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, remarkably, the specific scalar relationship between the rate of learning and informativeness aligns very closely 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' response rates during the (never-reinforced) inter-trial interval, which was directly proportional to the contextual rate of reinforcement (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 the data 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 completely complementary, such that 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 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 (Gibbon & Balsam, 1981). 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). Their data on pigeon autoshaping are well described by a regression model whose only parameter is the x-intercept, which is 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 about here ---------

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) termed the ratio of the CS rate of reinforcement to the contextual rate the *informativeness*, 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, Jeong, Wu, Lee, & Namboodiri, 2023; 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.* (2023) 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 which established that *C*/*T* determines when pigeons start to respond to a CS, 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).

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.*

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 (Harris & Carpenter, 2011; Harris, Gharaei, & Pincham, 2011).

RESULTS

Rats in all groups eventually responded at a higher rate to the CS than during the ITI. Figure 2 shows the mean CS and ITI response rates for each group across the 42 conditioning sessions. The equivalent of this figure for each rat is included in the Supplementary Material (pages 2 to 169). The last panel of Figure 2 (Plot 15) shows the mean response rate per second during the CS for each group, averaged from all trials of the final 5 sessions of the experiment. This shows how responding increased rapidly over the first few seconds of the trial and then remained constant across the trial, consistent with previous experiments using variable CS-US intervals (Harris & Carpenter, 2011; Harris et al., 2011).

--------- Figure 2 about here ---------

Trials to acquisition.

We used the cumulative response rates to identify the trial at which conditioned responding first emerged. The very earliest evidence for responding to the CS was the trial after which the cumulative response rate during the CS permanently exceeded the cumulative ITI response rate. Figure 3A plots this trial as a function of the CS informativeness (*C/T*) for each rat. The number of trials needed for rats to show evidence of conditioning systematically decreased as the informativeness increased. These data are superimposed on the data shown in Figure 1 that plots the results of a meta-analysis of many comparable experiments with pigeons (Gibbon & Balsam, 1981), revealing a surprising agreement between the present results and those previous results.

--------- Figure 3 about here ---------

The apparent agreement shown in Figure 3A between our data and those reported by Gibbon and Balsam (1981) is difficult to interpret because the criterion for learning used in our data is different from that used by Gibbon and Balsam. To investigate the agreement in our data sets more carefully, we have applied statistical tests to our data to match the criterion used by Gibbon and Balsam. Their criterion—that pigeons make at least one peck at the response key on at least 3 out of 4 consecutive trials—cannot be applied directly to our magazine entry data because rats make magazine entries during the ITI, which means that evidence for conditioning must involve comparison between the response rate during the CS

and the baseline response rate. (This was not necessary for Gibbon and Balsam because their pigeons did not peck the response key in the ITI.) We used two different statistical tests of the difference between CS and ITI response rates, each selected to achieve a statistical threshold equivalent to the criterion used by Gibbon and Balsam. These tests were applied to the cumulative data starting from the trial on which the cumulative CS rate permanently exceeded the ITI rate (as plotted in Figure 3A) to identify the trial on which CS responding was significantly greater than the baseline. One analysis used a non-parametric signed rank test for evidence that the CS response rate exceeds the pre-CS response rate. We applied a statistical criterion of $p < .3125$ to be equivalent to Gibbon and Balsam's criterion of at least 3-out-of-4 consecutive trials with a response. The trials-to-criterion data from this analysis are shown in Figure 3B; they confirm the close agreement between our data and theirs.

The second statistical analysis we conducted to compare our data with those of Gibbon and Balsam (1981) was based on the $n\mathrm{D}_{\mathrm{KL}}$, an information-theoretic statistic of the distance between two distributions (Gallistel & Latham, 2023). Here, we have used the $n\mathrm{D}_{\mathrm{KL}}$ to measure the distance between the cumulative CS response rate and ITI response rate (see the Methods section for a description of the Kullback-Leibler divergence, and the $n\mathrm{D}_{\mathrm{KL}}$). The first step in this analysis was to estimate the value of the $n\mathrm{D}_{\mathrm{KL}}$ for the criterion used by Gibbon and Balsam. If we assume that pigeons make no responses in the ITI, and that the probability of a response during the CS is at least 0.75 (their 3-out-of-4 criterion), this corresponds to an $n\mathrm{D}_{\mathrm{KL}}$ of 2.2 (and an odds ratio of 27:1). We then calculated, trial-by-trial, the $n\mathrm{D}_{\mathrm{KL}}$ for the cumulative CS and ITI response rates of each rat. Figure 3C plots the trial at which the $n\mathrm{D}_{\mathrm{KL}}$ for each rat exceeded 2.2, the threshold derived from Gibbon and Balsam's criterion. Once again, there is strikingly close correspondence between our data and theirs.

In the data plots shown in Figures 3A to C, it can be hard to appreciate the distribution of trials to acquisition data when the data for individual subjects (open orange diamonds) are superimposed. This is particularly the case for subjects learning after their first reinforcement, where individual data pile up on the x-axis, even when the informativeness is as low as 9. To display the distribution of data more clearly, Figure 3D re-plots the data from Figure 3C showing trials to acquisition as a function of the mutual information between CS and US (bottom axis; equal to $\log_2$ of the informativeness shown on the top axis). In 3D, the individual data for trials to acquisition data have been grouped in bins of approximately equal width on

a logarithmic scale. Thus, the tiles in the first 4 rows represent single numbers of trials (1, 2, 3, and 4 trials to acquisition) and tiles in subsequent rows represent bins with wider ranges (trials 5-6, 7-9, 10-15 etc). The darkness of each tile in Figure 3D represents the number of rats within that bin. The right-hand axis shows the learning rate, computed as the reciprocal of number of trials to acquisition. Figures 3C and 3D show the large individual differences in trials to acquisition; subjects within the same informativeness group may have trials-to-acquisition values that differ by two orders of magnitude. Nonetheless, overall, the learning rate increases as the informativeness increases. It is also clear that the number of rats learning after just one reinforced trial (bottom row of 3D) increases as the informativeness increases.

Other values for trials-to-criterion using statistical thresholds based on *p* values ($p < .05$, .01, and .001) derived from the $n\mathrm{D_{KL}}$ are included in the Trials2AcquisitionTable Excel file that can be downloaded from https://osf.io/vmwzr/

Correlational analyses, with α set at .017 to correct for multiple tests, were used to assess the relationships between the median log of trials to criterion (based on the $n\mathrm{D_{KL}} > 2.2$ criterion shown in Figure 3D) and log(*C*/*T*), log(*C*), and log(*T*). Log(*C*/*T*) was the strongest predictor of trials to criterion ($r = -0.92$, $p < .001$). Log(*C*) was also significantly correlated with the log of trials to criterion, $r = -0.84$, $p < .001$, but log(*T*) was not correlated with trials to criterion, $r = 0.35$, $p = .22$. Given that log(*C*) and log(*C*/*T*) were strongly correlated with one another (r = 0.90), their partial correlations with the log of trials to criterion were calculated to test whether each made independent contributions to the rate of acquisition. After partialling out the effect of log(*C*), the correlation with log(*C*/*T*) remained significant, $r = -0.70$, $p = .007$. In contrast, after partialling out the effect of log(*C*/*T*), the correlation with log(*C*) was not significant, $r = 0.08$, $p = .795$. In sum, the number of trials required for conditioned responding to emerge was strongly affected by the *C*/*T* ratio, and neither *C* nor *T* alone had any independent effect.

Terminal response rate.

As a measure of the terminal level of responding, the response rates during the CS were averaged over the last 5 sessions for each rat. We have conducted the same analysis described here using the data from the final 10 session, with very similar results. (Response rate data of individual rats averaged from both the last 5 and last 10 sessions can be

downloaded from https://osf.io/vmwzr/) The response rate was calculated as total number of responses (summed across all trials over the 5 sessions) divided by the total CS duration (summed across all trials over the 5 sessions). Our initial analysis compared this terminal response rate with the log of *T*, *C*, and *C/T* (α = 0.017 after correction for multiple comparisons). The rate of responding to the CS was marginally correlated with log(*T*) (see Figure 4A), $r = -0.53$, $p = .051$, but was not correlated with either log(*C*), $r = -0.3$, $p = .298$, or log(*C*/*T*), $r = -0.04$, $p = .894$.

More detailed analyses were conducted after segmenting the response rate into 3 separate components: (1) the latency to the first response in a trial; (2) the mean duration of each response (the time spent in the magazine); and (3) the inverse of the mean inter-response interval, 1/IRI, which equals the response rate after excluding the latency and response durations. None of these indices correlated significantly with log(*C*), largest $r = -0.35$, $p = .221$, or with log(*C*/*T*), largest $r = -0.31$, $p = .281$. On the other hand, log(*T*) correlated significantly with latency, $r = 0.77$, $p < .001$ (Figure 4B), and with 1/IRI, $r = 0.89$, $p < .001$ (Figure 4C). There was a marginal correlation between duration of responding and log(*T*), $r = -0.55$, $p = .043$, that did not survive correction for multiple comparisons (Figure 4D). As is evident in Figure 4B, the positive correlation between latency and log(*T*) was largely confined to groups with long CS-US intervals ($T > 25$ s), whereas latency varied little among groups with shorter CS-US intervals. This invariance at short CS-US intervals may have been due to a floor effect because mean latencies to first response did not decrease below 2 s (horizontal dotted line in Figure 4B). This is also consistent with the plots of response by time-in-CS, shown in plot 15 of Figure 2, where responding to the CS was low for the first few seconds after CS onset. This apparent floor effect might reflect a constraint on how quickly the rats can commence responding after CS onset or it might have arisen because 2 s was the minimum CS-US interval used in this experiment. Regardless, these analyses indicate that neither duration of responding nor latency to first response are good markers of what the rats learn about the rate of reinforcement of the CS. By contrast, the response rate, 1/IRI, varied systematically across the entire range of values of *T* (Figure 4C). This confirms earlier evidence that rats' response rates scale with the log of the reinforcement rate (Harris & Carpenter, 2011).

--------- Figure 4 about here ---------

The response rates in Figure 4A are computed by dividing the response count by the duration of the interval over which the count was made. However, when the CS reinforcement rate was high, subjects made from 2, to more than 5, pokes per second and the pokes lasted substantial fractions of a second. A rat cannot make another poke when its head is in the magazine. The correct computation of a response rate should use the time over which it was possible for pokes to be counted, that is, the cumulative time minus the cumulative head-in-magazine time. Also, the latency of the first poke was rarely shorter than 2 s (Figure 4B); whereas the inter-poke intervals after the first poke were much shorter when *T* was short (Figure 4C). To make a proper estimate of the rate at which a rat poked once it had begun, we divided the number of pokes made during the CS (excluding the first poke) by the remaining time in which it was possible to initiate a poke (i.e., the CS duration minus the latency to 1st poke and the cumulative head-in-magazine time). Likewise, to estimate the contextual rate of poking, we divided the ITI response count by the time in which an ITI response could be initiated (i.e., the length of the pre-CS interval minus the cumulative time with head in magazine). These methods for computing the CS and contextual rates of responding correspond closely to 1/IRI (shown in Figure 4C) except that they can still be computed when there are fewer than 2 responses (unlike the IRI). Figure 5 plots the terminal CS response rates (black dots) and contextual response rates (red dots) as functions of the CS/contextual reinforcement rate on double logarithmic coordinates.

--------- Figure 5 about here ---------

Figure 5 shows that the log response rate while in the experimental context and the log response rate during CSs are both described by the same linear regression when plotted against the relevant reinforcement rate (regression line fitted using Excel^{TM}). It accounts for 80% of the variance. Its slope is essentially 1, which means that the two poke rates are the same *scalar* function of the two reinforcement rates. The scalar that relates response rate to reinforcement rate is $10^{1.28}$ = 19, where 1.28 is the constant in the logarithmic regression. Thus, the average rates of poking are 19 times faster than the reinforcement rates. This maximally simple relation between reinforcement rate and poke rate holds from reinforcement rates almost as low as 1 in 100 min (1 in every .00016 s), and corresponding poke rates as slow as 5 or 6/min, up to reinforcement rates as high as 10/min (1 every 6 s) and corresponding poke rates of 90-110/min (1 every 0.5 s).

Put another way, the average time the rat waits between magazine entries is 1/19 (= .05) times the average time it expects to wait for food (also Harris & Carpenter, 2011). The same scalar relation between the behavioural wait and the expected wait for reinforcement applies when the rat is in an intertrial interval—when reinforcement is never delivered—and when it is in a CS interval. As the informativeness of a protocol gets lower, that is, as the CS occupies a greater and greater fraction of the subject's time in the chamber, the difference between the CS poke rate and the contextual poke rate gets smaller because the difference between the CS reinforcement rate and the contextual reinforcement rate gets smaller.

As is also apparent in Figure 5, the variability about the scalar relation between poke rate and reinforcement rate also scales with reinforcement rate. That is why the regression must be computed in the logarithmic domain.

A single regression line, shown as the black line in Figure 5, is the simplest possible model of the relationship between response rate and reinforcement rate and it explains approximately 80% of the variance in response rate. Indeed, treating this as a 1-parameter model by fixing the slope of the line at 1 produces negligible loss of explanatory power and is theoretically justified because it means that, in the non-log domain, the scalar relation between response rate and reinforcement rate passes through the origin. Put simply, it means that the response rate is zero when the reinforcement rate is zero. Nonetheless, if regression lines are fitted to the CS and ITI data separately, as with the red and grey dashed lines in Figure 5, there is a small increase in explained variance (R^2 = 0.82). We leave it to further research to determine whether such a complex model, with 4 parameters, is required. However, we do not think the present data warrant comparing the simplest possible model, with one parameter, to any more complex model.

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 4), our measure of the CS response rate was calculated by dividing the response count by the total time out of the magazine during the CS. As shown in Figure 6A, 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 6B).

--------- Figure 6 about here ---------

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, cum$R$', 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 6C. 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 6 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 7. 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 7). 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.

--------- Figure 7 about here ---------

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 7 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 7, 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 just 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 7). The correlation between slope and *T* remained significant after partialing 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 5).

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 an analysis (see description of response rate parsing using the Kullback-Leibler divergence, and the $n\mathrm{D}_{\mathrm{KL}}$, in Methods) that uses the $n\mathrm{D}_{\mathrm{KL}}$ to parse the response rates into segments that have significantly different rates (either higher or lower). In such cases, the path to the peak rate can be bumpy and the peak can be higher than the terminal response rate, as illustrated in Figure 8 for 6 rats from the group with ι = 1.5. By contrast, when the informativeness is very high, the peak is usually reached almost immediately and there is little subsequent variation in the poke rate (Figure 9). The parsing of response rates shown in Figures 8 and 9 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 in Figure 9, there are only between 0 and 2 steps 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 in Figure 8, 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.

--------- Figures 8 and 9 about here ---------

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 initial acquisition of responding was determined by the *C*/*T* ratio. Neither *T* nor *C* had any independent effect after partialing out the effect of *C*/*T*. This is the same conclusion reached by Gibbon and Balsam (1981) in their meta-analysis of the acquisition of autoshaped key-pecking by pigeons. As shown in Figure 3, responding to the CS emerged sooner as the *C*/*T* ratio increased, or, more precisely, the log of the number of trials to acquire a response scaled negatively with the log of *C*/*T*.

The dependence of trials to acquisition on *C*/*T* extends over the entire range from 1 to infinity. Contrary to what Gibbon and Balsam supposed, but in agreement with what Jenkins *et al*. (1981) showed, the regression continues to describe the data all the way to the abscissa, that is, to acquisition after 1 reinforcement. It is remarkable that the 1-parameter regression model that describes the pigeon acquisition data also describes the current rat magazine-poking data—and with essentially the same parameter value (Figure 3).

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 strongly related to *T* (but not *C* or *C*/*T*). More specifically, the terminal CS response rate is a scalar function of the CS reinforcement rate (red symbols in Figure 5). Related to this is the third, and unexpected, finding that the response rates during the never-reinforced ITIs is proportional to the contextual rates of reinforcement (black symbols in Figure 5). It is particularly noteworthy that the constant of proportionality relating the ITI response rates to contextual reinforcement rates <u>is the same</u> as that relating the CS response rates to CS reinforcement rates. To the best of our knowledge, the dependence of the rate of responding during the ITIs on the *contextual* rate of reinforcement—rather than on the *probability* of reinforcement or the *rate* of reinforcement during the ITIs—has not been previously reported.

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 (Figure 7). 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 (Figure 8). 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 (Figure 9).

The first finding confirms results reported previously in studies of Pavlovian conditioning in pigeons. Perhaps the most remarkable aspect of the observed relationship between trials to acquisition and *C*/*T* is that the data collected from conditioning experiments with pigeons and the data reported here are consistent with a simple regression function that has a slope of -1 over most of its range and an intercept of 255. As discussed later, this finding is strong evidence in favour of the notion that conditioning is directly related to how much information the CS provides about the US. That rats poking into a magazine would satisfy the same model with essentially the same value for its one parameter was not to be expected.

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. Holland observed more responding in the last 8 of 16 conditioning sessions and also found a steeper increase in responding over sessions for rats trained with short values of *T* despite matched *C*/*T*. 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 170-172). 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 170-172), produced smaller differences between groups and therefore a weaker (albeit significant) correlation between trials to acquisition and log(*C/T*). 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).

Had it been appreciated in the 1960s that the rate of responding during ITIs has the same scalar relation to the contextual rate of reinforcement as the rate of responding during CSs has to the CS rate of reinforcement, the results from Rescorla's (1967, 1968) truly random control and from Kamin's (1968) blocking experiments would have been considered entirely predictable. In both protocols, the reinforcement rate during the (target) CS is the same as the rate expected in the context in which it occurs. In the truly random protocol, the context is the test chamber; in a blocking protocol, it is the previously conditioned CS.

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 assignment 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.

The results in Figure 5 confirm that the subjects compute the contextual rate of reinforcement. 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. During the ITIs, subject's poke rate is 18 times the reinforcement rate expected when in the experimental context; during the CSs, it is 18 times to the reinforcement rate expected when in the context and in the presence of the CS.

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, $n\mathrm{D_{KL}}$, 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 $\mathrm{D_{KL}}$ in the $n\mathrm{D_{KL}}$ 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: $\mathrm{D_{KL}}(X||Y) \neq \mathrm{D_{KL}}(Y||X)$. 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),

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

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), $n\mathrm{D}_{\mathrm{KL}}(X||Y)_{\mathrm{exp}} \sim \Gamma(0.5,1)$, $n\mathrm{D}_{\mathrm{KL}}$'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? Figure 10 plots the cumulative distributions (CDFs) of the values for $n\mathrm{D}_{\mathrm{KL}}(\lambda_{\mathrm{R|CS}}||\lambda_{\mathrm{R|C}})$—the $n\mathrm{D}_{\mathrm{KL}}$ for the evidence that the CS rate of reinforcement, $\lambda_{\mathrm{R|CS}}$, differs from the contextual rate of reinforcement, $\lambda_{\mathrm{R|C}}$, when subjects' increased poke rates during CSs meets increasingly stringent evidentiary criteria for acquisition.

--------- Figure 10 about here ---------

The medians of the blue CDFs in Figure 10 generally fall at or below 3.2 nats worth of evidence, which corresponds to a *p* value of .01. This implies that they assess appropriately the strength of the accumulating evidence for reliable mutual information. The CDFs for the 10:1, 100:10, and 1000:1 evidentiary criteria cluster together well to the right, toward absurdly stringent criteria. (The odds against when the $n\mathrm{D}_{\mathrm{KL}}$ is greater than 10 are more than 10,000:1.) The strong rightward shift in the strength of the stimulus evidence at acquisition occurs because the behavioural effect size is small and noisy (Figure 10). The low behavioural effect size requires larger n's (more trials) to satisfy increasingly stringent evidentiary criteria (Figure 3). The Kullback-Leibler divergence of the CS reinforcement rate from the contextual rate is also small when the informativeness is low; when $\iota = 2$, it is only 0.2 nats. However, the effect of the small ratio of the poke rates on the n's required for increasingly stringent evidentiary criteria outweighs the small divergence in determining the amount of evidence for mutual information subjects have acquired after a given number of reinforcements.

The divergence is a decelerating function of informativeness, but its initial rise is steep: An informativeness of 10 yields a divergence of 1.5 nats. Thus, when a CS transmits 3.3 bits of mutual information per trial, 3 reinforced trials provide on average 4.5 nats of evidence, which corresponds to Odds of 370:1 in favour of the conclusion that the observed mutual information is reliable. This explains why CDFs for more stringent criteria in Figure 10 migrate leftward as informativeness increases. When informativeness exceeds the value required for 1-trial acquisition in the median subject, the median subject satisfies the Odds 100:1 criterion when there is only 2.4 nats of evidence for the mutual information, which corresponds to Odds of 35:1 ($p < .03$). Again, this a rational evidentiary criterion.

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. The poke rate, for example, is 18 times the reinforcement rate in the median subject (Figure 5) 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 example is the simple relation between the learning rate (the reciprocal of reinforcements to acquisition) and the informativeness of a Pavlovian protocol. 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 3). A third example is the parameter free rate estimation equation. In a 1-CS protocol, that equation is:

$$\begin{bmatrix} \hat{\lambda}_{\mathrm{R|C\&\sim CS}} \\ \hat{\lambda}_{\mathrm{R|CS}} \end{bmatrix} = \begin{bmatrix} 1 & 1/\iota \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \lambda_{\mathrm{R|C}} \\ \lambda_{\mathrm{R|C\&CS}} \end{bmatrix} \tag{6}$$

Equation (6) 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 (6) is the inverse of the informativeness, which is also the protocol parameter in the learning-rate equation. Equation (6) 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 behaviour 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. 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 3). 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 during the ITI, then converting these into cumulative records of

response rates by dividing the cumulative count at each trial by the cumulative time (CS or ITI) at that trial. To identify the earliest evidence of differential responding to the CS we found the trial, t_e, for which the cumulative response rate to the CS after that trial (starting from t_e + 1) was permanently greater than the cumulative ITI response rate starting from t_e + 1. This trial is plotted for each rat in Figure 3A. Next, we used two statistical tests to identify when, starting from trial t_e + 1, the cumulative CS response rate exceeded the ITI rate by a statistical threshold. One analysis used a Wilcoxon signed rank test to identify when the number of trials with a CS rate that exceeded the ITI rate was greater than expected by change with threshold p = .3215. (This threshold was chosen to match the criterion used by Gibbon and Balsam to identify acquisition of key peck responding in pigeons.) The second analysis used a new information-theoretic statistic, the $n\mathrm{D}_{\mathrm{KL}}$ (see next section), to identify when the rate of responding during the CS began to exceed that during the ITI. 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 contextual rate of responding (during the ITI) and the rate during the CS and 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 in which the CS occurs. 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 and 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 $c = 6\mathrm{nats}$.

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 Material gives one 4-plot figure for each subject (pages 2-169). Figures 11 and 12 are examples of these plots. The bottom row of plots in Figures 11 and 12 show the functions plotted over the entire 420 reinforcements (hence CS trials). In the top row, these complete plots have been right-cropped to better show what happens early in training.

--------- Figures 11 and 12 about here ---------

In the left column of Figures 11 and 12, three cumulative poke-rate functions are plotted in black against the left axes. The poke rate as of a given reinforcement is the cumulative number of pokes as of that reinforcement divided by the cumulative duration of the relevant interval. In other words, it is the average rate as of a given reinforcement. The CS average rate is a solid line, the Pre average rate a dashed line. The contextual average rate is a dotted line.

In the left column of Figure 11, all three averages rise rapidly over the first 40 trials and then decline. The Pre average (dashed line) is generally greater than the CS average (solid line) until about the 110th trial, 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 the right column of Figure 11, 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.

The signed $n\mathrm{D}_{\mathrm{KL}}$ function is plotted in red against the right axis in the left columns of Figures 11 and 12. 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 Pre rate and negative sign when the reverse is true. In the left column of Figure 11, 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 110. The upper limit on the $n\mathrm{D}_{\mathrm{KL}}$ axis (right axis) is set at 6 nats, because the evidence for acquisition is decisive when that value is exceeded. (When $n\mathrm{D}_{\mathrm{KL}} = 6$, the odds are greater than 2000:1 against the null hypothesis.)

The vertical red lines on the four plots are drawn at the values for various estimates of reinforcements to acquisition. The solid heavy red line is drawn at our earliest estimate of the onset of CS-conditional responding. In Figure 11, this is at the minimum of the signed $n\mathrm{D}_{\mathrm{KL}}$, which is more often than not what we take to be the best estimate of where conditioned poking to the CS first appeared. The thinner solid red line is for a generally more conservative estimate of first appearance, namely, the reinforcement after which the signed $n\mathrm{D}_{\mathrm{KL}}$ becomes permanently positive (last upward 0 crossing). The five vertical dashed red lines are

for increasingly stringent evidentiary criteria: odds against the null of 4:1, 10:1, 20:1 (p = .05), 200:1 (p = .005) and 2000:1 (p = .0005).

The subject whose results are plotted in Figure 12 made no pokes during the first 2 CSs, one during the 3rd CS, none during the 4**th**, one during the 5th and 16 on the 6th. It made no pokes during the first five 30s Pre intervals. Consequently, the average ITI poke rates and the average contextual poke rates are undefined over the first 5 trials and so is the $n\mathrm{D}_{\mathrm{KL}}$. On Trial 6, all 3 statistics become defined—and the $n\mathrm{D}_{\mathrm{KL}}$ is already off scale, which is why it does not appear in Figure 12a&c. The $n\mathrm{D}_{\mathrm{KL}}$ on Trial 6 is 20.7 nats, which corresponds to odds of 8 billion to 1 against the null. Consequently, the minimum in the signed $n\mathrm{D}_{\mathrm{KL}}$, the point where it became permanently positive, and the 5 increasingly stringent evidentiary criteria for acquisition after that, all fell at Trial 6, as indicated by the superposed red verticals at the right edge of Figure 12a. One might conclude that CS-conditional responding in this subject did not appear until Trial 5.

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 the right column of Figure 12 found there is no evidence for a changes in the poke rates at Trial 6 . 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 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 plot in Figure 12d. 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 plot in Figure 12d) 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.

Four-panel figures, equivalent to Figures 11 and 12, are included for each subject within the 6-panel figures on pages 2 to 169 of the Supplementary Materials. The data are available in the file "nDkl Acquisition Table.xlsx" that can be downloaded from the OSF site https://osf.io/vmwzr/. Columns 4, 5 and 6 of the table give the trial after which conditioned responding appeared as estimated in the above described three different ways— by the location of the minimum in the $n\mathrm{D}_{\mathrm{KL}}$, the last upward 0 crossing, and the CS parse consistently greater than the ITI parse, respectively. Column 3 in that table gives the minimum of the three estimates. Columns 7-11 of the table give the trial after which the evidence for acquisition exceeded increasingly stringent criteria (odds of 4, 10, 20, 100 and 1000:1). The different estimates may coincide: In 81 subjects, the minimum of the $n\mathrm{D}_{\mathrm{KL}}$ coincides with the earliest estimate; in 69 subjects the parse-based estimate does, and in 40 subjects the last upward 0 crossing does.

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. F., J. R., & Namboodiri, V. M. K. (2023). Few-shot learning: temporal scaling in 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.

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.

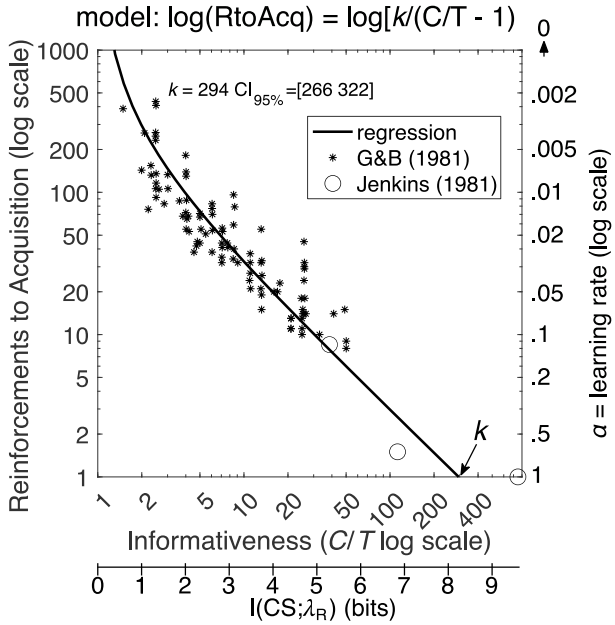
Thrailkill, E. A., Todd, T. P., & Bouton, M. E. (2020). Effects of conditioned stimulus (CS) duration, intertrial interval, and I/T ratio on appetitive Pavlovian conditioning. *Journal of Experimental Psychology: Animal Learning and Cognition, 46*, 243-255.

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.

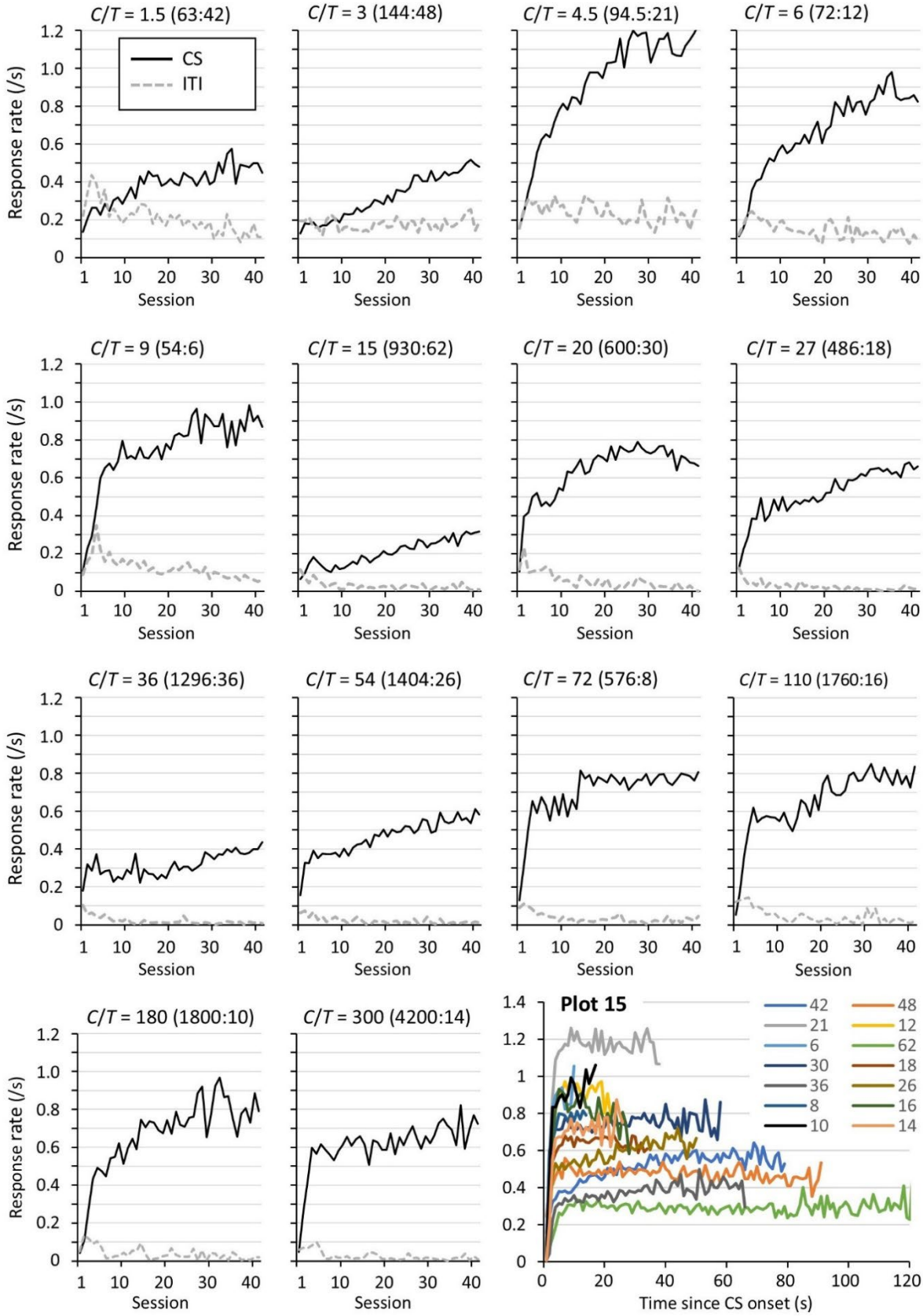
Ward, R. D., Gallistel, C. R., Jensen, G., Richards, V. L., Fairhurst, S., & Balsam, P. D. (2012). Conditioned stimulus informativeness governs conditioned stimulus-unconditioned stimulus associability. *Journal of Experimental Psychology: Animal Behavior Processes, 38*, 217-232.

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



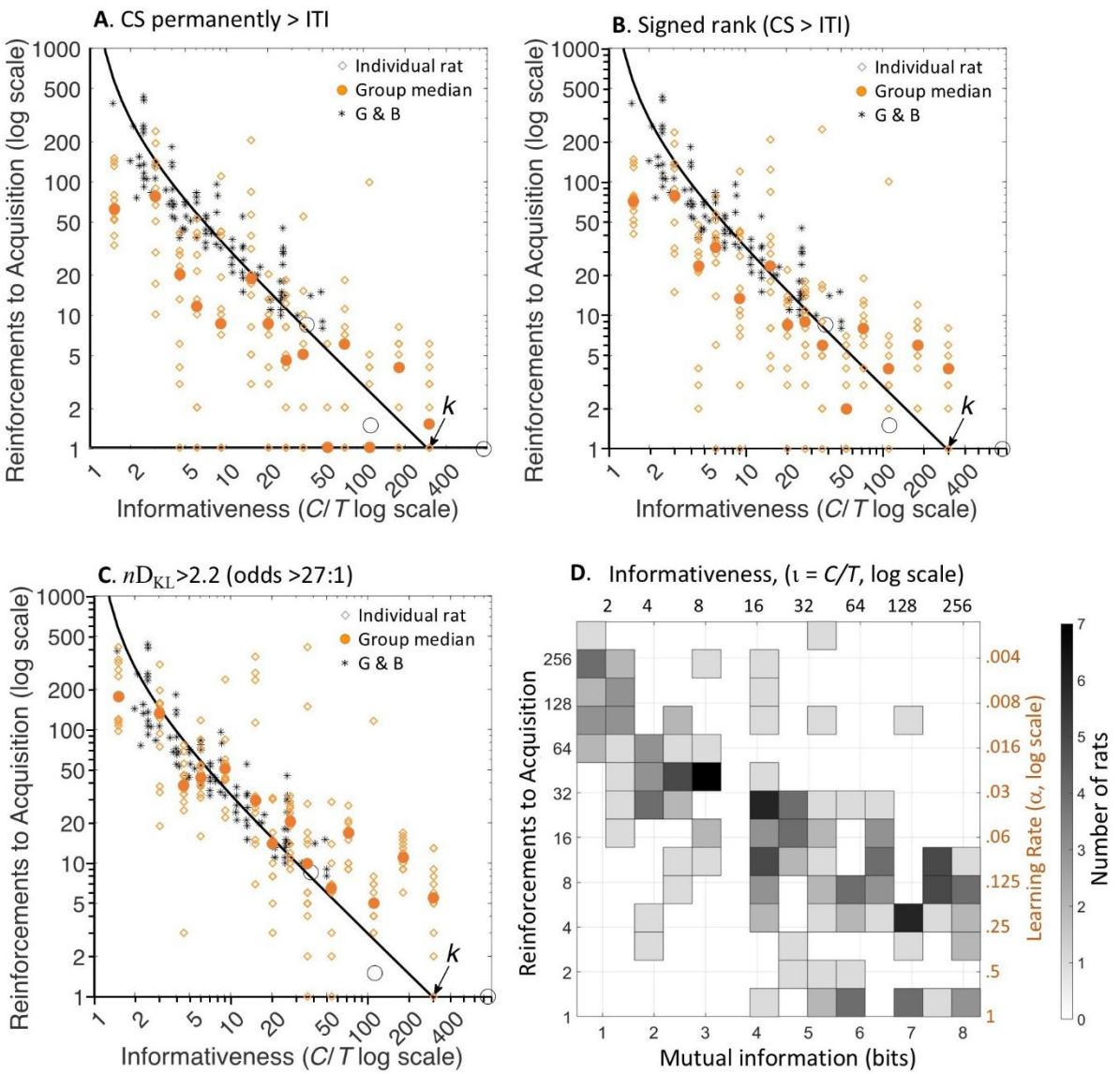
Note. *Asterisks show the data from Gibbon and Balsam's (1981) metanalysis of acquisition in pigeon acquisition. The two open circles are from Jenkins, Barnes and Barrera (1981) who used even bigger C/T ratios in some of their groups. The log of the informativeness is the mutual information, I, between 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).*

Figure 2. Mean response rate for each group across Sessions or across time since CS onset.



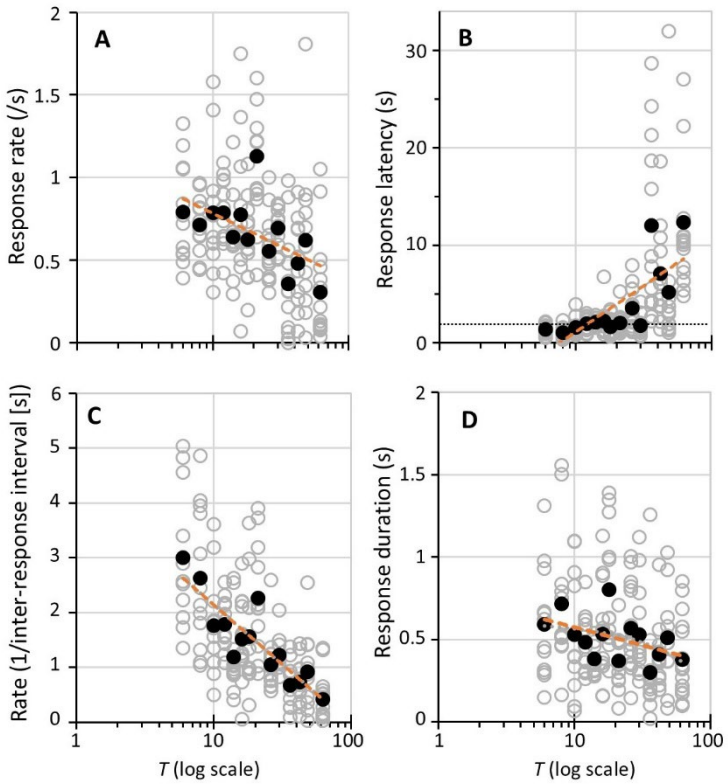
***Note**. There were 10 (groups with C/T ≤ 72) or 3 (groups with C/T ≥ 110) trials per session. Plot 15 (bottom right) shows the mean response rate per second during the CS, averaged from the last 5 sessions. Each group is identified by the length of the mean CS-US interval (T).*

Figure 3. Number of trials for acquisition of responding to the CS, plotted against CS informativeness (*C*/*T* ratio).



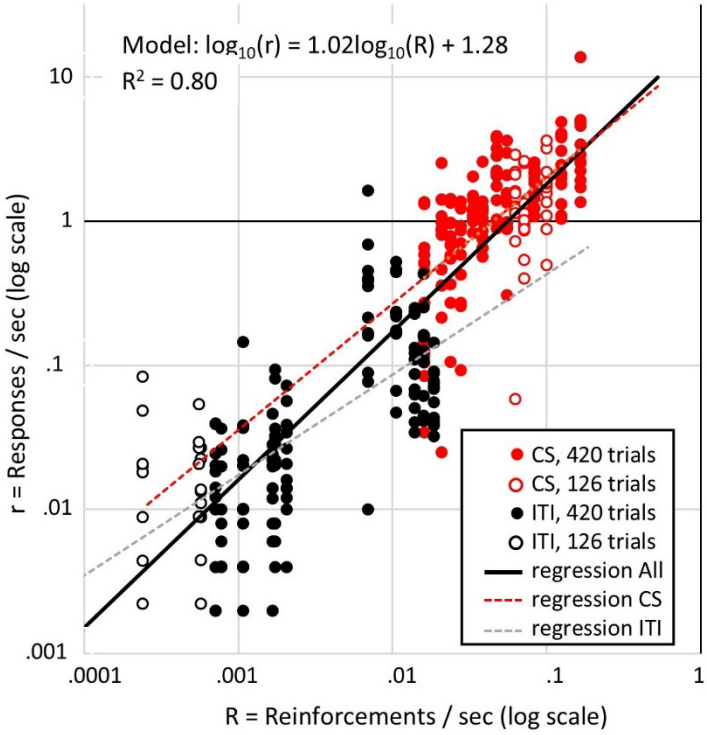
Note. *Number of trials to reach acquisition criterion for each group (filled orange circles) and each rat (unfilled orange diamonds) plotted against informativeness of the CS, superimposed onto the data (black asterisks) from Gibbon & Balsam (1981) and the regression line shown in Figure 1. The different plots show trials to acquisition using different acquisition criteria.* ***A*** *plots the trial after which the cumulative response rate during the CS is permanently greater than cumulative ITI response rate (not applying any statistical threshold for significance).* ***B*** *plots the trial on which the CS response rate is greater than the ITI response rate by a non-parametric signed-rank test with probability 0.3125 (equivalent to 3 out of 4 trials).* ***C*** *plots the trial on which the* nD_{KL} *is greater than 2.2, corresponding to odds more than 27:1 that the CS response rate was greater than the ITI rate.* ***D*** *is a tile plot of the same data as in C and includes the learning rate (right axis; reciprocal of the number of trials to acquisition) plotted against the mutual information (*$\log_2$ *of informativeness) for the 14 groups. The number of trials to acquisition has been divided into 14 bins of approximately equal width on a logarithmic scale represented by the 14 rows of tiles. The darkness of each tile shows the number of rats that reached criterion within that range of trials (see greyscale bar on right).*

Figure 4. Response metrics, averaged over the final 5 conditioning sessions, plotted against *T*.



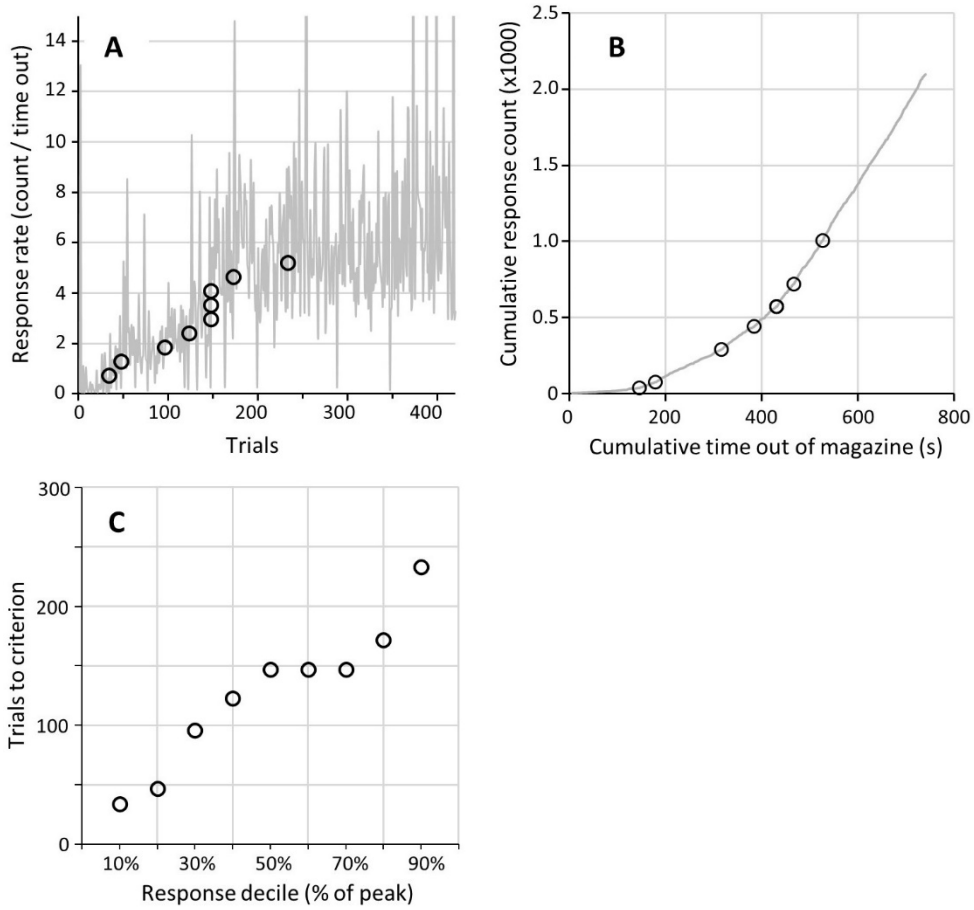
***Note**: Open grey circles show data for individual rats. Black filled circles show mean data for each group. The dashed orange lines are the best-fitting regression lines.*

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



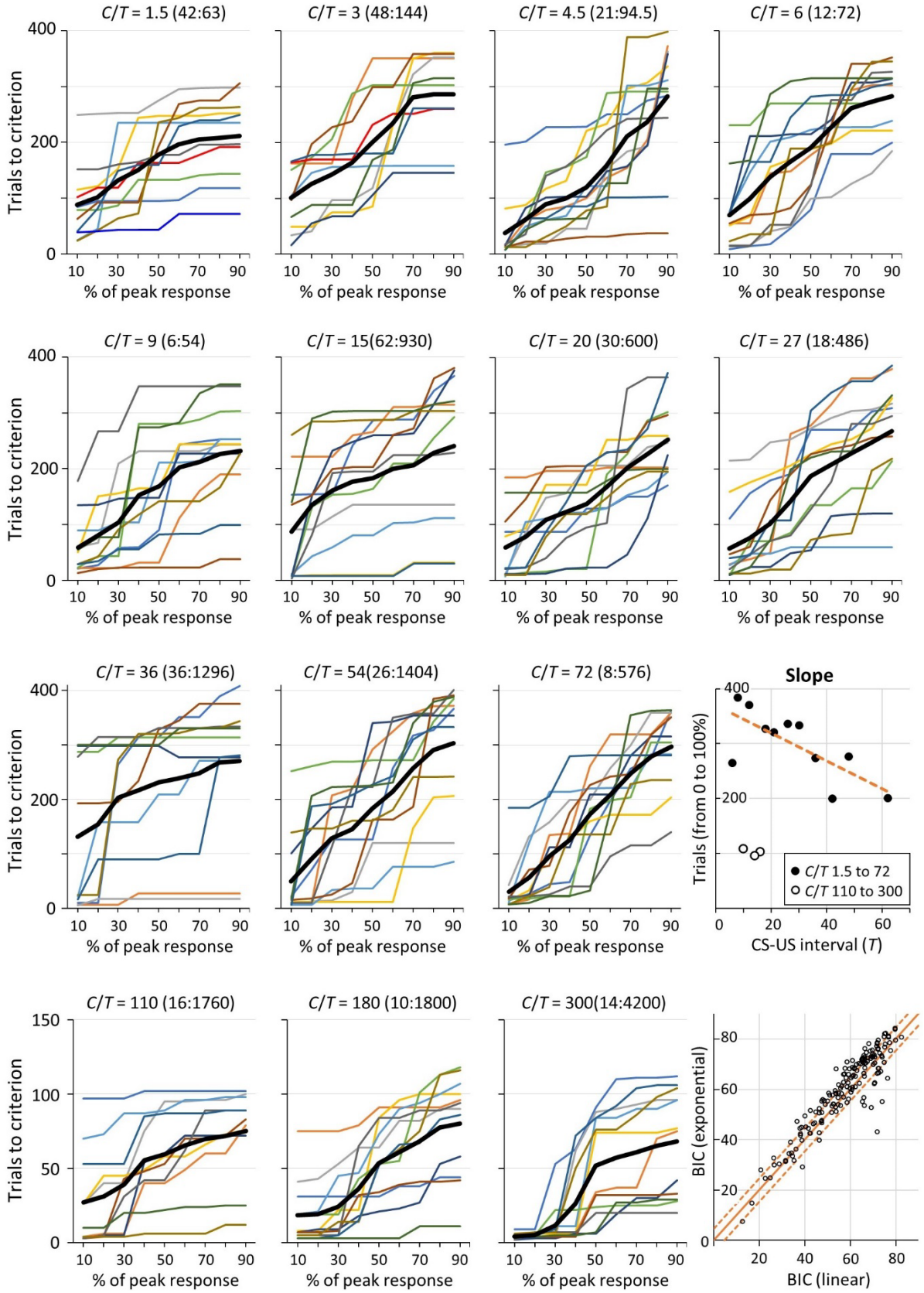
***Note**: Red circles show response rates during the CS (number of pokes after the first poke divided by the remaining CS time when the rat's head was out of the magazine) plotted against the CS reinforcement rate (1/T). Black circles show response rates during the inter-trial interval (ITI; calculated as number of pokes in the pre-CS interval divided by the pre-CS 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 of 72 or less), and open circles show data of rats given 126 reinforced trials (Groups with C/T ratios of 110 or more). The Model gives the equation for the solid black regression line (R^2 = 0.80). The slope of this line is virtually indistinguishable from 1 (R^2 = 0.80, when slope fixed at 1). The dashed red and grey lines show the regression lines fitted separately to the CS response rates and ITI responses rates.*

Figure 6. Response rate and trials to criterion for an individual rat.



Note. *The grey line in* ***A*** *shows the response rate on each trial for an individual rat (Rat 10 from Group C/T = 9). The response rate was calculated as the number of responses during the CS divided by the total time out of the magazine but excluding the latency to first response (i.e., the total time across which the animal could respond). In* ***B****, the cumulative response count across trials for the same rat is plotted against the cumulative opportunity to respond (cumulative time out of the magazine). The slope of this cumulative function was used to identify the trial on which the rat's response rate reached each decile (from 10% to 90%) of its peak response rate, as shown in* ***C****. These trials are also marked as circles on the response plots in A and B.*

Figure 7. 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.*

Figure 8. The parsed response rates during the CS (solid lines) and during the ITI (dashed lines) for the first 6 rats in the group with the lowest informative-ness (ι = 1.5).

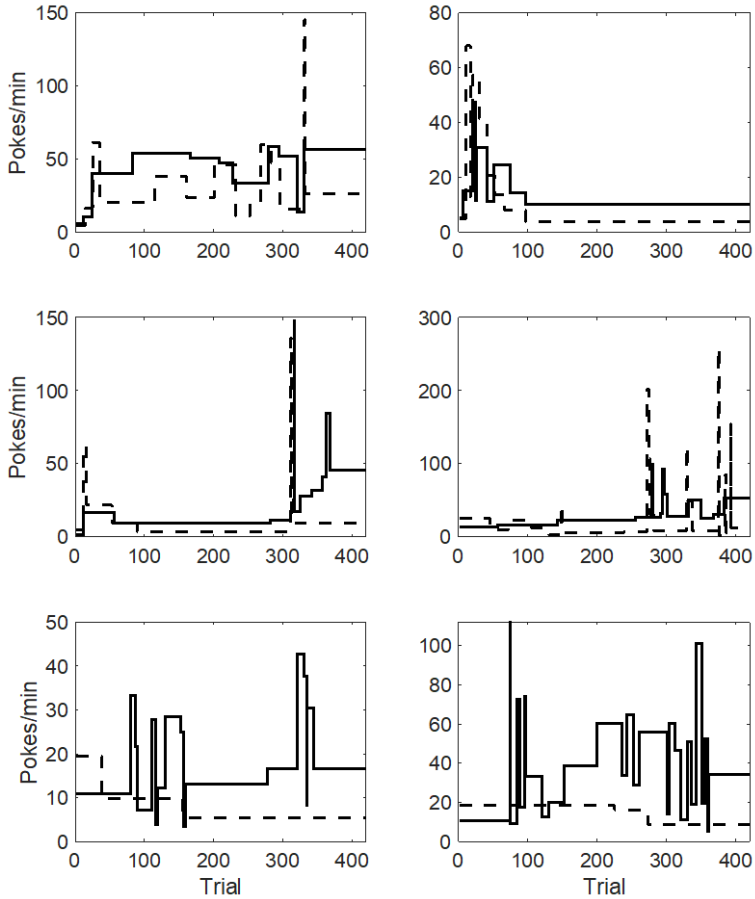


Figure 9. The parsed response rates during the CS (solid lines) and during the ITI (dashed lines) for the first 6 rats in the group with the highest informativeness (ι = 299).

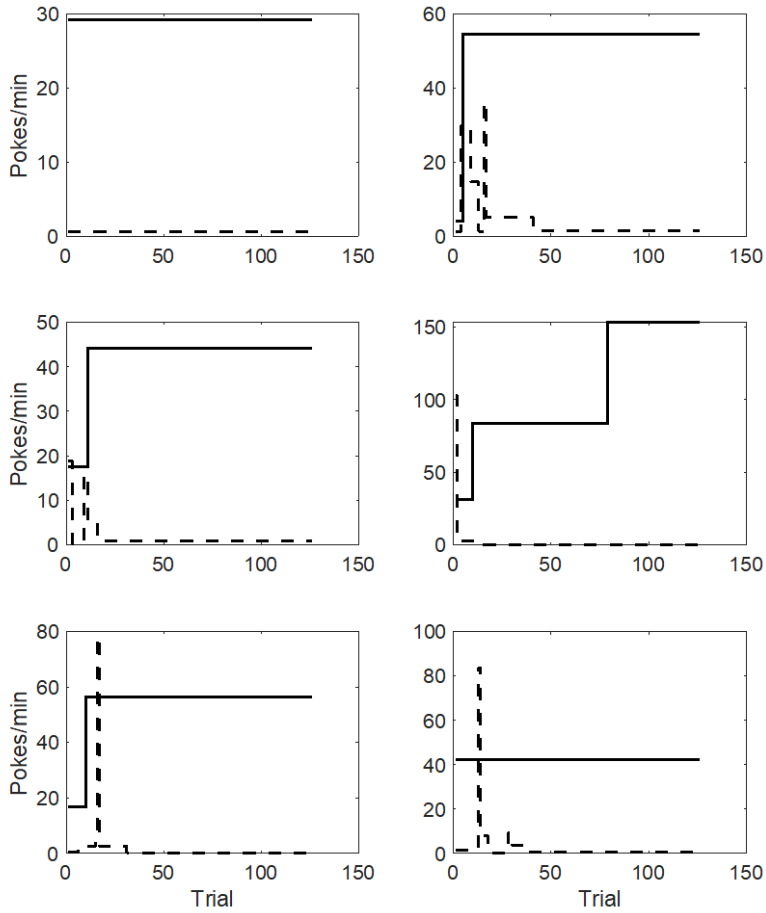
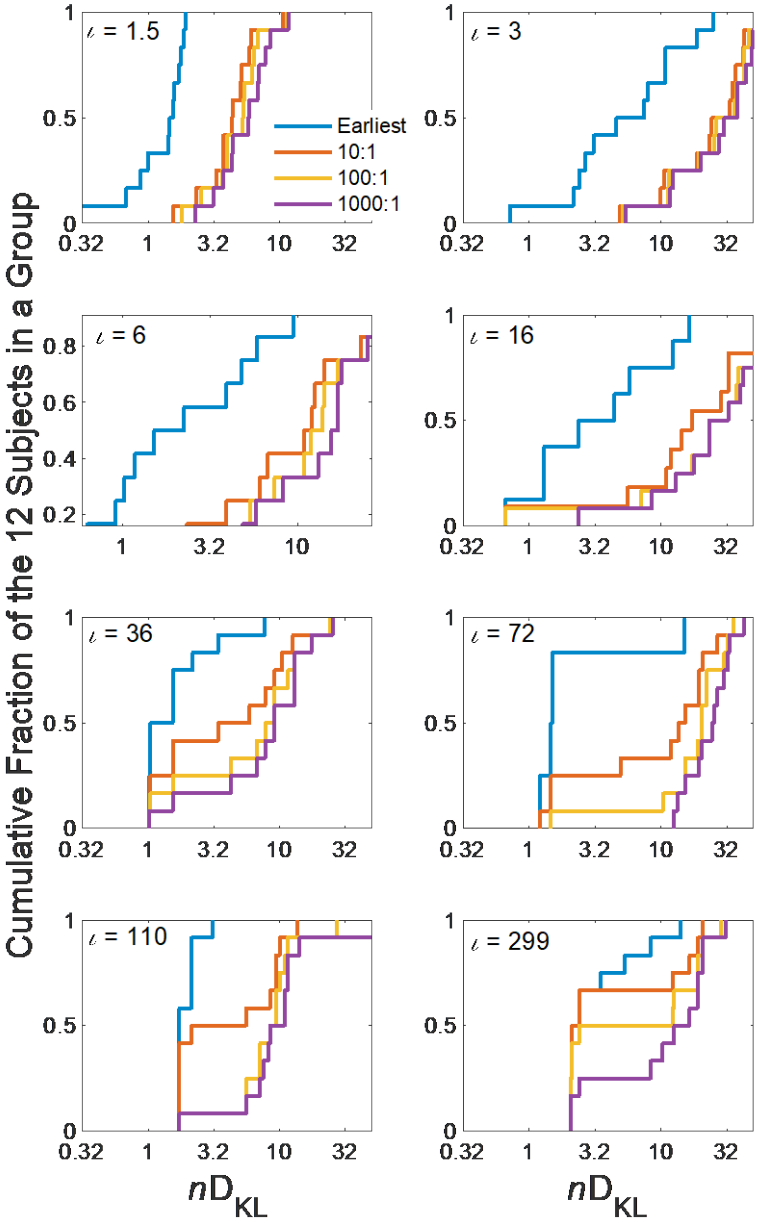
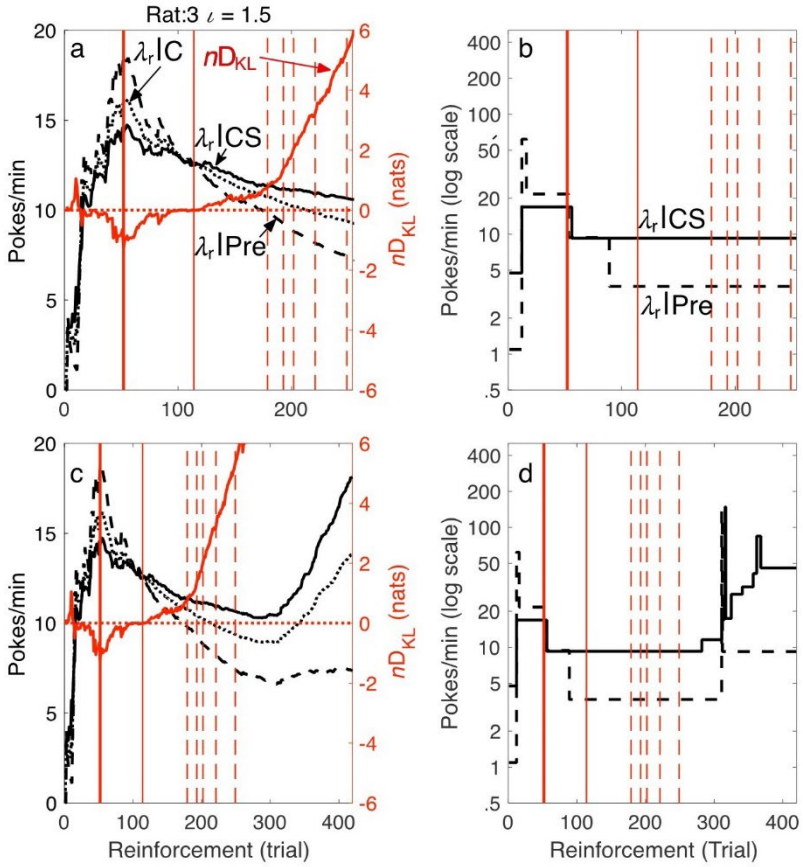


Figure 10. Cumulative distributions of the $n\mathrm{D}_{\mathrm{KL}}$ for the difference between the CS reinforcement rate and the contextual rate.



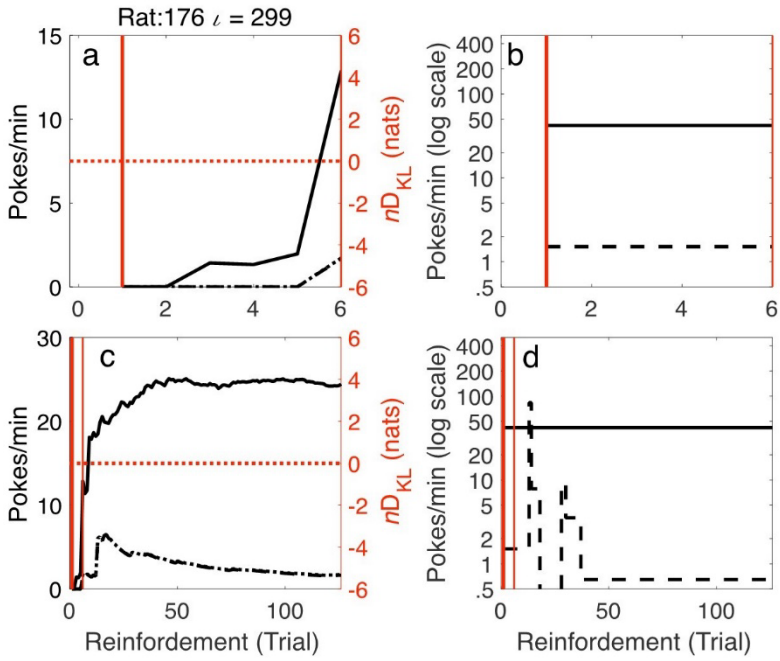
Note. *The cumulative distributions are plotted as of the reinforcement after which the behavioural evidence for acquisition satisfied increasingly stringent criteria (Earliest evidence, and Odds of 10:1, 100:1 and 1,000:1 in favour of the trustworthiness of the observed mutual information). The informativeness (ι) increases from left to right within a row of panels and top to bottom between panels.*

Figure 11.



Note. *In panels **a & c**, the rate of poking during the CS, denoted $\lambda_{r|CS}$, (solid black line), during pre-CS ITIs, $\lambda_{r|Pre}$ (dashed black line), and the contextual rate, $\lambda_{r|C}$ (dotted black line), plotted as functions of the number of reinforcements (1 R/trial). The red curve is the signed $n\mathrm{D}_{\mathrm{KL}}$ plotted against the right axis. The x-axis in a has been right-cropped to better reveal early changes. Panels **b & d** show the parsed estimates of $\lambda_{r|CS}$ and $\lambda_{r|Pre}$. In all plots, the vertical red lines mark different estimates of when CS-conditional responding appeared (see text).*

Figure 12.



Note. *Second example: from a subject in the group with the maximally informative protocol. The average pokes/min plots for the CS in Panels a and c (solid black curves) are cumulative poke count divided by cumulative CS duration; they include the single pokes made during CS 3 and CS 5. These first pokes were ignored by the parse algorithm.*