Variation of Bar-Press Duration: Where Do New Responses Come From?

Seth Roberts[a]*, Afshin Gharib[b]

[a]Department of Psychology, University of California, Berkeley, CA 94708 USA

[b]Department of Psychology, Dominican University of California, 50 Acacia Avenue,

San Rafael, CA 94901

*Corresponding author. Tel: +1 510 418 7753; fax: +1 267 222 4105.

E-mail address: twoutopias@gmail.com (S. Roberts).

Abstract

Instrumental learning involves both variation and selection: variation of what the animal does, selection by reward from among the variation. Four experiments with rats suggested a rule about how variation is controlled by recent events. Experiment 1 used the peak procedure. Measurements of bar-press durations showed a sharp increase in mean duration after the time that food was sometimes given. The increase was triggered by the omission of expected food. Our first explanation of the increase was that it was a frustration effect. Experiment 2 tested this explanation with a procedure in which the first response of a trial usually produced food, ending the trial. In Experiment 2, unlike Experiment 1, omission of expected food did not produce a large increase in bar-press duration, which cast doubt on the frustration explanation. Experiments 3 and 4 tested an alternative explanation: *A decrease in expectation of reward increases variation*. Both used two signals associated with different probabilities of reward. Bar presses were more variable in duration during the signal with the lower probability of reward, supporting this alternative. These experiments show how variation can be studied with ordinary equipment and responses.

Instrumental learning is usually thought of as a feedback system: The effect of a response controls its likelihood. Mackintosh (1983), for example, said that “instrumental conditioning [occurs] when [a] change in behavior is a consequence of a contingency between that behavior and the reinforcer” (p. 77). According to Dickinson (1994), “instrumental behavior refers to those actions whose acquisition and maintenance depend on their consequences for the animal” (p. 45). However, this way of thinking is incomplete. Instrumental learning involves both variation and selection (Hull, Langman, & Glenn, 2001; Staddon, 1983; Staddon & Simmelhag, 1971). The feedback view says nothing about variation – that is, it says nothing about where new responses come from.

Variation can be where, when, or how – variation of location, timing, or topography. Early studies of variation tended to measure topography. For example, Antonitis (1951) studied variation in the topography of nose pokes. The raw data was 6,600 photographs. Iversen (2002) did something similar with a digital camera. Recent studies of variation have usually involved location (e.g., Neuringer, 2002). For example, Page and Neuringer (1985) looked at sequences of responses in two locations. Variation was indexed by the number of different sequences that occurred.

Both sorts of studies have disadvantages. Photographs must be quantified by hand – at least, no one has quantified them by computer. This is very time-consuming. On the other hand, if you measure variation by number of sequences, the “response” whose variation is measured becomes more artificial (less like natural foraging) the longer the sequence; and the longer the sequence

the harder it is to generalize to individual responses. Machado (1997) concluded that when you increase sequence variability by rewarding it you cause the animal to adjust its probability of switching from one key to the other. Another problem with using sequences to measure variation is granularity. If you study sequences of length four with two keys, there are only 16 possible sequences.

This article reviews our work on variation (Gharib, Derby, & Roberts, 2001; Gharib, Gade, & Roberts, 2004). Our experiments do two things: (a) show how to study variation with standard equipment and a conventional response in a completely automated way; and (b) suggest a new rule about the control of variation. The most firmly established generalization about variation has been that it increases during extinction (e.g., Antonitis, 1951; Balsam, Deich, Ohyama, & Stokes, 1998; Iversen, 2002). But there are several differences between training and extinction: density of reward, rate of response, and expectation of reward, to name a few. It is unclear which difference or differences cause the increase in variability. Our work points to one of them: the decrease in expectation of reward.

We came to study variation by accident – at first, without realizing it. In experiments similar to Experiment 1 of this article, we noticed a large change in bar-press duration. Experiment 2 showed that our first explanation of this change was wrong. We eventually thought of another explanation, which Experiment 3 tested. The results supported the new explanation, but the experiment had a serious flaw. When the flaw was fixed, in Experiment 4, support for the new explanation was quite clear.

Experiment 1: Peak-Procedure Puzzle

This experiment used the peak procedure (Catania, 1970; Roberts, 1981), which resembles a discrete-trials fixed-interval schedule. Trials are indicated by the onset of a signal, such as a house light. On most trials, the first response after a fixed time (here, 40 s) is rewarded and the signal ends when food is given; these are called *food trials*. On the remaining trials (*empty trials*), the signal lasts a long time (much longer than 40 s) and ends without food independently of what the animal does.

The results shown in Figure 1 are from two groups of 18 rats each. The response was pressing a bar. One group was observed for 48 six-hour sessions (0.7 million responses), the other for 124 six-hour sessions (2.8 million responses).

Most uses of the peak procedure (e.g., Roberts, 1981) have focused on the time discrimination it produces – the change in bar-press rate as a function of time (upper panel of Figure 1). The lower panel of Figure 1 shows something else: bar-press *duration* – the length of time that the bar is held down – as a function of time. In all of these experiments, food was given after the lever was pressed down and released – so bar-press durations were unaffected by the delivery of food. We had no theoretical reason for looking at this graph; we just hoped it would tell us something new (Roberts, 1984).

The duration function (lower panel) was a great surprise – so different from the rate function (upper panel) that at first it seemed to be a mistake. The rate function was symmetric around the time of food (about 45 seconds after the

start of the trial). In contrast, the duration function was highly asymmetric around that time.

The most striking feature of the duration function was the sharp rise that began at about the time of food. An analysis showed that this increase was triggered by one event: the omission of expected food (Figure 2). When a rat pressed the bar after Second 40 (that is, more than 40 s after the start of the trial), it did not yet know whether the current trial was food or empty. If this particular bar press was not rewarded, there would be no food for the rest of the trial. Figure 2 divides responses into two groups: (a) all responses up to and including the first response after Second 40 (*before omission*) and (b) all the responses after that (*after omission*). Figure 2 shows that bar presses up to and including the first one after Second 40, no matter how late in the trial, did not increase in duration over time. It was only bar presses *after* the first bar press after Second 40 that had longer durations.

That the rise in duration was triggered by omission of expected food led to our first explanation of the cause of the rise: It was a frustration effect. Amsel and Roussel (1952) and many others had found that omission of expected food in the first goal box of a double runway caused rats to run faster on the second runway (Amsel, 1962). Perhaps omission of expected food also caused rats to press a bar for a longer period of time.

Experiment 2: Test of Frustration Hypothesis

If the frustration hypothesis is correct, any omission of expected food should produce a duration increase. In Experiment 2, we tested this prediction

with a simple procedure. Intertrial intervals were dark. Now and then a light went on. The next bar press turned off the light and, with probability 0.8, was rewarded with a food pellet. When no food was given – that is, when expected food was omitted – would there be a duration increase similar to Experiment 1's (lower panel of Figure 1)? Between trials, responses were rewarded on a random-interval 100 s schedule. (That is, reward was "primed" with probability 0.01 each second. When reward was primed, the next response produced reward.) We rewarded responses with a low probability (rather than give no reward at all) to make it easier to see a duration increase.

The results were not what we expected. The upper panel of Figure 3 shows bar-press rates before, during, and after trials; the lower panel shows bar-press durations. There was only a small and short-lived increase in response duration after the omission of expected food, in great contrast to the results of Experiment 1.

The frustration hypothesis – which we had believed (Roberts, 1998) -- was apparently wrong. Something unknown to us, triggered by a single unrewarded response, had produced a large and long-lasting change in response duration. This was a real mystery.

Re-examination of the results of Experiment 1 suggested a possible answer. Figure 4 shows the distribution of durations before and after food omission (with both *x* and *y* coordinates logarithmic). The increase in mean duration seen in Figure 1 was due to a large increase in the variability of a very skewed distribution. Figure 5, which shows percentiles of the distribution as a

function of time into the trial, makes the same point. The distance between the tenth percentile and the ninetieth percentile (a measure of spread) increased a great deal from early to late in the trial.

Thinking of the increase in bar-press duration as an increase in variability suggested a new explanation of it. A striking feature of the duration increase was how long it lasted: Until the end of the trial (or at least the measurement period). Non-delivery of food for the first response after Second 40 was not only a source of frustration but also a discriminative stimulus: It signaled no reward for the rest of the trial. If a rat learned this discrimination and remembered the event, the omission of expected food would cause a sharp drop in expectation of food for the rest of the trial.

Perhaps *reducing expectation of reward increases variability*. This proposal had three things in its favor. First, it explained the time course of the duration increase. Because the reduced expectation lasted the rest of the trial, so should the duration increase. Second, it made conceptual sense. Variation has costs and benefits. As the expectation of reward goes down, the cost of variation decreases – so the optimal amount of variation should increase. Nothing restricts this reasoning to animal behavior; the General Discussion describes similar empirical rules in genetics and business. Third, the novelty of the idea roughly matched the strength of the effect. The effect was very strong, $t(35) = 18$ for comparison of the two distributions in Figure 4. In an unstudied area, such as variation, there might be large effects not yet discovered. In a well-studied area, such as selection by reward, this was unlikely.

Other results of Experiment 2 gave the proposal more support. First, *trial versus intertrial interval.* The probability that a bar press would be rewarded was much higher during a trial (0.8) than between trials (0.07), so expectation of reward should have been greater during a trial than between trials. The new idea correctly predicted that durations were less variable during a trial (upper panel of Figure 6). Second, *after reward versus after nonreward.* Between trials, expectation of reward was surely greater after a rewarded response than after an unrewarded response. The new idea correctly predicted that durations were less variable after a rewarded response (lower panel of Figure 6). Third, *context specificity of the reward/nonreward effect*. When a trial ended with food, this should have increased expectation of food during future trials but not during the upcoming intertrial interval. As this predicts, the effect of reward/nonreward did *not* extend to a different context. There was no reliable difference in duration between intertrial responses after a trial that ended with food and intertrial responses after a trial that ended without food. The interaction (reward/nonreward by location of reward—end of trial or intertrial interval) was reliable.

Experiment 3: Test of Expected-Reward Hypothesis (First Try)

To test the new idea – less expectation of reward, more variation -- we varied expectation of reward in the simplest well-controlled way we could think of. Experiment 3 used a discrete-trials procedure with two signals. During both signals, the end of the trial was “primed” with probability 1/60 each second – like a random-interval schedule. Once the end of the trial was primed, the next bar

press ended the trial. The main comparison was between signals. *High-food* trials (one signal) always ended with food. *Low-food* trials (the other signal) ended with food with probability 1.0 during an initial baseline phase, with probability 0.50 during the first treatment phase, and with probability 0.25 during the second treatment phase. Between trials, no responses were rewarded.

Figure 7 shows the results. The duration results were in the predicted direction but weakly so. Only during the second treatment phase was the difference between the two signals reliable (one-tailed $p < 0.05$).

Puzzled that the results were not clearer, we realized there was a problem with the procedure: The more time since the previous response, the more likely that the end of the trial has been primed – thus the more likely the next response will be rewarded. If the rats measured the time since their last response, they could notice that the probability of reward increased as this time grew, and their expectation of food could grow. Although reward/time was less during the low-food signal than during the high-food signal, reward/response was not necessarily lower. Perhaps the expectation of reward that controlled variation depended on reward/response, not reward/time.

Analysis of the results according to interresponse time (Figure 8) supported this conclusion. The top panel shows that response duration depended on the time since the last response. As interresponse time increased, mean response duration decreased (meaning that variation decreased), supporting the idea that the rats were measuring time since their last response. The middle panel shows that how the probability of reward per bar press

increased with interresponse time. The bottom panel combines the top two panels to show that duration decreased as probability of reward increased. Figure 9 is the same as the bottom panel of Figure 8 except that it shows the 10th and 90th percentiles of the distributions rather than the means, in order to make clear the decrease in variation (here, the distance between the 10th percentile and the 90th percentile) as probability of reward increases. As Gharib et al. (2004) wrote, summarizing the bottom panel of Figure 8 and Figure 9, "duration differences between the conditions are almost entirely explained by differences in the probability of reward" (p. 275). We had varied the wrong thing. We had varied *density* of reward between the signals; we should have varied *probability* of reward.

Experiment 4: Test of Expected-Reward Hypothesis (Second Try)

Experiment 4 fixed the problem. It closely resembled Experiment 3. The main difference was that trials ended with a fixed probability (0.25) each response. This meant that, unlike Experiment 3, the probability of reward did not vary with interresponse time. Again, there were two types of trials, high food and low food, indicated by different signals (light and sound). All high-food trials ended with food. During a baseline phase, low-food trials ended with food with probability 1.0; during the treatment phase, they ended with food with probability 0.25.

This produced much clearer results (Figure 10). Reducing the probability of reward clearly increased response duration. The difference in duration between the two signals ($t(10) = 6.22$, based on the last 25 days of the treatment

phase) was as clear as the difference in rate ($t(10) = 6.07$). Figure 11 shows the distribution of bar-press durations. The lower the probability of reward, the greater the variation.

More support for an effect of reward expectation on variation comes from a within-signal comparison: low-food trials after many unrewarded bar-presses (less expectation) versus the same trials after zero or only a few unrewarded bar-presses (more expectation). To make this comparison, we calculated the number of consecutive unrewarded bar presses during the low-food signal at the beginning of every trial. This number "labels" the trial, whether low-food or high-food, and all bar presses during that trial are classified using that label. Whenever the most recent low-food trial has ended with reward, this number is reset to zero. It increases when a low-food trial ends without reward; if the low-food trial contained five responses (all unrewarded), then this number increases by five after the trial ends. It does not change during a trial and is unaffected by high-food trials. Suppose that this number is zero at the start of a session. The first trial (low-food) ends without reward after five bar presses. Now the number is five. The second trial (high-food) ends with reward after three bar presses. The number remains five. The third trial (low-food) ends without reward after ten bar presses. Now the number is 15. The fourth trial (low-food) ends with reward. Now the number is zero again. In general, the number measures how much "extinction" the low-food signal has recently experienced.

Figure 12 shows how recent extinction affects rate (upper panel) and duration (lower panel) during both signals. The most important finding is that

duration during the low-food signal increased as recent extinction increased, $t(10) = 2.88$, $p < 0.01$. The associative nature of the effect is shown by the fact that there was no reliable increase during the high-food signal and there was a reliable signal-by-recent-extinction interaction, $t(10) = 2.96$, $p < 0.01$. That the effect was associative supports the idea that it is due to expectancy.

General Discussion

Control of Variation by Reward Expectation

The results support the general rule that *reducing reward expectation increases variation of response form*. This rule predicted or explains nine differences in bar-press duration, summarized in Table 1, which gives in each case the two conditions that differed in reward expectation and the *t* value for the duration difference. None of the comparisons of Table 1 is “pure” (unconfounded); the two conditions always differed in other ways as well. But it seems safe to say the difference in reward expectation is the only difference present in every case.

The most plausible alternative explanation of Table 1’s results is probably that a decrease in rate increases variation. In most cases, expectation and rate were positively correlated, as you would expect: When expectation was lower, rate was lower. In Experiment 3, however, because of the interval schedule, longer interresponse times (= lower rate) were associated with a greater probability of reward. In this situation, the rule *lower rate, more variation* predicted the opposite of *lower reward expectation, more variation*. And it was the reward-expectation rule that predicted correctly: longer interresponse times

were associated with *less* variation (Figure 8). Gharib, Derby, and Roberts, (2001) describe several other reasons to believe that rate has little effect on variation.

The rule *less expectation of reward, more variation* explains the well-established finding that variation increases during extinction (Balsam et al, 1998; Neuringer, 2002) because expectation of reward decreases during extinction. Support for the rule comes from a wide range of responses, including rat nose pokes (Antonitis, 1951), sequences of responses made by rats (Balsam et al., 1998; Cherot, Jones, & Neuringer, 1996) and pigeons (Schwartz, 1982), and the location of pigeon key pecks (Eckerman & Lanson, 1969; Ferraro & Branch, 1968; Millenson, Hurwitz, & Nixon, 1961). In the peak procedure, lowering the probability of reward increases the spread of the response-rate function – that is, responses become more variable in time (Roberts, 1981). So the rule seems to apply to all three types of variation (location, timing, and topography).

One reviewer asked: What does your invocation of expectancy buy you? Why not just talk about the relation of probability of reward and variation? This is a fair question because in every case supporting the new rule expectancy of reward was varied by varying probability of reward. The answer is that you can vary expectancy while keeping probability of reward constant – by increasing delay of reward, for example. Conclusions about expectancy lead to predictions, in other words, that conclusions in terms of probability of reward do not.

Similar Rules in Other Areas

The rule *less expectation of reward, more variation* has parallels in genetics and business (Gharib, Gade, & Roberts, 2004). In genetics, Susan Lindquist and her colleagues have found that a similar rule holds true. It began with the discovery of what are called *heat shock proteins* in fruit flies. When a fly is exposed to warm temperature or other environmental stressors, these proteins function differently. In the case of one heat-shock protein, Hsp90, the change is an increase in phenotypic variation (Queitsch, Sangster, & Lindquist, 2002; Rutherford & Lindquist,1998). Usually this protein prevents mutations from producing phenotypic changes. In the absence of expression (and therefore selection), the genetic changes build up, forming a kind of reserve. During times of stress, when these mutations might be beneficial rather than harmful, they are expressed. “The buffered variation is therefore released precisely under those challenging conditions when selection is most stringent [i.e., the chance of successful reproduction is less than usual] and novelty might be most beneficial” (Sangster, Lindquist, & Queitsch, 2004, p. 356). Mustard plants show similar effects (Sangster, Lindquist, & Queitsch, 2004).

In business, Christensen (1997) noticed a puzzling pattern: Industry leaders commonly lost their lead when new technologies came along. For example, the companies that sold the most 14-inch disk drives did not become the companies that sold the most 8-inch drives. The leaders in 8-inch drives did not become the leaders in 5.25-inch drives. And so on. He found similar examples in retailing, computers, mechanical excavators, and steel mills.

Christensen's research suggested that the reason for this pattern was that industry-leading companies were "held captive by their customers" (p. 18)–that is, by their success and by "expected rewards" (p. 32). They did not vary enough what they did. Less successful companies were more adventurous. Surowiecki (2003, p. 46) gave a within-company example of the same thing: "Once Boeing became the dominant player in the aviation market, in the seventies, it lost its appetite for sporty bets [involving the manufacture of new airplanes]: why risk them, when profits were rolling in?" Battelle (2005) tells a similar story about the leaders of Excite, an early Web search business. "Because of their early success, they were closed-minded," he quotes someone as saying (Battelle, 2005, p. 55). "Nothing deceives like success," says someone else (Battelle, 2005, p. 55). To not invest enough in new technologies could be fatal; companies that lost their lead when a new technology arrived often went out of business. That technology changes and one must keep up is not a subtle point. Presumably most of the leaders in these companies knew this. They just didn't act on it, as if a hard-wired tendency overrode common sense.

Methodology

The main methodological contribution of this work is to show how to study variation using widely-available equipment and an "ordinary" response. No procedural changes are needed. Any bar-pressing experiment can study variation if bar-press durations are measured.

One reason to transform one's data (Behrens, 1997) – use log *x* instead of *x,* for example -- is to make regularities easier to see. In the analysis of

Experiment 1, transformation had this effect. Transformation of the histogram (*x* to log *x*, *y* to log *y*) in Figure 4 to simplify (linearize) the shape of the distribution made it much easier to see that the change from early to late in the trial was an increase in variation. Without transformation, this would have been much harder, perhaps impossible, to see because the change in the tail (high durations) would have been too small.

A Fuller View of Instrumental Learning

Instrumental learning is interesting because it occurs in nature. Lab studies try to recreate nature. From this point of view it isn't enough to say that instrumental learning involves variation as well as selection; it is important to produce a laboratory example where variation was clearly controlled by something likely to vary in nature. This is what Experiment 4 accomplished.

The usual view of instrumental learning (Dickinson, 1994; Mackintosh, 1983; "selection by consequences", Skinner, 1981, p. 501) is too simple. Instrumental learning involves variation as well as selection, and consequences control both. Because consequences control variation, whether a response is rewarded can have a big effect on the rate of *other* responses. Some of these connections are obvious: Yes, if the rate of one response goes down, the rate of all other behaviors (added together) will go up, if the animal remains equally active. Yes, if one response is not rewarded and its rate goes down, the rate of similar responses will go down (generalization). The dependency we saw was neither of these: It was that when a response was not rewarded, the rate of similar responses went *up* in some cases. This is not obvious. In Experiments 1-

4, almost all bar presses were short-duration (< 1 s). To reduce the probability of reward was, in effect, to reduce the probability of reward of short-duration bar presses. This *increased* the rate of long-duration (> 2 s) bar presses. From early in the trial to late in the trial, the rate of all bar presses decreased by a factor of about 10 (upper panel of Figure 1). The fraction of all bar presses that were long-duration (duration > 2 s) increased by a factor of about 100 (Figure 4). Thus the rate of long-duration bar presses increased by a factor of about 10.

The effects we observed also imply that some responses can be generated and maintained by the effects of *other* responses – making the usual definition of instrumental learning (responses maintained by their consequences, e.g., Dickinson, 1994) too narrow. During acquisition, the mean duration of bar presses decreased considerably (Gharib, Derby, & Roberts, 2001). If we think of short-duration bar presses as skilled bar presses, in the beginning it is the consequences of *long* (unskilled) bar presses that generated short bar presses. After the response was well-learned, the situation reversed: the consequences of *short* bar presses (frequent) maintained *long* bar presses (rare).

A more complete view of instrumental learning – including variation as well as selection – should lead to the discovery of new empirical rules (e.g., Experiments 1-4) and help connect animal learning with other areas of study (e.g., genetics, business).

Acknowledgements

We thank Steven Derby, Michelle Dokey, Christopher Gade, Jonathan Herberg, and Brian Louie for their help.

References

Amsel, A., 1962. Frustrative nonreward in partial reinforcement and discrimination learning: Some recent history and a theoretical extension. Psych. Rev., 69, 306-328.

Amsel, A., and Roussel, J., 1952. Motivational properties of frustration: I. Effect on a running response of the addition of frustration to the motivational complex. J. Exp. Psych., 43, 363–368.

Antonitis, J., 1951. Response variability in the white rat during conditioning, extinction, and reconditioning. J. Exp. Psych., 42, 273-281.

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

Battelle, J. (2005). The search: How Google and its rivals rewrote the rules of business and transformed our culture. New York: Portfolio,

Behrens, J. T., 1997. Principles and procedures of exploratory data analysis. Psych. Methods, 2, 131-160.

Catania, A. C., 1970. Reinforcement schedules and psychophysical judgments: A study of some temporal properties of behavior. In: Schoenfeld, W. N. (Ed.), The Theory of Reinforcement Schedules. New York: Appleton-Century-Crofts, New York, pp. 1-42.

Cherot, C., Jones, A., and Neuringer, A.,1996. Reinforced variability decreases with approach to reinforcers. J. Exp. Psych. Anim. Behav. Proc., 22, 497-508.

Christensen, C. M., 1997. The Innovator's Dilemma: When New Technologies Cause Great Firms To Fail. Harvard Business School Press, Cambridge, MA, 225 pp.

Dickinson, A., 1994. Instrumental conditioning. In: Mackintosh, N. J. (Ed.), Animal Learning and Cognition. Academic Press, San Diego, CA, pp. 45-79.

Eckerman, D. A., and 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.

Ferraro, D. P., and Branch, K. H., 1968. Variability of response location during regular and partial reinforcement. Psychol. Rep., 23, 1023-1031.

Gharib, A., Derby, S., and Roberts, S., 2001. Timing and the control of variation. J. Exp. Psych. Anim. Behav. Proc., 27, 165-178.

Gharib, A., Gade, C., and Roberts, S., 2004. Control of variation by reward probability. J. Exp. Psych. Anim. Behav. Proc., 4, 271-282.

Hull, D. L., Langman, R. E., and Glenn, S.S., 2001. A general account of selection: Biology, immunology and behavior. Behav. Brain Sci., 24, 511-573.

Iversen, I. H., 2002. Response-initiated imaging of operant behavior using a digital camera. J. Exp. Anal. Behav., 77, 283-300.

Machado, A., 1997. Increasing the variability of response sequences in pigeson by adjusting the frequency of switching between two keys. J. Exp. Anal. Behav., 68, 1-25.

Mackintosh, N. J., 1983. Conditioning and Associative Learning. Oxford University Press, New York, 316 pp.

Millenson, J. R., Hurwitz, H. M. B., and Nixon, W. L. B., 1961. Influence of reinforcement schedules on response duration. J. Exp. Anal. Behav., 4, 243-250.

Neuringer, A., 2002. Operant variability: Evidence, functions, and theory. Psychon. Bull. Rev., 9, 672-705.

Page, S., & Neuringer, A., 1985. Variability is an operant. J. Exp. Psych. Anim. Behav. Proc., 11, 429-452.

Queitsch, C., Sangster, T. A., and Lindquist, S., 2002. Hsp90 as a capacitor of phenotypic variation. Nature, 417, 618-624.

Roberts, S., 1981. Isolation of an internal clock. J. Exp. Psych. Anim. Behav. Proc., 7, 242-268.

Roberts, S., 1984. What then should we do? Behav. Brain Sci., 7, 532-533.

Roberts, S., 1998. The mental representation of time: Uncovering a biological clock. In: Scarborough, D., Sternberg, S. (Eds.), Methods, Models, and Conceptual Issues. In: D. N. Osherson (Ed.), An Invitation To Cognitive Science, vol. 4, MIT Press, Cambridge, MA, pp. 53-106.

Rutherford, S. L., and Lindquist, S., 1998. Hsp 90 as a capacitor for morphological evolution. Nature, 396, 336-342.

Sangster, T., Lindquist, S., and Queitsch, C., 2004. Under cover: Causes, effects, and implications of Hsp90-mediated genetic capacitance. BioEssays, 26, 348-362.

Schwartz, B., 1982. Interval and ratio reinforcement of a complex sequential operant in pigeons. J. Exp. Anal. Behav., 37, 349-357.

Skinner, B. F., 1981. Selection by consequences. Science, 213, 501-504.

Staddon, J. E. R., 1983. Adaptive Behavior and Learning. Cambridge University Press, Cambridge, England, 555 pp.

Staddon, J. E. R., and Simmelhag, V. L., 1971. The “superstition” experiment: A reexamination of its implications for the principles of adaptive behavior. Psych. Rev., 78, 3-43.

Surowiecki, J., 2003, December 15. Jet blues. The New Yorker, 46.

Table 1

Effects of Reward Expectation on Bar-Press Duration

		Reward Expectation		
Expt	Figure(s)	More	Less	
1	1, 4, 5	early in trial	late in trial	
2	3, 6	during trial	between trials	
2	6	after rewarded response	after unrewarded response	
3	7, 9	high-food signal	low-food signal (25%)	
3	7, 9	low-food signal	intertrial interval	
3	8	long after response	soon after response	
4	10, 11	high-food signal	low-food signal	
4	10, 11	low-food signal	intertrial interval	
4	12	few prior unrewarded responses	many prior unrewarded responses	

Note. In each case, the condition with more reward expectation of reward produced longer (more variable) bar-press durations. The *t* values are for the comparison of the two conditions. They come from Gharib, Derby, and Roberts (2001) and Gharib, Gade, and Roberts (2004).

Figure Captions

Figure 1. Experiment 1: Response rate (upper panel) and response duration (lower panel) as a function of time since the start of the signal. Points in the duration function are unequally spaced along the time axis so that each point will represent roughly the same number of responses. Each point in the lower panel is based on about 4,000 (Replication 1) or 8,000 (Replication 2) responses/rat.

Figure 2. Experiment 1: Response durations before and after omission as a function of time from the start of the signal. *After omission* = second and later response after 40 s on empty trials. *Before omission* = all earlier responses on empty trials, and all responses on food trials. The points are spaced so that each point is based on about 12,000 responses.

Figure 3. Experiment 2: Response rate (upper panel) and response duration (lower panel) as a function of time before, during, and after the signal. Each point is a 10% trimmed mean over 18 rats. Data during the signal is represented by triangles.

Figure 4. Experiment 1: How food omission changed the distribution of response durations. Each point is a median over 36 rats. Standard errors were computed using the jackknife.

Figure 5. Experiment 1: How food omission changed the spread of the distribution of response durations. Each point is a 10% trimmed mean over 36 rats. The percentiles were computed separately for each rat, then averaged across rats.

Figure 6. Experiment 2: Distribution of response durations before and during the signal (upper panel) and 30-40 s after rewarded and unrewarded responses during the intertrial interval (lower panel). Each point is a median over 18 rats.

Figure 7. Experiment 3: Bar-press rate (upper panel) and duration (lower panel) as a function of day and signal. Method of computation: 1. The days of each phase were divided into blocks close to 5 days long. 2. For each rat-signal-block combination, the mean of the logarithm of all rates (upper panel) or durations (lower panel) was computed. 3. For each signal-block combination, a 10% trimmed mean over rats was computed.

Figure 8. Experiment 3: Effect of interresponse time. Top panel: Duration as a function of interresponse time. Middle panel: Probability of reward as a function of interresponse time. Bottom panel: Duration as a function of probability of reward. Duration and interresponse times are on a log scale, probability of reward on a square-root scale. The 100%/50% data are from the last 15 days of that phase; the 100%/25% data are from the last 20 days of that phase. Method of computation: 1. For each rat-signal combination (14 rats, 4 signals), all interresponse times were gathered, ranked, and divided into 8 equal-sized bins by rank, a total of 448 bins. (To show as clearly as possible the correlation between interresponse time and probability of reward, the first bar press during a signal was given an interresponse time equal to the time since the signal began.) One group contained all interresponse times between the minimum and the 12.5% quantile; the next, all interresponse times between the 12.5% quantile and

the 25% quantile, and so forth. 2. For each group, the mean of the log interresponse times was computed (448 means, 32 per rat). 3. The 10% trimmed mean over rats was computed to summarize each bin (32 trimmed means). 4. For each rat-signal-bin combination, the mean log duration was computed, then averaged over rats to get 32 values. 5. For each rat-signal-bin combination, the probability of reward was computed, then averaged over rats to get 32 values. 6. Responses during intertrial intervals were not divided by interresponse time because there were too few of them. The average duration for interresponse time was computed by taking the mean within rats and a trimmed mean over rats.

Figure 9. Experiment 3: 10% and 90% quantiles of duration distributions as a function of probability of reward. These values were computed in the same way as the values in the bottom panel of Figure 8, except that each set of durations was summarized by the 10% and 90% quantiles rather than by the mean.

Figure 10. Experiment 4: Bar-press rate (upper panel) and duration (lower panel) as a function of day and signal. Same method of computation as Figure 7.

Figure 11. Experiment 4: Distribution of bar-press durations as a function of signal. Upper panel: Probability density functions. Method of computation: 1. The minimum and maximum duration over all rats and signals were used to determine an interval. 2. The interval was divided into ten (high-food function) or eight (low-food and intertrial functions) segments of equal width on a log scale. For each function, the number of segments was the maximum that would not produce any zero probabilities in the final result. A probability of zero is hard to

show on a log scale. 3. These bin definitions were used to determine frequencies for each rat. 4. Frequencies were converted to probabilities by dividing by sample size. 5. Probabilities were converted to probability densities by dividing by bin width on the untransformed scale (seconds, not log seconds). 6. For each bin, a median was computed over rats. Lower panel: Inverse cumulative distribution functions. Method of computation: 1. The function was computed for each rat separately. 2. For each abscissa value, a median was computed over rats.

Figure 12. Experiment 4: Effect of recent consecutive unrewarded bar-presses during the low-food signal on bar-press rate (upper panel) and duration (lower panel). Every trial was classified according to the number of consecutive unrewarded bar-presses during the low-food signals that immediately preceded it. Values on the abscissa are plotted according to log(1 + *x*), where *x* is the number of unrewarded bar presses. For example, the value for 0 is plotted above log 1, the value for 10 above log 11, and so on. Values within a bin are assigned the geometric mean of the edges of the bin. For example, the right-most bin contains data from trials preceded by 31 to 140 unrewarded bar presses during the low-food signal. The geometric mean of 31 and 140 is 66, which is plotted above log 67.

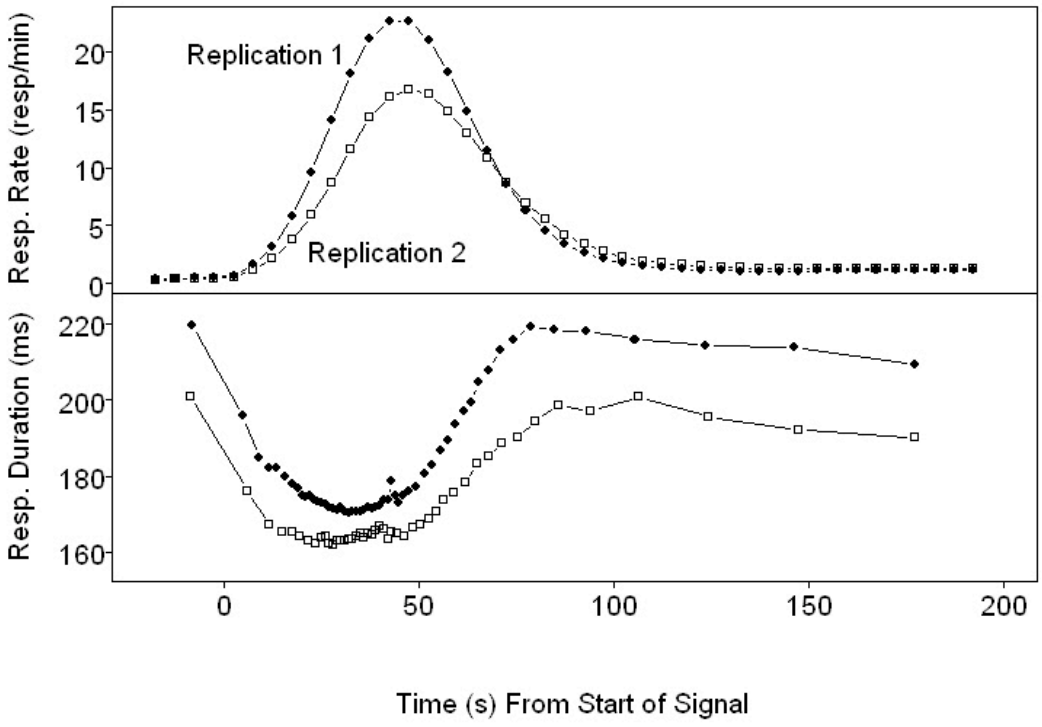

Resp. Rate (resp/min)
Replication 1
Replication 2
20
15
10
5
0
Resp. Duration (ms)
220
200
180
160
0
50
100
150
200
Time (s) From Start of Signal

Figure 1

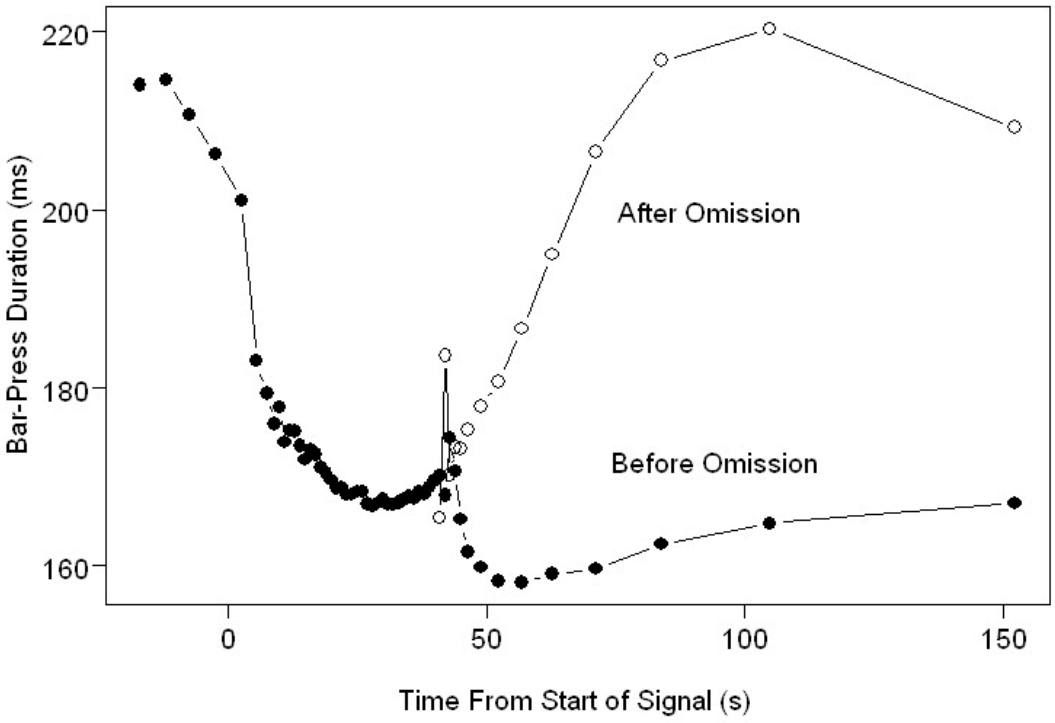

220
200
180
160
Bar-Press Duration (ms)
After Omission
Before Omission
0
50
100
150
Time From Start of Signal (s)

Figure 2

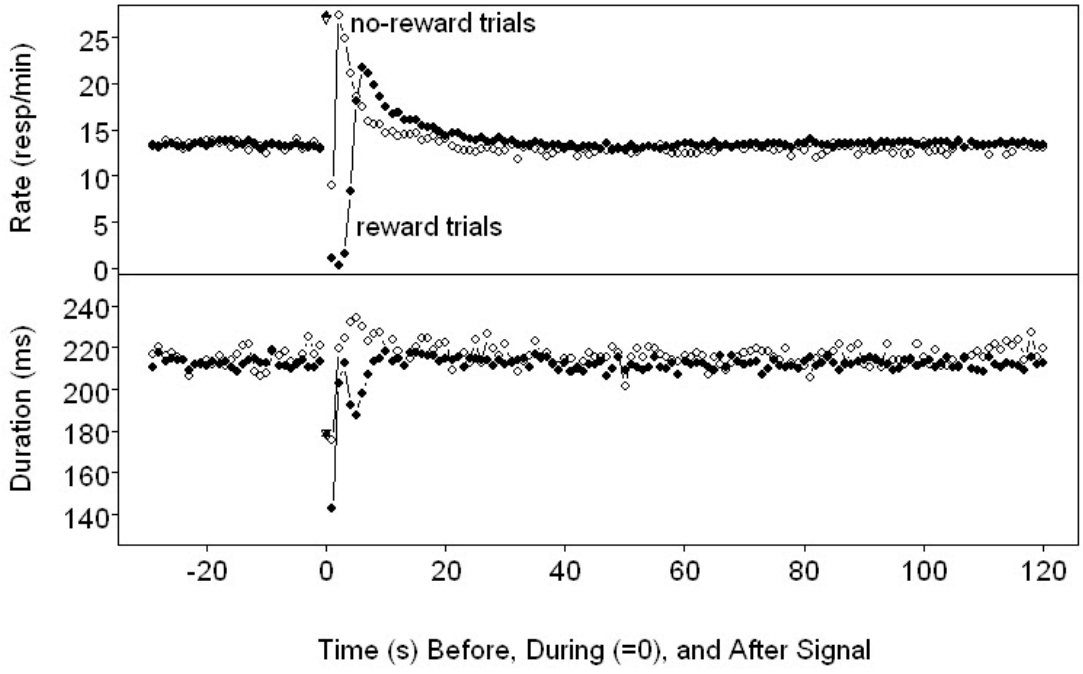

no-reward trials
reward trials
Rate (resp/min)
25
20
15
10
5
0
Duration (ms)
240
220
200
180
160
140
-20
0
20
40
60
80
100
120
Time (s) Before, During (=0), and After Signal

Figure 3

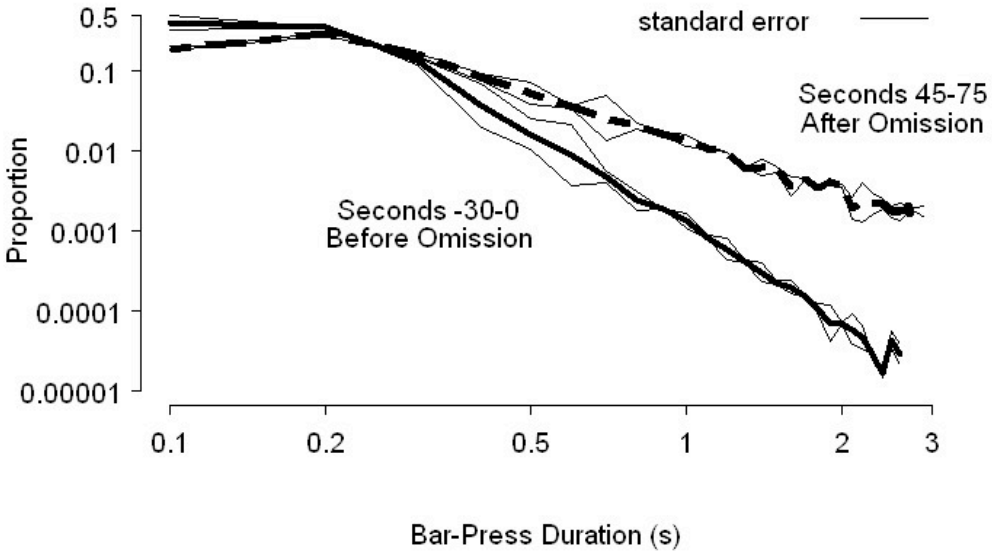

0.5
0.1
0.01
0.001
0.0001
0.00001
Proportion
standard error
Seconds 45-75
After Omission
Seconds -30-0
Before Omission
0.1
0.2
0.5
1
2
3
Bar-Press Duration (s)

Figure 4

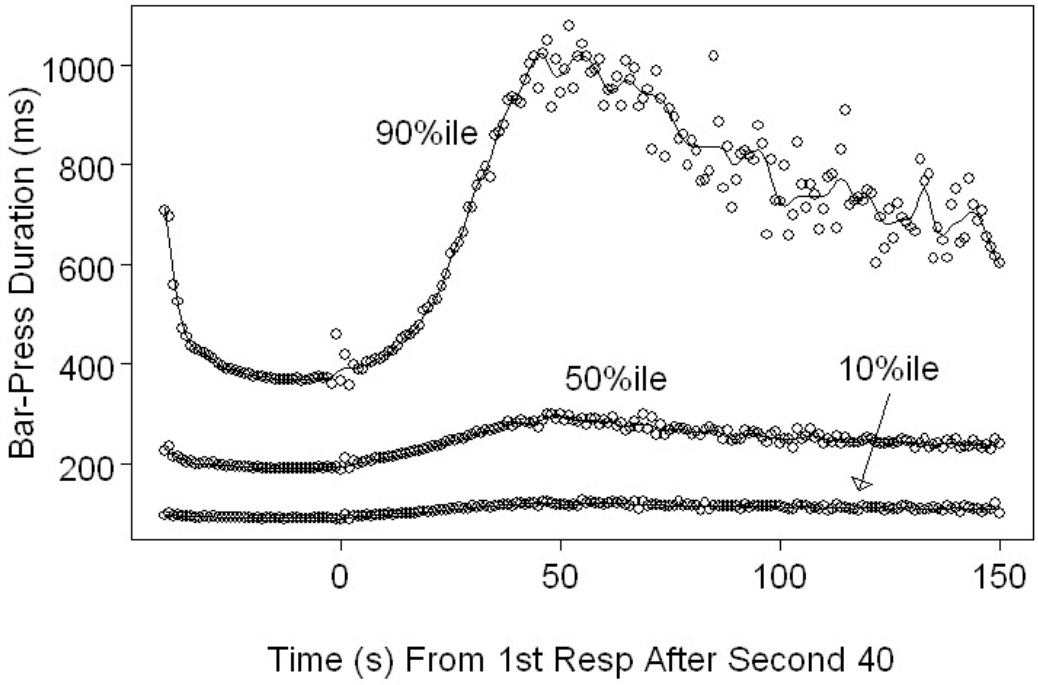

Bar-Press Duration (ms)
1000
800
600
400
200
90%ile
50%ile
10%ile
0
50
100
150
Time (s) From 1st Resp After Second 40

Figure 5

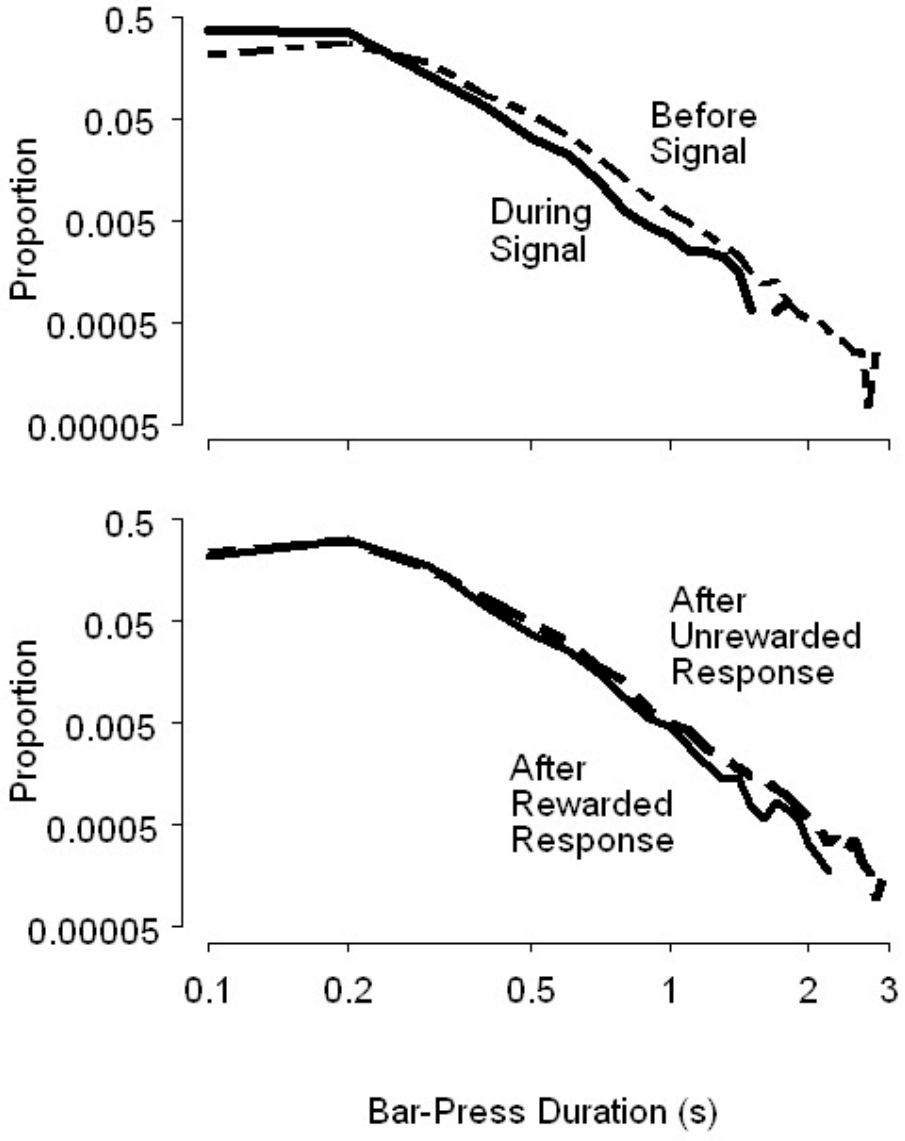

0.5
0.05
0.005
0.0005
0.00005
Proportion
Before Signal
During Signal
After Unrewarded Response
After Rewarded Response
0.1
0.2
0.5
1
2
3
Bar-Press Duration (s)

Figure 6

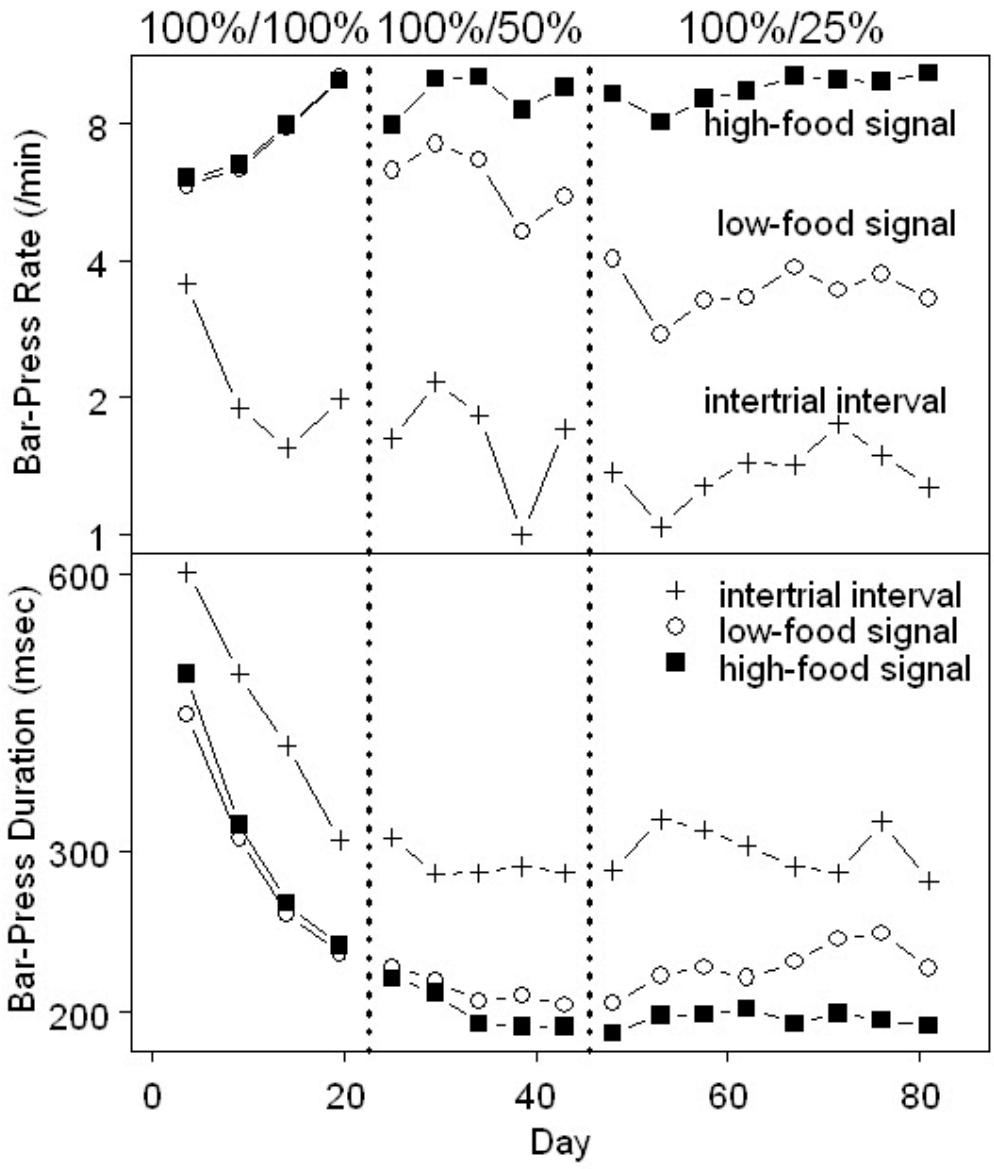

100%/100%
100%/50%
100%/25%
Bar-Press Rate (/min)
8
4
2
1
high-food signal
low-food signal
intertrial interval
Bar-Press Duration (msec)
600
300
200
intertrial interval
low-food signal
high-food signal
0
20
40
60
80
Day

Figure 7

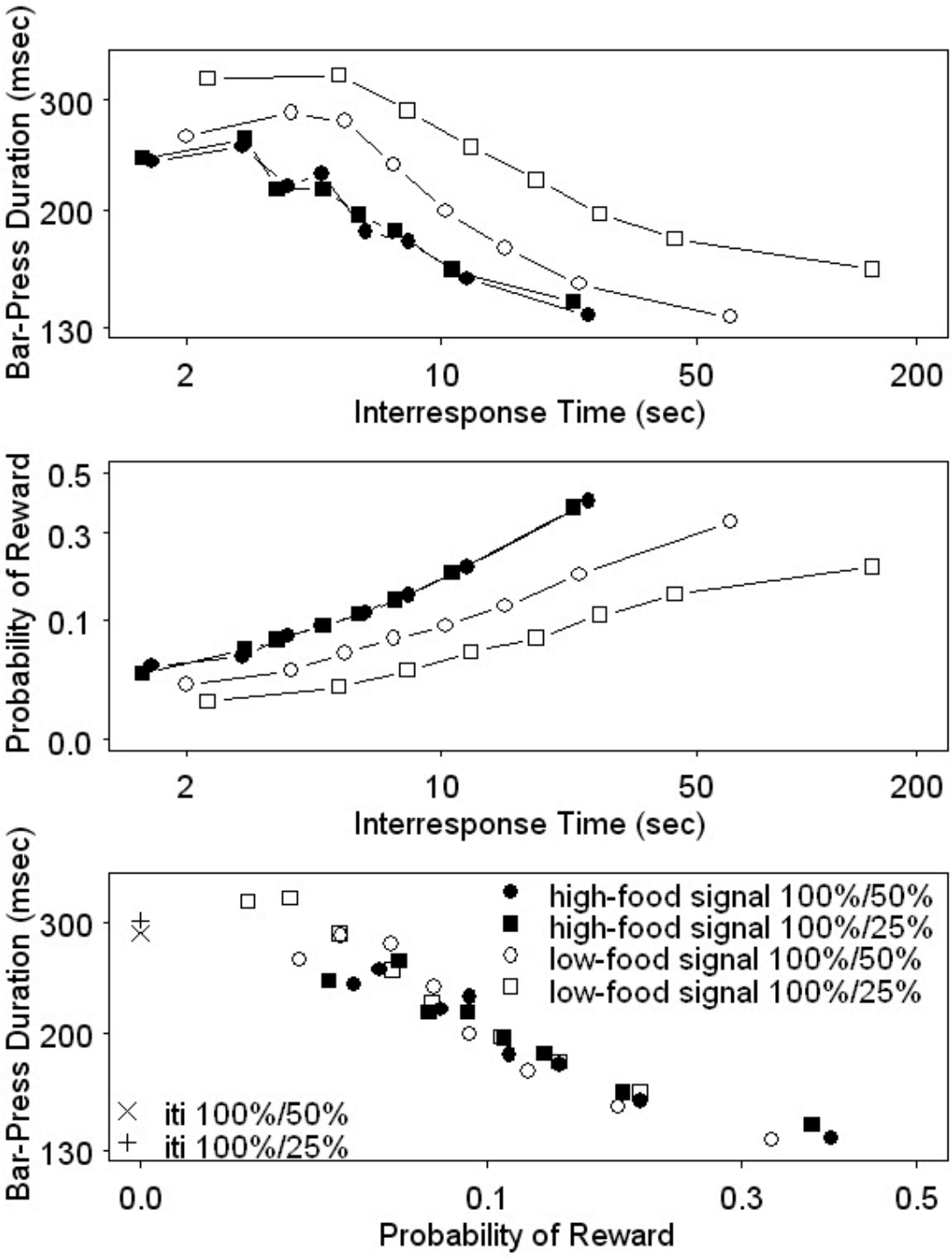

Bar-Press Duration (msec)
300
200
130
2
10
50
200
Interresponse Time (sec)
Probability of Reward
0.5
0.3
0.1
0.0
2
10
50
200
Interresponse Time (sec)
Bar-Press Duration (msec)
300
200
130
high-food signal 100%/50%
high-food signal 100%/25%
low-food signal 100%/50%
low-food signal 100%/25%
iti 100%/50%
iti 100%/25%
0.0
0.1
0.3
0.5
Probability of Reward

Figure 8

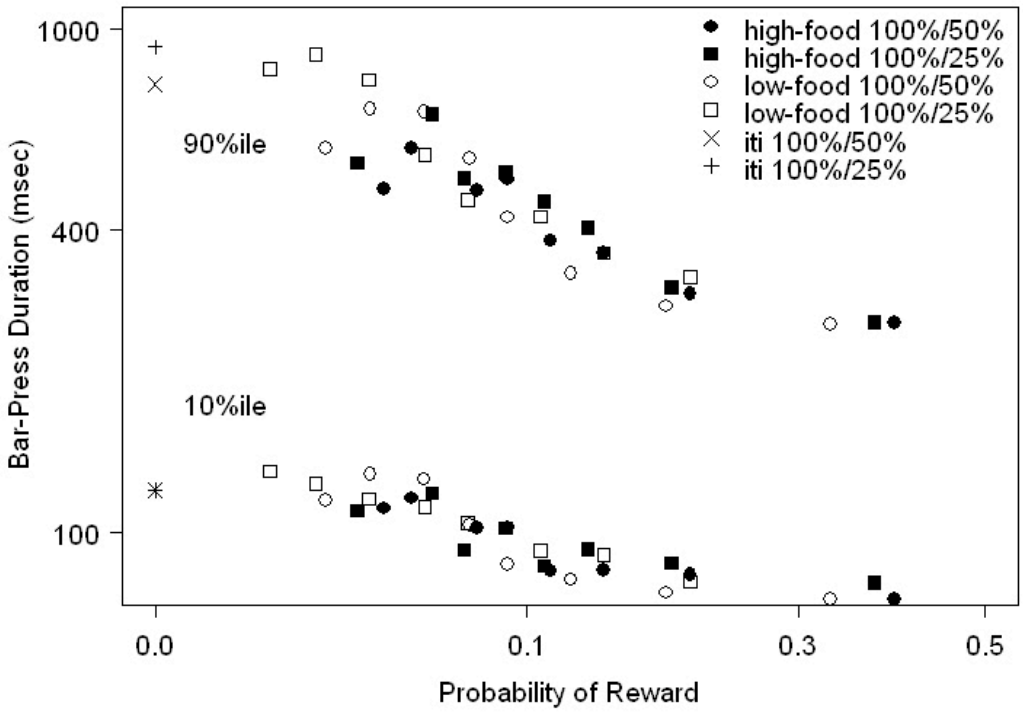

1000
400
100
Bar-Press Duration (msec)
90%ile
10%ile
high-food 100%/50%
high-food 100%/25%
low-food 100%/50%
low-food 100%/25%
iti 100%/50%
iti 100%/25%
0.0
0.1
0.3
0.5
Probability of Reward

Figure 9

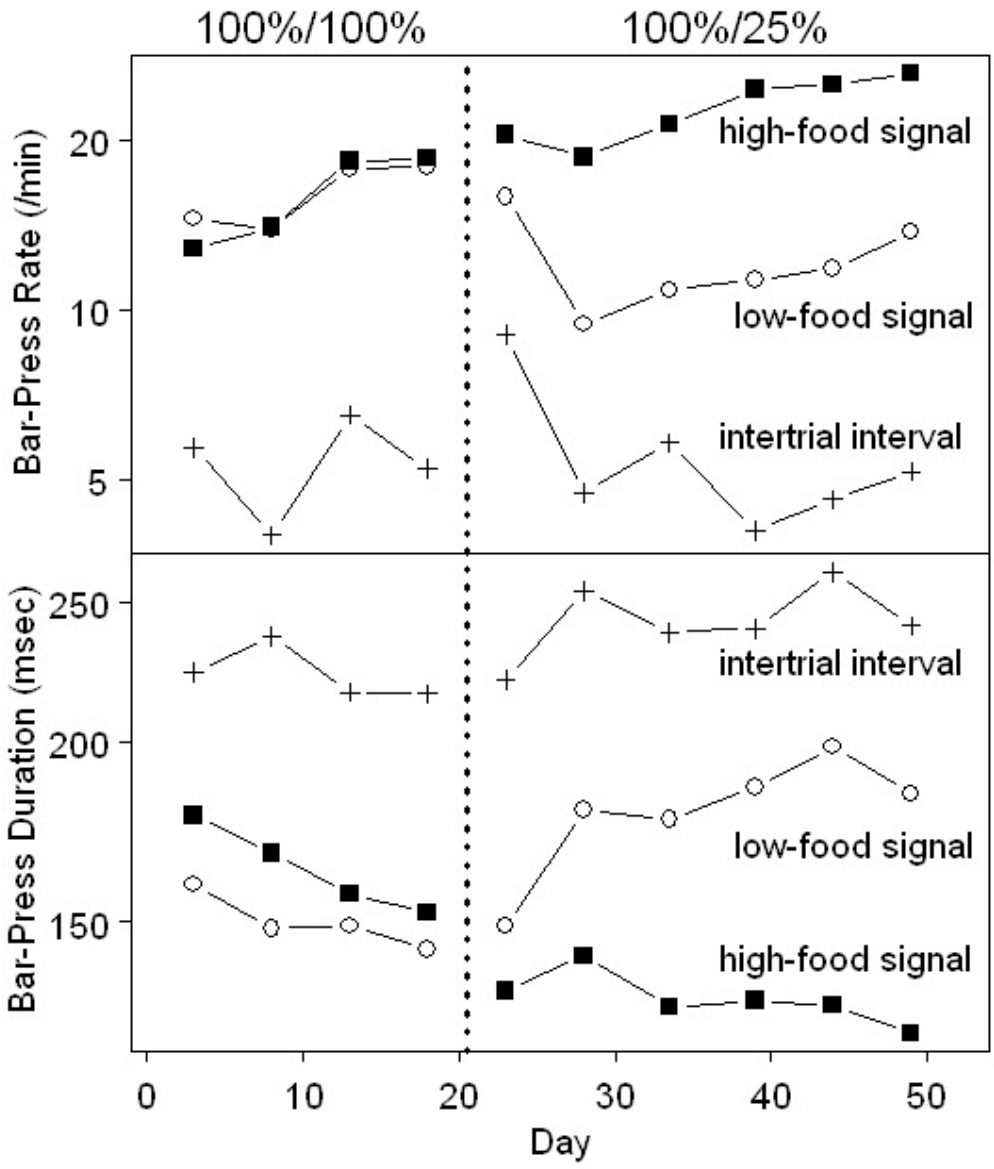
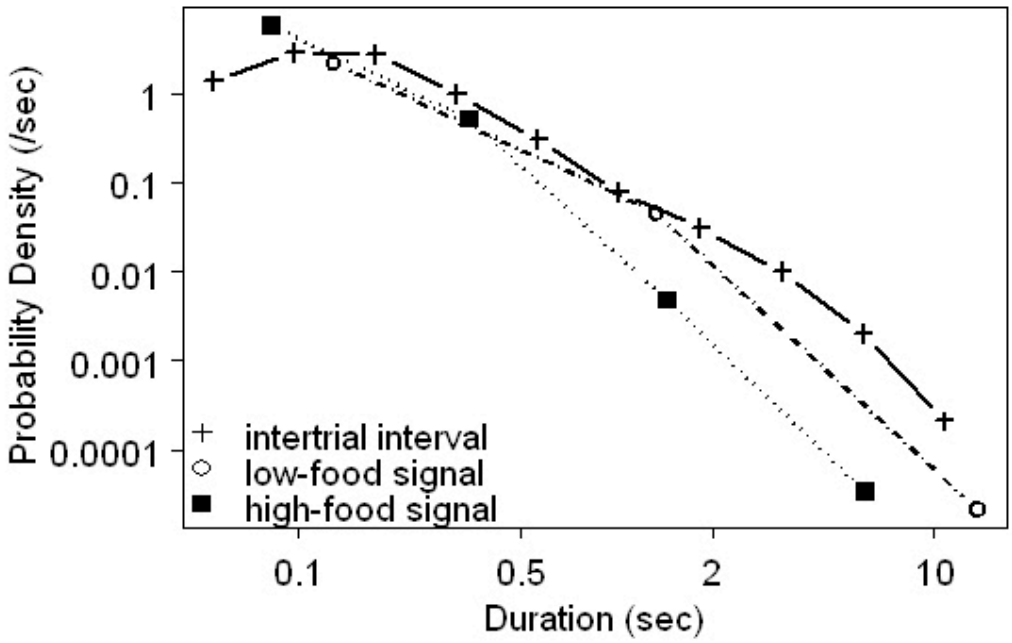

100%/100%
100%/25%
Bar-Press Rate (/min)
20
10
5
high-food signal
low-food signal
intertrial interval
Bar-Press Duration (msec)
250
200
150
intertrial interval
low-food signal
high-food signal
0
10
20
30
40
50
Day

Figure 10


Probability Density (/sec)
1
0.1
0.01
0.001
0.0001
intertrial interval
low-food signal
high-food signal
0.1
0.5
2
10
Duration (sec)

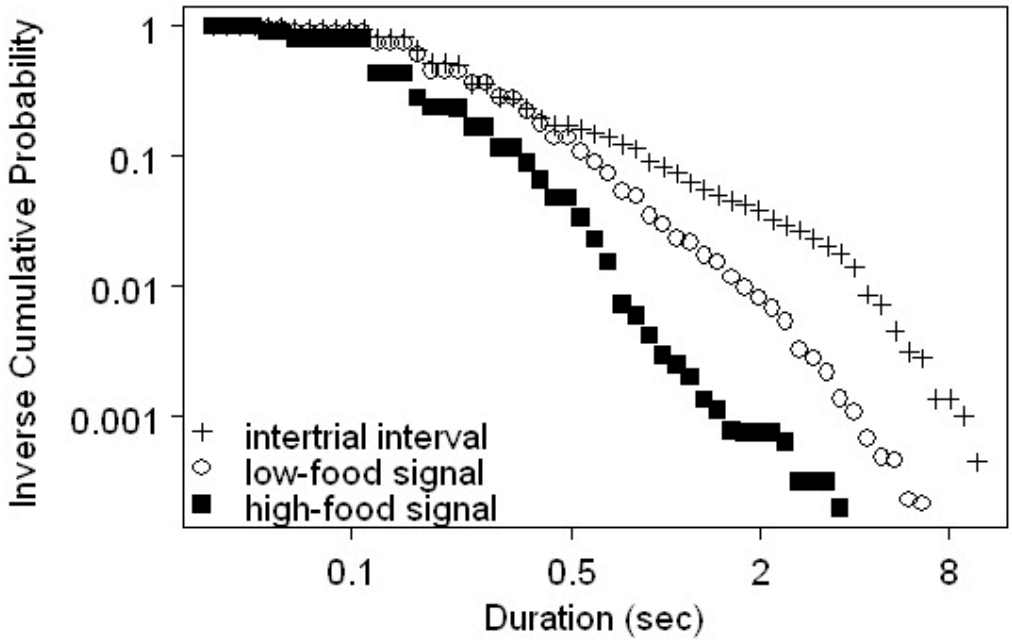

Inverse Cumulative Probability
1
0.1
0.01
0.001
intertrial interval
low-food signal
high-food signal
0.1
0.5
2
8
Duration (sec)

Figure 11

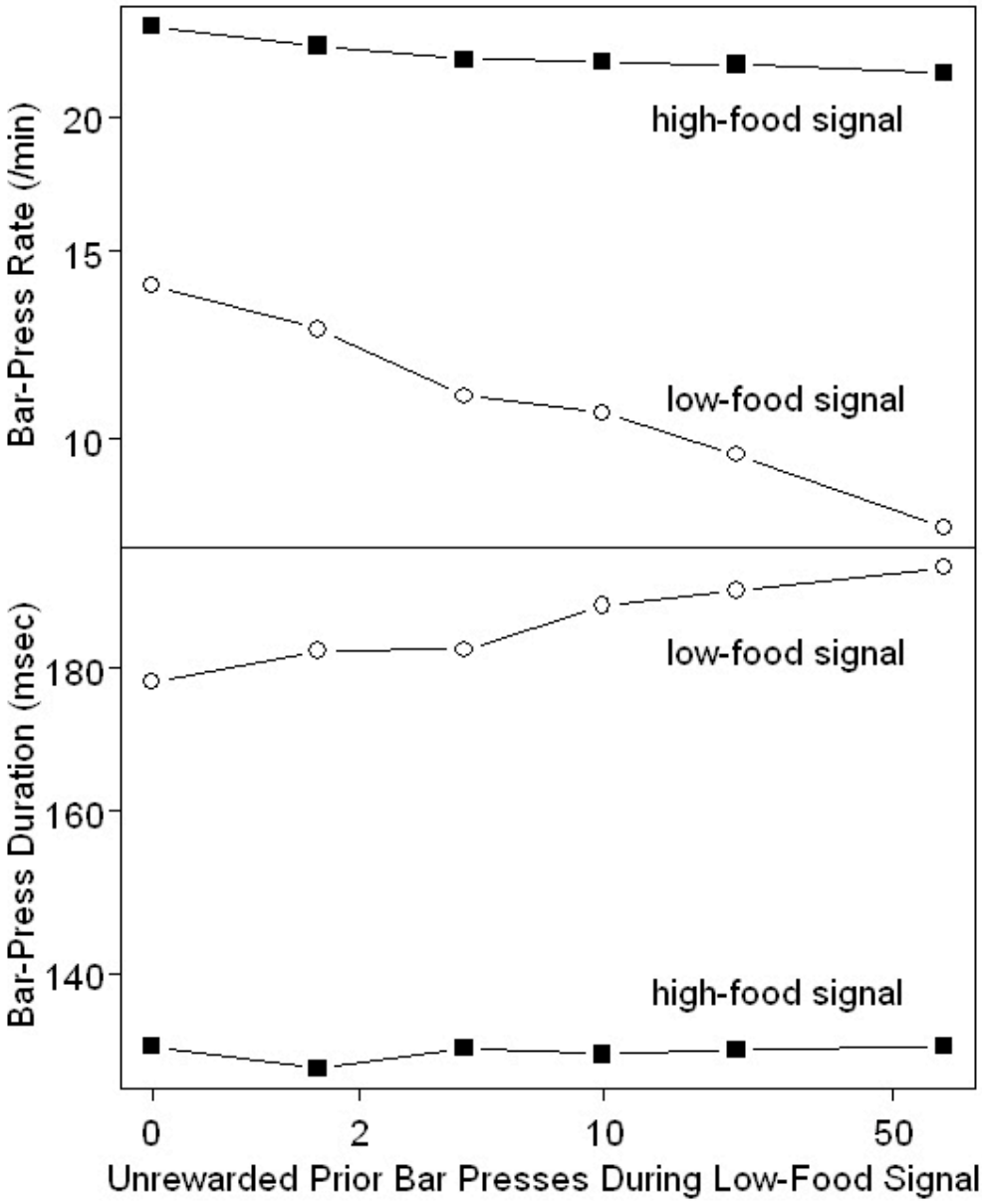

Bar-Press Rate (/min)
20
15
10
high-food signal
low-food signal
Bar-Press Duration (msec)
180
160
140
low-food signal
high-food signal
0
2
10
50
Unrewarded Prior Bar Presses During Low-Food Signal

Figure 12