Response Rates Track the History
of Reinforcement Times

Justin A. Harris, Saba Gharaei, & Hannah L. Pincham

School of Psychology, University of Sydney, Australia

Abstract

When conditioning involves a consistent temporal relationship between the conditioned stimulus (CS) and unconditioned stimulus (US), the expression of conditioned responses within a trial peaks at the usual time of the US relative to the CS. Here we examine the temporal profile of responses during conditioning with variable CS-US intervals. We conditioned stimuli with either uniformly distributed or exponentially distributed random CS-US intervals. In the former case, the frequency of each CS-US interval within a specified range is uniform but the momentary probability of the US (the hazard function) increases as time elapses during the trial; with the latter distribution, short CS-US intervals are more frequent than longer intervals, but the momentary probability of the US is constant across time within the trial. We report that, in a magazine approach paradigm, rats' response rates remained stable as time elapses during the CS when the CS-US intervals were uniformly distributed, whereas their response rates declined when the CS-US intervals were exponentially distributed. In other words, the profile of responding during the CS matched the frequency distribution of the US times, not the momentary probability of the US during the CS. These results are inconsistent with real-time associative models, which predict that associative strength tracks the momentary probability of the US, but may provide support for timing models of conditioning in which conditioned responding is tied to remembered times of reinforcement.

Address for correspondence:

Justin Harris
School of Psychology
University of Sydney
Sydney, 2006
Australia
Email: justin.harris@sydney.edu.au

Many models of conditioning, such as that proposed by Rescorla and Wagner (1972), are based on an associative mechanism that tracks changes in the probability of the unconditioned stimulus (US). This mechanism provides a succinct account of the growth and decay of conditioning strength, as well as the distribution of that strength among different conditioned stimuli (CSs). By and large these models define the operations underlying learning at the level of the trial – changes to conditioning strength are computed at the end of the trial, and expressed as a one-dimensional value (e.g., associative strength) that controls responding on subsequent trials. While this level of description has served associative models well, it limits their theoretical and empirical scope. The notion of discrete trials is a conceptual convenience for the experimenter, but it is unlikely that the subject (such as a rat) parses the continuity of its experience in the same way. Further, there is a great deal of evidence that the production of conditioned responses (CRs) tracks the timing of the US relative to the CS. For example, when animals are trained with a fixed duration CS whose offset coincides with the US, known as "delay conditioning", the frequency of CRs increases over the time course of the CS, typically peaking near the time of US presentation (e.g., Davis, Schlesigner, & Sorenson, 1989; Kehoe & Joscelyne, 2005; Pavlov, 1927; Roberts, 1981; Smith, 1968; Williams, Lawson, Cook, Mather, & Johns, 2008). Indeed, recent evidence suggests that information about the timing of the US can be more important than contingency in determining whether an animal learns a CS-US association (Williams et al., 2008). Such within-trial features of conditioned behavior are not within the explanatory realm of simple trial-based models.

Demonstrations of timing in conditioned responding have inspired new theoretical approaches to associative learning. Some theorists have argued that subjects explicitly keep track of time during each trial, and conditioned responding is governed by a memory for the precise time at which the US occurred (e.g., Balsam & Gallistel, 2009; Gallistel & Gibbon, 2000; Gibbon, 1977; Kirkpatrick & Church, 2003). An alternative approach maintains the central features of associative processes, but attributes the evidence for response timing to a representation of the CS that is distributed across time. The most popular version of this approach, dating back to Pavlov (1927), assumes that the CS is represented by multiple elements with different temporal characteristics (e.g., Desmond & Moore, 1988; Sutton & Barto, 1981, 1990; Wagner & Brandon, 2001). Each element serves as a "micro-CS", and the elements that are most active at the time of the US gain the majority of associative strength according to a competitive learning rule like that proposed by Rescorla and Wagner (1972). Thus the passage of time within a trial is coded by a temporally-distributed representation of the CS, and it is this distribution that accounts for the emergence of timed responses to fixed duration CSs. Such "real time" extensions of the Rescorla-Wagner model have proved very successful in accounting for experimental demonstrations of timing in delay conditioning (Joscelyne & Kehoe, 2007; Kehoe, Horne, Macrae, & Horne, 1993; Williams et al., 2008). They have also been used to great effect in modeling the behavior of dopaminergic neurons in substantia nigra of awake monkeys (Ludwig, Sutton, & Kehoe, 2008). The activity of these neurons tracks in real time the discrepancy between the monkey's expectation of reward and its

2

experience of that reward, as if to code prediction error as defined in the Rescorla-Wagner learning rule (Schultz, Dayan, & Montague, 1997; Waelti, Dickinson, & Schultz, 2001).

The present paper presents findings from experiments that examine the temporal distribution of responding when the time of the US relative to the CS onset is variable. Random schedules of reinforcement produce relatively uniform patterns of responding over time (Catania & Reynolds, 1968; Kirkpatrick & Church, 2003). Here, we are interested in comparing two different types of randomly varying schedule in delay conditioning. In a *uniform* distribution, the duration of each CS, and thus the time of the US relative to the onset of the CS, varies uniformly over a defined range. Therefore, each CS-US interval within the range is equally frequent and the mean interval is the midpoint of the range. This can be contrasted with an *exponential* distribution in which shorter CSs are more frequent than longer CSs. The relationship between the CS-US interval, i, and its frequency, f, is defined as $f(i) = k \cdot e^{-ki}$, where $1/i$ is the average duration. The exponential distribution arises from any random process in which the probability of the outcome is constant across time. In terms of conditioning, this means that the probability of the US is constant at each moment during the CS. This can be contrasted with a uniformly distributed CS in which the momentary probability of the US increases as time elapses during the CS (the "hazard function"). To illustrate, consider a CS whose length is uniformly distributed between 1 and 10 s. After 1 s has elapsed in a given trial, the probability that the US will occur at that instant is 10%, but after 10 s have elapsed the probability of the US has reached 100%. The two types of frequency

distribution and the momentary probability of the outcome (US) associated with each distribution are shown in Figure 1. This illustrates the key difference between the distributions: for the uniform CS, the frequency of each CS-US interval is the same across the range of possible intervals; for the exponential CS, the probability of the US is constant as time elapses during the CS.

Because frequency and probability dissociate in different ways for exponentially and uniformly distributed events, these distributions provide an opportunity to test distinct predictions about the pattern of responding across the course of a CS. As mentioned earlier, according to associative models conditioning strength tracks changes in the probability of the US. Therefore, real time associative models predict that the distribution of associative strength across the duration of a CS will track the momentary probability of the US. Thus,
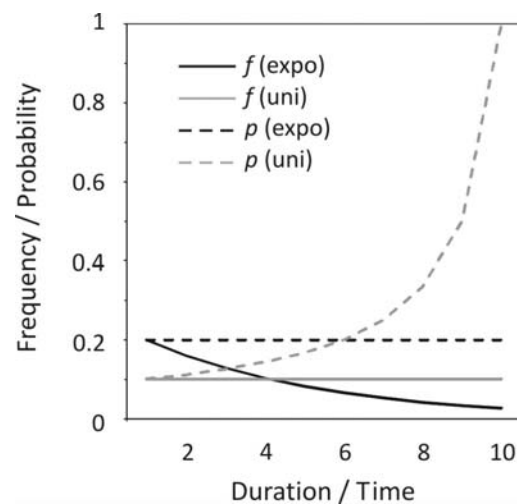

Figure 1. The frequency (*f*) of CS-US intervals conforming to either an exponential (expo) or uniform (uni) distribution, each with a mean duration equal to 5. The plot also shows the moment-by-moment probability (*p*) of the US as time elapses (the hazard function) during the exponential or uniform CS.

responding should be uniform across the duration of an exponential CS but should increase across the duration of a uniform CS. This is perhaps easiest to illustrate for a uniform CS, such as the one depicted in Figure 1. CS elements that are momentarily active at the beginning of the CS will be active at the beginning of every trial, but their activity will coincide with the US on only a small proportion of trials (whenever the trial is long, the activity of these elements will be low by the time the US occurs). Because these elements have a low reinforcement rate, they will acquire low associative strength. In contrast, elements that only become active when the CS has been on for a long time will be active rarely but will have a high rate of reinforcement, and will therefore acquire high associative strength. The continuous rise in reinforcement rate across the course of the uniform CS will mean that associative strength, and thus responding, will be progressively stronger across the duration of that CS. In contrast, the different elements of an exponential CS will have very different frequencies of activation, but will all have the same rate of reinforcement and, therefore, should acquire the same associative strength and support the same level of responding. In other words, the pattern of responding across the duration of a CS is predicted to resemble the probability (hazard) functions (dashed lines) in Figure 1.

To test the predictions described above, we have conducted four experiments that examine rats' responding to variable CSs in a magazine approach paradigm. In Experiment 1, rats were conditioned with CSs that had variable durations sampled randomly from uniform distributions with a mean duration of either 20 s or 60 s. Experiment 2 conditioned rats with variable CSs from distributions that

were approximately exponential and had a mean of either 20 s or 60 s. Experiments 3 and 4 made within-subjects comparisons between responding to a uniform and an exponential CS. The raw data from all four experiments can be downloaded from the website:

http://sydney.edu.au/science/psychology/staff/justinh/downloads/

Experiments 1 and 2

These experiments trained rats with four CSs, all with randomly varying durations. The durations in Experiment 1 were uniformly distributed, and those in Experiment 2 were from a distribution that was approximately exponential (these were only approximations of exponential distributions because we capped the maximum CS duration). In both experiments, two CSs had a mean duration of 20 s (ranging from 2 to 38 s in Experiment 1, and from 2 to 62 s in Experiment 2); the other two CSs had a mean duration of 60 s (ranging from 2 to 118 s in Experiment 1, and from 2 to 192 s in Experiment 2). The experiments included both shorter CSs (mean = 20 s) and longer CSs (mean = 60 s) so that we could determine whether the pattern of responding over the duration of the CS was independent of the actual level of responding. This meant we could tell whether the response pattern was constrained by performance factors such as ceiling or floor effects. In each experiment, one shorter and one longer CS were auditory stimuli, and one shorter and one longer CS were visual stimuli. In analyzing the data, we combined responses to the two shorter CSs together, and did likewise for the two longer CSs. In this way, the shorter and longer CSs were always matched for modality, and thus

hopefully they were well matched for salience or other factors that might influence conditioning and response rate.

Methods

Subjects

For each experiment, 16 experimentally naive male Hooded Wistar rats (Rattus norvegicus; approximately 12 weeks of age at the start of the experiment) were obtained from the Laboratory Animal Services breeding unit at Adelaide University. During the experiments, they were housed in groups of 8 in large white plastic tubs, measuring 26 x 59 x 37cm (height x length x depth), located in the animal colony maintained by the School of Psychology at the University of Sydney. They had unrestricted access to water in the home tubs. Three days prior to commencement of the experiment, their access to food was restricted to 2 hr per day (to commence half an hour after the end of the daily training sessions).

Apparatus

Rats were trained and tested in 16 Med Associates™ conditioning chambers measuring 28.5 x 30 x 25 cm (height x length x depth). The end walls of each chamber were made of aluminum; the sidewalls and ceiling were Plexiglas™. The floor of the chamber consisted of stainless-steel rods, 0.5 cm in diameter, spaced 1.5 cm apart. Each chamber had a recessed food magazine in the center of one end wall. A small metal cup measuring 3.5 cm in diameter and 0.5 cm deep was fixed on the floor of each food magazine. Attached to the food magazine was a dispenser delivering 45 mg food pellets (Noyes Formula P; Research Diets Inc, New Brunswick, NJ). Each chamber was enclosed in a sound- and light-resistant wooden shell. Throughout all sessions, fans located in the rear wall provided ventilation; the operation of these created a background level of noise measuring 70dB. Experimental events were controlled and recorded automatically by computers and relays located in the same room.

Two auditory and two visual stimuli were presented from four spatially separated sources. White noise (78dB) was presented from a speaker mounted on the wall of each operant chamber above and to the right of the food magazine. A tone (78dB and 2.9 kHz) was produced from a piezo buzzer positioned on the floor of the sound-attenuating shell behind each operant chamber. A flashing light (2 Hz; 3.0cd/m^2) was emitted by a 3x5 array of white LEDs, located on the floor of the sound-attenuating shell in front of the operant chamber. A steady light (30cd/m^2) was produced by an incandescent bulb mounted high on the back wall of the sound-attenuating shell.

Procedure

On the day before training began, the rats received a single 20-min magazine training session during which 20 food pellets were presented on a variable-time (VT) 1-min schedule, with no stimulus presentations. The rats then received daily conditioning sessions for the next 30 days (Experiment 1) or 36 days (Experiment 2). Each session consisted of intermixed presentations of four CSs, for a total of 48 presentations in Experiment 1 and 40 in Experiment 2. The order of presentations was randomized within each quarter of the session (i.e., each quarter contained equal numbers of all four CSs). The average inter-trial interval was 120 s (the intervals varied randomly according to an exponential distribution, but with a minimum of 40 s). The duration of each CS varied from trial to trial, but regardless of length, the termination of every CS coincided

with delivery of a food pellet. Two CSs, one auditory and one visual, had mean durations of 20 s; the other two CSs had mean durations of 60 s. The allocation of stimuli to these durations was counterbalanced across rats. In Experiment 1, the distribution of CS durations was uniform across the range 2 to 38 s for the two 20-s CSs, and was uniform across the range 2 to 118 s for the two 60-s CSs. In Experiment 2, the distribution of CS durations was approximately exponential across the range 2 to 62 s for the 20-s CSs, and 2 to 192 s for the 60-s CSs. In each session, the durations of each CS were randomly shuffled with the constraint that their mean duration within each quarter of the session approximated the overall mean duration for that CS (20 s or 60 s). Across all days the number of photo-beam interruptions by head entry into the magazine was recorded during each CS and during the 30-s pre-CS interval. Each session lasted approximately 120 min in Experiment 1 and 100 min in Experiment 2.

The data were analyzed as response rates during each second of the CS. All response rates were corrected for opportunity. This was done by counting the number of responses that a rat made in each 1-s time bin since the onset of a given CS, and dividing that by the number of CS presentations that lasted for at least that number of seconds.

Results

Experiment 1

Response rates to the CSs increased across the 25 days of conditioning, but changed little thereafter. The top panels of Figure 2 show the response rates to different portions of the 20-s and 60-s CSs across the course of training. Beyond the first 2-day block of training, response rates during the first 5-s of the 20-s
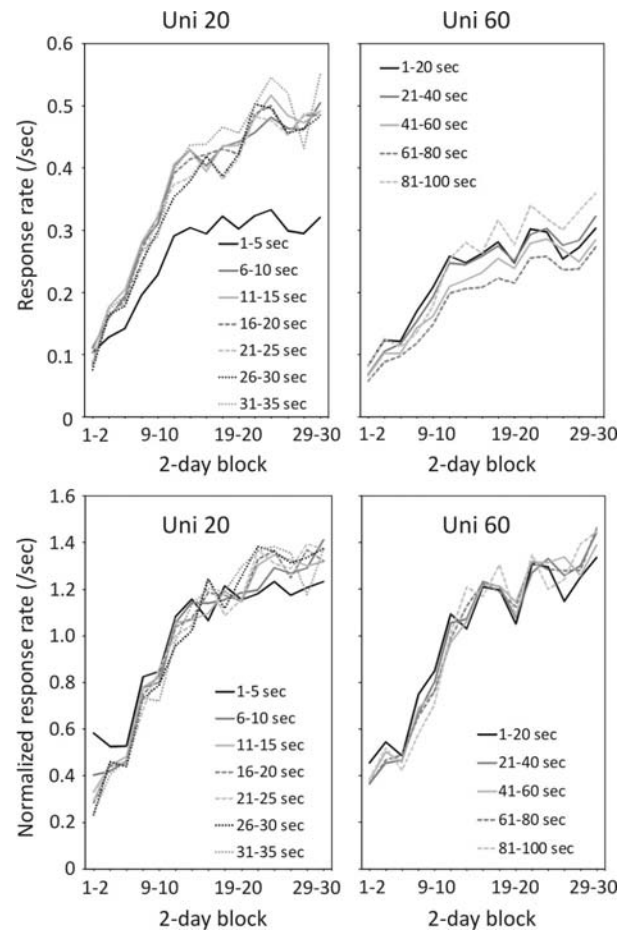

Figure 2. The two top plots show the mean response rates (responses/s) during different 5-s time periods within the 20-s uniform CSs (left) and during different 20-s time periods within the 60-s uniform CSs (right) across the 30 days of Experiment 1. In both cases, responding is acquired at similar rates for each of the different time periods. The similarity in shape of these acquisition functions is shown more clearly in the lower panels where response rates during each time period have been normalized by the mean response rate during that time period averaged over the entire 30 days of the experiment.

CSs were lower than at other time periods within the CS, but responding during all later periods (up to 35-s) were very similar across the entire experiment. Response rates during each 20-s period of the first 100 s of the 60-s CS were also very similar across the experiment. The similarity in level of responding across time

periods indicates that responding was uniform across time within the CS (as shown below in Figure 3). The similarity in the shape of the acquisition functions is shown more clearly in the lower panels of Figure 2. There, differences in overall level of responding have been removed by normalizing the response rates in each time period. This normalization was achieved by dividing the response rate in that period on each day by the mean response rate in that period averaged over all 30 days. Separate repeated measures ANOVAs were run on the normalized response rates to the 20-s and 60-s CSs to test for differences in the shape of the acquisition functions between each of the time periods within the CS. A significant day-by-time period interaction would establish that the shape of the acquisition functions was not the same for each time period. Here and elsewhere, Greenhouse-Geisser corrections were used whenever the data were found to violate the assumption of sphericity. This analysis confirmed that there was no interaction between the time period and days for either CS: $F(5,75) = 1.76$, $p = .13$, for the 20-s CS; $F(7.3,109.5) = 1.64$, $p = .13$, for the 60-s CS.

The left panel of Figure 3 plots the response rates per second during presentations of the 20-s CSs and the 60-s CSs, as well as during the pre-CS interval, averaged over all trials across the last 6 days of the experiment (Days 25-30). Response rates rose sharply from pre-CS baseline levels during the first 5 s of the CS presentations, reaching a higher rate for the 20-s CS than for the 60-s CS. Beyond the first 5 s of the CS, response rates remained relatively stable. To compare response rates over this part of the CS presentations, an ANOVA was conducted on the data from the 6th to the 35th second of the CSs. The analysis stopped at the 35th second because, up to but not beyond that
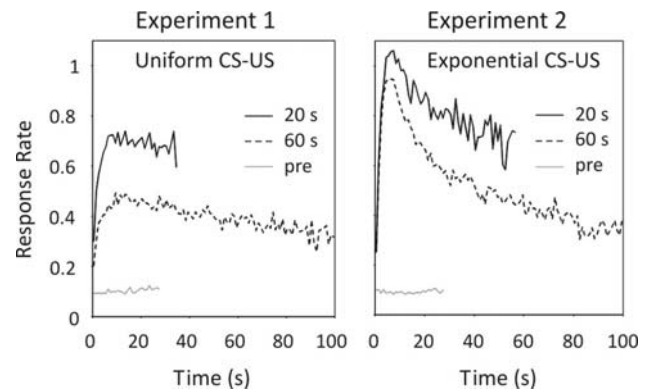

Figure 3. Mean response rates (responses/s) across time during CS presentations (black lines) or during the pre-CS interval (gray lines) in Experiments 1 and 2. The duration of the CSs varied randomly from trial-to-trial, with a mean of either 20 s (solid black line) or 60 s (dashed line). The frequency distribution of the CSs was either uniform (Experiment 1) or exponential (Experiment 2).

point, there was always at least one trial per day of each 20-s CS. This analysis confirmed that response rates to the 20-s CSs were significantly higher than to the 60-s CSs, $F(1,15) = 41.09$, $p < .001$. There was no significant main effect of time within the CS, $F(7,104.5) = 1.82$, $p = .091$, nor was there a significant interaction between time and CS duration, $F(8.6,129.5) = 1.10$, $p = .366$.

Experiment 2

Response rates to the CSs increased across the first 30 days of conditioning, after which they appeared to reach a plateau. The top panels of Figure 4 show the response rates to different portions of the 20-s and 60-s CSs in 2-day bins across the course of the whole experiment. In contrast to the results of Experiment 1 (Figure 2), by the end of training there were clear differences in the level of responding between each time period within the CSs. This is particularly evident for the 60-s CS, where response rates decreased for successively later portions of the CS. Despite

such differences in the level of responding, there was a similar shape to the acquisition function for each time period within the CS. This similarity is shown more clearly in the lower panels of Figure 4, where differences in overall level of responding have been removed by normalizing the response rates in each time period (the response rate in each time period on each day was divided by the mean response rate in that period averaged over all 36 days). To test for differences in the shape of the acquisition functions between each of the time periods within the CSs, separate repeated measures ANOVAs were run on the normalized response rates to the 20-s and 60-s CSs. These analyses confirmed that there was no interaction between the time period within the CS and the 2-day bins for either CS: $F(9.2,137) = 1.13$, $p = .344$, for the 20-s CS; $F(6.8,102) = 1.33$, $p = .244$, for the 60-s CS.

The right panel of Figure 3 plots the response rates during the 20-s and 60-s CSs, as well as the pre-CS interval, averaged over the last 6 days of Experiment 2 (Days 31-36). Response rates rose sharply during the first 5 s of the CS, but then began to decline thereafter. Response rates to the 20-s CS were higher than to the 60-s CS, though the evolution of responding over time was similar for the different CSs. To compare response rates between CSs, an ANOVA was conducted on the data from the 6th to the 55th second of the CS presentations. The analysis stopped at the 55th second because, up to but not beyond that point, there was always at least one trial per day of each 20-s CS. This analysis confirmed that response rates to the 20-s CSs were significantly higher than to the 60-s CSs, $F(1,15) = 58.10$, $p < .001$. There was a significant effect of time within the CS, $F(6.3,95) = 16.85$, $p < .001$, but there was not a significant interaction
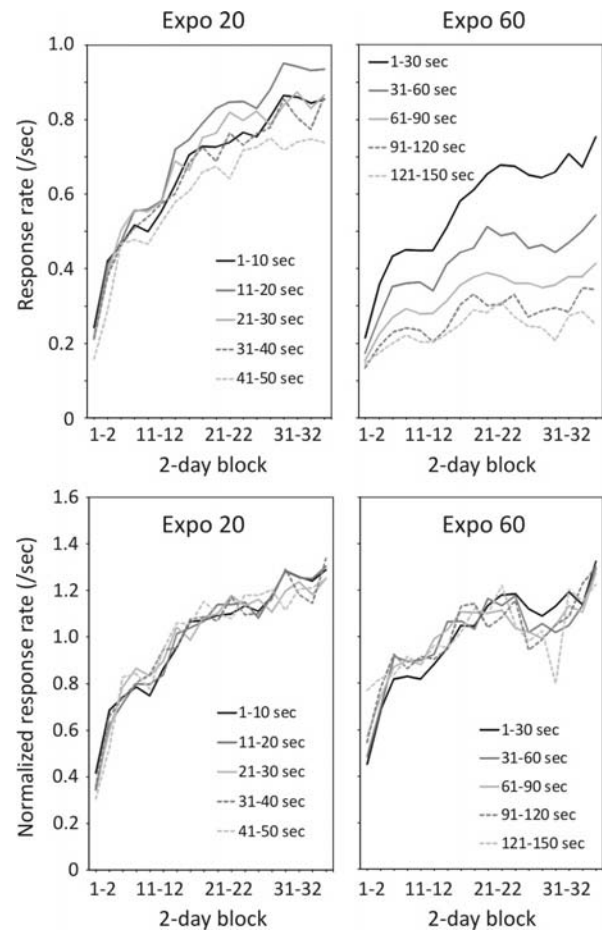

Figure 4. The two top plots show the mean response rates (responses/s) during different 10-s time periods within the 20-s exponential CSs (left) and during different 30-s time periods within the 60-s exponential CSs (right) across the 36 days of Experiment 2. In both cases, there are clear differences between the time periods in the level of responding by the end of training. Despite this, the shape of the acquisition function over days is similar for each of the different time periods. The similarity in shape of these acquisition functions is shown more clearly in the lower panels where response rates during each time period have been normalized by the mean response rate during that time period averaged over the entire 36 days of the experiment.

between the linear trend and CS, $F(9.4,141) = 1.24$, $p = .271$. Separate ANOVAs on the data from the two CSs confirmed there was a significant main effect of time within the 20-s

CS, $F(9.5,142) = 5.91$, $p < .001$, and within the 60-s, $F(4.4,66) = 19.12$, $p < .001$. Trend analyses within these ANOVAs (O'Brien & Kaiser, 1985) revealed that there were significant linear and quadratic trends over time within the 20-s CSs and within the 60-s CSs (smallest $F(1,15) = 7.75$, largest $p = .014$).

Fitting exponential functions to the response data from Experiment 2

In Experiment 2, response rates fell gradually as time elapsed during the exponential CSs, as confirmed by the significant linear trend over time for both CSs. Moreover, the rate of this decline also decreased progressively across the trial, as shown by the significant quadratic trend over time. Thus the shape of the response pattern across the CS resembled that of the underlying exponential distribution of CS-US intervals. As shown in Figure 5, the mean response rates to both CSs were well described by an exponential function that accounted for a large proportion of the variance (R^2s = .82 and .91). To test this conclusion, we compared how well exponential and linear functions could be fitted to the individual data from each rat. Functions were fitted to the response rates above baseline, calculated by subtracting each rat's mean response rate during the last 5-s of the pre-CS interval from its response rate for each second during the CS. To reduce noisy variability in these data, the response rates per second were collapsed into 4-s bins, starting from the 5th second after CS onset. The exponential function was defined as: $f(t) = A\lambda e^{-\lambda t}$, where t is time, λ is the rate parameter of the exponential function, and A scales the height of the function. For the 20-s CS, the R^2 value for the exponential function was higher than for the linear function in all but one rat (mean R^2s across rats = 0.55 and 0.51, for the exponential

and linear functions respectively), and this difference was statistically significant by paired-samples t-test, $t(15) = 3.51$, $p = .003$. For the 60-s CS, the R^2 for the exponential function was higher than for the linear function in all rats (mean R^2s = 0.68 and 0.60, for the exponential and linear functions respectively), and this difference was statistically significant, $t(15) = 7.71$, $p < .001$.

Discussion

These experiments confirmed that rats' response rates are sensitive to reinforcement rates: the rats responded more to CSs with higher mean reinforcement rates than to CSs with lower reinforcement rates (Kirkpatrick & Church, 2003). The more important observation concerns how response rates across the duration of a CS were affected by the distribution of the CS's reinforcement times. CS durations in Experiment 1 varied according to a
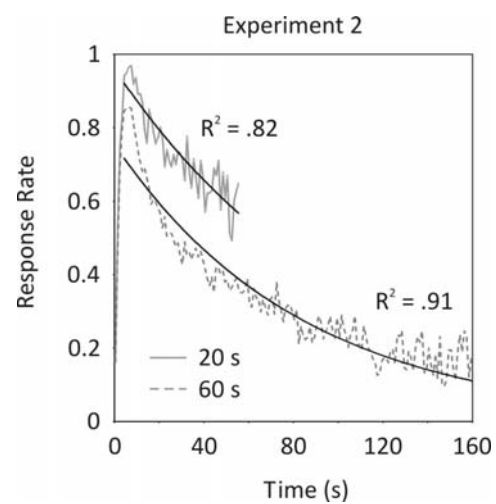

Figure 5. Dashed grey lines plot the mean response rate above baseline for each second during presentations of the 20-s and 60-s exponential CSs in Experiment 2. The solid black lines plot exponential functions fitted to these data beyond the 5th second of the CS. The proportion of variance explained by each exponential function is shown as R^2.

uniform distribution, and the rats' response rates to these CSs were largely uniform over time. By contrast, CS durations in Experiment 2 varied according to an exponential distribution, and the rats' response rates to these CSs changed over the course of the CS, initially rising sharply within the first 5 s, then declining gradually as the CS presentation continued. In both experiments, the pattern of responding over time was similar for the shorter (20-s) and longer (60-s) CSs, even though these CSs evoked different levels of responding. This is important because it shows that the pattern of responding over time was independent of its location on the response scale.

Taken together, the results of both experiments suggest that rats' response rates track the history of reinforcement times rather than the instantaneous probability of reinforcement during the CS presentation. These results are problematic for real-time associative models which predict that the associative strength of temporally-distributed CS elements will track the momentary probability of the US. That is, for a uniformly reinforced CS, elements that are active at the onset of the CS will have lower reinforcement rates, and thus lower associative strength, than elements active later in the CS. Similarly, for a CS with exponentially distributed reinforcement times, the reinforcement rate is the same for each element regardless of its time of activation, and thus the associative strength should be uniformly distributed across the duration of the CS. However, this logic rests on the assumption that conditioning of all elements was equivalently close to their terminal strength. If elements activated at the end of the CS had undergone less conditioning because there were fewer trials that activated those elements, this could lead to lower than

predicted response strength at later stages of the CS. In light of this, the data presented in Figures 2 and 4 are important in showing that the acquisition functions were equivalent for all time points within the trial. For example, in the top right plot of Figure 4 it is clear that response rates in the late time periods of the 60-s CSs (beyond 90 s) were as close to their asymptote as response rates at any other time point in those CSs. This observation was confirmed when overall differences in level of responding were removed, as shown in the lower right plot of the same Figure. Thus elements near the end of the CS were as close to their terminal level of conditioning as elements near the start of the CS.

Experiments 3 and 4

In the previous experiments, rats' response rates declined across the course of CS presentations when the frequency distribution of CS durations, and thus reinforcement times, was exponential (Experiment 2), but response rates were comparatively stable across the CS when the distribution of CS durations was uniform (Experiment 1). The aims of Experiments 3 and 4 were to provide within-experiment confirmation that exponential and uniform distributions lead to different patterns of responding. Thus each experiment investigated the temporal profile of responding when rats were conditioned with one uniformly distributed CS and one exponentially distributed CS. To facilitate successful discrimination, the CSs had different mean durations. In Experiment 3, the durations of one CS (Expo20) were exponentially distributed with a mean of 20 s, and the durations of the other CS (Uni60) were uniformly distributed with a mean of 60 s. In Experiment 4, the mean and shape of the CS distributions were switched; the exponentially

distributed CS had a mean of 60 s (Expo60) while the uniform CS had a mean of 20 s (Uni20).

The data of both experiments were initially analyzed in similar manner to that described for Experiments 1 and 2. We compared the trends in responding over time (starting 5 s after the onset of the CS) to determine whether responding to the exponential CS showed a steeper decline than responding to the uniform CS. This was followed by an analysis that aimed to test a simple explanation of the source of this difference between responding to the exponential and uniform CSs. According to this explanation, the difference could arise if the rats' responses to a CS on a given trial track the timing of food delivery on the previous trial with that CS. That is, if the peak response rate on trial n occurred at roughly the time when food was delivered on trial n-1, then response rates averaged over many trials would resemble the underlying distributions of the CS durations – average response rates would appear uniform if the timing of food on each previous trial was uniformly distributed, and average response rates would appear exponential if the timing of food on each previous trial was exponentially distributed. To test this hypothesis, the data on each trial were transformed so that, instead of being aligned to the start of the trial, they were aligned relative to the time at which food was delivered on the previous trial with that CS. If the rats' responses do track the timing of food delivery on the previous trial, we would expect to see a consistent increase in responding to peak at the time of food delivery on the previous trial.

Methods

Subjects and Apparatus

In each experiment, there were 16 experimentally naive male rats (approximately 3 months of age), of the same strain and source, and housed in the same manner as in Experiments 1 and 2. They were trained in the same chambers, but with only two stimuli: the white noise and the steady light. The allocation of these stimuli to the two CSs of the experiment was counterbalanced between rats.

Procedure

Rats received magazine training as in Experiment 1. They then received daily conditioning sessions with two variable-duration CSs, with each session lasting approximately 100 min. The sessions consisted of 20 presentations of one CS, with a mean duration of 20 s, intermixed with 20 presentations of a second CS with a mean duration of 60 s (the two CSs were randomly intermixed with the constraint that each quarter of the session contained an equal number of each CS). Every CS presentation was followed by the delivery of a single food pellet. The inter-trial interval varied exponentially with a mean of 120 s. In Experiment 3, the distribution of stimulus durations was approximately exponential (range 2 to 85 s) for the 20-s CS and uniform (2 to 118 s) for the 60-s CS. In Experiment 4, the distribution of stimulus durations was uniform for the 20-s CS (2 to 38 s) and approximately exponential for the 60-s CS (range 2 to 240 s). Rats in Experiment 3 received 30 days of conditioning; Rats in Experiment 4 received 34 days of conditioning. The number and duration of photo-beam interruptions by head entry into the magazine was recorded during each CS and during the 30-s pre-CS interval.

Results

Experiment 3

Response rates to both CSs increased steadily across conditioning, and, as for Experiments 1 and 2, the speed with which responding approached its asymptote across days was very similar for each time interval within the CS. Analyses were conducted on the data from the last 10 days of training (Days 21 to 30). The left panel of Figure 6 shows the response rates across time to both CSs. It is clear that the rats responded more vigorously to the CS with the higher reinforcement rate (Expo20), but that the pattern of responding across the trial differed between the two CSs, with response rates declining to Expo20 but remaining relatively constant to Uni60. These observations were supported by ANOVA. The analysis was confined to response rates within a 50-s window extending from the 6th to the 55th second of the CS presentations. Overall, response rates to Expo20 were significantly higher than to Uni60, $F(1,15) = 27.65$, $p < .001$. There was also a significant overall effect of time, $F(5.9,88) = 5.97$, $p < .001$, and significant interaction between CS and time, $F(9.9,149) = 3.31$, $p = .001$. Follow-up trend analyses (O'Brien & Kaiser, 1985) revealed a significant linear trend across time, $F(1,15) = 20.32$, $p < .001$, and a significant interaction between this linear trend and CS, $F(1,15) = 31.17$, $p < .001$. This interaction confirms that there was a much greater decline in responding to Expo20 than to Uni60 over the same time period. Analyses conducted on each CS separately established that there was a significant effect of time for both Expo20, $F(6.8,102) = 5.28$, $p < .001$, and Uni60, $F(8,121) = 3.27$, $p = .002$. The linear trend over time was significant for both CSs, smaller $F(1,15) = 15.68$, $p = .001$, but the quad-
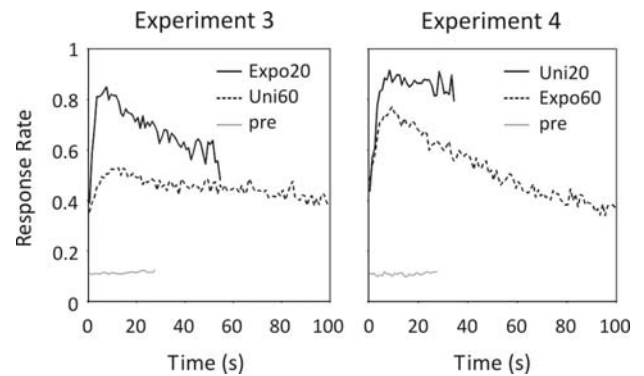

Figure 6. Mean response rates (responses/s) across time during CS presentations (black lines) or during the pre-CS interval (gray lines) in Experiments 3 and 4. In Experiment 3, one CS (Expo20) varied randomly according to an exponential distribution with a mean of either 20 s, and the other (Uni60) varied according to a uniform distribution with a mean of 60 s. In Experiment 4, the uniform CS had a mean of 20 s (Uni20) and the exponential CS had a mean of 60 s (Expo60).

ratic trend was not significant for either CS, larger $F(1,15) = 2.03$, $p = .175$.

Experiment 4

Response rates to both CSs increased across the first 20 days of conditioning, but remained relatively stable thereafter. As in the previous experiments, the speed with which responding approached its asymptote across training days was very similar for each time intervals during the CS. Analyses were conducted on the data from the last 10 days of training (Days 25 to 34). The right panel of Figure 6 shows the response rates to both CSs. Once again the rats responded more vigorously to the CS with the higher reinforcement rate (Uni20). As in Experiment 3, the pattern of responding across the trial differed between the two CSs, with response rates remaining relatively stable across presentations of Uni20 but declining across presentations of Expo60. These observations were supported by an ANOVA conducted on response rates over a 30-s

window extending from the 6th to the 35th second of the CS. Response rates to Uni20 were significantly higher than to Expo60, $F(1,15) = 13.73$, $p = .002$. There was also a significant overall effect of time, $F(7.1,106) = 5.60$, $p < .001$, and significant interaction between CS and time, $F(8.6,130) = 2.77$, $p = .006$ Follow-up trend analyses (O'Brien & Kaiser, 1985) revealed a significant overall linear trend across time, $F(1,15) = 23.39$, $p < .001$, and a significant interaction between linear trend and CS, $F(1,15) = 16.63$, $p < .001$. The interaction confirms that there was a steeper decline in responding to Expo60 than to Uni20. Analyses conducted on each CS separately established that there was a significant effect of time for Expo60, $F(5.5,83) = 24.68$, $p < .001$, but not for Uni20, $F(7.2,107) = 1.28$, $p = .27$. Both the linear and quadratic trends over time were significant for Expo60, smaller $F(1,15) = 21.88$, $p < .001$, whereas neither trend was significant for Uni20, larger $F(1,15) = 2.39$, $p = .143$.
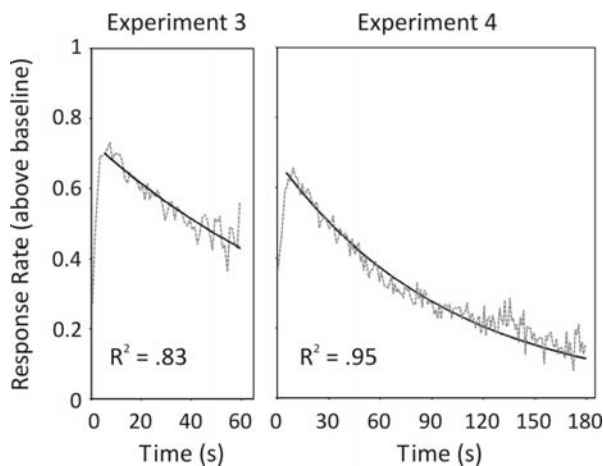

Figure 7. Dashed grey lines plot the mean response rate above baseline for each second during presentations of the 20-s exponential CS in Experiment 3 (left) and during presentations of the 60-s exponential CSs in Experiment 4 (right). The solid black lines plot exponential functions fitted to these data beyond the 5th second of the CS. The proportion of variance explained by each exponential function is shown as R^2.

Fitting exponential functions to the response data

In both experiments, the change in response rates as time elapsed during the two exponential CSs resembled the shape of the underlying exponential distribution of CS-US intervals. As for Experiment 2, we fitted exponential functions to the data from the two exponential CSs (Expo20 in Experiment 3, and Expo60 in Experiment 4). As shown in Figure 7, the mean response data to both CSs were well described by an exponential function that accounted for a large proportion of the variance (R^2s = .83 and .95). To test this conclusion, we compared the fit of an exponential and a linear function to the individual data from each rat, as described for Experiment 2. For Expo20 from Experiment 3, the R^2 values for the exponential function were significantly higher than for the linear function, $t(15) = 4.92$, $p < .001$ (mean R^2s across rats = 0.55 and 0.50, for the exponential and linear functions respectively). Similarly, for Expo60 from Experiment 4, the R^2 values for the exponential function were significantly higher than for the linear function, $t(15) = 4.63$, $p < .001$ (mean R^2s = 0.81 and 0.73, for the exponential and linear functions respectively).

Discussion

These two experiments have shown that rats' response rates to CSs are sensitive to the distribution of reinforcement times, corresponding here to the termination of the CS. When trained with two variable CSs, one from an exponential distribution and the other from a uniform distribution, the rats' response rates resembled the shape of the different distributions, showing stable response rates over time during the uniform CS and declining response rates during the exponential CS. The same difference was observed whether the

13

exponential CS had a shorter mean duration than the uniform CS, or vice versa.

General Discussion

The experiments presented here show that the strength of conditioned responding across the course of a variable CS tracks the frequency of the US at each moment during the CS. Response rates were uniform when the CS duration, and thus the timing of the US, was uniformly distributed; response rates decreased when the distribution of CS durations was exponential, such that longer durations were less frequent than shorter durations. Indeed, analyses conducted on the data from Experiments 2, 3 and 4 showed that response rates during each of the exponential CSs corresponded closely to an exponential function.

These observations are at odds with an existing literature investigating changes in the rate of conditioned responding during variable interval (VI) schedules of reinforcement. The common observation has been that response rates increase as time elapses during uniformly distributed VIs (Catania & Reynolds, 1968; Church, Lacourse, & Crystal, 1998; Harzem, Lowe, & Priddle-Higson, 1978; Lund, 1976), and remain at a constant level as time elapses during exponentially distributed VIs (Catania & Reynolds, 1968; Harzem et al., 1978; Kirkpatrick & Church, 2003). Most of these studies measured instrumental responses for food (bar pressing in rats, key pecking in pigeons) during VI schedules. The temporal control over those instrumental responses may differ from the magazine entry responses measured here that are, at least nominally, Pavlovian (but see Farwell & Ayres, 1979; Holland, 1979, for

evidence that rats can acquire magazine responding on an omission schedule). Consistent with this distinction, Schwartz (1978) compared pigeons' Pavlovian responses to a signal key with their instrumental responses to a separate operant key during VI schedules based on an exponential distribution. He observed that, while response rates on the operant key were stable across the interval, responses on the signal key declined. The same decline in response rate on the signal key was observed in a yoked group of pigeons that were not provided with an operant key.

While a distinction between instrumental and Pavlovian responses may go some way towards explaining the discrepancy between the current findings and those reported previously, this cannot be the whole story. One previous study to have reported stable response rates across exponentially distributed VIs (Kirkpatrick & Church, 2003) measured the same magazine responses in rats that we have measured here. Nonetheless, our observation that rats' magazine responses decrease as time elapses during an exponential CS is not without precedent. Kirkpatrick and Church (2000) and Church and Lacourse (2001) reported that rats' response rates to an exponentially varying CS were highest immediately after CS onset, and decreased as time elapsed during the CS. There is one important feature that distinguishes the experiments in these two studies and the present experiments from the earlier studies that reported stable response rates to exponentially varying CSs and increasing response rates to uniformly varying CSs. The experiments we present here, as well as those by Kirkpatrick and Church (2000) and Church and Lacourse (2001), measured responses during a discrete CS that was presented for a variable duration and followed by

14

Response rates track reinforcement times

reinforcement, and each CS presentation was separated by a variable inter-trial interval. In contrast, all of the other studies discussed above used a continuous conditioning schedule with no discrete CS, and response rates were measured across the full length of each inter-reinforcement interval. This distinction introduces two relevant differences between the experiments. First, in the continuous VI schedules (without a CS), time in the interval is set by a memory trace of the previous reinforcement. The mechanisms by which an animal keeps track of time during these trace intervals may well differ from its mechanism for tracking the passage of time during a delay interval that is defined by an ongoing CS. Second, the animal's behavior during the initial portion of the trace interval will be affected by its unconditioned responses to the previous US. For example, immediately after the delivery of reinforcement, instrumental responses will be suppressed as the animal retrieves and consumes the food pellet. Such unconditioned responses do not confound interpretation of conditioned responses during a discrete CS that is presented some time after the previous reinforcement. Therefore, the present observation that response rate tracks the history of reinforcement times may be specific to Pavlovian responses to a discrete CS during a delay conditioning procedure; it may not be true of instrumental responses during VI schedules or responses made across a variable trace interval.

The findings reported here are problematic for real-time associative models that attribute changes in responding across the course of a CS to differential conditioning of temporally distributed CS elements (Desmond & Moore, 1988; Sutton & Barto, 1981, 1990; Wagner & Brandon, 2001). The type of competitive learning rule (Rescorla & Wagner, 1972) typically used by these models distributes associative strength among the elements of the CS in proportion to their rate of reinforcement. Therefore, because associative strength across the CS should track changes in the probability of the US as time elapses, associative strength will increase across a uniform CS and will remain constant across an exponential CS. To confirm that associative models with temporally-distributed elements do indeed make these predictions, we have simulated this operation using the Rescorla-Wagner rule to update the associative strength of temporally distributed CS elements in real time. Figure 8 shows simulations for a uniform CS and an exponential CS. In these simulations, gamma functions were used to describe the rise and fall of activity of the CS elements (shown in the top panel in Figure 8). We have tested numerous other activation profiles, both symmetrical (e.g., a Gaussian distribution) and asymmetrical (e.g., an exponential distribution), which have all produced simulated outputs very similar to those shown here. We also obtain very similar simulated results if we completely remove the initial element (i.e., the element with an exponential activation profile shown in Figure 8). In short, our analysis indicates that, as long as CS elements are sufficiently distributed in time to account for the time course of responding to a fixed duration CS, this process will also distribute associative strength across a variable CS in proportion to the momentary change in probability of the US.

Our observation that conditioned responding tracks the distribution of US times, rather than US probability, supports an alternative class of conditioning models based on the general proposal that the production of conditioned responses is linked to a memory of
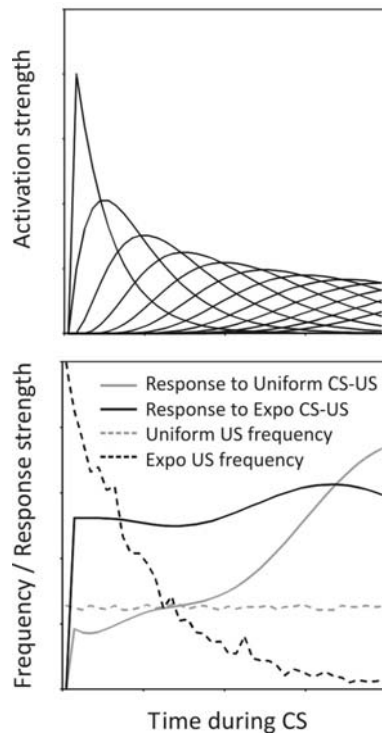
15

Figure 8. Top: Gamma functions describing activation strength of successive CS elements across time. Bottom: Simulations of response strength across the duration of a CS that has been conditioned with either a uniformly or exponentially varying CS-US delay interval. The frequency of the US at each time since CS onset is shown by the dashed line. The solid lines show simulated output strength for CSs comprised of temporally distributed elements with activation functions shown in top plot. During each trial, the associative strength of each element is updated in real time using a summed error term (Rescorla & Wagner, 1972). The output, representing simulated response strength, is the sum of each element's associative strength multiplied by its momentary activation strength.

the time of reinforcement on previous conditioning trials. A particularly influential example of such a theory, known as "Scalar Expectancy Theory" (SET), proposes that the decision to respond during a trial is based on a ratio between the current elapsed time in the trial and the remembered time of the US, such that the subject begins responding when this

ratio exceeds a threshold (Gibbon, 1977). One pertinent aspect of that theory is the suggestion that responding on a given trial is based on a randomly retrieved memory from all previous trials (Gibbon, Church, Fairhurst, & Kacelnik, 1988). In this case, the temporal pattern of response rates averaged over trials should resemble the frequency distribution of US times (the unbroken lines in Figure 1) because the probability of retrieving a particular memory of reinforcement will depend on the frequency of that reinforcement time. In other words, this process explains the current observation that response rates match the distribution of US times.

This apparent match between theory and data is weakened if, as typically assumed by SET, the error variance in the remembered time of reinforcement scales with the time interval itself, such that the error is greater for longer intervals but the ratio of variance to interval length is constant (Gibbon, 1977). If the ratio is large, then the scaling of error variance means that the distribution of remembered US times, and thus the predicted response rate, should be skewed to the right, favoring shorter intervals over longer ones (Brunner, Fairhurst, Stolovitzky, & Gibbon, 1997). In this case, response rates to both types of CS should decline across the length of the CS, but should decline more sharply for the exponential CS than the uniform CS. However, this prediction depends on the amount of error, such that, for a normally distributed source of error, the ratio between the variance and mean must be greater than 2 to produce an appreciable skew in the distribution. If the error variance is smaller than this, the distribution is not skewed, and the predicted response rates should be uniform for the uniformly distributed US times,

and decline for the exponentially distributed US times, as observed in the present experiments.

Rather than assuming that the rat learns all reinforcement times to which it has been exposed and randomly remembers one on a given trial, a simple alternative could be that it only remembers the time of reinforcement on the most recent trial. That is, on trial *n*, the rat expects food to arrive at the same time that it was delivered on trial *n*-1. This represents an extreme case of a more general account in which the rat's memory for the time of food is an average calculated over many trials but strongly weighted in favor of more recent trials. In either case, response rates averaged over many trials would resemble the underlying distributions of the CS durations, as we observed here. To test this explanation, the data from Experiments 3 and 4 were subjected to an additional analysis (not shown here) in which the response record for each trial was aligned according to the time of food delivery on the preceding trial with the same CS. If the expected moment of food on any trial approximated the time that food was delivered on the preceding trial, we would expect to see a rise in response rate across the trial as the time of food on the previous trial approached. However, this was not observed. Instead, response rates were uniform for each second leading up to the time of previous food delivery.

In conclusion, the present experiments show that rats' response rates to variable CSs do not track the momentary probability of the US across the course of the CS, but rather they match the frequency distribution of US times. This finding disconfirms the prediction of real time associative models, which distribute associative strength to individual elements within the CS according to their local reinforcement rate. The observations are consistent with models that attribute responding to retrieval of a remembered time of reinforcement, since the distribution of those remembered times will reflect the history of reinforcement times. The present data are most consistent with an account in which the remembered time of reinforcement is randomly retrieved from the collection of all memories of reinforcement.

References

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

Brunner, D., Fairhurst, S., Stolovitzky, G., & Gibbon, J. (1997). Mnemonics for variability: Remembering food delay. *Journal of Experimental Psychology: Animal Behavior Processes, 23*, 68-83.

Catania, A. C., & Reynolds, G. S. (1968). A quantitative analysis of the responding maintained by interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior, 11*, 327-383.

Church, R. M., & Lacourse, D. M. (2001). Temporal memory of interfood interval distributions with the same mean and variance. *Learning & Motivation, 32*, 2-21.

Church, R. M., Lacourse, D. M., & Crystal, J. (1998). Temporal search as a function of the variability of interfood intervals. *Journal of Experimental Psychology: Animal Behavior Processes, 24*, 291-315.

Davis, M., Schlesigner, L. S., & Sorenson, C. A. (1989). Temporal specificity of fear conditioning: Effects of different conditioned stimulus-unconditioned stimulus intervals on the fear-potentiated startle effect. . *Journal of Experimental Psychology: Animal Behavior Processes, 15*, 295-310.

Desmond, J. E., & Moore, J. W. (1988). Adaptive timing in neural networks: The conditioned response. *Biological Cybernetics, 58*, 405-415.

Farwell, B. J., & Ayres, J. J. (1979). Stimulus-reinforcer and response-reinforcer relations in the control of conditioned appetitive headpoking (goal tracking) in rats. *Learning and Motivation, 10*, 295-312.

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

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

Gibbon, J., Church, R. M., Fairhurst, S., & Kacelnik, A. (1988). Scalar expectancy theory and choice between delayed rewards. *Psychological Review, 95*, 102-114.

Harzem, P., Lowe, C. P., & Priddle-Higson, P. J. (1978). Inhibition function of reinforcement: Magnitude effects on variable-interval schedules. *Journal of the Experimental Analysis of Behavior, 30*, 1-10.

Holland, P. C. (1979). Differential effects of omission contingencies on various components of Pavlovian appetitive conditioned responding in rats. *Journal of Experimental Psychology: Animal Behavior Processes, 5*(2), 178-193.

Joscelyne, A., & Kehoe, E. J. (2007). Time and stimulus specificity in extinction of the conditioned nictitating membrane response in the rabbit (*oryctolagus cuniculus*). *Behavioral Neuroscience, 121*, 50-62.

Kehoe, E. J., Horne, P. S., Macrae, M., & Horne, S. J. (1993). Real-time processing of serial stimuli in classical conditioning of the rabbit's nictitating membrane response. *Journal of Experimental Psychology: Animal Behavior Processes, 19*, 265-283.

Kehoe, E. J., & Joscelyne, A. (2005). Temporally specific extinction of conditioned responses in the rabbit (Oryctolagus cuniculus) nictitating membrane preparation. *Behavioral Neuroscience, 119*, 1011-1022.

Kirkpatrick, K., & Church, R. M. (2000). Stimulus and temporal cues in classical conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 26*, 206-219.

Kirkpatrick, K., & Church, R. M. (2003). Tracking the expected time to reinforcement in temporal conditioning procedures. *Learning & Behavior, 31*, 3-21.

Ludwig, E. A., Sutton, R. S., & Kehoe, E. J. (2008). Stimulus representation and the timing of reward-prediction errors in models of the dopamine system. *Neural Computation, 20*, 3034-3054.

Lund, C. A. (1976). Effects of variations in the temporal distribution of reinforcements on interval schedule performance. *Journal of the Experimental Analysis of Behavior, 26*, 155-164.

O'Brien, R. G., & Kaiser, M. K. (1985). MANOVA method for analyzing repeated measures designs: an extensive primer. *Psychological Bulletin, 97*, 316-333.

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

Rescorla, R. A., & 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.

Roberts, S. (1981). Isolation of an internal clock. *Journal of Experimental Psychology: Animal Behavior Processes, 7*, 242-268.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*, 1593-1599.

Schwartz, B. (1978). Stimulus-reinforcer contingencies and local behavioral contrast. *Journal of the Experimental Analysis of Behavior, 29*(297-308).

Smith, M. C. (1968). CS-US interval and US intensity in classical conditioning of the rabbit's nictitating membrane response. *Journal of Comparative and Physiological Psychology, 66*, 679-687.

Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review, 88*, 135-171. 10.1037/0033-295X.88.2.135.

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.

Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature, 412*(6842), 43-48.

Wagner, A. R., & Brandon, S. E. (2001). A componential theory of Pavlovian conditioning. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories.* (pp. 23-64). Mahwah NJ, USA: Lawrence Erlbaum Associates, Inc.

Williams, D. A., Lawson, C., Cook, R., Mather, A. A., & Johns, K. W. (2008). Timed Excitatory Conditioning Under Zero and Negative Contingencies. *Journal of Experimental Psychology: Animal Behavior Processes, 34*(1), 94-105.