

Uncertainty Reduction versus Peck-to-Reinforcer Ratio
as Determinants of Suboptimal Choice in Pigeons

Ryan R. McDougle

California State University, Chico

UNCERTAINTY REDUCTION VERSUS PECK-TO-REINFORCER RATIO
AS DETERMINANTS OF SUBOPTIMAL CHOICE IN PIGEONS

An Empirical Research Manuscript

by

Ryan R. McDougle

Department of Psychology

California State University, Chico

Spring 2016

APPROVED BY THE GRADUATE ADVISORY COMMITTEE:

Daniel L. Worthen, Ph.D., Chair

Brian J. Oppy, Ph.D., Member

Abstract

When a stimulus is followed by informative stimuli signaling whether food will follow and another stimulus is followed by non-informative stimuli, pigeons prefer the stimulus followed by informative stimuli even when that stimulus leads to a lower overall probability of food (e.g., 20% vs. 50%). The determinants of this suboptimal choice behavior are unknown. One possibility is that pigeons peck less on informative trials, resulting in a more favorable peck-to-reinforcer ratio. Another possibility is that non-informative trials are steeply discounted due to delayed information regarding reinforcement. Those hypotheses were tested by manipulating the duration of informative and non-informative stimuli and independently manipulating the timing of information by either using chamber blackout to signal reinforcement on non-informative trials or not, while holding the timing of food delivery constant throughout. Preference for the suboptimal choice was apparent when the delay to information was 8s on non-informative trials but that preference was significantly weaker when the delay to information was 2s, even though the peck-to-reinforcer ratio was more favorable for the informative stimuli at both delays. Results indicate that information itself is reinforcing and suboptimal choice in this task reflects delay discounting of the greater probability of food due to the relative timing of information.

Uncertainty Reduction versus Peck-to-Reinforcer Ratio
as Determinants of Suboptimal Choice in Pigeons

When presented with a choice between two stimuli, where the probability of reinforcement is the same regardless of which stimulus is chosen, factors other than the probability of reinforcement can be manipulated to test their impact on preference. When the reinforcer is delayed and uncertain, animals will select a stimulus that precedes stimuli providing information regarding reinforcement or its absence, instead of a stimulus that precedes stimuli that do not provide information regarding reinforcement (Prokasy, 1956; Roper & Zentall, 1999). Moreover, the aforementioned preference is observed even when the overall likelihood of reinforcement is substantially larger for the non-informative stimuli (Gipson, Alessandri, Miller, & Zentall, 2009; Stagner & Zentall, 2010).

Stagner and Zentall (2010) developed a useful procedure for investigating this phenomenon in pigeons, as shown here in Figure 1. During each session of training, some trials presented the AB stimulus (e.g., a plus symbol) and some trials presented the CD stimulus (e.g., a circle). After a single peck (FR1) on this initial link, a terminal link which was either the A (e.g., green) or B (e.g., red) stimulus followed the AB stimulus, or a terminal link which was either the C (e.g., blue) or D (e.g., yellow) stimulus followed the CD stimulus. Each terminal link stimulus was displayed for 10 s with no response requirement (FT10s). The AB stimulus had an overall 20% reinforcement value because it was followed by stimulus A and food 20% of the time and followed by stimulus B and no food 80% of the time. Similarly, the CD stimulus had an overall 50% reinforcement value because it was followed by stimulus C 20% of the time and stimulus D 80% of the time, where both C and D resulted in food at a rate of 50%. Interspersed among those forced trials were choice trials where stimuli AB and CD were presented

simultaneously. A single peck to either stimulus darkened the other option and the trial proceeded according to contingencies identical to the forced trials. Pigeons exposed to this procedure rapidly and reliably developed a distinct preference for AB over CD, indicating a preference for predictive stimuli (A and B) over non-predictive stimuli (C and D), despite the lower probability of reinforcement in the predictive situation.

Considering that the subjects in these experiments sacrifice a substantial amount of food in choosing to experience the informative stimuli, the behavior is appropriately labeled “suboptimal” and calls for a theoretical explanation. Suboptimal choice behavior has been examined in terms of the reinforcing effects of information or uncertainty reduction (Arabian & Desiderato, 1975; Badia & Culbertson, 1972; Dinsmoor, 1983; Green & Rachlin, 1977; Wyckoff, 1952). Wyckoff (1952) found that, when exposed to un signaled alternating schedules of reinforcement and extinction (no reinforcement), pigeons frequently performed an observing response by activating a treadling lever to reveal the current schedule via a colored light. Although this did not change the existing condition, it effectively indicated whether the reinforcement schedule or extinction was in effect. Wyckoff hypothesized that the pigeons engaged in such behavior because information itself is reinforcing. This hypothesis was bolstered by the observation that when the treadle response produced uninformative stimuli, the rate of treadling was low. Although information played a direct role in pigeon behavior in this study, it could be the case that information about the schedule was beneficial only insofar as it allowed the pigeons to optimize pecking behavior in accordance with the prevailing schedule. This would suggest, in contrast to Wyckoff’s interpretation, that information about reinforcement is not itself intrinsically reinforcing.

Prokasy (1956) utilized an E-maze to test preference for informative vs. non-informative stimuli among rats. In this study, after choosing the left or right arm of the maze, the rats were held in a delay box for a short period of time before they were allowed to proceed to the goal box in that arm, which contained food on a random half of the trials. The color of the delay box alternated between black and white to signal pending reinforcement on the informative side of the maze, but the color alteration of the non-informative side did not signal forthcoming reinforcement. Although food was equally likely on either side, the rats preferred the side with the informative delay box. This choice preference, Prokasy claimed, may have been due to the opportunity provided by the informative stimuli to perform differential preparatory responses, such as salivating in anticipation of food but not doing so when food was not forthcoming. It follows that such a biological mechanism may serve to increase the subjective value of the reinforcement (Arabian & Desiderato, 1975). However, it seems that the results of this study are more parsimoniously explained under the assumption that information is intrinsically reinforcing rather than reinforcing only to the degree that it is conducive to initiating a preparatory response. After all, the assumption that there were differences in salivatory behavior was not substantiated by any tangible measure, and the principle manipulation in this study measured alternatives that varied only in their informative value.

The issue of whether informative stimuli are preferred because information is intrinsically reinforcing (Wyckoff, 1952) or whether it is instrumentally reinforcing because it allows adaptive behavioral adjustments (Prokasy, 1956) runs parallel to a theoretical debate concerning the essential qualities of conditioned reinforcers (Hendry, 1969). Indeed, recent research on suboptimal choice has focused on the relative differences between stimuli A, B, C, and D as conditioned reinforcers (Gipson et al., 2009; Roper & Zentall, 1999; Stagner, Laude,

and Zentall, 2012; Stagner & Zentall, 2010; Stagner & Zentall, 2011; Zentall, Laude, Stagner, & Smith, 2015). Those authors attributed the preference for the AB stimulus to how reliably the A stimulus preceded food (100%) when compared with the other stimuli. The fact that the B stimulus never resulted in food and appeared 80% of the time on AB trials was deemed of little consequence in light of the comparative power of stimulus A as a conditioned reinforcer—meaning that pigeons “gambled” on choosing the AB stimulus in hopes of an unlikely A stimulus jackpot. However, Smith and Zentall (2016) found that, when the AB stimulus (50% reinforcement) is pitted against a stimulus that has a single, 100% reinforcing conditioned stimulus as a consequence, pigeons were indifferent between the two options, a result which the authors also attributed to the relative power of the conditioned reinforcers. Unfortunately, this explanation of suboptimal choice could be considered circular and a mere restatement of the experimental parameters rather than a coherent explanation for the choice behavior. Although this literature still lends valuable findings, Zentall and coauthors may have overlooked potentially crucial temporal factors. For instance, Stagner et al. (2012) and Smith and Zentall (2016) both showed that (1) when both concurrent schedules were informative, choice trials resulted in indifference, (2) when only the A and B stimuli were informative, choice trials resulted in suboptimal choice, and (3) when no informative stimuli were present with either option, choice trials favored the optimal (CD) choice. When analyzing these findings, it became apparent that the relative onset of the informative epoch (the time at which an informative stimulus was presented) corresponded to differences in the overall AB preference across conditions, and that perhaps the informative value of the A and B stimuli could alter the perceived comparative value of the C and D stimuli which would cause a disparity in the overall

preference on choice trials, or produce indifference when the informative epochs were matched between alternatives.

Spetch, Mondloch, Belke, and Dunn (1994) suggested that a greater primary reinforcement associated with the optimal alternative (the CD stimulus) can be offset by what Fantino (1969) and Fantino and Abarca (1985) termed delay reduction. This is the notion that a stimulus becomes a conditioned reinforcer to the degree that it signals a reduction in delay to primary reinforcement. Applying this analysis to the suboptimal choice experimental configuration developed by Stagner and Zentall (2010), the A stimulus signaled sooner than average arrival of food when compared to the average delay to the next food on AB trials, making the A stimulus a powerful conditioned reinforcer. The B stimulus signals a delay to primary reinforcement that is only slightly longer than the average delay on AB trials, so any conditioned inhibition associated with the B stimulus is outweighed by the conditioned reinforcement associated with the A stimulus. In contrast, the C and D stimulus signal no change in the average delay to primary reinforcement, so they gain no value as conditioned reinforcers. Thus, only the A stimulus becomes a conditioned reinforcer, and since only AB leads to the A stimulus, AB becomes the preferred alternative. However, because the only discernible delay reduction is associated in the information value of the A stimulus, it suggests that information itself, along with conditioned reinforcement, may factor into the overall comparative reinforcement value of the AB stimulus.

Mazur (1991) has suggested that delay discounting could influence suboptimal choice. More specifically, in situations in which reinforcement is delayed, different reinforcement probabilities do not affect choice preference, but differences in delay to primary reinforcement have powerful effects on choice (Mazur, 1991; Mazur, 1996). However, because the delay to

food reinforcement is the same on the AB trials and the CD trials that do end with food, discounting should be equal across the two types of trials and would therefore result in indifference. So delay discounting of the primary reinforcer—food—does not account for pigeons' preference for AB. But if information about food is considered as the reinforcing event subject to discounting, then Mazur's account predicts that an AB preference would be observed under a wide variety of conditions. When applied to the first appearance of information rather than food, Mazur's delay discounting model has been supported. For example, when the onset of the informative epoch occurred earlier on AB trials than on CD trials, pigeons preferred AB, but when the onset of the informative epoch was held constant between AB and CD alternatives, pigeons were indifferent between the alternatives (Stagner et al., 2012). The delay discounting model, when applied to information, predicts suboptimal choice when (1) the A and B stimuli provide information about the subsequent availability of food, (2) the informative epoch on AB trials is temporally positioned near the initial link but well before food reinforcement or non-reinforcement, and (3) the C and D stimuli do not provide information and the informative epoch is delay until the arrival or non-arrival of the food reinforcement. In the aforementioned example, the delay to primary reinforcement is constant yet an AB preference developed. It seems that the discriminatory information present in the A and B stimuli serve as reinforcement in the form of information in its capacity to signal impending (or non-impending) food presentations by relatively reducing ambiguity among the AB and CD stimuli via information that has been temporally situated apart from the delivery of food.

Delay reduction and delay discounting models of suboptimal choice are not committed to a position on the issue of whether information is intrinsically reinforcing or reinforcing to the extent that it allows an adaptive behavioral adjustment prior to the primary reinforcer (or its

absence). However, it is well known that pigeons peck at stimuli followed by food even when pecking is not required (Brown & Jenkins, 1968). Therefore, pigeons should exhibit high rates of pecking at the A, C, and D stimuli, but not at the B stimulus. On that basis, Liedke and Worthen (unpublished) theorized that there would be a difference in the ratio of responses to reinforcers that causes pigeons to prefer stimulus AB over CD (the Liedke hypothesis). Because the A stimulus appears on only 20% of AB trials, it follows that the C and D stimuli collectively produce more total responses compared with A and B which therefore results in a less preferable ratio of responses to reinforcers on CD trials. Notably, the schedules reported in Roper and Zentall, 1999; Stagner and Zentall, 2011; Stagner et al., 2012; and Zentall et al., 2015, in which a single peck requirement (FR1) on the initial link stimuli was followed by 10-s fixed duration (FT10s) terminal link stimuli, should produce a more favorable peck to reinforcer ratio for AB than for CD over a wide range of typical pigeon pecking rates. Requiring an increased number of responses on the initial link or diminishing the duration of the terminal link stimuli would manifest a reduction in the response per reinforcer advantage theorized for AB trials. Therefore, if pigeons show less preference for the AB stimulus when schedule parameters are manipulated thusly, the hypothesis that “suboptimal choice” actually reflects an optimization strategy whereby pigeons economize their energy resources to an increased advantage within the situation that result in a capital advantage in terms of a response-to-reinforcer ratio. For such a hypothesis to be supported, the preference for AB would have to be shown to be an optimal choice with regard to the amount of pecking invested per reinforcer. Liedke and Worthen found support for this hypothesis by testing factorial combinations of two initial link response requirements (FR1 and FR20) and two terminal link durations (FT10s and FT3s). Pigeons trained first on FR1 FT10s developed an AB preference that persisted through their FR20 FT10s phase, but pigeons

trained first on FR20 FT10s did not have a consistent preference for either AB or CD until they were transferred to FR1 FT10s, which produced an AB preference. The preference for AB was then eliminated with the shortening of the terminal link to FT3s in the final two phases, and a CD preference developed for some pigeons, particularly on FR20 FT3s. Liedke and Worthen interpreted these results as strong evidence that suboptimal choice reflects a preference for more favorable peck-to-reinforcer ratios, since the manipulations that shifted preference toward CD were those that improved the peck-to-reinforcer ratio on CD trials relative to AB trials.

Upon closer inspection of these results, it is not at all clear that manipulations of the peck-to-reinforcer ratios were directly responsible for the shifting of preference because those manipulations also changed the delay to the information epoch in such a way that a delay discounting interpretation would also predict the shifts in preference that were observed. In other words, because changes to the information epoch confounded changes in the peck-to-reinforcer ratios, it is impossible to conclude which of those elements was instrumental in producing the change in preference: relative delay to information or relative peck-to-reinforcer ratios across the trial types. In any case, what may be the cause of this behavior has been illuminated in part by many of these experiments, but it seems that the enigma of suboptimal choice has yet to be fully unraveled.

Moreover, there may be some instances in which the findings need to be reinterpreted. For instance, the relationship between the preference for informative stimuli and the power of conditioned reinforcement associated with the terminal link stimulus should not be considered orthogonal findings. After all, the relative power of any conditioned reinforcer relies heavily on how reliably it predicts the impending arrival of food, which is for all intents and purposes a *de facto* measure of information accuracy. It is also apparent that an underlying thread common to

all of the aforementioned contemporary studies is the existence of a temporal disparity in regards to the timing of the informative epoch among concurrent schedules (e.g., Roper and Zentall, 1999; Stagner and Zentall, 2011; Stagner et al., 2012; and Zentall et al., 2015). The conclusion that, given a reinterpretation of the previous findings, certain avenues of experimentation have yet to be explored undoubtedly has merit.

The initial aim in this study was to manipulate the information epoch within subjects across two groups while maintaining a constant delay to reinforcement to test whether the relative timing of information among choice alternatives held sway over schedule preference. In addition, another aim was to measure an optimality index of peck-to-reinforcer ratio between groups. This course of investigation ultimately resulted in a 2 X 2 mixed measures, competing-hypothesis experimental design.

Method

Subjects

Twelve white carneau pigeons (*Columba Livia*) were obtained from Double T Farm in Glenwood, Iowa. The pigeons were approximately two years of age, of indeterminate sex, and were maintained at 80% of ad libitum weight with freely available water and health grit. Each pigeon was experimentally naive and individually housed under a 14:10 light/dark cycle.

Apparatus

Twelve BRS/LVE operant conditioning chambers were used for the experiment. Each chamber had inside measurements of 35.5 cm high, 30.5 cm long, and 35.0 cm across the response panel. The response panel in each chamber had a horizontal row of three response keys centered 6.75 cm above the floor. The circular keys (2.5 cm diameter) were separated from each other by 5.5cm. Inline projectors mounted behind the transparent surface of the key produced the

experimental stimuli. The experimental stimuli used included a white triangle, a white circle, a white plus, and blue, yellow, red, and green fields of color. The chambers also included a houselight, feeding mechanism, and ventilation fan. Experimental procedures were controlled and data were recorded by 12 Dell desktop PCs interfaced to the operant conditioning chambers by National Instruments, Inc. PCI boards and relays and programmed with National Instruments, Inc. LabVIEW 8.0 software. Each pigeon was trained and tested in the same chamber throughout.

Procedure

Pre-training. In the first phase of pre-training, four pigeons at a time were exposed to a manual feeding program, where the experimenter operated the feeding magazine via a computer mouse. Whenever the magazine was engaged for food presentation, the chamber houselight was simultaneously deactivated while a feeder light was activated. This allowed for the absence of the houselight (blackout) to become associated with food presentations. The feeder was initially engaged for up to 10s of feeding at a time, but this time was quickly reduced to 3.5s and the time between food presentations was incrementally increased until the pigeons responded quickly and reliably to each food presentation.

For the second phase of pretraining, all pigeons were run on automated 40-trial magazine training sessions in which a period of blackout preceding food presentations was gradually lengthened. Each trial started with the houselight on during an intertrial interval (ITI) programmed for a duration of 30s with extended time as needed to require 5s without pecking before beginning the blackout. The houselight was then turned off for the predetermined amount of time, the feeder and feeder light were then activated for 3.5s to allow the pigeon to feed before

the houselight was illuminated again to begin the next ITI. The blackout was increased gradually (1s, 2s, 4s, 6s, 8s) over five sessions culminating in 8s of blackout before food delivery.

In the third phase of pretraining, pigeons were autoshaped to peck at various stimuli: colors (green, red, blue, and yellow) and a triangle on the center key and a circle and a plus symbol on the left and right response keys. Every three sessions were comprised of two sessions of only the color stimuli and one session with only the form stimuli for a total of 12 sessions. Each session included 40 trials with an ITI of 50s, and a stimulus duration of 5s. Color stimuli were followed by 8s of blackout and then 3.5s. Responding during stimulus presentations was not required to receive food, but reliably pecking to all of the stimuli developed for all of the pigeons during this phase.

In the fourth stage of pretraining, an automatic program was used to introduce the pigeons to the various combinations of stimuli that would be used in the experiment. On the center key, a triangle with a one-peck requirement was followed by either the circle or the plus symbol (AB or CD) on the left or right key with a one-peck requirement. Next, on the center key, one of the colored stimuli (A, B, C, or D) appeared for either 2 or 8s, followed by a blackout for either 8s or 2s (to total 10s every trial), and then food was presented for 3s. The form-color sequences were counterbalanced across pigeons, such that each pigeon experienced only the sequences that would constitute its AB and CD trials during the actual experiment.

Training. Figure 2 shows the general configuration of trials during the experiment. Each session included 100 trials: 40 AB trials, 40 CD trials, and 20 AB vs. CD choice trials. Each trial was preceded by a 50s ITI which was extended when necessary so that the ITI did not end until 5s elapsed without a peck. Each trial began with the illumination of the triangle stimulus on the center key, with a one-peck (FR1) requirement. On an AB trial, the AB stimulus (circle or plus

symbol) was presented on one of the side keys (equally often on the left key and right key) with a FR1 requirement which was followed by the A stimulus 20% of the time and by the B stimulus 80% of the time. The A stimulus was always followed by blackout and food reinforcement; the B stimulus was followed by an equal period of timeout during which the houselight remained on and no food was presented. Thus, on AB trials, the A and B stimuli, along with the blackout or its absence, respectively, signaled impending reinforcement or non-reinforcement, respectively. On a CD trial, the CD stimulus was presented on one of the side keys (equally often on the left and right key) with an FR1 requirement which was followed by the C stimulus 20% of the times and the D stimulus 80% of the time. During the first phase of training, the C and D stimuli were both followed by blackout on every trial, but those blackouts followed by food only 50% of the time, so C and D and the blackouts were not predictive of reinforcement or non-reinforcement. Different durations of the terminal link stimuli (A, B, C, and D) and the blackouts formed two groupings of pigeons, with six pigeons in each group. Pigeons in one group, Group STL, always experienced a short terminal link of 2s followed by a blackout or a timeout of 8s, while pigeons in the other group, Group LTL, always experienced a long terminal link of 8s followed by a blackout or timeout of 2s. Pecks to terminal link stimuli were not required but were recorded.

On AB vs. CD choice trials, a single peck was required on the triangle stimulus on the center key, which was followed by both the AB and CD stimuli presented on the side keys (each equally often on each key) with a single peck requirement.. A single peck to either stimulus darkened the other key and presented the same probabilities and outcomes as on forced trials.

There were two phases in this experiment and all subjects experienced each of the phases. In the first phase, six pigeons experienced a FT 8s on the terminal link for the A, C, and D stimuli followed by a chamber blackout of FT 2s; stimulus B was not followed by blackout. By

design, in this condition the information about whether or not reinforcement would consequent was delivered for the AB stimulus as soon as the terminal link A or B was introduced (0s), but 10s later for the C or D stimuli. This condition also allowed for the subjects to peck at the terminal link for up to 8s, the entire length of the terminal link. The other six pigeons in this phase experienced the same amount of time before the delivery of reinforcement (10s) but for them the time of the terminal link and blackout were switched: 2s and 8s respectively. This manipulation allowed the group with the long terminal link (LTL) to perform much more work (pecks) in relation to their reward (food) than the group with the short terminal link (STL). The delay to information as to whether an AB trial would result in reinforcement was 0s (as soon as the terminal links were illuminated) but the delay to information as to whether a CD trial would result in reinforcement was 10s for both groups (terminal link + blackout). There were 40 sessions in this phase of the experiment.

In the second phase, the duration of the terminal link and blackout for both groups were identical to the first phase, but the presentation of the blackout was modified to function as a predictive stimulus on CD trials. In other words, during Phase 2, blackout occurred only on those trials in which food was provided. This subtle change effectively shortened the duration of time between the peck on the CD stimulus and the informative epoch, which was delivered for the C and D stimuli in both groups at the end of the terminal link. Between conditions, the information epoch was temporally changed thusly: from 10s to 8s after pecking CD for the LTL group, and from 10s to 2s after pecking CD for the STS group. Because terminal link stimulus durations were constant across the two phases of the experiment, changes in the peck-to-reinforcer ratios for AB and CD trials were not anticipated. Thus, any shift in preference for AB versus CD across the phases would be attributable to the change in the timing of information on CD trials, which

was a much greater change for Group STL than for Group LTL. There were 40 sessions in this phase of the experiment.

Results

Data collected during the last ten sessions of each phase were analyzed. A 2 (terminal stimulus duration: 8s vs. 2s) x 2 (blackout information: non-informative vs. informative) mixed factorial ANOVA was conducted to analyze the pigeon's behavior on choice trials. A significant between-subjects effect for terminal link stimulus duration, $F(1, 10) = 13.57, p = .004, \eta_p^2 = .58$, indicated that the AB stimulus preference was less for the 2s terminal link ($M = 77.7\%$, $SD = 20.14$) than for the 8s terminal link ($M = 99\%$, $SD = 1.54$). A significant within-subjects effect for blackout information, $F(1, 10) = 16.76, p = .002, \eta_p^2 = .63$, indicated that AB stimulus preference was less in the informative condition ($M = 65.1\%$, $SD = 15.83$) than the non-informative condition ($M = 90.3\%$, $SD = 16.2$). A significant interaction was also found, $F(1, 10) = 16.32, p = .002, \eta_p^2 = .62$ and is depicted in Figure 3. This shows that, with an 8s terminal link stimulus, AB preference differs very little across the non-informative ($M = 99.1\%$, $SD = .6$) and informative ($M = 99.0\%$, $SD = 2.2$) CD choice conditions. However, AB preference was significantly reduced for the 2s terminal stimulus group in the informative condition ($M = 65.1\%$, $SD = 15.83$) compared to the non-informative choice condition ($M = 90.3\%$, $SD = 16.2$).

Table 1 shows the relationship between two measured variables. One is the percentage preference for AB. The other is an "optimality index," which was determined by calculating the ratio of pecks to reinforcers separately for AB trials and for CD trials, then taking a ratio of those ratios (CD:AB). Calculated this way, an optimality index greater than 1.0 indicates that more pecks were made per reinforce on CD trials than on AB trials. An optimality index less than 1.0 indicates the opposite. Therefore, if peck to reinforcer ratio determines preference for AB or CD,

values of the optimality index should have been directly correlated with AB choice percentage. But they were not. A Pearson product-moment correlation coefficient was calculated to assess the relationship between the optimality index and AB choice preference. The results were nonsignificant, ($N = 24$) $r = .18$.

Discussion

The results in this study have replicated the strong AB preferences found in Stagner and Zentall, 2010, and Stagner et al., 2012. Although, in the current study, there were some small differences (larger ITI, and the use of a fixation stimulus, center key for terminal links, etc.), the findings were similar to those typical of the suboptimal choice literature. In contrast to previous studies, this experiment established that a manipulation of the information epoch resulted in a sizable reduction in AB preference.

Group STL's AB preference was significantly lower in the informative phase than in the non-informative phase. When the blackouts were made to be informative on CD trials, the AB preference diminished to the point of indifference for three out of six pigeons in Group STL while two other pigeons maintained a slight AB preference. The hypothesis that suboptimal choice occurs because the A stimulus becomes a powerful conditioned reinforcer while B, C, and D gain little value as either conditioned reinforcers or inhibitors does not predict these results. The frequency with which each terminal link stimulus was followed by food was a constant across the two conditions and the two phases of the experiment, so the dominance of the A stimulus should not have been affected by either manipulation. In other words, since the frequency with which each terminal link was followed by food was not altered by the manipulation of the informational epoch, and since the strength of each stimulus as a conditioned reinforcer or inhibitor should be a function purely of the frequency of pairings with food, altering

the timing of information should not have affected AB preference. Since AB preference was clearly affected when the timing of information was altered, the conditioned reinforcement hypothesis is ruled out as a plausible explanation of suboptimal choice.

The Liedtke hypothesis would have predicted a greater AB preference for Group LTL than Group STL, and that such a preference would persevere through phases. This is because the longer terminal links for Group LTL provided greater opportunity for pecking, and since the pigeons reliably pecked at A, C, and D, but not B, the result was a relatively more favorable peck-to-reinforcer ratio for AB than for CD. However, this predicted difference in AB preference was not observed between groups in the first phase of this experiment. Moreover, the AB preference was affected by manipulating the informativeness of the blackouts in Phase 2, a manipulation that should not have affected relative peck-to-reinforcer ratios and therefore should not have affected AB preference for either group. In fact, the pigeons exhibited a tendency to peck at C and D stimuli at greater rates in Phase 2 than in Phase 1, which resulted in optimality indexes that should have shifted the preference even more firmly toward AB, but the AB preference was actually attenuated for Group STL pigeons in Phase 2. Since the Liedtke hypothesis made predictions that were not supported, it has also been effectively discredited.

The pattern of results across phases cannot be accounted for by the conditioned stimulus hypothesis or the Liedtke hypothesis, but is supported by the uncertainty reduction (information) hypothesis. If information about forthcoming primary reinforcement is itself reinforcing, then it would be expected that its value would be discounted as it is delayed with respect to the response that produces it. Pecking the AB stimulus always produces immediate information, because A and B reliably signal the outcome. Pecking CD produces delayed information; in the standard task (e.g., Stagner and Zental, 2010), information about food is not available prior to the

operation of the feeder or the end of the trial. Since most of these experiments have used an FT10s schedule during the terminal links, this produces a 10s disparity in the timing of information on CD trials compared to AB, more than enough to produce a substantial discounting effect on CD trials. However, the closer the informative epoch follows the initial link, the more powerful its influence should be. Thus, when the informative epoch was shortened from 10s to 2s for Group STL by making the blackouts predictive of food, the disparity in timing of information was substantially reduced, and the preference for AB was attenuated accordingly. However, doing the same for Group LTL did not markedly change the disparity in the timing of information between AB and CD trials, so the AB preference was preserved for pigeons in that group. It appears that all of the results in the suboptimal choice literature can be explained by the relative timing of the informative epoch among choice alternatives. Simply put, the choice with the sooner information epoch is preferred over a later epoch, regardless of the probability of food. Therefore, it has been concluded that the delay-discounting of information is responsible for the data pattern.

As suggested by Wyckoff (1952) and later showcased by Prokasy (1956), information may be reinforcing when it signals forthcoming food presentations or their absence. The current findings suggest that such a preference is not because the information is merely instrumental in allowing the organism to adjust its behavior to optimize responding to gain more preferential outcomes, but because information has an intrinsic value. Prokasy (1956) demonstrated this with his E-maze study but attributed it to differential preparatory behavior rather than uncertainty reduction per se. The current study has upheld the intrinsic reinforcing value of information but also tested whether or not the overall probability of reinforcement predicted by that information altered the choice preference of food associated with said information. If the immediacy of

information is as valuable as the current study suggests, there remains much to be investigated in respect to the temporal mechanism of information regarding suboptimal choice.

The main strength of this study lies in the innovative technique utilized to measure the temporal element of suboptimal choice associated with information, while maintaining an experimental design that had been employed in the previous literature. A weakness of this study lies in the amount of carryover between phases. Although the STL group's AB preference was significantly diminished in the second phase of the experiment, it may be the case that complete indifference would have been achieved, or perhaps a slight CD preference, if the pigeons had not previously been exposed to a condition that produced such a strong AB preference. In Phase 2, information was still substantially delayed on CD trials for Group LTL, so the delay-discounting of information hypothesis would still predict a pronounced AB preference for that group. But considering the much higher probability of food on CD trials, Group STL might really have been expected to develop a CD preference with predictive blackouts occurring so early in CD trials, had they not already developed such a strong AB preference in Phase 1.

Future direction for study should include investigations into the exact nature of the delayed-discounting of information associated with the informative epoch at various temporal distances from the terminal link stimuli. It has been shown that when (1) the informative epochs are identical between optimal and suboptimal alternatives, and (2) that these epochs immediately follow the terminal link, indifference between the alternatives is the consequence regardless of the overall probability of food (Stagner et al., 2012). It would be interesting to uncover exactly what temporal distance would render information non-reinforcing, and under which conditions indifference among optimal and suboptimal alternatives would perpetuate when their informative epochs are matched.

References

- Arabian, J. M., & Desiderato, O. (1975). Preference for signaled shock: A test of two hypotheses. *Animal Learning and Behavior*, *3*, 191-195.
<http://dx.doi.org/10.3758/BF03213429>
- Badia, P., & Culbertson, S. (1972). The relative aversiveness of signaled vs. unsignalled escapable shock and inescapable shock. *Journal of the Experimental Analysis of Behavior*, *17*, 463-471. <http://dx.doi.org/10.1901/jeab.1972.17-463>
- Brown, P. L., & Jenkins, H. M. (1968). Auto-shaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior*, *11*, 1-8.
- Dinsmoor, J. A. (1983). Observing and conditioned reinforcement. *Behavioral and Brain Sciences*, *6*, 693-728. <http://dx.doi.org/10.1017/S0140525X00017969>
- Fantino, E. (1969). Choice and rate of reinforcement. *Journal of the Experimental Analysis of Behavior*, *12*, 723-730. <http://dx.doi.org/10.1901/jeab.1969.12-723>
- Fantino, E. & Abarca, N. (1985). Choice, optimal foraging, and the delay-reduction hypothesis. *Behavioral and Brain Sciences*, *8*, 315-330.
<http://dx.doi.org/10.1017/S0140525X00020847>
- Gipson, C. D., Alessandri, J. J., Miller, H. C., & Zentall, T. R. (2009). Preference for 50% reinforcement over 75% reinforcement by pigeons. *Learning and Behavior*, *37*, 289-298.
<http://dx.doi.org/10.3758/LB.37.4.289>
- Green, L., & Rachlin, H. (1977). Pigeons preferences for stimulus information: Effects of amount of information. *Journal of the Experimental Analysis of Behavior*, *27*, 255-263.
<http://dx.doi.org/10.1901/jeab.1977.27-255>

Hendry, D. P. (1969). Introduction. In D. P. Hendry (Ed.), *Conditioned reinforcement* (pp. 1-33).

Homewood, IL.: The Dorsey Press.

Mazur, J. E. (1991). Observing behavior in pigeons: Choice with probabilistic reinforcement:

Effects of delay and conditioned reinforcers. *Journal of the Experimental Analysis of Behavior*, *55*, 63-77. <http://dx.doi.org/10.1901/jeab.1991.55-63>

Mazur, J. E. (1996). Choice with certain and uncertain reinforcers in an adjusting delay procedure. *Journal of the Experimental Analysis of Behavior*, *66*, 63-73.

<http://dx.doi.org/10.1901/jeab.1996.66-63>

Prokasy, W. F. (1956). The acquisition of observing responses in the absence of differential external reinforcement. *Journal of Comparative and Physiological Psychology*, *49*, 131-134.

Roper, K. L., & Zentall, T. R. (1999). Observing behavior in pigeons: The effect of reinforcement probability and response cost using a symmetrical choice procedure.

Learning and Motivation, *30*, 201-220. <http://dx.doi.org/10.1006/lmot.1999.1030>

Smith, A. P., & Zentall, T. R. (2016). Suboptimal choice in pigeons: Choice is primarily based on the value of the conditioned reinforce than the overall reinforcement rate. *Journal of Experimental Psychology: Animal Learning and Cognition*. Advance online publication.

<http://dx.doi.org/10.1037/xan0000092>

Spetch, M. L., Mondloch, M V., Belke, T. W., & Dunn, R. (1994). Determinants of pigeons' choice between certain and probabilistic outcomes. *Animal Learning and Behavior*, *22*,

239-251. <http://dx.doi.org/10.3758/BF03209832>

- Stagner, J. P., Laude, J. R., & Zentall, T. R. (2011). Sub-optimal choice in pigeons does not depend on avoidance of the stimulus associated with the absence of reinforcement. *Learning and Motivation, 42*, 282-287. <http://dx.doi.org/10.1016/j.lmot.2011.09.001>
- Stagner, J. P., Laude, J. R., & Zentall, T. R. (2012). Pigeons prefer discriminative stimuli independently of the overall probability of reinforcement and of the number or presentation of the conditioned reinforcer. *Journal of Experimental Psychology: Animal Behavior Processes, 38*, 466-452. <http://dx.doi.org/10.1037/a0030321>
- Stagner, J. P., & Zentall, T. R. (2005). Suboptimal choice behavior by pigeons. *Psychonomic Bulletin & Review, 17*, 412-416. <http://dx.doi.org/10.3758/PBR.17.3.412>
- Wyckoff, L. B. (1952). The role of observing responses in discrimination learning: Part 1. *Psychological Review, 59*, 431-442. <http://dx.doi.org/10.1037/h0053932>
- Zentall, T. R., Laude, J. R., Stagner, J. P., & Smith, A. P. (2015). Suboptimal choice by pigeons: Evidence that the value of the conditioned reinforcer rather than its frequency determines choice. *The Psychological Record, 65*, 223-229. <http://dx.doi.org/10.1007/s40732-015-0119-2>

Table 1

Percentage of AB Choices and Optimality Indexes for Each Pigeon across Phases with Non-Informative and Informative Blackouts.

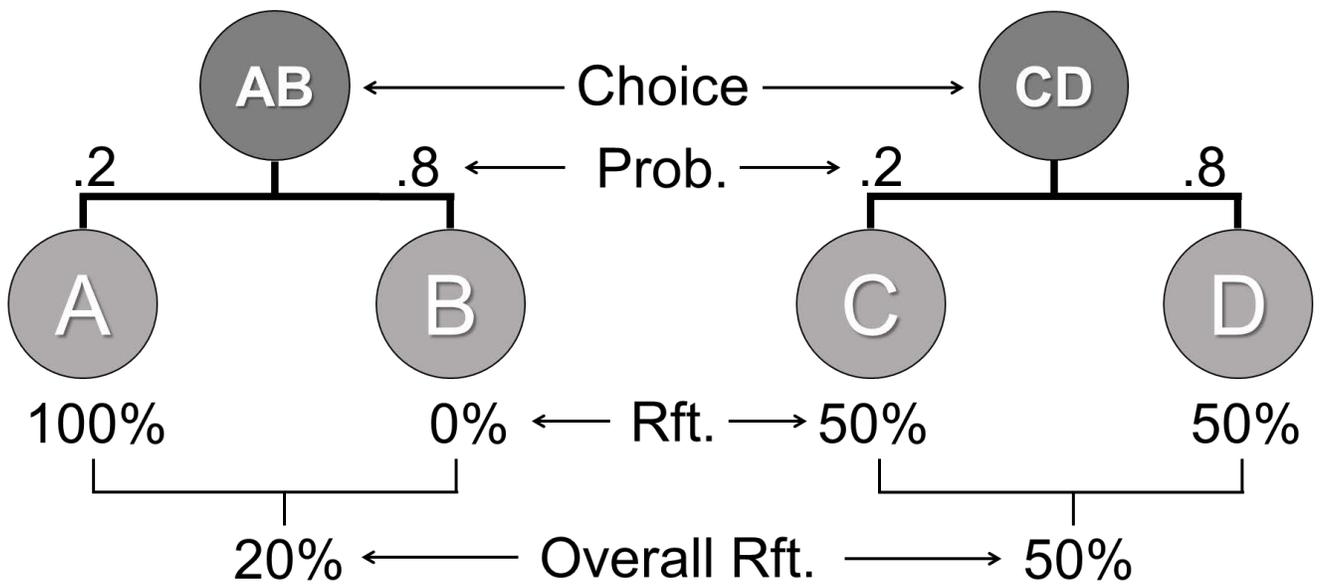
	Subject Number	Non-informative Blackout		Informative Blackout	
		% AB	Optimality Index	% AB	Optimality Index
Group STL	1	98.5	.83	77.0	1.95
	3	100	.77	66.5	1.43
	5	94.5	.78	52.5	1.09
	7	91.0	.77	53.0	1.43
	9	99.5	.57	90.0	1.06
	11	58.0	.63	51.5	.70
Group LTL	2	99.5	.59	100	1.73
	4	99.0	2.07	100	1.75
	6	99.0	.89	94.5	2.21
	8	98.5	1.62	99.0	2.61
	10	100	.63	100	4.76
	12	98.5	.97	100	1.26

Figure Captions

Figure 1. Configuration of stimuli and contingencies in Stagner and Zentall (2010).

Figure 2. Configuration of stimuli and contingencies in the present experiment. During the first phase, blackouts occurred on all CD trials, but during the second phase, blackouts occurred only on trials terminating with food reinforcement.

Figure 3. Percentage of AB (suboptimal) choices as a function of the duration of the terminal link and the informativeness of the blackouts.



Group STL



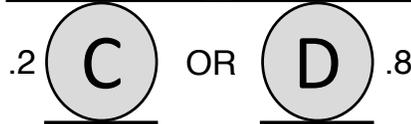
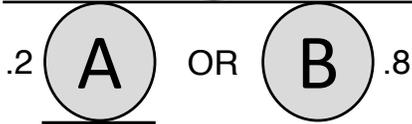
Center Key Fixation Stimulus – FR1



OR



Side Key Initial Link – FR1
(+ or O)



Center Key Terminal Link – FT2s
(Green, Red, Blue, Yellow)



100%



0%



50%



50%

Chamber Blackout – FT8s

Likelihood of Food Reinforcement

Group LTL



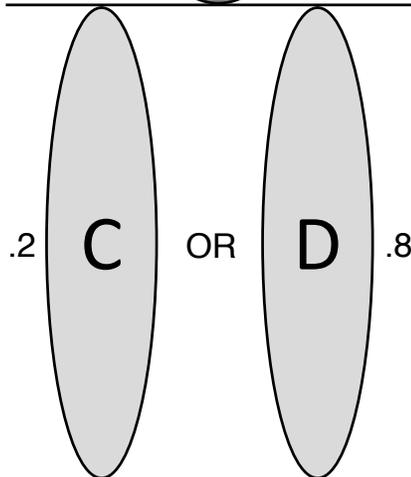
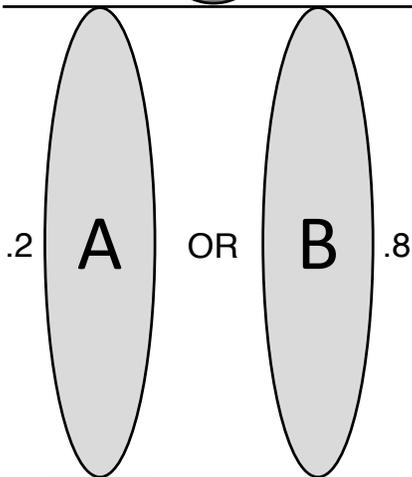
Center Key Fixation Stimulus – FR1



OR



Side Key Initial Link – FR1
(+ or O)



Center Key Terminal Link – FT8s
(Green, Red, Blue, Yellow)



100%



0%



50%



50%

Chamber Blackout – FT2s

Likelihood of Food Reinforcement

