

Resistance to Change Within Heterogeneous Response Sequences

Alliston K. Reid
Wofford College

Three experiments investigated how instrumental and Pavlovian contingencies contribute to resistance to change (RTC) in different ordinal response positions within heterogeneous response sequences in pigeons. RTC in the initial and terminal response positions of a three-response sequence were compared in Experiment 1, which presented three colored key lights in succession in each trial; and in Experiment 2, which severely degraded Pavlovian contingencies by presenting the lights simultaneously at each ordinal position. Experiment 3 eliminated the instrumental contingency in a high-order sign-tracking procedure. When the instrumental contingency was in effect, RTC of the initial position was greater than the terminal position (Initial RTC > Terminal RTC) when the Pavlovian contingencies were strong and when they were degraded. When the instrumental contingency was eliminated, RTC patterns reversed, producing a graded pattern of RTC (Initial < Middle < Terminal). Current theoretical approaches (e.g., behavioral momentum theory, conditioned reinforcement, and motivational control of instrumental conditioning) cannot account for these results. An alternative approach (a gradient model) shows that obtained measures of RTC in heterogeneous sequences may reflect a combination of three dissociable processes.

Keywords: resistance to change, response sequence, response strength, incentive learning, chained schedule

In many of his classic experiments, Thorndike (1898) studied how sequences of activities producing escape from a “puzzle box” would be strengthened, or become more likely the next time the animal was placed in that situation. Like many researchers of the day, he was concerned with how organisms learn to produce adaptive patterns of various behaviors that appear to be goal directed. For example, the response sequence of pressing a platform, clawing a string, and pushing a bar would open a door to allow his cats to escape the puzzle box, a consequence that Thorndike considered to be “satisfying” to the cat. Other researchers such as Tolman (1948) and Hull (1943) studied sequences of left and right turns by rats placed in mazes for food reward in the goalbox. Vertebrates and many invertebrates are able to learn similar heterogeneous response sequences as they adapt to conditions in their environment, forage or stalk prey, attract mates, and defend themselves from predators. Ironically, little is known about how the organism assigns credit (or blame) to the individual responses within the sequence (i.e., which activities in the behavior stream are “strengthened” and by how much). We have little understanding of the ways in which stimuli and various responses

become combined into an integrated sequence structure—that is, the rules of integration of responding and environmental cues to produce adaptive patterns of behavior.

The dominant theoretical framework for understanding such integrated sequences has been ‘chaining theory,’ which always characterizes the structure as stimulus-response chains. Although chaining theory has repeatedly been shown to be an inadequate explanation ever since Lashley’s (1951) classic critique, no other model has been accepted as an adequate replacement. The early successes in the study of simple reinforcement schedules may be one reason for this shortcoming, because emphasis shifted from ordinal patterns of heterogeneous responses (a categorical measure) to temporal patterns of repetitions of the same response (a quantitative measure; e.g., Ferster & Skinner, 1957; Skinner, 1938). Response rate became the dominant dependent variable. Researchers attempted to study response sequences and response chains by creating chained schedules, which allowed response rate in each component to reflect the strength of responding to that ordinal stimulus. Nevertheless, the response patterns produced by chained schedules are not the same as the integrated structures that Thorndike studied, and the reinforcement contingencies are not equivalent. Thorndike’s adaptive patterns required a variety of different responses in an ordinal sequence, rather than repetitions of the same response. Response rate could not serve as a measure of the strength of responding at each ordinal position of the sequence. Instead, the organism must learn which response to make, in which order, and avoid making the wrong response—issues that are not present in chained schedules. For example, William James (1890) provided examples, such as a man getting dressed or playing the piano, as chains in which behaviors are normally completed in a relatively fixed sequence, repeated responses would be mistakes (such as placing the right leg in pants,

Alliston K. Reid, Department of Psychology, Wofford College, Spartanburg, South Carolina.

Research awards provided by Wofford College supported the research reported in this article. I thank each of the students in my lab for their proficient help conducting the various experiments leading to this article. I especially thank Kara Vogelpohl, Rebecca Catoe, and Klaire Kiehne for their contributions to this research.

Correspondence concerning this article should be addressed to Alliston K. Reid, Department of Psychology, Wofford College, 429 North Church St., Spartanburg, SC, 29303. E-mail: Alliston.Reid@wofford.edu

twice), and interruptions often produce errors in the response sequence.

Response strength and conditioned value within response sequences are often assumed to depend on local delay-of-reinforcement gradients. In chained schedules, strength and value of a component are often inferred by its resistance to change (RTC) in response rate when a disrupter is provided. The central problem is how to measure the value of response strength at each ordinal position of a heterogeneous sequence when response rate is not an option. This study appears to be the first to measure RTC within heterogeneous response sequences to identify the relation between RTC and reinforcement contiguity.

RTC has served as a dependent variable in various research areas with differing goals, and measurements of RTC have been proposed to reflect different constructs or processes. Even though similar RTC procedures have been used in these different research areas, the areas have remained somewhat independent from each other with infrequent cross-referencing and little theoretical integration. These different theoretical approaches make different predictions of the patterns of RTC in the three experiments in this paper. Therefore, we begin by identifying briefly the relevant research areas. The main purpose of the current research was to see how patterns of RTC across ordinal positions of heterogeneous response sequences can inform us about the properties of these learned response structures. A secondary goal was to compare observations of RTC in heterogeneous response sequences to those obtained in chained schedules to answer the question, "What does resistance to change within sequences actually measure?"

The first theoretical approach to be summarized involves Behavioral Momentum Theory (BMT) (Nevin, 1992; Nevin & Grace, 2000, 2005). RTC is presumed to be an indicator of behavioral mass, reflecting response strength produced by reinforcement. BMT also proposes that preference between alternatives in concurrent chain schedules is another measure of behavioral mass, and thus is directly related to RTC (Nevin & Grace, 2000). Presumably, RTC is produced by the molar Pavlovian relation between a discriminative stimulus and the relative number of reinforcements delivered in its presence. Although instrumental contingencies were not originally assumed to influence RTC, several studies have recently demonstrated that local reinforcement delays do influence RTC in multiple schedules and preference in concurrent chain schedules (e.g., Bell, 1999; Bell, Scip, & Fitzsimmons, 2007; Doughty & Lattal, 2001; Grace, Schwendimann, & Nevin, 1998; Nevin, Grace, Holland, & McLean, 2001; Podlesnik, Jimenez-Gomez, Ward, & Shahan, 2006; Reed & Doughty, 2005; Shahan & Podlesnik, 2005). These findings are relevant for the current study because it explores the degree to which BMT, as a molar theory, is compatible with obtained RTC patterns reflecting local processes occurring within response sequences, such as delay-of-reinforcement gradients.

In a second research area, procedures used to identify the associative and incentive processes underlying Pavlovian and instrumental learning often measure RTC following devaluation of the reinforcer. Some reinforcer-devaluation procedures attempt to characterize the reinforcer representations created during Pavlovian conditioning procedures (e.g., Colwill & Motzkin, 1994; Holland, 1990, 2004; Konorski, 1967; Nairne & Rescorla, 1981; Ward-Robinson, 2004). RTC in this case measures the changes in the conditioned response, rather than changes in an instrumental

response. However, other reinforcer devaluation studies have measured RTC in instrumental responding to identify the Pavlovian (S-O), instrumental (R-O), and discriminative (SR) associative processes underlying instrumental conditioning (Adams & Dickinson, 1981; Colwill & Rescorla, 1988, 1990; Dickinson, Balleine, Watt, Gonzalez, & Boakes, 1995; Rescorla, 1990, 1992).

A related, third, method utilizing RTC is found in Pavlovian-instrumental transfer studies, which demonstrate that hedonic value obtained through Pavlovian conditioning contributes to the amount of instrumental responding (Colwill & Rescorla, 1990; Holland, 1990; Rescorla, 1990). Transfer-of-value studies are also used to identify how learned value is transferred from the unconditioned stimulus (US) to one or more conditioned stimuli (CS) that predict the US, such as in second-order autoshaping (e.g., Crawford & Domjan, 1995; Mazur & Wagner, 1982; Wagner & Brandon, 1989; Ward-Robinson, 2004). These procedures typically use reinforcer devaluation, so the procedures often appear identical to the reinforcer devaluation studies described above. Nevertheless, the goals tend to be different: transfer of value in this case, rather than identifying associative processes involving US representations. Research in operant conditioning has sometimes used reinforcer devaluation or extinction to measure transfer of value across different components of chained schedules (e.g., Williams, 1994, 1997, 1999; Williams, Ploog, & Bell, 1995). In these studies, RTC of operant response rates presumably reflects conditioned reinforcement in different components of the chained schedule and may reflect the backward propagation of value in these schedules.

The current study uses a fourth approach: subjects are trained to complete simple heterogeneous response sequences, and the amount of reinforcement or the reinforcement contingency is then changed to identify local properties of the learned response structure (e.g., Bachá-Méndez, Reid, & Mendoza-Soylovna, 2007; Neuringer, Kornell, & Olufs, 2001; Reed & Morgan, 2006; Reid, 1994). The procedures are similar to those used in BMT except the emphasis is on the local processes occurring at the level of responses within the sequence rather than on molar processes acting across the session. For example, Reid (1994) trained rats to complete a 3-response lever-press sequence and used RTC to demonstrate that the terminal response position in the sequence was more sensitive to the change in contingency (less RTC) than was the initial response position (greater RTC). Reid, Dixon, and Gray (2008) subsequently replicated and extended this finding with additional measures of RTC within sequences.

The general procedure of the current study compared RTC of the initial and terminal responses in a sequence of three pecks to different colored keys. Instead of switching to a different contingency, we provided free food during intercomponent intervals or prior to the experimental session in order to measure RTC. Thus, this procedure was similar to the multiple schedules frequently used in tests of BMT, but we were concerned with issues of response strength and stimulus value within the heterogeneous response sequence; that is, at a local level rather than at the molar level prescribed by BMT.

In three experiments we asked, "How do instrumental and Pavlovian contingencies contribute to RTC in learned response sequences?" The three panels of Figure 1 depict the response requirements and stimulus presentations for each trial of the three experiments. In Experiment 1, depicted by the top panel, pigeons were trained to complete a three-response sequence in which each

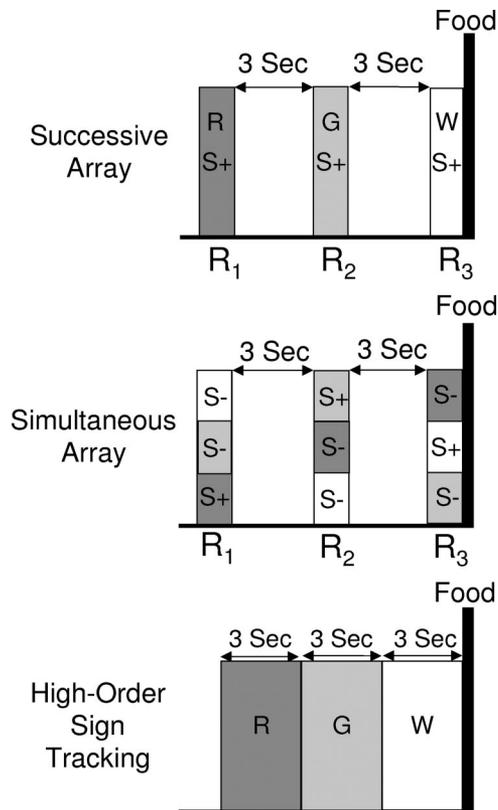


Figure 1. The three panels represent the different procedures used in Experiments 1–3.

stimulus (colored key lamp) was presented in strict succession. One and only one peck to each color in the correct order provided food, and pecks to the wrong color (e.g., darkened keys) produced timeout. In order to increase the difference in response strength or conditioned value between the initial and terminal responses of the sequence, delays of 3 s were imposed after the initial and middle responses, thus decreasing their temporal contiguity with reinforcement. Experiment 2 maintained similar instrumental contingencies, but degraded the Pavlovian contingencies by presenting all key colors simultaneously at each point in the sequence (middle panel). In this experiment the same pigeons were trained to complete a different three-response sequence in which all three colored key lamps were presented simultaneously and at randomized key locations at each ordinal position of the sequence (which Terrace and his colleagues have called a simultaneous chain; e.g., Straub & Terrace, 1981; Terrace, 1984, 2001, 2005). In Experiment 3 (bottom panel), the instrumental contingency was eliminated altogether in a high-order sign tracking procedure in which key pecks had no programmed consequences, yet the three key colors maintained similar Pavlovian relations with food delivery as in Experiment 1. Our goal in the three experiments was to identify the factors that establish the amount of RTC at each position within a response sequence.

Experiment 1: Successive Arrays

The main question for Experiment 1 was whether measures of RTC at each ordinal position of the sequence reflect the effects of

temporal contiguity between discriminated responding and reinforcement. A central claim of BMT is that RTC is an indicator of behavioral mass, which is a molar measure of response strength. If BMT can be extended to the local level of three-response sequences, then one would expect measurements of RTC to correspond to theoretical delay-of-reinforcement gradients. Thus, if RTC reflects response strength at the local level, then RTC of responding at each ordinal position should increase as the sequence progresses.

Because studies measuring RTC within unchanging heterogeneous response sequences are uncommon (but see Bachá-Mendez, Reid, & Mendoza-Soylovna, 2007; Reid, 1994; Reid et al., 2008), we must look elsewhere for relevant empirical results. Chained schedules of reinforcement provide a rich source of information relevant to our interest in simple response sequences. However, recall that chained schedules usually consist of many repetitions of the same response in each component rather than repeated “choices” between available heterogeneous responses. Because one of our goals is to compare RTC in chained schedules with RTC in simpler heterogeneous sequences, it is important to distinguish carefully between the two procedures: (A) First, the imposed contingencies are not the same. Reinforcement of a three-response sequence is not equivalent to a chained schedule containing three fixed-ratio 1 schedules. While both schedules do require a minimum of three pecks, reinforcement of response sequences in our discrete trials procedure requires the organism to respond exactly once (and only once) to each stimulus presented in succession. Responding twice to the same stimulus is an error, and pecking the wrong key (even a dark key) is also an error. Each response can be represented as a correct or incorrect choice between available alternatives. Errors immediately produce timeout and end the trial. Such errors and consequences are not possible with chained schedules. (B) Responding is not the same: Chained schedules typically involve repetitions of the same response, whereas the organism typically must learn the particular pattern of heterogeneous responses that produces reinforcement. Responding on chained schedules is typically more rapid than when response sequences are reinforced, which generates slower, more paced responding. (C) Stimulus control is different: Chained schedules provide distinctive discriminative stimuli that indicate the component of the chain currently in effect. When each component employs different schedules of reinforcement (e.g., fixed interval or fixed ratio), the temporal pattern of responding (the cumulative record) in each component is normally consistent with the temporal pattern of responding when that schedule is provided alone (Ferster & Skinner, 1957). These stimuli usually last considerably longer than the brief discriminative stimuli that set the occasion for each heterogeneous response in simple response sequences (minutes vs. seconds). The acquired value of the brief stimuli presented during response sequences and the strength of responding in the sequence are normally assumed to be affected by the local delay of reinforcement gradient, rather than the multiple patterns of cumulative records generated by chained schedules. And (D), the dependent variables are different in the two contingencies. Response rate during each component is the typical measure in chained schedules, but this measure is not possible with heterogeneous sequences because only one response can occur at each ordinal position in the sequence. Instead, response accuracy is the usual measure.

The three-component chained schedule used in the first experiment in Williams et al. (1995) compared the RTC of initial-link responding to that of middle-link responding when terminal-link responding was extinguished. They observed that RTC for the initial link was less than RTC for the middle link, a result compatible with an extended delay-of-reinforcement gradient but incompatible with a backward progression of extinction. They proposed that the more rapid extinction during the initial link occurred because the initial link was more strongly affected by a generalization decrement caused by removal of reinforcement from the situation. Therefore, in a second experiment they presented food in the middle of a 30-s intertrial interval (ITI) to ensure that the stimuli prior to initial-link responding were constant even when terminal-link reinforcement was removed. In contrast with their first experiment, they observed RTC for initial-link responding was greater than RTC for the middle link. This finding is important to the present experiment because it indicates that RTC in chained schedules is not simply a measure of local response strength or stimulus preference—RTC is also influenced by discriminative stimuli, particularly those affecting initial-link responding such as beginning-of-trial cues.

Our challenge was to measure RTC within a simple response sequence in which only one response can occur at each ordinal position in the sequence, and to ensure that the roles of discriminative stimuli (such as beginning-of-trial cues) were not altered in the process. Experiments testing BMT face similar constraints, so we adopted a similar procedure (although we could not utilize response rates at each ordinal position). Once pigeons were trained to peck once at three lighted keys presented in succession in a fixed sequence, this contingency was used as one component (called the SEQ component) of a two-component multiple schedule. The top panel of Figure 1 depicts the temporal characteristics of this contingency. For half of the subjects, a red keylight would turn on, and a single peck to this key would turn it off, followed by a 3-s delay in which no key lights were on. Then the green keylight would turn on, and a single peck to this key would turn it off, followed by another 3-s delay. Then the white keylight would turn on, and a single key peck would turn it off and activate the feeder. The next trial would begin immediately following reinforcement. The order of the lights was reversed for the other half of the subjects. Note that this procedure reinforced heterogeneous response sequences as described above. Pecking keys in the wrong order (even dark keys) produced timeout. It was not a chained schedule.

The goal was to compare RTC in the initial and terminal response positions. This was done in a separate preference component (PREF), which measured preference between the initial and terminal key colors presented simultaneously when responding to either color had identical consequences. Preference for one color over the other would reflect the preference created and maintained by completing the fixed response sequence in the SEQ component. RTC was measured by comparing responding to the initial and terminal colors in the PREF component when free food was provided during intercomponent intervals.

Method

Subjects. Two homing pigeons (P11 and P746) and two Silver King pigeons (P46 and P14) served as experimental subjects. Each

bird had previous experience in studies involving sequences of key pecks. They were maintained at 80% of their free-feeding weights by providing supplementary mixed grain after daily sessions. The pigeons were housed individually with free access to water and grit in a room that maintained relatively constant temperature and humidity under natural lighting conditions.

Apparatus. Two identical standard BRS140 pigeon chambers (24 cm wide, 35.5 cm long, and 29.5 cm high) were used. The front panel was equipped with three 2.5-cm response keys (Med Associates ENV-123 with tri-color displays) arranged horizontally, and a food hopper, 12.5 cm below the center key, provided access to mixed grain. Each key could be illuminated white, red, or green, and was located 15 cm above the floor. A force of approximately 0.15 N was necessary to operate each key. Each experimental chamber was enclosed in a sound-attenuating box equipped with a 7-W 120-V night light located behind the upper back wall of the chamber and a ventilation fan that provided air circulation and masked extraneous noise. Individual microcomputers programmed in Borland Pascal controlled each session and stored each event and its time of occurrence.

Procedure. Training: The purpose of the training phase was to train subjects to complete a sequence of individual pecks to three key colors presented one at a time in succession: red → green → white for Pigeons 11 and 14, and white → green → red for Pigeons 46 and 746. Green was always the middle color of the sequence, and it was always presented on the middle key of the chamber. To prevent subjects from anticipating the locations where the initial or terminal key colors would be presented, the initial color (red or white) was randomly assigned to the left or right key at the beginning of every trial. The terminal color was randomly assigned to the left or right key within each trial following the second response in the sequence.

A single peck to each key color immediately turned off that key lamp. A peck to the initial and middle colors of the sequence produced a 3-s period in which all key lights were off. Pecks to the terminal color of the sequence immediately produced 3-s food delivery without a delay, and the next trial began immediately afterward. These within-trial delays were included to ensure that the initial and terminal responses in the sequence differed substantially in their temporal contiguity with food presentation (more than 6 s for the initial color; approximately 0 s for the terminal color). We anticipated that differences in temporal contiguity would produce differences in response strength, stimulus value, or resistance to change in responding to the initial and terminal key colors, to be measured in later experimental phases. Infrequent errors (i.e., pecking darkened keys) produced an immediate 10-s timeout, and the next trial began.

This training phase was in effect for 20 sessions for all subjects. Experimental sessions occurred daily, seven days per week, at approximately the same time each day. Sessions lasted until 65 food presentations were delivered or 60 min expired.

Experimental baseline. The differences in responding to the initial and terminal key colors were measured in a two-component multiple schedule. The reinforcement contingencies in the SEQ component were identical to those in the training phase. Subjects were reinforced for completing a sequence of pecks to three key colors presented in succession, with a 3-s period with key lights off following the first and second responses. The purpose of this SEQ component was to maintain the differences (established during

training) in stimulus value, response strength, or resistance to change in responding to the initial and terminal key colors while these characteristics were measured in the other component of the multiple schedule. Each SEQ component lasted until five reinforcers were delivered.

The second component was a PREF component, used to measure preference between responding to the two key colors (red and white) representing the initial and terminal positions of the sequence. A single variable-interval (VI) 60-s reinforcement schedule was in effect for the PREF component. Red and white key colors were simultaneously presented on the left and right keys. Pecks to either key had identical consequences and probabilities of reinforcement on the VI schedule. That is, even though two key colors were simultaneously presented, a single VI schedule was in effect—it was not a concurrent VI–VI schedule. The red and white key colors were randomly assigned to the left and right keys at the beginning of the PREF component and after every food delivery. Because pecks to red and white had identical consequences and their positions were determined randomly, differences in choice should reflect differences in stimulus value or response strength that had developed when the response sequence was learned and practiced. Both key lights were turned off only during food delivery. Component duration was 3 min, and the single VI timer was stopped at the end of the component.

The PREF component was easily discriminated from the SEQ component. In the PREF component, two keys were simultaneously illuminated for the duration of the component (except during food delivery), and pecks to either key did not turn off that keylight. In the SEQ component, only one key was illuminated, and every peck immediately turned off that keylight for 3 s.

The two alternating components of the multiple schedule were always separated by intercomponent intervals lasting 90 s. During these periods, all key lights were off and pecks had no programmed consequences. The house light remained on. Sessions lasted until the two alternating components were each presented 10 times, typically about 85 min.

Resistance-to-change tests. As a RTC manipulation, free food was presented on a Variable Time (VT) schedule during the intercomponent intervals. Each RTC test reflected changes in the rates of responding to the two colors in the PREF component, not in the SEQ component. These tests involved no other changes in the multiple schedule used for the Experimental Baseline. Two separate RTC tests were done. The first test provided food on a VT 20-s schedule during the intercomponent intervals. Following recovery of the experimental baseline conditions, the second RTC test provided food on a VT 8-s schedule during the intercomponent intervals.

Table 1 describes the order of conditions and the number of sessions each subject was exposed to each condition.

Results

The PREF component provided the opportunity for subjects to respond differentially to the colors representing the initial and terminal positions of the sequence. The programmed consequences of responding to either key color were identical. Thus, preference for one color over the other should derive from its ordinal position in the SEQ component. For example, if greater response strength or conditioned value is correlated with closer temporal contiguity

Table 1
Order of Conditions and Number of Sessions in Each Condition for Each Subject in Experiment 1

Condition	Pigeon 11	Pigeon 46	Pigeon 14	Pigeon 746
1 Training	20	20	20	20
2 Baseline	45	29	93	83
3 VT 20-s	9	10	9	11
4 Baseline	9	9	7	9
5 VT 8-s	9	9	9	9

Note. Baseline refers to the experimental baseline condition utilizing the two-component multiple schedule. The two VT conditions refer to the two resistance-to-change tests in which free food was provided during the intercomponent intervals.

with food delivery, subjects should have responded at higher rates to the stimulus representing the terminal position in the sequence than to the stimulus representing the initial position. A repeated measures *t* test conducted on the last 10 sessions of the baseline condition indicated that two subjects, P14 and P746, did respond at higher rates to the terminal color in the PREF component during the baseline condition, P14: $t(9) = 3.16, p < .01$; P746: $t(9) = 3.91, p < .01$. However, P11 responded at a higher rate to the initial color, P11: $t(9) = 6.31, p < .001$, and P46 showed no preference for either color, P46: $t(9) = 0.24, p = .817$. Therefore, response rates across subjects were not systematically correlated with the initial and terminal positions in the PREF component during the baseline condition.

The main purpose of this experiment was to compare the resistance to change in responding to the colors representing the initial and terminal positions of the sequence. The procedure provided baseline response rates to each color, followed by an RTC manipulation in which free food was provided on a VT 20-s schedule during the periods separating the components of the multiple schedule. Following the first RTC manipulation, baseline was reestablished and subjects were shifted to a second RTC manipulation in which free food was delivered on a richer schedule (VT 8-s). This procedure provided two independent opportunities to measure RTC for each subject. Response rate in each condition was calculated by averaging daily response rates in the last five sessions for the initial and terminal colors in the PREF component (not the SEQ component). The left panels of Figure 2 show that for every subject, providing free food during the intercomponent interval produced greater reductions in responding to the colors representing the terminal position than to the initial position of the sequence. For all subjects, the initial position of the sequence was more resistant to change than the terminal position. The right panels of Figure 2 show that when free food was delivered a higher rate (VT 8-s), three of the four subjects showed the same effect: responding to the initial color was more resistance to change than was responding to the terminal color. P46 showed an unexplained reversal. Thus, in seven of eight independent tests, responding to the initial color was more resistant to change than to the terminal color.

We argue that the patterns of RTC during the PREF component reflect the effects of learning the response sequence during training and in the SEQ component. One potential alternative explanation is that the obtained reinforcement rates for each color during the

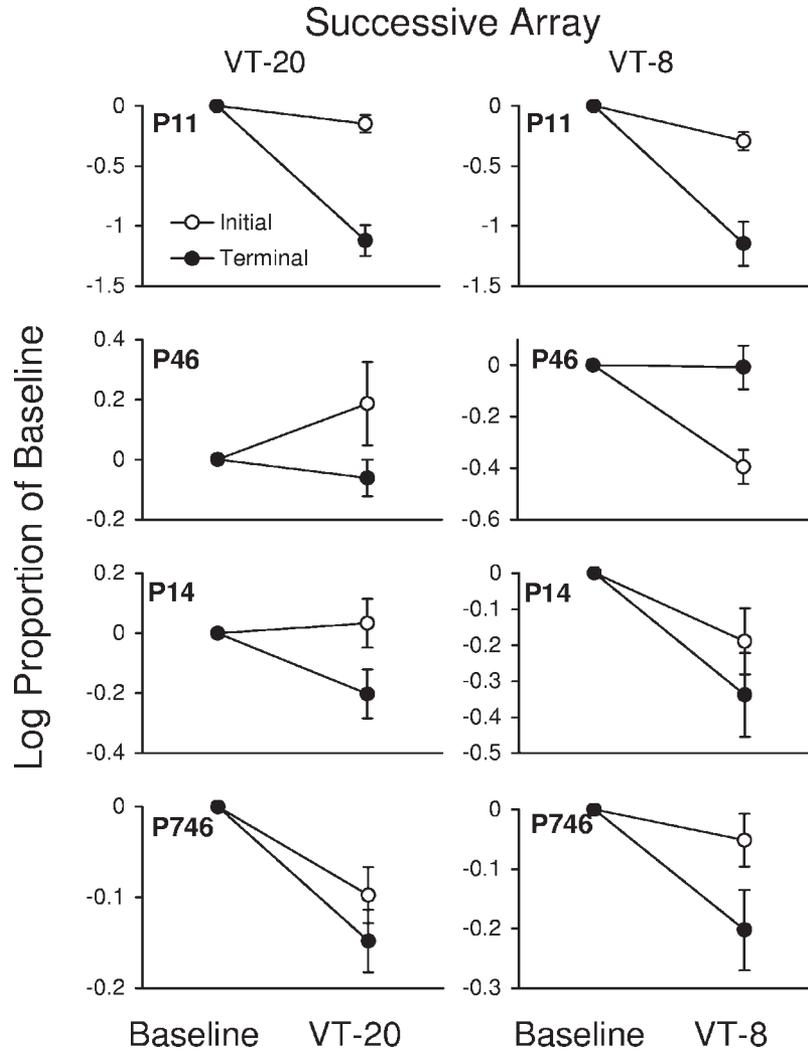


Figure 2. Each panel compares the resistance to change of responding in the initial versus terminal response positions of the response sequences in Experiment 1 when stimuli were presented as a successive array. Unearned food was provided on a VT 20-s schedule (left column) and on a VT 8-s schedule (right column) during the intercomponent intervals of the multiple schedule. Responding was measured during the Preference component. Error bars represent standard error of the mean.

PREF component varied with the subjects' response rates, and higher reinforcement rates for a particular color may have been systematically related to observed RTC. This potential explanation of our RTC patterns is easy to reject. The PREF component had a single VI schedule of reinforcement that treated responding to either color identically. More responding to one color would produce higher reinforcement rates for that color (recall that it was not a concurrent VI-VI schedule). Large differences in preference between the two colors would produce large differences in reinforcement rate. However, the obtained baseline response rates to the initial and terminal colors in the PREF component were not related to observed patterns of RTC. Pigeon 11 responded significantly more to the initial color ($M = 88.2/s$) than to the terminal color ($M = 19.5/s$) [$t(9) = 6.31, p < .001$]. Pigeon 46 responded nearly equally to the two colors (Initial = 57.1; Terminal = 54.8). Yet both Pigeons 14 and 746 responded significantly more to the

terminal color than to the initial color: Pigeon 14, Initial $M = 22.9$, Terminal $M = 38.6$ [$t(9) = -3.16, p < .01$]. Pigeon 746, Initial $M = 35.5$, Terminal $M = 75.3$ [$t(9) = -3.9, p < .01$]. Recall that Initial RTC > Terminal RTC in all subjects when food was provided by a VT-20 during intercomponent intervals. Therefore, even large differences in preference during baseline responding in the PREF component were unrelated to the observed patterns of RTC. Obtained relative reinforcement rates during the PREF component were not related to observed patterns of RTC, so the alternative explanation is invalid. The patterns of RTC during the PREF component must have reflected the effects of learning the response sequence during training and in the SEQ component.

Figure 3 shows the relationship between relative response rates and obtained reinforcement rate for both RTC manipulations, averaged across subjects. This figure represents the average effect observed in Figure 2 for the two RTC measurements. The baseline

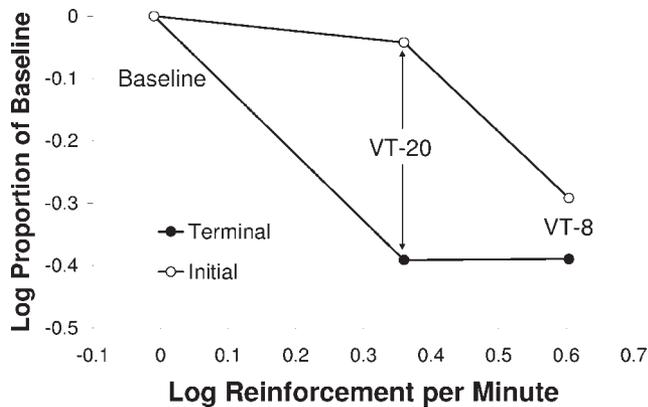


Figure 3. Comparison of resistance to change in initial versus terminal response positions in Experiment 1 as a function of reinforcement rate.

value was calculated by averaging across the last 5 days of both baselines for all subjects. The smaller separation between curves in the VT 8-s condition was produced by the inclusion of data from Pigeon 46, which showed a reversal in preference only in this condition. Overall, the figure shows greater persistence in responding to the colors representing the initial rather than the terminal response position in the sequence.

Discussion

Resistance to change in this experiment was not consistent with predictions based on theoretical delay-of-reinforcement gradients: the expectation of greater response strength or greater conditioned value due to greater temporal contiguity with reinforcement. We observed greater RTC in responding to the initial position rather than the terminal position. Thus, our results were not consistent with the behavioral momentum study of Nevin, Mandell, and Yarensky (1981) in which free food presented in a chained schedule produced greater RTC for the terminal component contiguous to reinforcement. The differences between their procedure and ours make the different findings difficult to interpret. For example, our procedure reflected our interest in the acquired structure of fixed response sequences: our subjects were required to learn a particular sequence of three responses to three different colors, and pecking colors in the wrong order produced timeout. The procedure used by Nevin et al. (1981) allowed subjects to peck keys repeatedly and in any order, and reinforcement produced by pecking an extra key was used as a disrupter. Thus, their procedure measured the effects of initial and terminal stimuli on RTC, but without the acquired structure of the response sequence that is the central issue of our experiments. We describe below other studies that have shared our emphasis on the structure of response sequences and have produced results entirely consistent with those of the current study (e.g., Reid, 1994; Reid et al., 2008).

Providing substantial amounts of free food during the intercomponent intervals can be considered an outcome devaluation procedure because each earned food delivery would have reduced value. Two lines of research using outcome devaluation procedures in chained schedules have produced results that are consistent with our RTC observations. One of these lines of research has centered on the transmission of value in chained schedules in the study of

the dynamics of conditioned reinforcement. For example, Williams et al. (1995) argued that the conditioned reinforcement value of stimuli in the terminal links of a chain must be lost before extinction of initial-link responding occurs. According to this argument, the backward progression of extinction often observed (Royalty, Williams, & Fantino, 1987; Williams, 1997, 1999; Williams et al., 1995; Williams & Royalty, 1990) is due to a backward transmission of reduced value. Our measures of RTC are consistent with a backward transmission of value to earlier positions in the sequence because the initial position showed greater persistence than did the terminal position.

The second line of research using chained schedules consistent with our results has centered on the incentive processes involved in instrumental conditioning. Balleine and his colleagues (Balleine, Garner, Gonzalez, & Dickinson, 1995; Corbit & Balleine, 2003) observed less RTC in the components of chained schedules closer to food delivery. Balleine (2001) proposed that the difference in RTC between components is due to differences in the associations formed between responding and certain features of the outcome representation. Because of differences in the salience of specific features of the outcome, responding proximal to food delivery is presumably associated with the sensory features of the outcome; whereas distal action (i.e., the initial link) is presumably associated with motivationally irrelevant features of the outcome. Thus, he claimed that outcome devaluation has a direct effect on responding in the terminal link but only reduced and delayed effects on the initial link.

Finally, Reid (1994) and Reid et al. (2008) have emphasized a third line of research consistent with our RTC results. Fortunately, this research utilized heterogeneous response sequences as in this study, rather than chained schedules. In both studies rats were trained to complete a sequence of three lever presses on two levers until accuracy was high and stable. Then the sequence producing reinforcement was changed, and this new contingency was maintained until the new sequence was learned. Using several measures of RTC, they observed that responding in the initial position of the sequence was substantially more persistent (greater RTC) than was responding in the terminal position. They concluded that the terminal position of a sequence is more sensitive to changes in reinforcement contingency than is the initial position.

The three converging explanations just described rely on different constructs to explain the same RTC results: conditioned value, specific associations, and sensitivity to contingency. In the two experiments that follow, we manipulated the relative contributions of Pavlovian and instrumental contingencies in a response sequence task to produce procedures that would generate divergent predictions from these explanations.

Experiment 2: Simultaneous Arrays

When response sequences involve stimuli presented in strict succession, as in chained schedules and in the successive array of Experiment 1, stimulus value is presumably mediated by Pavlovian associations (e.g., Balleine, 2001; Bell, 1999; Holland, 2004; Staddon, 2003; Ward-Robinson, 2004; Williams, 1997, 1999; Williams et al., 1995). Each stimulus in the sequence has a unique stimulus-reinforcer association because of the unique temporal contiguity between them. Pigeons can also learn to complete heterogeneous response sequences when all of the stimuli are

presented simultaneously at each position in the sequence. Terrace and his colleagues have called this procedure a “simultaneous chain” (Straub & Terrace, 1981; Terrace, 1984, 2001, 2005), but we prefer the term “simultaneous array” to avoid the theoretical implications of the term “chain” (particularly its historical implications about the mechanism of value transfer within response sequences). Pavlovian associations are severely degraded in a simultaneous array because all stimuli share the same combination of temporal contiguities with reinforcement, even though the subject must select only one key color at a time.

Because it is possible to make mistakes in the order of responding in sequences involving simultaneous arrays, the procedure involves contingencies unlike those of successive arrays and chained schedules. Pecks to the wrong key color immediately produce timeout and end the trial. The middle panel of Figure 1 depicts the simultaneous array used in Experiment 2. At each response position in the sequence, responses to two key colors would produce timeout (S⁻) and a response to only one key color would be correct (S⁺). In the initial position, the red key should be selected and the others avoided, but selection of red in the middle or terminal position would produce timeout. Therefore unlike in successive arrays, the value of each color must change with its ordinal position within the sequence, and each color must serve as S⁻ in two of the three positions in each trial. In this procedure, conditioned value should not be represented by the theoretical delay-of-reinforcement gradient generally ascribed to successive arrays and chained schedules.

Experiment 2 repeated the procedure used in Experiment 1, but a simultaneous array replaced the successive array. The same pigeons used in Experiment 1 were trained to complete a different three-response sequence, but the reinforcement contingency was otherwise the same as before. The central issue was how the degraded Pavlovian relations would influence RTC in the initial and terminal positions of the sequence.

Method

Subjects and apparatus. The subjects and apparatus were the same as those used in Experiment 1.

Procedure. The experimental procedure for Experiment 2 generally duplicated the conditions of Experiment 1 with one major exception. Subjects were trained to complete a three-response sequence of key pecks to three colors presented simultaneously on three keys, rather than in succession. The training phase was more involved in this experiment, because learning to complete response sequences on simultaneous arrays requires more training than sequences in which the stimuli are presented in succession. Table 2 describes the order of conditions and the number of sessions each subject was exposed to each condition. Experimental sessions occurred daily, 7 days per week, at approximately the same time each day.

Training. The purpose of the training phase was to train subjects to complete a sequence of individual pecks to three key colors presented simultaneously on three separate keys: white → green → red for Pigeons 11 and 14, and red → green → white for Pigeons 46 and 746. Thus, the target response sequences were reversed from those required for each subject in Experiment 1. Following the recommendations outlined by Straub and Terrace (1981) and Terrace (2001), subjects were trained to complete the sequence

Table 2
Order of Conditions and Number of Sessions in Each Condition for Each Subject in Experiment 2

Condition	Pigeon 11	Pigeon 46	Pigeon 14	Pigeon 746
1 Training: A	3	3	3	3
2 Training: AB	5	7	3	5
3 Training: ABC	27	50	24	37
4 Multiple baseline	24	21	21	21
5 VT 20-s	9	12	9	12
6 Multiple baseline	21	12	9	15
7 VT 8-s	9	9	9	12

Note. The three training conditions, labeled A, AB, and ABC, represent a forward conditioning training procedure. Multiple Baseline refers to the experimental baseline condition utilizing the two-component multiple schedule. The two VT conditions refer to the two resistance-to-change tests in which free food was provided during the intercomponent intervals.

using a forward chaining procedure. They were first trained to complete the initial response of the sequence, followed by training to complete the two-response sequence comprised of the initial and middle responses, followed by training to complete the entire three-response sequence. Sessions lasted until 65 food presentations were delivered or 60 min had elapsed.

To prevent subjects from anticipating the locations where the initial or terminal key colors would be presented, the initial and terminal colors were randomly assigned to the left or right key at the beginning of every trial and following each response within a trial. Green was always the middle color of the sequence, and it was always presented on the middle key of the chamber.

The middle panel of Figure 1 depicts the simultaneous array. A single peck to any key color immediately turned off all three key lamps. A correct peck to the initial and middle colors of the sequence produced a 3-s period in which all key lights were off and pecks had no programmed consequences. Correctly executed sequences of three pecks immediately produced 3-s food delivery without the delay, and the next trial began immediately afterward. As in Experiment 1, these within-trial delays were included to ensure that the initial and terminal responses in the sequence differed substantially in their temporal contiguity with food presentation. The temporal characteristics of stimulus presentation and the operant response contingencies of each trial were programmed to be identical to Experiment 1, although errors were more frequent. Errors produced an immediate 10-s timeout, and the next trial began. Errors were pecks to any color other than the current target at that ordinal position in the sequence. Therefore, each color served as S⁺ in one position, and S⁻ in the other two positions.

Experimental baseline. The differences in responding to the initial and terminal key colors were measured using the same two-component multiple schedule utilized in Experiment 1. However, the SEQ component now presented the key colors as a simultaneous array instead of the successive array of Experiment 1. The PREF component was programmed to be identical to that in Experiment 1, which measured preference between responding to the two key colors (red and white) representing the initial and terminal positions of the sequence. A single VI 60-s reinforcement schedule was in effect for the PREF component.

The PREF component was easily discriminated from the SEQ component. In the PREF component, two keys were simulta-

neously illuminated for the duration of the component (except during food delivery), and pecks to either key did not turn off that keylight. In the SEQ component, three keys were simultaneously illuminated, and every peck immediately turned off all three key lights for 3 s.

Resistance-to-change tests. The two RTC tests were the same as those used in Experiment 1. The first test provided food on a VT 20-s schedule during intercomponent intervals. Following recovery of the experimental baseline conditions, the second RTC test provided food on a VT 8-s schedule during intercomponent intervals.

Results

Figure 4 shows the response rates during the PREF component to the colors representing the initial and terminal positions of the response sequence with the simultaneous array. Examination of the eight baseline conditions shows greater response rates in five of the eight baseline conditions to the color representing the initial position than the terminal position of the sequence. Three baseline conditions showed no clear differences, but the terminal color was not favored in any baseline condition for any subject. Preference was sometimes transient (e.g., P11), but subjects always responded

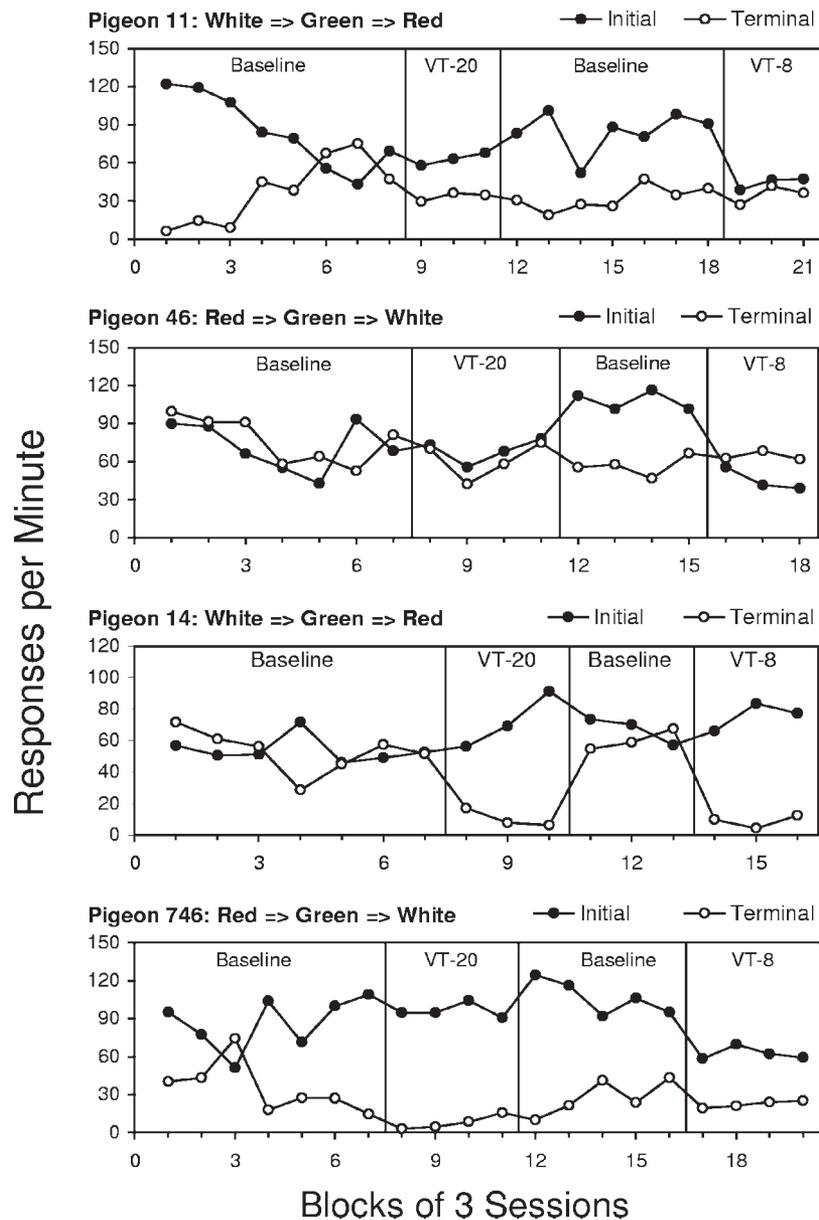


Figure 4. Comparison of average response rates to the key colors representing the initial and terminal response positions of the sequence. Response rates were measured during the Preference component of the multiple schedule.

more to the initial color when differences were observed. P14 did not show preferences toward either key color during the baseline conditions, but showed clear preference for the initial color when free food was provided on the VT schedules.

This procedure provided two independent opportunities to measure RTC for each subject. Response rate in each condition was calculated by averaging daily response rates for the initial and terminal colors in the last five sessions in each condition in the PREF component. The left panels of Figure 5 show that for every subject, providing free food during the intercomponent intervals produced greater reductions in responding to the colors representing the terminal position than to the initial position of the sequence. Thus, the initial position of the sequence was more resistant to change than the terminal position for all subjects. The right panels of Figure 5 show that when free food was delivered at a

higher rate, three of the four subjects showed the same effect: responding to the initial color was more resistance to change than was responding to the terminal color. P46 showed the same unexplained reversal that was observed in Experiment 1. Thus, in seven of eight independent tests, responding to the initial color was more resistant to change than to the terminal color.

Figure 6 shows the relationship between relative response rates and obtained reinforcement rate for both RTC manipulations averaged across subjects. This figure represents the average effect observed in Figure 5 for the two RTC measurements. The baseline value was calculated by averaging across the last 5 days of both baselines for all subjects. As observed in Experiment 1, the smaller separation between curves in the VT 8-s condition was produced by the inclusion of data from Pigeon 46, which showed a reversal in preference only in this condition. Overall, the figure shows

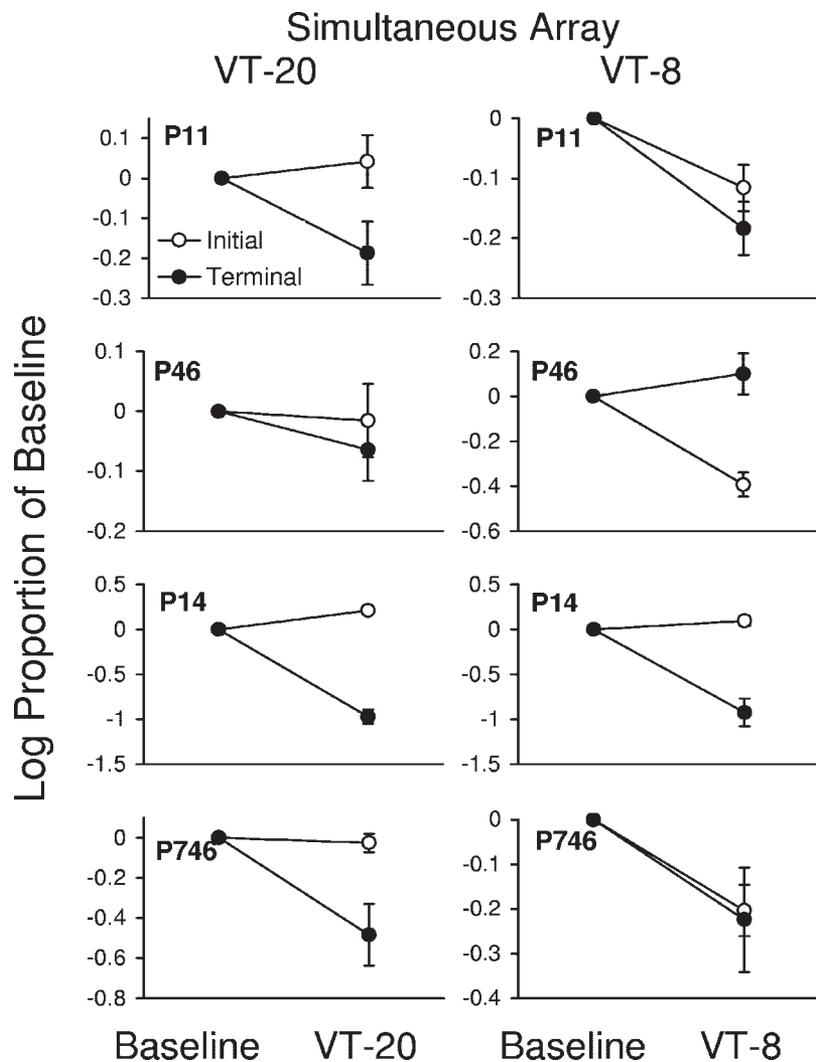


Figure 5. Each panel compares the resistance to change in responding in the initial versus terminal response positions of response sequences in Experiment 2 when stimuli were presented as a simultaneous array. Unearned food was provided on a VT 20-s schedule (left column) and on a VT 8-s schedule (right column) during the intercomponent intervals of the multiple schedule. Responding was measured during the Preference component. Error bars represent standard error of the mean.

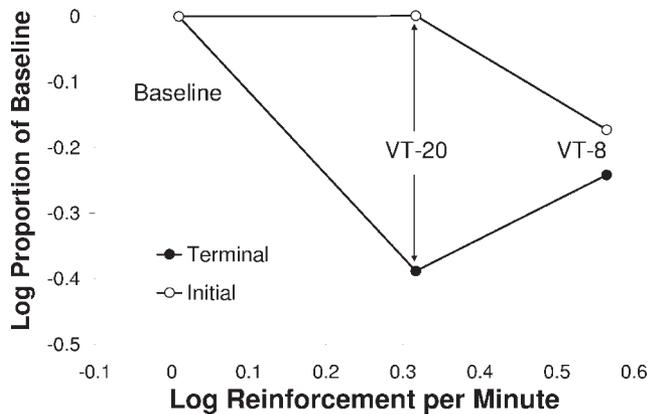


Figure 6. Comparison of resistance to change in initial versus terminal response positions in Experiment 2 as a function of reinforcement rate.

greater persistence in responding to the colors representing the initial rather than the terminal response position in the sequence.

Discussion

Two dependent variables, preference and RTC, were measured in this procedure in order to compare the characteristics of responding in the initial versus the terminal positions of learned response sequences. RTC was a more reliable measure across subjects than was preference, but preference was expressed consistently during baseline conditions for some of the subjects. When subjects did display preference in the PREF component, they responded at higher rates to the color representing the initial position of the sequence, rather than the terminal position. For some subjects, this preference was displayed throughout the experiment, and for other subjects, it was observed only in baseline or VT conditions. Only one subject (P46) in one condition (VT 8-s) responded at a greater rate to the terminal position. The question is, “Why would subjects respond at higher rates to the color representing the initial position of the sequence, given that all three colors were always presented simultaneously at each position?” Even if subjects attended most to the color that they actually pecked as they completed the sequence, it would be reasonable to expect the terminal color to be preferred because of its closer temporal contiguity with food delivery; but this was not observed.

One potential explanation for the observed preference might be that the subjects were not able to discriminate adequately between the PREF and SEQ components. That is, perhaps in the PREF component, subjects attempted to complete the learned sequence favoring the initial response, unaware that the contingencies had changed. This explanation is easily ruled out by comparison of the characteristics of each component and performance in each. Table 3 compares some of the discriminable characteristics of the two components and response rates in each. Both components were easily discriminated visually and by their response-produced feedback. Performance was dramatically different in the two components: response rates exceeded 60 pecks per min in the PREF component, but averaged about nine pecks per min in the SEQ component. All subjects demonstrated clear stimulus control. Therefore, confusion or lack of stimulus control was not responsible for the observations that

subjects responded at higher rates to the color representing the initial position in the sequence.

The second dependent variable, resistance to change, was reliably greater for the initial than the terminal position for all subjects in the VT 20-s condition, and for three of the four subjects in the VT 8-s condition. These RTC results closely replicate the surprising RTC results from Experiment 1, which presented the stimuli in succession—even P46 displayed the same reversal in the VT 8-s condition. The similarity between these data and those of Experiment 1 cannot be attributed to a continuation of preference or persistence acquired in Experiment 1 because the reinforced sequence was reversed for all subjects in Experiment 2. Thus, greater RTC for the initial position would reflect greater RTC for red in one experiment and for white in the other.

Even though presenting each stimulus simultaneously at each position in the sequence degraded the Pavlovian relations, the initial position of the sequence was more resistant to change than the terminal position. It is not clear how conditioned value would develop in a simultaneous array procedure. Unlike in successive arrays, the value of each color in a simultaneous array must change with its ordinal position within the sequence and each color must be actively avoided in two of the three positions in each trial. The color red, for example, could not acquire a fixed value within or across trials because its value must change dynamically second-by-second within each trial. Thus, conditioned value in this procedure could not be represented by the delay-of-reinforcement gradient in Experiment 1. Nevertheless, procedures produced greater RTC in the initial position of the sequence than in the terminal position, without regard to the alterations of the Pavlovian relations.

A similar line of reasoning holds for the concept of response strength, which has been shown to affect RTC in behavioral momentum studies. Straub and Terrace (1981), in a classic study of simultaneous arrays, evaluated several measures of responding to see if they conformed to a typical delay-of-reinforcement gradient. None of their measures (latency, the relative frequency of

Table 3
Comparison of Discriminable Differences Between the PREF and SEQ Components With the Simultaneous Array in Experiment 2

PREF component	SEQ component
2 keys lights on – only side keys used.	3 key lights on, green on the middle key.
Key lights remained on after every peck except during food delivery.	Key lights turned off for 3 or 10 seconds after every peck.
Colors remained on the same key until food delivery.	Colors were randomly assigned to keys after every peck.
No timeout was possible.	Timeout occurred for pecking wrong color.
Response rate averaged over 60 pecks per min.	Response rate averaged 9 pecks per min.
Subjects never pecked green because green was not available – subjects did not alternate pecks to red and white keys.	Subjects completed sequences accurately, including pecking green.

forward errors, and the relative frequency of repeat pecks) was consistent with a monotonic delay-of-reinforcement gradient. Therefore, our observations of greater RTC in the initial position did not depend upon the delay-of-reinforcement gradient, whether the gradient represents acquired response strength or conditioned value.

Experiment 3: High-Order Sign Tracking

Experiments 1 and 2 demonstrated that the differences in resistance to change between the initial and terminal positions of response sequences were essentially identical whether pigeons responded in a successive array or a simultaneous array, even though the Pavlovian relations between each color and food delivery differed substantially between the two procedures. Both procedures shared a common response-outcome relation: exactly three responses were always required for food delivery. Experiment 3 tested the importance of this instrumental relation on resistance to change by completely eliminating the response-outcome relation. Naïve pigeons were exposed to a high-order sign tracking procedure in which three key colors were presented in fixed succession for 3 s each, followed by food delivery and a 150-s ITI. Key pecks had no programmed consequences and did not influence the timing of food delivery. Once baseline response rates of elicited key pecking were established to each color, RTC was measured by prefeeding subjects prior to each daily session. Our key question was whether RTC for the initial color would be greater than that for the terminal color when the instrumental contingency was eliminated.

Method

Subjects. Four experimentally naïve male Silver King Pigeons were maintained at 85% of their free feeding weights by supplemental feeding after daily sessions. The housing conditions were the same as those in Experiments 1 and 2.

Apparatus. Four Gerbrands-style pigeon chambers 43 cm long, 35 cm wide, and 36 cm high were equipped with three 2.5-cm response keys (Med Associates ENV-124 with multicolored displays) in a row 26 cm above the floor. Each key required a force of approximately 0.15 N to operate, but only the center key was used. A BRS/LVE grain magazine was centered below the center key 10.5 cm from the floor, and the magazine was illuminated when wheat was presented. A 28-V house light (GE-1819) was located 6.5 cm above the center key. Ventilation fans masked extraneous noises. Each chamber contained a miniature color TV camera (PRO Video CVC-770PH) located at the top of the back wall of the chamber to monitor the subject. Individual personal computers programmed in Visual Basic.NET controlled the experiment and recorded every event and its time of occurrence.

Procedure. Baseline. After one session of magazine training, a high-order sign tracking procedure was used. Sessions consisted of 30 trials in which a 150-s ITI was followed by the presentation of three key colors in succession on the center key for 3 s each, followed by 3-s response-independent food delivery. Key pecks were recorded during the ITI and during each stimulus presentation, but they had no programmed consequences; that is, pecks did not terminate the stimulus or influence the timing of food delivery. The order of stimulus presentation for Pigeons 89 and 147 was

red → green → white, and for Pigeons 426 and 525 the order was white → green → red.

Prefeeding. After responding appeared stable in the presence of all three key colors, a prefeeding condition was implemented for 7 days. Pigeons were provided with mixed grain 1 hour before each experimental session. The procedure during each experimental session was not altered. The amount of prefeeding each day depended upon the subject's body weight. Pigeons 89 and 147 were given 20g on the first day of prefeeding, 40g on the second day, and 60g for the next 5 days. Pigeons 426 and 525 were given 15g the first day, 25g on the second, and 35g for the next 5 days. After the 7-day prefeeding condition was completed, subjects' 85% body weights were recovered, and the experimental baseline condition was reestablished.

Results

Figure 7 depicts the number of key pecks made to the colors at each ordinal position of the sequence averaged over sessions for each subject. All subjects responded less to the initial color than to the middle color. However, there was no systematic difference across subjects in responding to the middle and terminal colors: One subject (P525) responded more to the terminal color, one subject (P89) responded less, and the remaining two subjects responded the same amounts to both colors. Informal video observations of P147, P426, and especially P89 revealed fast pecking to the terminal color at the beginning of the 3-s period, followed by a lowering of the head toward the feeder just prior to food delivery. Thus, responding to the terminal color appeared to be reduced because of competition with activities related to food consumption.

Figure 8 depicts relative RTC when subjects were exposed to the prefeeding condition. For all subjects, persistence in responding was inversely related to ordinal position of the colors; thus RTC was directly related to temporal contiguity with reinforcement. Unlike the previous experiments, responding to the initial color was more disrupted than responding to the terminal color for all subjects.

The graded effect of prefeeding on responding to the three ordinal positions is more clearly evident in Figure 9. Points were calculated by averaging across the last five sessions in the baseline and prefeeding condition for each subject. The effect of prefeeding

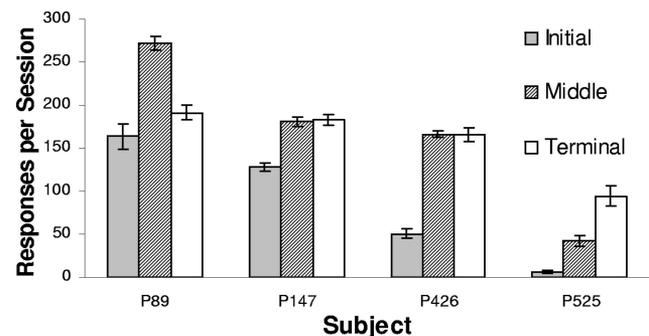


Figure 7. Amount of responding per session to the initial, middle, and terminal key colors in the high-order sign tracking study, Experiment 3, for each subject. Error bars represent standard error of the mean.

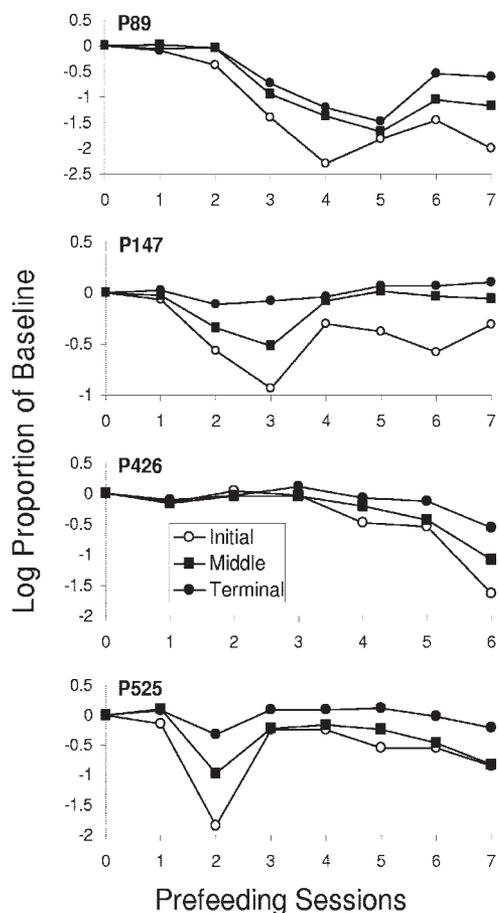


Figure 8. Comparison of the resistance to change in responding to the initial, middle, and terminal key colors measured during consecutive sessions of prefeeding in Experiment 3.

was the same for all subjects: RTC was greatest for the terminal color, less for the middle color, and least for the initial color.

Discussion

The high-order sign tracking procedure produced systematic, graded values of resistance to change in all subjects: Initial RTC < Middle RTC < Terminal RTC.

These results were directly opposite to those obtained in Experiments 1 and 2 in which an instrumental contingency was in effect. Presumably, elimination of the instrumental contingency in this procedure was responsible for the differences in results. Nevertheless, comparisons of instrumental and Pavlovian procedures necessarily involve multiple differences in the procedures of these experiments (such as different training histories, interreinforcement intervals, etc.) that could have contributed to these differences in results.

The obtained graded values of RTC are consistent with several theoretical frameworks. Within the Pavlovian conditioning literature, a large number of studies have examined the processes responsible for second-order conditioned responding. In a recent analysis of second-order autoshaping, Ward-Robinson (2004) con-

cluded that a CS2 \rightarrow US process is the causal mechanism for responding in second-order autoshaping, and devaluation of the US directly influences responding to CS2. Applied to the present study, this explanation would predict more responding to stimuli closer to food presentation. When prefeeding is used as an outcome devaluation procedure, RTC should be proportional to the strength of the Pavlovian relation. Thus, responding should decrease most with the stimuli farthest away from food delivery, as observed in the present study.

It is reasonable to expect conditioned value to increase with increasing temporal contiguity with food delivery. In their discussions of BMT, Nevin (1992) and Nevin and Grace (2000) have argued that Pavlovian relations are responsible for RTC, and RTC varies directly with conditioned value. The relevant Pavlovian relations in BMT, however, are molar relations measured by the relative number of reinforcers occurring in the presence of a particular stimulus (such as in a multiple schedule). In the present experiment, the Pavlovian relations exist within each trial as local relations that vary in temporal contiguity with food delivery. Therefore, if BMT were extended to the local level, it would predict the graded values of RTC obtained in this experiment.

General Discussion

When pigeons completed a sequence of three responses to stimuli presented in succession in Experiment 1, we observed greater RTC in responding to the initial position than in the terminal position of the sequence. RTC was not consistent with predictions based on theoretical delay-of-reinforcement gradients. Experiment 2 degraded the Pavlovian relation between each stimulus and reinforcement by presenting all three stimuli simultaneously at each point in the sequence. This procedure produced greater RTC in responding to the initial position than to the terminal position of the sequence, just as in Experiment 1. Experiment 3 reinstated the Pavlovian contingencies but completely eliminated the instrumental contingency by exposing naïve birds to a high-order sign tracking procedure using the same three stimuli as the other experiments. This procedure produced less RTC to the initial position than to the terminal position—the opposite of the other two experiments.

How can these different results be explained? We begin by examining the adequacy of three theoretical frameworks that make explicit predictions about RTC.

Behavioral Momentum Theory

BMT is perhaps the most well-developed explanation of RTC. Nevin and Grace (2000) claim that RTC is a measure of behavioral mass, reflecting acquired response strength. Preference between alternatives in concurrent schedules is another measure of behavioral mass. Thus, preference and RTC are purported to be correlated, converging measures of the same construct. Although BMT is a molar model, we can ask whether its predictions of RTC and preference at the local level of a response sequence would be compatible with the results of these experiments. Applied to response sequences, BMT would predict that RTC should vary directly with the height of a delay-of-reinforcement gradient, reflecting increasing response strength or conditioned value as the sequence progresses. The graded increase in RTC obtained in the

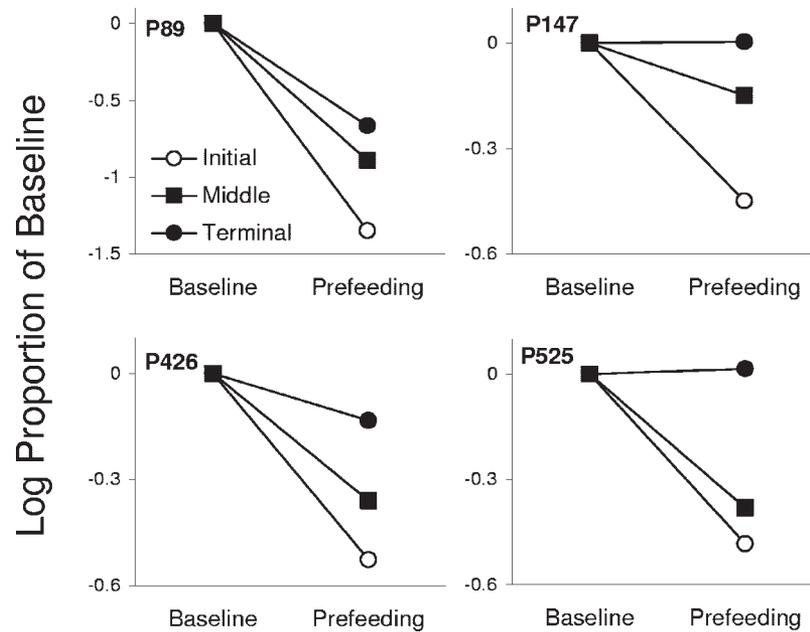


Figure 9. Each panel compares the resistance to change in responding to the initial, middle, and terminal key colors for each subject in the prefeeding condition of the high-order sign tracking procedure (Experiment 3).

high-order sign tracking experiment was consistent with this prediction. However, RTC in Experiments 1 and 2 varied in a direction opposite to predictions of a delay-of-reinforcement gradient. The observation of greater RTC for the initial than the terminal color was not consistent with the claim that RTC is a measure of response strength or preference within response sequences. Hence an extension of BMT to this local level cannot account for these results and would not be justified.

Dynamics of Conditioned Reinforcement

A second theoretical framework utilizing RTC has centered on the dynamics of conditioned reinforcement. Experiments have measured RTC following reinforcer devaluation or extinction in an attempt to measure the transfer of value across different components of chained schedules (e.g., Bell, 1999; Bell et al., 2007; Williams, 1997, 1999; Williams et al., 1995). In these studies, RTC of operant response rates presumably reflects conditioned reinforcement in different components of the chained schedule. Presumably, the pattern of RTC across components may reflect a backward propagation of reduced value following reinforcer devaluation or extinction (Williams et al., 1995). Experiments 1 and 3 presented stimuli in strict succession, so they are more similar to a chained schedule than is the simultaneous procedure used in Experiment 2. Our measures of RTC in Experiment 1 were consistent with a backward transmission of value to earlier positions in the sequence because the initial position showed greater persistence than did the terminal position. However, the opposite effect was observed with the high-order sign tracking experiment. Devaluation of the reinforcer by prefeeding the subjects produced a graded pattern of RTC compatible with a delay-of-reinforcement gradient, but incompatible with backward transmission of conditioned value. In one sense, this experiment was the most similar to

chained schedules because only in this experiment were subjects allowed to respond multiple times to each key color. It is not clear what role the absence of an instrumental contingency would have on the creation of conditioned reinforcement, given that traditional accounts of conditioned reinforcement claim that Pavlovian relations determine stimulus value. Thus, conditioned value in the early links of a chained schedule is presumably mediated by higher order conditioning—an opportunity clearly provided by this high-order sign tracking experiment.

The results of the simultaneous procedure of Experiment 2 provide another challenge to an explanation based on conditioned reinforcement and a backward transmission of reduced value. It is difficult to understand how conditioned reinforcement could develop when all stimuli were presented simultaneously at each point in the trial, and responding to the wrong color produced immediate timeout. In a simultaneous procedure, the value of a color must change on a moment-by-moment basis within each trial. The notion of conditioned reinforcement, stable from trial to trial, is inconsistent with a simultaneous procedure: Imagine pecking green in order to produce the red light that was avoided when green was chosen. Even though Experiment 2 in Williams et al. (1995) produced RTC effects that were in the same direction as those of Experiments 1 and 2 of the present study, their explanation cannot account for the results of the present three experiments.

Incentive Theory

Finally, researchers exploring the motivational control of instrumental conditioning often measure the effects of reinforcer (“outcome”) devaluation and extinction procedures on the persistence (i.e., RTC) of instrumental responding. Incentive theory makes explicit predictions about the relative amounts of RTC in response sequences and chained schedules. Balleine and his colleagues

(Balleine, 2001; Balleine et al., 1995; Corbit & Balleine, 2003) claim, “it has been demonstrated that a post-training shift in motivational state often has relatively little influence on subsequent instrumental performance unless animals have been allowed to consume the instrumental outcome in the new motivational state, that is, unless they have been given the opportunity for incentive learning” (Corbit & Balleine, 2003, p. 99). Accordingly, incentive theory predicts that initial RTC will be greater than terminal RTC in situations that do not provide the opportunity for incentive learning. Balleine et al. (1995) and Corbit and Balleine (2003) tested the effects of devaluation of the outcome in a single extinction session, before the opportunity for incentive learning was provided. They observed less RTC in the components of chained schedules closest to food delivery, just as we observed in Experiments 1 and 2. Balleine (2001) proposed that the difference in RTC between the initial and terminal components is due to differences in the particular associations formed between responding and certain features of the outcome representation. Chained schedules (and our Experiment 1) presumably create $R \rightarrow O$ associations in which responding to the terminal color overshadows responding to the initial color because of its greater temporal contiguity with food delivery. Overshadowing allows the terminal link to form associations with the most salient sensory features of the outcome, whereas the initial link forms associations with more diffuse, motivationally irrelevant features of the outcome. Thus, outcome devaluation purportedly has a direct effect on responding in the terminal link but only indirect and delayed effects on the initial link until incentive learning has occurred.

Incentive theory makes a different prediction, however, after incentive learning has occurred: once the animals are allowed to consume the devalued food in the new motivational state, incentive theory predicts no difference between initial and terminal RTC. Balleine et al. (1995, p. 207) explains that once incentive learning has occurred, the initial component of the chained schedule should be reduced “just as much” as the terminal component. The RTC procedure used in our Experiment 1 provided ample opportunity for incentive learning to occur. Our procedure allowed each subject to consume the devalued food in the new motivational state over 100 times per session for at least nine sessions. Therefore, incentive theory should have predicted no differences between initial and terminal RTC in Experiment 1, yet our results were more compatible with the theory’s predictions in the absence of incentive learning.

Explanation of our high-order sign tracking study is the most problematic for Balleine’s (2001) explanation. He argued that his model is consistent with Mackintosh’s (1983) model of value transfer in Pavlovian second-order schedules, in which the first-order and second-order CSs became associated with different attributes of the US. As an example, Balleine (2001) pointed out the compatibility in results between his chained schedules and those of Holland and Rescorla (1975), in which devaluation of the outcome in a Pavlovian second-order procedure immediately affected the conditioned response to the CS most contiguous to food delivery without affecting responding to the distal CS (but see the review by Ward-Robinson, 2004, for conflicting evidence). Therefore, Balleine (2001) argued that both instrumental and Pavlovian procedures share the same prediction that outcome devaluation should affect the terminal link more than the initial link. According to this logic, responding to the terminal color in our high-order

sign tracking study (Experiment 3) should have produced the least RTC, just as observed in the successive procedure of Experiment 1. Because Experiment 3 produced a clear, graded effect in the direction opposite to this prediction, Balleine’s (2001) explanation is not sufficient to account for our results.

An Alternative Approach

Since these three theoretical frameworks were unable to explain the results of all three experiments, is there a more satisfactory alternative? Unfortunately, the research literature on RTC in response sequences remains too sparse for the development of a comprehensive theory. We can, however, propose a tentative hypothesis that may serve as a step toward a more comprehensive theory of learned response sequences. Reid’s (1994) results (discussed earlier) and conclusions provide guidance toward a quantitative explanation. Recall that Reid trained rats to complete a three-response sequence. After extended exposure to the training sequence, subjects were shifted to a new target sequence that differed from the previous sequence by the response required in the first or last position of the sequence. Measurements of initial and terminal RTC demonstrated that the terminal response position was more sensitive to the change in contingency (less RTC) than was the initial response position (greater RTC) (see also Reid et al., 2008, which extended this procedure and obtained equivalent results with multiple measures of RTC). This finding is consistent with the results of Experiments 1 and 2.

Reid (1994) and Reid et al. (2008) argued that this procedure of switching the reinforcement contingency to extinction or devaluing the reinforcer often pits two separate processes against each other, and the resulting patterns of RTC will depend upon the quantitative details of each process: (1) After extended reinforcement of the response sequence, the response strength at each position of the sequence should be compatible with a positively accelerated delay-of-reinforcement gradient, such as Mazur’s (1984) hyperbolic decay function; and (2) The effects of extinction or reinforcer devaluation should produce a graded decrease at each response position with the greatest effect on the terminal response, usually described as a backward progression of extinction or a backward transfer of reduced value. To these two processes, we must add a third: (3) Cues such as food delivery often serve as powerful discriminative stimuli indicating the beginning of a new trial, and these cues differentially affect the probability (and thus the RTC) of the first response in the sequence (for justifications to this claim, see Reid et al., 2008; Reid, Chadwick, Dunham, & Miller, 2001; Staddon, 2003; Williams et al., 1995). Reid (1994) argued that if the changes in response strength due to nonreinforcement or reinforcer devaluation are steeper than the acquired response strength gradient, then RTC for the terminal position would be less than that for the initial position (as in Experiments 1 and 2).

By quantifying this verbal description, we can improve our understanding of the interactions between these processes and make predictions of RTC patterns. While we do not know the precise forms of the processes above, we can adapt Mazur’s (1984) hyperbolic decay function for each. [Note: this assumption is not critical—other monotonic decay functions such as reciprocals, exponential, and the hyperbolic with exponent produce similar predicted patterns of RTC]. Assume that the steady-state delay-of-

reinforcement gradient produced by repeated reinforcement of a fixed response sequence can be represented by:

$$V_i = \frac{A}{1 + k_r D_i} \quad (1)$$

in which V_i represents the value (or strength) of a response occurring D_i seconds before food delivery, A represents the value of the reinforcer at delay = 0, and k_r represents the rate of decay during reinforcement conditions.

For simplicity, assume that nonreinforcement and devaluation of the reinforcer affect the last response of the sequence more than the earlier responses in a way exactly consistent with the delay-of-reinforcement gradient, but the slope is inverted to produce an inhibitory gradient represented by I_i :

$$I_i = \frac{-A_e}{1 + k_e D_i} \quad (2)$$

The variable A_e represents the inhibitory influence of the disruptor (or nonreinforcement or reinforcer devaluation) at delay = 0. The variable k_e represents the rate of decay in nonreinforcement and devaluation conditions, which may differ from k_r in the reinforcement conditions described by Equation 1. These two variables may be manipulated independently by manipulating characteristics of the reinforcement conditions and devaluation conditions separately. For example, k_e can be manipulated by providing large or small amounts of free food (or "surprising" vs. "anticipated" free food), without changing k_r , the slope of the learned delay-of-reinforcement gradient.

Some procedures provide strong beginning-of-trial cues, such as food delivery in discrete trials procedures, which act as powerful discriminative stimuli that increase the probability of correctly producing the first response in the sequence. In Experiments 1 and 2, for example, food delivery signaled the immediate beginning of the next trial. Other procedures may not provide these beginning-of-trial cues. For example, Experiment 3 was a high-order sign-tracking study in which the beginning of each sequence followed a 150-s ITI, so it is unlikely that prior food delivery was a strong influence on responding to the first color. Assume that when such cues are provided, RTC for only the first response position is increased in proportion to its original value during reinforcement conditions:

$$\beta \frac{A}{1 + k_r D_1}, \quad (0 \leq \beta < 1) \quad (3)$$

in which D_1 represents the delay between the first response and reinforcement, and β represents the degree of influence of the beginning-of-trial cues on that first response.

Following extinction or devaluation conditions, the predicted RTC values of each response in a sequence can be calculated with or without the influence of beginning-of-trial cues, leading to two related equations. Ignoring beginning-of-trial cues, predicted RTC for each response (RTC_i) is simply the sum of Equations 1 and 2:

$$RTC_i = \frac{A}{1 + k_r D_i} - \frac{A_e}{1 + k_e D_i} \quad (4)$$

When the procedure does provide prominent beginning-of-trial cues that influence the first response of the sequence, then Equa-

tion 5 (which we call the gradient model of RTC) represents RTC only for that first response (RTC_1), and Equation 4 would continue to predict RTC for all other responses in the sequence (or equivalently, $\beta = 0$).

$$RTC_1 = \frac{A}{1 + k_r D_1} - \frac{A_e}{1 + k_e D_1} + \beta \frac{A}{1 + k_r D_1}, \quad (5)$$

Ignoring beginning-of-trial cues first, the gradient model predicts two different response patterns: when $k_r \ll k_e$, Initial RTC > Terminal RTC; otherwise, Terminal RTC > Initial RTC. The first case would occur with strong devaluation of the reinforcer (high k_e) or when procedures have produced high expectations of reinforcement (low k_r). If the changes due to nonreinforcement or reinforcer devaluation were not as steep, then RTC for the terminal position would be greater than the initial position. This should occur when nonreinforcement is not as surprising or when reinforcer devaluation is weaker. Therefore, the gradient model predicts two possible observations, even while holding A_e constant and ignoring all beginning-of-trial cues, depending upon the relative values of k_r and k_e (which may be dissociated experimentally).

The same two different response patterns are predicted when k_r and k_e are held constant and A_e (the strength of the disruptor or reinforcer devaluation) is manipulated. With low values of A_e , Initial RTC is predicted to be lower than Terminal RTC. However, with larger values of A_e , Initial RTC is predicted to be higher than Terminal RTC. Once again, the gradient model can predict both patterns of RTC even when the contribution of beginning-of-trial cues (Equation 3) is nonexistent or ignored. When the effects of beginning-of-trial cues (β) are strong, the model predicts that Initial RTC will exceed Terminal RTC in most experimental conditions.

This gradient model indicates that three processes, which may be easily dissociated experimentally, may determine the pattern of RTC in response sequences: (a) the relative sizes of the decay rates (k_r and k_e) of the distributions described by Equations 1 and 2; (b) the strength, A_e , of the inhibitory influence of nonreinforcement or devaluation of the reinforcer; and (c) the salience of cues that signal the beginning of a trial or response sequence, β , which influences the probability of the first response in the sequence. Each of these three processes may independently produce two different patterns of RTC (even while holding the other factors constant), predicting Initial RTC > Terminal RTC in some situations and the opposite in others. Even though each of these processes may be manipulated independently in experiments, the resulting patterns of RTC in response sequences will depend upon the combination of all three factors.

This model helps illuminate the combination of factors leading to the RTC patterns in the three experiments reported here. The top panel of Figure 10 shows how the two theoretical gradients can combine to produce RTC predictions for the successive array procedure. We assume that extended training to complete the response sequence would produce a delay-of-reinforcement gradient with a slow decay ($k_r = .05$), as depicted by the curve at the top of the figure. The two VT schedules in the RTC conditions provided high rates of free food ($A_e = 1$) that dramatically decreased the value of earned food. Because subjects were exposed to these free-food conditions abruptly and for a relatively short period of time, we assume that the inhibitory gradient would decay

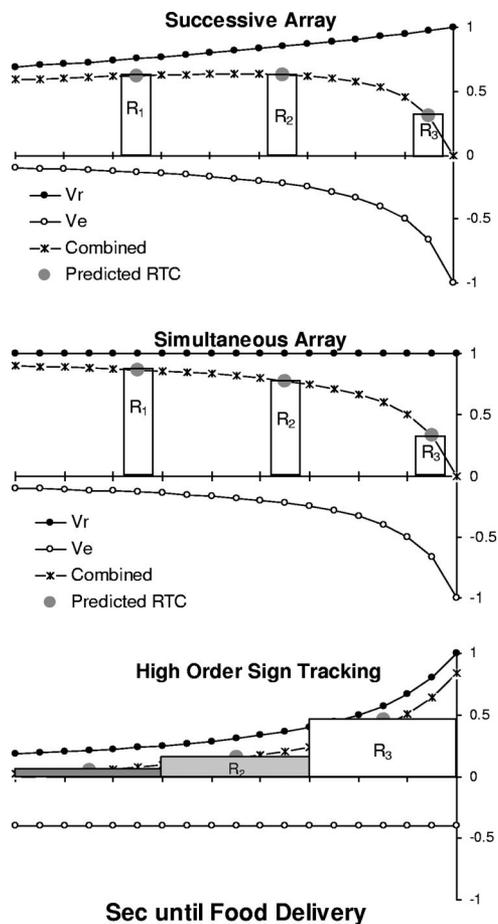


Figure 10. The three panels represent the model's predictions of patterns of RTC in Experiments 1–3. V_r represents the assumed delay-of-reinforcement gradient produced by the procedure, and V_e represents the devaluation gradient. The combined gradient represents the results of combining V_r and V_e according to Equation 4. Large filled circles represent predicted values of RTC for the initial, middle, and terminal response positions of the sequence. Possible effects of beginning-of-trial cues are ignored in this figure.

more rapidly ($k_e = 1$) than the delay-of-reinforcement gradient ($k_e \gg k_r$). When these two gradients are combined, both factors (high A_e and $k_e \gg k_r$) would lead to the prediction of Initial RTC > Terminal RTC, as observed. For simplicity, we did not include the potential influence of beginning-of-trial cues to generate the predictions of RTC depicted in Figure 10 (i.e., Equation 4 was used, not Equation 5). The successive array procedure did provide distinctive beginning-of-trial cues (food delivery immediately preceded the sequence, thus $\beta > 0$), so the predictions of greater Initial RTC than Terminal RTC using the complete model (Equation 5) would be somewhat more than the figure depicts. The figure demonstrates that predictions of Initial RTC > Terminal RTC need not rely on the influence of beginning-of-trial cues.

The middle panel of Figure 10 depicts the model's predictions (Equation 4) for the simultaneous array procedure of Experiment 2. Even though little is currently known about the Pavlovian and instrumental processes involved in simultaneous arrays, the model

makes simple predictions about the expected pattern of RTC across the sequence. Degradation of the Pavlovian relations between each key color and food delivery and punishment of response errors with timeout would prevent the development of the normal delay-of-reinforcement gradient described by Equation 1. The subjects had to select and avoid each key color on a moment-to-moment basis during each trial, so we assume that the gradient would be relatively flat ($k_r = 0$), as represented by the top curve. Thus the high rates of free food ($A_e = 1$) and strong inhibitory gradient ($k_e = 1$) would most strongly influence the terminal response, leading to the prediction of Initial RTC > Terminal RTC, just as observed. This prediction is strengthened by providing food delivery or timeout immediately preceding each trial, producing strong beginning-of-trial cues (high β) that would further increase Initial RTC.

The high-order sign tracking study (Experiment 3) provides an interesting contrast. The strong Pavlovian relations between key color and food delivery should produce a prominent delay gradient (Equation 1), depicted in the bottom panel of Figure 10. The interesting question concerns the role of Equation 2, the inhibitory influence of prefeeding prior to experimental sessions. How does reinforcer devaluation affect responding in the presence of strong $S \rightarrow O$ associations and the absence of $R \rightarrow O$ associations? Value transfer is immediate across all key colors, rather than producing a backward progression (see Ward-Robinson, 2004, for a review of value transfer in second-order autoshaping). Thus, prefeeding in the high-order sign tracking experiment would devalue the reinforcer and reduce the value of each color, presumably in proportion to its original value, which depended upon its position in the sequence. When combined with the fact that the long 150-s ITI would eliminate the contribution of beginning-of-trial cues (Equation 3), the model predicts a prominent graded increase in RTC across the ordinal positions of the sequence, just as observed.

The gradient model is not designed to account for the response patterns of chained schedules containing multiple or nonmonotonic delay gradients. However, when a chained schedule is composed of identical components based on VI or random interval (RI) schedules, then a single monotonic delay gradient may be more applicable (e.g., Balleine et al., 1995; Nevin et al., 1981; Williams et al., 1995).

We propose that the disparate findings from apparently similar procedures may be accounted for by our simple model based on monotonic delay gradients. This tentative model does not rely on differences in the types of $R \rightarrow O$ associations learned (cf. Balleine, 2001). It does not assume that RTC is a relative pure measure of any construct, such as response strength, at the local level of response sequences (cf. Nevin & Grace, 2000). Instead, the relative amount of RTC observed in a response sequence should be a function of three separate variables that influence transition probability: the relative slopes of the delay-of-reinforcement gradient and the gradient representing nonreinforcement or reinforcer devaluation, the intensity of the devaluation procedure, and the effectiveness of discriminative stimuli such as beginning-of-trial cues or end-of-trial cues. These factors have long been recognized as contributors to responding in chained schedules and response chains. This model simply shows that these factors can be represented quantitatively and combined to make explicit predictions about RTC patterns in response sequences.

RTC within response sequences will be influenced by any variable that affects transition probabilities. When considering only chained schedules and successive arrays, the number of variables involved may appear to be relatively small. Researchers using these procedures have often described RTC as a measure of acquired response strength (e.g., Nevin et al., 1981). With simultaneous arrays, it becomes clear that the measurement of response strength in a response sequence is always confounded with multiple factors that influence transition probability, including performance variables unrelated to acquired response strength. Accurate responding in simultaneous procedures is considerably more difficult and more easily disrupted because the environment does not provide the step-by-step stimuli to occasion the appropriate response. Subjects often make forward errors in the response sequence (“skipping a step”), and brief interruptions disrupt sequence accuracy (Terrace, 1984, 2005). Although RTC is a powerful tool useful for measuring the combined effects of the three processes described, RTC is not a direct measure of response strength or conditioned value in heterogeneous response sequences.

References

- Adams, C. D., & Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *Quarterly Journal of Experimental Psychology*, *33B*, 109–122.
- Bachá-Méndez, G., Reid, A. K., & Mendoza-Soylovna, A. (2007). Resurgence of integrated behavioral units. *Journal of the Experimental Analysis of Behavior*, *87*, 5–24.
- Balleine, B. W. (2001). Incentive processes in instrumental conditioning. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories*. Mahwah, N. J.: Erlbaum.
- Balleine, B. W., Garner, C., Gonzalez, F., & Dickinson, A. (1995). Motivational control of heterogeneous instrumental chains. *Journal of Experimental Psychology: Animal Behavior Processes*, *21*, 203–217.
- Bell, M. C. (1999). Pavlovian contingencies and resistance to change in a multiple schedule. *Journal of the Experimental Analysis of Behavior*, *72*, 81–96.
- Bell, M. C., Seip, K. M., & Fitzsimmons, K. S. (2007). Effect of signaling reinforcement on resistance to change in a multiple schedule. *Behavioural Processes*, *74*, 33–48.
- Colwill, R. M., & Motzkin, D. K. (1994). Encoding of the unconditioned stimulus in Pavlovian conditioning. *Animal Learning & Behavior*, *22*, 384–394.
- Colwill, R. M., & Rescorla, R. A. (1988). Associations between the discriminative stimulus and the reinforcer in instrumental learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 155–164.
- Colwill, R. M., & Rescorla, R. A. (1990). Evidence for the hierarchical structure of instrumental learning. *Animal Learning & Behavior*, *18*, 71–82.
- Corbit, L. H., & Balleine, B. W. (2003). Instrumental and Pavlovian incentive processes have dissociable effects on components of a heterogeneous instrumental chain. *Journal of Experimental Psychology: Animal Behavior Processes*, *29*, 99–106.
- Crawford, L. L., & Domjan, M. (1995). Second-order sexual conditioning in male Japanese quail (*Coturnix japonica*). *Animal Learning & Behavior*, *23*, 327–334.
- Dickinson, A., Balleine, B. W., Watt, A., Gonzalez, F., & Boakes, R. A. (1995). Motivational control after extended instrumental training. *Animal Learning & Behavior*, *23*, 197–206.
- Doughty, A. H., & Lattal, K. A. (2001). Resistance to change of operant variation and repetition. *Journal of the Experimental Analysis of Behavior*, *76*, 195–215.
- Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York: Appleton-Century-Crofts.
- Grace, R. C., Schwendimann, J. W., & Nevin, J. A. (1998). Effects of unsignaled delay of reinforcement on preference and resistance to change. *Journal of the Experimental Analysis of Behavior*, *69*, 247–261.
- Holland, P. C. (1990). Event representation in Pavlovian conditioning: Image and action. *Cognition*, *37*, 105–131.
- Holland, P. C. (2004). Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *Journal of Experimental Psychology: Animal Behavior Processes*, *30*, 104–117.
- Holland, P. C., & Rescorla, R. A. (1975). The effect of two ways of devaluing the unconditioned stimulus after first- and second-order appetitive conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *1*, 364–373.
- Hull, C. L. (1943). *Principles of behavior*. New York: Appleton-Century-Crofts.
- James, W. (1890). *The principles of psychology*. Vol. 1. New York: Holt.
- Konorski, J. (1967). *Integrative activity of the brain*. Chicago: University of Chicago Press.
- Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior. The Hixon Symposium* (pp. 112–146). New York: Wiley.
- Mackintosh, N. J. (1983). *Conditioning and associative learning*. New York: Oxford University Press.
- Mazur, J. E. (1984). Tests of an equivalence rule for fixed and variable reinforcer delays. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 426–436.
- Mazur, J. E., & Wagner, A. R. (1982). An episodic model of associative learning. In M. L. Commons, R. L. Herrnstein, & A. R. Wagner (Eds.), *Quantitative analyses of behavior: Acquisition* (pp. 3–39). Cambridge, MA: Ballinger Publishing.
- Nairne, J. S., & Rescorla, R. A. (1981). Second-order conditioning with diffuse auditory reinforcers in the pigeon. *Learning and Motivation*, *12*, 65–91.
- Neuringer, A., Kornell, N., & Olufs, M. (2001). Stability and variability in extinction. *Journal of Experimental Psychology: Animal Behavior Processes*, *27*, 79–94.
- Nevin, J. A. (1992). An integrative model for the study of behavioral momentum. *Journal of the Experimental Analysis of Behavior*, *57*, 301–316.
- Nevin, J. A., & Grace, R. C. (2000). Behavioral momentum and the law of effect. *Behavioral and Brain Sciences*, *23*, 73–130.
- Nevin, J. A., & Grace, R. C. (2005). Resistance to extinction in the steady state and in transition. *Journal of Experimental Psychology: Animal Behavior Processes*, *31*, 199–212.
- Nevin, J. A., Grace, R. C., Holland, S., & McLean, A. P. (2001). Variable-ratio versus variable-interval schedules: Response rate, resistance to change, and preference. *Journal of the Experimental Analysis of Behavior*, *76*, 43–74.
- Nevin, J. A., Mandell, C., & Yarensky, P. (1981). Response rate and resistance to change in chained schedules. *Journal of Experimental Psychology: Animal Behavior Processes*, *7*, 278–294.
- Podlesnik, C. A., Jimenez-Gomez, C., Ward, R. D., & Shahan, T. A. (2006). Resistance to change of responding maintained by unsignaled delays to reinforcement: A response-bout analysis. *Journal of the Experimental Analysis of Behavior*, *85*, 329–347.
- Reed, P., & Doughty, A. H. (2005). Within-subject testing of the signaled-reinforcement effect on operant responding as measured by response rate and resistance to change. *Journal of the Experimental Analysis of Behavior*, *83*, 31–45.
- Reed, P., & Morgan, T. A. (2006). Resurgence of response sequences

- during extinction in rats shows a primacy effect. *Journal of the Experimental Analysis of Behavior*, 86, 307–315.
- Reid, A. K. (1994). Learning new response sequences. *Behavioural Processes*, 32, 147–162.
- Reid, A. K., Chadwick, C. Z., Dunham, M., & Miller, A. (2001). The development of functional response units: The role of demarcating stimuli. *Journal of the Experimental Analysis of Behavior*, 76, 303–320.
- Reid, A. K., Dixon, R. A., & Gray, S. (2008). Variation and selection in response structures. In N. Innis (Ed.), *Reflections on Adaptive Behavior: Essays in Honor of J. E. R. Staddon* [pp. 51–86]. Cambridge, MA: MIT Press.
- Rescorla, R. A. (1990). Instrumental responses become associated with reinforcers that differ in one feature. *Animal Learning & Behavior*, 18, 206–211.
- Rescorla, R. A. (1992). Response outcome versus outcome response associations in instrumental learning. *Animal Learning & Behavior*, 20, 223–232.
- Royalty, P., Williams, B. A., & Fantino, E. (1987). Effects of delayed conditioned reinforcement in chain schedules. *Journal of the Experimental Analysis of Behavior*, 47, 41–56.
- Shahan, T. A., & Podlesnik, C. A. (2005). Rate of conditioned reinforcement affects observing rate but not resistance to change. *Journal of the Experimental Analysis of Behavior*, 84, 1–17.
- Skinner, B. F. (1938). *The behavior of organisms*. New York: Appleton-Century-Crofts.
- Staddon, J. E. R. (2003). *Adaptive behavior and learning*, 2003 Ed. [Electronic version]. Retrieved July 14, 2008 from <http://psychweb.psych.duke.edu/department/jers/abl/TableC.htm>
- Straub, R. O., & Terrace, H. S. (1981). Generalization of serial learning in the pigeon. *Animal Learning and Behavior*, 9, 454–468.
- Terrace, H. S. (1984). Simultaneous chaining: The problem it poses for traditional chaining theory. In M. L. Commons, R. J. Herrnstein, & A. R. Wagner (Eds.), *Quantitative Analyses of Behavior: Discrimination Processes* (pp. 115–138). Cambridge, MA: Ballinger.
- Terrace, H. S. (2001). Chunking and serially organized behavior in pigeons, monkeys, and humans. In R. G. Cook (Ed.), *Avian visual cognition*. Retrieved from <http://www.pigeon.psy.tufts.edu/avc/terrace/>
- Terrace, H. S. (2005). The simultaneous chain: A new look at serially organized behavior. In E. A. Wasserman & T. R. Zentall (Eds.), *Comparative cognition: Experimental explorations of animal intelligence* (pp. 481–514). New York: Oxford University Press.
- Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review Monograph Supplements*, 2, (No. 4).
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55, 189–208.
- Wagner, A. R., & Brandon, S. E. (1989). Evolution of a structured connectionist model of Pavlovian conditioning (AESOP). In S. B. Klein & R. R. Mower (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory* (pp. 149–189). New Jersey: Erlbaum.
- Ward-Robinson, J. (2004). An analysis of second-order autoshaping. *Learning and Motivation*, 35, 1–24.
- Williams, B. A. (1994). Conditioned reinforcement: Neglected or outmoded explanatory construct? *Psychonomic Bulletin & Review*, 1, 457–475.
- Williams, B. A. (1997). Conditioned reinforcement dynamics in three-link chained schedules. *Journal of the Experimental Analysis of Behavior*, 67, 145–159.
- Williams, B. A. (1999). Value transmission in discrimination learning involving stimulus chains. *Journal of the Experimental Analysis of Behavior*, 72, 177–185.
- Williams, B. A., Ploog, B. O., & Bell, M. C. (1995). Stimulus devaluation and extinction of chain schedule performance. *Animal Learning & Behavior*, 23, 104–114.
- Williams, B. A., & Royalty, P. (1990). Conditioned reinforcement versus time to reinforcement in chain schedules. *Journal of the Experimental Analysis of Behavior*, 53, 381–393.

Received July 24, 2007

Revision received August 26, 2008

Accepted August 26, 2008 ■

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <http://notify.apa.org/> and you will be notified by e-mail when issues of interest to you become available!