and black lines show the regression lines fitted separately to the CS response rates and overall responses rates.*

It is notable that the single regression line in Figure 3 has a slope very close to 1. Indeed, a regression line with slope fixed at 1 (and intercept = 1.36) accounts for almost as much variance as the 2-parameter regression model, both R^2 = 0.83. Moreover, the BIC for this 1-parameter regression model is -717, which constitutes strong evidence favouring this single parameter regression over the 2-parameter regression (estimated BF = 12). Fixing the slope at 1 in the log-log domain means that the response rate and reinforcement rate also scale linearly in the non-log domain, and this line passes through the origin (i.e., response rate = 0 when reinforcement rate = 0). The intercept in the log domain is the exponent of the slope in the non-log domain, and thus, in this case, the response rate is 23 (=$10^{1.36}$) times faster than the expected reinforcement rate. Put another way, the average time the rat waits between magazine entries is 23 times shorter than the average time it expects to wait for food (also Harris & Carpenter, 2011). It is particularly impressive that this simple relation holds across more than 2 orders of magnitude variation in reinforcement rate, from rates as low as one pellet every 70 min to CS rates as high as 10 pellets per minute. As is also apparent in Figure 3, the variability about the scalar relation between response rate and reinforcement rate also scales with reinforcement rate. That is why the regression must be computed in the logarithmic domain.

The contextual rate of responding plotted in Figure 3 (black circles) is the weighted average of the response rates during the ITI and CS. By the end of conditioning, the CS response rates were much higher than the ITI response rates, so the ITI rates themselves are lower than the contextual rates shown in Figure 3. We have not plotted the ITI response rates in Figure 3 because it is not clear what values of reinforcement rate they should be plotted against. Arguably, the ITI reinforcement rate is zero because food was never delivered in the ITI (in the absence of the CS). However, the fact that the rats consistently produce some responding during the ITI indicates that they did not learn to treat the ITI as a defined period in which reinforcement never occurs. Rather, their response rate during the ITI must reflect some expectation of a low rate of reinforcement attributed to the experimental context itself.

Trials from onset of responding to peak responding

Having established the relationship between response rate and reinforcement rate, we next analysed how response rate increased over trials towards its maximum value. Our first analysis assumes that there is a consistent (monotonic) increase in response rate starting

from the initial point of acquisition. This analysis followed a method recently described (Harris, 2022) that uses the slope of the cumulative response rate over trials to identify the trial on which the response rate had reached each decile (from 10% to 90%) of the peak response rate. Based on our earlier analysis (see Figure 3), our measure of the CS response rate was calculated by dividing the response count (excluding the first response in each CS presentation) by the total time out of the magazine during the CS (and excluding the latency to the first response). As shown in Figure S4A of Supplementary Materials, the response rate of an individual rat varies greatly from trial to trial. However, a clearer picture of the overall change in responding over trials can be obtained by plotting the cumulative response count against the cumulative opportunity to respond (cumulative time out of the magazine; Figure S4B).

The slope of the cumulative function can be used to estimate the rat's response rate across conditioning to find when responding had reached a given proportion of the peak response rate. To analyse how response rates changed across the course of conditioning, we extracted a segment of each rat's conditioning data starting from the trial, t_1, on which the response rate during the CS became reliably greater than the ITI response rate and finishing at the trial, t_{end}, on which the response rate reached its peak (according to a moving average with a window width of 3 sessions). The total change in responding across conditioning was calculated by subtracting the response rate at the start of this segment of trials, R_1, from the peak response rate, R_{max} (at the end of the segment): $\Delta R = R_{\text{max}} - R_1$. To identify when responding had increased by 10% of ΔR, we estimated what the cumulative response count, cumR', would be at each trial, t, if the rat maintained a fixed level of responding equal to the starting rate plus 10% of ΔR. Thus, $\text{cum}R'_t = R_1 + 0.1 \cdot \Delta R \cdot \text{cum}T_t$, where $\text{cum}T_t$ is the cumulative CS-US interval from trial 1 to t. We then calculated the difference between $\text{cum}R'_t$ and the observed $\text{cum}R_t$. The trial at which this difference was maximum was identified as the trial when the slope of $\text{cum}R_t$ had increased by at least 10% of ΔR. This process was repeated for all deciles up to 90%. An example of the values obtained for one rat is show in Figure S4C. In this example, the rats' response rate had increased by 10% of ΔR on Trial 34, by 50% of ΔR by Trial 147, and by 90% of ΔR by Trial 233.

The analysis illustrated in Figure S4 was conducted on the individual data of all rats (except those rats missing data from Session 1). We excluded rats with a ΔR less than 0.1

responses/s. The mean number of trials to reach each decile for every rat in the 14 groups is shown in Figure 4. With some exceptions, for most rats, the relationship between the number of trials and response decile was roughly linear, meaning that their response rate increased uniformly over trials as it approached the peak response rate. This is clearest in the averaged functions (thick black lines in Figure 4). To investigate more precisely the relationship between trials to criterion, t_c, and response decile, d, we compared 4 different functions for their fit to the data for each individual rat. Based on the apparent linear increase in trials across deciles, the first function tested was a straight line, $t_c = m.d + c$. The second was an exponential function, $t_c = c.e^{m.d}$, which has a continuously increasing slope and thus predicts a systematic increase across deciles in the number of trials between deciles. This function had most successfully captured the relationship between trials and response deciles in the data analysed by Harris (2022). The third function was an inverse cumulative Gaussian: $t_c = s.2^{½}.erf^{-1}[2.(m.d + 0.5)–1] + c$. This was used to model a stepwise increase in responding, as would be observed if responding were governed by a decision process when evidence for the CS-US relationship exceeded some threshold (Gallistel, Fairhurst, & Balsam, 2004). The fourth function was a log function, $t_c = -(log_e[1–d])/k + c$, derived as the inverse of a cumulative exponential function. This function models the relationship between trials and response criterion predicted by an error-correction learning algorithm such as used by the Rescorla-Wagner model (Rescorla & Wagner, 1972). All model fitting was conducted using Matlab™. To compare between these four models of the data, the Bayesian Information Criterion (BIC) was calculated as

$$BIC = n \cdot log_e\left(\frac{RSS}{n}\right) + p \cdot log_e n$$

where RSS is the residual sum of squares for the difference between each observed t_c and its corresponding point on the fitted function, n is the number of points being fitted (= 9) and p is the number of free parameters in the function.

The overall performance of the functions can be compared by summing the BICs obtained from all rats. The function with the smallest ΣBIC is the function with most evidence. According to this analysis, the linear function had the most evidence, ΣBIC = 9588, followed by the exponential function, ΣBIC = 9974. The difference between these (ΔBIC=386) constitutes overwhelming evidence in favour of the linear model: BF = $e^{\Delta BIC/2}$ = 7.3×10^{83}

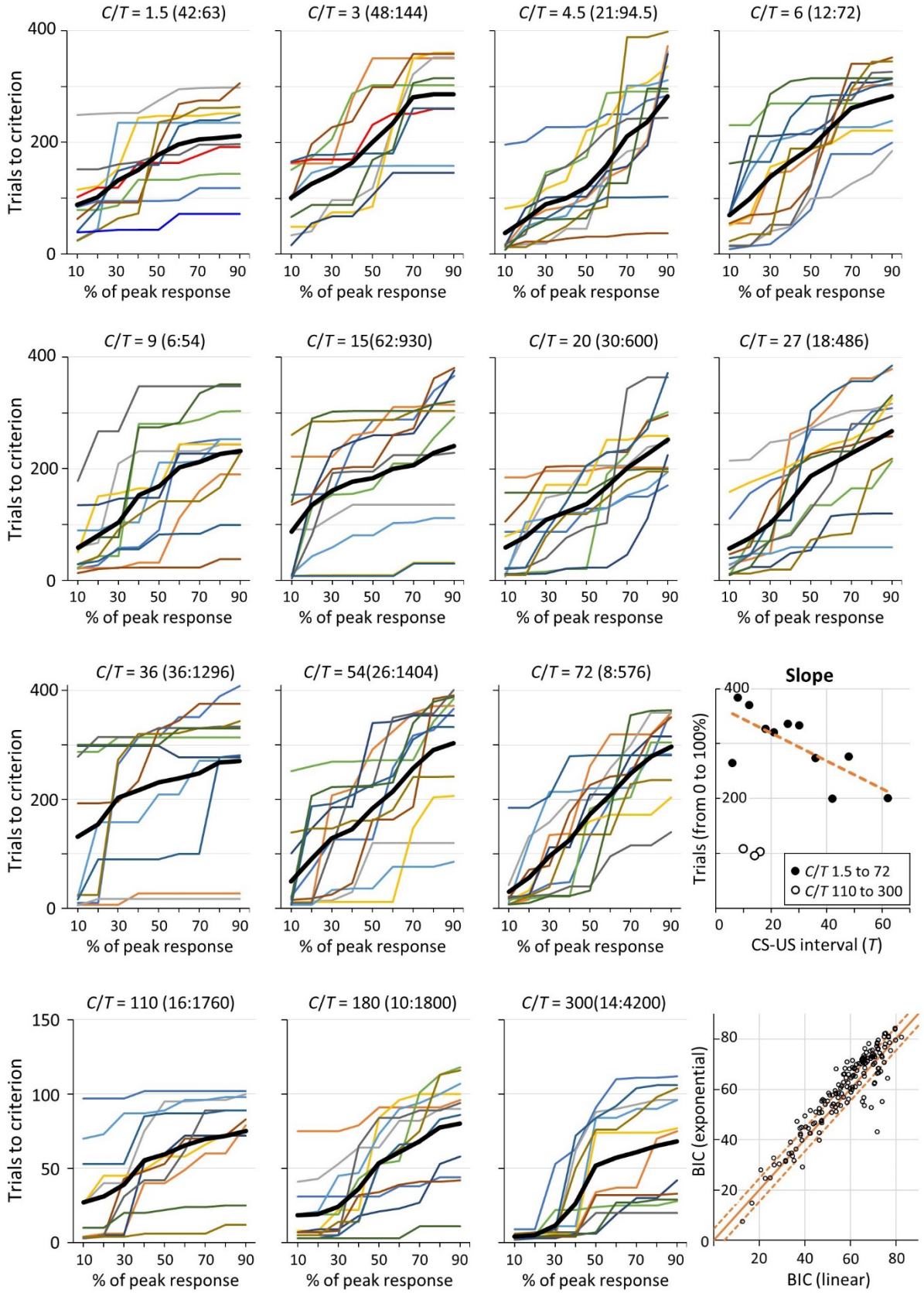
(Wagenmakers, 2007). The ΣBIC of the other two models were much higher again, 12128 and 10358, indicating that they had even less support from the data. In addition to this comparison of the aggregate BIC, a more specific comparison can be made by comparing the BIC for one model against the BIC for another for each rat. The scatter plot in the bottom right corner of Figure 4 plots, for each rat, the BIC for the exponential function against the BIC for the linear function (the two best-fitting functions). The orange line marks where these BICs are equal. The large majority of values (72%) sit above this line. These represent cases where the BIC for the linear function is lower than that for the exponential function, meaning that the evidence is stronger for the linear function. If we look at cases where the difference in BICs was greater than 4.6, corresponding to strong evidence in favour of one function over the other (a BF ≥ 10), there is strong evidence favouring the linear function over the exponential in 39% of cases, whereas only 9% of cases provide strong evidence in favour of the exponential function. The evidence favouring the linear function over either the inverse cumulative Gaussian or the log function is even stronger: 69% of cases provide stronger evidence (BF ≥ 10) in favour of the linear over the inverse cumulative Gaussian and 0% of cases favour the latter function; 59% of cases provide strong evidence in favour of the linear over the log function and only 8% of cases strongly favour the log function over the linear.

The above analyses reveal an overall tendency for the response rate to increase approximately linearly up to the point where the peak response rate is reached. To test how the rate of this increase varied across groups, we calculated the correlation coefficient between the slope of the line for each group, as shown in Figure 4, and the value of *C*, *T*, or *C*/*T*. The slope was negatively correlated with *C* and *C/T*, *r*s = -0.66 and -.67, *p*s <= .010, and not correlated with *T*, *r* = -0.05, *p* = .864. The correlations with *C* and *C/T* were substantially driven by the fact that the three groups with the largest values of *C* and *C/T* had much smaller slope values than the other 11 groups (slopes < 110 vs slopes >=200). It is likely that this difference in slopes was an artefact of the smaller number of trials given to those three groups (126 vs 420 trials) which limited the range of values that their slopes could take. In light of this, we calculated correlations after excluding the 3 groups given only 126 trials. For the other 11 groups, the slope was not significantly correlated with *C*/*T*, *r* = 0.51, *p* = .110, or with *C*, *r* = 0.08, *p* = .938, but was significantly negatively correlated with *T*, *r* = -0.73, *p* = .011 (see plot titled "Slope" in Figure 4). The correlation between slope and *T* remained significant after

partialling out the effect of *C/T*, r = -0.70, p = .025. This suggests that the response rate increased more quickly when the reinforcement rate of the CS (1/*T*) was higher, which is not surprising given that the peak rate of responding also increased systematically with reinforcement rate (Figure 3).

The preceding analysis assumes that there is an overall monotonic increase in responding over trials. However, this trend is not always apparent, particularly when the CS informativeness is low. Such irregular and non-monotonic changes in responding are revealed by the analysis that uses the $n\mathrm{D_{KL}}$ to parse the response rates into segments that have significantly different rates (either higher or lower). (See description of response rate parsing using the Kullback-Leibler divergence, and the $n\mathrm{D_{KL}}$, in Methods.) The parsed response rates are shown in the bottom two plots of the individual figures for each rat in the Supplementary Materials. These show that, in some cases, the path to the peak rate can be bumpy and the peak can be higher than the terminal response rate. This is evident in the parsed response rates of Rats 1 to 6 that were all trained with the lowest level of informativeness (ι = 1.5). By contrast, when informativeness is very high, the peak is usually reached almost immediately and there is little subsequent variation in the poke rate (e.g., see parsed rates for Rats 171 to 176 that were trained with the highest level of informativeness, ι = 300). The parsing of response rates shown in these figures used a stringent decision criterion; the algorithm found a step up or down only when the evidence for a divergence exceeded 6 nats. On the null hypothesis, a divergence that large would occur by chance about once in 1800 tests, well above the number of tests made when parsing the data over the 420 or 126 trials of the experiment. When the divergence between the contextual rate of reinforcement and the CS rate is high, as when ι = 300, there are very few steps (between 0 and 2) in the CS poke rate over 126 trials. By contrast, when the divergence between the contextual rate and the CS rate of reinforcement is low, as when ι = 1.5, there are many up and down steps, some lasting only a few trials. Thus, the linear increases to a peak revealed by the preceding analyses should not be taken to indicate that a steady increase is routinely seen in the individual subjects regardless of the informativeness. When informativeness is low, the post-acquisition rate of responding during CSs and the difference between it and the rate during ITIs are unstable but trend upwards. When informativeness is high, a stable asymptotic rate generally appears after only one or two early steps, the first of which is often the step after the first trial.

Figure 4. Number of trials to each decile of peak response rate.



***Note**. The coloured line plots show the mean number of trials to each decile for each rat in each of the 14 groups (identified by the C/T ratio). The thick black line in each plot is the average for that group. The slopes of those black lines for each group are plotted against each group's T in the rightmost plot of the third row. The final plot in the bottom right corner of the figure plots the Bayesian Information Criterion (BIC) for an exponential function against the BIC for a straight line when each was fitted to the trials-to-criterion data of individual rats (each open black circle shows the BICs for one rat). For each point in the top left of the plot (above the solid orange line),* $BIC_{line} < BIC_{exponential}$*, meaning that the data were better accounted for by the line than the exponential function (and vice versa for points in the bottom right half of the plot). The two dashed orange lines mark where the difference in BICs equals 4.6 which equates to odds of 10:1 in favour of the function with the lower BIC.*

One of us has previously argued that responding appears abruptly when the accumulated evidence that the CS reinforcement rate is greater than the contextual rate exceeds a decision threshold (Gallistel et al., 2004). The new more extensive data require a more nuanced view. Evidence about the manner in which responding changes over the course of training is to some extent dependent on the analytic method used to track those changes. One method we have used here suggests that responding rises steadily over trials. The other method we have used relies on an information theoretic measure of divergence to identify discrete points of change (up or down) in the response record. This method suggests the first increment in responding can be large and usually followed by further, often smaller, increments in responding. At the same time, there is marked within-subject variability in the response rate, characterized by large steps up and down in the parsed response rates following the initial increment, but this variability tends to decrease across further training, with fewer and smaller steps in both the ITI and CS response rates. We think that the initial large increment reflects an underlying decision process whose latency is controlled by diminishing uncertainty about the two reinforcement rates and hence about their ratio. It is possible that diminishing uncertainty can also explain the subsequent increments in responding (stepwise or gradual) as conditioning continues. We think that decreasing uncertainty about the true values of the estimated rates of reinforcement is also likely to be an important part of the explanation for the decreasing within-subject variability in response rates.

DISCUSSION

The present results have provided several important pieces of evidence about the nature of the conditioned response acquired by rats in an appetitive Pavlovian paradigm. The first is that the rate of learning, defined as the reciprocal of the number of trials needed to acquire a response to the CS, was determined by the CS's informativeness (equal to the C/T ratio). Neither the rate at which the CS itself was reinforced, $1/T$, nor the spacing of the trials, C, had any independent effect after partialling out the effect of C/T. This is the same conclusion reached by Gibbon and Balsam (1981) in their meta-analysis showing scalar invariance in the acquisition of autoshaped key-pecking by pigeons. Contrary to what Gibbon and Balsam supposed, but in agreement with what Jenkins *et al*. (1981) showed, the dependence of trials to acquisition on C/T extends all the way to the abscissa, that is, when informativeness is very high acquisition can occur after just 1 reinforcement.

The second finding is that, when each CS is reinforced so that reinforcement rate scales inversely with T, the response rate after extended conditioning is directly proportional to $1/T$ (but not C or C/T). More specifically, the CS response rate is a scalar function of the CS reinforcement rate (red symbols in Figure 3). Related to this is the third, and unexpected, finding that the overall response rates across the entire trial (comprising the CS and ITI) is proportional to the contextual rates of reinforcement, $1/C$ (black symbols in Figure 3). To the best of our knowledge, the dependence of the overall rate of responding on the *contextual* rate of reinforcement has not been previously reported. It is particularly noteworthy that the constant of proportionality relating the overall response rates to contextual reinforcement rates matches that relating the CS response rates to CS reinforcement rates. Thus, there is a single scalar relationship between response rates and reinforcement rates that spans several orders of magnitude.

The fourth finding is that the post-acquisition response rate trends linearly upward to a peak. The higher the rate of reinforcement of the CS, the steeper that linear increase. However, the path to the peak rate has several ups and downs, particularly when informativeness is low, and the terminal rate can be lower than the peak rate. When informativeness is high, the response rate generally rises to the peak in one or two steps. The first and often final step often occurs after the first trial.

The previous published evidence concerning the relationship between *C*/*T* and acquisition of Pavlovian conditioning with rodents has been mixed. Lattal (1999) and Holland (2000) reported evidence that conditioning in rats was related to *C*/*T*, as did Ward *et al*. (2012) in a series of experiments with mice. However, Kirkpatrick and Church (2000), and, more recently, Thrailkill *et al*. (2020), found no evidence for an effect of *C*/*T* on trials to acquisition in rats. Moreover, Lattal, Holland, and Thrailkill *et al*. all found evidence for an effect of *T* when *C*/*T* was held constant. We suggest that these inconsistencies relate to differences in how response acquisition was measured across the studies. Our results show that the point at which evidence for conditioned responding first emerges is directly related to *C*/*T*, and not to *C* or *T* alone, but the level of responding subsequently acquired is directly related to *T* and not to *C*/*T* or *C*. Therefore, both *C*/*T* and *T* will affect any index of conditioning that is sensitive to both the time when responding emerges and how much responding is subsequently acquired. Lattal's evidence for an effect of *T* between groups matched on *C*/*T* was obtained in a test session conducted after 4 conditioning sessions totalling 48 trials, and Holland's evidence came from the last 8 of 16 conditioning sessions. In both cases, the observed effects of *T* could have been due to its effect on the level of responding acquired rather than how quickly responding emerged. A similar argument can be made for the evidence provided by Thrailkill *et al*., based on the evidence that our rats took substantially longer to reach the Thrailkill *et al*. criterion for response acquisition than to reach the criterion we developed (see Supplementary Materials, pages 7-8). This suggests that, to satisfy their criterion for acquisition, the rats acquired a higher level of responding which would have been affected by *T*. Finally, the absence of evidence for an effect of *C/T* in the study by Kirkpatrick and Church may also have been due to the particular method they used to assess trials to acquisition. Their method, when applied to the present data (see Supplementary materials, pages 8-9), produced smaller differences between groups and the correlation between trials to acquisition and log(*C/T*) was not significant. Thus, their method may have been less sensitive to differences between groups in their rate of acquisition, which could explain why they failed to see an effect of *C/T* across the relatively limited range of *C*/*T* ratios they tested (from 1.5 to 12).

The results shown in Figure 3 confirm that the subjects compute the contextual rate of reinforcement. That is, the rats' overall rate of responding across the trial scales with the

overall rate of reinforcement with exactly the same constant of proportionality as that between responding and reinforcement during the CS. It determines how aroused they are when in that context. It has long been recognized that the contextual rate of reinforcement in appetitive operant conditioning has a scalar effect on foraging activities (Belke, 1992; Drew, Zupan, Cooke, Couvillon, & Balsam, 2005; Killeen, 2023; Killeen, Hall, & Bizo, 1999; Killeen, Hanson, & Osborne, 1984). The effect of the operant contingencies is to channel that activity into the activity or activities on which reinforcement is contingent (Gallistel & Shahan, 2024). It now appears that the same is true in Pavlovian conditioning. Reinforcement is contingent on poking into the magazine, and the rats' poke rate is 23 times the reinforcement rate.

The contextual rate of reinforcement plays a fundamental role in Rate Estimation Theory (RET, Gallistel & Gibbon, 2000). It is the first term in the vector of *uncorrected* rates. The uncorrected vector is the list of the *observed* rates of reinforcement for each possible predictor. The *corrected* rate vector is the list of reinforcement rates subjects *ascribe* to the actions of the different predictors. The matrix equation that does the ascription is based on the assumption they act independently, in which case the ascribed rates must sum to the observed rates when the predictors co-occur. The contextual rate of reinforcement also plays a fundamental role in the information-theoretic model of acquisition (Balsam, Fairhurst, & Gallistel, 2006; Balsam & Gallistel, 2009; Ward et al., 2012). In that model, the appearance of differential responding to the CS is determined by the informativeness of the protocol, which is the ratio between the CS reinforcement rate and the contextual reinforcement rate. The information transmitted to a subject by CS onset is the log of the informativeness.

Information-Theoretic Contingency

Contingency in Pavlovian and operant conditioning has long resisted a mathematical definition that made it measurable in all circumstances, particularly when there is no time at which reinforcement may be anticipated hence no time at which failures of reinforcement to occur can be counted (Donahaoe, 2006; Gallistel, 2021; Gibbon, Berryman, & Thompson, 1974; Granger & Schlimmer, 1986; Hallam, Grahame, & Miller, 1992; Hammond & Paynter, 1983). Gallistel and Latham (2023) have developed a generally applicable measure based on the trivially computable prospective and retrospective mutual information between CSs and reinforcements or between responses and reinforcements. Given two distinguishable event streams X and Y—for example a stream of CS onsets and a stream of reinforcements at CS

termination—there is prospective mutual information between the x events and the y events, $I(\overrightarrow{X};Y)$, when the expected wait to the next y, conditional on an x, ($\mu_{\mathrm{y}}|\vec{x}$), is reliably shorter than the expected wait between the y's (μ_{y}), as defined in Equation (1). There is retrospective mutual information, $I(X;\overleftarrow{Y})$, when the expected wait looking back from a y to the most recent x, ($\mu_{\mathrm{x}}|\overleftarrow{y}$), is shorter than the expected wait between the x's (μ_{x}), as in Equation (2):

$$I(\overrightarrow{X};Y) = \log\frac{\mu_{\mathrm{y}}}{\mu_{\mathrm{y}}|\vec{x}} = \log\frac{\lambda_{\mathrm{y}}|\vec{x}}{\lambda_{\mathrm{y}}} \tag{1}$$

$$I(X;\overleftarrow{Y}) = \log\frac{\mu_{\mathrm{x}}}{\mu_{\mathrm{x}}|\overleftarrow{y}} = \log\frac{\lambda_{\mathrm{x}}|\overleftarrow{y}}{\lambda_{\mathrm{x}}} \tag{2}$$

The arguments of the log function in Equations (1) and (2) are the ratio of the *unconditional wait* (in the numerator) to the *conditional w*ait (in the denominator), or equivalently, the ratio of the inverse of the waits, *conditional rate* in numerator and *unconditional rate* in denominator. Because its logarithm is the mutual information, Balsam *et al*. (2006) have termed this ratio the *informativeness* of a Pavlovian protocol.

Gallistel and Latham (2023) define *contingency* as the ratio of the *mutual information* to the *available information*, which is the amount that reduces subjective uncertainty to 0. Because measurement error scales with latency (Weber's Law, see Gibbon, 1977a), contingency equals one only when the x's and y's coincide. How to measure the available information is often unclear. Mutual information, however, is trivially computed, as in Equations (1) and (2).

Another information-theoretic measure, $n\mathrm{D}_{\mathrm{KL}}$, measures the degree to which computed mutual information may be trusted. Its role, relative to mutual information, is analogous to the role of *p* values relative to correlations. The D_{KL} in the $n\mathrm{D}_{\mathrm{KL}}$ is the Kullback-Leibler *divergence*. It is analogous to the effect size in conventional statistics. The effect size is the normalized *distance* between two distributions assumed to have the same variance. Distance is symmetric, i.e., D(X,Y) = D(Y,X), but divergence is not: $\mathrm{D}_{\mathrm{KL}}(X||Y) \neq \mathrm{D}_{\mathrm{KL}}(Y||X)$. The asymmetric information-theoretic divergence is arguably the better measure, because the amount of data required to determine whether an X distribution differs from a Y distribution is not the same as the amount required to determine whether the opposite is

true—and rats and mice are sensitive to this asymmetry (Kheifets, Freestone, & Gallistel, 2017; Kheifets & Gallistel, 2012).

The $\mathrm{D_{KL}}$ of one exponential distribution from another depends only on their rate parameters:

$$\mathrm{D_{KL}}(X||Y)_{\mathrm{exp}} = \ln\frac{\lambda_{\mathrm{X}}}{\lambda_{\mathrm{Y}}} + \frac{\lambda_{\mathrm{Y}}}{\lambda_{\mathrm{X}}} - 1 \tag{3}$$

The uncertainty regarding the values of estimates for λ_{X} and λ_{Y} depends on the sample sizes, n_{X} and n_{Y}. Peter Latham (Gallistel & Latham, 2023, see their Appendix) has recently shown that when $\lambda_{\mathrm{X}} = \lambda_{\mathrm{Y}}$ (the null hypothesis),

$$n_{\mathrm{e}}\mathrm{D_{KL}}(X||Y) \sim \Gamma\big(n_{\mathrm{p}}/2\,,1\big), \tag{4}$$

where n_{e} *is* the *effective* sample size; $n_{\mathrm{e}} = n_{\mathrm{X}}/(1 + n_{\mathrm{X}}/n_{\mathrm{Y}})$; and n_{p} is the size of the parameter vector. For the exponential, $n_{\mathrm{p}} = 1$. Therefore, as described earlier in Equation (2),

$$n\mathrm{D_{KL}}(X||Y)_{\mathrm{exp}} = \frac{n_{\mathrm{X}}}{1+n_{\mathrm{X}}/n_{\mathrm{Y}}}\left(\ln\frac{\lambda_{\mathrm{X}}}{\lambda_{\mathrm{Y}}} + \frac{\lambda_{\mathrm{Y}}}{\lambda_{\mathrm{X}}} - 1\right) \tag{5}$$

Equation (5) measures how "significant" an observed amount of mutual information is, how unlikely it is to have arisen by chance, and, by Equation (4), $n\mathrm{D_{KL}}(X||Y)_{\mathrm{exp}} \sim \Gamma(0.5,1)$, $n\mathrm{D_{KL}}$'s, so it may be converted to the more familiar *p* values. Both are measures of the strength of statistical evidence. Conversion is motivated only by sociological considerations.

Equations (1, 2 and 5) and the RET equation—$\hat{\boldsymbol{\lambda}}_{\mathrm{R}} = \mathbf{T}^{-1}\boldsymbol{\lambda}_{\mathrm{R}}^{\mathrm{T}}$ (Gallistel, 1990)—constitute a computationally simple, parameter-free model of associative learning. Equation (2) enables us to address the question, How sensitive are subjects to the strength of the evidence for observed mutual information? Put another way, Does their behavioural sensitivity to the accumulating mutual information suggest that they make a rational assessment of the extent to which they can trust the mutual information so far observed? In the next section, we propose a simple generative model of learning that incorporates answers to these questions.

Towards a Generative Model That Explains the Learning Rate Law

Our results confirm a theoretically important law: The median learning rate in a group, operationally defined as the reciprocal of the median number of reinforcements prior to the

appearance of the conditioned response, is an approximately scalar function of the protocol's informativeness, ι, that is, of the ratio between the rate of reinforcement expected during the CS and the rate expected in the context in which the CS is presented. This law is descriptive; it does not model an hypothesized process that generates the behavioural data. As such, the law stands in contrast to formal models of the associative process, such as the well-known Rescorla-Wagner model (1972) and its many descendants, that do propose a generative mechanism of the learning process. However, any generative model must explain the law described here. That necessity poses a seemingly insurmountable challenge to models in which reinforcements strengthen an associative bond (or augment a prediction) and non-reinforcements weaken the bond (or the prediction). The challenge arises from the fact that, when informativeness is substantial (≥ 5), a substantial fraction of subjects (23% or more) begin to respond appropriately to the CS after a single reinforcement (Figure 2D; see also Figure S2D in Supplementary Materials). By contrast, when informativeness is low (≤ 3), no subject begins to respond after only one reinforcement, and almost all subjects take at least 10 reinforcements to begin responding.

After a single reinforcement, both the CS and the context have been reinforced once. Delta-rule, or prediction-error, models for the updating of an hypothesized associative bond or prediction are event driven: the bond/prediction is augmented when a reinforcement occurs and decremented when an expected reinforcement fails to occur. The challenge is to explain why, when informativeness is high, the association/prediction between CS and reinforcement is already strong enough to be acted upon by many subjects after the first reinforcement. Decomposing the interval prior to first CS into hypothesized "trials" that weaken the context's association with reinforcement (cf. Rescorla & Wagner, 1972, p 88) cannot offer an explanation because there can be no expectation of reinforcement prior to its first occurrence; failures of reinforcement cannot be imagined to have occurred during imagined trials prior to the first reinforcement. By contrast, when informativeness is low, a strong majority of subjects will not have shown evidence of selective responding to the CS even when the first 10 reinforcements have all occurred during the CS. A generative model must explain both extremes of the learning rate law.

A model of the behaviour-generating process must also generate the distributions of reinforcements to acquisition, not simply their central tendencies. Looking at Figure 2, we can

see that, as informativeness increases, more and more data fall at the analytic lower limit of one reinforcement. Thus, the "central" tendency is not in fact central; in the limit as informativeness becomes large, all the data fall at or above it. Statistically speaking, the distributions of reinforcements to acquisition become strongly left skewed, with a mode at 1, the smallest possible value. The increasing left skew as ι increases in the distributions shown in Figure 2D is ascribable first to the analytic fact that statistically significant behavioural evidence that an association has been perceived can only exist when the number of reinforcements is ≥ 1. It is ascribable second to the empirical fact that the ratio between the CS response rate and the contextual rate goes to 1 as the ι goes to 1. That is, the rate at which the behavioural evidence accumulates goes to 0 as the CS and contextual rates of reinforcement approach parity, which is the truly random control condition (Rescorla, 1967).

The distributions plotted in Figure 2D suggest to us a generative model incorporating an information-theoretic learning rule and performance function. In such a model, subjects perceive an association between the CS and reinforcement as soon as there is non-trivial evidence for it. The strength of that association is measured by the CS's informativeness, ι_{CS}. A stochastic performance function then maps the strength of this perceived association into observable behaviour. The performance function can delay the appearance of measurable behavioural change, with the length of the delay decreasing as a monotonic function of the strength of the perceived association.

In the simplest model that predicts scaling between learning rate and informativeness, the probability that a subject changes its rate of responding after CS onset, $p(\Delta\lambda_{\mathrm{r}}|\mathrm{CS}\uparrow)$, as a function of the CS's informativeness, is described by the exponential cumulative distribution function

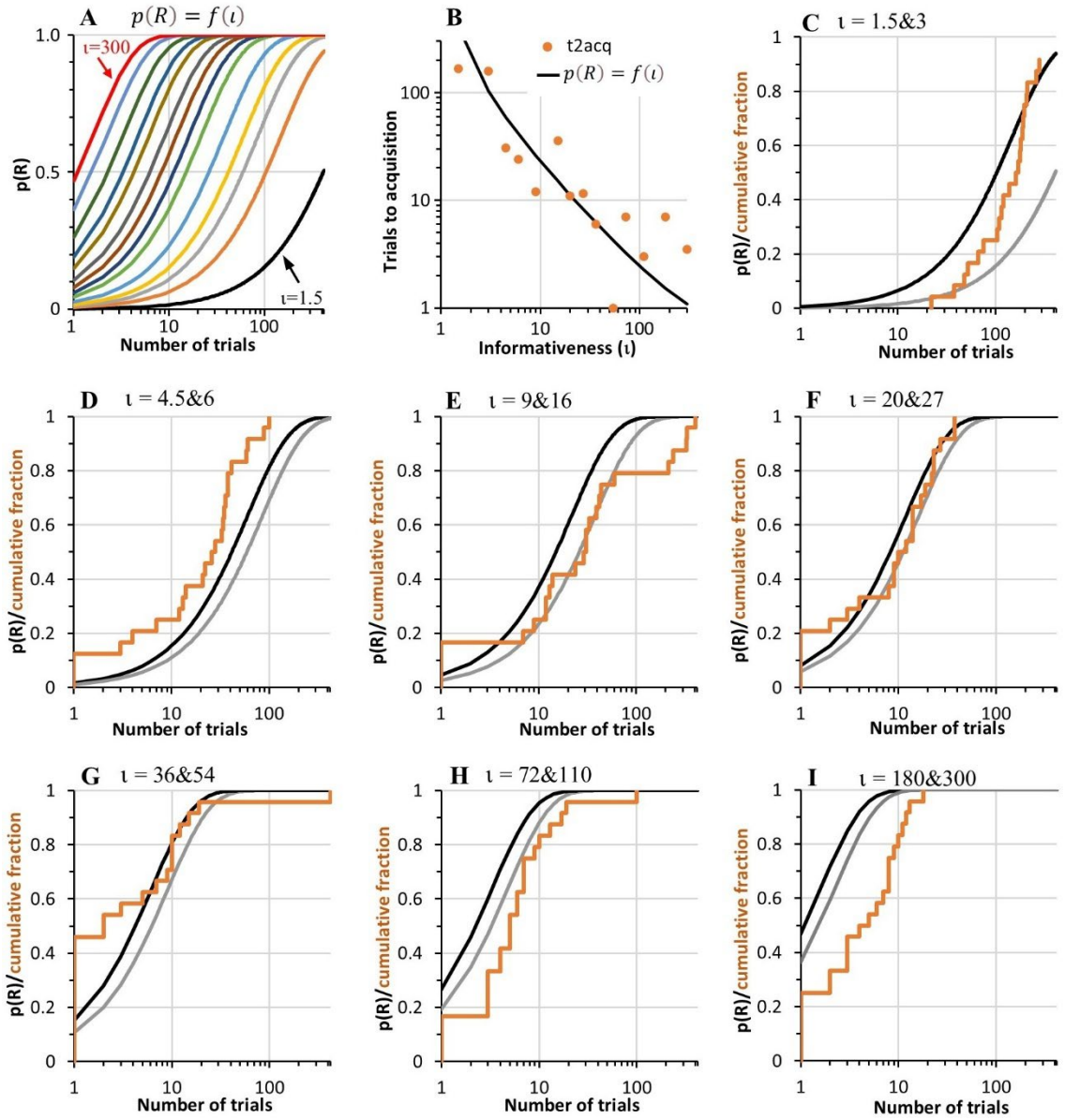
$$p(\Delta\lambda_{\mathrm{r}}|\mathrm{CS}\uparrow) = 1 - \mathrm{e}^{-k\cdot(\iota_{CS}-1)}\,. \tag{6}$$

This function describes how the probability of a response increases as ι_{CS} increases, and thus the median number of trials to see a response decreases with increasing ι_{CS}. As the contingency between CS and reinforcement goes to zero and ι_{CS}–1 goes to zero, $p(\Delta\lambda_{\mathrm{r}}|\mathrm{CS}\uparrow)$ goes to 0. In Figure 5A, the cumulative probability of seeing at least one response across trials is plotted for each of the 14 levels of informativeness in the current experiment, using k = 1/297 from the regression equation in Figure 1. The horizontal distance between the curves corresponds to the effect of informativeness on the number trials required to reach a

response criterion. For example, $p(\Delta\lambda_r|\mathrm{CS}\uparrow)$ = 0.5 after 1 trial when ι = 300 but after about 100 trials when ι = 3. These values, corresponding to the number of trials for the median subject to make at least one response, are plotted as a black line in Figure 5B. This is equivalent to the regression in Figure 1 with slope -1 in the loglog domain. In this model, the appearance of conditioned responding is delayed when ι is low because the probability of responding on any trial is low.

In the remaining panels in Figure 5 (C to I), the modelled probability of a response at different levels of informativeness is compared with the data from Figure 2 showing the proportion of rats that had acquired responding to the CS at the same informativeness, based on the criterion that the CS response rate was permanently greater than the overall response rate. In most cases the modelled probabilities of a response correspond well with the observed proportion of rats that had acquired a response. Thus, the simple generative model of responding in Equation 6, with only one parameter that we have fixed at k = 1/297 based on data reported previously by Gibbon & Balsam (1981), can explain with remarkable success the learning rates of the 14 groups in our experiment, predicting not only the central tendencies of the groups but also their distributions.

Figure 5. Modelling response probability across trials as a function of informativeness.



Note. ***A:*** *The cumulative probability of a response, p(R), calculated using Equation 6, as a function of number of reinforced trials. Each curve plots p(R) for one of the 14 levels of informativeness used in the current experiment, ι = 1.5 to 300.* ***B:*** *The black line plots the number of trials taken for p(R) to reach 0.5 as a function of the same 14 levels of informativeness. The plot includes the median trials to acquisition (t2acq) for the 14 groups in the current experiment, replotted from Figure 2A.* ***D to I:*** *The cumulative fraction of rats that had acquired responding (orange lines) for pairs of groups with similar informativeness, replotted from Figure 2D. The black and grey curves plot the cumulative p(R) using the same two values of ι (the black line is based on the higher of the two values).*

Conclusion

Associative learning obeys simple equations that map from measurable properties of subjects' experience to measurable properties of their behaviour. The equations have at most one free parameter, and it is a scale factor with a data-anchored interpretation. According to one of these rules, the response rate is 23 times the reinforcement rate in the median subject (Figure 3) over a range that covers 3 orders of magnitude. The variability about this regression is also multiplicative, with a scale factor of $10^{.39}$ = 2.5. A second rule is the simple scalar relation between the learning rate (the reciprocal of reinforcements to acquisition) and the informativeness of a Pavlovian protocol. In its simplest form, the log of the learning rate is -1 times the mutual information; the constant, *k*, in this regression is the informativeness that produces 1-trial acquisition (Figures 1 and 2). A third example is the parameter free rate estimation equation. In a 1-CS protocol, that equation is:

$$\begin{bmatrix} \hat{\lambda}_{\mathrm{R|C\&\sim CS}} \\ \hat{\lambda}_{\mathrm{R|CS}} \end{bmatrix} = \begin{bmatrix} 1 & 1/\iota \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \lambda_{\mathrm{R|C}} \\ \lambda_{\mathrm{R|C\&CS}} \end{bmatrix} \tag{7}$$

Equation (7) maps the *observed* contextual rate of reinforcement, $\lambda_{\mathrm{R|C}}$, and the *observe*d rate of reinforcement when the CS is also present, $\lambda_{\mathrm{R|C\&CS}}$, to the rates of reinforcements that subjects *ascribe* to these predictors, $\hat{\lambda}_{\mathrm{R|C\&\sim CS}}$ and $\hat{\lambda}_{\mathrm{R|CS}}$. It does so by way of the inverse of a simple matrix. The only element not equal to 1 in the matrix in Equation (7) is the inverse of the informativeness, which is also the protocol parameter in the learning-rate equation. Equation (7) explains the results in the cue-competition literature (a.k.a. the assignment-of-credit literature). (For review and comparison to the Rescorla-Wagner model, see Gallistel, 1990, Chapter 13).

These equations bring out the formal structure of the *data* from associative learning experiments. They map an always defined and computable aspect of subjects' experience—*rate* of reinforcement—to measurable properties of subjects' behaviour. They contrast with equations that map from the often-undefinable *probability* of reinforcement to a *hypothetical construct* like associative strength or expected value (Honey, Dwyer, & Iliescu, 2020; Ludvig, Sutton, & Kehoe, 2012; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Sutton & Barto, 1990; Vogel, Ponce, & Wagner, 2019).

That the rates of reinforcement rather than the probabilities are the inputs to the equations that map perceived associations to behavior calls into question the basic assumption in the common understanding of associative processes, the assumption that associative learning is mediated by the incrementing and decrementing of some scalar brain quantity (e.g., the conductivity of a plastic synapse) by reinforcement and non-reinforcement, respectively. This assumption is central to all models of Pavlovian and operant learning that map *probability* of reinforcement to associative strength and is central to all reinforcement learning models that map it to a weight state in a neural net (supervised and unsupervised learning) or to an expected long-term reward (reinforcement learning). They all fail to "leverage the statistical and computational structure of [the] problem" (Russo, Roy, Kazerouni, Osband, & Wen, 2018, p. 6), because they must discretize time into unobserved trials or states with unspecified durations (Namboodiri, 2022). They must do this because, unlike a rate that has a temporal unit, a probability is unitless.

A probability is the ratio between a count of reinforcements and the sum of the count of reinforcements and non-reinforcements. It is impossible to map from probability of reinforcement to behaviour in the general case because the non-reinforcements in the denominator are unobserved events with no physical properties, and, in the general case, often uncountable. They are countable only when reinforcement fails to occur at an expected time of reinforcement. In our protocol and many others the reinforcements occur at an unpredictable time or times following CS onset. The durations of our CS were drawn from a uniform distribution and the reinforcements coincided with its termination. In Rescorla's (1967, 1968) experiments with truly random controls (and the many follow-ons), the reinforcements were programmed by Poisson processes. The defining feature of a Poisson process is its flat hazard function: there is no moment at which a reinforcement is any more likely than at any other moment.

Unlike models that take probability of reinforcement as the essential aspect of subjects' experience, all of which have at least two free parameters, the model of associative learning provided by the simple equations we present here directly explains quantitative facts about associative learning that have gone unexplained for decades: One such fact is that reinforcements to acquisition are unaffected by partial reinforcement (Balsam & Gallistel, 2009; Gallistel, 2003; Gallistel, Craig, & Shahan, 2014; Gibbon, Farrell, Locurto, Duncan, &

Terrace, 1980; Gottlieb, 2004, 2005). This follows immediately from the equation for the learning rate as a function of informativeness (top of Figure 1), because partial reinforcement does not alter informativeness. A second such fact is that deleting reinforced trials while retaining the spacing of the remaining reinforced trials does not alter the temporal progress of acquisition (Bouton & Sunsay, 2003; Gottlieb, 2008). The number of reinforced trials in a given amount of training time is irrelevant, given only that there is at least one. This, too, follows from the equation that relates the learning rate to the informativeness or to its log, the mutual information (Figures 1 and 2). Because the slope on a loglog plot is essentially –1 over most of the useable range of values for informativeness, halving the number of trials doubles the informativeness and that doubles the learning rate.

We conclude that models of associative learning based on probability of reinforcement cannot be correct because the rate of responding, the learning rate and the assignment of credit are simple functions of rate of reinforcement, a quantity with temporal units. Only models that leverage the metric temporal structure of subject's experience can be neurobiologically realisable.

METHODS

Subjects

A total of 176 experimentally-naive female albino Sprague Dawley rats (8 to 10 weeks of age) were obtained from Animal Resources Centre, Perth, Western Australia. They were housed in groups of 4 in split-level ventilated plastic tubs (Techniplast™), measuring 40 x 46 x 40cm (length x width x height), located in an animal research facility at the University of Sydney. They had unrestricted access to water in their home tubs. Three days before commencing the experiment, they were placed on a restricted food schedule. Each day, half an hour after the end of the daily training session, each tub of rats received a ration of their regular dry chow (3.4 kcal/g) equal to 5% of the total weight of all rats in the tub. This amount is approximately equal to their required daily energy intake (Rogers, 1979), and took at least 2 h to be eaten (but was usually finished within 3 h). Rats on this schedule do not typically lose weight (and never more than 10%) but gain weight only very slowly. All experimental procedures were approved by the Animal Research Authority of the University of Sydney (protocol 2020/1840).

Apparatus

Rats were trained and tested in 32 Med Associates™ conditioning chambers distributed equally across four rooms. Twenty-four chambers (Set A) measured 28.5 x 30 x 25 cm (height x length x depth) and the other eight (Set B) were 21 x 30.5 x 24 cm (height x length x depth). Each chamber was individually enclosed in a sound- and light-resistant wooden shell (Set A) or PVC shell (Set B). The end walls of each chamber were made of aluminum; the sidewalls and ceiling were Plexiglas™. The floor consisted of stainless-steel rods, 0.5 cm in diameter, spaced 1.5 cm apart. Each chamber had a recessed food magazine in the center of one end wall, with an infra-red LED and sensor located just inside the magazine to record entries by the rat. A small metal cup measuring 3.5 cm in diameter and 0.5 cm deep was fixed on the floor of each food magazine either in the center (Set A) or offset to the left of center (Set B). Attached to the food magazine was a dispenser delivering 45-mg food pellets (purified rodent pellets; Bioserve, Frenchtown, NJ). Illumination of an LED (Med Associates product ENV-200RL-LED) mounted in the ceiling of the magazine served as the CS. Experimental events were controlled and recorded automatically by computers and relays located in the same

room. Throughout all sessions, fans located in the rear wall of the outer shell provided ventilation and created background noise (between 61 and 66 dB, depending on the chamber).

Procedure

The rats were allocated to the 14 groups shown in Table 1. The experiment was run with three separate cohorts of 64, 80, and 32 rats, so that by the end of the experiment each of the 14 groups had at least 12 rats (Groups 6 and 10 had 16 rats, as described next). In the first cohort, the data for 4 rats in each of Groups 6 and 10 were not saved in Session 1 due to human error. These 8 rats continued through the entire experiment but only their data for the final 5 sessions were used (to calculate their terminal response rate). An extra 4 rats were run in both groups in the second cohort.

The rats were not given magazine training. The experiment commenced with the first conditioning session and continued for 42 sessions over 42 consecutive days. Within each group, the CS-US interval varied from trial-to-trial according to a uniform distribution centred on the value of T for that group (the interval varied from a minimum of 2 s to a maximum of $2xT$–2 s). The ITI in each group also varied from trial-to-trial as a uniform distribution with a minimum of 15 s. Between groups, T varied from 6 s to 62 s, and C varied from 63 s to 4,200 s (70 min) as summarised in Table 1. The combinations of C and T gave rise to 14 distinct C/T ratios that were approximately evenly distributed on a log scale ranging from a ratio equal to 1.5 (63/42) up to 300 (4200/14). For the first 11 groups (those with ratios from 1.5 to 72), there were 10 trials per session; for the remaining three groups (with C/T ratios of 110, 180 and 300), each session contained only 3 trials so that the total session time remained within a manageable length (less than 4 h). All groups were trained with one session per day for a total of 42 sessions. Photo-beam interruptions by entry into the magazine were recorded during each CS and each ITI (recorded during the 10-s period immediately before CS onset).

Data analysis.

Several different indices were used to identify when responding to the CS first appeared. The first index involved creating, for each rat, cumulative records of response counts during the CS and overall response rates (across the CS and ITI), then converting these into cumulative records of response rates by dividing the cumulative count at each trial by

the cumulative time (CS or CS+ITI) at that trial. Since the response rate during the CS should not differ from the overall rate before the rat has learned that the CS signals food, we identified the point of acquisition as the trial, t_i, for which the cumulative response rate to the CS after that trial (starting from t_i + 1) was permanently greater than the cumulative overall response rate starting from t_i + 1. This trial is plotted for each rat in Figure 2A. Next, we used a new information-theoretic statistic, the $n\mathrm{D_{KL}}$ (see next section), to identify when, starting from trial t_i + 1, the cumulative rate of responding during the CS began to exceed the overall rate by a statistical threshold. Figures 2B shows the trials to acquisition when the computed $n\mathrm{D_{KL}}$ exceeded a threshold of 0.82 which correspond to odds of 4:1 in favour of a difference between the CS and ITI response rates. Figure 2C shows the trials to acquisition when the $n\mathrm{D_{KL}}$ exceeded 1.92, corresponding to probability $p < .05$ that the CS and ITI response rates are the same. Additional analyses were run adopting measures used in previous studies (Kirkpatrick & Church, 2000; Thrailkill et al., 2020) to identify trials to a learning criterion. These analyses and results are described in the Supplementary Materials (pages 170-172).

Analyses were also conducted to assess how *T*, *C*, and *C*/*T* affect conditioned responding after the point when it first emerges. These analyses examined how the response rate increased over trials and what level of responding was reached after extended training. The level of responding ultimately acquired by each group was computed as the mean response rate over the last 5 conditioning sessions. Further analyses broke down the response rate into separate components: the latency to first response; the mean duration of each response (time in the magazine); and the interval between responses (time out of the magazine). The increase in responding over sessions was assessed using a method described by Harris (2022) that uses the slope of the cumulative response record to identify the number of trials required for responding to reach successive response milestones corresponding to each decile of the rat's peak response rate. Based on findings relating response rates to reinforcement rates (1/*T*) using this paradigm (Harris & Carpenter, 2011), we hypothesised that responding at the end of conditioning would be related to *T*, and indeed would scale linearly with log(*T*), but would not be related to *C* or *C*/*T*. We had no clear hypotheses about whether *T*, *C*, or *C*/*T* would affect how quickly responding increases after it has emerged.

The Kullback-Leibler divergence, and the $n\mathrm{D_{KL}}$.

Comparing the CS poke rate to the ITI poke rate reinforcement-by-reinforcement is problematic in cases where one or both rates are undefined because the subject has not yet made a poke. When a subject has made 5 pokes during the first 2 CSs and no pokes during the much longer ITIs, one does not want to conclude that the subject has not yet acquired a conditioned response to the CS. A second problem when using conventional statistics like the *t*-test is that one is required to specify in advance the sample size.

We circumvented these difficulties by reformulating the null hypothesis as a comparison between the rate of responding during the CS and the contextual rate of responding across the whole trial (during the ITI and CS) by using an information-theoretic statistic to measure the strength of the evidence that the CS rate of responding differs from the contextual rate (Gallistel & Latham, 2023). The information theoretic statistic measures the strength of the evidence against the hypothesis that the distribution of inter-poke intervals during CSs is the same as the distribution in the context overall. Put another way, the null hypothesis is that poking during CSs does not differ from the poking expected because the subject pokes into the magazine in response to the fact that pellets sometimes drop there without regard to the signal value of the CS.

The CS poke rate, $\lambda_{\mathrm{r|CS}}$, at any point in training is the number of pokes made during the CSs divided by their cumulative duration: $\lambda_{\mathrm{r|CS}} = n_{\mathrm{r|CS}}/D_{\mathrm{CS}}$. And likewise for the contextual poke rate: $\lambda_{\mathrm{r|C}} = n_{\mathrm{r|C}}/D_{\mathrm{C}}$. In these calculations, D_{CS} is the cumulative duration of the CS and D_{C} is cumulative training time. Assume for example that cumulative CS time as of the 2nd reinforcement is 20 s, total training time is 1000 s, and the subject has made 5 pokes during the two CS intervals and no pokes during the ITIs. Then, $\lambda_{\mathrm{r|CS}} = 5/20 = .25/\mathrm{s}$ and $\lambda_{\mathrm{r|C}} = 5/1000 = .005/\mathrm{s}$. The ratio of the two estimates is .25/.005 = 50:1, that is, the subject's observed poke rate at this point in training is 50 times faster during the CSs than would be expected given the estimate of how frequently it pokes in the training context. This discrepancy is unlikely to have arisen by chance.

An information theoretic statistic, the $n\mathrm{D_{KL}}$, measures the unlikeliness (Gallistel & Latham, 2023). The $\mathrm{D_{KL}}$ is the Kullback-Leibler divergence, a measure of the distance between two distributions. The divergence of one exponential distribution from another depends only

on the rate parameters of the distributions. Therefore, as per Equation (3), the divergence of $\lambda_{\mathrm{r|CS}}$ from $\lambda_{\mathrm{r|C}}$ is

$$\mathrm{D_{KL}}\left(\lambda_{r|CS}||\lambda_{r|C}\right)_{exp} = log_e\left(\frac{\lambda_{\mathrm{r|CS}}}{\lambda_{\mathrm{r|C}}}\right) + \frac{\lambda_{\mathrm{r|C}}}{\lambda_{\mathrm{r|CS}}} - 1$$

It is the information-theoretic measure of the extent to which an exponential distribution with rate parameter $\lambda_{\mathrm{r|CS}}$ diverges from an exponential distribution with rate parameter $\lambda_{\mathrm{r|C}}$. The divergence is, roughly speaking, the equivalent of the effect size in a conventional analysis. The equivalence is rough because the effect size—the normalized distance between the means—is symmetric, whereas the divergence is not: $\mathrm{D_{KL}}\left(\lambda_{\mathrm{r|CS}}||\lambda_{\mathrm{r|C}}\right) \neq \mathrm{D_{KL}}\left(\lambda_{\mathrm{r|C}}||\lambda_{\mathrm{r|CS}}\right)$.

The $n\mathrm{D_{KL}}$ measures the additional cost of encoding n data drawn from the exponential distribution with parameter $\lambda_{\mathrm{r|CS}}$ on the assumption that they come from an exponential distribution with rate parameter $\lambda_{\mathrm{r|C}}$ (Cover & Thomas, 1991). It multiples the $\mathrm{D_{KL}}$ by the effective sample size, $n = n_{\mathrm{r|CS}}/\left(1 + n_{\mathrm{r|CS}}/n_{\mathrm{r|C}}\right)$. Therefore, as per Equation (5),

$$n\mathrm{D_{KL}}\left(\lambda_{r|CS}||\lambda_{r|C}\right) = \frac{n_{\mathrm{r|CS}}}{1 + n_{\mathrm{r|CS}}/n_{\mathrm{r|C}}}\left(log_e\left(\frac{\lambda_{\mathrm{r|CS}}}{\lambda_{\mathrm{r|C}}}\right) + \frac{\lambda_{\mathrm{r|C}}}{\lambda_{\mathrm{r|CS}}} - 1\right).$$

When there is no divergence, the $n\mathrm{D_{KL}}$ is distributed gamma(.5,1) (for proof, see Appendix in Gallistel & Latham, 2023). Thus, we can convert the information-theoretic measure of the strength of the evidence for divergence to the more familiar *p* value measure.

In the illustrative example, $\mathrm{D_{KL}}(.25||.005) = 2.93$ nats[1], and n = 5/(1+5/5) = 2.5. Therefore, the $n\mathrm{D_{KL}} = 2.5 \times 2.93 = 7.32$ nats and $1 - \mathrm{gamcdf}(7.32, .5, 1) \cong .0001$. The evidence against the null hypothesis is very strong; the odds against the assumption that all 5 pokes have occurred only during the CSs only by chance are on the order of 10,000 to 1. This example illustrates one approach to the statistical comparison of poke rates following each of the first few variable duration CSs, each terminating in a reinforcement.

[1] The nat is a unit of information equal to the base of the natural logarithm. The Kullback-Leibler divergence must be computed using the natural logarithm, but the result in nats may be converted to bits by multiplying by $\log_2 e$ = 1.44.

Parsing response rates. Another approach to estimating the onset of conditioned poking comes from parsing the CS and ITI inter-poke interval vectors into segments with significantly different rate parameters. Parsing of the inter-poke interval vectors was also motivated by a desire to capture differences between subjects in the often bumpy evolution of their post-acquisition rate of responding.

Our parsing algorithm recursively extends the length, n_{e}, of the vector of inter-poke intervals one interval at a time. After each extension, it compares the rate estimate for each successively longer sub-sequence to the rate estimate for the full sequence, using the $n\mathrm{D}_{\mathrm{KL}}$ statistic. These comparisons generate the function $n\mathrm{D}_{\mathrm{KL}}(n_{\mathrm{s}})$ for $n_{\mathrm{s}} \leq n_{\mathrm{e}}$. Whenever $\max[n\mathrm{D}_{\mathrm{KL}}(n_{\mathrm{s}})\ ] > c$, the parser truncates the inter-poke interval vector at the location of the maximum. The value for the decision criterion, c, is user supplied and generally falls between 2 and 6 nats. The rate estimate for the segment truncated is the number of pokes up to the truncation divided by the sum of the intervals up to the truncation. The algorithm operates recursively on the post-truncation portions of the vector until there is no significant sub-sequence. The results for values of 2, 4 and 6 nats, corresponding to *p* values .05, .005 and .0005, are contained in an Excel™ file, ParseTable.xlsx, that can be downloaded from https://osf.io/vmwzr/ along with the Matlab™ code for the custom expparser.m function. Because parsing inescapably involves multiple comparisons, highly conservative decision criteria are generally to be preferred. When reporting results, we give only the results for *c* = 6 nats.

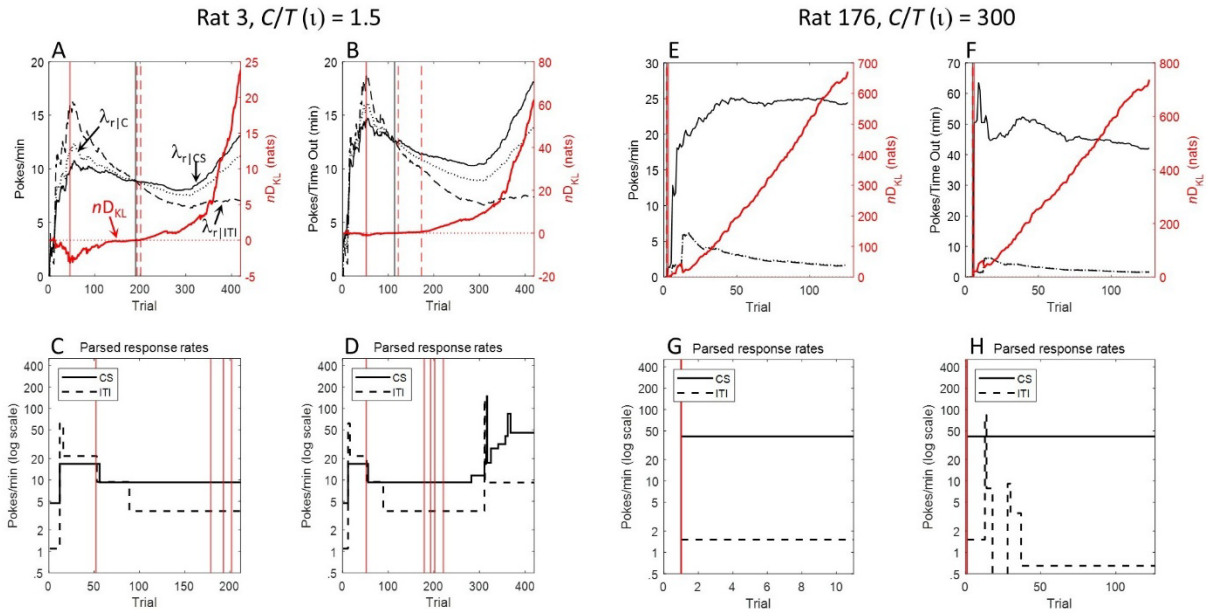
In computing the parses, we did not include the first CS poke in any given CS in our estimate of the CS poke rate because the latencies to the first pokes clearly came from a different distribution than the distribution of inter-poke intervals. In protocols with a short mean CS, the mean inter-poke interval (the reciprocal of the poke rate) was as short as 0.1 s; whereas the mean latency to the first poke in a CS was rarely shorter than 2 s. The explanation for the slow first pokes is probably the fact that the duration of a CS—hence the minimum reinforcement latency—was never shorter than 2 s. In the denominator of our poke rate estimates, we excluded the interval to the first poke. We also excluded the intervals when the head was in the hopper, because a poke cannot be made when the head is in the hopper. CS durations with no pokes were also excluded because there was no way to estimate the first-poke latency. Thus, our estimates of the poke rates included only CSs where there were pokes

and the *n*'s in the numerators of those estimates were only the counts after each first poke. In the denominator was the cumulative duration of the CSs that had at least one poke minus the cumulative latencies to the first pokes minus the cumulative duration of the hopper entries. The estimates of the poke rates during the Pre periods were the cumulative number of pokes in those intervals divided by their cumulative duration. The estimates of the contextual poke rates were the Pre poke rate estimate and the CS poke rate estimate weighted respectively by the cumulative ITI duration divided by the cumulative training duration and by the cumulative CS duration divided by the cumulative training duration.

The initial rate estimate in a parse extends back to the beginning of observation, which makes it possible to plot, for example, an initial ITI poke rate that starts at the beginning of training even though the first poke during a Pre interval may not have occurred until the 5th ITI. The parsing gives an alternative way of comparing rates of poking as of each successive reinforcement for the first few reinforcements. This becomes important when informativeness is high because then conditioned poking appears very early in training, as early as the second CS presentation.

Given the novelty of the methods used to estimate the onset of conditioned poking, we thought it essential to provide plots of the results on which these estimates are based. The Supplementary Materials (pages 10-185) include, for every rat, a plot of the $n\mathrm{D}_{\mathrm{KL}}$ and cumulative response rates during the CS, ITI, and context, as well as plots of the parsed response rates during the CS and ITI. Figure 6 provides examples of these plots for two rats: Rat 3, trained with the lowest level of informativeness (ι = 1.5), and Rat 176, trained with the highest level of informativeness (ι = 300). The top row of plots shows the cumulative response rates during the CS, $\lambda_{r|CS}$, during the ITI, $\lambda_{r|ITI}$, and the overall rate in the context, $\lambda_{r|C}$, as well as the signed $n\mathrm{D}_{\mathrm{KL}}$ for the divergence between CS and context response rats. The bottom row of plots shows the parsed response rates for the CS and ITI.

Figure 6. $n\mathrm{D_{KL}}$ and parsed response rates for two rats.



Note. *Panels A to D show data from Rat 3; panels E to H show the corresponding data from Rat 176. For each rat, the top two panels plot the cumulative rate of poking during the CS,* $\lambda_{r|CS}$ *(solid black line), during pre-CS ITIs,* $\lambda_{r|ITI}$ *(dashed black line), and the contextual rate,* $\lambda_{r|C}$ *(dotted black line), as functions of the number of trials (1 reinforcement per trial). In each plot, the red curve is the signed* $n\mathrm{D_{KL}}$ *plotted against the right axis. The black vertical line marks when the cumulative CS response rate permanently exceeded the cumulative context rate. The two red dashed vertical lines to the right of the black line mark when the* $n\mathrm{D_{KL}}$ *reached 0.82 (Odds 4:1 that CS rate > Context rate) and 1.92 (*$p < .05$ *that CS rate = Context rate), and the unbroken red line marks the minimum of the* $n\mathrm{D_{KL}}$*. In A and E, response rates were calculated conventionally, as number of responses divided by total time. In B and F, response rates were calculated as the number of responses excluding the first response in each CS divided by the remaining time (after the 1st response) out of the magazine. The bottom two panels for each rat show the parsed estimates of* $\lambda_{r|CS}$ *and* $\lambda_{r|ITI}$*. In C and G, the x-axis has been right-cropped to better reveal early changes. In all plots, the vertical red lines mark estimates of acquisition based on the* $n\mathrm{D_{KL}}$ *as the Earliest estimate (leftmost), and when the odds against the null hypothesis that the parsed CS and ITI rates were equal reached 4:1, 10:1, 20:1, and 100:1.*

In the top two plots of data from Rat 3, all three response rates rise rapidly over the first 40 trials and then decline. The ITI rate (dashed line) is generally greater than the CS rate (solid line) until about the 200th trial in Figure 6A (or 110th trial when using the adjust measure of response rate in 6B), after which it is consistently lower. Both averages steadily decline up to about Trial 300, implying a persistent drop in both rates following their peak at around Trial 40. At about Trial 300, the average CS poke rate begins to rise rapidly, implying that the CS

poke rate increased at the start of the rise. The slight rise in the dashed curve indicates a slight increase in the Pre poke rate as well. The successive changes in the poke rates inferred from these plots of the cumulative rate estimates are confirmed by the parses plotted in Figure 6C and 6D, where the solid black line plots the parse of the CS poke rate and the dashed black line the parse of the Pre rate. The Pre poke rate started lower than the CS rate but jumped after about 10 reinforcements to a higher value. At about Trial 40, both rates became the same, but at about Trial 85, the Pre poke rate dropped permanently below the CS rate. At about Trial 300, there is a sequence of increases in the CS poke rate, accompanied by a single smaller increase in the Pre poke rate. The decision criterion used in this and all the plotted parses was 6 nats, so one may have substantial confidence that the changes are statistically significant. (When $n\mathrm{D}_{\mathrm{KL}}$ = 6, the odds are greater than 2000:1 against the null hypothesis.)

The signed $n\mathrm{D}_{\mathrm{KL}}$ function is plotted in red against the right axis in the top row of Figure 6. The $n\mathrm{D}_{\mathrm{KL}}$ is always positive because the magnitude of a divergence cannot be negative. The direction of a divergence may, however, vary; the Pre rate may be greater than or less than the CS rate. To make the direction of divergence apparent, we give the $n\mathrm{D}_{\mathrm{KL}}$ positive sign when the CS rate is greater than the Context rate and negative sign when the reverse is true. For Rat 3, there is a short initial positive spike in the signed $n\mathrm{D}_{\mathrm{KL}}$ followed by an interval of several tens of reinforcements when it is negative. After about 40 reinforcements, it hits its minimum and began a more or less steady climb. Its last upward crossing of the 0 line is at about Trial 200 (Trial 110 for the plot to the right).

Rat 176, whose data are plotted in Figure 6E to 6H, made no pokes during the first 2 CSs, one during the 3rd CS, none during the 4$^{\mathbf{th}}$**,** one during the 5th, and 16 on the 6th. It made no pokes during the first five 30-s Pre intervals. Consequently, the average ITI poke rates and the average contextual poke rates are undefined over the first few trials and so is the $n\mathrm{D}_{\mathrm{KL}}$. On Trial 3, the CS response rate and Context response rate can be defined—and the $n\mathrm{D}_{\mathrm{KL}}$ is already greater than 2 (thus $p < .05$). By Trial 6, the $n\mathrm{D}_{\mathrm{KL}}$ is 11.7 nats, which corresponds to odds of 760 million to 1 against the null. Despite this, one might conclude there was no evidence of CS-conditional responding in this subject until Trial 3.

However, as already remarked, the first segment of a rate parse always extends back to the onset of observation. In the parses of the CS poke rate and the ITI poke rate, which are plotted in Figure 6G and H, there is no evidence for a changes in the poke rates at Trial 3.

Indeed, the parse of the CS poke rate shows no change over all 126 CSs, while the parse of the ITI poke rate finds the first change to be at Trial 13.

The failure to find changes in the poke rates at Trial 3 or even 6 is not a consequence of the high value for the decision criterion. Lowering it from 6 nats to 4 did not change the parse; lowering it to 2 nats produced a parse with a single change, a drop from 52 pokes/min to 38 pokes/min at Trial 40. Likewise for the parse of the ITI rate of poking: lowering the decision criterion from 6 nats to 2 nats did not alter it. Nor does this failure occur because the algorithm cannot find very short segments. Parses of the contextual poke rate at all three values for the decision criterion find the same 1-trial long segment at Trial 13—see upward blip in dashed last plot in Figure 6. In other subjects, upward or downward blips 1 or 2 trials long are sometimes found at the outset of training.

There are significantly fewer pokes during the first 5 CSs than expected given the initial rate estimate for the CS poke rate. Much of this is attributable to the initially slow reaction to CS onset. The poke on Trial 3 came 20.3s into the 23s long CS; the poke on Trial 5 came 14.45s into the 16s long CS. The latency to make the first poke dropped rapidly over the first 9 trials from a mean greater than 11s for the first 5 trials to a mean of 3.6s for the Trials beyond 10. During the first 5 trials, the observation intervals during which it was possible to register a poke that would go into the parsing algorithm totaled only 4.18s. Given an initial poke rate estimate of 0.7/s, the expected number of pokes in that interval is 2.9 and there is a 5% chance of observing no pokes.

All considered, an argument can be made in this and several other cases where the informativeness was large that the parse results are a better indicator of the onset of CS-conditional poking, particularly when informativeness is high and the ITI poke rate very low. Blips in the ITI poke rate often put it momentarily greater than the CS poke rate (see dashed line in Figure 6H) are common in post-acquisition protocols. Therefore, our parse-based estimate of the onset of conditioning is the trial after which the parsed CS poke rate is greater than the parsed ITI poke rate on 95% of the trials.

REFERENCES

Balsam, P. D., Fairhurst, S., & Gallistel, C. R. (2006). Pavlovian contingencies and temporal information. *Journal of Experimental Psychology: Animal Behavior Processes, 32*, 284-294.

Balsam, P. D., & Gallistel, C. R. (2009). Temporal maps and informativeness in associative learning. *Trends in Neurosciences, 32*, 73-78.

Belke, T. W. (1992). Stimulus preference and the transitivity of preference. *Animal Learning and Behavior, 20*, 401-406.

Bouton, M. E., & Sunsay, C. (2003). Importance of Trials Versus Accumulating Time Across Trials in Partially Reinforced Appetitive Conditioning. *Journal of Experimental Psychology: Animal Behaviour Processes, 29*, 62-77.

Burke, D. A., Jeong, H., Wu, B., Lee, S. A., loeder, J. R., & Namboodiri, V. M. K. (2024). Reward timescale controls the rate of behavioral and dopaminergic learning. *bioRxiv*. doi:10.1101/2023.03.31.535173

Cover, T. M., & Thomas, J. A. (1991). *Information theory*. New York: Wiley.

Donahaoe, J. W. (2006). Contingency: Its meaning in the experimental analysis of behavior. *European Journal of Behavior Analysis, 7*, 111-114.

Drew, M. R., Zupan, B., Cooke, A., Couvillon, P. A., & Balsam, P. (2005). Temporal control of conditioned responding in goldfish. *Journal of Experimental Psychology: Animal Behavior Processes, 31*, 31-39.

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: Bradford Books/MIT Press.

Gallistel, C. R. (2003). Conditioning from an information processing pespective. *Behavioural Processes, 62*, 89-101.

Gallistel, C. R. (2021). Robert Rescorla: Time, Information and Contingency. *Revista de historia de la psicología, 42*, 7-21.

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

Gallistel, C. R., Fairhurst, S., & Balsam, P. (2004). The learning curve: Implications of a quantitative analysis. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 13124-13131.

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

Gallistel, C. R., & Latham, P. (2023). Bringing Bayes and Shannon to the study of behavioral and neurobiological timing. *Timing & TIME Perception., 11*, 29-89. doi:10.1163/22134468-bja10069

Gallistel, C. R., & Shahan, T. A. (2024). Time-scale invariant contingency in reinforcement learning with extremely long delays to reinforcement. *under review*.

Gibbon, J. (1977a). Scalar expectancy theory and Weber's Law in animal timing. *Psychological Review, 84*, 279-335. doi:doi.org/10.1037/0033-295X.84.3.279

Gibbon, J. (1977b). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84*, 279-325.

Gibbon, J., Baldock, M. D., Locurto, C. M., Gold, L., & Terrace, H. S. (1977). Trial and intertrial durations in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes, 3*, 264-284.

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

Gibbon, J., Berryman, R., & Thompson, R. L. (1974). Contingency spaces and measures in classical and instrumental conditioning. *Journal of the Experimental Analysis of Behavior, 21*, 585-605. doi:10.1901/jeab.1974.21-585

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.

Gottlieb, D. A. (2004). Acquisition with partial and continuous reinforcement in pigeon autoshaping. *Learning & Behavior, 32*, 231-334.

Gottlieb, D. A. (2005). Acquisition with partial and continuous reinforcement in rat magazine approach. *Journal of Experimental Psychology: Animal Behavior Processes, 31*, 319-333.

Gottlieb, D. A. (2008). Is the number of trials a primary determinant of conditioned responding? *Journal of Experimental Psychology: Animal Behavior Processes, 34*, 185-201.

Granger, R. H. J., & Schlimmer, J. C. (1986). The computation of contingency in classical conditioning. .). . In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 20, pp. 137-192). New York: Academic Press.

Hallam, S. C., Grahame, N. J., & Miller, R. R. (1992). Exploring the edges of Pavlovian contingency space: An assessment of contingency theory and its various metrics. *Learning and Motivation, 23*, 225-249.

Hammond, L. J., & Paynter, W. E. (1983). Probabilistic contingency theories of animal conditioning: A critical analysis. *Learning and Motivation, 14*, 527-550. doi:10.1016/0023-9690(83)90031-0

Harris, J. A. (2022). The learning curve, revisited. *Journal of Experimental Psychology: Animal Learning and Cognition, 48*, 265-280.

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

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

Holland, P. C. (2000). Trial and intertrial durations in appetitive conditioning in rats. *Animal Learning and Behavior, 28*(2), 121-135.

Honey, R. C., Dwyer, D. M., & Iliescu, A. F. (2020). A model for Pavlovian learning and performance with reciprocal associations. *Psychological Review, 127*, 829–852.

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 Press.

Kamin, L. J. (1968). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: aversive stimulation* (pp. 9-31). Miami: Miami University Press.

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*, 39-72. doi:doi.org/10.1002/jeab.261

Kheifets, A., & Gallistel, C. R. (2012). Mice take calculated risks. *Proceedings of the National Academy of Sciences of the United States of America, 109*, 8776-8779. doi:doi.org/10.1073/pnas.1205131109

Killeen, P. R. (2023). Theory of reinforcement schedules. *Journal of the Experimental Analysis of Behavior, 120*, 289-319. doi:doi.org/10.1002/jeab.880

Killeen, P. R., Hall, S., & Bizo, L. A. (1999). A clock not wound runs down. 45, 129-139. . *Behavioural Processes, 45*, 129-139.

Killeen, P. R., Hanson, S. J., & Osborne, S. R. (1984). Arousal: its genesis and manifestation as response rate. , 85, doi:10.1037/0033-295X.85.6.571. *Psychological Review, 85*, 571-581. doi:10.1037/0033-295X.85.6.571

Kirkpatrick, K., & Church, R. M. (2000). Independent effects of stimulus and cycle duration on conditioning: The role of timing processes. *Animal Learning and Behavior, 28*, 373-388.

Lattal, K. M. (1999). Trial and intertrial durations in Pavlovian conditioning: Issues of learning and performance. *Journal of Experimental Psychology: Animal Behavior Processes, 25*(433-450).

Ludvig, E. A., Sutton, R. S., & Kehoe, E. J. (2012). Evaluating the TD model of classical conditioning. *Learning & Behavior, 40*, 305-319.

Namboodiri, V. M. K. (2022). How do real animals account for the passage of time during associative learning? *Behavioral Neuroscience, 136*, 383-391.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review, 87*(6), 532-552.

Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review, 74*(1), 71-80.

Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology, 66*(1), 1-5.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory.* (pp. 64-99). New York: Appleton-Century-Crofts.

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

Russo, D. J., Roy, B. V., Kazerouni, A., Osband, I., & Wen, Z. (2018). A tutorial on Thompson Sampling. *Foundations and Trends in Machine Learning, 11*, 1-96.

Sutton, R. S., & Barto, A. G. (1990). Time-derivative models of Pavlovian reinforcement. In M. Gabriel & J. Moore (Eds.), *Learning and computational neuroscience: Foundations of adaptive networks* (pp. 497-537). Cambridge, MA: Bradford Books/MIT Press.

Thrailkill, E. A., Todd, T. P., & Bouton, M. E. (2020). Effects of conditioned stimulus (CS) duration, intertrial interval, and I/T ratio on appetitive Pavlovian conditioning. *Journal of Experimental Psychology: Animal Learning and Cognition, 46*, 243-255.

Vogel, E. H., Ponce, F. P., & Wagner, A. R. (2019). The development and present status of the SOP model of associative learning. *Quarterly Journal of Experimental Psychology, 72*, 346-374. doi:10.1177/1747021818777074

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of *p* values. *Psychonomic Bulletin & Review, 14*, 779-804.

Ward, R. D., Gallistel, C. R., Jensen, G., Richards, V. L., Fairhurst, S., & Balsam, P. D. (2012). Conditioned stimulus informativeness governs conditioned stimulus-unconditioned stimulus associability. *Journal of Experimental Psychology: Animal Behavior Processes, 38*, 217-232.