Published in final edited form as:

Behav Processes. 2009 June ; 81(2): 195–204. doi:10.1016/j.beproc.2009.01.007.

Reaction Times Identify a Pavlovian Component in a Two-Choice Discrimination

Donald S. Blough

Brown University, Providence, RI 02912, USA

Abstract

Six pigeons discriminated on discrete trials between two colors. In Experiment 1, two luminous spots were both either blue or green and the reinforced responses were “peck left” for blue and “peck right” for green. In Experiment 2, the hue of a center spot controlled subsequent choice pecks to left or right. In both experiments response bias was manipulated in two ways. During stimulus frequency (“SF”) sessions correct responses brought food on 40% of trials; in “imbalanced” blocks of sessions one hue appeared on 80% of trials and the other on 20%. During reinforcement probability (“RNF”) sessions the hues appeared equally often, but in imbalanced blocks the hues signaled different reinforcement probabilities, either 64% or 16%. In “balanced” control blocks the hues appeared equally often and were both reinforced at 40%. The experiments gave similar results. When bias was computed from choice percentages the imbalanced conditions yielded substantial response bias, and the amount of bias was about the same under RNF and SF treatments. However, reaction times (RTs) gave a different outcome. RNF imbalance slowed responses directed at the less reinforced stimulus, but SF imbalance had little RT effect (Experiment 1) or no effect (Experiment 2). These results suggest that choice was controlled by an instrumental stimulus-response-reinforcement association, whereas RTs were controlled by a Pavlovian stimulus-reinforcement association.

Keywords

bias; choice; discrimination; reaction time; pigeon

1. Introduction

The experiments reported here expand some intriguing results from earlier research, and they also address theoretical issues related to discriminative processing. They were initially prompted by previous studies that recorded patterns of response frequency and RT during search and detection tasks (e.g. Blough, 1992, 1996, 1998, 2000, 2004, 2007). One study in particular prompted the present work. In that experiment (Blough, 2004) pigeons confronted a single luminous spot that varied in hue from trial to trial. Pecks to a particular red hue sometimes brought food, whereas slightly dissimilar reds went unreinforced; the differing reds constituted the “similarity” variable. Pecks to a green spot were often reinforced during some blocks of sessions, less often reinforced during other blocks: this was the “reinforcement”

²The interpretation of these data adopts the well-supported assumption that the key variable determining both kinds of association is reinforcer probability - that is, reinforcers per occurrence of a particular stimulus or response - rather than the absolute frequency of these pairs of events (e.g. Blough, 1992; Gibbon, 1981; Herrnstein, 1970).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

variable. The data showed that stimulus similarity strongly affected how often a hue was responded to, but it had only a small impact on RT. In contrast, changes in the base rate of reinforcement had a large impact on RT but a relatively small effect on response frequency. Together with related results these findings suggested that response frequency and RT reflect somewhat different aspects of discriminative processing.

The present experiments were designed to further explore the disparate response frequency and RT outcomes by moving from a go/no-go to a choice procedure and from variable stimulus discriminability to variable stimulus frequency. The two-choice design was conceptually related to some previous operant experiments that measured response frequencies in two-choice situations as a function of the same independent variables (e.g. McCarthy & Davison, 1979; Johnstone & Alsop, 1999). Those studies dealt, in part, with the analysis of factors affecting response bias. Like the present experiments, this previous work manipulated bias by varying: (1) the relative probability of reinforcement associated with stimulus-response alternatives and (2) the relative frequency with which alternative stimuli appear. In a human detection experiment the resulting biases are exemplified by an observer's increased tendency to respond "yes" ("stimulus detected") either because the signal often precedes a "payoff" event, or because the signal is relatively likely to appear (e.g. Elsmore, 1972; See et al, 1997; Macmillan & Creelman, 2005; Terman & Terman, 1972).

Such attempts to measure the effects of both stimulus frequency and reinforcement frequency encounter a procedural entanglement. Suppose, for example, that "blue" signals "peck left" and "green" signals "peck right". If reinforcement probability is held constant, reducing the number of blue stimuli will reduce the number of reinforcements for "peck left" just as if blue and green appeared equally often and reinforcement probability was less for blue than for green. This covariation has prompted the hypothesis that the effects of stimulus frequency on response bias are actually mediated by the accompanying changes in reinforcement frequency. Several experiments have supported this idea, though the results are not entirely consistent (e.g. McCarthy & Davison, 1979, Alsop & Davison, 1991; Johnstone & Alsop, 1999). If this single-factor hypothesis is correct, models of two-choice discrimination need not have a separate parameter for stimulus probability, and, in fact, behavioral detection models have continued to develop without such a parameter (e.g. Davison & Nevin, 1999).

A deeper theoretical implication surfaced after the present experiments were largely complete, and this implication prompted the title of this paper. This matter relates to Skinner's "three-term contingency", also known as the "stimulus-response-outcome" ("S-R-O") triad, which is the empirical foundation for most models of stimulus discrimination. The repeated occurrence of this triad of events is often assumed to result in an instrumental chain in which the stimulus sets the occasion for responses to be reinforced (Skinner, 1938). In addition, a number of accounts assume that the triad may yield a Pavlovian "S-O" association that has a separable role in the control of behavior. Variations on this idea appear, for example, in classical two-process theories (e.g. Rescorla & Solomon, 1967), accounts of behavioral contrast (e.g. Gamzu & Schwartz, 1973), analyses of autoshaping (e.g. Hearst & Jenkins, 1974), and the behavioral momentum theory of Nevin and his associates (e.g. Nevin & Grace, 2000). Often in such accounts a combination of instrumental and Pavlovian associations is reflected in a pattern of data derived from a single behavioral measure such as the rate of a key peck or lever press, or the percentage of choice between two such responses. Some studies also explore the details of associative structure with designs employing multiple stimuli, responses and outcomes together with measures such as transfer tests or outcome devaluation (e.g. Dickinson & deWit, 2003, Rescorla, 1994; cf. Hall, 2002). As will be explained in due course, the present experiments find evidence for separable instrumental and Pavlovian associations in a rather simple choice paradigm that provides two measures, choice percentages and reaction times (RTs).

The experiments used a trial-wise two-choice color discrimination that a human might describe as: “if blue, go left, if green, go right.” Each experiment was designed to bias this choice through two manipulations. (1) A stimulus frequency (“SF”) treatment examined the effect of changing the relative frequency with which either green or blue appeared as the stimulus; in this treatment, each hue signaled reinforcement on 40% of trials. (2) A reinforcement frequency (“RNF”) treatment examined the effect of changing the relative frequency of reinforcement assigned to each hue; in this treatment, each hue appeared on 50% of trials. Both of the treatments generated large and roughly equal effects on response bias as conventionally measured by response percentages, but they produced quite different effects on bias as measured by RT. The ensuing discussion relates this pattern of response percentages and RTs to the theoretical issues described above.

2. Experiment 1

2.1. Materials and methods

2.1.1 Subjects—Six White Carneaux pigeons served as subjects; they had previously served in a discrimination experiment similar to that used here. The birds were maintained at approximately 85% of their free-feeding weight, and they received most of their food in the form of 45 mg Noyes pigeon pellets during experimental sessions; if necessary, this ration was supplemented by mixed grain given after the experimental session.

2.1.2. Apparatus and Stimulus Displays—Three experimental boxes contained 34 cm × 30 cm × 34 cm subject chambers. An opening in the front panel of each chamber gave access to an LCD color video monitor; the visible part of the monitor screen measured 18 cm wide by 11 cm high. Infra-red emitters and detectors (Smart-Frame, Carol Touch, Round Rock, Texas) detected the position of pecks to the screen. A pellet dispenser was located immediately below and to the right of the screen. A loudspeaker near the chambers provided white masking noise. Each reinforcement consisted of a single 45 mg Noyes pigeon food pellet. Three IBM 486 computers controlled the experiments and recorded response data. Reaction times were recorded to the nearest 0.03 sec, a limitation imposed by the scan rate of the touch screens.

The task used three circular stimulus spots 5mm in diameter, all appearing on a dark grey background. One spot was an achromatic “ready” signal, which appeared in the center of the screen. The two stimulus spots, blue or green, appeared 3 cm to the left and right of the ready spot position. The saturation of the hues was adjusted as specified below. In what follows “a peck to a spot” means that the peck was directed toward a vertical band on the screen approximately 1 cm wide, centered on the target spot. For pecks to the ready spot, the response was defined by the end of the pecking motion, when the bird’s head had moved back about 1 cm from the screen and was no longer sensed by the infrared detectors in the touch frame. For pecks to the choice spot, the response was recorded when the birds head approached the screen. Figure 1 shows how these events fit into the time-line of a single experimental trial. The luminance of the screen remained at the background level during the inter-trial interval (ITI); it rose to a bright white for 2 sec to signal the delivery of food pellets.

2.1.3. Procedure—Each session included 800 two-choice trials. On each trial the achromatic “ready” spot appeared and stayed on until the bird had pecked at it. At this point the ready spot went off, the RT timer was started, and the two discriminative stimulus spots appeared to left and right; both spots were blue on some trials and green on others. The spots stayed on until the bird had pecked one of them, after which they disappeared and an RT was recorded. The RT was defined as the time between the appearance of the two stimulus spots and a peck to one of them, as defined above and illustrated in Figure 1. Extremely rare RTs greater than 10 sec were recorded as 10 sec. While a stimulus was displayed responses to any part of the screen outside the target area were ignored. The bird was reinforced on a percentage of trials for a

peck to the left if the two spots were blue and for a peck to the right if the spots were green. If a response was correct the screen briefly flashed white, and reinforcement followed if scheduled as described below. For reinforcement, the feeder light came on for 2 sec and a pellet dropped into the feeder. In either case the ITI then began; the ITI interval was randomly chosen from the interval between 4 and 6 sec.

An incorrect response produced the ITI, and a correction trial followed in which only the stimulus on the correct side was visible. This invariably led to a correct response to that spot, which produced a white screen flash, reinforcement if scheduled, and the next ITI. Data from these correction trials were not analyzed. To make the data for different birds as comparable as possible, the saturation of the hues was adjusted for individual birds by mixing more or less gray into the hues, aiming to induce a response accuracy for each bird that averaged about 90% during each session. For most of the birds the resulting hues were extremely desaturated, such that they could only be discriminated with great difficulty by a human observer. Sessions lasted approximately two hours and were run daily, except for infrequent breaks of from one to four days.

Four experimental phases were run. Each phase consisted of a series of blocks; during each block a particular set of stimulus and reinforcement conditions prevailed for 4 sessions (3200 trials). Blocks with “balanced” conditions alternated with “imbalanced” blocks. Balanced blocks were the same throughout the experiment; each hue appeared on half of the trials and correct pecks to each brought food on 40% of trials. Imbalanced blocks of trials differed for the two treatments. During SF sessions (Phases 1 and 4) one hue appeared four times as often as the other, but reinforcement was assigned to each on 40% of their appearances. The sequence of “green frequent” and “blue frequent” blocks was counterbalanced across birds. During RNF sessions (Phases 2 and 3) the hues appeared equally often, but the relative frequency of reinforcement assigned to each differed, such that one hue signaled reinforcement on 16% of trials and the other one 64%. Hues were again counterbalanced across birds. The first, third, and fifth blocks of each phase were balanced, and the second and fourth blocks were imbalanced, first favoring one hue and location and then favoring the other.

Figure 2 displays a sample set of trials from imbalanced blocks favoring the blue hue (left response) for each treatment. Each pair of spots represents the stimulus display on a single trial. Trials from the RNF treatment appear on the left, those from the SF treatment on the right. White stars represent reinforced trials. Note that reinforcements for RNF and SF are assigned to the same left and right locations on comparable trials. Although this identity would not be true in general, it underscores the fact that the treatments do not differ in the number of reinforcements they allocate to the two locations. The samples also show that the treatments differ their relative number of blue and green trials and in the assignment of reinforcements to each hue. In the RNF trials the stimuli appear equally often but blue has four reinforcements and green has only one. In the SF trials there are eight blues and two greens, but both hues are reinforced at 50%. (The reinforcement rates are not quite the same as those programmed for entire sessions, but the diagram accurately characterizes the treatments. See Appendix section A.2. for further details.)

As described earlier, Phase 4 replicated Phase 1, and Phase 2 replicated Phase 3. As the phases were run in numerical order, this sequence provided some control for possible effects of the order in which the conditions were run. The replications differed only in their treatment of correction trials. In Phases 1 and 2 the response on each correction trial was followed by reinforcement. In the replication Phases 3 and 4, responses on correction trials were followed by reinforcement only if that reinforcement had been scheduled for the previous trial, which elicited an error. The reasons for these procedures are discussed in Appendix A.3. As it turned

out, the data showed little if any differential impact of these correction procedures, perhaps because error rates were low.

2.2. Results and Discussion

Data analysis included results from the last 2 sessions of each 4-session block. Mean percent responses and mean RTs across all balanced-block sessions were computed for each bird; they were also computed separately for each imbalanced block¹. This process yielded for each bird three sets of percent response values and three sets of mean RTs: “balanced”, “imbalanced favoring left (blue)”, “imbalanced favoring right (green)”.

2.2.1. Percent choice—The mean percent correct responses averaged 92% across all the balanced conditions and birds, close to the intended 90%. This ranged from 89% to 96% for individual birds across all phases. Bias and sensitivity measures were computed separately for each bird and condition within each phase. These measures were defined as: bias = $.5 * (z(R1|S1) - z(R2|S2))$ and sensitivity (d') = $z(R1|S1) - z(R1|S2)$, where z represents the normal standard score, R1 and R2 are responses to the left and right spots, and S1 = blue stimulus, S2 = green stimulus (Macmillan & Creelman, 2005).

Mean bias was computed by averaging the absolute bias values from the two imbalanced conditions for each experimental phase within and then across birds. Figure 3 summarizes the result. The columns to the left show data from Phase 2 (RNF) and Phase 1 (SF), the columns to the right show the results of the replications, Phases 3 (RNF) and 4 (SF). In this and all other figures the error bars show the standard error of the mean. It is evident that for both treatments the unbalanced conditions induced substantial bias; this bias was about the same in all four phases and t-tests showed no significant difference between these for either pair. The difference in the correction procedure between the first and last two phases also had no significant effect. The fact that the two treatments yielded comparable biases is in agreement with most previous research, and it suggests that the controlling variable is the distribution of reinforcements to left and right responses regardless of the stimulus.

Sensitivity data were combined and averaged in the same way as the bias data just presented. Figure 4 compares the mean d' scores of the two imbalanced conditions with the mean d' score for the balanced conditions. Error bars show standard error of the mean. The key finding is evident on inspection: imbalanced stimulus frequency (SF) reduced d' somewhat below the balanced control (Phases 1 and 4, left and right pairs of bars), whereas imbalanced reinforcement (RNF) had little effect on d' (middle two pairs of bars). This impression was confirmed statistically: non-significant for RNF treatments ($df=5$, $t=0.86$ and $t=0.42$, $p=.43$ and $p=.69$), significant for SF treatments ($df=5$, $t=4.57$ and $t=4.24$, $p<.01$ and $p<.01$).

This pattern of sensitivity results is difficult to compare with earlier data because of procedural differences. Changing presentation probabilities, as in the SF procedure, has previously been found to influence d' in studies with both humans (e.g. Markowitz & Swets, 1967) and animals (e.g. Johnstone and Alsop, 1999). When the slope of iso-sensitivity (“ROC”) curves is not unity, as it usually is not (e.g. Blough, 1967; Swets, 1986), sensitivities differ when computed at different bias points, and this may relate to the inconstancy of d' here. The lack of effect of the RNF procedure on d' receives consideration in the General Discussion below.

¹Analysis of median RTs gave results very similar to that for means, and only mean RTs will be considered here. Upon analysis, the data from the first balanced block of Phase 1 of Experiment 1 were found to be highly variable, partly due to apparatus problems, and those data were omitted from the average of balanced blocks. An alternative normalization of bias measures was done by taking the ratio of the imbalanced RTs, rather than the difference between them; the two measures yielded essentially the same picture. RT distributions were also determined, but their extreme variability precluded analysis.

Of secondary interest is an overall increase in d' , expressed in Figure 4 as the increasing height of the columns from left to right. The phases were run in the order shown, so it is likely that this increase in d' reflects a gradual improvement in discriminative accuracy with practice, a common occurrence in a long experiment of this sort. Finally, as with bias, there is no sign that the difference in correction procedure between the early and late phases had any effect here.

2.2.2. Reaction time—Biases based on RTs were computed separately for each bird. RTs were first normalized within each phase by subtracting the bird's mean RT across the balanced conditions from the mean RT from each of the imbalanced conditions. Figure 5 displays the result, derived from correct responses made during each of the four experimental phases. The left bar in each pair shows the extent to which mean RT lengthened for the less favored stimulus as compared with the balanced condition, whereas the right bar shows the extent to which mean RT changed for the more favored stimulus, becoming shorter than that from the balanced conditions in all but Phase 4.

The data in Figure 5 show a relatively large effect of the reinforcement differential on RTs in Phases 2 and 3, and a considerably smaller effect of the SF changes in Phases 1 and 4. The error bars are somewhat misleading, as they include a substantial amount of variance caused by the fact that RTs for the two imbalanced conditions tended to move together, but moved differently for different birds. When the low and high conditions were compared directly with each other, rather than in comparison with the balanced condition, all 4 comparisons were found to be significant or nearly so. For SF (Phases 1 & 4) $df=5$, $t=5.8$, $p<.01$ and $t=2.3$, $p<.07$ ($p<.04$ one-tailed); For RNF (Phases 2 & 3), $df=5$, $t=3.5$ and 3.6 , $p<.02$

Mean RTs for errors were relatively variable, probably because there were many fewer errors than correct responses. In Phases 1 and 2, and unlike the result for correct responses, the imbalanced condition yielded no significant difference in error RT, whereas data from Phases 3 and 4 were very similar to those for correct response RTs shown in the right side of Figure 5. This difference probably stems from fact that in Phases 1 and 2, correction responses after errors were always reinforced, whereas in Phases 3 and 4 correction responses were only reinforced if reinforcement had been scheduled for the trial on which an error was made. In bias conditions, where more errors tend to be made, extra reinforcement of correction responses would tend to oppose the bias, and this may be the reason for the obtained difference.

3. Experiment 2

In Experiment 1 the biases calculated from response percentages could be attributed to the relative probability of reinforcement for the alternative *responses*, left or right, because bias appeared when this probability was imbalanced, and the equal imbalances in the two treatments produced equal amounts of bias. In contrast, the biases calculated from RTs could be attributed to the relative probability of reinforcement for the alternative *stimuli*, blue or green. The RNF treatment imbalanced this probability and produced RT bias; the SF treatment did not imbalance this probability and produced little bias. These outcomes suggest that bias in the left-right choice was governed by an instrumental response-reinforcement association whereas RT bias was governed by a Pavlovian stimulus-reinforcement association.

In Experiment 1 a single response on each trial expressed both the putative instrumental and Pavlovian components because each peck to left or right was also directed at a spot carrying the stimulus hue. The procedure of Experiment 2 was designed to further separate the components by requiring first an "observing" peck directed at a centrally-located blue or green spot and then a second "choice" peck between two identical red circles, one on the left and one on the right.

3.1. Materials and methods

3.1.1. Subjects—The six pigeons that served in Experiment 1 also served in Experiment 2, and they were maintained as before.

3.1.2. Apparatus and Stimulus Displays—The apparatus was unchanged from Experiment 1. As before, the task used three circular stimuli 5mm in diameter. However, there was no ready signal. In its place was the blue or green discriminative stimulus, centered on the display screen. A peck to this stimulus produced two identical red outline circles that differed strikingly in appearance from the discriminative stimuli. These circles marked the same left and right locations as the choice stimuli in Experiment 1. The background and other aspects of the display were as in Experiment 1.

3.1.3. Procedure—Each session included 800 two-choice trials. Each trial began with the appearance of the stimulus spot, either green or blue, in the center of the screen. A peck to this spot turned it off and produced the red choice circles to either side. The bird was reinforced for a peck to the left circle if the center spot had been blue and for a peck to the right circle if the center spot had been green. The circles stayed on until the bird had pecked one of them, after which they disappeared. Correction trials were run as in Phases 3 and 4 of Experiment 1; that is, responses on correction trials were reinforced if reinforcement had been scheduled for the preceding trial. Two RTs were recorded. The first or “observing” RT was the time between the onset of the center spot and a peck to that spot, as defined below. The second or “choice” RT was the time between the peck to the center spot and a peck to one of the choice circles. In other respects experimental trials were run as in Experiment 1. A comparison of Figures 1 and 6 shows the essential differences between the procedures of Experiment 1 and Experiment 2. Further explanation of Figure 6 is given in the procedure section below.

The experimental treatments were similar to those of Experiment 1. During balanced blocks each hue appeared on half of the trials and correct pecks to each brought food on 40% of trials. These balanced blocks alternated with imbalanced blocks, which differed for the two treatments. During the SF treatments the stimulus frequency with which either green or blue appeared differed within a session, but reinforcement was assigned to both on a constant 40% of their appearances. During RNF treatments the relative frequency of reinforcement assigned to each hue differed within imbalanced blocks, but each hue appeared on a constant 50% of trials.

3.1.3.1. Preliminary sessions: An initial series of 13 to 17 sessions was run, beginning with 4 to 6 balanced sessions during which the birds learned the new procedure and saturation of the discriminative hues was adjusted. The task was quickly acquired, but the birds produced a great many errors when tested with the hue saturation levels used in Experiment 1. The 10% error performance approximated in Experiment 1 was achieved only by substantially increasing hue saturation such that, to a human observer, the stimuli were decidedly “blue” or “green”, unlike the barely discriminable hues of Experiment 1. (One bird required so many training and adjustment sessions that it did not participate in the RNF treatment to be described next.)

The preliminary sessions continued with the RNF treatment, first a shift to unbalanced reinforcement and then a return to balanced reinforcement. These sessions had been intended to begin the experiment proper, but the nature of their data prevented this and prompted an adjustment in the definition of the response to the discriminative stimulus. During these preliminary sessions the touch-frame signaled a response to the center spot when the bird’s head moved away from the screen at the end of the pecking motion. Prior to this head withdrawal the beak was near the spot for a brief interval, or “dwell time,” and this dwell time

was included in the recorded RT to the center spot and also in the stimulus “on” time. A time-line for this preliminary procedure appears at the top of Figure 6 (see also Appendix A.1.).

Unexpectedly, the data suggested that dwell time varied systematically between balanced and unbalanced conditions, providing useful information that is discussed below. However, this dwell time varied in such a way as to offset changes in the RT to stimulus onset, and in order to remove this complication from the analysis of the RTs the touch-screen was re-programmed to signal an observing response on head entry to the sensor field, rather than head exit. Thus, during the main experimental sessions the hue spot disappeared as the beak approached it, and dwell time was included in the second, or “choice” RT. A time-line showing this response definition appears at the bottom of Figure 6. The rest of the main procedure is described next.

3.1.3.2. Main experimental sessions: The experiment included four phases that each contained a series of 4-session blocks. During each block a particular set of stimulus and reinforcement conditions prevailed. As in Experiment 1, “balanced” blocks of sessions alternated with “imbalanced” blocks throughout all phases. In balanced blocks each hue appeared on half of the trials and correct pecks brought food on 40% of trials. Imbalanced blocks differed between phases. During Phase 1 the RNF treatment was in effect. The stimuli appeared equally often, but correct choices following green brought food on 64% of trials and following blue on 16% for three birds, with the hue assignments reversed for the other three birds. During Phase 2 the SF treatment was in effect: each hue signaled reinforcement on 40% of trials, but green appeared on 80% of trials and blue on 20% for three birds, with the hues reversed for the other three birds. Phase 3 continued the SF treatment with hue assignments reversed, and Phase 4 returned to the RNF treatment with hue assignments reversed from those in Phase 1.

3.2. Results and Discussion

Data from the last two sessions of each treatment block were analyzed. Each of the two treatments (RNF and SF) yielded three sets of mean response percentages and three sets of mean RTs, namely “balanced”, “imbalanced favoring blue (left)” and “imbalanced favoring green (right)”. The RTs were further subdivided into observing RTs for responses to the green or blue center spot and choice RTs for responses to one of the red circles that marked the left and right choice locations.

3.2.1. Percent choice—The percentage of correct choice responses averaged 91% across all balanced conditions and birds, close to the intended 90%; for individual birds this ranged from 88% to 95%. Bias and sensitivity measures were computed from the percent data for each bird as in Experiment 1, and then absolute bias from the two imbalanced conditions was averaged for each experimental treatment within and then across birds. Figure 7 displays the results. As in Experiment 1, the unbalanced conditions induced substantial biases, but these did not differ significantly between the two treatments ($df=5$, $t=1.12$). As discussed above, this outcome is expected if the source of bias is simply the relative reinforcement associated with the left and right responses, which was the same for the two treatments.

Sensitivity scores (d') were computed from the same percentage data, but there was no hint of a difference between the two treatments or between the balanced and unbalanced conditions: for RNF: balanced $d' = 2.73$, unbalanced $d' = 2.77$, and for SF: balanced $d' = 2.84$, unbalanced $d' = 2.80$.

3.2.2. Observing response RTs—RT biases were computed from observing-response RTs on correct-choice trials, yielding for each experimental phase a set of differences between each bird’s balanced and unbalanced RTs. These were averaged for each treatment (RNF or

SF) across birds and phases for the high probability (frequent reinforcement or stimulus) and low probability (infrequent reinforcement or stimulus) conditions. The RNF treatment produced a substantial lengthening of RT for the less reinforced stimulus ($df=5$, $t=5.76$, $p<.01$) together with a small decrease in RT for the more reinforced stimulus ($df=5$, $t=2.52$, $p<.05$). There was no significant effect of the SF treatment on observing RTs. The left pair of columns in Figure 8 summarizes these results in terms of “net bias,” which is the difference between the biases for the less and more reinforced stimuli.

These results are quite similar to those from Experiment 1 (Figure 5), even though here the responses involved are simply “observing” pecks to the discriminative stimulus, not choice pecks. The findings support the idea that the RT in both experiments reflects a stimulus-reinforcer association that is separable from the control of choice. This matter receives further discussion below.

3.2.3. Choice response RTs—As stated earlier, a choice RT was defined as the interval between the observing peck and the peck to one of the two red circles defining choice locations. The analysis of these RTs paralleled that described above for the observing pecks. No significant or suggestive biases appeared in these RTs under any condition; this result is summarized in the right pair of columns in Figure 8. This lack of any choice RT bias in either treatment adds to the evidence that RT biases were a function of the stimuli to which pecks were directed.

3.2.4. Observing RTs on error trials—Figure 9 shows the results of the same analysis for RTs of responses that were followed by choice errors. The results for the low probability condition were similar to those for correct responses, which were displayed as net bias in the leftmost bars of Figure 8. That is, as compared to the balanced condition there was a significant lengthening of RT for the RNF treatment but not for the SF treatment (Figure 9, left bars in each pair). However, for both treatments the observing response RTs were also significantly increased when made to the *high* probability stimulus. This suggests that an important reason for error (although probably not the only reason) was confusion between the stimuli. That is, if the bird “mistook” the high-probability hue for the low-probability hue, it would respond slowly and also make choice error, as the data suggest. (For other possible reasons for error, see Blough, 1996.) However, it is unclear why this bias appears under the SF treatment here, and not also in error data from Experiment 1.

3.2.5. Dwell times—“Dwell time” was defined above as the time during a peck to the center spot that the bird’s head was near the display screen. Recall that during the preliminary procedure the stimulus stayed on until the head withdrew from its vicinity and the observing RT included that interval (top panel, Figure 6), whereas during the main experimental procedure the stimulus stayed on only until the head approached it, and the observing RT was measured to that approach (bottom panel, Figure 6). Samples of the RNF treatment were available from both the preliminary and main experimental sessions. Thus, as shown to the right in Figure 6, the difference between the two observing RTs provides an estimate of RNF dwell time.

Five birds provided a set of mean observing RTs from the balanced-imbalanced-balanced series of blocks in the preliminary RNF sessions. From these were subtracted mean observing RTs from the same birds and reinforcement conditions obtained during the main body of the experiment. These provided estimates of dwell times for the balanced and unbalanced conditions. The difference between balanced and unbalanced RTs then gave an estimate, for each bird, of the effect of the reinforcement variation on dwell time. The mean across birds of these differences for trials resulting in correct responses appears in Figure 10. The two means

were significantly different from each other ($df = 4$, $t=5.11$, $p<.01$); each was also significantly different from zero ($df = 4$, $t=3.7$ and 4.0 , $p<.02$).

Figure 10 shows that, relative to balanced conditions, dwell times were long to the more reinforced stimulus, and short to the less reinforced stimulus. Because the RNF treatment provided imbalanced reinforcement to the choice locations as well as to the stimuli (e.g. Figure 2), a response-reinforcement association might predict the opposite result, because increased dwell time induces a slight delay in the receipt of reinforcement. That is, seeing the more reinforced stimulus, the bird might shift with special rapidity to the choice response. Instead, the result suggests that dwell time as well as RT reflects a Pavlovian association between stimulus and reinforcement. This accords with previous observations that pigeons peck more rapidly at sample stimuli that signal a relatively favorable outcome (e.g. Santi & Roberts, 1985). Indeed, “dwell time” here may well include rapid “nibbling” that kept the bird’s head near the stimulus spot (see Appendix A1.)

4. General Discussion

It may be helpful to review some essential items of procedure. In Experiment 1 each trial allowed a single choice peck to left or right stimulus spots; the hue of both spots was the same, with blue signaling left and green signaling right. In Experiment 2 the stimulus hue appeared on a central spot; an observing peck to this spot preceded a choice peck to left or right. In both experiments (1) the stimulus frequency (SF) and reinforcement (RNF) treatments were identical when reinforcements and stimuli were balanced; (2) when imbalanced, both treatments reinforced one choice response (e.g. “peck left”) on 64% of trials and the other choice response (e.g. “peck right”) on 16% of trials; (3) when imbalanced, the treatments differed with respect to the relative frequency with which each hue appeared, and, consequently, also with respect to the probability of reinforcement associated with each hue.

In important respects the two experiments gave similar results. When bias was computed from choice response percentages the imbalanced conditions in both experiments yielded substantial bias, and the amount of bias was about the same under RNF and SF treatments (Figures 3 and 7). However, when bias was computed from RTs of responses directed at the discriminative stimulus, RNF imbalance slowed response to the less reinforced stimulus substantially, but SF imbalance had little or no effect (Figures 5 and 8). Data from the RNF treatments in Experiment 2 showed that, compared with balanced conditions, dwell time, when the bird’s head was near the hue stimulus, was longer to a more reinforced stimulus and shorter to a less reinforced stimulus (Figure 10).

These findings suggest that bias calculated from response percentages is primarily determined by the relative density of reinforcement following the alternative *responses* to left or right; this may be thought to reflect an instrumental response-reinforcement association. In contrast, bias calculated from RTs appears to be primarily determined by the relative density of reinforcement paired with the alternative *stimuli*, blue or green; this may be thought to reflect a Pavlovian stimulus-reinforcement association. The estimates of observing response dwell times suggest that the Pavlovian association also governed this measure. The main reason for errors, in this view, is confusion of the stimuli due to the difficulty of the task, which many theorists might call a generalization decrement.

If this analysis is correct, it remains to explain the relatively small but significant RT bias produced by the SF treatment in Experiment 1 (Figure 5), which did not occur in Experiment 2 (Figure 8). This might be accounted for if, in Experiment 1, the effective stimulus was a compound comprised of the location as well as the hue of the pecked spot. If the governing Pavlovian association was between such a “hue-location” compound and reinforcement, the

differential reinforcement of the location part of that compound may have sufficed to yield the partial RT effect under the SF treatment. In Experiment 2 the discriminative stimulus was in one location, so only a hue-reinforcement association would affect RTs of observing responses made to that location; hence there was no basis for an observing response RT bias under the SF treatment, and none was found.

As remarked in the introduction, the idea that Pavlovian associations play a role in discriminated instrumental behavior is far from new, but the present results seem to provide a particularly clear example of an S-O influence because it is revealed in one response measure but not in the other. Just how this influence works remains in question. An idea prominent in the extensive literature on the subject is that the Pavlovian component plays a motivational or energizing role (c.f. Hall, 2002), which is sometimes viewed as a modulation of the incentive value of the discriminative stimulus. This notion helped to account for a previous data set of mine (Blough 2007), and the regulation of RT and duration (dwell time) in the present experiments both fit well with this motivational interpretation. The literature on differential reinforcement suggests two possible alternatives to the motivational interpretation. One alternative is that the Pavlovian component modulates the strength of attention directed at the discriminative stimulus (e.g., Nevin, Davison & Shahan, 2005). The second alternative is that the Pavlovian component provides an expectancy cue that acts like other cues to control instrumental behavior (Trapold, 1970; see review by Urcuioli, 2005). However, neither of these accounts has obvious implications for RT and both seem to imply that differential reinforcement should enhance choice performance, which it did not in either of the present experiments.

We can now return to the relationship between the present results and those operant experiments mentioned in the introduction that explored the effects of reinforcement and stimulus-frequency variations on choice (e.g., McCarthy and Davison, 1979, Alsop & Davison, 1991). As I mentioned, several such experiments have found that manipulations comparable to the RNF and SF treatments produced approximately equal bias effects on response percentages, just as in the present data. The implication is that the operant results and the behavioral detection models associated with them are based on the instrumental component of the discrimination and thus they may be correct as far as they go, but they do not account for RT effects that might have been discovered had they been measured.

Behavioral detection models are generically related to standard signal-detection theory (SDT) (e.g. Alsop, 1991), and SDT also predicts that both the RNF and SF treatments should induce response bias because the imbalanced reinforcement in both would tend to shift a response criterion along a decision axis (e.g. Macmillan & Creelman, 2005). SDT also has little to say about RT. The major empirical finding associated with human detection is that RTs are relatively high for discriminations that are difficult (e.g. Espioza-Varas & Watson, 1994), while an inverted U-shaped function describes RTs over the full range of difficulties. This pattern is predicted by other models (e.g. Ratcliff & Rouder, 1998) and pigeons also have produced such data, (e.g. Blough, 2007), but these observations are beyond the scope of the present discussion.

The identification of a Pavlovian component in the present experiments contributes significantly to the interpretation the findings of an earlier study that I described in the introduction (Blough, 2004). In that experiment only a single hue spot appeared on each trial; it was found that (1) the reinforcement density previously associated with the visible hue strongly affected the RT of pecks directed at it, and also (2) that the similarity of the spot to a previously reinforced hue strongly affected the frequency with which the spot was pecked. One can now hypothesize that the first finding reflected a Pavlovian stimulus- reinforcement association whereas the second reflected an instrumental association that generalized to an extent dependent on stimulus similarity. This new interpretation suggests a revised understanding of

parameters in a random walk model of discrimination that I am currently constructing, a model related to the rudimentary account presented in the earlier paper (Blough, 2004) and also to the extensively tested model of Ratcliff (e.g. Ratcliff & Rounder, 1998). More generally, these results underscore the explanatory usefulness of multiple measures of response, and they suggest continued mindfulness of the S-O correlation contained in the classic S-R-O contingency.

References

- Alsop, B. Behavioral models of signal detection and detection models of choice. In: Commons, ML.; Nevin, JA.; Davison, Mc, editors. *Quantitative Analyses of Behavior*. Vol. 11. Hillsdale, NJ: Erlbaum; 1991. p. 39-55. Signal detection: Mechanisms, models and applications
- Alsop B, Davison M. Effects of varying stimulus disparity and the reinforcer ratio in concurrent schedule and signal-detection procedures. *J Exp Anal Behav* 1991;56:67–80. [PubMed: 1940764]
- Blough DS. Stimulus generalization as a signal detection in pigeons. *Science* 1967;158:940–941. [PubMed: 6054168]
- Blough DS. Reaction times of pigeons on a wavelength discrimination task. *J Exp Anal Behav* 1978;30:133–137.
- Blough DS. Effects of stimulus frequency and reinforcement variables on reaction time. *J Exp Anal Behav* 1992;57:47–50. [PubMed: 16812648]
- Blough DS. Error factors in pigeon discrimination and delayed matching. *J Exp Psychol: Anim Behav Process* 1996;22:118–131.
- Blough DS. Context reinforcement degrades discriminative control: A memory approach. *J Exp Psychol: Anim Behav Process* 1998;24:1–15.
- Blough DS. Effects of priming, discriminability and reinforcement on reaction-time components in pigeon visual search. *J Exp Psychol: Anim Behav Process* 2000;26:50–63. [PubMed: 10650543]
- Blough DS. Reaction time signatures of discriminative processes: Differential effects of stimulus similarity and incentive. *Learn Behav* 2004;32:157–172. [PubMed: 15281388]
- Blough DS. Reaction times in discriminations of varying difficulty: Decision modulated by arousal. *Behav Process* 2007;74:234–243.
- Davison M, Nevin JA. Stimuli, reinforcers, and behavior: an integration. *J Exp Anal Behav* 1999;71:439–482. [PubMed: 16812904]
- Dickenson A, deWit S. The interaction between discriminative stimuli and outcomes during instrumental learning. *The Quart J Exp Psychol* 2003;56B:127–139.
- Elmsore TF. Duration discrimination: Effects of probability of stimulus presentation. *J Exp Anal Behav* 1972;18:465–469. [PubMed: 16811638]
- Espinoza-Varas B, Watson CS. Effects of decision criterion on response latencies of binary decisions. *Percept Psychophys* 1994;55:190–203. [PubMed: 8036100]
- Gamzu E, Schwartz B. The maintenance of key pecking by stimulus-contingent and response-independent food presentation. *J Exp Anal Behav* 1973;19:65–72. [PubMed: 16811654]
- Gibbon, J. The contingency problem in autoshaping. In: Locurto, CM.; Terrace, HS.; Gibbon, J., editors. *Autoshaping and conditioning theory*. New York: Academic Press; 1981. p. 285-308.
- Hall, G. Associative structures in Pavlovian and instrumental conditioning. In: Pashler, H., editor. *Stevens' Handbook of Experimental Psychology*. Vol. 3. New York: Wiley; 2002. p. 1-45.
- Hearst, E.; Jenkins, HM. Sign-tracking: The stimulus-reinforcer relation and directed action. *The Psychonomic Society*; Austin, TX: 1974.
- Herrnstein RJ. On the law of effect. *J Exp Anal Behav* 1970;13:243–266. [PubMed: 16811440]
- Johnstone V, Alsop B. Stimulus presentation ratios and the outcomes for correct responses in signal-detection procedures. *J Exp Anal Behav* 1999;72:1–20. [PubMed: 16812906]
- Markowitz J, Swets JA. Factors affecting the slope of empirical ROC curves: Comparison of binary and rating responses. *Percept Psychophys* 1967;2:91–100.
- McCarthy D, Davison M. Signal probability, reinforcement, and signal detection. *J Exp Anal Behav* 1979;32:373–386. [PubMed: 512570]

- Macmillan, NA.; Creelman, CD. Detection theory: A user's guide. New York: Cambridge Univ; 2005.
- Nevin JA, Davison M, Shahand TA. A theory of attending and reinforcement in conditional discriminations. *J Exp Anal Behav* 2005;84:281–303. [PubMed: 16262190]
- Nevin JA, Grace RC. Behavioral momentum and the Law of Effect. *Behav Brain Sci* 2000;23:73–130. [PubMed: 11303339]
- Ratcliff R, Rouder JN. Modeling response times for two-choice decisions. *Psych Sci* 1998;9:347–356.
- Rescorla RA. Control of instrumental performance by Pavlovian and instrumental stimuli. *J Exp Psychol: Anim Behav Process* 1994;20:44–50. [PubMed: 8308492]
- Rescorla RA, Solomon RL. Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psych Rev* 1967;74:151–172.
- Santi A, Roberts WA. Reinforcement expectancy and trial spacing effects in delayed matching-to-sample by pigeons. *Anim Learn & Behav* 1985;13:274–294.
- See JE, Warm JS, Dember WN, Howe SR. Vigilance and signal detection theory: an empirical evaluation of five measures of response bias. *Human Factors* 1997;39:14–29.
- Skinner, BF. The behavior of organisms. New York: Appleton-Century; 1938.
- Swets JA. Indices of discrimination or diagnostic accuracy: Their ROCs and implied models. *Psych Bull* 1986;99:100–117.
- Terman M, Terman JS. Concurrent variation of response bias and sensitivity in an operant-psychophysical test. *Percept Psychophys* 1972;1972:423–432.
- Trapold MA. Are expectancies based upon different positive reinforcing event discriminably different? *Learn & Motiv* 1970;1:129–140.
- Urcuoli PJ. Behavioral and associative effects of differential outcomes in discrimination learning. *Learn & Behav* 2005;33:1–21.

Appendix A: A.1. Response definition and dwell time

The infra-red touch frame sensors could be set to report either of two events: (1) *entry*, the approach of the beak to the screen (beam broken) or (2) *exit*, the beak's departure (beam unbroken). The original reason for using head exit to define the peck to the ready spot in Experiment 1 was that some birds occasionally “swiped” their heads from the ready location to the choice location in such a way as to truncate the measured RT. However, results of the preliminary procedure in Experiment 2 necessitated a shift to head entry, for the reasons given in the procedure section of that experiment. As shown in Figure 6, the serendipitous result was that RT differences between the “head entry” and “head exit” methods provided a measure of dwell time. It should be understood that dwell time is not quite the same as “peck duration,” because it might involve multiple partial pecks or other activity near the stimulus spot. Also, it is important to note that the stimulus stayed on during dwell, and this probably affected the pattern of recorded dwell times.

A.2. Reinforcements allocated to the left and right responses

It is explained in the text that imbalanced stimulus presentations were in the ratio of 80% to 20%, with reinforcement for both stimuli at 40%, whereas the imbalanced reinforcement percentages were 16% and 64%. It was stated that that both of these imbalanced conditions yielded the same ratio of reinforcements delivered for responses to each key. One way to appreciate this is to compute the number of reinforcements scheduled for each response during a single session. For the imbalanced reinforcement condition (RNF, Phases 2 & 3) green and blue each appeared 400 times in an 800 trial session. One was reinforced on 16% of its appearances, yielding $400 \times 0.16 = 64$ reinforcements associated with its location (left for blue, right for green); the other was reinforced on 64% of its appearances, yielding $400 \times 0.64 = 256$ reinforcements associated with its location. For the imbalanced stimulus condition (SF, Phases 1 & 4), one stimulus appeared on 20% of trials, or $800 \times 0.20 = 160$, and this was reinforced 40% of the time, for $0.4 \times 160 = 64$ reinforcements for its location. The other stimulus appeared

on 80% of trials, or $800 \times 0.80 = 640$, again reinforced 40% of the time, for $0.4 \times 640 = 256$ for its location. Thus, the imbalanced conditions for each treatment yielded the same distribution of reinforcements to the two locations. This would be exactly true if responses were always correct. Errors caused it to be approximately true; this approximation is improved to the extent that one counts reinforcements on correction trials, remembering, however, that the display on correction trials (a single stimulus spot) might cause one to question the relevance of responses and reinforcements on those trials.

A.3. Reinforcement of responses on correction trials

If a discrete-trials choice procedure has no correction trials birds are apt to drift into a key preference which becomes exaggerated as reinforcements to the non-favored choice are missed (e.g. Elsmore, 1972); correction trials effectively suppress such key preferences. An alternative procedure arranges that if an error is made to a given stimulus the missed reinforcement will be delivered on a later presentation. If reinforcements are not scheduled too often, this can keep the obtained reinforcement per stimulus near its programmed level, but with the reinforcement rate used here (40%) the occurrence of several errors can cause reinforcements to pile up waiting to be delivered, resulting in a distorted pattern of reinforcement.

Figure 1.

Time-line for a single trial of Experiment 1. "Head entry" occurred when pigeon's head first broke the infra-red sensor beams; "head exit" occurred when this break ended. Here RT timing begins at exit for the "ready" peck and entry for the choice peck. Pecks and RTs were defined differently in Experiment 2 (see text and Figure 6.)

Figure 2.

An illustrative sample of 10 trials from an imbalanced session with each bias treatment in Experiment 1. Trials from the RNF treatment appear to the left, those from the SF treatment to the right. White stars indicate trials terminating in reinforcement. In this example imbalanced reinforcement favors the left response, which is correct for the blue stimulus. Each treatment provides 4 reinforcements for left responses and 1 for right responses; the treatments differ in the relative number of stimuli (equal numbers of blue and green for RNF, unequal for SF) and, consequently, in the probability of reinforcement being assigned to a stimulus (unequal for RNF, equal for SF).

Figure 3.

Response biases computed from choice response percentages in Experiment 1. Note that the abscissa labels reverse the position of Phases 1 and 2 because the first RNF treatment occurred during Phase 2, but for clarity the results of both RNF treatments are shown here in the dark columns at left. All biases were substantial, but they did not differ significantly from each other.

Figure 4.

Mean sensitivity (d') scores for the balanced and imbalanced conditions in each treatment phase in Experiment 1. Sensitivity was slightly but significantly reduced in the imbalanced condition of the SF treatment, where there were four times as many trials with one hue as there were with the other.

Figure 5.

Response biases computed from RTs in Experiment 1. Bias, shown on the ordinate, is determined from the difference between RTs on unbalanced trials and RTs on balanced trials. “Low reinf” columns come from trials on which the correct choice was reinforced with probability 0.16; “high reinf” columns come from trials on which the correct choice was reinforced with probability 0.64. The RNF treatment yielded large RT biases, but the SF treatment did not.

Figure 6.

Time-lines for single trials during two procedures in Experiment 2. The preliminary procedure differed from the main procedure in the time at which the observing peck was detected; detection occurred at the end of the peck for the preliminary procedure and at the beginning of the peck for the main procedure. This difference affected stimulus duration and also enabled an estimation of dwell time, as shown at right. See text.

Figure 7. Mean response biases computed from choice response percentages in Experiment 2; data from phases with identical treatment conditions (RNF or SF) have been averaged together. As in Experiment 1, biases induced by both treatments were substantial but did not differ significantly.

Figure 8.

RT response biases for Experiment 2, computed separately for observing responses (left columns) and choice responses (right columns). Net bias, shown on the ordinate, is the difference between the bias from trials with the less reinforced stimulus and that from trials with the more reinforced stimulus. The RNF treatment yielded a large observing-response RT bias, but the SF treatment did not. In contrast, neither the RNF nor the SF treatment affected RTs of the choice responses.

Figure 9.

RT biases for Experiment 2 as in Figure 8, except that these come from observing responses that were followed by choice *errors*. Such errors were made on about 10% of trials on average. The significant slowing of responses to high probability stimuli (white bars) suggests that the subsequent errors resulted, at least in part, from confusion between the stimuli (see text).

Figure 10.

Mean dwell times during which the beak was near screen on observing pecks (RNF treatments, preliminary procedure, Experiment 2). Dwell time was computed from head-entry data from the preliminary portion of Experiment 2 and head-exit data from the main experiment (see Figure 6 and Appendix A.1.). Even though a long dwell delayed reinforcement slightly, dwell times were longer to more reinforced stimuli than they were to less reinforced stimuli.