



The Midsession Reversal Task with Pigeons Does a Brief Delay Between Choice and Reinforcement Facilitate Reversal Learning?

Megan A. Halloran, Thomas R. Zentall*

University of Kentucky, United States



ARTICLE INFO

Keywords:
midsession reversal
delay
of reinforcement
attention
pigeons

ABSTRACT

In a midsession reversal task, the session begins with a simple simultaneous discrimination in which one stimulus (S1) is correct and the other stimulus (S2) is incorrect (S1+/S2-). At the midpoint of the session, the discrimination reverses and S2 becomes the correct choice (S2+/S1-). When choosing optimally, a pigeon should choose S1 until the first trial in which its choice is not reinforced and then it should shift to S2 (win-stay/lose-shift). With this task, pigeons have been shown to respond suboptimally by anticipating the reversal (making anticipatory errors) and continuing to choose S1 after the reversal (making perseverative errors). This sub-optimal behavior may result from a pigeon's relative impulsivity due to the immediacy of reinforcement following choice. In other choice tasks, there is evidence that the introduction of a short delay between choice and reinforcement may decrease pigeons' impulsivity. In the present experiment, a delay was introduced between stimulus selection and reinforcement to assess whether it results in a decrease in anticipatory and perseverative errors. Pigeons that had a delay between choice and reinforcement were a bit slower in acquiring the midsession reversal task compared to those without a delay, but showed no decrease in either anticipatory or perseverative errors. It is likely that the pigeons' natural tendency to use time from the start of the session to the reversal as a cue to reverse prevented the delay from increasing accuracy on this task.

1. Introduction

Cognitive flexibility or the ability to adapt one's behavior to environmental change depends on several factors including learning that the reward value of the stimuli can suddenly change. If a previously rewarded stimulus ceases to be rewarding, an organism must adjust its behavior accordingly.

The number of trials that it takes for an animal to reverse a discrimination sometimes has been taken as a measure of its flexibility. In fact, Bitterman (1965) has attempted to order several species according to how much they can benefit from a repeated series of reversals (i.e., a serial reversal task). The measure of the flexibility being the asymptotic number of errors made as the number of reversals increases, relative to original learning.

2. The Midsession Reversal Task

The mid-session reversal task involves a simple simultaneous discrimination in which one stimulus (S1) is correct (S1+) for the first half of a session and the other stimulus (S2) is incorrect (S2-). At the midpoint of the session, the stimuli switch roles so that S2 is now correct

(S2+) and S1 is incorrect (S1-). Following many sessions of training, the midsession reversal can be used to assess how an animal adapts to the feedback it receives following each trial. Assuming that the animal cannot accurately count the trials to the reversal, to respond optimally, it should develop a win-stay/lose-shift strategy. For example, if on the preceding trial, reinforcement was provided (win), the same stimulus should be chosen on the next trial (stay). However, if reinforcement did not occur (lose), on the next trial the other stimulus should be chosen (shift). Several animals have been shown to use a win-stay/lose-shift strategy when presented with a mid-session reversal task including humans (Rayburn-Reeves et al., 2011), apes and monkeys (Beran et al., 2008), and rats (Rayburn-Reeves et al., 2013a, 2013b). However, under most conditions pigeons have shown suboptimal performance with this task (e.g., McMillan et al., 2014; Rayburn-Reeves et al., 2011; Rayburn-Reeves & Cook, 2016).

A study examining the mid-session reversal task in both pigeons and humans was conducted by Rayburn-Reeves et al. (2011). Experiment 1 with pigeons involved a color discrimination between red and green lights. Subjects were trained for 50 sessions and each session consisted of 80 trials, with a reversal occurring after trial 40. S1 was correct for the first 40 trials (S1+/S2-) and S2 was correct for the last 40 trials

* Corresponding author at: Department of Psychology, University of Kentucky, Lexington, KY, 40506-0044, United States
E-mail address: zentall@uky.edu (T.R. Zentall).

<https://doi.org/10.1016/j.beproc.2020.104150>

Received 5 March 2020; Received in revised form 19 May 2020; Accepted 22 May 2020

Available online 26 May 2020

0376-6357/ © 2020 Elsevier B.V. All rights reserved.

(S2 + /S1-). At the beginning of each trial, one key was illuminated red and the other was illuminated green. A single peck to either key turned off both keys and produced reinforcement for correct choice together with an intertrial interval (ITI) or just an ITI for incorrect selections. Over the course of the session, the position of the colors were counterbalanced over the left and right keys. The results indicated that pigeons chose S1 almost exclusively during the early trials of a session, but choice of S1 began to decline as the reversal approached. Choice of S1 continued to decline until subjects were almost exclusively choosing S2 towards the end of the session. Subjects were approximately 70% accurate during the five trials preceding the reversal (36-40) and approximately 55% accurate during the five trials immediately following the reversal (41-45). These results indicate that pigeons were anticipating the reversal, by choosing S2 while S1 was still correct (anticipatory errors) and they continued to choose S1 even after the reversal when S2 was correct (perseverative errors). Thus, the pigeons did not use the feedback from the preceding trial(s) sufficiently. Instead there is evidence that the pigeons relied on the passage of time from the start of the session, as a cue to reverse (McMillan & Roberts, 2012; Smith et al., 2017).

In an attempt to get the pigeons to avoid using the passage of time as the basis for reversing their choice, Rayburn-Reeves et al. (2011) used a procedure in which the reversal could occur randomly at different points in the session. With the reversal unpredictable, the pigeons made few anticipatory errors when the reversal occurred early in the session, but many perseverative errors after the reversal. Conversely, when the reversal occurred late in the session, the pigeons made few perseverative errors, but they made a large number of anticipatory errors. In fact, the worst performance was found when the reversal occurred near the end of the session. Thus it appears that even when the reversal was unpredictable, pigeons also relied on timing as opposed to only relying on immediate feedback from the preceding trial(s).

Further evidence that time from the start of the session to the reversal generally serves as the cue for pigeons to reverse was obtained by Smith et al. (2017). When pigeons were trained with a 5-s ITI and shifted to a 2.5-s ITI, they showed very few anticipatory errors but continued to choose the S1 stimulus for many trials after the reversal. Conversely, when the ITI was shifted to 10 s, the pigeons began reversing very early, making many more anticipatory errors but fewer perseverative errors.

There are conditions, however, under which pigeons come close developing a win-stay/lose-shift strategy. Rayburn-Reeves et al. (2013a), conducted a mid-session reversal study with pigeons using a spatial discrimination with different ITI durations, between groups, 1.5 s, 5 s, or 10 s. After 40 sessions of training, the pigeons in the 5-s and 10-s ITI condition showed the typical number of anticipatory and perseverative errors. However, the 1.5-s ITI group showed almost no anticipatory errors and no additional perseverative errors.

These results suggest that pigeons can develop something like a win-stay/lose-shift strategy when the ITI is very short. However, with a spatial discrimination, a very short ITI, equal to reinforcement time, allows the bird to anticipate an immediate move from the feeder to the correct location (right or left). This repeated anticipatory movement may account for the decreased anticipatory errors with the short ITI because the improvement in accuracy was not found with a short ITI when the discrimination involved colors, the location of which could not be anticipated (Laude et al., 2014) or when the ITI was 5 s.

The question that remains is why pigeons continue to make anticipatory and perseverative errors with a color discrimination even after many sessions of training. Smith et al. (2017) hypothesized that pigeons may use a time-based strategy to predict the reversal because they are unable to remember what stimulus was chosen on the previous trial and what the outcome of the selection was. In that study they tested whether cueing pigeons, with the use of houselights during the ITI, to indicate the stimulus that was selected on the previous trial and leaving

the feeder light on when the last response was reinforced. Although the relevant cues did improve accuracy relative to a randomly cued control group, both groups continued to show anticipatory and perseverative errors. That is, neither group showed evidence of a win-stay/lose-shift choice pattern, suggesting that even the cued group was still relying on a time-based strategy. To determine the extent to which the two groups were relying on temporal cues, the ITI was either doubled to 10 s or halved to 2.5 s. The results indicated that both groups were still relying on temporal cues. The general structure of the midsession reversal task appears to invariably encourage behavior to be guided by inaccurate temporal cues rather than by the feedback from the immediately preceding trial(s).

2.1. Delay between choice and reinforcement

Interestingly, when the choice response involves a single response followed immediately by reinforcement, a number of studies have found that tasks that should be relatively easy to learn have proven to be quite difficult. In those studies, learning has been facilitated, paradoxically, by increasing the time between the choice response and reinforcement. Zentall (2019) hypothesized that this increase in time may encourage pigeons to choose more carefully and thus reduce impulsive choice.

The ephemeral reward task. In the ephemeral reward task, an animal is given a choice between two stimuli, S1 and S2, choice of either one results in an identical reinforcer. If the animal chooses S1, it gets the reinforcer associated with that choice and the trial is over. However, if the animal chooses S2, it gets the reinforcer associated with that choice but the S1 option remains and it can also get the reinforcer associated with S1. Thus, in this task, it is optimal to choose S2.

Research has found that wrasse (cleaner fish) quickly learn to choose optimally (Salwiczek et al., 2012), an ability attributed to their natural foraging strategy. Salwiczek et al. (2012) found, however, that most primates were not able to learn to choose optimally within the same number of trials.

Interestingly, grey parrots did learn to choose optimally as quickly as the wrasse (Pepperberg & Hartsfield, 2014), even though their natural foraging behavior is quite different from that of wrasse and more like that of primates. Pepperberg and Hartsfield hypothesized that wrasse and parrots both make selections with their mouths, while primates make selections with their hands. This hypothesis was tested with pigeons, which also choose with their beak, but the pigeons failed to learn this discrimination even with considerably more training (Zentall et al., 2016). hypothesized that pigeons might not be associating their first choice with the second reinforcement because of the immediacy of reinforcement associated with either choice.

Using a technique developed by Rachlin and Green (1972), Zentall et al. (2016) introduced a delay between the choice response and reinforcement (what Rachlin & Green referred to as *making a prior commitment*). For the delay group, after an initial observing response to a white key, choice between a yellow and blue stimulus started a 20-s timer and the first response following the 20-s delay provided reinforcement (i.e., a fixed interval schedule). If the optimal choice had been selected, after the pigeons received their first reinforcement, they were able to select the other stimulus and immediately receive a second reinforcement. If initially, the other choice had been selected, a single reinforcement was provided (after 20 s) and the trial ended. For the control group, to control for the extended duration of the trial for the delay group, the white-key observing response involved a fixed-interval 20-s schedule prior to the choice and a single peck to the yellow or blue stimulus provided immediate reinforcement.

After many sessions of training, the control group was not significantly different from chance, whereas the experimental group was at 90% correct. It was proposed that the pigeons learned the optimal response with an inserted delay between choice and reinforcement because the delay reduced the pigeons' tendency to respond impulsively

to the initial choice.

Object permanence. A similar finding was reported by Zentall and Raley (2018) for pigeons trained on an object permanence task using a procedure similar to that used previously in research with dogs (Miller et al., 2009). The study used a rotating beam with a cup on either end of the beam. Initially, pigeons were trained on a visible displacement task in which a reinforcer evident to the subject was placed inside one of the cups. The subject was then tested to determine if it would choose the cup with the reinforcer. Surprisingly, the pigeons chose at chance levels over the initial session of 30 testing trials. More surprisingly, over many sessions of training, although the pigeons performed better than chance, they never showed a high degree of choice accuracy.

As previously suggested by Zentall et al. (2016) the immediacy of reinforcement following choice may have resulted in impulsive choice. To test this hypothesis, in a follow-up experiment, a 5-s delay was introduced between when the cup was baited and the pigeon's access to the cups. Interestingly, the 5 s delay facilitated visible displacement learning and that learning then transferred to invisible displacement in which the cup with the reinforcer moved to a new position.

The suboptimal choice gambling-like task. Zentall, Andrews & Case (2017) also found evidence that introducing a delay can increase optimal choice in a task in which pigeons were choosing suboptimally without a delay. The experiment involved a gambling-like task in which pigeons were presented with a spatial discrimination (left and right). Choosing optimally resulted in a signal indicating reinforcement 100% of the time, while the suboptimal selection resulted in one of two equally probable signals: one indicating reinforcement and the other indicating no reinforcement (thus, overall, 50% reinforcement). For the experimental group there was a 20 s delay (a fixed interval 20-s schedule) between the initial selection and presentation of the signal for reinforcement or its absence. For the control group there was no delay between initial selection and the signal for reinforcement. The results indicated that the pigeons in the experimental group generally chose optimally, whereas those in the control group generally chose suboptimally.

Thus, it appears that with several different procedures inserting a delay between choice and reinforcement (or in the latter case a conditioned reinforcement) can facilitate learning. This brings us back to the midsession reversal task. It is possible that in the midsession reversal task, the immediacy of reinforcement following choice results in impulsive errors, especially in the region around the reversal, that could be reduced by introducing a delay between stimulus choice and reinforcement.

It is also possible, however, that the errors made by pigeons performing the midsession reversal are not produced by impulsive choice. Because pigeons appear to time the interval from the start of the sessions to the reversal, errors may result solely from the variability in the pigeons' ability to time the occurrence of the reversal from the start of the session as well as from the competition between the response strength associated with the S1 and S2 stimuli. In the present experiment, for the experimental group, on each trial, a 5-s delay was inserted between stimulus choice and reinforcement. For the control group there was no delay but because pigeons are known to time the interval from the start of the session to the reversal, 5 s was added to the ITI to control for time from the start of the session to the reversal.

3. Methods

3.1. Subjects

The subjects were 12, 8-12 year old, unsexed pigeons that had been purchased from the Palmetto Pigeon Plant (Sumter, SC). They had all participated in unrelated color discrimination experiments. During the experiment, the birds were kept at 85% of their free-feeding weight to ensure motivation in the experiment. They were individually housed in wire cages (28 × 38 × 30.5 cm) with free access to water and grit in a

colony room that was maintained on a 12:12-hr 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.

3.2. Apparatus

The experiment was conducted in a BRS/LVE (Laurel, MD) sound-attenuating standard operant test chamber with inside measurements 34 cm high, 30 cm from the response panel to the back wall, and 35 cm across the response panel. Three circular response keys (2.5 cm diameter) were aligned horizontally on the response panel and were separated from each other by 6.0 cm. Only the left and right-side keys were used in the experiment. 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 could project blue and yellow hues (Kodak Wratten Filters Nos. 38 and 9, 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 5.1 × 5.7 cm aperture horizontally centered and vertically located midway between the response keys and the floor of the chamber. Reinforcement consisted of 1.5 s access to mixed grain. White noise was generated from outside the chamber and the apparatus was controlled by a computer in the adjacent room running Med-PC IV (Tatham & Zurn, 1989) with a 10-ms resolution. The houselight was not used in this experiment.

3.3. Procedure

Each experimental session began with one side key illuminated blue and the other yellow. For the experimental group ($n = 6$), a single response to either stimulus turned off the stimulus not selected and started a 5-s delay. Following the delay, a correct response resulted in 1.5-s access to grain followed by a 3.5-s ITI. An incorrect response resulted in a 5-s ITI, with no reinforcement. For the control group ($n = 6$), reinforcement occurred immediately following a correct response. To account for the increased trial duration resulting from the delay, the control group had a 10-s ITI. Each session consisted of 80 trials, in which during the first 40 trials (1-40) S1 was correct (S1+/S2-) and during the last 40 trials (41-80) S2 was correct (S1-/S2+). For half of the subjects the blue hue was S1 and for the other half of the subjects the yellow hue was S1. The location of the blue and yellow hues was counterbalanced over trials. All pigeons completed 60 sessions of training.

3.4. Analysis

Data from each bird was averaged over blocks of ten sessions (Sessions 1-10, 11-20, 21-30, 31-40, 41-50, and 51-60). Overall accuracy (Trials 1-80), first half accuracy (Trials 1-41), and second half accuracy (Trials 42-80) for each group for each ten-session block, was assessed. Because individual differences in asymptotic accuracy at the start and end of each session may be unrelated to the ability to detect the reversal, the data in close proximity to the reversal were analyzed as well. Anticipatory errors just prior to the reversal were analyzed using trial-by-trial data for the 4 trials prior to the reversal (Trials 38 to 41; note, the feedback from the reversal occurred only after choice on Trial 41, so choice on Trial 41 was included in anticipatory errors) while perseverative errors just after the reversal were analyzed using the 4 trials immediately following the reversal (Trials 42 to 45). The four trials immediately preceding and following the reversal were chosen because these trials were most likely to be affected by impulsive choice and therefore, most likely to be affected by the delay between choice and reinforcement.

A two-way 2 (Session Half: first-half or second-half) × 2 (Group:

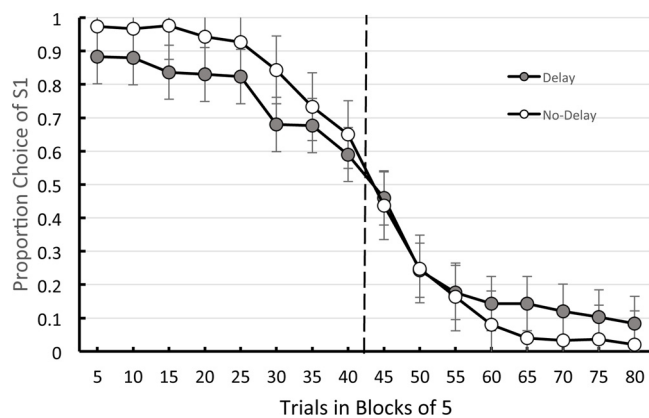


Fig. 1. The average percentage choice of S1 grouped by blocks of five trials for Sessions 11-20 for the delay and no-delay groups. The midsession reversal point is denoted by a dashed vertical line. Error bars are standard errors of the mean.

delay or no-delay) mixed ANOVA on percent correct choice was conducted for each ten-session block. Another two-way 2 (Error Type: pre-reversal errors or post-reversal errors) \times 2 (Group: delay or no-delay) mixed ANOVA on percent correct was conducted for each block of ten sessions on the errors immediately before and after the reversal. Additionally, an independent samples *t*-test was conducted to compare the delay group with the no-delay group for criterion accuracy (60%, 70%, 80%, 90% correct). Additionally, the number of sessions to various criteria 60%, 70%, 80% and 90% were analyzed for each pigeon. Statistical significance was set at $p = .05$ level for all statistical tests.

4. Results

A difference between the delay and no-delay groups emerged in the second block of Training Sessions (11-20). The data from Sessions 11-20 plotted by choice of S1 as a function of trial number appears in Fig. 1. Normality was assessed with a Shapiro-Wilk test. The overall accuracy measure for the two groups, was found to be not statistically different from normally distributed, $w(6) = 0.87$, $p = 0.23$, for the delay group and $w(6) = 0.97$, $p = 0.90$, for the no-delay group. Analysis of the accuracy data indicated that there was a significant difference between the delay group ($M = 0.795$, $SD = 0.069$) and no-delay group ($M = 0.873$, $SD = 0.027$), $F(1,10) = 6.53$, $p = 0.029$, partial eta squared = .0395, with the no-delay group performing more accurately than the delay group. But there was no significant difference between first-half session accuracy and second-half session accuracy $F(1,10) = 3.59$, $p = 0.088$, partial eta squared = 0.264, and there was no significant interaction between session half and group type, $F(1,10) = 2.270$, $p = 0.163$, partial eta squared = 0.185. Errors just before and just after the reversal were also analyzed but as can be seen in Fig. 1, none of the effects was statistically significant.

Similar but decreasing group differences were found for Sessions 21-30, 31-40, 41-50, and 51-60, however, none of the effects was statistically significant. The data from Sessions 51-60 plotted as choice of S1 as a function of trial number appears in Fig. 2. A can be seen in the figure, by the end of training, accuracy by the delay and no-delay groups was essentially the same. For comparison purposes, data from (Rayburn-Reeves et al., 2013a, 2013b) with the more typical single-peck choice with a 5-s ITI also appears in Fig. 2.

There were no significant differences between the delay and no-delay groups in the mean sessions to 60%, 70%, 80%, or 90% correct criterion, all $ts < 1$.

5. Discussion

Although a significant difference in accuracy between the delay and

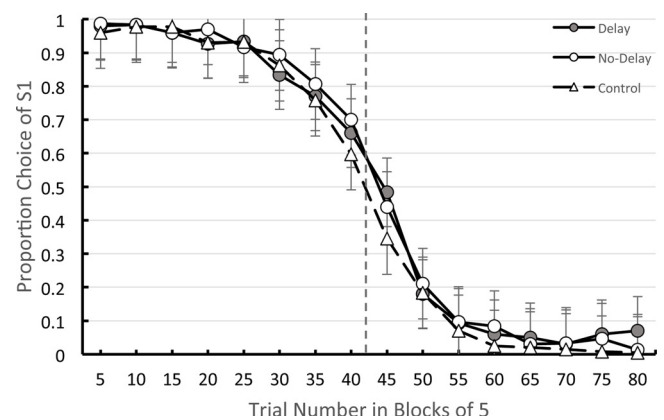


Fig. 2. The average percentage choice of S1 grouped by blocks of five trials for Sessions 51-60 for the delay and no-delay groups as well as the 5 s ITI no-delay control group from Rayburn-Reeves et al. (2013a), 2013b. The midsession reversal point is denoted by a dashed vertical line. Error bars are standard errors of the mean.

the no-delay groups was found in Sessions 11-20, by the end of training, both groups reached the same level of accuracy. This difference can be attributed to slower learning by pigeons in the delay group. Importantly, there was no difference in accuracy close to the reversal between the two groups across any of the ten-session blocks. This result implies there was no effect of the inserted delay on the assumed impulsivity that pigeons exhibit on this task. These results are inconsistent with the results from a number of tasks, including the ephemeral choice task (Zentall et al., 2016), the object permanence task (Zentall & Raley, 2018) and the gambling-like task (Zentall et al., 2017), that have been shown to increase optimal choice behavior by pigeons when a short delay was inserted between the choice response and reinforcement.

Why the delay inserted between choice and reinforcement in the midsession reversal task did not improve accuracy as it did in the ephemeral choice task, the object permanence task, and the gambling-like task is not clear. An important difference between those other tasks and the midsession reversal task is that in those other tasks, time into the session is not a factor, whereas pigeons perform the midsession reversal task by using the time from the start of the session to the reversal as a cue to reverse. Thus, it may be that in the midsession reversal task, the pigeons are not choosing impulsively but are making errors in their timing of the reversal.

Zentall, Peng, House, and Yadav (2019) have suggested that anticipatory and perseverative errors in the midsession reversal task result from competition between the response strength of the S1 and S2 stimuli in the region of the reversal. This hypothesis was originally proposed by McMillan et al. (2015) and later confirmed by Smith et al. (2018). They argued that the response strength associated with the S1 stimulus is high at the start of the session and as the session progresses, the response strength associated with the S2 stimulus increases to the point that it competes with that of the S1 stimulus, resulting in anticipatory errors. Following the reversal, the response strength associated with the S1 stimulus gradually decreases but while it is still strong, it competes with the response strength associated with the now strong response strength of the S2 stimulus. Importantly, the response strength associated with the two stimuli is controlled largely by the passage of time. In support of this model, Smith et al. found that choice latency was consistent with a horse-race model of interference rather than a tug-of-war model. That is, choice latency was found to be relatively constant throughout the session instead of increasing as the reversal approached.

According to this model, anticipatory and perseverative errors result from interference between the response strengths of the S1 and S2 stimuli in the region around the reversal. One way that the competition between the two response strengths could be reduced would be to

reduce the validity of one of those stimuli. Santos, Soares, Vasconcelos, and Machado (2019) reduced the validity of the S2 stimulus by reducing the probability of reinforcement for correct choice of the S2 stimulus to 20% while maintaining the probability of reinforcement for correct choice of the S1 stimulus at 100%.

Consistent with the competition hypothesis, this manipulation virtually eliminated anticipatory errors, but it did not increase perseverative errors. That is, it resulted in a net increase in midsession reversal accuracy. Zentall et al. (2019) replicated the effect of the decrease in the probability of reinforcement for correct choice of the S2 stimulus found by Santos et al. (2019) and showed that one could find a similar effect by increasing the response requirement for choice of the S2 stimulus, from 1 to 10 pecks.

Interestingly, by devaluing the S2 stimulus, not only did it virtually eliminate anticipatory errors, but it appeared to make the reversal more discriminable. Consider the possibility that with the elimination of anticipatory errors, nonreinforcement for choice of the S1 stimulus may be easier to detect, and once detected, faced with a choice between 0% reinforcement for continued choice of the S1 stimulus and 20% reinforcement for correct choice of the S2 stimulus there was no increase in perseverative errors.

The question remains, why rats do so much better on the standard midsession reversal task than pigeons. The fact that they have been trained on a spatial version of the task but not a visual midsession reversal may account for some of the difference in accuracy on this task. Furthermore, the fact that pigeons do as well as rats when the ITI is reduced to 1.5 s may provide additional clues. When the ITI is very short and the task is spatial (e.g., choose left for the first 40 trials and right for the last 40 trials). The pigeons can move immediately from the feeder to the correct stimulus without looking for it. That is, they can anticipate where the next correct response will be and on Trial 41 when reinforcement does not follow, it breaks up the repetitive key-feeder-key response pattern. When the ITI is just a bit longer, 5 s, that pattern of behavior is not as easily maintained for pigeons because other behavior is likely to intervene. However, the rats may be able to mediate that 5-s interval, because they are less active than the pigeons. This hypothesis is supported by the finding by McMillan et al. (2014) that when rats were tested on a T-maze version of the midsession reversal task, anticipatory and perseverative errors were similar to pigeons tested on a visual discrimination. In the T-maze task, the rats were brought back to the start box before each new trial so they could not easily orient to the correct location between trials.

It would be informative to test rats on a non-spatial version of this task. For example, rats could be provided with a discrimination in which a light is illuminated over one lever, but not the other. For the first half of the session, for example, the lever with the light would be the S1 stimulus whereas the lever without the light would be the S2 stimulus. The location of the light would vary from trial to trial so its spatial location would not be predictable.

5.1. Limitations

Although the typical control group in a midsession reversal experiment involves a 5-s ITI with no delay, a condition that was not included in the present study, the 10-s ITI no-delay group from the present experiment looked quite similar to the 5-s ITI no-delay group (see Fig. 2) from Rayburn-Reeves et al. (2013a), 2013b. Importantly, by the end of training, the data from all three conditions looked quite similar, in spite of the fact that the duration of the session for the 5-s no-delay condition was considerably shorter than the session duration for both of the groups in the present experiment, and therefore it should have been somewhat easier to time the reversal from the start of the session. It should be noted, however, that Rayburn-Reeves et al. (2013a), 2013b found little difference in accuracy of pigeons with a 5-s ITI compared with a 10-s ITI.

Finally, in both the ephemeral choice task (Zentall et al., 2016) and

the gambling-like task (Zentall et al., 2017) the choice to reinforcement delay was 20 s, whereas it was only 5 s in the present midsession reversal study. However, the delay was only 5 s in the object permanence experiment and that delay proved to be sufficient to facilitate acquisition.

6. Conclusions

Inserting a delay between choice and reinforcement was not effective in reducing perseverative and anticipatory errors in pigeons performing the mid-session reversal task. A 5-s delay modestly slowed down learning during earlier session blocks as compared to the no-delay group, but the two groups performed similarly upon the completion of 60 sessions. The present results suggest that mechanisms different from those responsible for the suboptimal performance of the ephemeral reward task, the object permanence task, and the gambling-like task may be involved in the midsession reversal task. Ultimately, further research should be conducted to properly evaluate the ability of pigeons with this task and what its implications are for the behavioral flexibility of this species.

CRedit authorship contribution statement

Megan A. Halloran: Data curation, Formal analysis, Writing - original draft. **Thomas R. Zentall:** Conceptualization, Methodology, Resources, Writing - review & editing.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.beproc.2020.104150>.

References

- Beran, M.J., Klein, E.D., Evans, T.A., Chan, B., Flemming, T.M., Harris, E.H., et al., 2008. Discrimination reversal learning in capuchin monkeys (*Cebus apella*). *Psychological Record* 58, 3–14.
- Bitterman, M.E., 1965. The evolution of intelligence. *Scientific American* 212, 92–100.
- Laude, J.R., Stagner, J.P., Rayburn-Reeves, R., Zentall, T.R., 2014. Midsession reversals with pigeons: visual versus spatial discriminations and the intertrial interval. *Learning & Behavior* 42 (1), 40–46.
- McMillan, N., Roberts, W., 2012. Pigeons make errors as a result of interval timing in a visual, but not a visual-spatial, midsession reversal task. *Journal of Experimental Psychology: Animal Behavior Processes* 38 (4), 440–445.
- McMillan, N., Kirk, C., Roberts, W., 2014. Pigeon (*Columba livia*) and rat (*Rattus norvegicus*) performance in the midsession reversal procedure depends upon cue dimensionality. *Journal of Comparative Psychology* 128 (4), 357–366.
- McMillan, N., Sturdy, C.B., Spetch, M.L., 2015. When is a choice not a choice? Pigeons fail to inhibit incorrect responses on a go/no-go midsession reversal task. *Journal of Experimental Psychology: Animal Learning and Cognition* 41, 255–265. <https://doi.org/10.1037/xan0000058>.
- Miller, H.C., Gipson, C.D., Vaughan, A., Rayburn-Reeves, R., Zentall, T.R., 2009. Object permanence in dogs: Invisible displacement in a rotation task. *Psychonomic Bulletin and Review* 16, 150–155. <https://doi.org/10.3758/PBR.16.1.150>.
- Pepperberg, I.M., Hartsfield, L.A., 2014. Can grey parrots (*Psittacus erithacus*) succeed on a Bcomplex foraging task failed by nonhuman primates (*Pan troglodytes*, *Pongo abelii*, *Sapajus apella*) but solved by wrasse fish (*Labroides dimidiatus*)? *Journal of Comparative Psychology* 128, 298–306.
- Rachlin, H., Green, L., 1972. Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior* 17, 15–22.
- 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 (2), 125–137.
- Rayburn-Reeves, R.M., Laude, J.R., Zentall, T.R., 2013a. Pigeons show near-optimal win-stay/lose-shift performance on a simultaneous-discrimination, midsession reversal task with short intertrial intervals. *Behavioural Processes* 92, 65–70.
- Rayburn-Reeves, R.M., Stagner, J.P., Kirk, C.R., Zentall, T.R., 2013b. Reversal learning in rats (*Rattus norvegicus*) and pigeons (*Columba livia*): Qualitative differences in behavioral flexibility. *Journal of Comparative Psychology* 127, 202–211. <https://doi.org/10.1037/a0026311>.
- Salwiczek, L.H., Prôtôt, L., Demarta, L., Proctor, D., Essler, J., Pinto, A.I., et al., 2012. Adult cleaner wrasse outperform capuchin monkeys, chimpanzees, and orangutans in a complex foraging task derived from cleaner-client reef fish cooperation. *PLOS ONE* 7, e49068.
- Santos, C., Soares, C., Vasconcelos, M., Machado, A., 2019. The effect of reinforcement

- probability on time discrimination in the midsession reversal task. *Journal of the Experimental Analysis of Behavior* 111, 371–386.
- Smith, A.P., Beckmann, J.S., Zentall, T.R., 2017. Mechanisms of midsession reversal accuracy: Memory for preceding events and timing. *Journal of Experimental Psychology: Animal Learning and Cognition* 43 (1), 62–71.
- Smith, A.P., Zentall, T.R., Kacelnik, A., 2018. Midsession reversal task with pigeons: Parallel processing of alternatives explains choices. *Journal of Experimental Psychology: Animal Learning and Cognition* 44 (3), 272–279.
- Tatham, T.A., Zurn, K.R., 1989. The Med-PC experimental apparatus programming system. *Behavioral Research Methods, Instruments, and Computers* 21, 294–302.
- Zentall, T.R., 2019. What suboptimal choice tells us about the control of behavior. *Comparative Cognition & Behavior Reviews* 14, 1–17.
- Zentall, T.R., Case, J.P., Berry, J.R., 2016. Early commitment facilitates optimal choice by pigeons. *Psychonomic Bulletin & Review* 24 (3), 957–963.
- Zentall, T.R., Andrews, D.M., Case, J.P., 2017. Prior commitment: its effect on suboptimal choice in a gambling-like task. *Behavioral Processes* 145, 1–9.
- Zentall, T.R., Raley, O.L., 2018. Object permanence in the pigeon (*Columba livia*): insertion of a delay prior to choice facilitates visible- and invisible-displacement accuracy. *Journal of Comparative Psychology*.