



Modulation of Variation by Response-Reward Spatial Proximity

Kenneth J. Leising & Chad M. Ruprecht

Texas Christian University, U.S.A.

W. David Stahlman

University of Mary Washington, U.S.A.

There has been a recent surge in the experimental investigation of the control of behavioral variability. Currently, it is understood that variability in behavior is predictably modulated by reinforcement parameters (e.g., a probability of reward delivery and reward magnitude). In two experiments, we investigated how spatial proximity between response and reward locations impacts the production of behavioral variability in both response rate and lever press duration. Rats were trained to lever press on two levers in a standard operant chamber that only differed from one another in their proximity to a food niche (i.e., Near vs. Far); a second experimental factor, the probability of reward, was signaled by an auditory cue. In Experiment 1, trials with a high-probability stimulus terminated with reward on 100% of trials, while trials with a low-probability stimulus terminated with reward 25% of the time. We used a similar procedure in Experiment 2, but reduced the likelihood of reward on low-probability trials to 10% and collected data in a post-acquisition extinction test. Overall, proximity and probability were inversely related to variation of response rate, whereas only the probability factor affected variation in lever press duration. Proximity also interacted with probability to influence variation in response rate. These findings extend the factors modulating behavioral variability to include the spatial proximity between a response and reward.

The degree of behavioral variability during periods in which an expected reward was omitted (i.e., an extinction trial) has been the focus of much previous research (e.g., Antonitis, 1951; Balsam, Deich, Ohyama, & Stokes, 1998; Devenport, 1984; Eckerman & Lanson, 1969; Herrick & Bromberger, 1965; Morgan & Lee, 1996; Millenson & Herwitz, 1961; Neuringer, Kornell, & Olufs, 2001; Notterman, 1959; Stebbins & Lanson, 1962). During extinction trials, response variability (e.g., in the rate of lever pressing) has been shown to be higher relative to that observed during acquisition. However, response rate also decreases during extinction trials, thus limiting the amount of data that can be collected without retraining.

Many studies of reward-expectation induced variability have utilized steady-state procedures as alternatives to extinction. During steady-state procedures, response rates are stable and variability can be measured across many sessions of training (e.g., Stahlman, Roberts, & Blaisdell, 2010a). For example, Gharib, Gade, and Roberts (2004) trained rats that two signals, high-probability and low-probability, ended with food with a probability of 100% or 25%, respectively. Over 31 sessions of training, rats emitted hundreds of responses each day. Interresponse times (IRTs) and lever press durations (DURs) during the low-probability signal demonstrated greater variability as compared to the high-probability signal. Because the delivery of reward during steady-state instrumental procedures is contingent solely upon a lever press, additional dimensions (e.g., IRT or Duration) of the response (related to the expected density of reward) are free to vary. These results supported a rule proposed by Gharib et al. (2001): a reduction in the expectation of reward increases variation of behavioral form (see also Roberts & Gharib, 2006).

More recent results have supported this rule. Stahlman and Blaisdell (2011b) manipulated both the magnitude of the reward and the temporal delay to reward with pigeons pecking to a visual stimulus displayed on a touchscreen-equipped monitor. Presenting a reward of lower magnitude instantiated higher amounts of behavioral variability (e.g., standard deviation of IRTs and standard deviation of X, Y coordinates); similarly, the insertion of a delay between pecking and the delivery of the reward increased variability of behavior. These results suggest that probability and magnitude of reward, as well as the temporal proximity of the response to reward modulated the pigeons' expectation of reward, and that lower expectations of reward led to greater variation in behavior (Stahlman & Blaisdell, 2011b). Previous research investigating the production of temporal and spatial behavior has found similar psychophysical principles governing each (see Cheng, 1992 for a review). These findings raise the question of whether manipulations of the spatial proximity between response and reward would impact variation in the same manner as temporal proximity.

Pecoraro, Timberlake, and Tinsley (1999) reported increases in the variability of sampling behavior within a plus-maze associated with a decrease in reward quality, which is another factor that modulates reward expectation. On an elevated plus maze, rats were trained to seek out a beaker of 32% sucrose solution always located on one arm of the maze. Rats in a downshift-group (32%/4%) then received trials in which the sweet solution was switched to a considerably less sweet sucrose solution; importantly, these rats decreased consummatory behavior, and increased their nose-down locomotion, location entries, and sampling of alternatives in comparison to an unshifted (4%/4%) group. Thus, in the spatial domain, drops in reward quality spur variable behavior (see Wiegmann & Smith, 2009, for a similar demonstration in honeybees).

More relevant to signaled manipulations of reward probability, Stahlman and Blaisdell (2011a) trained rats to use landmarks to search for food buried within a 4 x 4 array of response locations. On separate trials, the two landmarks were paired with a hidden food reward with similar spatial relations, but they differed in the probability (HI = 100%, LO = 20%) that reward would be found at the expected reward location. Foraging variability was considerably higher during trials with the LO landmark, as measured by the standard deviation of search frequencies. These results suggest that spatial variability in foraging is also modulated by reward probability.

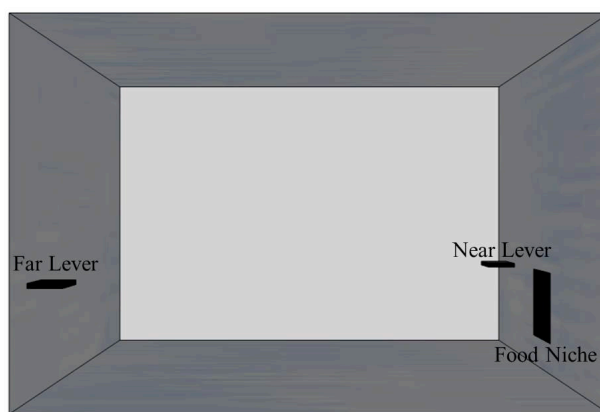


Figure 1. A diagram of the interior of the experimental apparatus. The food niche is on the right-hand side of the image in the center of the chamber. The near lever is adjacent to the food niche. The far lever is positioned on the opposite wall of the chamber from the food niche. Both levers are shown in the active position (i.e., inserted), but during experimental trials only one lever was active. This image is drawn to scale.

One factor that has not yet been studied is how the spatial proximity of the response to the reward influences the production of variability. Classic models of response learning state that responses, such as pressing a lever for food, strengthen as a function of their proximity (spatial and/or temporal) to reward (Hull,

1943; Silva, Timberlake, & Gont, 1998; Staddon & Horner, 1989). According to classic optimal foraging theory (MacArthur & Pianka, 1966) successful foragers are sensitive to shifting environmental contingencies. In general, animals are required to behave flexibly, varying their responses to accommodate changes in reward probability, as well as conservatively, minimizing travel time and energy expenditure by favoring proximal response locations over distal response locations (Kacelnik & Todd, 1992). However, it is unclear how spatial proximity will influence reward expectation, and consequently, response variation. Therefore, the current research investigated whether the spatial proximity of the response location (a lever) to the food niche would influence temporal variation in the behavior of rats. We also sought to examine how reward probability and spatial proximity interact to modulate reward expectation.

In two experiments, rats were trained to lever press on two levers in a standard operant chamber; one lever was positioned adjacent to the food niche (i.e., Near), while the other was placed on the opposite side of the chamber (i.e., Far, see Figure 1). During discrimination training, rats were given separate trials of the Near and Far lever paired with one auditory cue associated with a High (100%) probability of reward and the other a Low (i.e., 25% in Experiment 1 and 10% in Experiment 2) probability. In Experiment 2, all subjects were also tested during post-acquisition extinction sessions. We expected to replicate prior findings of higher variability with decreased reward expectation, manifest as less temporal variation in lever press IRTs and DURs on the Near lever compared to the Far lever. Furthermore, based on the results of Stahlman and Blaisdell (2011b), we predicted that reward probability and spatial proximity would interact to control variability. We conducted the procedures and analyses below to conform closely to prior published work (e.g., Gharib et al., 2004).

Experiment 1

Method

Subjects. The subjects were 12 female hooded Long-Evans rats bred in the TCU vivarium from parents obtained from Harlan Laboratories (Indianapolis, IN). Subjects were pair-housed in translucent plastic tubs with a substrate of wood shavings in a vivarium maintained on a 12 hr dark/12 hr light cycle. All experimental manipulations were conducted during the light portion of the cycle. A progressive food restriction schedule was imposed over the week prior to the beginning of the experiment, until each rat received 15 g of food each day. All animals were maintained at 85% of their free-feeding weight. The rats weighed, on average, 270 grams and were six months old at the start of the experiment. All animals were handled daily for 30 s, during the week prior to the initiation of the study.

Apparatus. Figure 1 displays the experimental chamber. Each of six experimental chambers measuring 30 cm x 25 cm x 20 cm (l x w x h) was housed in a sound and light-attenuating environmental isolation chest (Med Associates). The walls and ceiling of the chamber were constructed of clear Plexiglas and the floor was constructed of stainless-steel rods measuring 0.5 cm in diameter, spaced 1.5 cm center-to-center. One wall of the chamber was equipped with a dipper that delivered a 16% sucrose solution. When in the raised position, a small well (0.05 cc) at the end of the dipper arm protruded up into the food niche. An infrared photo-detector was positioned across the entrance to the feeding niche. The near lever was positioned on the same wall as the food niche, at a distance of 12 cm center-to-center. The far lever was positioned on the wall opposite the food niche, at a distance of 25 cm center-to-center.

A ventilation fan in the enclosure and a white-noise generator on a shelf outside of the enclosure provided a constant 74-dB (A) background noise. One speaker on the outside walls of the chamber could deliver a high-frequency tone (3000 Hz) 8 dB(A) above background and the other a white noise (N) stimulus 8 dB (A) above background. The tone and white noise served as discriminative stimuli and were counterbalanced across groups. A diffuse houselight was located 13 cm above the floor on the wall opposite the food magazine.

Procedure.

Pretraining. Pretraining consisted of 10 sessions. All rats were initially trained to drink sucrose from the feeding niche. Once available, a disruption in the infrared beam in the feeding niche would initiate a 10-s access period before the lever retracted. This was also true for non-differential and discrimination training. Rats were then trained to respond on the levers in the near and far positions. On each trial, one of the two levers was randomly selected to be inserted into the chamber. Initially, responses on the lever were reinforced according to a continuous reinforcement schedule, followed by variable-ratio 2 (VR-2), VR-4, and then random-ratio 4

(RR-4) schedules of reinforcement. All trials were separated by a 60 s inter-trial interval (ITI). All sessions were 90 minutes in length. Rats were advanced together to the next reinforcement schedule only after all of them lever pressed at least 50 times in a session.

Non-differential training. All rats were given separate trial types of two stimuli (tone and noise), each paired equally often with the near and far levers (i.e., Near-Tone, Far-Tone, Near-Noise, Far-Noise). Trial type was chosen randomly without replacement, such that the number of trials per trial type would be roughly equal within each session. Each trial began with simultaneous insertion of a lever and onset of a discriminative stimulus. All trials terminated according to a RR-4 schedule and the probability of sucrose reward following trial termination was 100%. All trials were separated by a 60 s ITI and sessions were 90 minutes in length. Non-differential training consisted of 4 sessions. One rat failed to show reliable bar pressing after 4 additional sessions and so training for that rat was discontinued.

Discrimination training. In this phase, the probability of reward at a trial's termination for one of the two auditory stimuli was lowered from 100% to 25% for all subjects. Which stimulus was paired with the reduction in reward probability was counterbalanced across subjects. Each stimulus was presented equally often with the near or far lever and reinforced with the assigned probability irrespective of which lever was present. All trials were separated by a 60 s ITI. The manipulation of probability and proximity resulted in four within-subject trial types – High-Near, High-Far, Low-Near, and Low-Far. Trial type was chosen randomly without replacement in blocks with all four trials represented in each block. All sessions were 90 minutes in length. Discrimination training consisted of 30 sessions.

Measures.

Individual durations were measured to an accuracy of 20 msec. The duration from the onset of the discriminative stimulus until the first response (pre-response time, PRT) and subsequent interresponses times (IRTs) were recorded. The duration (DUR) of each lever press was also recorded. Because IRT and DUR were strongly positively skewed, we performed a log transform of the data to normalize it prior to statistical analysis (cf. Gharib et al., 2004). We then calculated the standard deviation of log IRT and log DUR within session as our measures of temporal variation. This is a common transformation of these kinds of data and has been employed in a number of recent reports (e.g., Stahlman & Blaisdell, 2011b).

PRTs and IRTs have been shown to exhibit different patterns of variability (Gharib, Gade, & Roberts, 2004; Stahlman et al., 2010a; Stahlman et al., 2010b). Differences between PRTs and IRTs would likely result in differences between DURs of the first vs. subsequent lever press as well. Consequently, we analyzed the PRTs and DURs from first responses separately from the IRTs and DURs from subsequent responses.

The number of discrimination training sessions and the probability of reward for the low and high-probability stimuli were chosen to match those reported in Gharib et al. (2004). As in previous reports (Gharib et al., 2004; Stahlman et al., 2010a), we were interested in the effects of probability and contiguity following the acquisition of steady-state responding. Consequently, we analyzed the last 20 sessions of discrimination training.

Results and Discussion

One subject failed to reach the requisite number of lever presses within four additional sessions of non-differential training, thus analyses were conducted with the remaining subjects ($n = 11$). The standard deviations of the first responses (PRTs and DURs) are reported followed by the same measures (IRT and DUR) collected from subsequent responses. All of the statistical results from Experiments 1 and 2 can be found in Table 1. Figure 2 shows the standard deviation of PRTs (left panel) and DURs (right panel) as a function of probability (High vs. Low) and proximity (Near vs. Far) during the last 20 sessions of training. Though the differences appear small, Figure 2 indicates that both low-probability trials (Near-Low and Far-Low) produced higher levels of variation in PRT than those with the high-probability stimulus. This observation was supported by a repeated-measures ANOVA conducted on PRTs with Probability (High vs. Low) and Proximity (Near vs. Far) as factors, which found a main effect of Probability, $F(1,10) = 16.86$, $p = 0.002$, and a marginally non-significant interaction of Proximity x Probability, $F(1,10) = 4.51$, $p = 0.06$. To determine the source of the near-interaction, we conducted separate post-hoc Scheffe tests. These tests revealed greater temporal variation during Far-Low than either trial type with the high-probability stimulus (Far-High and Near-High), $ps < .04$, but no reliable difference between the two low-probability trial types (see Figure 2), $p = 0.86$. Near-Low also produced higher levels of variation than Far-High, $p = 0.005$. For the

DUR of first responses (see Figure 2, right-panel), we conducted the same analysis but found no reliable main effects or the interaction, $F_s(1,10) < 3.45$, $p_s > 0.09$.

The data from subsequent responses (IRT and DUR) are also shown in Figure 2. Small differences in the low- vs. high-probability trial types are noticeable in IRT and DUR, but particularly in the DUR measure. These observations were again supported by separate repeated measures ANOVAs with Probability (High vs. Low) and Proximity (Near vs. Far) as factors revealing a main effect of Probability in both IRTs, $F(1,10) = 5.17$, $p = 0.05$, and contrary to the results with first responses, the DUR of subsequent responses, $F(1,10) = 12.33$, $p = 0.006$.

The results of the last 20 sessions of discrimination training provide evidence of increased variation of PRTs/IRTs with a lower probability of reward. Experiment 1 also provided evidence for a marginally-nonsignificant interaction of probability and proximity, but only on PRTs. With respect to temporal variation in DURs, the standard deviations provided evidence only for increased variation with a lower probability of reward.

Table 1
F values and effect sizes from ANOVAs in Experiments 1 and 2

	IRT				DUR			
	First		Subsequent		First		Subsequent	
	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2	<i>F</i>	η_p^2
Experiment 1								
Proximity (Prox)	.37	.04	.21	.02	3.49	.26	3.41	.25
Probability (Prob)	16.86**	.63	5.17*	.34	3.45	.26	12.33**	.55
Prox * Prob	4.51	.31	1.28	.11	.01	.01	.06	.01
Experiment 2								
<i>Post-Acquisition</i>								
Prox	12.55**	.56	3.97	.38	.01	.01	1.46	.13
Prob	18.08**	.64	6.11*	.38	.87	.08	8.80*	.47
Prox * Prob	15.62**	.61	6.17*	.28	.10	.01	.01	.01
<i>Post-Acq vs. Extinction</i>								
Prox	28.20***	.74	11.35**	.53	2.34	.19	3.24	.24
Prob	20.67**	.67	11.42**	.53	.03	.01	7.37*	.42
Prox * Prob	1.58	.14	4.02	.29	.29	.03	.02	.01
Treatment	16.27**	.62	16.86**	.63	27.51***	.73	9.38**	.48
Treatment *Prox	.01	.14	3.54	.26	.55	.05	1.57	.14
Treatment *Prob	6.31*	.39	11.59**	.54	.01	.01	1.23	.11
Treatment *Prox*Prob	3.12	.24	.35	.03	.77	.07	.85	.08

Note. Exact *p*-values are provided in the text. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Experiment 1 failed to provide evidence for the proximity of the response-reward spatial relationship to increase the standard deviation of response rate (i.e., IRTs) and lever press DURs. Furthermore, we report only weak evidence of an interaction between proximity and probability. In Experiment 2, we decreased the probability of reward associated with the low-probability stimulus, extended the number of sessions in both non-differential and discrimination training, and included post-acquisition extinction sessions. The first two manipulations have been employed before in the investigation of reward probability and its relationship to variability in behavior (e.g., Gharib et al., 2004). Extending the number of sessions allowed us to analyze temporal variation during a period of time following a greater number of discrimination training sessions. Lastly, we included extinction training to detect how both factors (i.e., proximity and probability) would influence temporal variation in the absence of reward.

Experiment 2

Method

Subjects. The subjects were 12 experimentally-naïve female hooded Long-Evans rats bred in the TCU vivarium from parents obtained from Harlan Laboratories (Indianapolis, IN). All subjects were housed and handled as in Experiment 1. The rats weighed, on average, 265 grams and were approximately four months old at the start of the experiment.

Apparatus. The apparatus was the same as that used in Experiment 1.

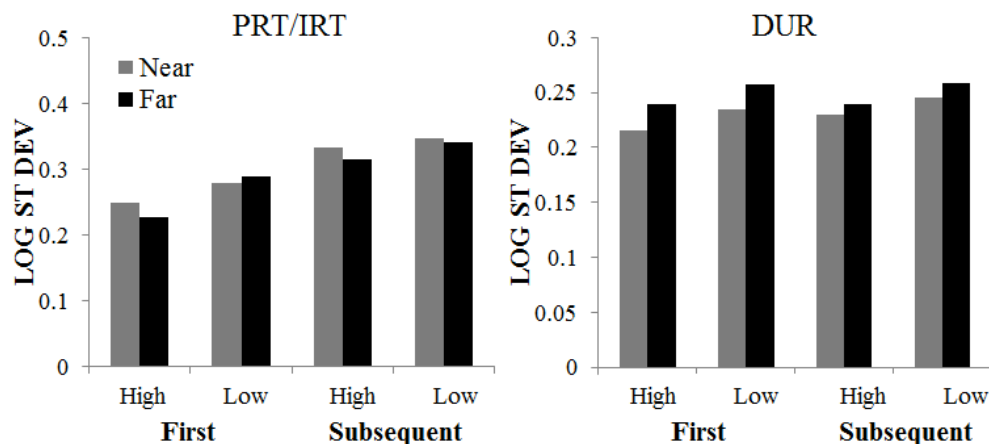


Figure 2. Left panel: The standard deviation of log PRTs and log IRTs from the last 20 sessions of discrimination training in Experiment 1. Right panel: Log DURs from the same period of training. The data from the first and subsequent responses on a trial are displayed separately for each trial type. *Far* and *Near* refer to the location of the lever relative to the food niche, and *High* (100%) and *Low* (25%) refer to the probability that a trial will terminate in reward. All units are measured in seconds.

Procedure. Pretraining and non-differential training were the same as in Experiment 1, with the exception that non-differential training in this experiment continued for 15 sessions.

Discrimination training. Discrimination training was similar to that of Experiment 1, but the probability of reward at a trial's termination for the low-probability auditory stimulus was reduced to 10% and discrimination training continued for 45 sessions. Each stimulus was again presented equally often with the near or far lever and reinforced with the assigned probability irrespective of which lever was present. The manipulation of probability and proximity resulted in four within-subject trial types – High-Near, High-Far, Low-Near, and Low-Far.

Extinction. In this phase, all sessions and trials were identical to that of Discrimination training, except that no responses were reinforced irrespective of discriminative stimulus or lever position. Extinction continued for 6 sessions.

Measures. The measures were the same as those used in Experiment 1.

Results and Discussion

Discrimination Training

One subject died during discrimination training; the analyses below were conducted with the remaining subjects ($n = 11$). We again report the standard deviations from the first response (PRT and DUR) followed by the same measures (IRT and DUR) collected from subsequent responses (see Table 1 for effect sizes). Figure 3 displays data from the last 20 sessions of discrimination training for first and subsequent responses. Large group differences appear in the IRT measure of variation, with the greatest variation always

in Group Far-Low. To evaluate differences in the first response, we conducted a repeated measures ANOVA with Probability (High vs. Low) and Proximity (Near vs. Far) as factors on PRTs and found significant main effects of Proximity, $F(1,10) = 12.55, p = 0.005$, Probability, $F(1,10) = 18.08, p = 0.002$, and the interaction of Proximity x Probability, $F(1,10) = 15.62, p = 0.003$. Separate post-hoc tests revealed variability during Far-Low ($M = 0.36, SD = 0.13$) and Near-Low ($M = 0.26, SD = 0.08$) were greater than all other trials types, $ps < .03$, but variation on Far-Low trials was also reliably greater than Near-Low, $p = 0.0001$. Separate tests also indicated no reliable difference between Far-High ($M = 0.21, SD = 0.04$) and Near-High ($M = 0.19, SD = 0.04$), $p = 0.89$. A similar analysis conducted on DURs of first responses revealed no main effects or interactions, $Fs(1,10) < 1.0$.

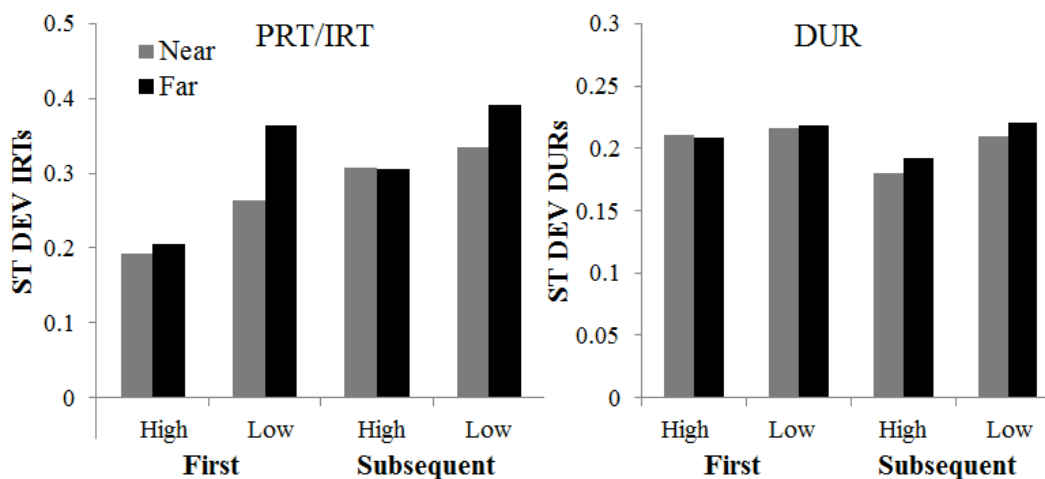


Figure 3. Left panel: The standard deviation of log PRTs and log IRTs from the last 20 sessions of discrimination training in Experiment 2. Right panel: Log DURs from the same period of training. The data from the first and subsequent responses on a trial are displayed separately for each trial type. “Far” and “Near” refer to the location of the lever relative to the food niche, and “High” (100%) and “Low” (10%) refer to the probability that a trial will terminate in reward. All units are measured in seconds.

Similar group differences appear in Figure 3 for subsequent responses. These differences were confirmed by a repeated-measures ANOVA with Probability (High vs. Low) and Proximity (Near vs. Far) as factors conducted on IRTs which revealed a main effect of Probability, $F(1,10) = 6.11, p = 0.03$, a significant Probability x Proximity interaction, $F(1,10) = 6.17, p = 0.03$, as well as a marginally non-significant main effect of Proximity, $F(1,10) = 3.97, p = 0.07$. Separate post-hoc tests revealed that variability during Far-Low ($M = 0.39, SD = 0.08$) was greater than all other trials types, $ps < 0.05$ ($M = 0.39, SD = 0.02$), which did not reliably differ from one another, $ps > 0.43$. Contrary to the results with DUR of first responses, analysis of subsequent responses found a main effect of Probability, $F(1,10) = 8.80, p = 0.01$, but no other main effects or the interaction, $Fs(1,10) < 1.46$.

The results of the last 20 sessions of discrimination training provide strong evidence for increased variation of PRTs/IRTs with a lower probability of reward and a greater distance between response and reward (i.e., proximity). Additionally, probability and proximity interacted to further modulate temporal variation in PRTs/IRTs. With respect to temporal variation in DUR, however, we again have evidence only for increased variation with a lower probability of reward. These results support and extend previous research. Researchers investigating the relationship between reward probability and variation of behavior have reported greater variation in IRTs for pigeons and rats on trials with a low-probability stimulus (Gharib, Derby, & Roberts, 2001; Gharib, Gade, & Roberts, 2004; Stahlman et al., 2010a; Stahlman et al., 2010b; Stahlman et al., 2011b). Variation in IRTs has also been shown to be sensitive to manipulations of reward expectation other than probability of reward, including the temporal proximity of a response to reward and the magnitude of reward.

Variation in response DURs, however, was not found with pigeons and has thus far only been shown to be sensitive to reward probability in rats. Gharib et al. (2001) suggests that IRT and DUR reflect different neural mechanisms which control the decision to respond and the form of the response, respectively. In support, they refer to prior research demonstrating the independence of IRT and DUR (e.g., Roberts, 1987). The implication of this is that treatments that change IRT may leave DUR unchanged, and vice versa. In the current research, proximity and probability influence the decision to respond (IRT), whereas, response form is only affected by the probability of reward signaled by the stimulus. These results support the claim by Gharib et al. (2001) of the independence of IRTs and DURs.

Extinction Training

In order to evaluate whether extinction interacts with one or both factors (probability and proximity), we compared responding from the last 5 sessions of discrimination training with those conducted during extinction (see Figure 4). For first responses, Figure 4 indicates that the differences between trials with a High- or Low-probability discriminative stimulus appear exaggerated during extinction. Our observations were supported by a repeated-measures ANOVA conducted on PRTs with Treatment (Discrimination training vs. Extinction), Proximity (Near vs. Far) and Probability (High vs. Low) as factors which revealed significant main effects of Treatment, $F(1,10) = 16.27, p = 0.002$, Proximity, $F(1,10) = 28.20, p = 0.0003$, Probability, $F(1,10) = 20.67, p = 0.001$, and importantly, an interaction of Treatment x Probability, $F(1,10) = 6.31, p = 0.03$. Separate post-hoc tests of the Treatment x Probability interaction revealed that variation between High-probability and Low-probability trials differed reliably during both discrimination and extinction, $ps < .01$. However, variation on the Low-probability trials (Near-Low and Far-Low) was greater during extinction ($M = 0.43, SD = 0.14$) than discrimination training ($M = 0.32, SD = 0.10$), $p < 0.01$, while variation during High-probability trials (Near-High and Far-High) did not differ reliably from training to extinction, $p > 0.10$ (see Figure 4). A similar analysis conducted on DURs of responses revealed only a main effect of Treatment, $F(1,10) = 27.51, p = 0.0003$, but no other main effects or interactions, $F_s(1,10) < 2.34, ps > 0.16$.

Figure 4 indicates a similar pattern of results during the IRTs and DURs of subsequent responses. This observation is supported by a repeated-measures ANOVA conducted on IRTs for subsequent responses, which revealed effects of Treatment, $F(1,10) = 16.86, p = 0.002$, Proximity, $F(1,10) = 11.35, p = 0.007$, Probability, $F(1,10) = 11.42, p = 0.007$, and the interaction of Treatment x Probability, $F(1,10) = 11.59, p = 0.007$. Separate post-hoc tests of the Treatment x Probability interaction revealed variation during the Low-probability and High-probability trial types did not differ reliably during discrimination training, $p > 0.10$, but did support greater variation on Low-probability trials during extinction, $p < 0.001$. The analysis of DURs for subsequent responses again found a main effect of Probability, $F(1,10) = 7.37, p = 0.02$, and Treatment, $F(1,10) = 9.38, p = 0.01$.

Recent research on variability as a function of reward expectation has either reported measures of variation during a steady-state procedure or during extinction. This is the first report, of which we know, that examines temporal variation during extinction following a steady-state procedure. It is unlikely that extinction would influence variation uniformly throughout each session [e.g., considering factors such as the rate of extinction during a session, the magnitude of recovery at the beginning of each extinction session, and the rate of change of this magnitude over consecutive sessions (see Bouton, 2004)]. Nevertheless, we can make some general conclusions regarding the comparison of data during steady-state responding and across the six sessions of extinction. Not surprisingly, the most robust finding during extinction was an increase in variation for all trial types. However, when an exception occurred (e.g., PRTs in extinction), it was due to no change in temporal variation on Near-High trials; during which both factors (near lever and high probability) supported lower baseline levels of variation. We reported Treatment x Probability interactions for both PRTs and IRTs,

indicating a larger effect of probability than proximity on temporal variation during extinction. Temporal variation in DURs was again affected only by the probability of reward.

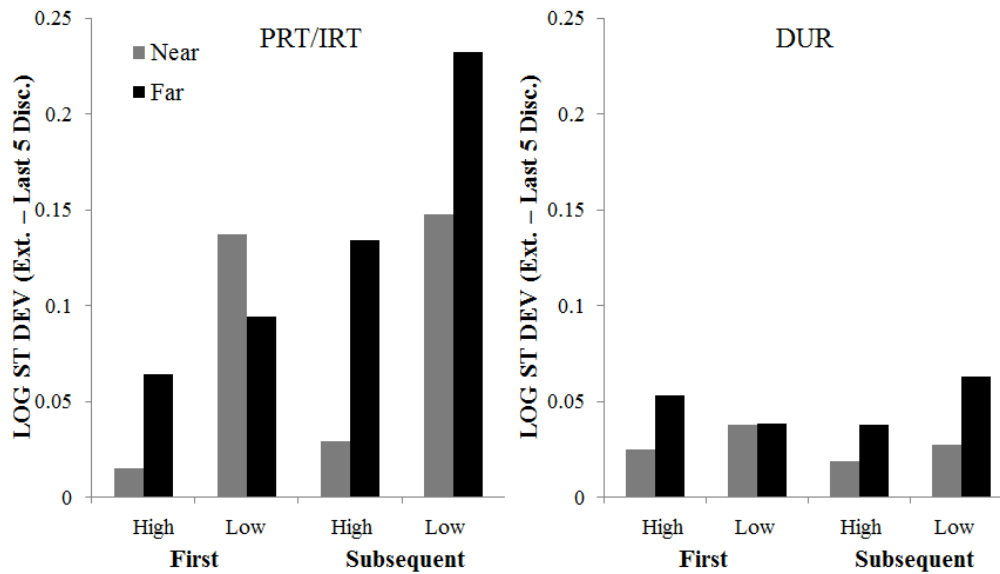


Figure 4. Difference score of the standard deviation of log PRTs, log IRTs, and log DURs during extinction minus the mean for those values over the last five sessions of discrimination training in Experiment 2. The data from the first and subsequent responses on a trial are displayed separately for each trial type. “Far” and “Near” refer to the location of the lever relative to the food niche, and “High” (100%) and “Low” (10%) refer to the probability that a trial will terminate in reward. All units are measured in seconds.

General Discussion

On the whole, our results support and extend prior work on the control of behavioral variability by associative features. On the basis of recent research, we had predicted that the variability of rats’ lever pressing would increase with a decreased proximity of the lever to the goal and with a decreased reward probability, and that these factors should interact with each other to produce variation in operant behavior. Our results largely support our predictions.

The Effects of Proximity and Probability

In Experiment 1, we found that the standard deviation of IRT and DUR scores was greater with decreased probability of reward at the termination of trials. This factor weakly interacted with spatial proximity between the lever and reward, such that the effect of reduced probability increased with reduced proximity.

We conducted Experiment 2 to examine the effect of increasing the amount of training and of increasing the discrepancy between the levels of reward probability. As expected, we found that proximity and reward probability both modulated behavioral variability, and that the factors interacted to produce their effects. Specifically, in multiple instances, lever pressing behavior on Low-Far trials was more variable than behavior exhibited on the other three trial types; it is as though the combination of reduced proximity to the goal location and reduced probability of reward crosses a threshold that triggers higher amounts of variability in behavior. These findings corroborate those of Stahlman et al. (2011b), who found that a similar interaction

between low levels of multiple associative factors (e.g., probability, magnitude of reward) produced an increase in variability that was not realized when either of the factors was low by themselves. In other instances (Experiment 1, PRTs), temporal variation on Far-Low was greater than all other trials types, whereas Near-Low did not differ reliably from Far-High, but variation was greater than Near-High. This suggests that in some cases, the effects of increased proximity and low probability of reward produce similar levels of variability, which are greater still than that observed during Near-High.

Of course, we cannot identify whether the additional sessions of non-differential and discrimination training or increasing the discrepancy between levels of the reward probability in Experiment 2 resulted in the reliable interaction of proximity by probability on PRTs and IRTs. There is reason to believe both manipulations would contribute to stronger effects in Experiment 2. More training in Experiment 2 would result in more reliable differences during the last 20 sessions of discrimination training, but increasing the discrepancy between levels of probability may exceed the previously mentioned threshold for higher amounts of variability during Low-Far trials.

The Impact of Extinction

In Experiment 2, we followed our training procedures with an extinction test. Prior research has typically found that extinction produces an increase in behavioral variability (e.g., Antonitis, 1951; Neuringer et al., 2001). We corroborated these findings, with variability generally greater during extinction than it was during training. We found that the variability of both lever press PRTs and IRTs was modulated by the proximity and by the probability of stimuli ending with reward during training. Interestingly, probability interacted with treatment for PRTs/IRTs and durations of subsequent responses, but we did not find any effects of probability, proximity, or an interaction with treatment on the durations of first responses. This dissociation is not unprecedented; for example, Stahlman et al. (2010a) found that the variation of pre-response times for pigeons' operant pecking behavior was modulated by reward probability, but only for responses subsequent to the initial peck. It is probable that there are distinct processes that contribute to the form of the response, depending on whether the response is the first or a subsequent response in a sequence. These processes appear to be differentially impacted by factors contributing to expectation (e.g., reward probability).

A final result was that at least two differences in behavior (in variation of both IRT and DUR across trial types) only became evident during the extinction procedure in Experiment 2, after being non-existent during the training period. This further suggests that behavior does not conform to reinforcement parameters linearly, but instead that there is a functional calculation of reward expectation that is compared to a threshold that determines whether the animal will respond with relatively large or small amounts of variability in behavior. The parameters that contribute to this calculation have been shown to be, at minimum: the probability (Gharib et al., 2004; Stahlman et al., 2010a), magnitude (Stahlman et al., 2011b), and quality (Flaherty et al., 1979) of reward, the proximity (spatial and temporal) of the reward to the location of the response (Stahlman et al., 2011b), and almost certainly the strength of the animal's relevant motivational state (Stahlman et al., 2010b). At present, the relative weights of these factors in the determination of variability in behavior are unknown, but it is evident that each holds some sway in the production of variability. Determining the functional relationship of these factors with respect to each other and with respect to behavioral variability remains an important set of questions to be answered in the future¹.

¹ It is highly likely that the value of reinforcement parameters with respect to variability production is going to be species-specific, at minimum. For example, there are data that suggest that reducing reward probability to 25% from 100% in lever-pressing rats is enough to engender consistent and enduring increases in behavioral variability (e.g., Gharib et al., 2004), but that a similar decrease in reward probability is insufficient to increase operant key-pecking variability in pigeons (e.g., Stahlman et al., 2010b).

One Potential Limitation

One possible interpretation of our experimental design and results is that the spatial relationship between the operant location and the reward location is merely an incidental factor that is not, in fact, in control of behavioral variability. After all, our manipulation of the spatial relationship between response and reward is confounded with the temporal proximity between these events. If our effects were driven by temporal proximity this study would, in essence, replicate with rats the temporal proximity effect demonstrated in pigeons (Stahlman & Blaisdell, 2011b) and extend previous results by also reporting differences during a post-acquisition extinction test. Of course, it is difficult to know which of the two possible factors (i.e., temporal vs. spatial proximity) is the controlling factor, or indeed, whether either is solely in control of behavioral variability. Future experiments that explicitly manipulate either the spatial or temporal proximity between response and reward could certainly illuminate the nature of the causal relationship that controls behavioral variability in tasks such as these.

Conclusions and Future Directions

These data strengthen the case that elements of an expected reward modulate the production of behavioral variability. A reduction in expectation, as instantiated by a reduction in probability and/or spatiotemporal proximity of response and reward, is sufficient to increase behavioral variability as compared to conditions with comparatively higher levels of reward expectation. This general finding has been replicated and verified numerous times using a variety of measures, suggesting that it is in fact a common process by which animals engage with the world around them. Future studies should focus on empirically determining the functional aspects of the relationship between expectation and variability – how does it benefit the organism to modify respondent variability in the face of fluctuations in reward expectancy?

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.
- Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory*, *11*, 485-494.
- Cheng, K. (1992). Three psychophysical principles in the processing of spatial and temporal information. In W. K. Honig & J. G. Fetterman (Eds.), *Cognitive aspects of stimulus control* (pp. 69-88). London, England: Erlbaum.
- 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.
- Flaherty, C. F., Powell, G., & Hamilton, L. W. (1979). Septal lesion, sex, and incentive shift effects on open field behavior of rats. *Physiology & Behavior*, *22*, 903-909.
- 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.
- Herrick, R. M., & Bromberger, R. A. (1965). Lever displacement under a variable ratio schedule and subsequent extinction. *Journal of Comparative & Physiological Psychology*, 59, 392-398.
- Hull, C. L. (1943). *Principles of behavior*. New York, NY: Appleton-Century-Crofts.
- Kacelnik, A., & Todd, I. A. (1992). Psychological mechanisms and the marginal value theorem: Effect of variability in travel time on patch exploitation. *Animal Behaviour*, 43, 313-322.
- MacArthur, R. H., & Pianka, E. R. (1966). On optimal use of a patchy environment. *The American Naturalist*, 100, 603-609.
- 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.
- Notterman, J. M. (1959). Force emission during bar pressing. *Journal of Experimental Psychology*, 58, 341-347.
- Pecoraro, N. C., Timberlake, W. D., & Tinsley, M. (1999). Incentive downshifts evoke search repertoires in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, 25, 153-167.
- Roberts, S. (1987). Evidence for distinct serial processes in animals: The multiplicative-factors method. *Animal Learning and Behavior*, 15, 135-173.
- Roberts, S., & Gharib, A. (2006). Variation of bar-press duration: Where do new responses come from? *Behavioral Processes*, 75, 215-223.
- Silva, F. J., Timberlake, W., & Gont, R. S. (1998). Spatiotemporal characteristics of serial CSs and their relation to search modes and response form. *Animal Learning & Behavior*, 26, 299-312.
- Staddon, J. E. R., & Horner, J. M. (1989). Stochastic choice models: A comparison between Bush-Mosteller and a source-independent reward-following model. *Journal of the Experimental Analysis of Behavior*, 52, 57-64.
- Stahlman, W. D., & Blaisdell, A.P. (2011a). Reward probability and the variability of foraging behavior in rats. *International Journal of Comparative Psychology*, 24, 168-176.
- Stahlman, W. D., & Blaisdell, A. P. (2011b). The modulation of operant variation by the probability, magnitude, and delay of reinforcement. *Learning & Motivation*, 42, 221-236.
- Stahlman, W. D., Roberts, S., & Blaisdell, A. P. (2010a). Effect of reward probability on spatial and temporal variation. *Journal of Experimental Psychology: Animal Behavior Processes*, 36, 77-91.
- Stahlman, W. D., Young, M. E., & Blaisdell, A.P. (2010b). Response variability in pigeons in a Pavlovian task. *Learning & Behavior*, 38, 111-118.
- Stebbins W. C., & Lanson, R. N. (1962). Response latency as a function of reinforcement schedule. *Journal of the Experimental Analysis of Behavior*, 5, 299-304.
- Wiegmann, D. D., & Smith, B. H. (2009). Incentive relativity and the specificity of reward expectations in honey bees. *International Journal of Comparative Psychology*, 22, 141-152.

Financial Support: This research was conducted at Texas Christian University. Financial support was provided by Texas Christian University.

Conflict of Interest: K. J. Leising and C. M. Ruprecht declare no conflicts of interest. W. D. Stahlman was a co-editor of the IJCP special issue on variability but the editorial decisions regarding this manuscript were handled by A.P. Blaisdell.

Submitted: August 30th, 2013
Resubmitted: December 9th, 2013
Accepted: January 21st, 2014