Pigeons' use of cues in a repeated five-trial-sequence, single-reversal task

Rebecca M. Rayburn-Reeves · Thomas R. Zentall

Abstract We studied behavioral flexibility, or the ability to modify one's behavior in accordance with the changing environment, in pigeons using a reversal-learning paradigm. In two experiments, each session consisted of a series of five-trial sequences involving a simple simultaneous color discrimination in which a reversal could occur during each sequence. The ideal strategy would be to start each sequence with a choice of S1 (the first correct stimulus) until it was no longer correct, and then to switch to S2 (the second correct stimulus), thus utilizing cues provided by local reinforcement (feedback from the preceding trial). In both experiments, subjects showed little evidence of using local reinforcement cues, but instead used the mean probabilities of reinforcement for S1 and S2 on each trial within each sequence. That is, subjects showed remarkably similar behavior, regardless of where (or, in Exp. 2, whether) a reversal occurred during a given sequence. Therefore, subjects appeared to be relatively insensitive to the consequences of responses (local feedback) and were not able to maximize reinforcement. The fact that pigeons did not use the more optimal feedback afforded by recent reinforcement contingencies to maximize their reinforcement has implications for their use of flexible response strategies under reversal-learning conditions.

Keywords Pigeons · Serial learning/memory · Behavioral flexibility · Reversal learning · Discrimination learning · Probability learning · Timing

Behavioral flexibility, described by Bond, Kamil, and Balda (2007), is the ability to respond rapidly to environmental

change and to be ready to seek out alternative solutions to problems encountered if the strategies initially used are not effective. Stenhouse (1974) made use of the term *flexibility* when he described intelligence as "the built-in flexibility that allows individual organisms to adjust their behavior to relatively rapidly changing environments" (p. 61). The ability of an animal to behave flexibly and to adjust rapidly to environmental change is largely dependent on the ability of the animal to use the available environmental cues. Of particular interest is the degree to which an animal can base responding on the most relevant cues provided. The ability of animals to track changes in reinforcement rate over time has received considerable study. For example, studies using concurrent variable-interval schedules of reinforcement in which the overall rate of reinforcement differs between two choice alternatives have revealed that subjects will tend to allocate their responses in proportion to the relative rate of reinforcement obtained from the two alternatives, thereby matching their relative rate of responding to that of the relative rate of reinforcement (Baum, 1974; Herrnstein, 1961). Furthermore, the rate at which animals learn to adjust their response rate to the changing rate of reinforcement is related to the difference in reinforcement rates between the two alternatives (Bailey & Mazur, 1990).

There is also evidence that animals can use local cues on which to base their behavior. For example, pigeons will tend to repeat a just-reinforced response (sometimes referred to as a *preference pulse*), and this tendency also depends on the richness or leanness of the schedule that is in effect (Bailey & Mazur, 1990; Davison & Baum, 2003). Thus, it appears that pigeons can be sensitive to the local contingencies (e.g., what response was just made and whether it was reinforced), while also being sensitive to the overall changes in reinforcement rate between two alternatives over time. A question that may arise from this finding is whether sensitivity to global cues (e.g., the overall probability of reinforcement) is

R. M. Rayburn-Reeves · T. R. Zentall (✉)
Department of Psychology, University of Kentucky,
Lexington, KY 40506-0044, USA
e-mail: zentall@uky.edu

greater than sensitivity to the information provided by local cues (e.g., recent reinforcement contingencies) when both are available as reliable cues. In one study, Krägeloh, Davison, and Elliffe (2005) manipulated the number of sequential reinforcers on the same key using a concurrent variable-interval schedule while keeping the overall rate of reinforcement constant over the two alternatives; they found a tendency for pigeons to repeat the just-reinforced response when the same-key contingencies were long, and that this effect dissipated as the string of reinforcement was reduced. This research suggests that, although pigeons are sensitive to the local information provided by recent reinforcement contingencies, their behavior is also controlled by the overall rate of reinforcement.

Cowie, Davison, and Elliffe (2011) discussed the fact that reinforcement will increase the probability of the response that leads to it only if the "past and future contingencies are the same" (p. 63). Furthermore, they argued that when past and future contingencies differ, control of choice is maintained by future contingencies. Therefore, although traditional theories of learning suggest that reinforcing a behavior strengthens the likelihood of that behavior occurring in the future, reinforcement can also signal that a different behavior will lead to reinforcement.

The passage of time can also serve as a cue for when and where to respond. Wilkie and Willson (1992) found that if a variable-interval schedule of reinforcement was programmed for periods of 15 min sequentially on each of four response keys, pigeons used the passage of time to move from key to key. Although the pigeons could have used either the absence of reinforcement for responding to the current key for an extended period or the reinforcement obtained from occasionally testing the next key in the sequence (i.e., local cues), they apparently primarily used the passage of time (i.e., a global cue) as the predominant cue to switch keys.

With intermittent schedules of reinforcement, however, the absence of reinforcement may provide an ambiguous cue as to whether to continue with the current key or switch to the next. To determine whether animals can use the local feedback (reinforcement or its absence) from their behavior as an unambiguous cue to stay or to switch, researchers have used discrete trials with continuous reinforcement for a correct choice. In *serial reversal-learning task*s, one alternative is defined as correct until consistent responding is found, and then the other alternative is defined as correct. When consistent responding is again obtained, the original alternative once again becomes correct, and so on (see Mackintosh, McGonigle, Holgate, & Vanderver, 1968). The degree to which the number of trials to criterion decreases with successive reversals can be taken as a measure of behavioral flexibility (Bitterman, 1975), and it also suggests a sensitivity to local cues.

The improvement in the number of trials needed to acquire a reversal has been attributed to the ability of an animal to use feedback from the previous trials as a cue to reverse behavior. With these serial reversal tasks, in general, the only cue that the animal has that the previously reinforced response is no longer correct is the feedback from current responding.

A variation of the serial reversal task that has been used recently with pigeons (Cook & Rosen, 2010; Rayburn-Reeves, Molet, & Zentall, 2011; Rayburn-Reeves, Stagner, Kirk, & Zentall, 2012) is the *midsession reversal task*, in which for the first half of each session, choices of one discriminative stimulus (S1) are reinforced (S+), but not those of S2 (S–), and for the remainder of the session, choices of the other discriminative stimulus (S2) are reinforced, but not choices of S1. This reversal task provides two different cues for the occurrence of the reversal: a global cue, in the form of time or trial number estimation to the midpoint of the session, and a local cue, in the form of the outcome of the previous trial or trials.

Rayburn-Reeves et al. (2011) trained pigeons on a midsession reversal task involving red and green hues. During the first half of each 80-trial session, responses to S1 were reinforced and responses to S2 were not (S1+, S2–). During the last half of the session, the contingencies were reversed. After the pigeons' performance had stabilized (50 sessions of training), it appeared that the pigeons were using the passage of time (or trial number) into the session as the primary cue to reverse, rather than the feedback from the preceding trial(s). That is, they began to respond to S2 prior to the change in contingency (they made anticipatory errors) and also maintained responding to S1 after the change in contingency (they made perseverative errors). That is, they developed what can be described as a psychophysical function when plotting the probability of a response to S1 as a function of trial number. Specifically, the pigeons showed a decline in accuracy between Trials 31 and 40, and similarly poor accuracy between Trials 41 and 50. Importantly, there was a more valid cue available in the outcome of the previous trial (the local cue). Had the pigeons used the outcome of their choice on the previous trial as a cue (a win-stay/lose-shift rule), they could have obtained reinforcement on every trial except for the first reversal trial, which would have resulted in almost 99 % reinforcement. Although once performance had stabilized (after about 20 sessions of training) the overall accuracy was about 90 % correct, the fact that anticipatory and perseverative errors persisted suggests that the use of temporal (or number estimation) cues may be difficult to abandon. It appears that under these conditions, the more global cue (time or trial number estimation) may be a more natural cue for pigeons than the outcome of the most recent trial.

Surprisingly, even when timing cues were made unreliable by varying them across sessions (e.g., in Session 1, reverse on Trial 11; in Session 2, reverse on Trial 56; etc.), pigeons tended to average the point of the reversal over sessions (Rayburn-Reeves et al., 2011, Exp. 2). That is, they made an especially large number of perseverative errors when the reversal occurred early in the session, and they made an especially large number of anticipatory errors when the reversal occurred late in the session. In fact, the pigeons were most accurate when the reversal occurred in the middle of the session.

It may be that the pigeons continued to use molar cues to reverse rather than the more efficient local cues because the cost of using those cues was not sufficiently great. That is, the loss of approximately 10 % reinforcement was not sufficiently great to encourage the pigeons to switch from using the passage of time to relying more heavily on feedback from the preceding trial. Specifically, each incorrect response resulted in only a small reduction in the probability of reinforcement. Perhaps pigeons would be encouraged to use the feedback from the preceding trial as a cue to stay or shift if (1) the number of trials per session was reduced because the cost of an error, as a proportion of the trials in a session, would be greater, and (2) the point of the reversal was unpredictable.

In the present experiments, each session consisted of a series of five-trial sequences, to emphasize the use of cues provided by the outcome of the preceding trial (i.e., the local history of reinforcement). With a five-trial sequence, an incorrect response would result in a reduction of 20 % of the possible reinforcements, a change that might increase the saliency of the reinforcement contingencies on each trial. To make the sequences highly discriminable, we inserted relatively long intervals between sequences. With 16 five-trial sequences per session, and with one reversal per sequence, it was possible to increase the number of reversal events while still keeping the number of trials per session at 80. In this research, we asked whether the increase in exposure to reversal events would possibly allow the subjects to learn to use the change in contingencies as an immediate cue to reverse.

Experiment 1

The purpose of Experiment 1 was to ask whether decreasing the number of trials per sequence from 80 to 5 (to make the cost of an error greater) and varying the point in the sequence at which the reversal would occur (to make timing or counting a less valid cue) would encourage the use of the feedback from the previous trial as a cue to reverse. In Experiment 1, S1 was always correct on Trial 1 of every sequence, and S2 was always correct on Trial 5. Thus, the

reversal point could occur at one of four different locations in each sequence (between Trials 1 and 2, 2 and 3, 3 and 4, or 4 and 5).

Method

Subjects Four White Carneaux pigeons (*Columbia livia*) and two homing pigeons (*Columbia livia*) served as subjects. The White Carneaux pigeons ranged in age from 2 to 12 years old, while the homing pigeons were approximately 1 year old at the start of the experiment. All of the subjects had experience in previous, unrelated studies involving successive discriminations of the colors red, green, yellow, and blue, but they had not been exposed to a reversal learning task. The pigeons were maintained at 85 % of their free-feeding weight and were housed individually in wire cages, with free access to water and grit in a colony room that was maintained on a 12-h/12-h light/dark cycle. The pigeons were maintained in accordance with a protocol approved by the Institutional Animal Care and Use Committee at the University of Kentucky.

Apparatus The experiment was conducted in a BRS/LVE (Laurel, MD) sound-attenuating operant test chamber measuring 34 cm high, 30 cm from the response panel to the back wall, and 35 cm across the response panel. Three circular response keys (3 cm in diameter) were aligned horizontally on the response panel and were separated from each other by 6.0 cm, but only the side response keys were used in these experiments. The bottom edge of the response keys was 24 cm from the wire-mesh floor. A 12-stimulus in-line projector (Industrial Electronics Engineering, Van Nuys, CA) with 28-V, 0.1-A lamps (GE 1820) that projected red and green hues (Kodak Wratten Filter Nos. 26 and 60, respectively) was mounted behind each response key. Mixed-grain reinforcement (Purina Pro Grains—a mixture of corn, wheat, peas, kafir, and vetch) was provided from a raised and illuminated grain feeder located behind a horizontally centered 5.1 × 5.7 cm aperture, which was located vertically midway between the response keys and the floor of the chamber. The reinforcement consisted of 2-s access to mixed grain. A white houselight, which provided general illumination between sequences, was located in the center of the ceiling of the chamber. The experiment was controlled by a microcomputer and interface located in an adjacent room.

Procedure At the start of each sequence, one side key was illuminated red, and the other green. The locations of the hues (left vs. right) varied randomly from trial to trial. The red and green hues were randomly assigned to each subject as S1 and S2, with the constraint that for half of the subjects, red was designated as S1 and green as S2, and for the other

half, green was S1 and red was S2. In a given sequence, the reversal point occurred randomly on one of four trials in each five-trial sequence (on Trial 2, 3, 4, or 5). Over the course of each session, the probabilities of reinforcement for choices of S1 and S2 were equated (see Table 1). A single response to the correct stimulus resulted in both stimuli turning off and in 2-s access to grain, followed by a 3-s dark intertrial interval (ITI), whereas a single response to the incorrect stimulus turned off both stimuli and resulted in a 5-s dark ITI. The next trial started immediately following the ITI. Each five-trial sequence was separated by a 1-min intersequence interval, during which the houselight was illuminated. Subjects were trained with this variable-reversal procedure for 16 sequences per daily session (four of each sequence type in random order) for a total of 60 sessions (240 sequences at each reversal point, for a total of 960 reversals).

Results and discussion

It was anticipated that by reducing the number of trials per sequence, the pigeons' behavior might be more sensitive to the outcome of the preceding trial. However, the results of Experiment 1 indicated that when pigeons were given five-trial sequences in which the point of the reversal was unpredictable but S1 was always correct on Trial 1 and S2 was always correct on Trial 5, regardless of the trial on which the reversal occurred, the percentage of S1 choices by the pigeons declined systematically with the trial number in the sequence (see Fig. 1). Figure 1 shows the percentages of choices of S1 as a function of trial number in the sequence, averaged across subjects for Sessions 41–60 combined (a total of 80 reversals for each pigeon at each reversal point; 240 reversals per subject). The data were pooled over Sessions 41–60 because choices as a function of sequence type had stabilized. Whether the reversal occurred on
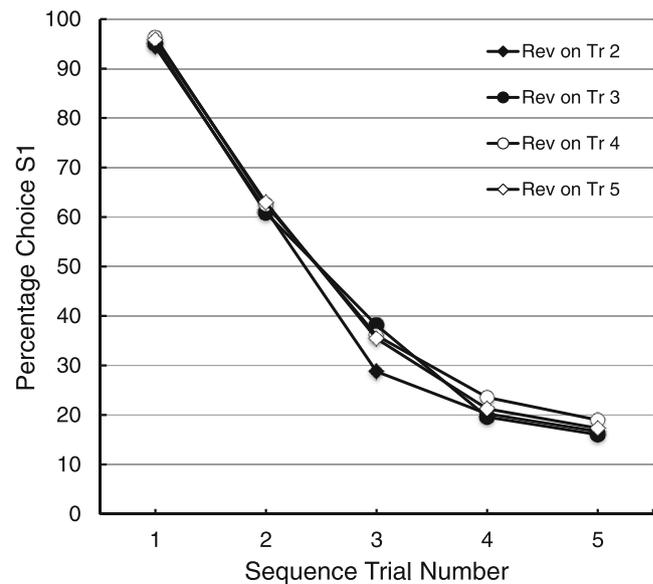
Fig. 1 Experiment 1: Percentage choices of Stimulus 1 (S1) as a function of trial in the sequence, averaged over subjects and over Sessions 41–60 and plotted separately for sequences on which the reversal occurred on Trial 2, 3, 4, or 5

Trial 2, 3, 4, or 5, the shapes of the choice functions across reversal locations were almost identical. A repeated measures analysis of variance (ANOVA) was conducted on the four reversal locations as a function of the five trials within a sequence, indicating a significant main effect of trial number $F(4, 20) = 51.22$, $p < .01$, but no significant effect of reversal location, $F(3, 15) = 2.04$, $p = .15$, nor was there a significant Trial Number × Reversal Location interaction, $F < 1$. Thus, the pigeons' behavior was relatively insensitive to the relation with the outcome of the preceding trial, as evidenced by their choices on the current trial. Overall, for Sessions 41–60, the mean accuracy for the pigeons was 70.0 % (range 63.1 % to 75.5 %), significantly worse than the 80 %

Table 1 Experiment 1: Probability of Stimulus 1 (S1) being correct as a function of sequence type and trial number

Sequence Type	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Percent S1 Correct by Sequence	Mean Percent Correct
1	S1	S1	S1	S1	S2	80 %	
2	S1	S1	S1	S2	S2	60 %	
3	S1	S1	S2	S2	S2	40 %	
4	S1	S2	S2	S2	S2	20 %	
Percent S1 Correct	100	75	50	25	0		50 %
Ideal Per. Corr. Dep.	80	80	80	80	80		80 %
Ideal Per. Corr. Ind.	100	75	50	75	100		80 %
Ideal Per. Match	100	62	50	62	100		75 %

S2, Stimulus 2; Ideal Per. Corr. Dep., ideal percentage correct if pigeon is sensitive to the win-stay/lose-shift contingencies; Ideal Per. Corr. Ind., ideal percentage correct if pigeon is sensitive to the likelihood of S1 being correct at each trial in the sequence, independent of the outcome on the preceding trial; Ideal Per. Match, ideal percentage correct if pigeon is matching its choice to the likelihoods of S1 and S2 being correct

that could have been obtained with a win-stay/lose-shift rule $t(5) = -4.9$, $p < .01$.

Because of the apparent independence of the choice functions resulting from where the reversal occurred within a sequence, we pooled the data over reversal locations to assess the mean percentages of S1 choices on each trial (see Fig. 2, open circles) and compared them with the probabilities of reinforcement given the choice of S1 on each trial, independent of the feedback from the preceding trial (Fig. 2, open squares). A single-sample t test was conducted on the data from each trial, relative to the mean overall probability that S1 was correct, independent of the location of the reversal in the sequence. Although the difference was relatively small, the mean choice of S1 (mean 95.4 %, *SEM* 0.46 %) was significantly lower than the ideal mean for Trial 1 (100 %), given that S1 was always correct on Trial 1, $t(5) = 10.09$, $p < .01$; however, the mean choice of S1 on Trial 2 (mean 61.9 %, *SEM* 7.28 %) was not significantly different from the probability that choice of S1 was correct (75 %) on Trial 2, $t(5) = 1.80$, $p = .13$. On Trial 3, the mean choice of S1 (mean 34.6 %, *SEM* 5.26 %) was again significantly lower than the probability that choice of S1 was correct (50 %), $t(5) = 2.93$, $p = .03$; however, on Trial 4, the mean choice of S1 (mean 21.2 %, *SEM* 4.23 %) was not significantly different from the probability that S1 was correct (25 %), $t(5) = 0.91$, $p = .40$. Finally, on Trial 5, the average choice of S1 (mean 17.24 %, *SEM* 4.42 %) was significantly higher than the probability that S1 was correct (0 %), $t(5) = 3.90$, $p = .01$.
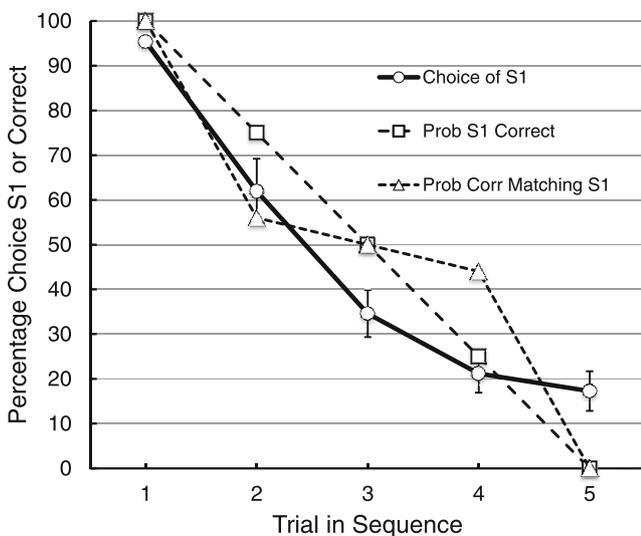
Fig. 2 Experiment 1: Percentage choices of Stimulus 1 (S1) as a function of trial in the sequence, pooled over sequence types and averaged over subjects for Sessions 41–60. Squares represent the probabilities of reinforcement for a choice of S1 as a function of trial number, independent of reversal locations. Triangles represent the probabilities of reinforcement for a choice of S1 as a function of trial number, if subjects were matching their choices of S1 to the probability of reinforcement for that choice

Also presented in Fig. 2 is the probability correct if subjects had been matching, leading to choices of S1 in proportion to the probability of reinforcement for choosing that stimulus (open triangles).

In Fig. 2, one can see that the greatest deviations from probability matching were on Trials 2 and 3 (although the difference was not significant on Trial 2). On these trials, the pigeons were choosing S1 less often than expected if they had been matching the probability that S1 would be correct. On these trials, it appears that the pigeons were anticipating the reversal (see Rayburn-Reeves et al., 2011, for a similar effect). The only other trial on which there was a large deviation from probability matching was Trial 5, on which the pigeons chose S1 more than they should have, given that S1 was never correct on Trial 5. However, it is quite likely that the pigeons were not able to estimate the trial number accurately enough to be sure when they were on Trial 5. Previous research on counting by pigeons has shown that under similar conditions, pigeons have difficulty discriminating more than three sequential events (Rayburn-Reeves, Miller, & Zentall, 2010). In fact, in the present study, the similarity in choices of S1 on Trials 4 and 5 suggests that a lack of discriminability may well have been the case.

If subjects had begun each sequence with a response to S1 and had used the feedback from the preceding trial as a cue to how to respond on the current trial (i.e., they had adopted a win-stay/lose-shift response), they would have received reinforcement on 80 % of the trials (refer to Table 1), and the functions for each of the reversal locations would not have fallen on top of each other. On the other hand, probability matching would have reduced the probability of reinforcement to 74.8 %, accuracy that was just a bit better than what the pigeons actually received (70.0 %).

Another interesting finding is that errors tended to be more anticipatory in nature, rather than perseverative. That is, subjects tended to respond to S2 prior to the reversal (as has been reliably shown in the midsession and variable-reversal tasks previously reported), thereby not profiting from their errors in terms of the feedback that an error provided. Unlike perseverative errors, in the present task, anticipatory errors did not indicate how to respond on the subsequent trial; they only indicated whether the previous response was correct. Perseverative errors, on the other hand, would indicate not only that the previous response was incorrect, but that choice of the other alternative would now be correct for the remainder of the sequence.

The results of Experiment 1 indicated that with five-trial sequences, when the reversal location varied in an unpredictable manner and when S1 was always correct on Trial 1 and S2 always correct on Trial 5, pigeons did not use the information afforded by the local contingencies of reinforcement as a basis for choice of responses to S1 or S2 on

subsequent trials. Instead, it appears that pigeons judge the likelihood that S1 is correct by estimating the average probability of reinforcement of S1 over many training sequences, and then use the passage of time or trial number as a basis for that choice.

Experiment 2

In Experiment 1, because a reversal always occurred during a sequence, the likelihood of a reversal increased from 0 % to 100 % as the pigeon progressed through the sequence (if it had not occurred by Trial 4, it would certainly occur on Trial 5; see Table 1). The results of Experiment 1 suggested that pigeons began to anticipate the reversal by choosing S2 more often than S1 as the number of trials in a sequence increased. Thus, pigeons in Experiment 1 did not adopt a win-stay/lose-shift response strategy. In Experiment 2, we included two additional sequence types in which a reversal did not occur: one in which a reversal never occurred (i.e., S1 remained correct throughout the entire sequence), and a second in which S2 was correct for the entire sequence (see Table 2). By adding these two sequence types, neither S1 nor S2 was correct with certainty on any trial in the sequence. With this change, we hypothesized that subjects might be more inclined to use the feedback from the preceding trial as a basis for the choice on the current trial, because neither response would produce reinforcement with certainty on any trial. In addition, we thought that this change might help to reduce the number of anticipatory errors while also equating the overall probabilities correct of S1 and S2. With the addition of these two sequence types, if pigeons were to adopt a win-stay/lose-shift response strategy, reinforcement would be provided on 83.3 % of the trials.

Method

Subjects Four White Carneaux pigeons (*Columbia livia*), ranging in age from 2 to 12 years, and two homing pigeons (*Columbia livia*), which were approximately 1 year old at the start of the experiment, served as the subjects. All subjects had had previous experience similar to that of the pigeons in Experiment 1. The subjects were housed and maintained in the same manner as in Experiment 1.

Apparatus The experiment was conducted using the same apparatus as in Experiment 1.

Procedure The procedure was the same as in Experiment 1, with the exception that two sequence types were added: one in which the sequence started with S2 correct and S2 remained correct throughout the sequence, and the other in which S1 remained correct throughout the sequence. Subjects were tested on the variable-reversal paradigm for 18 sequences per session (three of each of the six sequence types in each session), for a total of 60 sessions (180 sequences of each sequence type; a total of 1,080 sequences).

Results and discussion

The results of Experiment 2 indicated that when pigeons were trained with sequences in which the point of the reversal was made variable, with no certainty of S1 or S2 being correct on any given trial, the patterns of choices of S1 were still quite similar to those of Experiment 1 (see Fig. 3). Once again, choice of S1 on Trial 1 was similar to the probability of reinforcement associated with that alternative. And once again, on Trials 2 and 3, the pigeons chose S1 less often than

Table 2 Experiment 2: Probability of Stimulus 1 (S1) being correct as a function of sequence type and trial number

Sequence Type	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Percent S1 Correct by Sequence	Mean Percent Correct
1	S1	S1	S1	S1	S1	100 %	
2	S1	S1	S1	S1	S2	80 %	
3	S1	S1	S1	S2	S2	60 %	
4	S1	S1	S2	S2	S2	40 %	
5	S1	S2	S2	S2	S2	20 %	
6	S2	S2	S2	S2	S2	0 %	
Percent S1 Correct	83	67	50	33	17		50 %
Ideal Per. Corr. Dep.	83	83	83	83	83		83 %
Ideal Per. Corr. Ind.	83	67	50	67	83		70 %
Ideal Per. Match	72	56	50	56	72		61 %

S2, Stimulus 2; Ideal Per. Corr. Dep., ideal percentage correct if pigeon is sensitive to the win-stay/lose-shift contingencies; Ideal Per. Corr. Ind., ideal percentage correct if pigeon is sensitive to the likelihood of S1 being correct at each trial in the sequence, independent of the outcome on the preceding trial; Ideal Per. Match, ideal percentage correct if pigeon is matching its choice to the likelihoods of S1 and S2 being correct
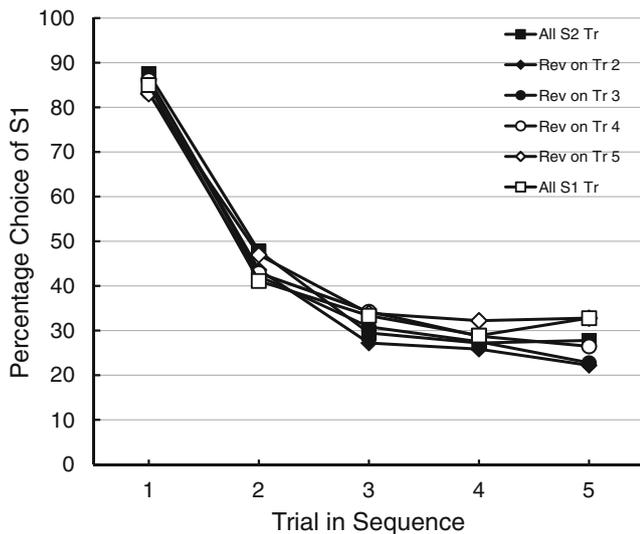
Fig. 3 Experiment 2: Percentage choices of Stimulus 1 (S1) as a function of trial in the sequence, averaged over subjects and over Sessions 41–60 and plotted separately for each sequence type

Fig. 4 Experiment 2: Percentage choices of Stimulus 1 (S1) as a function of trial in the sequence, pooled over sequence types and averaged over subjects for Sessions 41–60. Squares represent the probabilities of reinforcement for a choice of S1 as a function of trial number, independent of reversal locations. Triangles represent the probabilities of reinforcement for a choice of S1 as a function of trial number, if subjects were matching their choices of S1 to the probability of reinforcement for that choice

the overall probabilities of reinforcement associated with those alternatives. Consistent with the results of Experiment 1, on Trial 4, the pigeons chose S1 at about the same rate as the probability of reinforcement associated with that alternative, and on Trial 5, the pigeons chose S1 less often than would have been expected had they been matching the probability of reinforcement associated with that alternative.

A repeated measures ANOVA conducted on the six sequence types as a function of the five trials within a sequence indicated, once again, a significant main effect of trial number, $F(4, 20) = 28.76$, $p < .01$, but no significant main effect of reversal location, $F(5, 25) = 2.10$, $p = .10$, nor a significant Sequence Type × Reversal Location interaction, $F(20, 100) = 1.4$, $p = .13$.

Again, the data were pooled over reversal locations to assess the average percentage choices of S1 on each trial, as compared with a measure of the overall probability that S1 was correct as a function of trial number (see Fig. 4). The probabilities for correct S1 responses on each trial in the sequence, pooled over the six sequence types (i.e., without regard for the feedback from reinforcement or its absence on preceding trials in the sequence), are indicated by the open squares in Fig. 4. A single-sample *t* test was conducted on the data from each trial relative to the percentages of the time that S1 was correct. The mean percentage choice of S1 (mean 85 %, *SEM* 6%) was not significantly different from the mean percentage of reinforcement associated with Trial 1 (83 %), $t(5) = 0.30$, $p > .10$; however, on Trial 2, the mean percentage choice of S1 (mean 44 %, *SEM* 5 %) was significantly lower than the overall percentage that S1 was correct on that trial (67 %), $t(5) = 4.53$, $p < .01$ and on Trial
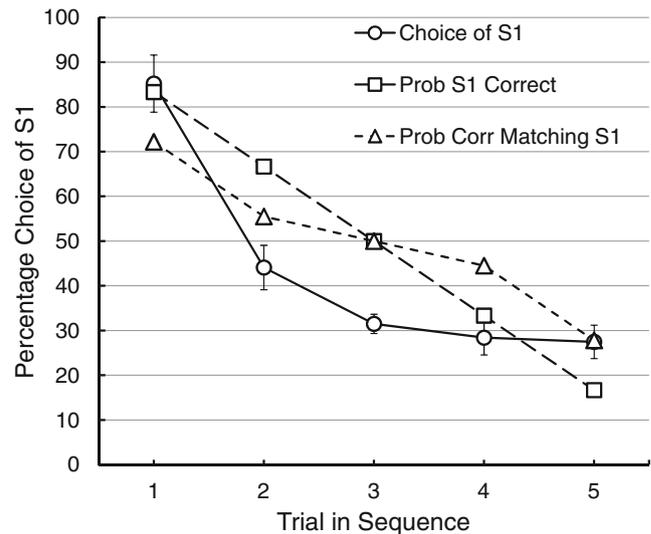
3, the mean percentage choice of S1 (mean 31 %, *SEM* 2 %) was also significantly lower than the overall percentage that S1 was correct on that trial (50 %), $t(5) = 8.61$, $p < .01$. On Trial 4, the mean percentage choice of S1 (mean 28 %, *SEM* 4 %) was not significantly different from the overall percentage that S1 was correct on that trial (33 %), $t(5) = 1.28$, $p > .10$. Finally, on Trial 5, the mean percentage choice of S1 (mean 27 %, *SEM* 4 %) was significantly higher than the overall percentage that S1 was correct on that trial (17 %), $t(5) = 2.89$, $p < .05$. Also presented in Fig. 4 are the percentages correct if subjects had been matching, and thus choosing S1 in proportion to the probability of reinforcement for that choice (open triangles).

In Experiment 2, if subjects had adopted the win-stay/lose-shift rule with this task, they would have received a mean of 83.3 % correct; however, the pigeons' performance was significantly worse (59.5 %), $t(5) = 16.7$, $p < .01$. Thus, when the reinforcement associated with S1 was less certain on Trial 1 (by including sequences with S2 correct on Trial 1) and when the reinforcement associated with S2 was less certain on Trial 5 (by including sequences with S1 correct on Trial 5), the use of a win-stay/lose-shift rule was not encouraged. Instead, the pigeons continued to use the overall probability of reinforcement associated with each trial in the sequence, together with a consistent tendency to anticipate the reversal (on Trials 2 and 3) rather than using the feedback (reinforcement or its absence) from the preceding trial. Thus, for each subject the percentage of S1 choices

over trials for the different sequence types varied very little. Therefore, as in Experiment 1, subjects were not treating the sequence types independently of one another.

The average choice of S1 on Trial 1 (85.2 %) was not significantly different from the overall percentage that S1 was correct (83.3 %). Thus, subjects were matching the overall percentage of correct reinforcement on this trial. The average choices of S1 on Trials 2 and 3 showed significant anticipation of S2, relative to the percentage that S2 was actually correct on those trials, indicating a significant anticipatory choice of S2. On Trial 4, the average choice of S2 did not differ significantly from the overall percentage correct for S2 over sequence types. Finally, the average choice of S2 on Trial 5 was significantly higher than the overall percentage correct for S2 choices, indicating a possible lack of discriminability of trial number at that point in the sequence.

The results of Experiment 2 indicated that, when the location of the reversal was both variable and uncertain on all trials, pigeons still did not adopt a win-stay/lose shift rule, but instead five out of the six pigeons appeared to learn that S1 was correct on the first trial and that S2 became correct soon thereafter.

General discussion

The purpose of the present reversal experiments was to explore procedures that might enhance pigeons' sensitivity to the consequences of prior choices in a sequence. Specifically, the goal was to increase the importance of the response–outcome association on the preceding trial as a cue for responding. That is, could we encourage pigeons to use the local reinforcement history (i.e., the response and outcome from the preceding trial) as the basis for their choice on the current trial? In previous research, we had trained pigeons with 80-trial sessions. With that procedure, errors prior to the reversal and following the reversal had only a small effect on the overall rate of reinforcement (pigeons were often correct on 90 % of the trials). In the present research, with five-trial sequences, each error represented a 20 % error rate. In both of the present experiments, pigeons showed little use of the cues provided by the local history of reinforcement, as evidenced by the fact that their choices were largely independent of sequence type (i.e., regardless of where or whether the reversal occurred in the sequence). The pigeons appeared to base their choices on the overall probability of reinforcement for a given trial in the sequence (or on how much time had elapsed since the start of the sequence). Furthermore, they also anticipated the reversal, as they had tended to do with 80-trial sequences. That is, in addition to matching the overall percentages correct over sequences, the pigeons also tended to choose S2 earlier than

they should have, according to the percentage of reinforcement associated with choice of S1 on that trial.

Surprisingly, the results showed that the local reinforcement cues, which provided reliable information about the reinforcement contingencies, did not control choice behavior. Thus, whether or not reinforcement occurred for a particular response did not affect the pigeon's choice on the following trial. It appears that pigeons make minimal use of local cues that signal a reversal, but instead appear to rely on cues such as the passage of time into the session (with 80-trial sessions; Rayburn-Reeves et al., 2011; Rayburn-Reeves et al., 2012) or the percentage of reinforcement associated with each trial (with five-trial sequences, in the present experiments), which involve averaging over sessions/sequences (i.e., the use of global cues).

One might note that in both Experiments 1 and 2, the likelihood of a reversal was not constant as the pigeon proceeded through the sequence. In Experiment 1, if a reversal had not occurred yet, the likelihood of a reversal on Trial 5 was 100 %, on Trial 4 it was 50 %, on Trial 3 it was 33 %, and on Trial 2 it was 25 %. In Experiment 2, if a reversal had not occurred yet, the likelihood that it would occur on Trial 5 was 50 %; on Trial 4, 33 %; on Trial 3, 25 %; and on Trial 2, 20 %. The changing likelihood of a reversal as the pigeons proceeded through the sequence was something that we not only considered, but actually tested in a separate experiment, by making the likelihoods of reversals equal for each trial. We have not included those data, because that procedure produced its own problem: To have equal likelihoods of S1 being correct on Trial 5, one needs to have at least one S1 and one S2 sequence on Trial 5. To have equal likelihoods of S1 being correct on Trial 4, one needs to have at least two S1 and two S2 sequences on Trial 4. To have equal likelihoods of S1 being correct on Trial 3, one needs to have at least four S1 and four S2 sequences on Trial 3. To have equal likelihoods of S1 being correct on Trial 2, one needs to have at least eight S1 and eight S2 sequences on Trial 2. And to have equal likelihoods of S1 being correct on Trial 1, one needs to have at least 16 S1 and 16 S2 sequences on Trial 1. Keeping in mind that once a reversal occurs, S2 is correct from then on, that would mean that in 32 sequences there would be $16 + 8 + 4 + 2 + 1 = 31$ correct S1 trials, and $16 + 24 + 28 + 30 + 31 = 129$ correct S2 trials (or, S2 would be correct on more than 80 % of the trials). When we exposed the pigeons to this procedure, they responded to that asymmetry by producing a large S2 bias. As it turned out, the pigeons were only somewhat sensitive to the equal likelihoods of S1 and S2 on Trial 1 (they chose S1 16 % of the time), but then they responded optimally by choosing S2 consistently on all succeeding trials (they chose S2 99 % of the time).

Although in the present experiments pigeons did not appear to use the presence or absence of reinforcement from

the previous trial, pigeons can certainly learn conditional discriminations, such as in matching-to-sample and oddity-from-sample tasks, in which the choice of a comparison stimulus is contingent on the most recently presented sample stimulus. Pigeons can also learn to choose a comparison stimulus on the basis of hedonic samples (i.e., food or the absence of food) after being trained to associate one stimulus with the presentation of food and another with the absence of food (Zentall, Sherburne, & Steirn, 1992). However, in the present research, the pigeon would have had to remember not only the outcome of the previous trial, but also the stimulus that was chosen. If that is how the procedure used in the midsession reversal task should be viewed, a better analogy would be serial biconditional matching such as that used by Edwards, Miller, and Zentall (1985), in which the houselight (on or off) signaled whether the pigeons should match the previously presented stimulus (the sample) or mismatch it.

However, other procedures specifically designed to assess the ability of pigeons to use feedback from the previously reinforced response to maximize reinforcement have had some success. Williams (1972) used a procedure in which the overall probabilities of reinforcement associated with two stimuli were .50, but the local probability of reinforcement for repeating the same response was varied over trials, depending on the outcome of the previous trial. Specifically, the probability of reinforcement for repeating a response that had been reinforced on the previous trial was .80, whereas switching to the alternative stimulus was reinforced with a probability of .20 (a probabilistic win-stay). However, if the response on the previous trial was not reinforced, this indicated that switching to the alternative stimulus would result in reinforcement with a probability of 1.0, whereas repeating the same response would never be reinforced (lose-shift). The results showed that even though both components were learned, the lose-shift component was learned faster and more accurately than the win-stay component; however, the probability of reinforcement associated with the lose-shift component was better differentiated (1.0 vs. 0) that the win-stay component (.8 vs. .2), and the delay between trials in the lose-shift component was half as long as in the win-stay component (3 vs. 6 s).

Shimp (1976) noted these differences and conducted a similar study in which the delay between trials was varied (2.5, 4.0, and 6.0 s) over trials for both components, to determine the effects of the delay between trials on subjects' ability to use the stimulus and outcome from the previous trial as a cue. Additionally, he used a correction procedure in which incorrect responses resulted in a 5-s time out and, at the end of the interval, the trial was repeated until the subject made the correct response. Thus, all trials ended in reinforcement. He found that subjects performed well on both components, but as the delay between trials increased,

accuracy decreased. Thus, control by reinforcement (and its absence) of choice on the following trial was evident, but it was less effective as the delay between trials increased. Shimp interpreted this finding as the ability of pigeons to use memory for recent events to predict the likelihood of reinforcement for future behavior. The two previous studies indicated that pigeons are able to use the outcome of the previous trial as a basis for subsequent behavior, but that this ability is greatly affected by the delay between trials.

The fact that pigeons in our experiments did not appear to use the feedback provided by the outcome of the most recent trial, but seemed to rely on the percentage of reinforcement associated with each trial in the sequence (averaged over trials), suggests that our procedures may not have sufficiently encouraged the pigeons to use win-stay/lose-shift as a cue for responding. Although the average percentage of reinforcement associated with each trial in the sequence did not lead to optimal choice accuracy, apparently the difference in the overall percentages of reinforcement between win-stay/lose-shift and overall percentage matching was not sufficient to encourage pigeons to use the more relevant cue. In this regard, it should be noted that in the Williams (1972) and Shimp (1976) studies, the only cues that could have been used to improve choice accuracy were the previous choice and its outcome. However, it also may be that it is easier for pigeons to use a more global cue (the probability of reinforcement accumulated over sequences associated with each trial in the sequence) rather than a local cue (the response and outcome from the preceding trial) when performing this task.

References

Bailey, J. T., & Mazur, J. E. (1990). Choice behavior in transition: Development of preference for the higher probability of reinforcement. *Journal of the Experimental Analysis of Behavior, 53,* 409–422.

Baum, W. M. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior, 22,* 231–242.

Bitterman, M. E. (1975). The comparative analysis of learning. *Science, 188,* 699–709.

Bond, A. B., Kamil, A., & Balda, R. P. (2007). Serial reversal learning and the evolution of behavioral flexibility in three species of north American corvids (*Gymnorhinus cyanocephalus, Nucifraga Columbiana, Aphelocoma californica*). *Journal of Comparative Psychology, 121,* 372–379.

Cook, R. G., & Rosen, H. A. (2010). Temporal control of internal states in pigeons. *Psychonomic Bulletin & Review, 17,* 915–922.

Cowie, S., Davison, M., & Elliffe, D. (2011). Reinforcement: Food signals the time and location of future food. *Journal of the Experimental Analysis of Behavior, 96,* 63–86.

Davison, M., & Baum, W. M. (2003). Every reinforcer counts: Reinforcer magnitude and local preference. *Journal of the Experimental Analysis of Behavior, 80,* 95–129.

Edwards, C. A., Miller, J. S., & Zentall, T. R. (1985). Control of pigeons' matching and mismatching performance by instructional cues. *Animal Learning & Behavior, 13,* 383–391.

Herrnstein, R. J. (1961). Relative and absolute strength of responses as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4,* 267–272. doi:10.1901/jeab.1961.4-267

Krägeloh, C. U., Davison, M., & Elliffe, D. M. (2005). Local preference in concurrent schedules: The effects of reinforce sequences. *Journal of the Experimental Analysis of Behavior, 84,* 37–64.

Mackintosh, N. J., McGonigle, B., Holgate, V., & Vanderver, V. (1968). Factors underlying improvement in serial reversal learning. *Canadian Journal of Psychology, 22,* 85–95.

Rayburn-Reeves, R. M., Miller, H. C., & Zentall, T. R. (2010). "Counting" by pigeons: Discrimination of the number of biologically relevant sequential events. *Learning & Behavior, 38,* 169–176.

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.

Rayburn-Reeves, R. M., Stagner, J. P., Kirk, C. R., & Zentall, T. R. (2012). Reversal learning in rats (*Rattus norvegicus*) and pigeons (*Columba livia*): Qualitative differences in behavioral flexibility. *Journal of Comparative Psychology.* doi:10.1037/a0026311

Shimp, C. P. (1976). Short-term memory in the pigeon: The previously reinforced response. *Journal of the Experimental Analysis of Behavior, 26,* 487–493.

Stenhouse, D. (1974). *The evolution of intelligence: A general theory and some of its implications*. New York, NY: Barnes & Noble.

Wilkie, D. M., & Willson, R. J. (1992). Time–place learning by pigeons, *Columbia livia*. *Journal of the Experimental Analysis of Behavior, 57,* 145–158.

Williams, B. A. (1972). Probability learning as a function of momentary reinforcement probability. *Journal of the Experimental Analysis of Behavior, 17,* 363–368.

Zentall, T. R., Sherburne, L. M., & Steirn, J. N. (1992). Development of excitatory backward associations during the establishment of forward associations in a delayed conditional discrimination by pigeons. *Animal Learning & Behavior, 20,* 199–206.