



The importance of the reinforcer as a time marker

David M. Freestone*, Russell M. Church

Brown University, United States

ARTICLE INFO

Article history:

Received 1 September 2009

Received in revised form 16 January 2010

Accepted 18 January 2010

Keywords:

Pacemaker

Extinction

Random interval

Time perception

Time marker

ABSTRACT

Shifts in the psychometric function in the Free Operant Psychophysical Procedure (FOPP) have traditionally been explained by suggesting the pacemaker rate is proportional to the reinforcement rate (Behavioral Theory of Timing; BeT), or by direct associative competition between the two responses (Learning to Time; LeT). The application of these models assumes that the stimulus onset is the relevant time marker. Head-entries were reinforced in the second half of a stimulus (s1) and in the first half of a different stimulus (s2). During extinction, response times shifted earlier in s2 only, contrary to BeT. Competition between responses is an unlikely cause of the shift, contrary to LeT. We found a single-cycle correlation between the last food delivery and the time the rat stopped responding on s2. This correlation was also present in FOPP data. These results are consistent with the interpretation that the reinforcer, more than the stimulus onset, acted as the relevant time marker on this task.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

In most timing studies the reinforcer plays a central role in determining both when and how much the animal responds. For example, in the fixed interval procedure, the animal typically starts responding vigorously about two-thirds of the way into the stimulus-reinforcer interval. How vigorously the animal responds depends on the duration of the interval: the shorter the interval, the higher the response rate (Schneider, 1969). A property of fixed interval procedures is that the reinforcement rate is the reciprocal of the interval that the animal must time. Thus, it is unclear whether the response rate is due to the reinforcement rate or to the interval. And more interestingly, it is unclear whether the animal's temporal performance is even due to the interval itself.

Ironically, this property allows the Behavioral Theory of Timing (BeT) to account for fixed interval performance (Killeen and Fetterman, 1988; Bizo and White, 1997). The fundamental equation is:

$$P(t, n) = \frac{\lambda t^n e^{-\lambda t}}{n!}$$

BeT posits that a hypothetical pacemaker (λ) drives the animal through behavioral states (n). Each state is associated with the response. The reinforcement rate determines the strength of the association (and by a scale transformation, the response rate). In addition, the reinforcement rate also determines the pacemaker

rate. Put another way, the reinforcement rate determines temporal perception.

Evidence that the pacemaker rate is related to the reinforcement rate comes primarily from studies using the free operant psychophysical procedure, FOPP (Bizo and White, 1994, 1995a,b,c). In this procedure, developed by Stubbs (1980), a discriminative stimulus starts the cycle. In the first half of the stimulus, a response is reinforced on a Variable Interval (VI) schedule (an early response). In the second half of the stimulus, a different response is reinforced on a (sometimes) different VI schedule (a late response). The typical results are as follows. When the VI schedules are equal, the animal switches from the early response to the late response about halfway into the stimulus interval. When the VI schedules are not equal, the animal favors the response associated with the more reinforcing portion of the interval. This procedure is appealing because, at least on the surface, it separates reinforcement rate and the interval to be timed. That is, the animal must time the stimulus, but it must do so in the presence of reinforcers. Thus, this procedure seems ideal for testing the reinforcement rate–pacemaker rate relationship.

BeT is not, however, the only explanation of the animal's behavior on this procedure that the literature offers. More recently, Machado and colleagues have viewed this procedure as a choice procedure (Machado and Guilhaudi, 2000; Guilhaudi et al., 2007). They used the Learning to Time model (LeT). Like BeT, associations form between responses and the behavioral states in which reinforcement is received. These associations are most competitive about halfway through the interval, where the response that reinforces changes. When the rates are not equal in the two halves, the more reinforcing response accrues more associative strength around the middle of the interval, biasing the response times. In support of this, Machado and colleagues report that both rats and

* Corresponding author at: Department of Psychology, Box 1853, 89 Waterman St., Providence, RI 02912, United States. Tel.: +1 401 863 3739; fax: +1 401 863 1300.
E-mail address: David.Freestone@brown.edu (D.M. Freestone).

pigeons are sensitive to reinforcement rate changes around the middle of the interval, but not at the extreme portions of the interval.

The following experiment was conducted in an attempt to empirically dissociate BeT's explanation from LeT's. In addition, we test an alternative interpretation in which the reinforcer itself, not merely the rate at which it occurs, is an invaluable component in explaining animal behavior on the FOPP. In short, we question the assumption that this task does not confound reinforcement rate with the interval to be timed.

1.1. The reinforcer as a time marker

Our entire set of predictions is based on the assumption that reinforcers do not merely increase or decrease associative strength, nor do they adjust the pacemaker. Instead, they act as time markers that give useful information about how to perform this task. Our explanation relies on the single-cycle result that many temporal procedures (including the FOPP) produce an abrupt transition from a low to high (or a high to low) rate of responding (Guilhardi et al., 2007). We hypothesize that animals will modify their behavior based on when they were last reinforced; each reinforcer delays the changeover time on the FOPP by an approximately constant amount.

2. Materials and methods

2.1. General procedure and model predictions

We modify the procedure as follows: instead of one stimulus and two responses, our procedure uses two stimuli and one response. During one discriminative stimulus, responses in the first half of the interval are reinforced on an Random Interval (RI) schedule. Because the feeding schedule stops halfway through, we call these stop cycles. In a different discriminative stimulus, responses in the second half of the interval are reinforced on an RI schedule. These are start cycles. This procedural manipulation was made to test LeT; if a change in response times occurs in the absence of multiple response possibilities, response competition cannot be the cause. In addition, we report an extinction phase. This tests BeT. Since the reinforcement rate reduces to zero during extinction, BeT predicts a rightward shift in the response times in both start and stop cycles. Killeen et al. (1999) reported a shift in the FOPP under these conditions, but did not report the early and late responses individually (only the proportion of late responses was reported). Thus, it is unclear where the shift took place. Our hypothesis, on the other hand, predicts a leftward shift in the stop responses on stop cycles (analogous to the early responses), but no shift in the start responses on start cycles (analogous to the late responses). This latter prediction is again derived from the use of reinforcers as time markers; there are no reinforcers on start cycles until the animal has started responding. Thus, the only time marker is the stimulus onset.

2.2. Subjects

Twenty-four male Sprague Dawley rats (Taconic Laboratories, Germantown, NY) were used in this experiment. The rats were kept in a colony room on a 12:12 light–dark cycle (lights off at 8:30 am). Dim red lights provided illumination in the colony room and testing rooms. Upon arrival, the rats were 8 weeks of age and weighed between 75 and 100 g ($\mu = 89.6$, $\sigma = 7.5$). During the first week, the rats were on a free-feeding schedule. After a week, their daily food (FormuLab 5008) was rationed to maintain their weight at 85% of the free-feeding weight of the ad lib control (function obtained from Taconic Laboratories). During the experimental session, the

rats were fed 45 mg Noyes pellets (Improved Formula A) as a reinforcer. Water was available ad libitum in both the home cage and the testing chamber.

2.3. Apparatus

Twenty-four experiment chambers (Med Associates, dimensions 25 cm × 30 cm × 30 cm) were situated in two separate experiment rooms (12 in each room). Each chamber was contained in a sound-attenuating box (Med Associates, dimensions 74 cm × 38 cm × 60 cm) with a fan for ventilation. Each experimental chamber was equipped with a pellet dispenser (Med Associates, ENV-203) on the front wall that delivered the reinforcer into a food cup. A head-entry into this cup interrupted a photo beam (Med Associates, ENV-254). On the opposite wall, a water bottle protruded into the chamber allowing ad libitum access to water during the session. A lick on the spout of the water bottle completed an electric circuit. A white noise generator (Med Associates, ANL-926), and a house light (Med Associates, ENV-227M) were mounted on the back wall (near the water bottle). Four Gateway Pentium III/500 computers running Med-PC for Windows (version 1.15) controlled the experiments and recorded the data. The interruption of the photo beam, the completion of the lick circuit, and the noise and light onsets and terminations were recorded in time–event format with 2-ms accuracy.

2.4. Procedure

There were two cycle types for each rat, signaled by different 120 s stimuli (a light or white noise). The stimulus that signaled each cycle type was counterbalanced across rats. In the first cycle type (start cycles), a head-entry response was reinforced on an RI schedule in the second half of the interval (60–120 s). If food was available (primed) but not delivered by the end of the stimulus interval (120 s), the food delivery was cancelled. In the second cycle type (stop cycles), a head-entry response was reinforced on an RI schedule in the first half of the stimulus interval (0–60 s). If food was primed but not delivered by the time 60 s had elapsed, the food delivery was cancelled. If a prime occurred while the rat's head was in the food cup, the rat needed to exit and re-enter to obtain a pellet. The experiment started with a fixed 25 s timeout period in which no stimuli were presented. There was also a fixed 25 s timeout period that separated cycle presentations (the inter-trial interval). Cycles were presented randomly with equal probability. Each session had 48 total cycles, and lasted approximately 110 min. There were two phases.

2.4.1. Phase 1: training (sessions 1–30)

The rats were trained on the experiment described above for 30 daily sessions. Overfeeding after session 25 produced poor performance on session 26, and that session was not included in the analysis.

2.4.2. Phase 2: extinction (sessions 31–40)

The rats were given the same procedure as above, except responses did not produce food.

There were four groups. Each group was defined by the RI schedule in the two cycle types, written as “start cycle–stop cycle”. Group 30–30, for example, was fed on an RI30 schedule for the start cycles, and an RI30 schedule for the stop cycles. Groups 30–15 were fed on an RI30 schedule for start cycles, and an RI15 schedule for stop cycles. There were also groups 15–15 and 15–30, likewise named.

2.5. Secondary data analysis

We performed secondary data analysis on two FOPP data sets; one with pigeons (Machado and Guilhardi, 2000) and one with rats

(Guilhardi et al., 2007). The rats pressed two levers as responses and the pigeons pecked two keys. For both experiments, we analyzed the baseline phase in which all the subjects received an RI60 schedule for both early and late responses. The stimulus interval lasted 60 s.

This analysis is included in order to extend our explanation to the FOPP. The data was obtained from the Machado studies mentioned above. These data sets can be accessed on the Timing Laboratory website at <http://www.brown.edu/Research/Timelab>.

2.6. Analysis

To obtain the start and stop times, an algorithm was used which identifies change points in discrete time-series data. The algorithm finds the point that is maximally deviant from the mean of the data (Church et al., 1994; Ploberger and Kramer, 1992; Gallistel et al., 2004). In symbols, the change point, cp , is the time that maximizes the equation:

$$cp = \max_t (Tn - Nt)$$

where n is the cumulative number of responses and t is the time of the n th response. N is the total number of responses and T is the time of the last response. Since this only finds a change point over the range $0-T$, and thus misses a lot of true stops, we mimic a response at stimulus termination. In addition to finding starts and stops, we also found the time of the last food on stop cycles. In the secondary data analysis, this is defined as the time of the last food that the early lever delivered (analogous to stop cycles in our experiment).

3. Results

Fig. 1 shows the typical responding on this task. These panels show one full, random, session of responding for a single, random, rat. Each horizontal line is a head-entry response; the length of the line matches the duration of the head-entry (note that the top panels of Fig. 2 give a better representation of response rate). The solid circle is the change point on the cycle. A one-way ANOVA did not reveal a significant difference among the groups in the mean response rate in either start or stop cycles [$F(3,23) = .971$, $p = .426$, $F(3,23) = .842$, $p = .487$, respectively]. There was also no significant difference among the groups in the start or stop times in either training or extinction [all $F(3,20) < 3.03$, all $p > .05$]. This is not a surprising result because relative responses between the discrete, successive discrimination cycles are generally insensitive to the relative reinforcement rates in those cycle (see McLean and White, 1983). Because of this null result, the data was combined across groups.

During extinction, all rats significantly reduced their responding to the stimulus [$F(1,20) = 86.3$, $p < .01$]. This result is shown in the top panels of Fig. 2. While the rats maintained the same general pattern of responding, their temporal performance differed based on cycle type (bottom panels of Fig. 2). The data in the bottom panels of Fig. 2 was normalized using the equation $(y - \min(y)) / (\max(y) - \min(y))$, where y is the vector of response rates per one second bin (it is the normalization of the data in the corresponding top panels). This allows for a visual pattern comparison between training and extinction. Since extinction data is noisy and thus harder to visualize, the data in Fig. 2 has been smoothed using a moving average 5 s wide (note that this for visualization

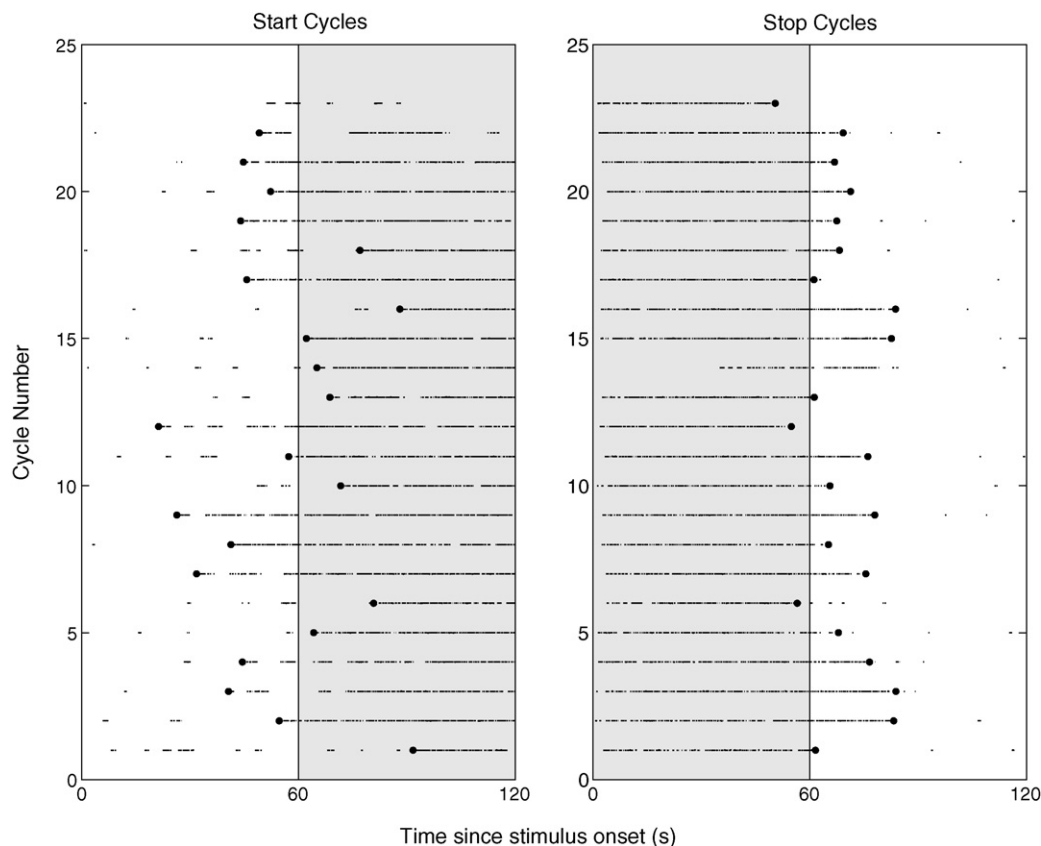


Fig. 1. Raster of response pattern during training. The two panels show a head-entry raster for start cycles (left) and stop cycles (right) for a randomly selected rat on a randomly selected session. Each response is represented as a very short horizontal line on a single cycle; the length of the line matches the duration of the response. The solid circles show the change point calculated for that cycle. The shaded regions show the interval in which the RI schedule was in effect.

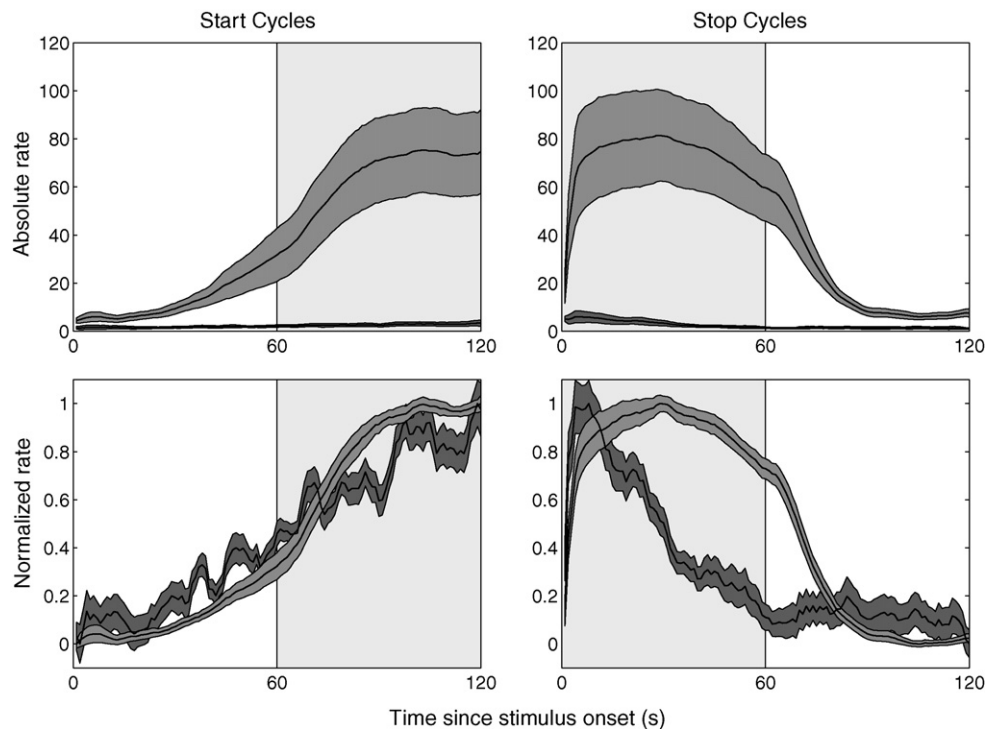


Fig. 2. Response gradients in training and extinction. The top panels show the start (left) and stop (right) absolute response rates in training (lighter shading) and extinction (darker shading). The dark solid line is the mean and the shaded region around the mean is the confidence interval. The bottom panels show the normalized response rate. The shaded regions show the interval in which the RI schedule was in effect.

only; the original, unsmoothed, data was used for all statistics). The rats stopped significantly earlier on stop cycles in extinction than in training [$F(1,20)=38.9$, $p<.01$], but did not significantly change their start time on start cycles between the two phases [$F(1,20)=2.01$, $p=.171$]. Importantly, the mean stop time during extinction did not differ from the mean stop time on cycles in training in which no food was delivered [$F(3,20)=1.18$, $p=.341$].

Our alternative explanation predicts this pattern of results. But it does so only based on the assumption that food deliveries act as time markers that tell the rat when to stop (but cannot tell it when to start). In order to test this specific hypothesis, we correlated the time of the last food with the time the rat stopped on a stop cycle (Fig. 3 shows the correlation for all data sets analyzed). A strong positive correlation was found in our experiment [lowest group $r^2=.903$, $p<.01$]. Further, this behavior was learned. Averaged across rats, the slope of the regression line follows the traditional negatively accelerated learning curve as a function of training [$F(1,28)=8.90$, $p<.01$]. The intercept linearly decreased as a function of training [$F(1,28)=5.56$, $p<.05$]. In other words, the rats stopped earlier (the intercept) and reacted more strongly to the times of the food deliveries (the slope) as training progressed. Interestingly, there was no difference in the slopes or intercepts across groups [$F(3,23)=.630$, $p=.604$, $F(3,23)=1.15$, $p=.354$, respectively].

We found the last food–stop time correlation in both rats and pigeons on the FOPP as well [$r^2=.960$, $p<.01$, $r^2=.962$, $p<.01$, respectively]. For all data sets, the slope of the best fitting line was greater than zero [smallest slope in pigeons on the FOPP, $t(16)=7.18$, $p<.01$], and less than one [largest slope in rats on the FOPP, $t(11)=-4.63$, $p<.01$]. The correlation is clearly less pronounced in the pigeon.

BeT and LeT are applied such that the animal stops around some mean value on every stop cycle, regardless of the food times. Thus they predict a last food–stop time correlation to be zero with an intercept at the mean stop time. The correct null hypothesis against which our obtained correlations should be tested is thus a slope of

0. However, since there is a region in which the data cannot lie, if the last food times and stop times were randomly placed with respect to each other, a small positive correlation would be obtained. In order for our correlations to be more than just random pairings of last food and stop times, we need to compare our correlations to the corrected null slope. To obtain the value of this slope, we did the following: for every cycle the rats experienced, we simulated (using Matlab) new food primes. The actual response times of the rat from that cycle were then used to determine the food deliveries. We obtained a last food–stop time correlation of .17. In all data sets, the correlation was significantly higher than this value [rats, $t(23)=13.77$, $p<.01$, rats on FOPP, $t(11)=20.74$, $p<.01$, pigeons on FOPP, $t(16)=5.28$, $p<.01$].

4. Discussion

BeT rests on the assumption that temporal perception is roughly proportional to the reinforcement rate. This assumption fails in our experiment. While BeT predicts a later (rightward) shift in stop times, Fig. 2 shows an earlier (leftward) shift. Killeen, Hall, and Bizo report a similar leftward shift under extinction in the FOPP (Killeen et al., 1999). They show that this is still consistent with a changing pacemaker rate provided all of BeT's parameters are free to vary. Put another way, not only the pacemaker (λ), but also BeT's state parameter (n), are affected by the changing reinforcement rate. However, according to BeT, the same general pattern of results should have been observed in the start times in our experiment. This was not the case; the rat's temporal performance did not appreciably change on the start cycles in extinction.

LeT cannot account for the data either, but for different reasons. LeT predicts that the response pattern (i.e. timing) in extinction will be the same as training; only the response rate will change. This was seen in the start cycles, but not in the stop cycles (where a temporal shift was seen). Moreover, since there was only one response (a head-entry), no competition between responses that

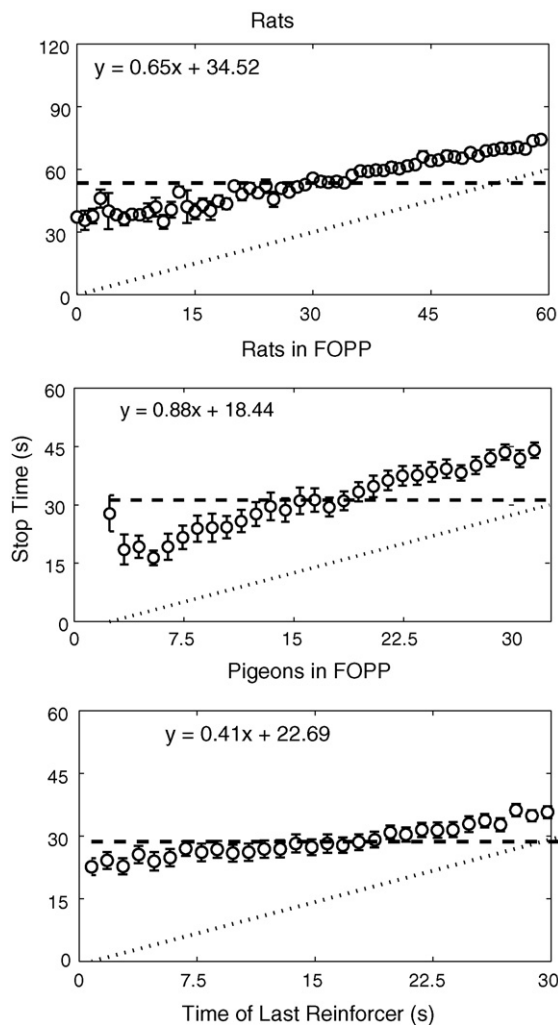


Fig. 3. Stop time as a function since the last reinforcer. Each circle is the mean (± 1 SEM) stop time for rats on this task (top panel), rats on the FOPP (middle panel), and pigeons on the FOPP (bottom panel). The dashed horizontal line is drawn through the mean of the data (BeT and LeT's prediction for the correlation). The diagonal line represents the forced correlation line; no data can lie underneath this line.

deliver a food reinforcer could have taken place that may have biased the response times. It is, however, always possible that some other behavior (i.e. exploring, grooming, drinking, etc.) competed with the head-entry response (which differentially affected stop cycles but not start cycles). To this end, we attempted to analyze drinking behavior in relation to the head-entries. However, during extinction, the rats seldom (less than 25% of the time) both drank and made head-entries in the same cycle. While this suggests that drinking did not compete with head-entries, there were too few data points to reliably test this.

It is useful to note here that it is the application of these models to the FOPP (and by extension, to our experiment) that cannot account for the observed correlation observed in Fig. 3. Any model which is applied in such a way that the stimulus onset is the only relevant time marker (and the individual food times do not matter) will fail to predict our correlation results. BeT and LeT served as examples because they were the only two theories that have been applied to the FOPP, but the same general application of the highly influential Scalar Expectancy Theory (Church et al., 1984) succumbs to the same criticism (for another discussion of SET on the FOPP see Machado et al., 2009).

One of the most interesting and striking aspects of the data is that the animals (both rats and pigeons) used a seemingly sub-

optimal strategy; the rats in our experiment, for example, stopped earlier than the halfway point on nearly 45% of the cycles. In fact, they stopped a full 15 s earlier than the halfway point about 20% of the time. In other words, on a non-trivial portion of the cycles, the animals likely lost many of the reinforcers they would have received had they kept responding. This finding is surprising because the reinforcement history is not a likely cause. The distribution of programmed reinforcer times is uniform over the half interval. So too, is the distribution of reinforcers on cycles in which only one reinforcer was programmed. Thus, the animals could not have exploited any systematic bias in the programmed reinforcement rate.

While our explanation that the animals used the reinforcer as a time maker fits much of the data we report, this explanation is incomplete. The slopes of the regression lines for the correlations shown in Fig. 3 were significantly less than one. The animals stopped relatively earlier if the last food time was later in the cycle. This may suggest simultaneous temporal processing; the animals may have been timing both the time since stimulus onset and the time since last food. Simultaneous timing is a trivial task: rats can time up to three intervals easily (Meck and Church, 1984). Moreover, rabbits have been shown to simultaneously time two intervals despite the fact that one would have sufficed (Desmond and Moore, 1991).

The major issue in this article is the role that the reinforcement times play in mediating behavior. Church and Lacourse (2001) showed that rats are sensitive to the distribution form of variable interval schedules. That is, despite the constant level of responding typically seen in variable interval procedures (Ferster and Skinner, 1957), and in the reinforcing half in the cycles in our experiment, it is likely the animals store the inter-reinforcement interval and can differentially behave based on its characteristics. In contrast, the application of both BeT and LeT to the FOPP is such that the reinforcement times play a peripheral role; they only enter in as an averaged quantity. This quantity mediates the perceived reinforcement rate and the associative strength of the states in which the reinforcement occurs. But the times of the reinforcers on a cycle-by-cycle basis do not matter. In our alternative explanation, the reinforcement times play a central role. They act as time makers that tell the animal to continue responding for another (roughly) fixed period of time.

References

- Bizo, L.A., White, K.G., 1994. Pacemaker rate in the behavioral theory of timing. *J. Exp. Psychol. Anim. Behav. Process.* 20, 308–321.
- Bizo, L.A., White, K.G., 1995a. The behavioral theory of timing: reinforcer rate determines pacemaker rate. *J. Exp. Anal. Behav.* 61, 19–33.
- Bizo, L.A., White, K.G., 1995b. Biasing the pacemaker in the behavioral theory of timing. *J. Exp. Anal. Behav.* 64, 225–235.
- Bizo, L.A., White, K.G., 1995c. Reinforcement context and pacemaker rate in the behavioural theory of timing. *Anim. Learn. Behav.* 23, 376–382.
- Bizo, L.A., White, K.G., 1997. Timing with controlled reinforcer density: implications for models of timing. *J. Exp. Psychol. Anim. Behav. Process.* 23, 44–55.
- Church, R.M., Lacourse, D.M., 2001. Temporal memory of interfood interval distributions with the same mean and variance. *Learn. Motiv.* 32, 2–21.
- Church, R.M., Meck, W.H., Gibbon, J., 1984. Scalar timing in memory. *Ann. N.Y. Acad. Sci.* 423, 52–77.
- Church, R.M., Meck, W.H., Gibbon, J., 1994. Application of scalar timing theory to individual trials. *J. Exp. Psychol. Anim. Behav. Process.* 20, 135–155.
- Desmond, J.E., Moore, J.W., 1991. Altering the synchrony of stimulus trace processes: tests of a neural-network model. *Biol. Cybern.* 65, 161–169.
- Ferster, C.B., Skinner, B.F., 1957. *Schedules of Reinforcement*. Appleton-Century-Crofts Inc., New York, NY, p. 741.
- Gallistel, C.R., Fairhurst, S., Balsam, P., 2004. The learning curve: implications of a quantitative analysis. *Proc. Natl. Acad. Sci.* 101, 13124–13131.
- Guilhardi, P., Macinnis, M., Church, R.M., Machado, A., 2007. Shifts in the psychophysical function in rats. *Behav. Process.* 75, 167–175.
- Killeen, P.R., Fetterman, J.G., 1988. A behavioral theory of timing. *Psychol. Rev.* 95, 274–295.
- Killeen, P.R., Hall, S., Bizo, L.A., 1999. A clock not wound runs down. *Behav. Process.* 45, 129–139.

- Machado, A., Guilhardi, P., 2000. Shifts in the psychometric function and their implications for models of timing. *J. Exp. Anal. Behav.* 74, 25–54.
- Machado, A., Malheiro, M.T., Erhagen, W., 2009. Learning to time: a perspective. *J. Exp. Anal. Behav.* 92, 423–458.
- McLean, A.P., White, K.G., 1983. Temporal constraint on choice: sensitivity and bias in multiple schedules. *J. Exp. Anal. Behav.* 39, 405–426.
- Meck, W.H., Church, R.M., 1984. Simultaneous temporal processing. *J. Exp. Psychol. Anim. Behav. Process.* 10, 1–29.
- Ploberger, W., Kramer, W., 1992. The CUSUM test with OLS residuals. *Econometrica* 60 (2), 271–285.
- Schneider, B.A., 1969. A two-state analysis of fixed-interval responding in the pigeon. *J. Exp. Anal. Behav.* 12, 677–687.
- Stubbs, D.A., 1980. Temporal discrimination and a free-operant procedure. *J. Exp. Anal. Behav.* 33, 167–185.