Journal of Experimental Psychology:
Animal Behavior Processes
2010, Vol. 36, No. 1, 77–91

0097-7403/10/$12.00 DOI: 10.1037/a0015971

Effect of Reward Probability on Spatial and Temporal Variation

W. David Stahlman
University of California at Los Angeles

Seth Roberts
University of California, Berkeley and Tsinghua University

Aaron P. Blaisdell
University of California at Los Angeles

Gharib, Derby, and Roberts (2001) proposed that reducing reward expectation increases variation of response form. We tested this rule in a new situation and asked if it also applied to variation of response location and timing. In 2 discrete-trial experiments, pigeons pecked colored circles for food. The circles were of 6 possible colors, each associated with a different probability of reward. Reducing reward expectation did not affect peck duration (a measure of form) but did increase horizontal variation of peck location and interpeck-interval variation. The effect of reward probability on the standard deviation of interpeck intervals was clearer (larger *t* value) than its effect on mean interpeck interval. Two datasets from rats had similar interresponse-interval effects.

Keywords: behavioral variation, reward probability, expectation

Variation and selection are the essence of instrumental learning. An animal acts in different ways (variation); actions followed by reward become more frequent than those not followed by reward (selection). We did the experiments reported here to study variation but they also shed light on selection.

Variation

The best-established generalization about variation is that it increases during extinction. This has been observed with pigeons (Eckerman & Lanson, 1969; Schwartz, 1980), rats (Antonitis, 1951; Balsam, Deich, Ohyama, & Stokes, 1998; Beck & Loh, 1990; Crow, 1978; Devenport, 1984; Frick & Miller, 1951; Mechner, 1958; Neuringer, Kornell, & Olufs, 2001), and humans (Mechner, Hyten, Field, & Madden, 1997; Morgan & Lee, 1996).

However, a training/extinction difference is hard to interpret. One problem is that training and extinction differ in many ways, both associative (expectation of reward, context conditioning) and nonassociative (density of reward, overall response rate). Another problem is that variation usually has been measured on the selection dimension—the dimension used to decide which responses are rewarded. For example, Devenport (1984) trained rats to choose the odd-numbered arms of an eight-arm radial maze. During extinction, variation of arm choice increased: the rats chose even-numbered arms more often than during training. Perhaps this was because a constraint—something pushing the rats toward odd-numbered arms, namely reward—was removed.

Constraints on behavior can have nonobvious effects (Stokes, 2005), but a very common effect is that they reduce variation. When constraints are removed, things spread out.

To make this explanation less plausible requires measuring variation on a dimension different than the selection dimension. If variation increases on a nonselection dimension during extinction, it is unlikely that this happened because a constraint was removed. For example, if birds are rewarded for pecking in one small place during training, it would be unsurprising if they pecked more widely during extinction. It would be far more interesting if other features of pecks, such as timing, varied more during extinction.

Gharib, Derby, and Roberts (2001), using the peak procedure with rats, avoided these two problems. During part of the peak-procedure trial (Phase A), there was a substantial probability of reward; during the rest of the trial (Phase B) the probability of reward was zero. Variation of bar press duration increased sharply after the change from A to B—a within-trial version of extinction increasing variation. Gharib et al. avoided the first problem because within-trial comparisons keep roughly constant several nonassociative factors that vary greatly between training and extinction. For example, recent density of reward (e.g., density of reward during the last 30 min) is much better equated by the within-trial comparison. Rescorla had made a similar point that within-subject comparisons are better than between-groups comparisons (e.g., Rescorla, 2004). Gharib et al. avoided the second problem because the duration of a detected bar press had no effect on if it was rewarded (any nonzero duration was acceptable). This made a release-from-constraint explanation unlikely.

Gharib et al. (2001) considered two associative explanations of the variation increase: (a) frustration or (b) reduced reward expectation. A test of the frustration explanation suggested it was incor-

---

W. David Stahlman and Aaron P. Blaisdell, Department of Psychology, University of California Los Angeles; Seth Roberts, Department of Psychology, University of California, Berkeley and Tsinghua University.

This research was supported by Grant MH066855 (Blaisdell). We thank Paulo Guilhardi for unpublished data, and William Baum, Nicholas Mackintosh, Allen Neuringer, and Federico Sanabria for helpful comments. This research was conducted following the relevant ethics guidelines and approved by University of California at Los Angeles' Animal Care and Use Committee.

Correspondence concerning this article should be addressed to Aaron P. Blaisdell, UCLA Department of Psychology, 1285 Franz Hall, Box 951563, Los Angeles, CA 90095-1563. E-mail: blaisdell@psych.ucla.edu

rect, so Gharib et al. proposed the alternative: "variability of the form of responses [depends on] the expected density of reward" (p. 177)—in particular, less expectation causes more variation. Cherot, Jones, and Neuringer (1996) reached a similar conclusion.

That variation depends on reward expectation makes sense. A cost of variation is that varying Response X—pressing the bar in a new way, for example—may cause the new version of X to fall outside the range of what is rewarded. The less the reward, the less the loss if this happens. A benefit of variation is that it helps find new sources of food. The more depleted the current source, the more valuable a new source. Of course, this reasoning applies to more than animal behavior. Sutton and Barto (1998) made a similar point about problem solving. Roberts and Gharib (2006) pointed out similar empirical rules in other areas.

Did the Gharib et al. (2001) rule predict correctly? Gharib, Gade, and Roberts (2004) tested it in a discrete-trials bar-pressing experiment with rats. Trials were signaled by one of two events. One signal (high probability of reward) meant a response on the current trial would always end with food; the other signal (low probability of reward) meant a response on the trial would end with food with a 0.25 probability. Trials were randomly mixed, which meant that nonassociative factors such as overall density of food, overall response rate, time since most recent food, and time since most recent response were equal for the two types of trials. The result: Bar-press durations were more variable when reward was less likely, supporting the rule.

We report two experiments similar to the Gharib et al. (2004) experiment. Pigeons were rewarded for pecking a small circle, the color of which varied across trials and indicated the probability of reward at the end of the trial. We measured three dimensions of pecks: duration, location, and what we will call *prepeck time* (PPT), the length of peck-free time before a peck. For the first peck of a trial, PPT is the time from the start of the trial; for later pecks, it is the time since the previous peck. Our main question was: Would the Gharib et al. (2001) rule correctly predict the results?

Selection

Analysis of our results unexpectedly found an effect of methodological and theoretical interest. Starting with Skinner (Ferster & Skinner, 1957), many researchers have used pigeon peck rates to draw conclusions about instrumental learning. The discovery of autoshaping (Brown & Jenkins, 1968) called these conclusions into question. If birds may peck a key simply because it has been paired with food, then changes in peck rate may be due to Pavlovian learning rather than instrumental learning.

How much does Pavlovian learning control peck rate in an ordinary instrumental learning experiment? Schwartz and Williams (1972) found that pecks generated by negative automaintenance, a steady-state form of autoshaping, were much shorter (less than 20 ms) than pecks generated by other schedules of reinforcement (more than 40 ms). This was reassuring; it suggested that instrumental responses were different from Pavlovian pecks. It was reasonable to conclude that Pavlovian learning exerts less control in ordinary schedules of reinforcement than it does in an autoshaping procedure. However, does Pavlovian learning exert no control in other situations? That was hard to conclude.

We found that time until the first peck of a trial seems to be more controlled by Pavlovian learning than interpeck intervals. This means a researcher can get a purer measure of instrumental learning by measuring only interpeck intervals. We suggest that Pavlovian learning draws an animal to a food patch; instrumental learning controls how it looks for food when it gets there.

Experiment 1

In this experiment, pigeons received food for pecking a colored circle at the center of a touchscreen. The circle was one of six colors. Pecks inside the circle—the target—ended the trial with a fixed probability that did not depend on the color of the circle. The probability of reward at the end of the trial did depend on the color: It varied from 100% to 0.6%.

An earlier, unpublished experiment from our lab had used only two colors that differed in reward probability by a factor of 8. It had not found clear differences in variation, perhaps because the difference in reward was too small. So this experiment used a wider range (a factor of 170).

The Gharib et al. (2001) rule predicted that peck duration, a measure of form, would vary more during trials with colors associated with less reward. Generalized to spatial and temporal variation, it predicted that the location and timing of pecks would vary more during those trials.

Method

Subjects. Four male racing homer and two male white Carneaux pigeons (*Columba livia*; Double T Farm, Iowa) were the subjects. They had been subjects in the preliminary experiment mentioned above. They were maintained at 80 to 85% of their free-feeding weight and were individually housed with a 12-hr light–dark cycle and free access to water and grit. The experiment was done during the light portion of the cycle.

Apparatus. The birds were tested in a flat-black Plexiglas chamber (38 cm wide × 36 cm deep × 38 cm high). Stimuli came from a color monitor (NEC MultiSync LCD1550M) visible through a 23.2 × 30.5 cm viewing window in the middle of the front panel of the chamber. The stimuli were red, blue, white, yellow, green, and magenta circles, 5.5 cm in diameter. The bottom edge of the viewing window was 13 cm above the chamber floor. Pecks to the monitor were detected by a transparent infrared touchscreen (Carroll Touch, Elotouch Systems, Fremont, CA) mounted on the front panel. A 28-V house light located in the ceiling of the chamber was on at all times. A food hopper (Coulbourn Instruments, Allentown, PA) was located in the center of the front panel, its access hole flush with the chamber floor. Experimental events were controlled and recorded by a Pentium III computer (Dell, Austin, TX). We measured the time of day of events (such as pecks) to the nearest 100 ms; peck durations were measured to the nearest millisecond. We did not measure time of day to the nearest millisecond because some of the numbers would have been exceedingly large. A videocard controlled the monitor in SVGA graphics mode (800 × 600 pixels). The resolution of the touchscreen was 710 horizontal by 644 vertical units. The horizontal units were 0.045 cm apart; the vertical units 0.035 cm apart.

Procedure. A trial consisted of the presentation of one of the six colored stimuli. Each color was equally likely. During a trial, a peck ended the trial with a probability of 20%. When the trial ended, the bird was given 3-s access to food with probability

100%, 35.3%, 12.5%, 4.4%, 1.5%, or 0.6%, depending on the color of the trial stimulus. In the preliminary experiment with two colors, we nearly balanced assignments across the birds: each of the two colors was equally often associated with each of the two probabilities. In this experiment we maintained those assignments, which made it impossible to perfectly balance the assignments in this experiment. By bird, the assignments were rwbgym (Bird 1), bgrmyw (Bird 2), rmbywg (Bird 3), bmrgwy (Bird 4), bmrgwy (Bird 5), and bwrgym (Bird 6). The string rwbgym means that red was assigned to 100%, white to 35.3%, blue to 12.5%, green to 4.4%, yellow to 1.5%, and magenta to 0.6%. Between trials was a 10-s intertrial interval. Each session lasted 65 min. There were 66 sessions, 1 per day.

Data analysis. We wanted to measure the effect of reward probability on pecks. The birds sometimes activated the touchscreen by brushing against it, rather than pecking it. Because of this, we disregarded all "pecks" on the lower quarter of the screen.

To measure peck rate we measured the time since the previous peck or, for the first peck of a trial, the time since the trial started. Because raw PPTs are highly skewed, we used a log (base *e*) transformation to make them nearly symmetric and used log PPTs in all analyses (Gharib et al., 2004; Tolkamp & Kyriazakis, 1999). Likewise, to make the distribution of peck durations more symmetric, we analyzed the reciprocals of peck durations. Tukey (1977) and Roberts (2008) explained the rationale of transformations.

A tiny fraction of PPTs, about 2 in 1,000, were 0; these were always associated with the first peck in the trial and meant that the time from the start of the trial until the first peck was less than 100 ms. This latency was too short for the associated peck to be controlled by the color of the stimulus. The start of the trial had just happened to coincide with a random peck of the screen. These pecks were disregarded.

PPTs were measured to the nearest 100 ms, which created many ties. This created problems when computing percentiles. For example, suppose you want to compare the medians of two conditions, each with 101 observations. In one condition, there is 1 observation at 100 ms, 50 at 200 ms, and 50 at 300 ms. In the other, there are 50 observations at 100 ms, 50 at 200 ms, and 1 at 300 ms. The usual computation of the median would give 200 ms for both conditions. A better computation would give medians that differ by about 100 ms. To solve this problem, when computing percentiles a small random duration, with a uniform distribution between 0 and 100 ms, was added to each PPT. This is a standard way to deal with ties.

To get a summary value for each measure, such as mean peck duration or standard deviation of peck duration, we computed the measure separately for each bird/day combination and then took a weighted mean over all days (the value for each day weighted by the number of pecks that day) to get a measure for each bird. Weighted means were used because the number of pecks per session could vary greatly.

Measuring the effect of probability of reward had two steps. Step 1: Fit a straight line to the cumulated-over-days results for each bird (*y*) as a function of probability of reward (*x*) assuming that the probabilities of reward are equally spaced (which on a log scale they are). The slope of the line measures the effect. This provides one number (the slope) per bird. Step 2: To decide if the slope is nonzero—that is, probability of reward had an effect—conduct a *t* test ($n = 6$ birds) using the six slopes (one/bird). We used such a test because it gave equal weight to each level of probability of reward and because it is especially sensitive to the monotonic changes predicted by the idea we wanted to test (reducing probability of reward increases variation).

Unless stated otherwise *p* values are one-tailed. We used one-tailed *p* values when testing predictions; we used two-tailed *p* values when assessing the strength of a pattern that we did not predict.

Results and Discussion

There were about 80,000 pecks per bird (range 31,000 to 121,000) during the experiment. Each bird received about 10,000 trials (range 3,000 to 11,000).

Form variation. Probability of reward had no detectable effect on peck duration. Figure 1 shows the mean (top panel) and standard deviation (bottom panel) of peck duration as a function of probability of reward. The two numbers in the upper right of each graph (e.g., 0.6, 0.4) are one-tailed *p* values from a test of the slopes. The first is for the hypothesis that the slope is positive (the relationship predicted for the standard deviations), the second for the hypothesis that the slope is negative. Neither the mean nor the standard deviation reliably increased with probability of reward, $ts < 1$.

Spatial variation. Variation of peck location—horizontal and vertical—did depend on reward probability (see Figure 2). We measured the location of pecks using Cartesian (*xy*) coordinates; the origin was the lower left-hand portion of the screen. The upper graphs show means of the horizontal and vertical values; the lower graphs, standard deviations.

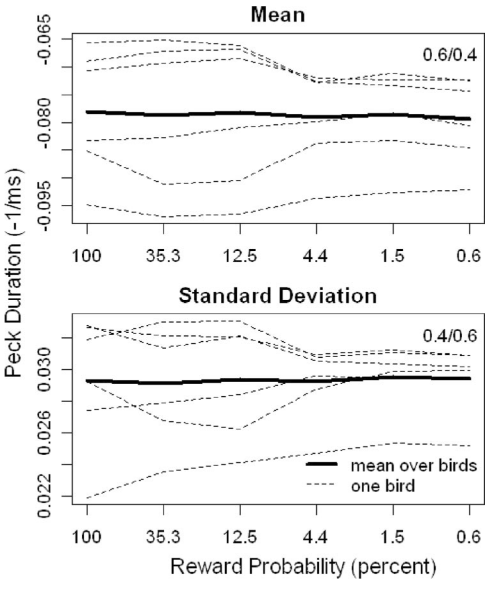

Figure 1. Experiment 1: Peck duration as a function of probability of reward. Means (upper panel) and standard deviations (lower panel) were computed after a negative reciprocal transformation. Each point is a mean across all sessions of the within-session means or the within-session standard deviations. Each dotted line is from a different bird. The solid line shows the mean over birds. The two numbers in the upper right of each graph are *p* values from one-tailed *t* tests of the hypothesis that the values are increasing (first number) or decreasing (second number).

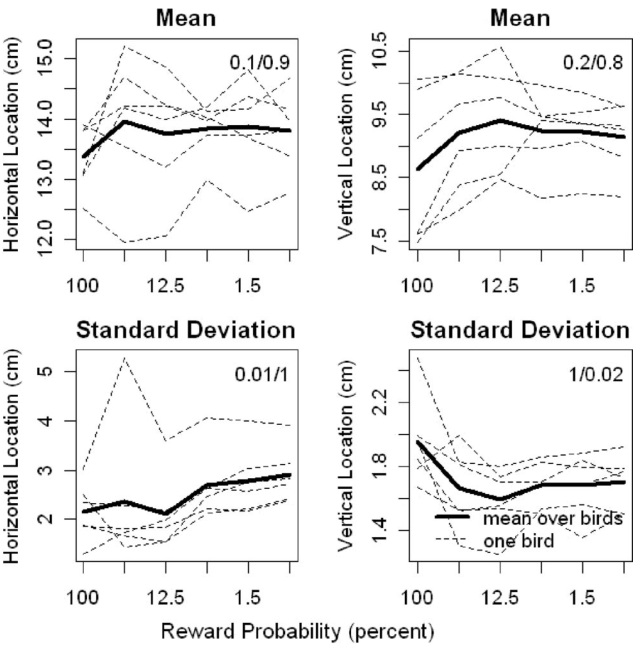

Figure 2. Experiment 1: Peck location as a function of probability of reward. The left panels show horizontal position, the right panels, vertical position, both measured from the lower left-hand corner of the touchscreen. Each dotted line is from a different bird. The solid line shows the mean over birds. The two numbers in the upper right of each graph are *p* values from one-tailed *t* tests of the hypothesis that the values are increasing (first number) or decreasing (second number).

Horizontal variation (left graphs) showed the predicted effect. The standard deviation of horizontal location (bottom left of Figure 2) increased as probability of reward decreased, *t*(5) = 3.14, *p* = .01. There was no detectable change from 100% to 12.5%, *t*(5) = −0.15, *p* = .56; a very clear increase from 12.5% to 4.4%, *t*(5) = 7.28, *p* = .0004; and a weaker change from 4.4% to 0.6%, *t*(5) = 2.24, *p* = .04.

Vertical location changed differently. Its standard deviation (bottom right of Figure 2) decreased as reward probability decreased, *t*(5) = 2.96, two-tailed *p* = .03. The decrease was nearly significant from 100% to 35.3%, *t*(5) = 2.25, two-tailed *p* = .08. There was no detectable change from 35.3% to 0.6%, *t*(5) = 0.61, two-tailed *p* = .60.

As we mentioned in the introduction, reducing probability of reward might increase variation due to release from constraint. In this experiment, peck location was constrained by reward—only pecks within the target were rewarded. If release from constraint is why horizontal variation increased as reward probability decreased (lower left graph of Figure 2), then we would expect to see fewer on-target pecks as probability of reward decreased. This did not happen. Figure 3 shows the fraction of on-target pecks as a function of reward probability. The fraction of on-target pecks did not vary systematically as probability of reward decreased.

Another way to ask if the variation increase was due to release from constraint is to look at off-target pecks. Their location was not constrained by reward—they were not rewarded, no matter where they occurred. When the analysis is restricted to these pecks, the horizontal variation effect is even clearer than in Figure 2. As reward probability decreased, horizontal variation of off-target pecks increased, *t*(5) = 3.45, *p* = .009.

The vertical variation effect in Figure 2 (lower right graph) was probably due to the pigeons' tendency to peck at food sources (Hearst & Jenkins, 1974). The food cup was located directly below the screen. The vertical mean (upper right) graph in Figure 2 shows where the pecks were centered vertically. It had the same shape as the standard-deviation function (lower right graph in Figure 2): big change from 100% to 35.3%, little change from 35.3% to 0.6%. The similarity of shape suggests a common cause. A decrease in vertical location meant that the birds pecked closer to the food cup and further from the target. With low probabilities of reward, the birds pecked mostly at the level of the target. With a high (100%) probability of reward, their pecks shifted downward toward the food cup, apparently anticipating food presentation. Variation in the amount of anticipation raised the standard deviation at 100% reward probability. Apparently the location of the food cup exerted a downward pull that was substantial at 100% probability of reward and negligible at lower probabilities.

Temporal variation. Reward probability had a strong effect on PPTs. For the first peck of a trial, the PPT is the time to first peck, the latency of the first peck measured from the start of the trial; for later pecks, it is the interpeck interval, the time since the previous peck. Figure 4 shows how both were affected by reward probability. For time to first peck, means showed a big effect, *t*(5) = 4.37, *p* = .003 (upper left); standard deviations showed no effect, *t*(5) = −0.08, *p* = .53 (lower left). In contrast, for interpeck intervals, the effect on standard deviations (lower right) was remarkably strong, *t*(5) = 6.65, *p* = .0006; even stronger than the effect on means (upper right), *t*(5) = 1.77, *p* = .07. In other words, reward probability had a very clear effect on temporal variation (*p* = .0006) after the first peck and no detectable effect (*p* = .53) before the first peck.

The sizes of the effects (measured by slopes) showed a crossover interaction. With means, the time-to-first-peck slope was more than the interpeck-interval slope, *t*(5) = 3.24, two-tailed *p* = .02. With standard deviations, the time-to-first-peck slope was less than the interpeck-interval slope, *t*(5) = 3.29, two-tailed *p* = .02.

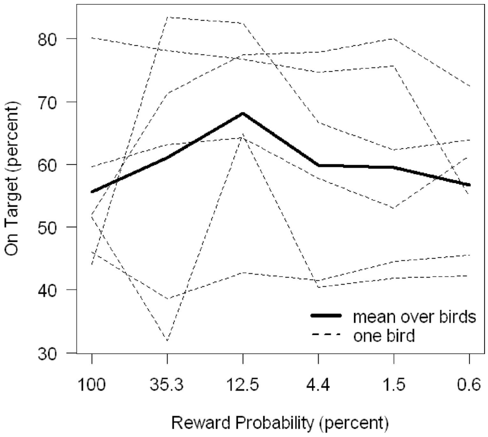

Figure 3. Experiment 1: Fraction of pecks on the target as a function of probability of reward. A percentage of 60% means that 60% of pecks were within the colored circle and 40% were outside the circle. Each dotted line is from a different bird. The solid line shows the mean over birds.

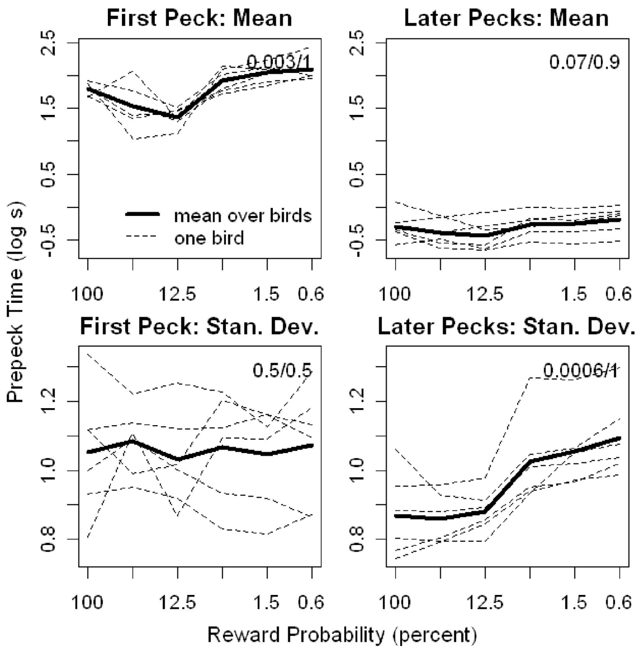

Figure 4. Experiment 1: Prepeck times as a function of probability of reward. First pecks are the first pecks during a trial; later pecks are all later pecks during that trial. The two numbers in the upper right of each graph are *p* values from one-tailed *t* tests of the hypothesis that the values are increasing (first number) or decreasing (second number). Each dotted line is from a different bird. The solid line shows the mean over birds.

Apparently reward probability acted differently before and after the first peck of a trial.

The law of effect, taken as an empirical statement, predicts that the mean slopes will be positive (longer PPTs with lower reward probabilities). Although the time-to-first-peck mean slope (upper left) is positive overall, it is negative between 100% and 12.5%—PPTs were shorter with 12.5% than with 100%, $t(5) = 5.04$, two-tailed $p = .003$. Interpeck intervals did not show this difference, $t(5) = 1.55$, two-tailed $p = .18$. An interaction between probability of reward (100% vs. 12.5%) and peck (first vs. later), $F(1, 5) = 17.94$, $p = .008$; indicates that the anomaly was smaller with later pecks.

To better understand these effects, we looked at PPT distributions. Figure 5 shows percentiles of the distributions. Its four panels separate PPTs by order of the peck in the trial: first (upper left), second (upper right), third (lower left), and fourth and later pecks (lower right). Each line shows the value of a percentile (from 1st percentile to 99th percentile) averaged over the six birds. Figure 5 shows two patterns of change. For second and later pecks, the effect of reward probability was remarkably localized, in two ways. First, it happened only between reward probabilities of 12.5 and 4.4%. Before (from 100% to 12.5%) and after that (from 4.4% to 0.6%) there was little change. Second, it happened with the 95th percentile and 99th percentile, but not with the 75th percentile and lower percentiles. First pecks show a different effect, localized in the first way but not the second. With first pecks, the 25th percentile and 50th percentile, in addition to the higher percentiles, appear to be influenced by reward probability.

Figures 6 (times to first peck) and 7 (interpeck intervals) show whole distributions of PPTs. In each graph, one distribution includes results from the higher probabilities of reward (100%, 34.5%, and 12.5%), the other, from the lower probabilities (4.4%, 1.5%, and 0.6%). In Figure 7 (interpeck intervals), the difference between the two distributions is, as Figure 5 suggests, very localized. For all six birds, reducing reward probability substantially increased the chances of long interpeck intervals but left the rest of the distribution virtually unchanged. In Figure 6 (times to first peck), the pattern of difference was different. The two distributions differ in their main bodies, not just in their tails. In Figure 6 the leftmost portions of the two distributions overlap (as in Figure 7) but this is because they derive from pecks so early in the trial that the color of the stimulus had little control over behavior.

To quantify the difference between Figures 6 and 7, we compared medians of the two distributions (high and low reward). For Figure 6 (times to first peck), the two medians were reliably different, $t(5) = 2.90$, two-tailed $p = .03$. For Figure 7 (interpeck intervals), the two medians were not reliably different, $t(5) = 0.24$, two-tailed $p = .81$. The Figure 6 differences were more than the Figure 7 differences, $t(5) = 3.37$, two-tailed $p = .02$.

Experiment 2

In Experiment 1 the birds often pecked near the edge of the target. This may reflect a tendency to peck at areas of high contrast. We wanted to increase target peck rate so—assuming this explanation was correct—we added a small black circle to the middle of the target. This increased contrast inside the target.

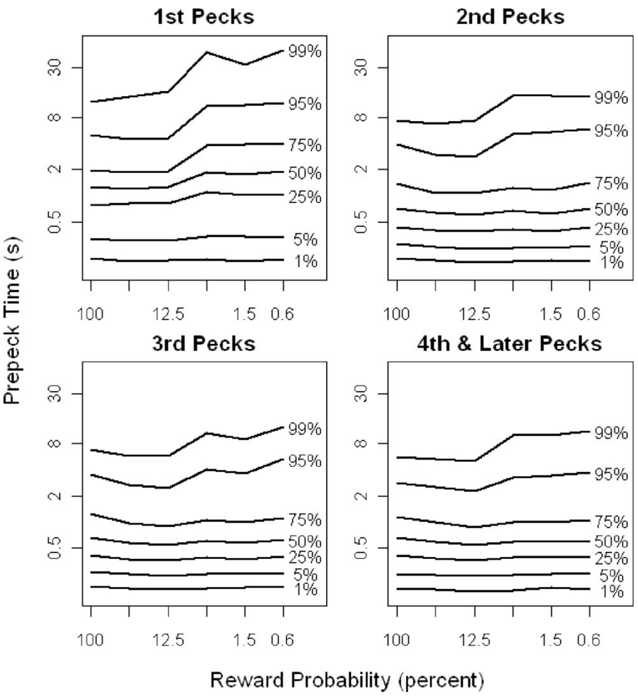

Figure 5. Experiment 1: Distribution of first and later prepeck times as a function of probability of reward. Each line is a mean over the six birds.

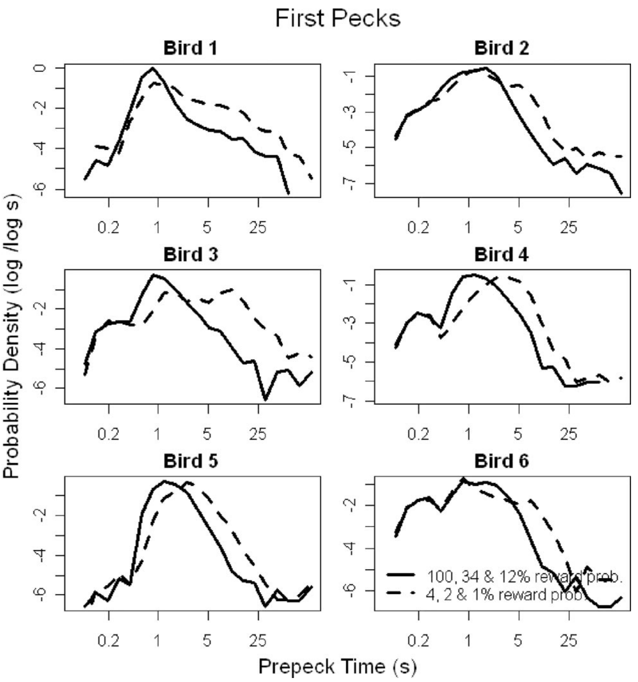

Figure 6. Experiment 1: Distribution of times to first peck as a function of probability of reward. Solid lines show data from the three signals with the higher probabilities of reward; broken lines show data from the other three signals.

Method

Subjects and apparatus. The birds used in Experiment 1 were the subjects. They were maintained as previously stated. The apparatus of Experiment 1 was used except for the addition of a 2-mm black dot to the center of each stimulus.

Procedure and data analysis. The procedure and data analysis were the same as Experiment 1. The experiment lasted 26 days, one 65-min session per day.

Results and Discussion

We measured about 32,000 pecks per bird (range 25,000 to 50,000) during about 8,000 trials per bird (range 6,000 to 9,000).

Form variation. Peck duration was unaffected by reward probability. There was no reliable change in mean and the standard deviation with reward probability, ts < 1.

Location variation. Variation of the horizontal location of pecks increased when reward probability decreased (see Figure 8), $t(5) = 2.23$, $p = .04$. Again, there was the flat/rise/flat pattern as seen in Figure 5. There was no reliable change from 100% to 12.5%, $t(5) = 0.32$, $p = .38$; a reliable increase from 12.5% to 4.4%, $t(5) = 2.39$, $p = .03$; and no reliable change from 4.4% to 0.6%, $t(5) = 0.17$, $p = .43$.

Variation of the vertical location of pecks decreased when reward probability decreased (see Figure 8), $t(5) = 3.56$, $p = .008$. There was a reliable decrease from 100% to 12.5%, $t(5) = 3.17$, $p = .01$; and no reliable change from 12.5% to 0.6%, $t(5) = 1.46$, $p = .11$.

Was the increase in horizontal variation due to release from constraint? The fraction of on-target pecks increased from 100% reward (49% on-target) to 0.6% reward (61% on-target)—opposite to what the release-from-constraint hypothesis predicts. When only pecks outside the target are considered, the horizontal variation increase also occurred with this subset of pecks, $t(5) = 3.55$, $p = .008$.

Temporal variation. Again, Experiment 1's important findings were repeated. Figure 9 shows three important results: First, crossover interaction in terms of p values. With times to first peck, means ($p = .02$) were more sensitive to the effect of reward probability than standard deviations ($p = .4$); with interpeck intervals, standard deviations ($p = .004$) were more sensitive than means ($p = .07$). Second, crossover interaction in terms of slopes: With means, time-to-first-peck slopes were greater than interpeck interval slopes, $t(5) = 2.31$, $p = .03$; with standard deviations, the reverse was true, $t(5) = 3.47$, $p = .01$. Third, partial failure of the law of effect: Times to first peck were shorter at 12.5% than at 100%, $t(5) = 5.01$, two-tailed $p = .004$. For interpeck intervals, there was no reliable difference, $t(5) = 0.47$, two-tailed $p = .66$. There was a reliable interaction between reward probability (100% or 12.5%) and peck (first or later), $F(1, 5) = 25.32$, two-tailed $p = .004$.

Figures 10 and 11 show distributions of times to first peck and interpeck intervals. Again, probability of reward affected these measures differently. For times to first peck, the median was greater with low probabilities of reward, $t(5) = 2.81$, two-tailed $p = .04$. For interpeck intervals, there was no difference, $t(5) = 0.61$, two-tailed $p = .57$. The time-to-first-peck difference was

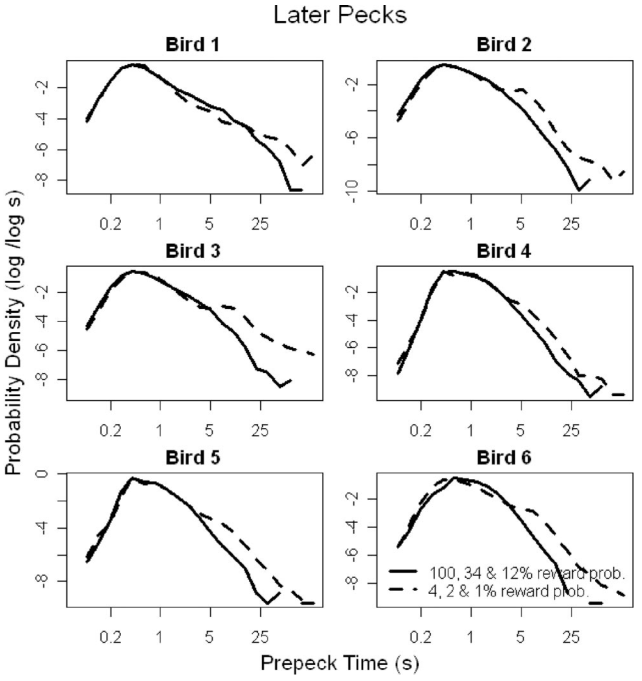

Figure 7. Experiment 1: Distribution of interpeck intervals as a function of probability of reward. Solid lines show data from the three signals with the higher probabilities of reward; broken lines show data from the other three signals.

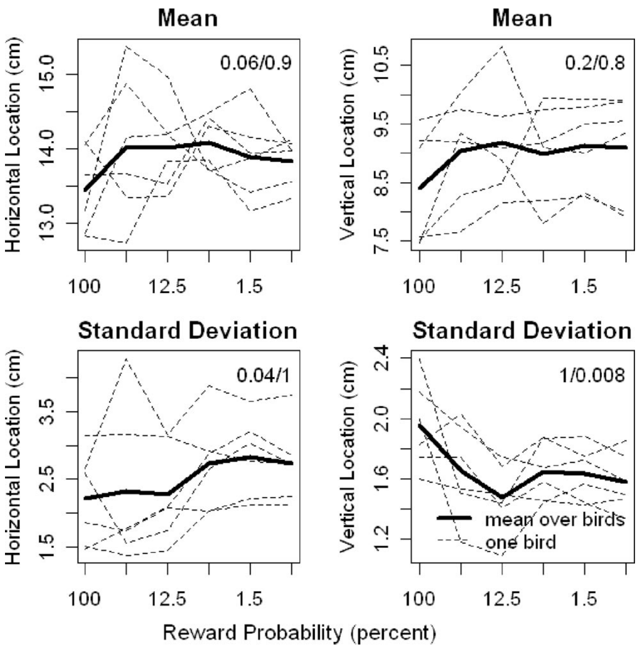

Figure 8. Experiment 2: Peck location as a function of probability of reward. The left panels show horizontal position, the right panels, vertical position, both measured from the lower left-hand corner of the touchscreen. Each dotted line is from a different bird. The solid line shows the mean over birds. The two numbers in the upper right of each graph are *p* values from one-tailed *t* tests of the hypothesis that the values are increasing (first number) or decreasing (second number).

reliably more than the interpeck-interval difference, $t(5) = 2.74$, two-tailed $p = .04$.

General Discussion

Summary of Results

We wanted to test the idea that reducing reward expectation increases variation. Relevant to that prediction:

1. Peck duration was unaffected by reward probability (Figure 1).
2. The horizontal location of pecks changed as predicted (Figures 2 and 8). Two observations argued against a release-from-constraint explanation: (a) The fraction of off-target pecks did not increase as reward probability decreased (Figure 3). (b) The same result was found when only off-target pecks were considered.
3. The vertical location of pecks changed opposite to prediction (Figures 2 and 8).
4. Variation of time from the start of the trial until the first peck was unaffected by reward probability (Figures 4 and 9).
5. Variation of interpeck intervals changed as predicted (Figures 4, 7, 9, and 11).

Both types of PPT—time to first peck and interpeck intervals—obeyed the law of effect in the sense that they increased as the probability of reward decreased. However, the details differed:

6. With times to first peck, the increase—measured by *p* values or by slopes—was clearer in the means than the standard deviations; with interpeck intervals, the reverse was true (Figures 4 and 9).
7. With times to first peck, reducing reward probability shifted the whole distribution (Figures 6 and 9); with interpeck intervals, reducing reward probability only made the tails heavier (Figures 7 and 11).
8. Times to first peck showed an anomaly: shorter latency (i.e., higher peck rate) with 12.5% reward probability than with 100%. Interpeck intervals did not show this anomaly (Figures 4 and 9).

What Controls Variation?

Gharib et al. (2001) proposed that reward expectation controls variation—the less expectation, the more variation. In a well-controlled experiment, Gharib et al. (2004) found results that supported the proposal. Our results both (a) increase the generality of the supporting evidence and (b) find cases where the proposal is wrong.

The previous well-controlled evidence came entirely from rat bar-press durations. Our results increase the generality of the

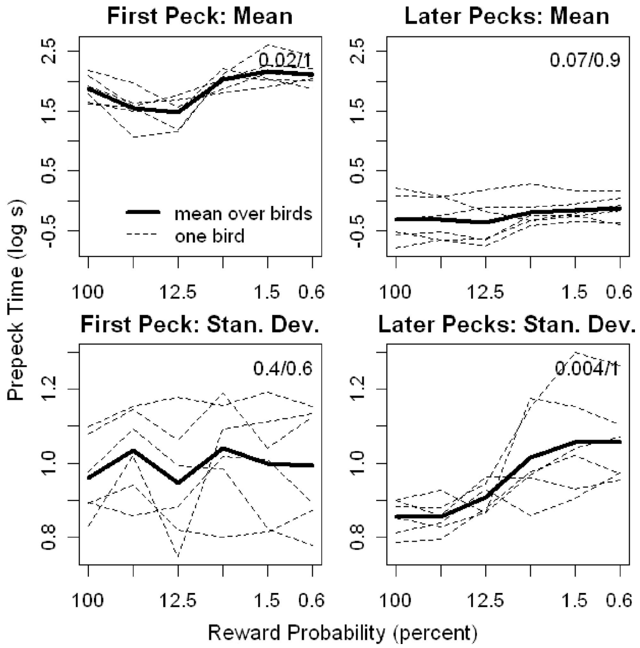

Figure 9. Experiment 2: Prepeck times as a function of probability of reward. First pecks are the first pecks during a trial; later pecks are all other pecks during that trial. The two numbers in the upper right of each graph are *p* values from one-tailed *t* tests of the hypothesis that the values are increasing (first number) or decreasing (second number). Each dotted line is from a different bird. The solid line shows the mean over birds.

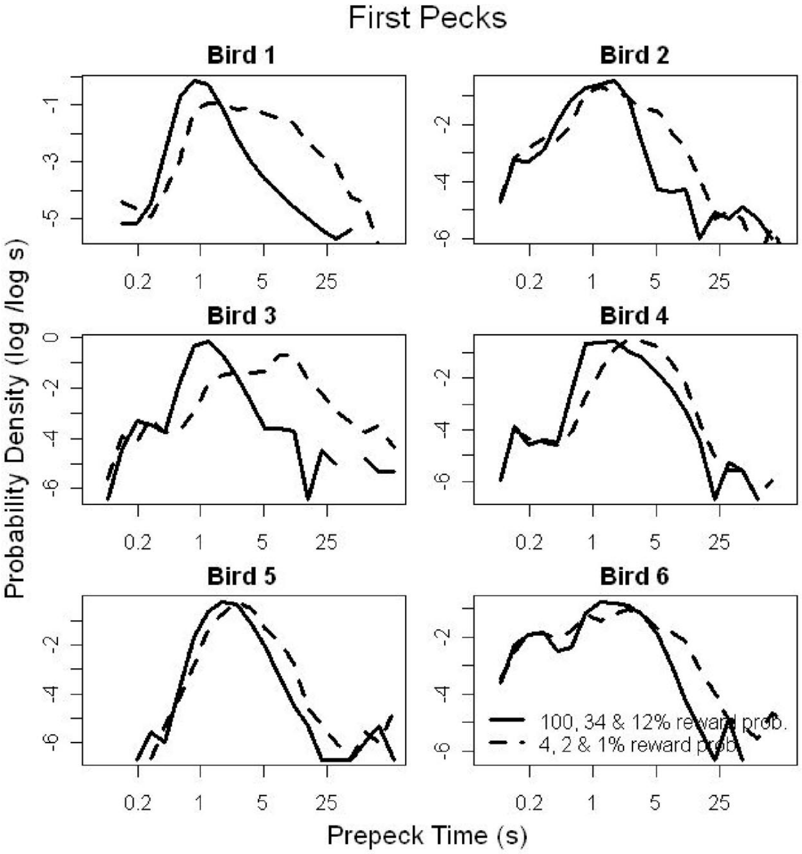

Figure 10. Experiment 2: Distribution of times to first peck as a function of probability of reward. Solid lines show data from the three signals with the higher probabilities of reward; broken lines show data from the three signals with lower probabilities of reward. Each point is a mean over six birds.

well-controlled evidence in three ways: (a) new taxonomic class (birds); (b) new response (pecks)—pecks are consummatory; bar presses are not—; and (c) new response dimensions (horizontal location and interpeck interval).

In three ways the Gharib et al. (2001) rule did not correctly predict the results: First, peck duration was unaffected by reward probability. Like us, Holt, Green, and Muenks (2004) found that peck duration was unchanged by a factor that changed rate. On the other hand, a few studies (Jenkins, 1981; Ziriax & Silberberg, 1978) have found peck duration sensitive to experimental treatments. The lack of change we saw is not surprising. The rationale behind the Gharib et al. rule is that variation will help animals uncover new sources of food. Varying the form of pecks would not do that. Whether food is scarce or plentiful, the best way to eat it remains the same. Second, vertical location: this measure shifted the wrong way. Its variation decreased when reward probability decreased. As mentioned earlier, this is surely because the food hopper was located directly beneath the screen. Pecks were drawn toward the food hopper on high-probability trials. Breland and Breland (1961) noticed that the form of actions trained with food drifted toward the unlearned responses made to consume the food. We observed location drift. Not only average location but variation changed the "wrong" way because the greater the amount of drift, the more variation in that amount. Third, time from start of trial until the first peck: Variation in this measure was unaffected by reward probability. Between trials, we assume birds wandered away from the screen because there was nothing to be gained from pecking it. It was no longer a food patch. Times to first peck reflect time to re-approach the screen. Hearst and Jenkins (1974) showed how Pavlovian learning controls approach behavior in a wide range of settings. In general, Pavlovian learning occurs in situations where it is clear what to do (e.g., salivate when eating, approach a food source); instrumental learning occurs in situations where it isn't clear what to do. An animal should approach a poor food source less than a good food source, but to vary how it approaches a food source depending on the quality of the source makes no sense (it should always go in a straight line). Treating latency to first peck as a measure of how quickly a bird approaches a food patch makes sense of our failure to find a variation effect with that measure.

In summary, our results suggest that reward expectancy controls variation of the middle of the foraging process. It does not control variation in approaching food patches. It does control variation in how a patch or food source is searched or exploited. It does not control variation of how food is eaten.

A regularity in behavior—such as the Gharib et al. (2001) rule—might be due to evolution or physics (e.g., tires become thinner with use; this is due to physics, not tire design.). Instru-

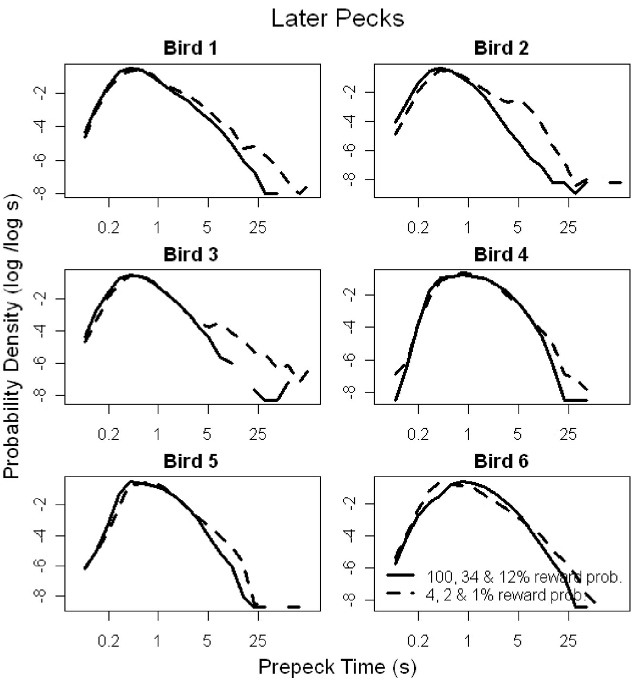

Figure 11. Experiment 2: Distribution of interpeck intervals as a function of probability of reward. Solid lines show data from the three signals with the higher probabilities of reward; broken lines show data from the three signals with lower probabilities of reward. Each point is a mean over six birds.

mental learning is surely due to evolution because of the complexity of the machinery. Variation, however, is everywhere. Systematic changes in variation occur in countless situations not involving neural mechanisms (e.g., thermodynamics). So it is conceivable that the Gharib et al. rule is a lucky accident. If so, then exceptions to the rule should not make evolutionary sense. However, as explained above, they do—which suggests that the Gharib et al. rule is true for adaptive reasons. Just as many biological variables (e.g., body temperature) are regulated, so is variation in several dimensions of behavior. It is pushed up and down in beneficial ways by natural selection.

What Causes the Law of Effect?

Both types of PPT—time to first peck and interpeck interval—followed the law of effect, in the sense that they increased when probability of reward decreased (that is, the slope of the fitted line was positive). However, how they increased—how the distributions changed (Figures 6, 7, 10, and 11) and the nonmonotonic change in one but not the other (upper panels of Figures 4 and 9)—imply different underlying mechanisms.

Let’s start with interpeck intervals. Their distributions (Figures 7 and 11) suggest a mixture model. The model is that times come from one of two distributions, short and long. The short distribution is what is observed on trials with a high probability of reward. It has a median of about 1 s. The long distribution is what is “added” to the short distribution on trials with a low probability of reward. Its shape is revealed by the “bump” in the low-reward distribution at long interpeck interval. It has a median of about 10 s. With high-reward probability, the overall distribution of PPTs is 100% short/0% long; with low-reward probability, the distribution is about 98% short/2% long (the mixing probabilities are just rough estimates). Presumably the two distributions correspond to two possible internal states. Many other studies have found similar results (Shull, Gaynor, & Grimes, 2001); the two states have been called “engagement” and “disengagement” (Bowers, Hill, & Palya, 2008, p. 345).

Figures 7 and 11 show that the distribution of short interpeck intervals—what the bird does when engaged—was not affected by probability of reward. Blough (1963) was the first to notice the insensitivity of short interpeck intervals to many treatments. Pecks after short interpeck intervals should probably be considered part of the act of feeding. Short interpeck intervals are insensitive to probability of food for the same reason peck duration is insensitive: The best way to eat does not depend on the scarcity of food.

The interesting feature of the mixture-model explanation is its discreteness: The bird is in one state or the other (engaged or not), nothing in between. The states are quite different because the two distributions are quite different. The notion of distinct states is part of behavior systems theory (e.g., Timberlake, 1983; Timberlake & Lucas, 1989), which assumes that animals shift between various states or modes, such as “focal search mode,” “general search mode,” and “quiescence” (Timberlake & Lucas, 1989, p. 262). Strength of food anticipation controls what mode an animal is in. With no anticipation of food, the animal is in quiescence; weak anticipation of food moves it to general search mode; strong anticipation moves it to focal search mode. An animal can be in only one mode at a time and there are no blends of modes. These ideas were inspired partly by Staddon and Simmelhag’s (1971) results, in which pigeons showed predictable patterns of interfood behavior when response-independent food was given at regular intervals. When food was distant in time birds did *X*; when food was imminent they did *Y*. They could not do *X* and *Y* at the same time nor could they do something between *X* and *Y*.

The time-to-first-peck results suggest a different mechanism. The simplicity that needs explaining is that as reward probability decreased, the median of the distribution increased but its spread remained the same (lower left panels of Figures 4 and 9). This suggests that the same mechanism controlled behavior in both cases. The change in median can be explained as due to a change in a parameter of the mechanism—the attractiveness of pecking or associative strength.

The notion that the pecks in instrumental experiments may come from two different sources became popular after the discovery of autoshaping (Brown & Jenkins, 1968), as mentioned in the introduction. Schwartz and Williams (1972) found that pecks generated by a negative automaintenance procedure, in which pecks were never followed by food, were quite short (20 ms or less), whereas pecks generated by common instrumental procedures, such as fixed-interval reinforcement, were longer. After many sessions of a conventional procedure, the duration difference was large; there was almost no overlap between the peck durations generated by negative automaintenance (20 ms or less) and the peck durations generated by ordinary procedures (more than 40 ms). The results implied that Pavlovian learning (keylight-food pairing) generated short-duration pecks whereas instrumental learning (pecks followed by food) generated long-duration pecks. Schwartz (1977) provided more evidence for this conclusion.

Does our distinction (first vs. later pecks) relate to the Schwartz and Williams (1972) distinction (short Pavlovian pecks vs. long instrumental pecks)? First pecks do have a Pavlovian aspect. Autoshaping and other sign-tracking results imply that Pavlovian learning causes the bird to approach the keylight (Hearst & Jenkins, 1974). As discussed above, the first peck of a trial involves approach more than later pecks, so time to the first peck of a trial may reflect Pavlovian learning more than later PPTs (interpeck intervals). That time to first peck is a relatively Pavlovian measure is supported by an anomaly we noticed: Time to first peck was less with 12.5% probability of reward than with 100% probability of reward. A series of Pavlovian autoshaping studies in pigeons by Collins and Pearce (1985) bears on this anomaly. They first pointed out that keypecking by pigeons was both a conditioned response (CR) and an orienting response (OR), and that ORs reflected the level of attention paid to the predictive stimulus by the subject. They found in their studies that partial reinforcement resulted in stronger responding to an initial element of a serial compound CS than did continuous reinforcement. Their results are consistent with the Pearce–Hall theory of Pavlovian learning (Pearce & Hall, 1980) in showing that partial reinforcement results in a weaker CR but a stronger orienting response (OR) than continuous reinforcement. Because both CRs and ORs consist of keypecking in an autoshaping procedure, an increase in the OR translates to an increased rate of keypecking. This application of the Pearce–Hall theory readily explains our results as well: Unpredictable reward maintains attention to the signal, and greater attention leads to a greater OR which, like the CR, are screen pecks. With 12.5% probability of reward, reward is less predictable than with 100% probability of reward. So far as we know, no theory of instrumental learning easily explains this anomaly.

If Pavlovian learning draws the bird to the screen to make the first peck of a trial, then perhaps first pecks are more Pavlovian in nature in other ways than later pecks. If so, the research (Schwartz, 1977; Schwartz & Williams, 1972) suggests that first pecks will be shorter than later pecks. We tested this prediction. Figure 12 shows the mean durations of first and later pecks. In both experiments, the means differed reliably in the predicted direction.

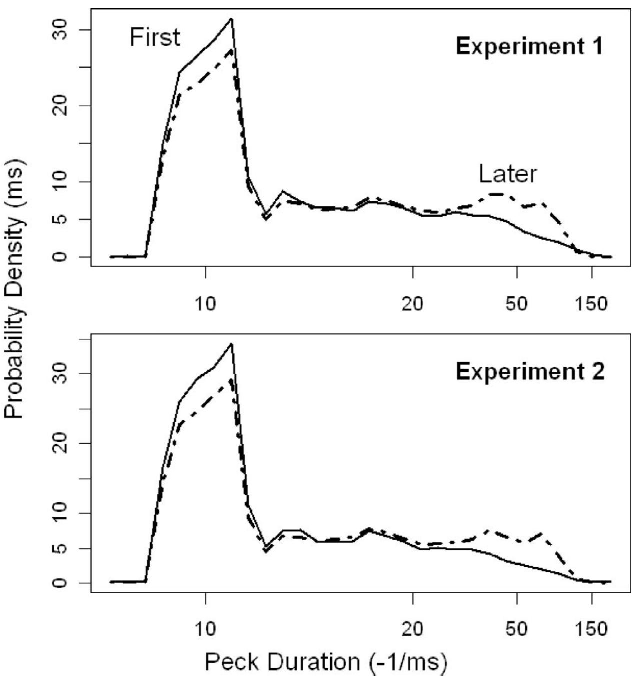

Figure 13. Distributions of peck duration as a function of peck order. First pecks are the first pecks during a trial; later pecks are all pecks after the first peck during the trial. Each point is a mean over six birds.

Figure 13 shows distributions of peck duration. The distributions have a sharp early peak and a wide later section, which suggests more than one mechanism. A reasonable conclusion

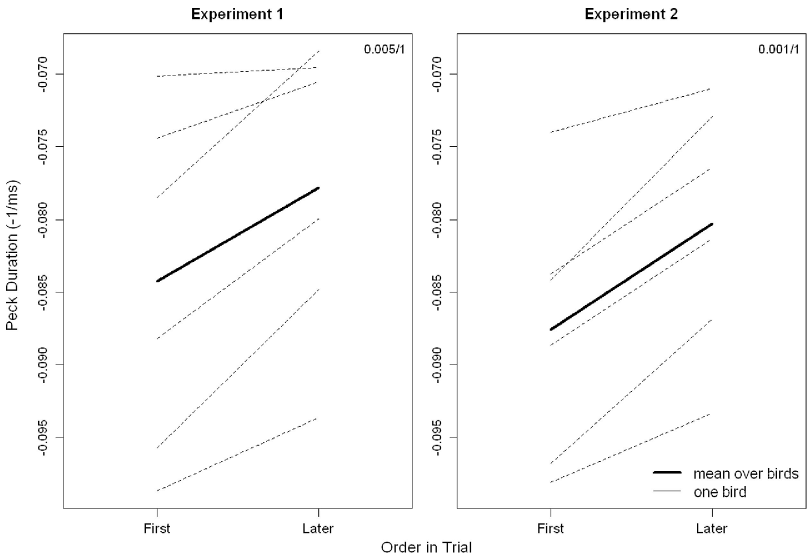

Figure 12. Peck duration as a function of peck order. First pecks are the first pecks during a trial; later pecks are all later pecks during that trial. Each dotted line is from a different bird. The solid lines show the mean over birds. The two numbers in the upper right of each graph are p values from one-tailed t tests of the hypothesis that the values are increasing (first number) or decreasing (second number).

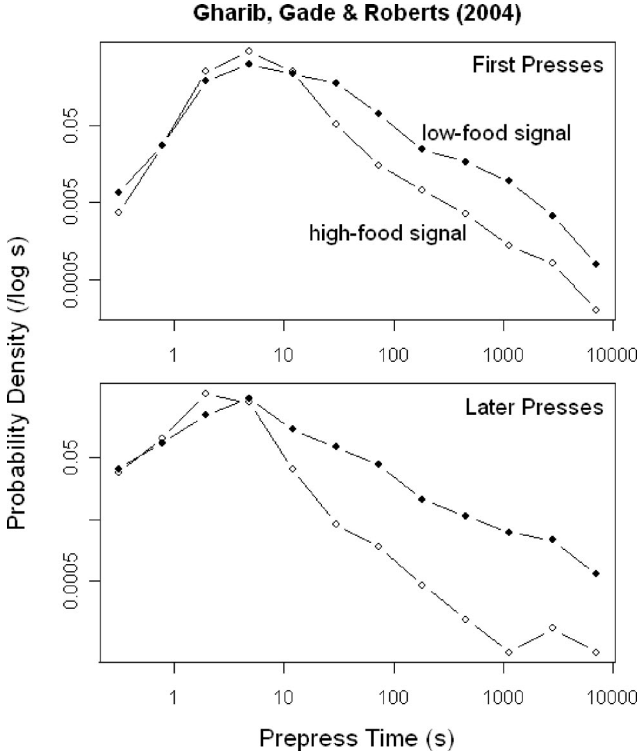

Figure 14. Preresponse time as a function of response order. Both axes are logarithmic. Data from Gharib, Gade, and Roberts (2004), Experiment 2, 100%/25% phase. First = data from the first bar presses during a trial. Later = data from all later bar presses. Each point is a mean over 11 rats.

would be that one mechanism generated short-duration pecks and another mechanism generated longer duration pecks, as Schwartz and Williams (1972) and Schwartz (1977) concluded.

The peak of short-duration (presumably more Pavlovian) pecks was reduced by the shift from first to later pecks. The necessary increase elsewhere (the area of the distribution always equals one), however, did not occur equally throughout the rest of the distribution. It was concentrated in the longer portion of the distribution (pecks 30 ms and more). The distributions of Figure 13 suggest there are three types of pecks: those reduced by the first-to-later shift (duration 12 ms or less); those unaffected by the first-to-later shift (duration 12 to 30 ms), and those increased by the first-to-later shift (duration 30 ms or more). Based on other peck-duration data, Ziriax and Silberberg (1978) likewise questioned a binary division of pecks. As they put it, “one could argue not just for two kinds of key pecks but for four” (p. 15).

The broad point is that the time until the first peck of a trial appears to be under Pavlovian control, whereas interpeck intervals are under instrumental control. This is the first evidence we know of that Pavlovian learning exerts substantial control over steady-state behavior in ordinary reinforcement schedules.

Generality of the PPT Results

The PPT results surprised us in two ways. First, the effect of probability of reward on variation (the standard deviation of the PPTs) was remarkably clear, even clearer (larger *t* value) than the mean effect, which is the law of effect applied to rate. Second, the first-peck/later-peck differences showed the importance of a factor (order in trial) that had gone unnoticed.

To assess the generality of these results, we examined two datasets from rats. One was from the second phase (100%/25%) of Experiment 2 of Gharib et al. (2004), the experiment on which our experiments were modeled. It had 11 subjects. The response was pressing a bar. There were two types of trials. During one (high food), trials always ended with food; during the other (low food), trials ended with food with a probability of 0.25.

Figure 14 shows the PPT distributions for first and later bar presses. As we found, reducing the likelihood of reward greatly increased variation. For both first and later responses, the difference between the two conditions was clearer with standard deviations than with means. For first bar presses, standard deviation $t(10) = 8.05$, mean $t(10) = 3.00$; for later bar presses, standard deviation $t(10) = 8.08$, mean $t(10) = 6.63$. Unlike our present results, in Figure 14 first- and later response distributions are similar.

We also reanalyzed data from Guilhardi and Church (2006), who measured the effect of extinction on fixed-interval performance. They trained rats on fixed-interval schedules of three different lengths (30, 60, and 120 s). The rewarded response was a head poke into a food cup. There were two experiments with two separate groups of 24 rats. Training lasted 20 or 30 sessions; extinction lasted 30 or 50 sessions. Guilhardi and Church noted that extinction “increases the [relative] frequency of short interre-

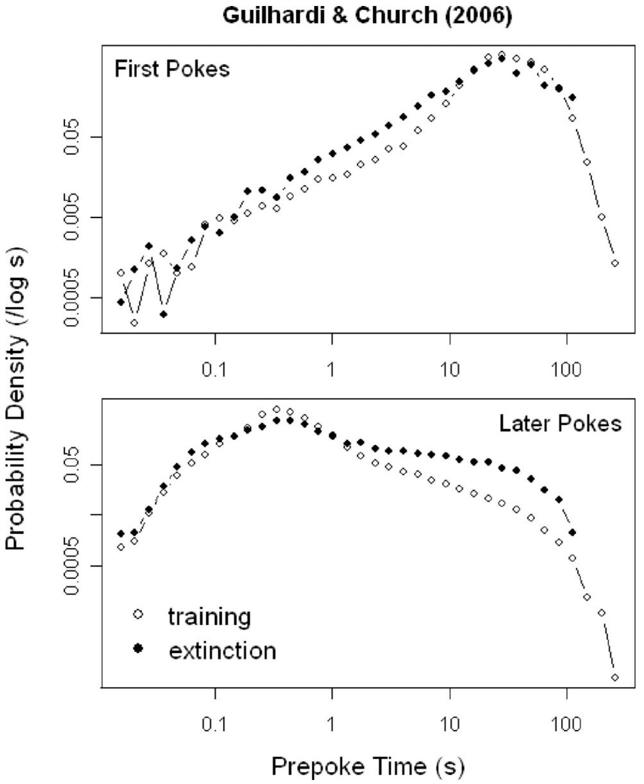

Figure 15. The distribution of prepoke times as a function of training condition. Data from Guilhardi and Church (2006). Each point is a mean over 24 rats, three fixed-interval values, and two experiments.

sponse times" (p. 269), which might mean the spread of the whole distribution increased.

Figure 15 shows the prepoke-time distributions from the last 10 training sessions and the last 30 extinction sessions averaged over both experiments and the three fixed-interval durations. For both first and later pokes, the distribution was much wider during extinction than training. The first-poke prepoke-times (upper panel of Figure 15) were shorter during extinction than during training (opposite to what the law of effect predicts). In three of the six cases, the difference was reliable (see Figure 17 for the *t* values). Perhaps this was due to loss of the time discrimination. In any case, it is interesting that the standard deviation increased during extinction at the same time the mean decreased. For later pokes, which did obey the law of effect (longer prepoke times in extinction than in training), the training/extinction difference was clearer with standard deviations than with means (see Figure 16). As in our results, first and later responses were quite different (compare the upper and lower panels of Figure 15).

Guilhardi and Church (2006) proposed that the increase in the relative frequency of short interresponse times during extinction be explained with "the concept of response packets" (p. 282). "A bout of head entry responses [a packet] may be defined as a sequence of responses separated by less than 2.5 sec," they wrote (p. 283). The increase of the relative frequency of short interresponse times, however, was one manifestation of a widening of the whole distribution (see Figure 15). The response-packet idea does not explain why the whole distribution became wider during extinction.

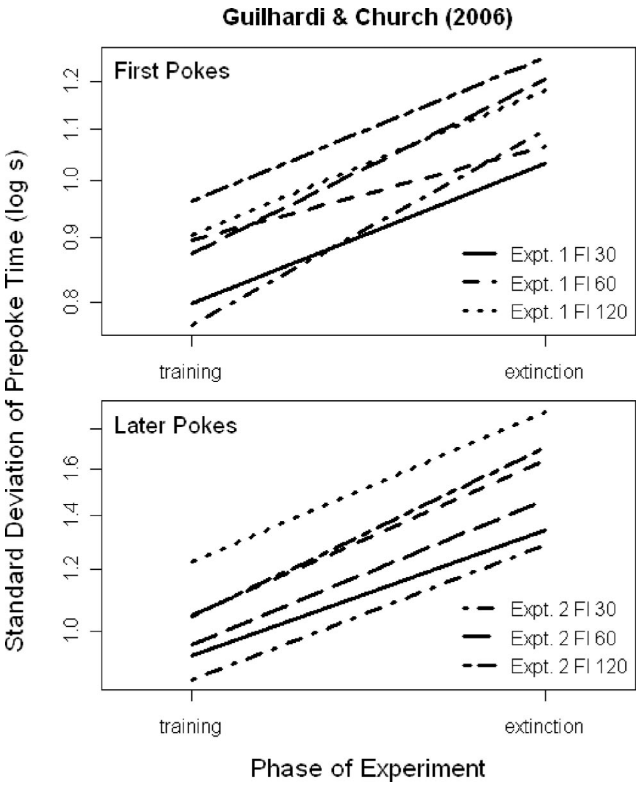

Figure 16. Standard deviations of prepoke times as a function of peck order and experimental phase. Data from Guilhardi and Church (2006). The legend identifying the various lines is in two parts: half in the upper panel, half in the lower panel. Each point is a mean over 24 rats.

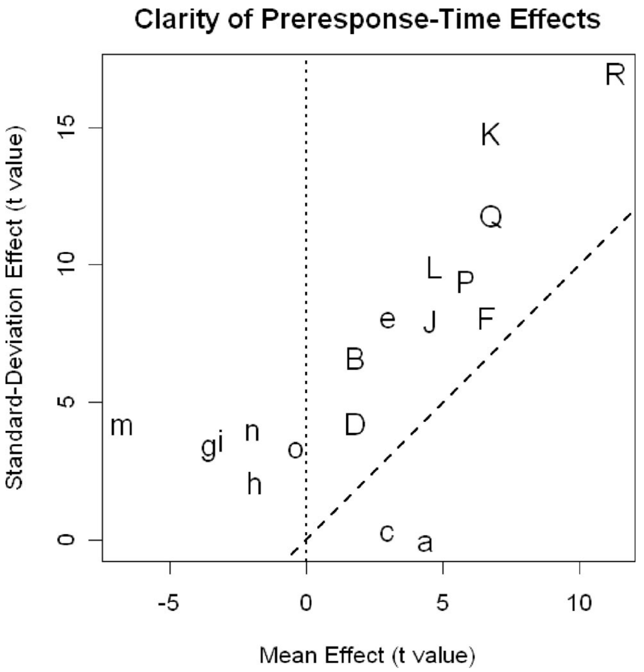

Figure 17. The relative clarity of standard deviation and mean effects. Each plotted value indicates two *t* values derived from the comparison of log preresponse times from two conditions that differed in the probability of reward. The *x* axis shows the *t* value of the effect of probability of reward on means; the *y* axis shows the *t* value of the effect on standard deviations. Lower-case letters come from the first responses of a trial; upper-case letters from the later responses. Each point is a different data set: a = Experiment 1, first pecks; B = Experiment 1, later pecks; c = Experiment 2, first pecks; D = Experiment 2, later pecks; e = Gharib et al. (2004), first bar presses; F = Gharib et al. (2004), later bar presses. The rest are from Guilhardi and Church (2006): g = Experiment 1, first pokes, fixed interval 30 s; h = Experiment 1, first pokes, fixed interval 60 s; i = Experiment 1, first pokes, fixed interval 120 s; J = Experiment 1, later pokes, fixed interval 30 s; K = Experiment 1, later pokes, fixed interval 60 s; L = Experiment 1, later pokes, fixed interval 120 s; m = Experiment 2, first pokes, fixed interval 30 s; n = Experiment 2, first pokes, fixed interval 60 s; o = Experiment 2, first pokes, fixed interval 120 s; P = Experiment 2, later pokes, fixed interval 30 s; Q = Experiment 2, later pokes, fixed interval 60 s; R = Experiment 2, later pokes, fixed interval 120 s.

Our finding of first-peck/later-peck differences was not true of both new datasets, but in both cases probability of reward did have a strong effect on preresponse-time variation, often stronger than its effect on means. Figure 16 compares the clarity (*t* value) of the standard deviation effect to the clarity of the mean effect for all three datasets— ours, Gharib et al. (2004), and Guilhardi and Church (2006). Each plotted letter is a different set of data, identified in the figure caption. The *t* value for an effect of probability of reward on standard deviations is plotted against the *t* value for an effect of probability of reward on means.

The points in Figure 16 divide into three clusters. One cluster (10 points) falls above the main diagonal. These are results where the standard-deviation effect was stronger than the mean effect.

The points in this cluster come from both rats and pigeons and from three different types of responses (pecks, bar presses, and head pokes). We take this to be normal instrumental learning. A second cluster (two points) is cases in which there was no standard-deviation effect. These were the first-peck results in our experiments. We take this cluster to indicate Pavlovian conditioning at work. The third cluster (six points) is cases where the mean effect was in the wrong direction, all from the first-poke results of Guilhardi and Church (2006). We suspect this is due to the loss of the time discrimination during extinction.

Our failure to find first-/later-response differences with the Gharib et al. (2004) data supports the Pavlovian/instrumental interpretation of the difference made earlier. Approach responses can be autoshaped; the food-getting responses that an animal makes between approach and consumption, such as bar pressing, cannot be autoshaped. That our PPT variation results correctly predicted results from other situations supports the idea that they are an example of a widespread feature of instrumental learning.

The Value of Transformations

A reviewer said that our use of transformations (e.g., analyzing log PPTs instead of raw PPTs) "distorted" the data and "underestimated" variation. Data transformation is good practice (Roberts, 2008; Tukey, 1977) but these comments may reflect common beliefs. We transformed our data for two reasons.

First, it made our statistical tests more sensitive (Roberts, 2008). Our analyses of peck duration failed to find an effect of reward probability. Because these failures came from more sensitive tests (on transformed data), they are more impressive than would be failures from less sensitive tests (on raw data). Transformations also helped us detect the interaction between peck order (first vs. later pecks) and measure (mean vs. standard deviation) on PPTs. With first pecks, means were more sensitive to reward probability (p = .004) than were standard deviations (p = .53); with later pecks, standard deviations (p = .0006) were more sensitive than were means (p = .07), using the p values of Experiment 1. Dividing the standard deviation of p by the mean p to get a sensitivity ratio, and then dividing the first-peck sensitivity ratio (0.53/0.004) by the later-peck sensitivity ratio (0.0006/0.07) gives a result of about 17,000. If the sensitivities were equal, this number would average 1. For Experiment 2, the same calculation gives a result of about 600. Those numbers come from log PPTs. Using raw PPTs, the same calculations give a result of 1.14 for both Experiments 1 and 2. Without transformation, in other words, the interaction is invisible.

Second, transformations made our graphs more informative (Tukey, 1977). Figure 18 illustrates this. It shows the data of one

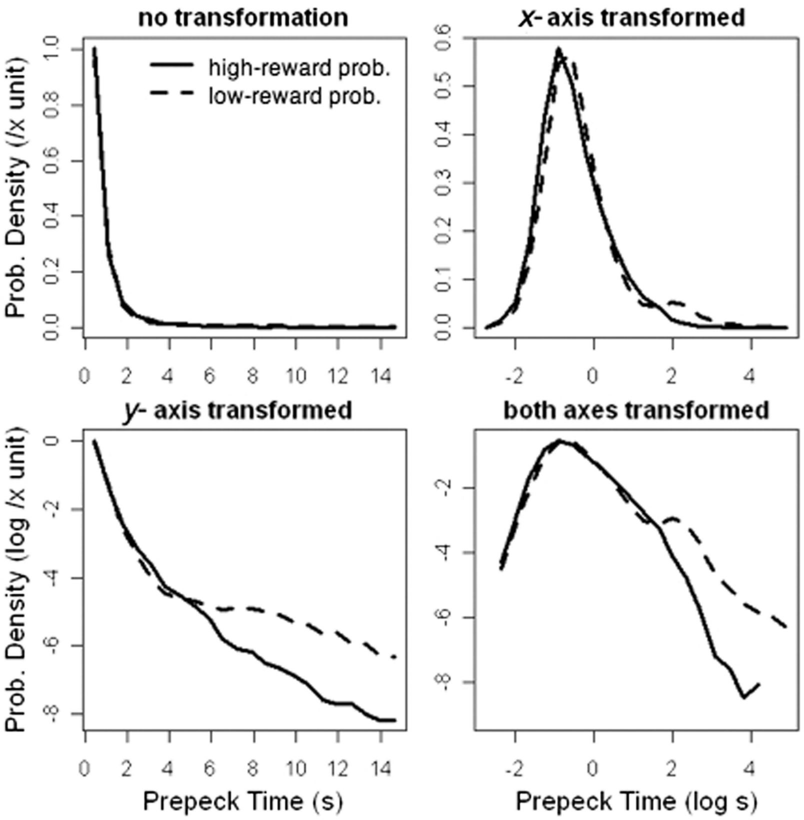

Figure 18. The value of transformations. The same data with and without log transformation of the x- and y-axes. The units of the y-axes are per second for the left-hand graphs and per log second for the right-hand graphs.

panel of Figure 7 (distributions of interpeck intervals) plotted four ways: with and without transformation of the *x*- and *y*-axes. With neither axis transformed (upper left), no difference is clear. With only the *x*-axis transformed (upper right), the overall shape of the distribution is clearer but the difference between the two distributions is barely visible. With only the *y*-axis transformed (bottom left), the difference between the two distributions at long intervals is clear but the similarity at short intervals—where most of the data are—is unclear. With both axes transformed (bottom right), both the similarity (at short intervals) and the difference (at long intervals) are easy to see.

Summary of Conclusions

Empirical

1. Status of the Gharib et al. (2001) rule: Two measures (horizontal location, interpeck interval) obeyed the rule; three (peck duration, vertical location, time to first peck) did not.

2. Large temporal variability effect: The effect of probability of reward on the variation of interpeck intervals was remarkably clear, even clearer than the effect of that factor on mean interpeck interval. The Gharib et al. (2001) and Guilhardi and Church (2006) results showed the same result.

Theoretical

3. First and later pecks controlled by different mechanisms: PPTs suggest big differences in what birds did before the first peck of a trial and what they did between pecks. The variation of times to first peck was insensitive to probability of reward; the variation of interpeck intervals was highly sensitive. The function relating time to first peck to probability of reward showed an anomaly; the same function for interpeck intervals did not. Interpeck-interval distributions suggested a two-state model with states engagement and disengagement; distributions of time to first peck did not. First and later pecks also differed in duration.

4. Support for regulation of variation: The exceptions we found to the Gharib et al. (2001) rule make evolutionary sense. Reward expectation controls variation in how a food patch is explored and exploited, but does not control variation in travel to the patch, nor does it control variation in the actual consumption of food.

References

Antonitis, J. J. (1951). Response variability in the white rat during conditioning, extinction, and reconditioning. *Journal of Experimental Psychology, 42,* 273–281.

Balsam, P. D., Deich, J. D., Ohyama, T., & Stokes, P. D. (1998). Origins of new behavior. In W. O'Donohue (Ed.), *Learning and behavior therapy* (pp. 403–420). Boston: Allyn & Bacon.

Beck, C. H., & Loh, E. A. (1990). Reduced behavioral variability in extinction: Effects of chronic treatment with the benzodiazepine, diazepam or with ethanol. *Psychopharmacology, 100,* 323–327.

Blough, D. S. (1963). Interresponse time as a function of continuous variables: A new method and some data. *Journal of the Experimental Analysis of Behavior, 6,* 237–246.

Bowers, M. T., Hill, J., & Palya, W. L. (2008). Interresponse time structures in variable-ratio and variable-interval schedules. *Journal of the Experimental Analysis of Behavior, 90,* 345–362.

Breland, K., & Breland, M. (1961). The misbehavior of organisms. *American Psychologist, 16,* 681–684.

Brown, P. L., & Jenkins, H. M. (1968). Auto-shaping the pigeon's key peck. *Journal of the Experimental Analysis of Behavior, 11,* 1–8.

Cherot, C., Jones, A., & Neuringer, A. (1996). Reinforced variability decreases with approach to reinforcers. *Journal of Experimental Psychology: Animal Behavior Processes, 22,* 497–508.

Collins, L., & Pearce, J. M. (1985). Predictive accuracy and the effects of partial reinforcement on serial autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes, 11,* 548–564.

Crow, L. T. (1978). A comparison of the effects of extinction and satiety on operant response duration in the rat. *Bulletin of the Psychonomic Society, 11,* 86–88.

Devenport, L. D. (1984). Extinction-induced spatial dispersion in the radial arm maze: Arrest by ethanol. *Behavioral Neuroscience, 98,* 979–985.

Eckerman, D. A., & Lanson, R. N. (1969). Variability of response location for pigeons responding under continuous reinforcement, intermittent reinforcement, and extinction. *Journal of the Experimental Analysis of Behavior, 12,* 73–80.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement.* New York: Appleton-Century-Crofts.

Frick, F. C., & Miller, G. A. (1951). A statistical description of operant conditioning. *American Journal of Psychology, 64,* 20–36.

Gharib, A., Derby, S., & Roberts, S. (2001). Timing and the control of variation. *Journal of Experimental Psychology: Animal Behavior Processes, 27,* 165–178.

Gharib, A., Gade, C., & Roberts, S. (2004). Control of variation by reward probability. *Journal of Experimental Psychology: Animal Behavior Processes, 30,* 271–282.

Guilhardi, P., & Church, R. M. (2006). The pattern of responding after extensive extinction. *Learning & Behavior, 34,* 269–284.

Hearst, E., & Jenkins, H. M. (1974). *Sign-tracking: The stimulus-reinforcer relation and directed action.* Austin, TX: Psychonomic Society.

Holt, D. D., Green, L., & Muenks, W. M. (2004). Biological and economic effects on responding: Rate and duration of the pigeon's key peck. *International Journal of Comparative Psychology, 17,* 203–221.

Jenkins, P. E. (1981). The determiners of keypeck duration. *Animal Learning & Behavior, 9,* 501–507.

Mechner, F. (1958). Sequential dependencies of the lengths of consecutive response runs. *Journal of the Experimental Analysis of Behavior, 1,* 229–233.

Mechner, F., Hyten, C., Field, D. P., & Madden, G. (1997). Using revealed operants to study the structure and properties of human operant behavior. *Psychological Record, 47,* 45–68.

Morgan, D. L., & Lee, K. (1996). Extinction-induced response variability in humans. *Psychological Record, 46,* 145–159.

Neuringer, A., Kornell, N., & Olufs, M. (2001). Stability and variability in extinction. *Journal of Experimental Psychology: Animal Behavior Processes, 27,* 79–94.

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,* 532–552.

Rescorla, R. A. (2004). Spontaneous recovery. *Learning and Motivation, 11,* 501–509.

Roberts, S. (2008). Transform your data. *Nutrition, 24,* 492–494.

Roberts, S., & Gharib, A. (2006). Variation of bar-press duration: Where do new responses come from? *Behavioral Processes, 75,* 215–223.

Schwartz, B. (1977). Studies of operant and reflexive key pecks in the pigeon. *Journal of the Experimental Analysis of Behavior, 27,* 301–313.

Schwartz, B. (1980). Development of complex, stereotyped behavior in pigeons. *Journal of the Experimental Analysis of Behavior, 33,* 153–166.

Schwartz, B., & Williams, D. R. (1972). Two different kinds of key peck in the pigeon: Some properties of responses maintained by negative and positive response-reinforcer contingencies. *Journal of the Experimental Analysis of Behavior, 18,* 201–216.

Shull, R. L., Gaynor, S. T., & Grimes, J. A. (2001). Response rate viewed as engagement bouts: Effects of relative reinforcement and schedule type. *Journal of the Experimental Analysis of Behavior, 75,* 247–274.

Staddon, J. E. R., & Simmelhag, V. L. (1971). The "superstition" experiment: A reexamination of its implications for the principles of adaptive behavior. *Psychological Review, 71,* 3–43.

Stokes, P. (2005). *Creativity from constraints.* New York: Springer.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction.* Cambridge, MA: MIT Press.

Timberlake, W. (1983). The functional organization of appetitive behavior: Behavior systems and learning. In M. D. Zeiler & P. Harzem (Eds.), *Advances in analysis of behavior: Vol. 3. Biological factors in learning* (pp. 177–221). Chichester, England: Wiley.

Timberlake, W., & Lucas, G. A. (1989). Behavior systems and learning: From misbehavior to general principles. In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theory: Instrumental conditioning theory and the impact of biological constraints on learning* (pp. 237–276). Hillsdale, NJ: Erlbaum.

Tolkamp, B. J., & Kyriazakis, I. (1999). To split behaviour into bouts, log-transform the intervals. *Animal Behaviour, 57,* 807–817.

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison Wesley.

Ziriax, J. M., & Silberberg, A. (1978). Discrimination and emission of different key-peck durations in the pigeon. *Journal of Experimental Psychology: Animal Behavior Processes, 4,* 1–21.

Received January 4, 2008
Revision received February 27, 2009
Accepted March 2, 2009 ■