

EXAMINING THE DISCRIMINATIVE AND STRENGTHENING EFFECTS OF REINFORCERS IN CONCURRENT SCHEDULES

NATHALIE BOUTROS, DOUGLAS ELLIFFE, AND MICHAEL DAVISON

THE UNIVERSITY OF AUCKLAND

Reinforcers may increase operant responding via a response-strengthening mechanism whereby the probability of the preceding response increases, or via some discriminative process whereby the response more likely to provide subsequent reinforcement becomes, itself, more likely. We tested these two accounts. Six pigeons responded for food reinforcers in a two-alternative switching-key concurrent schedule. Within a session, equal numbers of reinforcers were arranged for responses to each alternative. Those reinforcers strictly alternated between the two alternatives in half the conditions, and were randomly allocated to the alternatives in half the conditions. We also varied, across conditions, the alternative that became available immediately after a reinforcer. Preference after a single reinforcer always favored the immediately available alternative, regardless of the local probability of a reinforcer on that alternative (0 or 1 in the strictly alternating conditions, .5 in the random conditions). Choice then reflected the local reinforcer probabilities, suggesting some discriminative properties of reinforcement. At a more extended level, successive same-alternative reinforcers from an alternative systematically shifted preference towards that alternative, regardless of which alternative was available immediately after a reinforcer. There was no similar shift when successive reinforcers came from alternating sources. These more temporally extended results may suggest a strengthening function of reinforcement, or an enhanced ability to respond appropriately to “win–stay” contingencies over “win–shift” contingencies.

Key words: reinforcer effect, alternation, switching-key, choice, key peck, pigeon

The microstructure of choice has recently been extensively analyzed. Although the seminal research was conducted with pigeons key-pecking for grain in two-alternative concurrent schedule procedures (e.g., Davison & Baum, 2000; 2002; Landon & Davison 2001; Landon, Davison, & Elliffe, 2002), consistent and reliable response patterns have been found across species, response type, reinforcer type and number of response alternatives (Aparicio & Baum, 2006; 2009; Davison & Baum, 2006; 2010; Elliffe & Davison, 2010; Krägeloh, Zapanta, Shepherd, & Landon, 2010; Landon et al., 2007; Lie, Harper, & Hunt, 2009; Rodewald, Hughes, & Pitts, 2010a). One often-obtained regularity is the *preference pulse*: immediately after a reinforcer, choice tends to favor the just-productive alternative before settling at a level representative of the overall

distribution of reinforcers (Davison & Baum, 2002; Landon, Davison, & Elliffe, 2003; Menlove, 1975). This short-lived postreinforcer preference (both its magnitude and its duration) varies in subtle and informative ways related to experimental manipulations (Elliffe, 2006). How should these local changes in choice be interpreted? The transient period of increased responding on the just-productive alternative might, on the one hand, reflect some response-strengthening properties of the just-delivered reinforcer (Skinner, 1938). Alternatively, this momentary preference for the just-productive alternative may result from some discriminative properties of the reinforcer (Davison & Baum, 2006; 2010). For example, due to asymmetrical reinforcer ratios, changeover requirements and bodily position within the chamber, a reinforcer often signals a shorter time to subsequent reinforcement on the just-productive alternative (relative to the not-just-productive alternative), at least momentarily (Boutros, Davison, & Elliffe, 2011).

Krägeloh, Davison and Elliffe (2005) uncoupled the response more likely to produce reinforcement in the postreinforcer period from the prereinforcer response. They accomplished this by manipulating the probability that the next reinforcer would be contingent

This experiment was conducted by the first author in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Auckland. We thank Mick Sibley for his care of the birds and members of the Auckland University Operant Laboratory who helped conduct the experiments.

Address correspondence to Nathalie Boutros, Department of Psychiatry, School of Medicine, University of California San Diego, 9500 Gilman Drive, M/C 0603, La Jolla, California 92093-0603 (e-mail: ntboutros@ucsd.edu). doi: 10.1901/jeab.2011.96-227

on repeating the just-productive response. As this (local) probability of a same-alternative reinforcer increased, the postreinforcer preference for the just-productive response (the preference pulse) became more extreme and enduring. This finding supports a discriminative interpretation of local reinforcer effects: Preference tended towards the alternative more likely to provide subsequent reinforcement rather than always tending towards the just-productive alternative. Preference after a nonfood stimulus (magazine-light or keylight illumination) also varies directly with the probability of a subsequent same-alternative food (Boutros, *et al.*, 2011; Davison & Baum, 2006, 2010).

Although local preference did change as the probability of subsequent same-alternative reinforcement changed, Krägeloh *et al.* (2005) also found consistent tendencies to repeat the just-productive response, regardless of the local probability of subsequent reinforcement for that response. Even when the probability of a same-alternative reinforcer was 0 (indicating strict alternation of reinforcers across the alternatives), the probability of staying at the just-productive alternative immediately after a reinforcer was greater than 0 (Krägeloh *et al.*'s Figure 5). This finding suggests some direct response-strengthening reinforcer properties. Such difficulty with strict alternation procedures is not uncommon (Davison, Elliffe & Marr, 2010; Hearst, 1962; Randall & Zentall, 1997; Shimp, 1976) and may result from some tendency to repeat the just-productive response regardless of whether this response is locally likely to be reinforced (Williams, 1971a; 1971b).

Boutros *et al.* (2011) reported local obtained reinforcer ratios along with the local response ratios (preference pulse) after response-contingent events. Even when the arranged local food ratio in a period was 1:1, the local obtained food ratio in the first few seconds after a response-contingent event was at times extremely biased towards the just-productive alternative. This was attributed to the influence of the changeover requirement: Responses to the not-just-productive alternative could never be reinforced until after the changeover requirement was completed, even if a reinforcer on the not-just-productive alternative was immediately arranged. However, if a reinforcer was arranged on the just-

productive alternative, it could be immediately obtained. Even when the relative probability of arranging a reinforcer on the just-productive alternative was low, the local probability of receiving a reinforcer on that alternative was still higher than the local probability of receiving a reinforcer on the other, not-just-productive-alternative. In this way, changeover requirements may cause the local obtained reinforcer ratio to enter into a positive-feedback relation with the local behavior ratio, and thus maintain extreme local preference.

The present experiment directly investigated the influence of postreinforcer changeover requirements and response availability on preference after a single reinforcer and after specified reinforcer sequences. We arranged a two-key concurrent schedule in which only one side-key was available at a time. A peck to a center switching-key was required to make the other side-key available (Boutros *et al.*, 2011; Davison & Baum, 2006; 2010). Using this procedure, we probabilistically varied which alternative became available immediately after a reinforcer. Specifically, we varied $p_{(P\ on)}$, or the probability that the just-productive alternative (P), as opposed to the not-just-productive-alternative (N), would be available immediately after a reinforcer. In some conditions, the just-productive alternative always came on after a food ($p_{(P\ on)} = 1$), while in other conditions the not-just productive alternative always came on ($p_{(P\ on)} = 0$). In yet another set of conditions, there was no consistent relationship between the keylight pecked prior to the food delivery and the keylight illuminated immediately after that food delivery: The just-productive and not-just-productive alternatives were equally likely to come on ($p_{(P\ on)} = .5$). Using this procedure we hope to elucidate what role response availability has on postreinforcer preference. In addition to varying $p_{(P\ on)}$ across conditions, we also manipulated the local probability of a reinforcer on the just-productive alternative. In four of the eight conditions reinforcers strictly alternated across the left and right response keys: The local probability of a left food delivery was 0 after a left food delivery and 1 after a right food delivery. The four remaining conditions were random alternation conditions where the relative probability of a left food delivery was .5 at all times.

METHOD

Subjects

Six homing pigeons, all experimentally naïve at the start of training and numbered 141 through 146, served as subjects. Pigeons were maintained at $85\% \pm 15\%$ of their of their free-feeding body weights by postsession supplementary feeding of mixed grain when required. Water and grit were freely available in the home cages at all times. The home cages were situated in a room with about 80 other pigeons.

Apparatus

Each pigeon's home cage also served as its experimental chamber. Each cage was 385 mm high, 370 mm wide and 385 mm deep. Three of the walls were constructed of metal sheets and the fourth wall and floor were metal bars. Two wooden perches were mounted 60 mm above the floor, one parallel and the other at a right angle to the back wall. Three 20-mm diameter circular translucent response keys were on the right wall. These keys were 85 mm apart and 220 mm above the perches. The keys required a force exceeding approximately 0.1 N to register an effective response when illuminated. The food magazine was also on the right wall 100 mm below the center key. It measured 50 mm high by 50 mm wide and was 40 mm deep. A food hopper, containing wheat, situated behind the magazine, was raised and the magazine was illuminated during food presentations. All experimental events were arranged and recorded on an IBM-PC compatible computer running MED-PC software which was in a room adjacent to the colony room.

Procedure

As the pigeons were experimentally naïve prior to this experiment, preliminary training, which consisted of magazine training, auto-shaping, and VI training, was first implemented. These procedures were identical to those we have used previously (Boutros, et al., 2011).

Data collection commenced when the pigeons were all responding reliably to the VI 27-s (nominally a random-interval 27 s) schedule of food reinforcement. The two-key concurrent schedule employed a changeover response requirement (Davison & Baum, 2006) rather than a changeover delay: Two pecks on

the switching key were required to switch from one response alternative to the other: The first peck to the center key turned off the side-key that had been lit, and the second peck to the center key turned on the other key and turned off the center key. The first peck to the newly lit side key turned the center key back on, once again allowing changeover responses. In Condition 1, food deliveries randomly alternated across the left and right keys. After any food (from the left or right), and at the start of the session, the next food delivery was randomly allocated to the left or right key with $p = .5$. A probability gate was interrogated every 1 s to determine when food would be scheduled. Once food was set up on an alternative, the next peck to that key turned off that keylight and the red center key, illuminated the food magazine, and raised the food hopper for 3 s. After food delivery, the hopper was lowered, the magazine darkened and both the keylight that had been on prior to the reinforcer, as well as the centre switching key, were illuminated.

Across conditions, we varied whether the food deliveries strictly or randomly alternated. When food deliveries strictly alternated, a left food delivery was always followed by a right food delivery and vice versa. We also varied which alternative became illuminated immediately after a food delivery. When the just-productive alternative always came on, the events after a food delivery were as described for Condition 1. When the not-just-productive alternative always came on, it was the keylight that had not been pecked prior to the last food delivery that came on after the food. In a further set of conditions, the just-productive alternative came on after a food delivery with $p = .5$. Table 1 presents the sequence of conditions. Interspersed among some conditions of this experiment were various conditions of an unrelated experiment, the results of which are not reported here. All sessions lasted 60 minutes or for 60 reinforcers, whichever came first, and each condition lasted 65 sessions.

RESULTS

Because there was some evidence of continuing behavior change relatively late into some conditions (Rodewald, Hughes, & Pitts, 2010b), only data from the last 20 sessions of

Table 1

Sequence of conditions, food alternation (strict or random), and the probability that the just-productive alternative would be lit and available after food.

Cond.	Food alternation	$p_{(P\ on)}$
1	Random	1.0
2	Strict	1.0
3	Strict	0.5
4	Random	0.5
5	Random	0.0
6 (Replication of Cond. 1)	Random	1.0
7	Strict	0.0
8 (Replication of Cond. 2)	Strict	1.0

each condition will be presented. Using these data, the local response ratio after each response-contingent food delivery was plotted as a function of time since that food in 2-s bins. After Baum and Davison (2004), we used the log ratio of responses to the just-productive alternative (P) over responses to the not-just-productive alternative (N). Additionally, we removed any time spent engaging in the changeover response from these analyses. The preference pulses may thus more accurately depict changing local preference rather than procedural constraints.

To create these preference pulses, responses to the just-productive and not-just-productive alternative were separately tallied in each 2-s time bin after a reinforcer. Two preference pulses were calculated for each individual pigeon in the two conditions where $p_{(P\ on)} = .5$ (Condition 4 when the reinforcers randomly alternated and Condition 3 when the reinforcers strictly alternated): one when the just-productive alternative was immediately available, and another when the not-just-productive alternative was immediately available. If a time bin contained fewer than 20 responses in total, no response ratio was calculated for that pigeon in that time bin. If preference was exclusive to one alternative or another, a value of ± 3.5 was used. Figure 1 presents all 6 individual subject P/N log response ratios as well as the group mean (thick line) response ratio in the random alternation conditions (Conditions 1, 6, 4 and 5).

Although Condition 6 was a procedural replication of Condition 1, postreinforcer preference was more extreme in the latter condition, indicating some effect of extended experience. Generally, the same trends were present for all individual pigeons and the

group mean preference pulses were representative of the individual subject preference pulses. In all of these random-alternation conditions, postreinforcer preference was towards the immediately available alternative, whether that was the just-productive alternative (Condition 6 and to a lesser degree Condition 1 and the left Condition 4 panel), or the not-just-productive alternative (Condition 5 and the right Condition 4 panel). In the former panels, the log (P/N) response ratios are all above the horizontal line (which indicates indifference) after a food delivery. This demonstrates a preference for the just-productive (P) over the not-just-productive (N) alternative. In the latter panels the log (P/N) response ratios are below this indifference line, indicating a relative preference for the not-just-productive (N) alternative. There did not appear to be any independent effect of the last food delivery's location.

Figure 2 presents local preference in conditions where food deliveries strictly alternated. These preference pulses were calculated in the same way as were the preference pulses in Figure 1. The Condition 2 preference pulses were largely replicated in Condition 8, suggesting that Condition 6 did not precisely replicate Condition 1 (Figure 1) because Condition 1 was the first experimental condition these initially naive pigeons experienced. In all of the strict alternation conditions (Figure 2) there was always some detectable preference for the not-just-productive alternative as indicated by a log (P/N) response ratio below indifference (the horizontal line at 0). The alternative made available immediately after a food delivery appeared to determine just how extreme this preference was: Preference for the not-just-productive alternative (N) was most extreme in Condition 7 ($p_{(P\ on)} = 0$). Preference was also extreme when the not-just-productive alternative was made immediately available in Condition 3 ($p_{(P\ on)} = .5$; middle right panel). When the just-productive alternative was made immediately available in this condition (left middle panel), preference for the not-just-productive alternative was moderate and short-lived, but was more extreme (further below the indifference line) than in either Condition 2 or Condition 8 ($p_{(P\ on)} = 1$).

A reinforcer's effects extend beyond the immediate postreinforcer period. Landon *et al.* (2002; 2003) reported that choice

Random Alternation Conditions

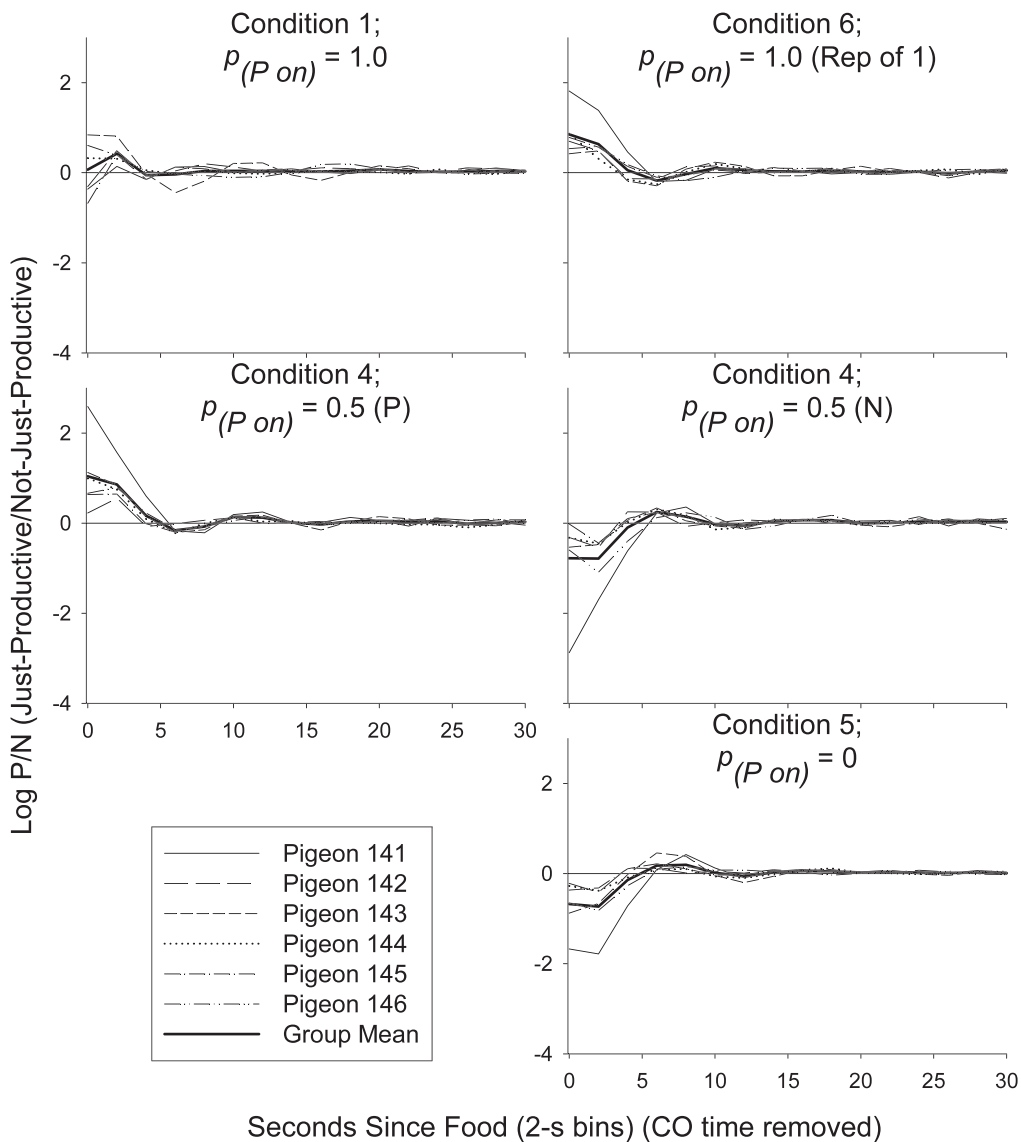


Fig. 1. Log (just-productive/not-just-productive) response ratio in successive 2-s time bins after a food for all individual pigeons in all conditions with random alternation of foods where $p(P_{on})$ was varied across conditions. A value of ± 3.5 indicates exclusive preference in that time bin. The thick line plots the group mean preference pulse and the horizontal line is at 0, indicating the indifference point.

responding in a steady-state concurrent schedule was influenced by each reinforcer in the preceding eight-reinforcer sequence. Landon et al.'s analysis was applied to the present data. Preference was modeled as a function of the sequence of the previous eight food locations. This analysis could only be done on conditions arranging

random alternation of food locations. First, the log (left/right) response ratio was calculated for each of the 256 possible eight food sequences of left and right foods. Then, the contribution of the location of each of the preceding eight food deliveries to the response ratio was calculated using the general linear model:

Strict Alternation Conditions

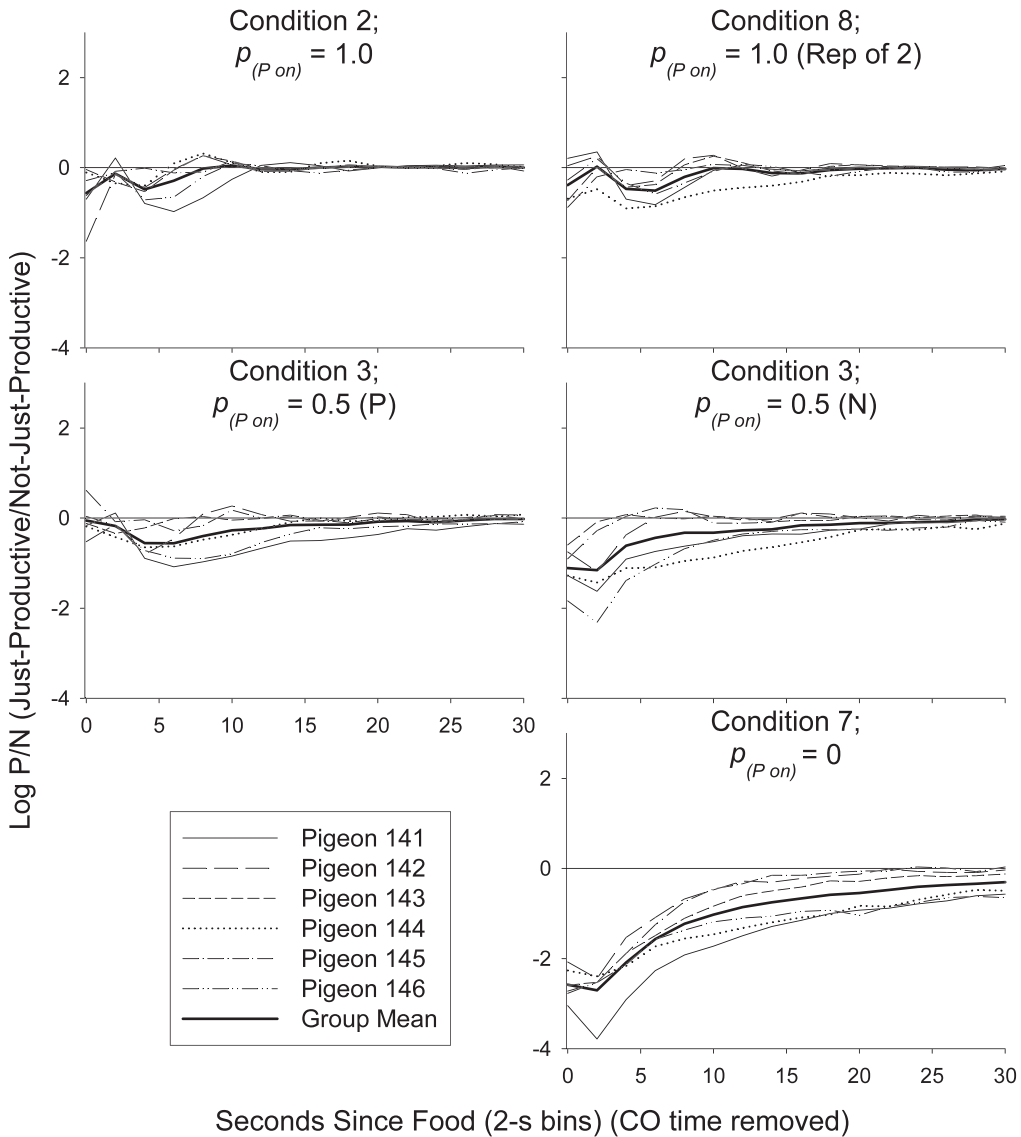


Fig. 2. Log (just-productive/not-just-productive) response ratio in successive 2-s time bins after a food for all individual pigeons in all conditions with strict alternation of foods where $p_{(P\ on)}$ was varied across conditions. A value of ± 3.5 indicates exclusive preference in that time bin. The thick line plots the group mean preference pulse and the horizontal line is at zero, indicating the indifference point.

$$\log\left(\frac{B_L}{B_R}\right) = \log k + \sum_{j=0}^7 \left\{ \begin{array}{l} R_j = L : +b_j \\ R_j = R : -b_j \end{array} \right\} \quad (1)$$

where the subscript j indicates reinforcer lag so that R_0 is the most recent reinforcer and R_7 is the eighth reinforcer back. The effect of a reinforcer

at a given lag is measured by the coefficient b_j . If a reinforcer was from the left, b_j was multiplied by +1; if the reinforcer was from the right, it was multiplied by -1. The constant $\log k$ is residual preference not due to any of the previous eight reinforcers. These values were calculated by finding the best-fitting least-squares estimates of

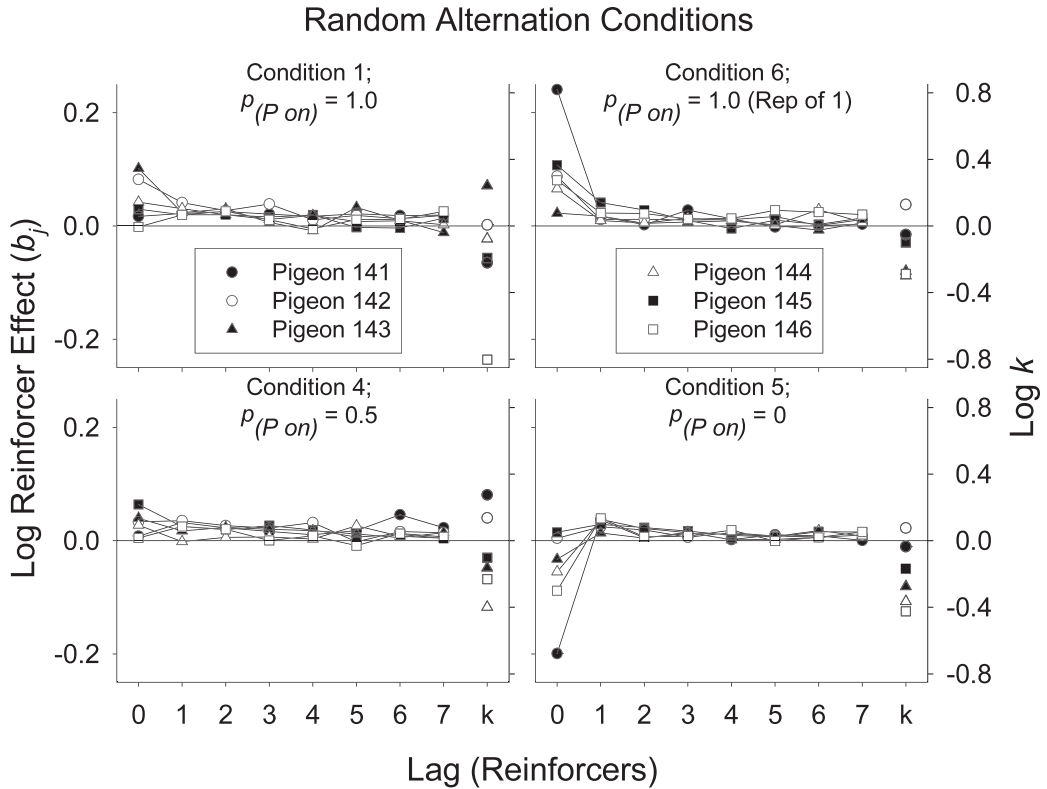


Fig. 3. Lag reinforcer effects and log k (preference not due to any of the 8 preceding reinforcers) for all individual pigeons in all random-alternation conditions where $p_{(P on)}$ was varied across conditions. The lag reinforcer effects (connected by a straight line) are relative to the left ordinate and log k (single point, not connected to others) is relative to the right ordinate.

b_j and log k using Excel's solver function. These b_j and log k values are presented in Figure 3 for each individual subject in each of the random-alternation conditions.

In Figure 3, any log reinforcer effect greater than 0 indicates a preference for the alternative which provided that reinforcer. Any log reinforcer effect less than 0 indicates a preference for the alternative that did not provide that reinforcer. A 0-valued log reinforcer effect indicates that the reinforcer in question had no measurable effect on preference. The effect of any reinforcer, even the most recent one (lag = 0), was always small (consistent with the lag effects Landon et al., 2002, reported when the food ratio was 1:1). As in the preference pulses, reinforcers had generally more extreme effects in Condition 6 ($p_{(P on)} = 1$) than in Condition 1 ($p_{(P on)} = 1$). The most recent reinforcer had a larger effect in Condition 6 than in Condition 1 (though not for all subjects) and log k (inherent bias or

bias due to reinforcers further back than the eighth) generally (though not for all subjects) decreased in the later condition. The effect of the most recent reinforcer (lag = 0) was rather smaller in Condition 4 ($p_{(P on)} = .5$) compared to when $p_{(P on)} = 1$ (Conditions 1 and 6). In Condition 5 ($p_{(P on)} = 0$), preference was shifted away from the alternative that provided the most recent reinforcer (lag₀ log reinforcer effect < 0). This preference for the not-just-productive alternative may be better understood as a preference to the immediately available alternative. The preference for the just-productive-alternative in Conditions 6 and 1 ($p_{(P on)} = 1$) may also be better understood as a preference for the immediately available alternative.

Although the reinforcer lag effect did decrease somewhat after the first reinforcer (lag₁₋₇), it always remained consistently greater than zero even out to lag₇ in all conditions. Additionally, while the log reinforcer effect for

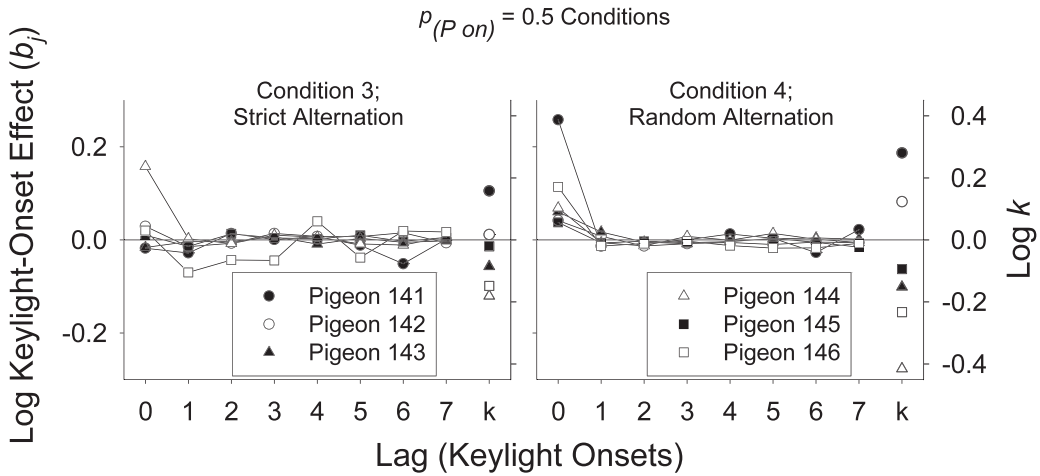


Fig. 4. Lag keylight-onset effects and log k (preference not due to any of the eight preceding keylight onsets) for all individual pigeons in Conditions 3 and 4 where $p_{(P\ on)} = 0.5$. In Condition 3 foods strictly alternated across the side-keys and in Condition 4 they randomly alternated. The lag keylight-onset effects (connected by a straight line) are relative to the left ordinate and log k (single point, not connected to the others) is relative to the right ordinate.

the most recent reinforcer (lag₀) in Condition 5 ($p_{(P\ on)} = 0$) was negative, the log reinforcer effect for reinforcers prior to the most recent (lag₁₋₇) was always greater than zero in all conditions, including Condition 5. These earlier reinforcers consistently appeared to shift preference towards the alternative which provided that reinforcer, regardless of which alternative was made available immediately after a food delivery. Figure 4 presents a similar lag analysis except that preference is a function of the sequence of preceding postreinforcer keylight availabilities rather than reinforcer locations. This analysis could only be done on conditions arranging random alternation of the postreinforcer response availabilities ($p_{(P\ on)} = .5$, Conditions 3 and 4). In Condition 3 food deliveries strictly alternated across the left and right keys and in Condition 4 food deliveries randomly alternated.

Figure 4 shows that while the location of the most recent postreinforcer response availability had a rather large effect (lag₀ > 0), the location of keylight onsets prior to the most recent had no effect (lag₁₋₇ ≈ 0). The effect of earlier keylight onsets was smaller than the effects of previous foods (Figure 3) which although also small, were consistently greater than zero. Thus, while the alternative available immediately after the most recent reinforcer had a rather large influence on current behavior (as seen in the preference pulses,

Figures 1–2), these effects were short-lived, such that there was no detectable effect at the level of preference across successive reinforcer deliveries (Figure 4). Conversely, while there was no apparent immediate effect of the most recent reinforcer's location, there was some indication of small, enduring and cumulative effects of prior reinforcer locations (Figure 3).

Reinforcer-sequence effects can also be investigated by plotting the log response ratio in successive IRIs for selected reinforcer sequences. Figure 5 plots group mean reinforcer trees for sequences of continuations and single discontinuations for the conditions depicted in Figure 3. We present only group mean analyses here due to the low intersubject variability at this level of temporal acuity (Figures 3 and 4).

Figure 5 replicates a number of findings previously reported in both frequently changing and steady-state arrangements. Conditions 1 and 6 were most similar to typical concurrent schedule procedures: In those conditions, foods randomly alternated (overall food ratio of 1:1), and the just-productive alternative were always available immediately after a reinforcer ($p_{(P\ on)} = 1$). Successive same-alternative foods (continuations; indicated in Figure 5 by an unbroken line) progressively shifted preference to the alternative providing those foods. A discontinuation (a reinforcer

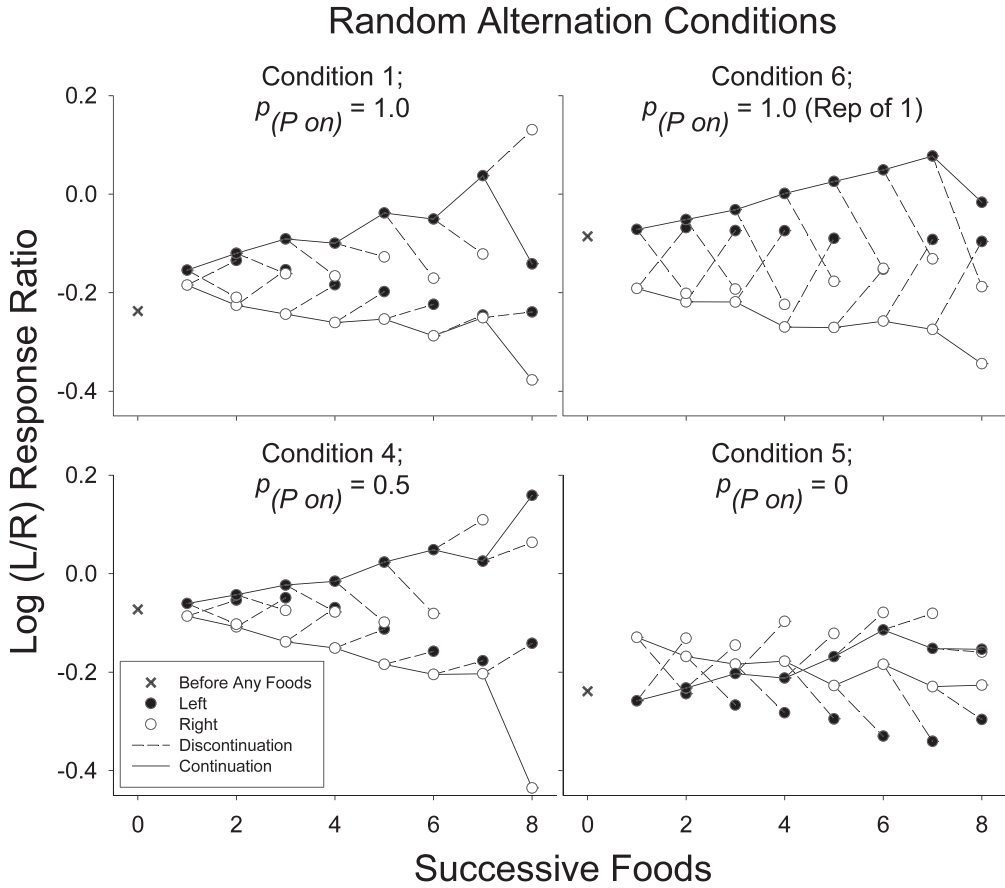


Fig. 5. Group mean log (L/R) response ratio in successive interreinforcer intervals in all random-alternation conditions where $p(P_{on})$ was varied across conditions.

from the alternative that did not provide the prior reinforcers; indicated in Figure 5 by a broken line) shifted preference towards the source of that most recent discontinuing reinforcer. Discontinuations did not, however, completely reverse the extreme preference engendered by prior continuations: There was a detectable residual effect of earlier reinforcers, and preference after a discontinuation was clearly influenced by the number of preceding other-alternative reinforcers. These general trends and patterns were present in both Condition 1 and its replication Condition 6 ($p(P_{on}) = 1$), although reinforcer effects were generally larger in Condition 6. The preference tree for Condition 4 was also largely consistent with these conclusions. Thus, even when the alternative available immediately after a reinforcer was random ($p(P_{on}) = .5$), the preference-increasing effects of reinforcer

location cumulated across successive reinforcer deliveries.

The Condition 5 ($p(P_{on}) = 0.0$) preference tree was rather unlike any previously-published preference tree in that the line tracking preference after successive left continuations (filled circles connected with an unbroken line) and the line tracking preference after successive right continuations (unfilled circles connected with an unbroken line) distinctly crossed paths after approximately four same-alternative reinforcers. Preference after four or fewer food deliveries from a single source was (relative to the right-key bias) toward the not-just-productive alternative (the alternative made immediately available after those reinforcers). However, there was also a general trend for preference to move towards the alternative providing the (continuation) reinforcers. Thus, after five or more same-alternative

reinforcers, preference was toward the alternative providing those reinforcers. As in the other conditions of Figure 5, preference continually shifted towards the alternative providing reinforcers. There appears to be some carryover of preference from the prereinforcer period into the postreinforcer period regardless of what alternative was available immediately after that reinforcer. Although this preference carryover is not visible at the level of the preference pulse, it becomes detectable at the level of preference across successive reinforcer deliveries (Figure 5 above and Figure 3).

There remains some question as to whether the increased preference for the alternative which provided the previous sequence of reinforcers was due to the strengthening effects of the previous reinforcers, or whether this was a case of discriminative control by perceived contingencies. While the experimenter-arranged reinforcer ratio remained 1:1 throughout the random alternation conditions, an extended sequence of same-alternative reinforcers uninterrupted by any reinforcers from the other alternative may have mistakenly signaled a locally rich patch. Thus, the question becomes: Is preference being driven towards the alternative providing successive continuation reinforcers (a response-strengthening process), or is preference tending towards the alternative perceived to be locally richer (a discriminative process)? Figure 6 examined this by plotting preference as a function of successively alternating reinforcers in the random-alternation conditions. Just as extended sequences of successive same-alternative reinforcers may have occurred by chance in the random alternation conditions (as depicted in Figure 5), extended sequences consisting of a left-alternative reinforcer followed by a right-alternative reinforcer followed by a left-alternative reinforcer (and so on) may also have occurred in these random alternation conditions. If the continually increasing preference for the just-productive alternative after successive continuations in the random alternation conditions (Figure 5) was due to the these successive same-alternative reinforcers mistakenly signaling a rich left or right patch, we may expect a sequence of alternation reinforcers (i.e., left, right, left, right, etc) of the same length to produce a continually increasing preference for the not-just-productive alternative, or equivalently a continually

decreasing preference for the just-productive alternative. Thus, (assuming that continuation and alternation sequences are equally discriminable) an exclusively discriminative account predicts ever-increasing preference for the not-just-productive alternative as a function of increasing numbers of discontinuation reinforcers.

Preference was largely uninfluenced by the number of preceding alternation reinforcers. There was never any apparent trend for choice to be further towards the left after a right food or further towards the right after a left food as the number of preceding alternations increased. A discontinuation shifted preference towards the just-productive alternative and away from the previously-reinforced alternative regardless of the preceding number of alternations. Thus, the preference trees which plot preference in these conditions as a function of a succession of same-alternative reinforcers (Figure 5) were not likely a result of continually increased responding to the alternative perceived as locally richer (a discriminative process). Such an exclusively discriminative account might predict that just as an increasingly long sequence of same-alternative reinforcers may inadvertently signal a locally rich patch, a similarly long sequence of alternating reinforcers (left, right, left, etc.) may inadvertently signal a period of very small (i.e., one reinforcer long) patches. This did not appear to be the case; Figure 6 may thus suggest a strengthening interpretation to the trends present in Figure 5. Preference appeared to shift continually toward the alternative that provided the preceding reinforcers.

Preference trees can also be plotted as a function of successive postfood keylight onsets. These are presented in Figure 7 for the group data for the two conditions which arranged random alternation of the postfood keylight onset (as in Figure 4). Generally, preference was towards the alternative available after the most recent reinforcer in both Condition 3 (strict alternation) and Condition 4 (random alternation), consistent with the preference pulses. There was no apparent trend for preference to increase as an alternative was illuminated after an increasing number of reinforcers (consistent with the conclusions drawn from the lag analysis, Figure 4). Also, preference after a stimulus onset from the alternative which did not provide the

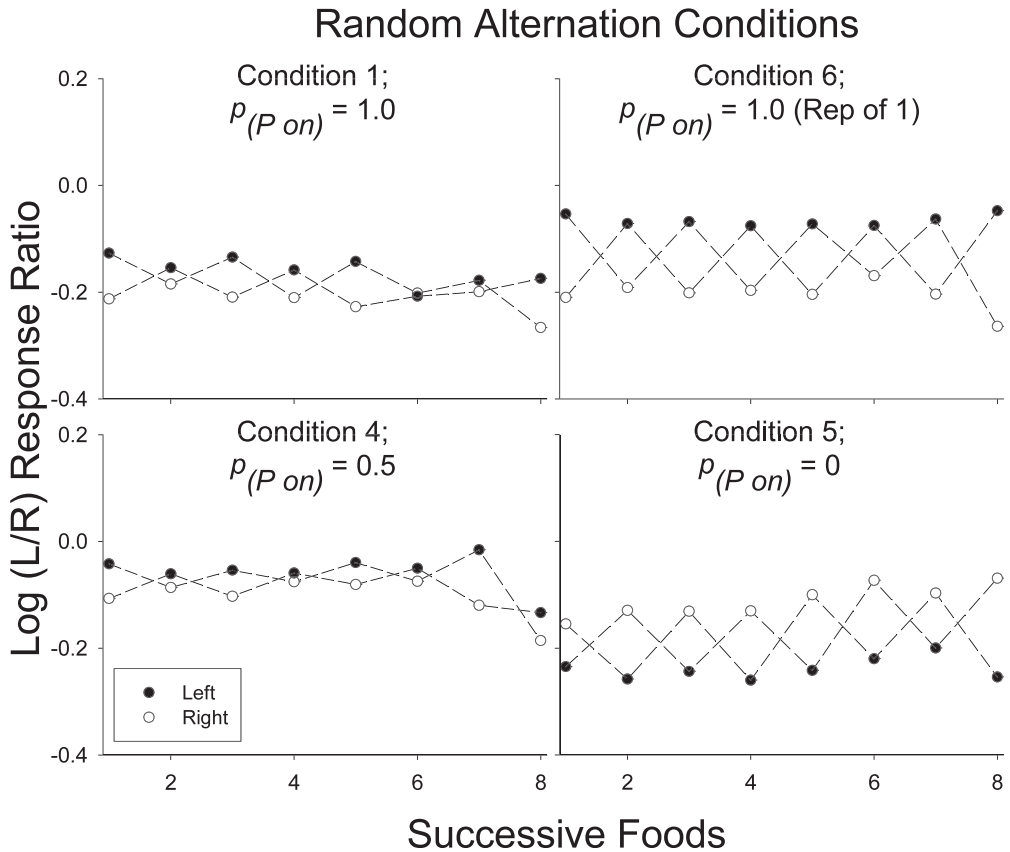


Fig. 6. Group mean log (L/R) response ratio as a function of the number of successive discontinuations in conditions with random alternation of foods where $p(p_{on})$ was varied across conditions.

preceding sequence of stimulus onsets (a stimulus discontinuation), was equal to preference after a stimulus onset from that alternative preceded by a number of same-alternative onsets. Thus, while food deliveries had cumulative and long-lasting effects on behavior, the effect of postfood response availability was constrained to the local postfood level (the preference pulse).

DISCUSSION

Using a switching-key procedure, we varied which key was available immediately after each reinforcer: In some conditions the just-productive alternative was always available ($p(p_{on}) = 1$); in other conditions the not-just-productive alternative was always available ($p(p_{on}) = 0$), and in yet other conditions the immediately-available alternative was unrelated to the last reinforcer's location ($p(p_{on}) = .5$).

Local preference shortly after a food delivery always favored the immediately available alternative (Figures 1 and 2), whether this was the just-productive or the not-just-productive alternative, and regardless of the local probability of further food on that alternative (1, .5, or 0).

In the strict alternation conditions, there was always some postreinforcer preference for the not-just-productive alternative in addition to the initial preference for the immediately-available alternative (Figure 2). This response pattern may have arisen from a simple rule (e.g., "always stay at the first-lit alternative"; "always switch from the first-lit alternative"). However, such a rule would be ineffective when reinforcers strictly alternated and there was no consistent relation between the reinforcer-productive response and the alternative available immediately after reinforcement (Condition 3). The local preference for the not-just-productive alternative that was present

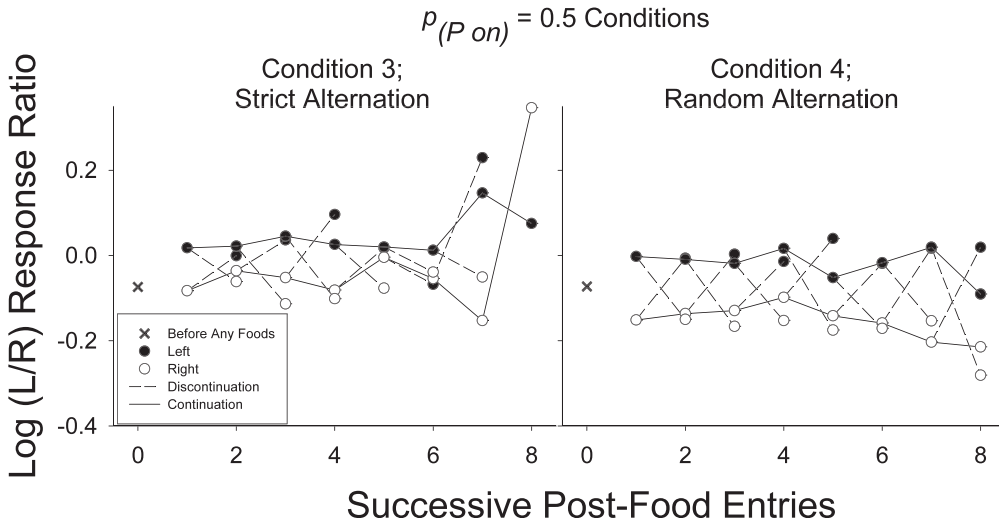


Fig. 7. Group mean log (L/R) response ratio after a series of reinforcers (from either alternative) followed by the same keylight onset or by a keylight onset which differed from the preceding keylight-onset sequence for all of the conditions where $p(P_{on}) = 0.5$.

in Condition 3 suggests some control by the local P/N response ratio.

Why then did the preference pulse vary with $p(P_{on})$ in the strict alternation conditions? In these conditions food was never contingent on a response to the just-productive alternative (i.e. the P/N reinforcer ratio was always 0 at all values of $p(P_{on})$). Perhaps the pigeons failed to discriminate the two different local reinforcer rates, and instead responded to the immediately-available alternative and switched only when food was not forthcoming. While we cannot rule out this interpretation of these particular results, it is inconsistent with a number of other findings. First, as noted above, there was always some preference for the not-just-productive alternative when reinforcers strictly alternated. This implies some control by, and thus discrimination of, the local relative reinforcer rates. Second, the preference trees (discussed below) suggested control by sequences of successive same-alternative reinforcers. Such control implies discrimination of relative reinforcer rates. The effect of $p(P_{on})$ in the strict alternation conditions may alternatively have resulted from imperfect discrimination of the strict-alternation contingencies. Some reinforcers delivered contingent on a response to the not-just-productive alternative may have been misattributed to the just-productive

alternative (Davison & Jenkins, 1985; Davison & Nevin, 1999), thus maintaining responding on that alternative, especially when such responding was favored by the postreinforcer changeover requirements.

When reinforcers randomly alternated, the local reinforcer ratios obtained in each post-reinforcer period likely differed from the experimenter-arranged reinforcer ratios. In the random-alternation conditions, each reinforcer was as likely to be on the just-productive as on the not-just-productive alternative (P/N reinforcer ratio = 1). However, any reinforcer arranged at the shortest IRIs could only be immediately obtained on the first available alternative. A reinforcer arranged on the other alternative could not be obtained until after the changeover requirement was completed. Thus, the local obtained reinforcer ratios did not necessarily equal the arranged 1:1 reinforcer ratio. Rather, they initially favored the alternative not requiring a changeover response access (see Boutros *et al.*, 2011). The postreinforcer preference for the alternative available immediately after each reinforcer (Figure 1) thus reflected these obtained reinforcer ratios. The transient preference for the alternative available immediately after reinforcement seems to indicate a discriminative reinforcer function, or a tendency for the pigeon to respond to the alternative experienced as locally richer.

While the preference pulse seems to reflect a discriminative reinforcer effect, preference at a somewhat more extended level (the reinforcer lag effects, Figure 3 and the reinforcer trees, Figure 5) appears to suggest a strengthening function of reinforcement. These analyses showed increased preference for the just-productive alternative as the number of preceding same-alternative reinforcers increased. This effect was present regardless of which alternative was made available immediately after each reinforcer. There was no similar effect of making the same alternative available after successive reinforcers (Figure 7). Perhaps most compelling was the preference tree in Condition 5 ($p_{(P\ om)} = 0$). Although the preference pulse indicated immediate choice of the not-just-productive (though immediately available) alternative (Figure 1), the preference tree depicted an increasing preference for the reinforcer-productive alternative. This continually increasing preference for the alternative providing successive reinforcers appears to suggest a response-strengthening reinforcer effect.

The preference shifts seen in Figure 5 might alternatively reflect a discriminative reinforcer process: Sequences of same-alternative reinforcers may have (inadvertently) signaled a locally rich patch. While we cannot completely eliminate this possibility, the absence of any trends or tendencies for preference to shift in Figure 6 argues against such an interpretation of the trends in Figure 5. If preference for the just-productive alternative increased as a function of continually increasing numbers of same-alternative reinforcers because those reinforcers (incorrectly) signaled that the just-productive alternative was locally more likely to provide further reinforcement (a discriminative interpretation of these effects), then increasing numbers of alternation reinforcers (i.e., a right reinforcer followed by a left reinforcer followed by a right reinforcer, etc.) should similarly signal that each reinforcer predicts a (momentary) absence of subsequent same-alternative reinforcement and that the not-just-productive alternative is locally more likely to provide the next reinforcer. However, there was no suggestion of increased choice of the not-just-productive alternative with increasing numbers of preceding alternation reinforcers in the random alternation conditions (Figure 6). Thus, the increased choice of the just-productive

alternative with increasing successive reinforcers on that alternative (Figure 5) likely does not reflect the effects of inadvertently signaling a locally richer patch.

This of course assumes that successive alternation reinforcers were as discriminable as successive same-alternative reinforcers. A number of findings suggest that impaired contingency-discriminability cannot itself, account for the asymmetry in alternation versus continuation reinforcer effects. When the not-just-productive alternative was available immediately after a reinforcer in a random-alternation condition (after all reinforcers in Condition 5 and after half of the reinforcers in Condition 4), alternation reinforcers should have been at least as discriminable as stay reinforcers. Choice tracked successive continuation reinforcers and not successive alternation reinforcers even in these conditions where discriminability of alternation sequences should have been enhanced by postreinforcer response availability. Thus, asymmetrical contingency discriminabilities is not likely responsible for the enhanced tracking of successive continuations relative to successive alternation reinforcers, supporting a strengthening interpretation of Figure 5.

These putative response-strengthening properties of the reinforcers were apparently masked by the rather larger discriminative effects when preference was considered only as a function of the most recent reinforcer (the preference pulses; Figure 1). These potential response-strengthening properties, however small and thus not visible at the level of preference after a single reinforcer, were cumulative and enduring, becoming plainly visible in the lag analyses and preference trees. In contrast, the discriminative effects, while larger in the short term (the preference pulses), were short-lived: There was no effect of the alternative illuminated after successive reinforcers at the level of choice in successive IRIs (Figure 7).

Davison and Baum (2006; 2010) and Shahan (2010) recently suggested that both primary and conditioned reinforcer effects may be regarded as entirely discriminative. The increasing tendency to repeat the previously-reinforced response with increasing same-alternative reinforcers (Figure 5) might indicate a direct strengthening function of reinforcement that would contradict a strong

version of this discriminative or signpost hypothesis. Even when the changeover requirements favored the not-just-productive alternative (Condition 5), successive same-alternative reinforcers increased preference to the just-productive alternative. There was no similar increase in choice of the not-just-productive alternative as the number of preceding discontinuations increased (Figure 6), even when choice of this not-just-productive alternative was favored by the postreinforcer changeover requirements. These results suggest a direct strengthening effect of reinforcement, in addition to the discriminative function demonstrated in the preference pulse. Any interpretation of these results purely in terms of a discriminative reinforcer function must account for the absence of any similar discriminative effect when successive reinforcers strictly alternated.

It may be noted that pigeons, as a species, have a known bias for win-stay strategies over win-shift strategies (Randall & Zentall, 1997; Shimp, 1976). Such a bias may be responsible for the different effects of successive continuations (Figure 5) versus successive discontinuations (Figure 6). The local preferences predicted to arise from a win-stay bias overlap considerably with those predicted to arise from a response-strengthening reinforcer effect: A tendency to repeat the prereinforcer response is predicted as a result of both processes. Pigeons may be in some sense "prepared" by their phylogenetic history (Seligman, 1970) to learn that the just-productive response is likely to produce further reinforcement. Thus, behavior appropriate to such contingencies becomes apparent after only a few same-alternative continuations (which occurred simply by chance, Figure 5). On the other hand, pigeons may be "unprepared" to learn that the not-just-productive response is more likely to produce the next reinforcer. Over a period of (again coincidental) discontinuations, there was no suggestion of behavior approximating strict alternation contingencies (Figure 6). However, this should not be taken to imply that pigeons cannot learn to adopt win-shift type strategies. Preference after a reinforcer was clearly towards the not-just-productive alternative in all strict alternation conditions.

While pigeons may be consistently more accurate on win-stay tasks than on win-shift tasks, the opposite is true for nectar-feeding

hummingbirds, which require fewer trials to reach criterion accuracy on win-shift than on win-stay tasks (Cole, Hainsworth, Kamil, Mercier, & Wolf, 1982). This was attributed to the fact that, while pigeons' natural prey tends to appear in "clumped" distributions (one prey predicts others in close proximity), once a flower has been visited, it is unlikely to immediately provide further nectar. A test on an omnivorous species (noisy miners; Sulikowski & Burke, 2010) found that initial (Session 1 of Condition 1) and asymptotic accuracy on a win-shift task were both higher when nectar reinforcers were used than when invertebrate reinforcers were used. This suggests a historical-ecological influence on the tendency either to repeat or avoid a previously reinforced response: Foods typically arranged in clumped arrays support a tendency to repeat the just-productive response in species who feed on these foods, while foods typically arranged singly support a tendency to avoid the just-productive response. Learning the opposite response is not impossible however: In the above experiments, the animals always performed well above chance when the contrary response was required. In the present experiment too, the pigeons always preferred the not-just-productive alternative in all strict-alternation conditions. Thus, the tendency to stay or to shift is by no means immune from modification by consequences in the individual's lifetime. Rather, the historical, species-typical environment may create some predisposition or preparedness, making some contingencies easier to learn than others (Johnston, 1981). The role of these evolutionary contingencies can be understood within a framework emphasizing the discriminative properties of reinforcers: Choice after a reinforcer reflects consequences that have previously followed such reinforcers, both in the individual's personal as well as its evolutionary history.

REFERENCES

- Aparicio, C. F., & Baum, W. M. (2006). Fix and sample with rats in the dynamics of choice. *Journal of the Experimental Analysis of Behavior*, *86*(1), 43–63.
- Aparicio, C. F., & Baum, W. M. (2009). Dynamics of choice: relative rate and amount affect local preference at three different time scales. *Journal of the Experimental Analysis of Behavior*, *91*, 293–317.
- Baum, W. M., & Davison, M. (2004). Choice in a variable environment: Visit patterns in the dynamics of choice. *Journal of the Experimental Analysis of Behavior*, *81*, 85–127.

- Boutros, N., Davison, M., & Elliffe, D. (2011). Contingent stimuli signal subsequent reinforcer ratios. *Journal of the Experimental Analysis of Behavior*, 96, 39–61.
- Cole, S., Hainsworth, F. R., Kamil, A. C., Mercier, T., & Wolf, L. L. (1982). Spatial-learning as an adaptation in hummingbirds. *Science*, 217, 655–657.
- Davison, M., & Baum, W. M. (2000). Choice in a variable environment: Every reinforcer counts. *Journal of the Experimental Analysis of Behavior*, 74(1), 1–24.
- Davison, M., & Baum, W. M. (2002). Choice in a variable environment: Effects of blackout duration and extinction between components. *Journal of the Experimental Analysis of Behavior*, 77(1), 65–89.
- Davison, M., & Baum, W. M. (2006). Do conditional reinforcers count? *Journal of the Experimental Analysis of Behavior*, 86(3), 269–283.
- Davison, M., & Baum, W. M. (2010). Stimulus effects on local preference: stimulus-response contingencies, stimulus-food pairing, and stimulus food correlation. *Journal of the Experimental Analysis of Behavior*, 93, 45–59.
- Davison, M., Elliffe, D., & Marr, M. J. (2010). The effects of a local negative feedback function between choice and relative reinforcer rate. *Journal of the Experimental Analysis of Behavior*, 94, 197–207.
- Davison, M., & Jenkins, P. E. (1985). Stimulus discriminability, contingency discriminability, and schedule performance. *Animal Learning & Behavior*, 13, 77–84.
- Davison, M., & Nevin, J. A. (1999). Stimuli, reinforcers, and behavior: an integration. *Journal of the Experimental Analysis of Behavior*, 71, 439–482.
- Elliffe, D. (2006). *Preference pulses: new analyses, new questions*. Paper presented at the 32nd Annual Convention of the Association for Behavior Analysis.
- Elliffe, D., & Davison, M. (2010). Four-alternative choice violates the constant-ratio rule. *Behavioural Processes*, 84, 381–389.
- Hearst, E. (1962). Delayed alternation in the pigeon. *Journal of the Experimental Analysis of Behavior*, 5, 225–228.
- Johnston, T. D. (1981). Contrasting approaches to a theory of learning. *Behavioral and Brain Sciences*, 4, 125–138.
- Krägeloh, C. U., Davison, M., Elliffe, D., & M. (2005). Local preference in concurrent schedules: The effects of reinforcer sequences. *Journal of the Experimental Analysis of Behavior*, 84(1), 37–64.
- Krägeloh, C. U., Zapanta, A. E., Shepherd, D., & Landon, J. (2010). Human choice behaviour in a frequently changing environment. *Behavioural Processes*, 83, 119–126.
- Landon, J., & Davison, M. (2001). Reinforcer-ratio variation and its effects on rate of adaptation. *Journal of the Experimental Analysis of Behavior*, 75(2), 207–234.
- Landon, J., Davison, M., & Elliffe, D. (2002). Concurrent schedules: Short- and long-term effects of reinforcers. *Journal of the Experimental Analysis of Behavior*, 77(3), 257–271.
- Landon, J., Davison, M., & Elliffe, D. (2003). Choice in a variable environment: Effects of unequal reinforcer distributions. *Journal of the Experimental Analysis of Behavior*, 80(2), 187–204.
- Landon, J., Davison, M., Krägeloh, C. U., Thompson, N. M., Miles, J. L., Vickers, M. H., & ...Breier, B. H. (2007). Global undernutrition during gestation influences learning during adult life. *Learning & Behavior*, 35, 79–86.
- Lie, C., Harper, D. N., & Hunt, M. (2009). Human performance on a two-alternative rapid-acquisition task. *Behavioural Processes*, 81, 244–249.
- Menlove, R. L. (1975). Local patterns of responding maintained by concurrent and multiple schedules. *Journal of the Experimental Analysis of Behavior*, 23, 309–337.
- Randall, C. K., & Zentall, T. R. (1997). Win-stay/lose-shift and win-shift/lose-stay learning by pigeons in the absence of overt response mediation. *Behavioural Processes*, 41, 227–236.
- Rodewald, A. M., Hughes, C. E., & Pitts, R. C. (2010a). Choice in a variable environment: Effects of *d*-amphetamine on sensitivity to reinforcement. *Behavioural Processes*, 84, 460–464.
- Rodewald, A. M., Hughes, C. E., & Pitts, R. C. (2010b). Development and maintenance of choice in a dynamic environment. *Journal of the Experimental Analysis of Behavior*, 94, 175–195.
- Seligman, M. E. P. (1970). On the generality of the laws of learning. *Psychological Review*, 77, 406–418.
- Shahan, T. A. (2010). Conditioned reinforcement and response strength. *Journal of the Experimental Analysis of Behavior*, 93, 269–289.
- Shimp, C. P. (1976). Short-term memory in the pigeon: the previously reinforced response. *Journal of the Experimental Analysis of Behavior*, 26, 487–493.
- Skinner, B. F. (1938). *The Behavior of Organisms: An Experimental Analysis*. New York: Appleton-Century-Crofts.
- Sulikowski, D., & Burke, D. (2010). Reward type influences performance and search structure of an omnivorous bird in an open-field maze. *Behavioural Processes*, 83, 31–35.
- Williams, B. A. (1971a). Color alternation learning in the pigeon under fixed-ratio schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 15, 129–140.
- Williams, B. A. (1971b). Non-spatial delayed alternation by the pigeon. *Journal of the Experimental Analysis of Behavior*, 16, 15–21.

Received: September 19, 2010

Final Acceptance: April 25, 2011