

Midsession shifts in reward probability and the control of behavioral variability

W. David Stahlman¹ · Kenneth J. Leising²

Published online: 19 February 2016
© Psychonomic Society, Inc. 2016

Abstract Extensive research has shown that the variability of organismal behavior is great when contingent reinforcement is delayed, small, or improbable. This research has generally employed stable response–reinforcer relationships, and therefore is limited in its explanatory scope with respect to a dynamic environment. We conducted two experiments to investigate whether pigeons’ conditioned pecking behavior shows anticipatory or perseverative patterns of behavioral variability when the reinforcement probability reliably changes within experimental sessions. In Experiment 1, three pigeons received alternating sessions in which the reinforcement probability (35% or 4.2%) was shifted at the midpoint of each session in the presence of the same discrete cue. Experiment 2 featured a similar design, but with the inclusion of a discrete visual cue that changed at the session midpoint, and thus unambiguously indicated reinforcement probability. Local reinforcement rates only reliably controlled response variability when a discrete visual cue was available. Without this, pigeons did not discriminate between trial types in the first halves of sessions, and showed evidence of perseveration of response variability following a within-session shift. Critically, this is the first experimental demonstration that the relationship between reinforcement probability and behavioral variability is moderated by another factor (i.e., trial position within a session). This study thus expands our understanding of the control of behavioral variability as a function of experiential factors.

Keywords Reinforcement · Anticipation · Perseveration · Behavioral variability

Organisms that accurately predict future events have an advantage over those that do not. Observed stimuli that inform an animal about future outcomes can provide prescriptions for adaptive behavior to exploit appetitive resources or to avoid dangerous events. However, the world is, in practice, dynamic. Previous predictive stimuli may no longer signal an available outcome (i.e., extinction) or may signal that a novel variation of a response is required (e.g., after “low-hanging fruit” has been consumed). Thus, we should expect to find that organisms have evolved systems that strike a balance between exploiting known advantages and producing new behaviors (Pecoraro, Timberlake, & Tinsley, 1999; Roberts & Gharib, 2006) that may potentially result in greater benefits. The generation of novel behavior is a critical component for an animal to profitably engage with different environments, or simply for it to increase its efficiency within a given, known environment (Epstein, 1990). New forms of behavior, emitted in an external environment, are then differentially strengthened according to the processes of reinforcement and extinction.

Researchers have recently examined the relationship between predictive learning and the production of behavioral variability in animals (e.g., Griffith, Farnsworth, & Stahlman, 2015; Jensen, Stokes, Paterniti, & Balsam, 2014; Stahlman, Roberts, & Blaisdell, 2010a). Almost invariably, researchers observe a relationship between predictive learning of the probability of reinforcement for behavior and the degree of behavioral variability—when the reinforcement (e.g., food) probability is low, behavioral variability is high. Similarly, when the reinforcement probability is high, behavioral variability is comparatively low (e.g., Gharib, Gade, & Roberts,

✉ W. David Stahlman
wdstahlm@umw.edu

¹ Department of Psychological Science, University of Mary Washington, 1301 College Ave., Fredericksburg, VA 22401, USA

² Texas Christian University, Fort Worth, TX, USA

2004; Stahlman, Roberts, & Blaisdell, 2010a).¹ An analogy here to Darwinian processes is obvious. With respect to biological evolution, environmental pressures may alter the variability of traits in a population over time (Darwin, 1859) by restricting the individuals that successfully produce offspring (e.g., by sexual selection; see Emlen & Oring, 1977). With respect to behavior, reinforcement functions similarly to that of biological selection pressures (Skinner, 1974). That is to say, a high level of behavioral reinforcement (or reproductive success) generally produces restricted variation in the resultant distribution of behaviors (or traits) relative to conditions in which reinforcement (or reproduction) is unlikely. Gharib, Derby, and Roberts (2001) suggested that this may represent a means by which behavior is adjusted for current reinforcement conditions. When an animal can predict that positive reinforcement is unlikely, there is very little cost to engaging in different and unusual behaviors. A subset of the consequent, novel behaviors may then prove more beneficial to the organism than those of its earlier repertoire. However, it is wasteful to engage in the generation of variable behavior when reinforcement delivery is otherwise likely. Doing so may result in missed reinforcement opportunities and may increase vulnerability to predators and other dangers.

A large amount of recent work has linked reinforcement to the control of respondent variability. The evidence includes work performed with various species (e.g., Doughty, Giorno, & Miller, 2013; Jensen et al., 2014; Stahlman & Blaisdell, 2011b), various response classes (e.g., Griffith et al., 2015; Stahlman, Young, & Blaisdell 2010b), and various dimensional manipulations of the reinforcer (e.g., Stahlman & Blaisdell 2011a, b). The accumulated evidence indicates that the relationship between reinforcement and behavioral variability is ubiquitous in and fundamental to learned animal behavior.

However, this is not to say that behavior always conforms immediately or optimally to reinforcement likelihood (for example). Thomas Zentall and his colleagues have investigated a striking example of inefficient choice behavior in nonhuman animals (e.g., Laude, Stagner, Rayburn-Reeves, & Zentall, 2014; Rayburn-Reeves, Molet, & Zentall, 2011; Rayburn-Reeves, Stagner, Kirk, & Zentall, 2013; Rayburn-Reeves & Zentall, 2013; Stagner, Michler, Rayburn-Reeves, Laude, & Zentall, 2013; see also Cook & Rosen, 2010; Daniel, Cook, & Katz, 2015; McMillan & Roberts, 2015; McMillan, Sturdy, & Spetch, 2015). An early article in this experimental series presents several experiments that are illustrative of the basic effect. Rayburn-Reeves

et al. (2011) trained pigeons on a task in which two possible response options were indicated by readily discriminable color cues: simultaneously presented red and green discs. For the first half of the training sessions, pigeons were reinforced with grain for only responding to one of the two color cues; beginning with the first trial of the second half, the reinforced response shifted to the other disc. For example, if a pigeon were reinforced for pecking the green disc in the first half of a session, it would be reinforced for pecking the red disc in the second half. Interestingly, the pigeons learned to anticipate the shift in response requirements, making incorrect responses to the nonreinforced disc prior to the midpoint of the session (i.e., anticipatory errors); additionally, for some time after the session's midpoint, the pigeons continued responding to the originally reinforced disc (i.e., perseveratory errors). This pattern of behavior therefore generated a large number of "errors," or responses that did not result in the delivery of grain reinforcement and that persisted even after extensive training. McMillan, Kirk, and Roberts, (2014) reported similar results in rats with respect to perseverative and anticipatory errors in a spatial task requiring navigation of a T-maze. Taken together, these results represent an effect whereby animals' choice behavior is imprecisely controlled and therefore generates many missed opportunities for reinforcement.

Behavioral topography (i.e., *how* actions are carried out) is as worthy of experimental analysis as "choice" (i.e., whether and which action occurs) or rates of response (i.e., the likelihood of an action). "Mutations" (Skinner, 1974), or novel variations, of prior organismal behavior may be a primary driver in establishing new adaptive responses (Epstein, 1990). Therefore, it is important to study not just, for example, choice behavior, but also the degree to which it is similar to (or different from) other emitted responses. For example, Rayburn-Reeves et al., (2011) could have also collected beak aperture, the distance of responses from the center of the color cues, or the interpeck interval. These kinds of measures would reveal whether other aspects of pecking behavior changed as a function of the reinforcement contingencies. The overwhelming majority of the research performed on midsession reversals has used multiple concurrently presented visual stimuli as the targets (but see McMillan et al., 2015). In contrast, examinations of behavioral variability have typically employed static reinforcement conditions with respect to single predictive stimuli. For example, in a study using pigeons, Stahlman, Young, and Blaisdell, (2010b) assigned unique and stable probabilities of grain delivery to each of six distinct, individually presented stimulus targets. As predicted, they found that low reinforcer probabilities produced greater amounts of variability in pigeons' pecking behavior to those targets. This sort of task and this sort of result are typical within the recent literature on the study of behavioral variability. An analysis

¹ This is known as *respondent variability* (Neuringer, 2004) and is distinguishable from other sources of variability in animal behavior. Another means by which variability is controlled, for example, includes procedures in which novel or unusual behavioral variations are explicitly reinforced (e.g., Pryor, Haag, & O'Reilly, 1969). This sort of variability in behavior is known as *operant variability* (Neuringer, 2002; Page & Neuringer, 1985).

of the effects of predictable shifts in reinforcement probability within a session on behavioral variability, however, has not yet been conducted.

It is conceivable that behavioral variability, like choice, might anticipate a scheduled change in reinforcement probability. We might expect, for example, that a pigeon that is frequently reinforced for pecking a target in the first halves of sessions and rarely reinforced for similar responses in the second halves of sessions might demonstrate elevated levels of behavioral variability prior to the midpoint of sessions (i.e., anticipation). Alternatively, variability may lag at a previously established level following a shift in reinforcement probability. In other words, we might observe that a rich reinforcement schedule in only the first halves of sessions would depress variability even after a scheduled shift in reinforcement likelihood at the midpoint of a session (i.e., perseveration). However, to this point, no such examinations have yet been conducted.

The interesting issue suggested by the literature on behavioral variability and midsession reversals is whether the degree of respondent behavioral variability would exhibit anticipation or perseveration in the same manner as choice behavior. Using pigeons, we conducted two experiments in which the probability of reinforcement for pecking a single discriminative visual target was reliably shifted at the midpoint of every session (see Tables 1 and 2 below). In Experiment 1, this shift was not signaled by any discrete cue. A within-subjects comparison in a group of pigeons (Group Shift) was used to determine how a midsession shift from a high probability of reinforcement (35%) to trials with a low probability of

reinforcement (4.2%) influenced variability in the topography of pecking. Shifts alternated daily between high-to-low probability and low-to-high probability. It was conceivable that behavioral variability, like choice, might shift in advance of the scheduled change in reinforcement probability (i.e., anticipation). Variability might also lag at a previously established level following a shift in reinforcement probability (i.e., perseveration). With no visual cue to indicate reinforcement likelihood, we predicted that animals would express an anticipatory shift (i.e., prior to the session's midpoint) and/or perseveration (i.e., after the midpoint of sessions) in spatial behavioral variability.

In Experiment 2, the color of the discriminative stimulus reliably indicated the probability of reinforcement. Prior research has shown that the insertion of a reliable color cue dramatically improves performance in a variation of the midsession reversal task (Cook & Rosen, 2010). We predicted that a consistent relation between color cue and reinforcement probability would produce stronger control of variability about the session's midpoint than we found in the Shift birds of Experiment 1.

Experiment 1

In Experiment 1, we were principally interested in whether the variability of pigeons' pecking behavior would (1) anticipate a coming shift in reward probability or (2) perseverate at a prior level after reward probability had shifted values at the midpoint of the session.

Method

Subjects

Six male White Carneaux pigeons (*Columba livia*; Double T Farm, Iowa) served as subjects. All of the birds had previously been trained via a standard autoshaping procedure to peck to the touchscreen and retrieve mixed grain from a hopper. All stimuli used in this experiment were novel with respect to the birds' histories. Pigeons were maintained at 80%–85% of their free-feeding weight. They were individually housed in a vivarium in metal cages with a 12-h light–dark cycle and had free access to water and grit. The experiment was conducted during the light portion of the cycle.

Apparatus

Training and testing were conducted in a flat-black Plexiglas chamber (45 cm wide × 41 cm deep × 46 cm high). All stimuli were presented by computer on a color LCD monitor (Model L1750, HP, Palo Alto, CA) visible through a 33 × 40 cm viewing window in the middle of the front panel of the

Table 1 Experimental design for one hypothetical bird from each of Groups Stay and Shift in Experiment 1

Group	Trial	Probability	Session Type A		Session Type B	
			1–48	49–96	1–48	49–96
Stay	MAX	1.000	Pink	Pink	Pink	Pink
	H	.350	Aqua	Aqua	–	–
	L	.042	–	–	Maroon	Maroon
Shift	MAX	1.000	Pink	Pink	Pink	Pink
	H	.350	Aqua	–	–	Maroon
	L	.042	–	Aqua	Maroon	–

Session types A and B were alternated daily for both groups. The numbers beneath the session types indicate how each 96-trial session was divided into halves. The probability for each trial type indicates the probability that reinforcement will be delivered following the end of a trial. Pink, aqua, and maroon are the possible colors for the discriminative stimulus displayed during each trial type. For Group Stay, MAX trials were included in both sessions, but H versus L trials depended on the session type. The color of the discriminative stimulus in each trial type did not change from early (1–48) to late (49–96) during each session type. For Group Shift, both H and L trials were presented in each session, but when (either Trials 1–48 or 49–96) depended on session type. Each peck on a trial had a 20% probability of terminating the trial

chamber. The bottom edge of the viewing window was 12 cm above the chamber floor. Pecks to the monitor were detected by an infrared touch screen (Model EZ-170-WAVE, ezscreen, Houston, TX) mounted on the front panel. A 28-V houselight located in the ceiling of the box was illuminated 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 floor; when activated, it allowed 3 s of grain access. All experimental events were controlled and recorded with a computer. A video card controlled the monitor in the SVGA graphics mode (800 × 600 pixels). All experimental events were controlled via Microsoft Visual Basic 6.0 software. The stimuli that could be displayed included three colored discs, 6.5 cm in diameter, with a 2-mm black circle at the center of the disc. The stimulus colors were pink (RGB: 255, 192, 192), aqua (RGB: 128, 255, 255), and maroon (RGB: 92, 28, 28).

Procedure

Pigeons were equally divided into two groups: Group Stay and Group Shift. The present experimental parameters were chosen because they have been shown in previous research to produce differences in the variability of pigeons' pecking behavior (Stahlman & Blaisdell 2011a, b; Stahlman, Roberts, & Blaisdell 2010a; Stahlman, Young, & Blaisdell 2010b). This work has consistently demonstrated that pigeons will respond with greater behavioral variability in response to an approximately 4% (or less) probability of reinforcement than they do when there is a 35% (or greater) probability. Therefore, in the present experiment the probability of reinforcement on any particular trial varied between values of 100% (*MAX*), 35% (high, or *H*), and 4.2% (low, or *L*), as we describe below.

A *MAX* trial could be delivered with 20% likelihood at any point within a session for all birds, regardless of group. Though not a focus of our examination, these trials were included with uniform probability throughout each session to maintain consistency with past studies of variability (e.g., Stahlman & Blaisdell 2011a, b) and to maintain relatively high densities of reward. Pilot work had demonstrated that this is a crucial component in maintaining pigeons' performance in these kinds of studies.

The rest of the trials in each session were composed of *H* and/or *L* trials, determined by the session type. Table 1 illustrates some possible sessions for pigeons in each of our two groups (Stay and Shift). A total of 73 daily sessions were conducted, each consisting of 96 trials. A trial began with the presentation of a single disc at the center of the computer touchscreen. Each peck to the disc had a random 20% likelihood of immediately terminating the trial; each peck that did not terminate a trial (i.e., the other 80% of target-directed responses) were recorded but had no effect. A trial's

termination resulted in the disc disappearing from the display, the immediate availability of grain (for 3 s, and subject to the probability assignments below), and the inception of a 6-s intertrial interval (ITI). Off-target pecks were recorded but did not terminate any trials.

Pigeons received a training regimen over the first 14 sessions to gradually acclimate them to the final experimental parameters and to avoid ratio strain (Ferster & Skinner, 1957; D. M. Thompson, 1964). During each of the first six sessions, reaching response criterion on each trial resulted in the delivery of grain. During the next two sessions, the reinforcement probabilities on all non-*MAX* trials were reduced to 50%; during the next six sessions, the reinforcement probabilities on all non-*MAX* trials were reduced to 35%. Parameters for the remaining sessions are described below.

Group Stay The assignment of reinforcement probabilities to target colors was consistent within birds for this group (e.g., a bird might have a 100% likelihood of reinforcement on pink trials, 35% on aqua, and 4.2% on maroon), with the color assignments for *H* and *L* trials being balanced across birds. Two kinds of sessions comprised Group Stay's experience. The first type (*HH*, in which each letter indicates a session half) consisted of trials composed entirely of *MAX* and *H* trials, whereas the second session type (*LL*) consisted of *MAX* and *L* trials. On each of the 96 trials of an *HH* session, pigeons were 20% likely to be presented with a *MAX* trial and 80% likely to receive an *H* trial; on *LL* sessions, each trial was 20% likely to be a *MAX* trial and 80% likely to be an *L* trial. These session types were alternated over days (e.g., *HH*, *LL*, *HH*, etc.).

Group Shift Group Shift also received two kinds of sessions, alternating over days. However, unlike in Group Stay, the color targets were not consistently assigned to experimental reinforcement probabilities. Instead, *H* and *L* probabilities were signaled not simply by a target's color, but also by the trial number within a session. For example, if a maroon disc indicated an *H* reward probability within the first 48 trials of a session, it would indicate an *L* probability for Trials 49–96. Therefore, for this bird, maroon targets would denote an *HL* session. Similarly, an aqua disc would indicate an *L* probability in the first half of a session, but an *H* probability in the second half (i.e., indicating an *LH* session). A pink disc would consistently indicate *MAX* probability, regardless of trial half. See Table 1 for example assignments for birds from each of Groups Stay and Shift. These session types were alternated over days (e.g., *HL*, *LH*, *HL*, etc.).

Measures

We recorded the spatial location of each response to the touchscreen during experimental trials. We then calculated

the standard deviation of the horizontal response location for each subject as a function of the independent factors. The horizontal location of pecks on a touchscreen has been shown to be more sensitive to changes in reward probability than is vertical location in typical experimental situations with a vertically oriented response target; this is likely due to at least one of two reasons. First, the location of the grain hopper was directly below the computer monitor in this and in conceptually similar studies (e.g., Stahlman, Roberts, & Blaisdell 2010a). Therefore, response variation along the vertical axis of the screen would necessarily be meaningful with respect to reinforcement in a way that response variation along the horizontal axis would not. Second, kinesthetically speaking, it is likely easier for birds to freely vary their pecking in the horizontal dimension than in the vertical; pigeons naturally demonstrate restricted variability in the vertical dimension, and preferentially respond to lower regions (e.g., Racey, Young, Garlick, Pham, & Blaisdell, 2011) when pecking at a vertically oriented target.

A common procedure when working with data germane to the topic of the maintenance of behavioral variability is to omit early sessions, in which the animals would not yet have had the opportunity to learn about the relevant aspects of potentially controlling stimuli (e.g., McMillan et al., 2015; Rayburn-Reeves et al., 2011; Stahlman & Blaisdell 2011a, b). Therefore, we report analyses using the final 20 sessions of collected data.

Results and discussion

We grouped the raw data from each session into blocks of eight trials; a total of 12 blocks represents the 96 trials that were presented during every session. Because our experimental hypotheses were concerned only with animals' performance on *H* and *L* trials only, we excluded *MAX* trials from consideration.² Thus, analyses were performed with two levels of session type (e.g., *HL* vs. *LH*) within groups, and 12 levels of block.

Goodness-of-fit tests revealed that the normality assumption for analysis of variance (ANOVA) was not met, thereby necessitating transformation of the raw standard deviation scores prior to analysis. The rank transform (RT) procedure (Conover & Iman, 1981) involves assigning ranks to each of the raw dependent scores within

a dataset, yielding a dependent measure that may be probed via ANOVA. The RT procedure has been validated for nonparametric factorial ANOVA for small sample sizes (Sawilowsky, Blair, & Higgins, 1989; G. L. Thompson, 1991) and is thus appropriate for our dataset.

A repeated measures ANOVA on ranks for Group Shift revealed that pigeons responded with significantly greater variability during *HL* than during *LH* sessions, $F(1, 2) = 116.10$, $p = .009$. Importantly, we also obtained a significant Session Type \times Block interaction, $F(11, 22) = 2.71$, $p = .022$, but no main effect of trial block, $F < 1$. In Fig. 1a, nonparametric density heat maps illustrate the distribution of responses as a function of block, whereas Fig. 1b shows both the group means and individuals' variability in behavior. Planned comparisons showed that variability in *x*-location was significantly greater in *HL* than in *LH* sessions during Blocks 8, 9, 10, 11, and 12 ($ps < .05$), with no significant differences for all other blocks ($ps > .05$).

A repeated measures ANOVA for Group Stay revealed no significant effects of session type [$F(1, 2) = 7.20$, $p = .115$] or trial block ($F < 1$), nor a Session Type \times Block interaction [$F(11, 22) = 1.63$, $p = .160$]. This failure to find a significant effect of session type (i.e., *H* vs. *L*) initially seems at odds with prior work that indicated greater variability with a low reinforcement probability. However, an examination of Figs. 2a and b suggests that our failure to observe the expected effect of reinforcement probability on variation is likely a Type II error, due to low power. Nonparametric density heatmaps (Fig. 2a) suggest a greater spread of responses across the touchscreen on *LL* sessions than on *HH* ones; Fig. 2b illustrates that raw variation is also greater on *LL* than on *HH* sessions for all birds within nearly all trial blocks, including every block from 5 to 12. In fact, individual birds responded with greater raw variability on *L* trials on a total of 33 out of 36 blocks (an average of 11 out of 12 per bird). A sign test (a nonparametric analysis) revealed this to be significantly different from chance, $p = .006$. Thus, we did replicate the predicted effect of reinforcement probability on behavioral variability.

These data indicate a number of important points. It is important that we found that reinforcement probability controlled variability in the predicted direction with stable target probability assignments (as in Group Stay), corroborating a great deal of recent work (e.g., Stahlman, Roberts, & Blaisdell 2010a). A second, and novel, result is found in the group for whom reinforcement probability shifted at the midpoint of sessions. In this task, we observed that reinforcement probability only controlled behavioral variability within the second halves of trials (i.e., after a shift in probability), and even then only after a number of trials following the shift. Pigeons responded with equivalent variability in the first halves of sessions, *despite* reinforcement being dramatically different between trial types. The same difference (i.e.,

² There were a number of good reasons to exclude *MAX* trials from the analysis, besides the nature of our a priori predictions: There were only half as many *MAX* trials as of either the *H* or the *L* trials; *MAX* trials were included merely to maintain consistency with past research and to maintain responding; and there is good reason to expect that the variability on *MAX* trials would not differ from variability on *H* trials (see Stahlman, Young, & Blaisdell 2010b). For these reasons, prior work has also excluded these trials from data analysis (Stahlman & Blaisdell 2011a).

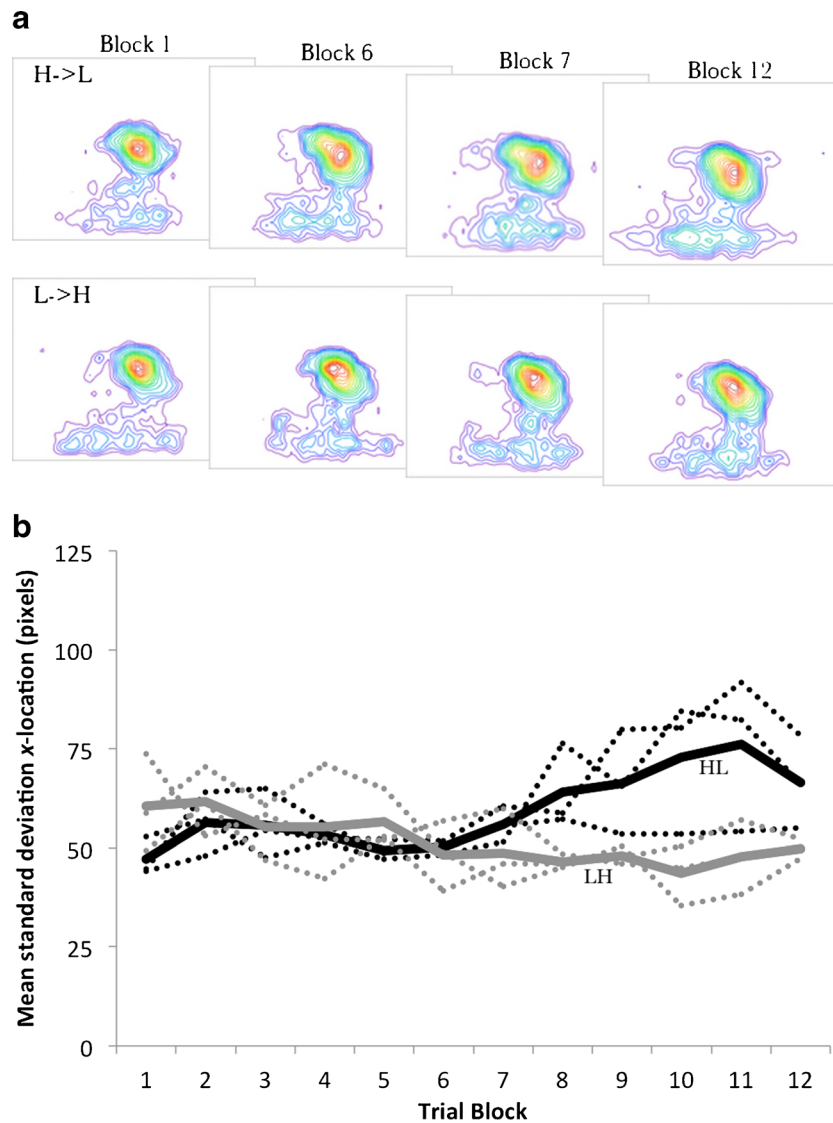


Fig. 1 Experiment 1, Group Shift. (a) Nonparametric density heat maps illustrating the distribution of responses across the touchscreen as a function of session type and trial block. Data depicted are collapsed over all birds and are from the first and last blocks of each of the two halves of sessions. Plots were created with the use of the JMP (Version 11,

SAS Institute, Cary, NC) bivariate nonparametric density function. (b) Mean standard deviations of x -locations of the responses of birds as a function of trial type and trial block. Solid lines indicate the means over birds; dotted lines indicate data from individual birds. H = high probability of grain delivery, L = low probability of grain delivery

35% vs. 4.2%) in local reinforcement probability that differentially controlled behavioral variability in the second halves of sessions did not control variability in the first halves. Clearly, this result cannot be explained by appeals regarding the discriminability between local reinforcement probabilities. Rather, it appears that variability in the first halves of *LH* trials is controlled not by the local reinforcement conditions, but at least in part by the training the animals have had with respect to an upcoming shift to a high probability of reinforcement. In other words, the animals experiencing a low probability of reward in the first half of a session respond with a level of variability more commensurate with a high probability of

reward. This ultimately drives a significant effect of session type in the control of variability. Despite the overall reinforcement probabilities being equivalent within all session types for Group Shift, pigeons respond with greater variability on *HL* than on *LH* sessions.

Critically, this is the first experimental demonstration that the relationship between reinforcement probability and behavioral variability is moderated by another factor (i.e., trial position within a session); we have not previously observed, nor has anyone else reported, animals responding with equal spatial variability despite dramatically different reinforcement probabilities.

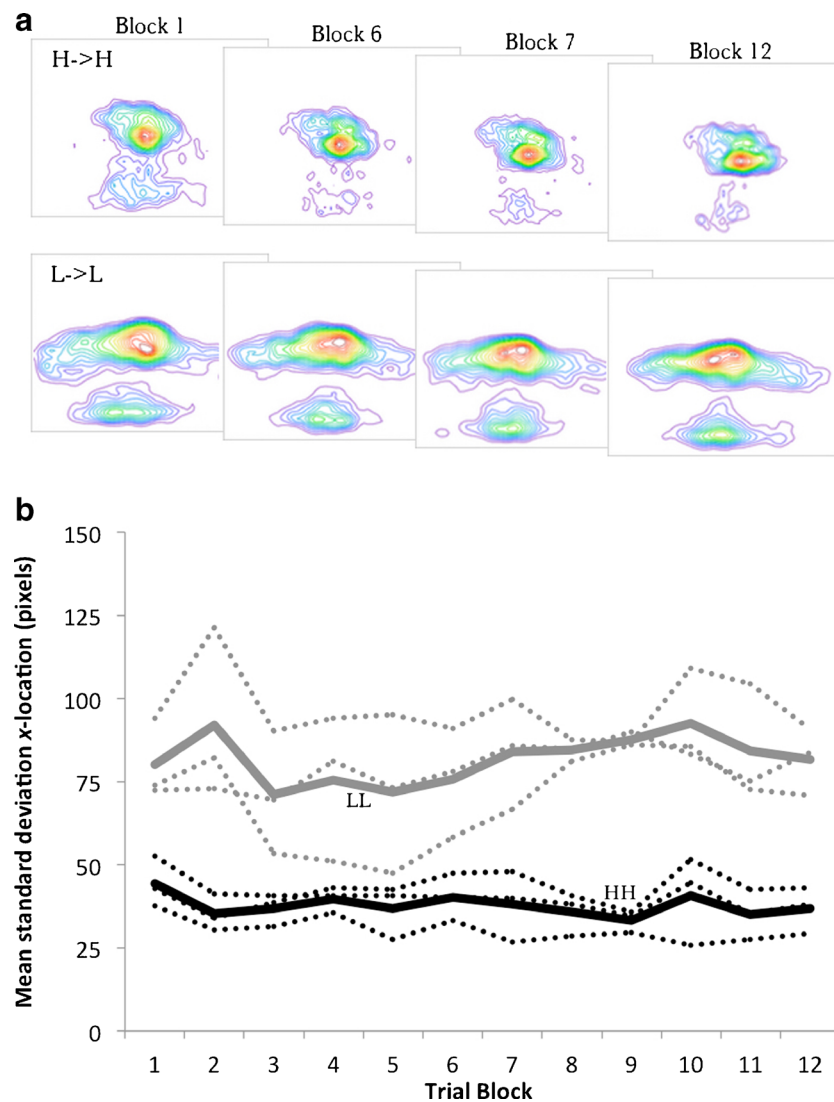


Fig. 2 Experiment 1, Group Stay. (a) Nonparametric density heat maps illustrating the distribution of responses across the touchscreen as a function of session type and trial block. Data depicted are collapsed over all birds and are from the first and last blocks of each of the two halves of sessions. Plots were created with the use of the JMP (Version 11,

SAS Institute, Cary, NC) bivariate nonparametric density function. (b) Mean standard deviations of x -locations of the responses of birds as a function of trial type and trial block. Solid lines indicate the means over birds; dotted lines indicate data from individual birds. H = high probability of grain delivery, L = low probability of grain delivery

Experiment 2

In Experiment 2, we investigated the effect of confounding color cue and midsession shift on stimulus control of behavioral variability. Using a two-choice midsession reversal task, Cook and Rosen (2010) found that pigeons expressed far fewer anticipatory and perseverative errors when they were given an unambiguous, discrete colored cue that indicated the reinforced response. Therefore, we predicted that consistently delivering a discrete, salient visual cue of reward probability would strengthen the control of behavioral variability in Group Shift. In other words, we expected to find significant differences in variability between trial types, irrespective of

session half, and without any anticipatory or perseverative shifts within a session half.

Method

Subjects

Six male White Carneaux pigeons (*Columba livia*; Double T Farm, Iowa) served as the subjects. The pigeons were naïve with respect to the experimental procedures and were maintained in the same manner as those that had served in Experiment 1.

Table 2 Experimental design for one hypothetical bird from each of Groups Stay and Shift in Experiment 2

Group	Trial	Probability	Session Type A		Session Type B	
			1–48	49–96	1–48	49–96
Stay	MAX	1.000	Pink	Pink	Pink	Pink
	H	.350	Aqua	Aqua	–	–
	L	.042	–	–	Maroon	Maroon
Shift	MAX	1.000	Pink	Pink	Pink	Pink
	H	.350	Aqua	–	–	Aqua
	L	.042	–	Maroon	Maroon	–

Session types A and B were alternated daily for both groups. The numbers beneath the session types indicate how each 96-trial session was divided into halves. The probability for each trial type indicates the probability that reinforcement will be delivered following the end of a trial. Pink, aqua, and maroon are the possible colors for the discriminative stimulus displayed during each trial type. Group Stay was identical to the same group in Experiment 1. Group Shift was similar to Group Shift in Experiment 1, except that the discriminative stimuli during *H* and *L* trials in each session were unique colors. When the *H* and *L* trials occurred in each session (Trials 1–48 or 49–96) was dependent on the session type. Each peck on a trial had a 20% probability of terminating the trial

Apparatus

The apparatus was the same as that of Experiment 1.

Procedure

The procedure in Experiment 2 was similar to that employed in Experiment 1, with one primary exception. As before, animals were evenly divided into Groups Stay and Shift, with Stay animals receiving *MAX* and either *H* or *L* trials, alternating over sessions, and Shift animals receiving *MAX* and both *H* and *L* trials within a session. However, unlike in Experiment 1, particular colors were unambiguously predictive of the likelihood of reinforcement. For example, a pink disc might reliably indicate to an animal in Group Shift an *H* probability of reinforcement on a particular trial, no matter whether the trial took place before or after the midpoint of a session. In this way, color would be confounded with trial half. See Table 2 for example assignments for birds from each of Group Stay and Group Shift. Training took place over a total of 73 sessions. As in Experiment 1, pigeons received a training regimen to reduce their reinforcement probabilities gradually over sessions. During each of the first six sessions, reaching response criterion on all trials was reinforced with grain delivery. During the next six sessions, the reinforcement probabilities on (to-be) *H* and (to-be) *L* trials were reduced to 50%; during the next four sessions, the reinforcement probabilities on both *H* and *L*

trials were reduced to 35%. Sessions 17–73 were as described above with the complete experimental design.

Measures

The same measures were utilized in Experiment 2 as were reported in Experiment 1. We report analyses of the final 20 sessions of data collection.

Results and discussion

As in Experiment 1, we grouped the raw data from each session into blocks of eight trials, so that 12 blocks represented all 96 trials. As before, we include analysis of the *H* and *L* trial types only.

Group shift

Critically, a repeated measures ANOVA on ranks for Group Shift revealed the predicted Session Type \times Block interaction, $F(11, 22) = 6.93, p < .0001$; we also observed a significant effect of block, $F(11, 22) = 2.72, p = .022$, but no significant effect of session type, $F(1, 2) = 3.69, p = .190$. Planned comparisons revealed that behavioral variability was significantly different at every level of trial block ($ps < .05$), excepting Blocks 1 and 4 ($ps = .37$ and $.15$, respectively). Greater variability was observed within blocks in the first halves of *LH* sessions and the second halves of *HL* sessions. Figures 3a and b illustrate this strong control of variability by reinforcement probability, in accordance with our predictions—greater reinforcement probability produced less variability in behavior throughout the sessions, irrespective of session half, and without any discernible anticipation or perseveration about sessions' midpoints.

Group stay

A repeated measures ANOVA for Group Stay revealed a main effect of session type, $F(1, 2) = 25.96, p = .036$. No effect of trial block was apparent, $F < 1$, but there was a Session Type \times Block interaction, $F(11, 22) = 4.40, p = .002$. We replicated prior results on the control of behavioral variability, with greater variability occurring in situations with a reduced probability of reinforcement (see Fig. 4a for nonparametric density heat maps, and Fig. 4b for both the group means and individual birds' data).

General discussion

We found strong support for our hypotheses regarding the effect of midsession reinforcement probability reversals on the production of behavioral variability. With no discrete

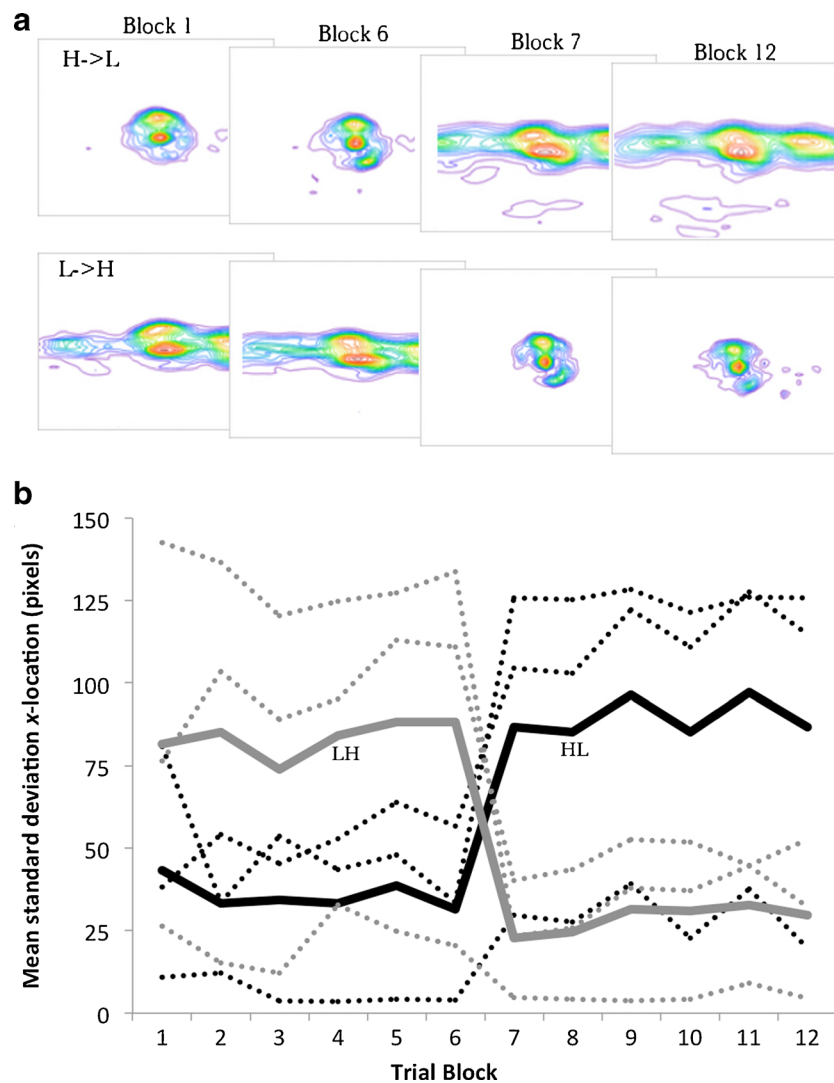


Fig. 3 Experiment 2, Group Shift. (a) Nonparametric density heat maps illustrating the distribution of responses across the touchscreen as a function of session type and trial block. Data depicted are collapsed over all birds and are from the first and last blocks of the each of the two halves of sessions. Plots were created with the use of the JMP

(Version 11, SAS Institute, Cary, NC) bivariate nonparametric density function. (b) Mean standard deviations of x -locations of the responses of birds as a function of trial type and trial block. Solid lines indicate the means over birds; dotted lines indicate data from individual birds. H = high probability of grain delivery, L = low probability of grain delivery

cue that signaled a shift in reinforcement probability (Exp. 1), we obtained evidence for anticipation of a coming upward shift in reinforcement probability, as indicated by equivalent levels of response variability despite substantially different local reinforcement rates. Animals also tended to perseverate at an established level of variability past the point at which the reinforcement probability had shifted. We found that animals' variability in behavior fell under more precise stimulus control when discrete color cues indicated, unambiguously, the probability of reinforcement on any given trial (Exp. 2).

Recent research has demonstrated that pigeons both anticipate upcoming shifts in the location of reinforced responses and perseverate on nonreinforced responses that were

reinforced prior to said shifts (e.g., Cook & Rosen, 2010; Rayburn-Reeves et al. 2011). Lacking from the body of research was an examination of whether variability in behavior would show either an anticipatory or a perseverative bias in an analogous task in which the reinforcement probability changed in the middle of a session. In Experiment 1, we observed that Shift animals failed to discriminate between trial types in the first halves of sessions, and behaved similarly across trials even after a shift in reinforcement probability had occurred (i.e., Block 7), at which point local reinforcement probabilities gained control over behavioral variability in the predicted direction (i.e., high variability with a low reinforcement probability, and vice versa). This was in contrast to pigeons that experienced no shift in reward probability:

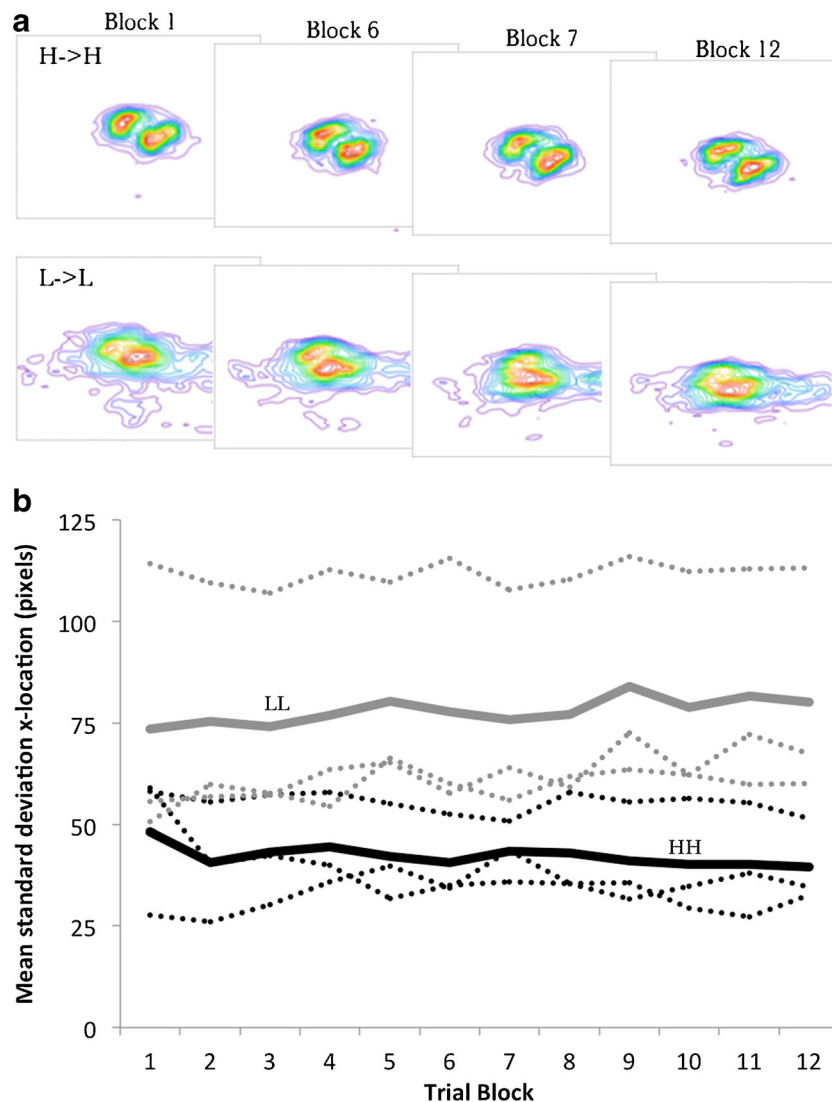


Fig. 4 Experiment 2, Group Stay. (a) Nonparametric density heat maps illustrating the distribution of responses across the touchscreen as a function of session type and trial block. Data depicted are collapsed over all birds and are from the first and last blocks of the each of the two halves of sessions. Plots were created with the use of the JMP

(Version 11, SAS Institute, Cary, NC) bivariate nonparametric density function. (b) Mean standard deviations of x -locations of the responses of birds as a function of trial type and trial block. Solid lines indicate the means over birds; dotted lines indicate data from individual birds. H = high probability of grain delivery, L = low probability of grain delivery

Within-session variability was consistently controlled by reinforcement probability for the animals in Group Stay.³

Experiment 1 indicated that behavioral variability is controlled by more than the mere local reinforcement conditions, since no differences between types of trials emerged for Group

³ A reviewer was curious to know whether *MAX* (100%) trials showed any reliable change within a session for either group, and whether *MAX* trials reliably differed from *H* (35%) trials. As predicted, ANOVAs revealed no differences between the groups in either Experiment 1 or 2 with respect to *MAX* trial variability, nor any differences in variability between *MAX* trials and *H* trials as a function of session block (all $ps > .09$). This was expected, since prior work in this paradigm has consistently shown no differences in response variability in pigeons from 100% probability down to as low as 12.5% probability with 3 s access to grain.

Shift prior to the midpoints of sessions. If local reinforcement conditions were the only important factor in the control of behavioral variability, animals ought to present an emerging difference in behavioral variability as a function of trial type over the first halves of sessions, as they do in the second halves (and as they did in the first halves of Exp. 2). This did not occur. Instead, performance during the first halves of sessions was equivalent across the two types of sessions. This is striking, because it represents a novel finding in the study of respondent variability. A wealth of evidence from our lab within this paradigm (Stahlman & Blaisdell 2011a, b; Stahlman, Roberts, & Blaisdell, 2010a; Stahlman, Young, & Blaisdell, 2010b) has consistently revealed that the reinforcement probability parameters chosen in this study (e.g., 35%

vs. 4.2% likelihood of 3 s access to grain) produce reliable differences in pigeons' response variability. Indeed, these experimental parameters were chosen *because* they have successfully engendered differences in variability. Given the history of this paradigm, and given that reinforcement probability *did* reliably produce differences in response variability in the predicted direction for all animals at all other points of interest, it is doubtful that the lack of first-half difference for Group Shift in Experiment 1 is attributable to factors such as a lack of power, an insensitive measure of variability, or carryover effects from prior sessions. The lack of difference is instead explicable by the fact that the Shift animals in Experiment 1 had neither discrete cues to signal within-session shifts in reinforcement likelihoods, nor discrete cues that unambiguously predicted reinforcement probabilities. The asymmetry about the midpoint that we found with respect to the variability in Group Shift is interesting, and may be explicable in similar terms to that of choice behavior. Rayburn-Reeves et al. (2011, p. 130) suggested that anticipatory errors in midsession reversal tasks should be more likely than perseverative ones, since only the latter provide valid information regarding future contingencies of reinforcement within a session. The relatively low variability on first-half *L* trials may be considered an effect akin to anticipation; compare this to the Shift pigeons' performance in Experiment 2, in which discrete cues for reinforcement probability were present. On the whole, these results support the explanatory narrative indicated by the prior literature regarding the control of variability by Pavlovian expectation (e.g., Stahlman, Young, & Blaisdell, 2010b). The animals in Group Stay consistently demonstrated greater variability in behavior on low-reinforcement trials and reduced variability on high-reinforcement trials, in perfect accordance with recent literature investigating respondent variability (e.g., Stahlman & Blaisdell 2011a, b). When the animals in Group Shift responded with differential variability across trial types (i.e., to the exclusion of the first halves of sessions in Exp. 1), they consistently did so in the predicted direction. That they did not in certain circumstances (i.e., Blocks 1–6) reflects the control of behavior by another factor other than local reinforcement, the first such demonstration of this effect.

Experiment 2 demonstrated the role that an unambiguous discrete predictive cue (i.e., color) plays in the control of behavioral variability in the midsession reversal task. By giving Shift animals a second, more salient (e.g., Donovan, 1978; Lazareva, Vecera, Levin, & Wasserman, 2005; Varela, Palacios, & Goldsmith, 1993) predictor of reinforcement probability (i.e., rather than the timing of a session's midpoint), we found strong control of behavioral variability by reinforcement probability and the elimination of the lack of difference in the first halves of sessions for Group Shift. This research supplements work from recent years on the control of variability by respondent factors (e.g., Leising et al., 2014; Stahlman & Blaisdell, 2011a). Additionally, these results are

precisely what one should expect given the literature regarding midsession reversals in pigeons. Similarly to choice behavior, variability in performance is controlled precisely by reinforcement probabilities when these probabilities are unambiguously indicated by a discrete, salient visual cue (but see McMillan & Roberts, 2012).

Of course, some limitations of the present research could inform future work in this paradigm. We employed a relatively small number of animals (i.e., six per experiment). Though this is a common issue for researchers who work with pigeons, it does promote issues with respect to statistical power in performing parametric analyses. It is also important to note that these experiments do not represent an exhaustion of the possibilities of how we may have manipulated the relationship between targets and the midsession reversal. For example, it could be worthwhile to implement a condition in which target colors did not map consistently to reinforcement probabilities, but instead were discretely predictive of the point at which the reversal takes place (N. McMillan, personal communication): For a single bird, an *HL* session could feature *either* an aqua or a maroon target in the first half of sessions, with the color switching with the reinforcement probability at the midpoint. Another informative manipulation would be to replicate the session types within the Shift condition reported here, but to replace discriminative target cues altogether in favor of a single target (A. Neuringer, personal communication). Manipulations like these might further isolate the control of variability by a timing process, by eliminating a second source of potential control (i.e., the relationship between a specific color and reinforcement).

Future work should also include comparative examination of mammalian species within this paradigm. As with pigeons, reduced reinforcement probability consistently produces greater variability in both rats' (e.g., Leising et al., 2014) and humans' (e.g., Morgan & Lee, 1996) behavior. Though the general relationship between behavioral variability and reinforcement is consistently observed across species (see Stahlman, Leising, Garlick, & Blaisdell, 2013, for a brief review), there are notable differences with regard to specifics. For example, unsignaled within-session shifts in reinforced variability rapidly produce appropriate shifts in response variability in humans (Jensen, Miller, & Neuringer, 2006), a result that appears at odds with the presently reported results. Also, both rats' and humans' performance has been found, under certain circumstances, to differ from that of pigeons on a two-response midsession reversal task. Although humans may tend to anticipate an upcoming reversal prior to the midpoint of a session, they do not demonstrate perseveration (Rayburn-Reeves et al., 2011). Rats have similarly demonstrated a reduced tendency to make errors in variations of the two-choice midsession reversal task (Rayburn-Reeves et al. 2013). It would therefore be worthwhile to investigate whether rats, for example, demonstrate a different pattern

of variability from pigeons on a single-response midsession probability reversal task similar to the one reported here.

The recent history of the experimental investigation of organismal behavior has shown overwhelming evidence for the selective powers of environmental effects (e.g., reinforcement) on the control of behavioral variability. The present work extends scientific knowledge of the relevant factors to include an important factor (trial within a session) in addition to local, static dimensions of reinforcement (e.g., probability, magnitude) to those that control variability in behavior. This is the first such demonstration and suggests productive avenues for future research.

Author note All experimental procedures were done in correspondence with the relevant ethical guidelines and were approved by TCU's Institutional Animal Care and Use Committee. This research was supported by the Department of Psychology at Texas Christian University. We give deep thanks to Jonathon Crystal, Geoff Hall, Neil McMillan, Allen Neuringer, and two anonymous reviewers for their helpful comments that allowed us to improve the manuscript. We thank Chad Ruprecht, Josh Wolf, and undergraduate members of the Comparative Cognition laboratory at Texas Christian University for their assistance with data collection.

References

- Cook, R. G., & Rosen, H. A. (2010). Temporal control of internal states in pigeons. *Psychonomic Bulletin & Review*, *17*, 915–922. doi:10.3758/PBR.17.6.915
- Conover, W. J., & Iman, R. L. (1981). Rank transformations as a bridge between parametric and nonparametric statistics. *American Statistician*, *35*, 124–129.
- Daniel, T. A., Cook, R. G., & Katz, J. S. (2015). Temporal dynamics of task switching and abstract-concept learning in pigeons. *Frontiers in Psychology*, *6*, 1334. doi:10.3389/fpsyg.2015.01334
- Darwin, C. (1859). *On the origin of species*. London, UK: John Murray.
- Donovan, W. J. (1978). Structure and function of the pigeon visual system. *Physiological Psychology*, *6*, 403–437.
- Doughty, A. H., Giorno, K. G., & Miller, H. L. (2013). Effects of reinforcer magnitude on reinforced behavioral variability. *Journal of the Experimental Analysis of Behavior*, *100*, 355–369.
- Emlen, S. T., & Oring, L. W. (1977). Ecology, sexual selection, and the evolution of mating systems. *Science*, *197*, 215–223.
- Epstein, R. (1990). Generativity theory and creativity. In M. A. Runco & R. S. Albert (Eds.), *Theories of creativity* (pp. 116–140). Newbury Park, CA: Sage.
- Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York, NY: Appleton-Century-Crofts.
- Gharib, A., Derby, S., & Roberts, S. (2001). Timing and the control of variation. *Journal of Experimental Psychology: Animal Behavior Processes*, *27*, 165–178. doi:10.1037/0097-7403.27.2.165
- Gharib, A., Gade, C., & Roberts, S. (2004). Control of variation by reward probability. *Journal of Experimental Psychology: Animal Behavior Processes*, *30*, 271–282. doi:10.1037/0097-7403.30.4.271
- Griffith, K. A., Farnsworth, E. M., & Stahlman, W. D. (2015). Reward expectation modulates variability in path choice in rats. *Animal Cognition*, *18*, 131–138.
- Jensen, G., Miller, C., & Neuringer, A. (2006). Truly random operant responding: Results and reasons. In T. R. Zentall & E. A. Wasserman (Eds.), *Comparative cognition: Experimental explorations of animal intelligence* (2nd ed., pp. 652–673). New York, NY: Oxford University Press.
- Jensen, G., Stokes, P. D., Paterniti, A., & Balsam, P. D. (2014). Unexpected downshifts in reward magnitude induce variation in human behavior. *Psychonomic Bulletin & Review*, *21*, 436–444. doi:10.3758/s13423-013-0490-4
- Laude, J. R., Stagner, J. P., Rayburn-Reeves, R., & Zentall, T. R. (2014). Midsession reversals with pigeons: Visual versus spatial discriminations and the intertrial interval. *Learning & Behavior*, *42*, 40–46. doi:10.3758/s13420-013-0122-x
- Lazareva, O. F., Vecera, S. P., Levin, J., & Wasserman, E. A. (2005). Object discrimination by pigeons: Effects of object color and shape. *Behavioural Processes*, *69*, 17–31.
- Leising, K. J., Ruprecht, C. M., & Stahlman, W. D. (2014). Modulation of variation by response-reward spatial proximity. *International Journal of Comparative Psychology*, *27*, 326–337.
- McMillan, N., Kirk, C. R., & Roberts, W. A. (2014). Pigeon (*Columba livia*) and rat (*Rattus norvegicus*) performance in the midsession reversal procedure depends upon cue dimensionality. *Journal of Comparative Psychology*, *128*, 357–366.
- McMillan, N., & Roberts, W. A. (2012). Pigeons make errors as a result of interval timing in a visual, but not a visual-spatial, midsession reversal task. *Journal of Experimental Psychology: Animal Behavior Processes*, *38*, 440–445.
- McMillan, N., & Roberts, W. A. (2015). A three-stimulus midsession reversal task in pigeons with visual and spatial discriminative stimuli. *Animal Cognition*, *18*, 373–383.
- McMillan, N., Sturdy, C. B., & Spetch, M. L. (2015). When is a choice not a choice? Pigeons fail to inhibit incorrect responses on a go/no-go midsession reversal task. *Journal of Experimental Psychology: Animal Learning and Cognition*, *41*, 255–265.
- Morgan, D. L., & Lee, K. (1996). Extinction-induced response variability in humans. *Psychological Record*, *46*, 145–159.
- Neuringer, A. (2002). Operant variability: Evidence, functions, & theory. *Psychonomic Bulletin & Review*, *9*, 672–705.
- Neuringer, A. (2004). Reinforced variability in animals and people: Implications for adaptive action. *American Psychologist*, *59*, 891–906.
- Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*, 429–452. doi:10.1037/0097-7403.11.3.429
- Pryor, K. W., Haag, R., & O'Reilly, J. (1969). The creative porpoise: Training for novel behavior. *Journal of the Experimental Analysis of Behavior*, *12*, 653–661.
- 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. doi:10.1037/0097-7403.25.2.153
- Racey, D., Young, M. E., Garlick, D., Pham, J. N.-M., & Blaisdell, A. P. (2011). Pigeon and human performance in a multi-armed bandit task in response to changes in variable interval schedules. *Learning & Behavior*, *39*, 245–258. doi:10.3758/s13420-011-0025-7
- Rayburn-Reeves, R. M., Molet, M., & Zentall, T. R. (2011). Simultaneous discrimination reversal learning in pigeons and humans: Anticipatory and perseverative errors. *Learning & Behavior*, *39*, 125–137. doi:10.3758/s13420-010-0011-5
- Rayburn-Reeves, R. M., Stagner, J. P., Kirk, C. R., & Zentall, T. R. (2013). Reversal learning in rats (*Rattus norvegicus*) and pigeons (*Columba livia*): Qualitative differences in behavioral flexibility. *Journal of Comparative Psychology*, *127*, 202–211. doi:10.1037/a0026311
- Rayburn-Reeves, R. M., & Zentall, T. R. (2013). Pigeons' use of cues in a repeated five-trial-sequence, single-reversal task. *Learning & Behavior*, *41*, 138–147.

- Roberts, S., & Gharib, A. (2006). Variation of bar-press duration: Where do new responses come from? *Behavioral Processes*, *75*, 215–223.
- Sawilowsky, S. S., Blair, R. C., & Higgins, J. J. (1989). An investigation of the Type I error and power properties of the rank transform procedure in factorial ANOVA. *Journal of Educational Statistics*, *14*, 255–267.
- Skinner, B. F. (1974). *About behaviorism*. New York, NY: Knopf.
- Stagner, J. P., Michler, D. M., Rayburn-Reeves, R. M., Laude, J. R., & Zentall, T. R. (2013). Midsession reversal learning: Why do pigeons anticipate and perseverate? *Learning & Behavior*, *41*, 54–60. doi:10.3758/s13420-012-0077-3
- Stahlman, W. D., & Blaisdell, A. P. (2011a). The modulation of operant variation by the probability, magnitude, and delay of reinforcement. *Learning and Motivation*, *42*, 221–236.
- Stahlman, W. D., & Blaisdell, A. P. (2011b). Reward probability and the variability of foraging behavior in rats. *International Journal of Comparative Psychology*, *24*, 168–176.
- Stahlman, W. D., Leising, K. J., Garlick, D., & Blaisdell, A. P. (2013). There is room for conditioning in the creative process: Associative learning and the control of behavioral variability. In O. Vartanian, A. S. Bristol, & J. C. Kaufman (Eds.), *The neuroscience of creativity* (pp. 45–67). Cambridge, MA: MIT Press.
- 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.
- Thompson, D. M. (1964). Escape from S^D associated with fixed-ratio reinforcement. *Journal of the Experimental Analysis of Behavior*, *7*, 1–8. doi:10.1901/jeab.1964.7-1
- Thompson, G. L. (1991). A note on the rank transform for interactions. *Biometrika*, *78*, 697–701.
- Varela, F. J., Palacios, A. G., & Goldsmith, T. H. (1993). Color vision of birds. In P. H. Zeigler (Ed.), *Vision, brain, and behavior in birds* (pp. 77–98). Cambridge, MA: MIT Press.