*Journal of Comparative and Physiological Psychology*
1966, Vol. 62, No. 2, 184–192

# VALUE OF KNOWING WHEN REINFORCEMENT IS DUE[1]

GORDON BOWER, JIM McLEAN, AND JACK MEACHAM
*Stanford University*

In 4 experiments concerning preference for knowing when reinforcement is to be delivered—although that information has no apparent instrumental value—pigeons chose between informative and noninformative stimulus sequences. Following an informative choice, the stimulus was correlated with the prevailing interval before reinforcement; following a noninformative choice, it was not. Strong preference for the informative sequence, regardless of the proportion of short vs. long reinforcement intervals, was unaffected by addition of a reliable cue at the moment of reinforcement. Cue strength of the various stimuli was assessed. A control experiment excluded explanations in terms of preference for unpredictable sequences, variety, or specific stimuli. The results are discussed in terms of secondary reinforcement.

These experiments inquire whether a pigeon is reinforced by getting forewarning about the delivery time of a reinforcement. The advance information is "useless" in the sense that the delivery time is predetermined and *S* can do nothing to modify it.

Technically speaking, this is a study of the learning of an observing response ($R_o$) which has no obvious instrumental value. It follows a tradition begun by Wyckoff (1952) and Prokasy (1956) and since supplemented by many others. The present studies follow Prokasy's design: on each trial *S* has an explicit forced choice between two alternatives (like the two arms of a **T** maze). One alternative is designated as the observing response, $R_o$, and the other as the nonobserving response, $R_n$. Following $R_o$, *S* receives at random one of two discriminative stimuli, the two correlated with presence or absence, respectively, of reinforcement for a subsequent goal-directed response. Following $R_n$, *S* receives a stimulus which has no correlation with reinforcement. However, following either choice, the distribution of reinforcements is the same. The only differential is that, following $R_o$, the animal receives information about whether it will be rewarded on this trial.

In such experiments *S*s typically learn a consistent preference for $R_o$. A variety of parameters which influence the level of preference and how quickly it develops have been investigated. Mitchell, Perkins, and Perkins (1965) found faster learning with greater primary reward, with practically no $R_o$ preference developing when a very small reward was used. Wehling and Prokasy (1962) found that $R_o$ was learned faster when *S*s were trained under higher levels of deprivation for the reinforcer. Lutz and Perkins (1960) varied the forewarning period between $R_o$ and the occurrence of reward or nonreward, and found that $R_o$ was learned for all periods except their minimum (0-sec. delay), where the forewarning occurred at best only a fraction of a second before the reward itself. In a second experiment, Mitchell et al. (1965) found that *S*s chose an $R_o$ followed by immediate information in preference to another $R_o$ followed by delayed information. In related work on punishment, Knapp, Kause, and Perkins (1959), Lockard (1963), and Perkins, Levis, and Seymann (1963) have shown that rats prefer a situation in which they receive a forewarning stimulus before an unavoidable electric shock to an alternative situation in which there is no such forewarning before the shock.

In the studies done with the Prokasy paradigm, the informative stimuli given following $R_o$ were always correlated with presence or absence of reward for subsequent instrumental responding. But this extreme presence-absence differential is probably not necessary for $R_o$ to be

[1] Supported by National Institute of Child Health and Human Development Research Grant HD-00954.

learned. A plausible conjecture is that $R_o$ would be learned whenever the cues following $R_o$ provide reliable information about any sizable differential in subsequent reinforcement, whether in its magnitude, delay, quality, or probability of occurrence. The present experiments test this generalization with a difference in intervals (delays) of reinforcement. Besides assessing preference for $R_o$, an additional aim was to collect data regarding *S*'s response strength to the several discriminative stimuli involved. In previous studies with the Prokasy paradigm, these strengths have typically been inferred and not assessed directly.

Our experiments adapted Prokasy's forced-choice maze procedure to a discrete-trial procedure for pigeons in a Skinner box. The paradigm for a choice trial is shown in Figure 1. A choice trial started with two response key lights coming on, one with a vertical bar, the other with a horizontal bar, the two in positions randomly alternating from trial to trial. When *S* pecked one of these keys, the other key darkened and became inoperative so that *S* had to finish the trial on the key first chosen. The initial response to the informative key turned it to red or green with probabilities $\pi$ and $1 - \pi$, respectively. Responding to the green key was reinforced on a 10-sec. fixed interval (FI) schedule; that is, the first response occurring after 10 sec. had elapsed was rewarded with food. When the key was red, reinforcement was set up on a 40-sec. FI. Let us call green and red the "short" and "long" stimuli, referring to the delay before reward associated with them.

Consider now the other half of Figure 1. If the noninformative stimulus was chosen first on a trial, then it always changed to yellow; this was on FI 40 or FI 10 with the same probabilities as before, $\pi$ and $1 - \pi$, except *S* had no way of knowing which delay was in effect on any trial. Call yellow the "mixed" stimulus since it was associated with a random mixture of 10-sec. and 40-sec. delays. On each trial, one reinforcement was given, followed by a 10-sec. time-out, followed by the next trial.

In this setup, as in Prokasy's, there is no

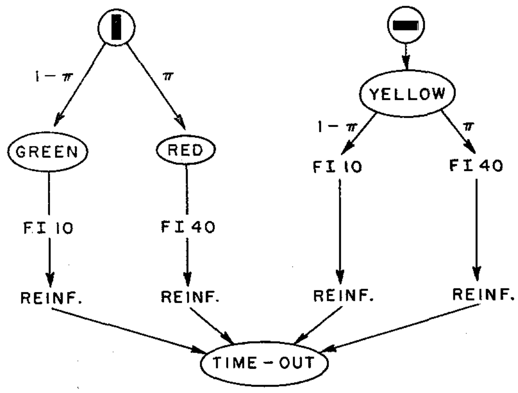


Fig. 1. Schematic representation of events on a free-choice trial, showing the stimuli projected onto the response keys and the reinforcement conditions in each.

obvious reinforcement differential favoring the information sequence. The vertical and horizontal bars are always followed by reinforcement, in the same amount, and the distribution of delay times before reinforcement is the same following each stimulus.[2] It is therefore of interest to see whether *S* will prefer to get this information which has no obvious instrumental value.

## Experiment 1

The first study used $\pi = .50$; with the two delays equally likely, the information conveyed by the green and red colors was at a maximum. In addition to assessing preference for the informative stimulus, auxiliary procedures were run to assess preference orderings of the long-, short-, and mixed-delay stimuli, and also to measure extinction responding to the informative and noninformative stimuli.

### Method

*Subjects.* The *S*s were three male White Carneau pigeons maintained at 80% ad-lib body weight throughout the experiment.

*Apparatus.* The chamber was a two-key pigeon box purchased from Lehigh Valley Electronics, with associated programming equipment located in an adjoining room. Throughout each session white noise was continuously present in the box. Reinforcement was provided by 3-sec. access to a

[2] Due to the characteristic response pattern induced by FI schedules, most reinforcements were obtained within 1 sec. of their setup time.

lighted grain hopper located between and below the two side keys. Between trials (reinforcements) the box lights and key lights were turned off for 10 sec.

*Forced trials.* After preliminary shaping of key pecking, *S*s had forced exposures to the four sequences of Figure 1, i.e., vertical-red FI 40, vertical-green FI 10, horizontal-yellow FI 40, and horizontal-yellow FI 10. Each sequence occurred in scrambled order 10 times per session, 5 trials on the left key, 5 on the right. Such training continued for 16 daily sessions.

*Free choices.* There followed 5 sessions of 40 trials, each permitting a choice between the informative and noninformative sequences, and each sequence being reinforced in customary fashion.

*Color preferences.* The next session began with 30 nonreinforced preference tests with the three color pairs: red vs. green, red vs. yellow, and green vs. yellow. The pairs occurred 10 times each in scrambled order with left-right positions randomized. The first choice response terminated the trial (no reward), leading to a 10-sec. blackout. Following these 30 tests, the session ended with 40 reinforced free choices between the informative vs. noninformative sequences.

*Extinction.* On the final day extinction tests were run with the informative (vertical bar) and noninformative (horizontal bar) stimuli. These were presented singly in random alternation for 15-sec. periods separated by 10-sec. blackouts, with left and right keys used in a random order over successive trials. A response to either stimulus had no effect. Trials continued until *S* failed to respond to each stimulus for 3 consecutive presentations.

### *Results*

The results are best summarized by saying that *S*s overwhelmingly preferred to get the information. The percentage choices of the informative stimulus for the three *S*s over the 200 free-choice trials were 95, 85, and 98, averaging to 93. No trends over the test days were apparent.

Response rates were well controlled by the color stimulus present. The average response rates to the short-, mixed-, and long-delay stimuli were 2.97, 1.42, and 1.24 responses per second, respectively. The short-delay stimulus commanded the highest response rate, and the rate to the mixed-delay stimulus was only slightly but consistently ahead of that to the long-delay stimulus. Each *S* was very consistent from day to day in this ordering of the three rates.

On the color pair comparison tests, the stimuli were preferred in the order of short, mixed, then long delay. The average percentage choices were: short over long, 100; mixed over long, 100; short over mixed, 77. The 77 score would probably have been higher with more observations or with a less confusing test series (i.e., not intermixing the three test pairs). It was observed (not measured) that speed of response on the choice test was fast when the short-delay stimulus was available, but slower on the mixed- vs. long-delay test pair.

Behavior to the vertical and horizontal bars—the stimuli at the beginning of the informative and noninformative sequences—was much as expected from the preference score. Responses during extinction to the informative and noninformative stimuli, respectively, for the three *S*s were 145, 70; 222, 29; and 417, 142 (averages of 261 to the informative stimulus and 80 to the noninformative stimulus). Also on the earlier forced trials it was observed (not measured) that initial response speed was fast to the informative stimulus and slower to the noninformative stimulus.

These results all tie together in a consistent package. The color stimuli were preferred and responded to according to their associated delays, and the informative stimulus was preferred and responded to more than the noninformative stimulus.

## Experiment 2

In the next experiment we sought conditions (other than unbalancing reinforcement) that might force a preference for the noninformative stimulus. In analogy to newspaper reading, we thought that if the information obtained is usually bad news, then *S* might prefer getting no news at all. Accordingly, this experiment was run with $\pi = .80$, where the long-delay condition predominated.

### *Method*

The general conditions and procedures of Experiment 1 were used with three new pigeons. The differences were: (*a*) the long-delay condition (FI 40) occurred 4 out of every 5 trials; (*b*) more sessions (10) and more trials per session (60) were given on the free-choice tests; (*c*) more sessions (4) and more trials per session (60) were given on color pair comparison tests, each with customary reinforcement and with choice latencies recorded; and (*d*) the final extinction test of Experiment 1 was deleted.

*Results*

The results were much the same as in Experiment 1, so we failed to induce a preference for the noninformative stimulus. Percentage choices of the informative stimulus over 600 trials were 93, 95, and 91 for the three *S*s, averaging to 93, the same as in Experiment 1. The average responses per second to the short-, mixed-, and long-delay colors were 2.23, 1.14, and 0.91, respectively, showing the same pattern as before. On the color pair tests, the percentage choices on the three types were: short over long, 100; short over mixed, 92; mixed over long, 90. The average choice latencies on these three types of tests (as ordered above) were 1.00, 0.95, and 3.52 sec., respectively. All *S*s consistently showed this pattern of latencies. Thus choice was quick when the short-delay stimulus was available (and usually chosen) but slow when it was not available. Thus despite the change in the proportion of long-delay trials, the results essentially duplicated the consistent pattern found in the first experiment.

## Experiment 3

The previous results appear to demonstrate that it is reinforcing to get information that reduces uncertainty about time delays. However, the short- and long-delay stimuli reduced uncertainty only partially; some uncertainty remained because *S* still must time out 10 or 40 sec., and there was some error or uncertainty in its doing so. But suppose that a universally reliable cue is provided which always tells *S* the exact moment when reinforcement sets up and is due. The question is whether the addition of such a reliable cue now preempts and negates the value of being forewarned about when reinforcement is due.

This condition was arranged in a $\pi$ = .50 situation by always projecting a white cross onto the key at the moment reinforcement was set up, so that the next peck on the crossed key would be reinforced. This cue was always added onto the short-, long-, and mixed-delay stimuli, whenever reinforcement was due. It would seem that an efficient strategy here would be to start the choice trial by pecking either the informative or noninformative cues, then sitting and waiting for the cross to appear, and then pecking it for the reward. If followed, *S*s should show no preference between the informative and noninformative stimuli that begin the trial.

*Method*

The *S*s from Experiment 2 were run with the following changes in procedure: (*a*) $\pi$ was set at .50; (*b*) the informative stimulus was changed to a 45° right-tilted white line, the noninformative stimulus to a 45° left-tilted line—both being midway between the vertical and horizontal lines used in Experiment 2; (*c*) a white cross was projected onto the colored key at the moment reinforcement set up; (*d*) 48 trials per session were run. Forced trials to the single stimulus sequences were run for 12 daily sessions, followed by 2 sessions of free choices between the informative and noninformative sequences. No pair comparison tests were run.

*Results*

The initial conjectures were wrong; the reliable cue did not negate the value of being forewarned about when reinforcement was due. Percentage informative choices were 99, 99, and 85 for the three *S*s, averaging to 94—about the same as obtained without the cross. The first two *S*s listed gave evidence of noticing and using the cross during the mixed- and long-delay stimuli, since they tended more than before to wait until the cross appeared before pecking. This effect showed up in their lower response rates with the cross available compared to their rates in Experiment 2 without it: in the mixed stimulus their average response rate went from .90 responses per second in Experiment 2 down to .47 in Experiment 3; in the long stimulus, from .61 down to .14 responses per second. For unknown reasons the third *S* actually increased its response rate to all stimuli in Experiment 3.

In sum, *S*s still preferred to get advance information about the delay time despite the availability of a reliable cue to mark the exact moment when reinforcement became available. In retrospect, this result may perhaps be viewed as a corollary of the previous results. Regarding its temporal contiguity to reinforcement, the cross is much like the sound of the feeder in the former studies; i.e., both cues elicit short behavior chains that soon bring food into the mouth. Yet we know from the previous work that the feeder sound does not preempt the value of prior information. A related result is that by Weyant (1957), who found that a regular signal sequence (an "external clock") preceding a constant-delay reward did not enhance its reinforcing value relative to the same constant delay without an external clock. Either Weyant's *S*s' internal clock was as accurate as the external one provided, or else exact timing of reward delivery gave no advantage over approximate timing. The implication of Weyant's result for the present study is that addition of the cross would not alter the preferences found in the prior experiments.

## Experiment 4

This was a control experiment designed to rule out several uninteresting explanations of the previous results. In the prior studies, *S* saw one of two different stimuli following unpredictably upon an informative choice, but only the same predictable color following a noninformative choice. Conceivably, choice of the former reflects only a preference for unpredictable over predictable stimulus sequences, or a preference for variety, or simply innate liking for the particular colors involved. Such factors were ruled out in the following experiment wherein the same two colors followed an informative or a noninformative choice. The difference was that when projected onto the informative key, the colors were consistently correlated with the prevailing delay (FI 10 or 40), whereas on the noninformative key the colors did not correlate with the delay. This, in fact, is the control condition usually run with the Prokasy paradigm.

### *Method*

The three *S*s from Experiment 3 were used again. The right-hand key was the informative side, and its sequence started with the 45° right-tilted line; the left-hand key was the noninformative side, and its sequence started with the 45° left-tilted line. Following a peck to either key, green or red was put on the chosen key a random half of the trials. On the informative side, green indicated FI 10 and red, FI 40. On the noninformative side there was no correlation, with FI 10 or 40 occurring unpredictably half of the time with each color. Forced trials on the possible stimulus sequences and delays occurred at a total of 48 a day for 12 days. There followed 1 session of 48 free-choice trials between the informative and noninformative sequences. Finally, there were 2 sessions of pair comparison tests on the colors with their customary reinforcement. The four test types were the combinations of red and green on the informative vs. noninformative side. Each test type occurred 12 times per session randomly mixed among the others.

### *Results*

Due to their training in Experiment 3, *S*s began with a high response rate to green and a low rate to red, irrespective of the side of key on which these colors occurred. For example, on the first day, rate to green on the noninformative side averaged 2.42 times faster than the rate to red on the noninformative side. However, over the course of training, *S*s learned appropriate patterned discriminations based on color plus side of key. By the last 4 days of training all *S*s had come to respond differently to green vs. red on the informative side (ratio of average rates = 2.22), but not very differently to green vs. red on the noninformative side (rate ratio = 1.07). Green-on-right commanded the highest response rate (2.09 responses per second) and red-on-right the lowest (.94), whereas green-on-left (1.76) and red-on-left (1.64) commanded approximately equal and intermediate response rates.

During the free-choice tests, all *S*s chose the informative key in preference to the noninformative key on all 48 trials (i.e., 100%).

To summarize the color pair comparison tests, let us use R and G to abbreviate the colors, and I and N the informative and noninformative sides. From most to least preferred, the stimuli ranked in the order of GI, GN, RN, RI. The average pairwise choice percentages over the 24 test trials were: GI over GN, 81; GI over RN, 92; GN over RI, 99; and RN over RI, 82.

These choice percentages are qualitatively in line with expectations from the response rates listed above for the various stimuli (e.g., GN is somewhat stronger than RN).

These results rule out explanations of the previous "information" preference as due to a preference for unpredictable sequences, or variety, or particular colors. The informative side came to be valued simply because it provided reliable information about the prevailing interval of reinforcement. The results demonstrate that stimulus compounds (color + location) can acquire discriminative control over responding, hardly a surprising fact; but they help not at all in understanding the reason for the information preference.

## Discussion

The results support the initial conjecture, at least for differentials in interval of reinforcement: *S*s prefer to know *when* a reward will be delivered. The intriguing puzzle is why this should be so. It is important to realize that the obvious reward and performance variables that could differentiate the aftermath of $R_o$ and $R_n$ were either equated or even overbalanced against $R_o$ in these experiments. For example, because responding to the long-delay stimulus was occasionally depressed (about 3% of trials), probably due to a "contrast" effect, reinforcement was postponed on such trials, producing a slightly longer average interval of reinforcement following $R_o$ than following $R_n$. Also, the response rate data presented earlier show that the total response output following $R_o$ often equaled or even exceeded that following $R_n$, so explanations of the $R_o$ preference in terms of some law of least effort would seem ruled out. The challenge of the puzzle is that the $R_o$ preference developed regularly and consistently in each *S*, despite our inability to understand why it should.

Although it is convenient to describe these results in terms of information and the reduction of uncertainty about the point of reinforcement, better continuity with related results is established if the discussion is cast in terms of the utility or secondary-reinforcement value of the short-, long-, and mixed-delay stimuli. It is supposed that these values are induced, in turn, onto the I and N stimuli which precede the colors, thus determining the preferential choice of the I stimulus.

The relevant considerations are shown in Figure 2. Since short delays were preferred to long delays, the stimulus associated with the short delay is assigned a high value, and that associated with the long delay a low value. To establish the standard for comparison, it will be assumed that the value induced onto the informative stimulus (vertical bar) is the objective average of the values of the short- and long-delay stimuli which follow it. This average value is depicted in Figure 2 for the $\pi = .50$ condition. The only question is, what value are we to assign to the mixed stimulus (and hence N) which sometimes has short and sometimes long delays? The objective expected value of the mixed stimulus would be the sum of the short and long values weighted by their probabilities in the mixture. This rule generates the straight line in Figure 2, depicting the decreasing value of the mixture as the percentage of long delays is increased. But this rule obviously will

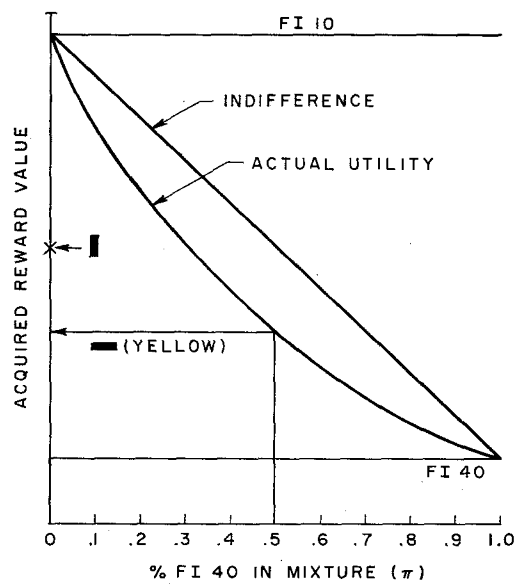


FIG. 2. Hypothetical curve relating utility of the mixed stimulus to the percentage of long delays in the mixture.

not suffice because it implies that the *Ss* should have shown no preference for the informative stimulus.

Our data instead require that the value of the mixture be given by a concave-upward utility curve such as that shown in Figure 2. This is one previously proposed by Wyckoff (1959), who was interested in essentially the same problem. For the concave function, the value of the mixed (yellow) stimulus is always less than the objective average value of the long- and short-delay stimuli considered separately.[3] This value of the mixed stimulus is assumed to be induced onto the noninformative stimulus which precedes it, thus making it lower than the value of the informative stimulus. The concave utility curve in Figure 2 implies that the informative stimulus will be chosen in preference to the noninformative stimulus for all values of $\pi$, except at the end points of 1 or 0 where no uncertainty is involved. This agrees with the results of Experiment 2.

It is easy to show that this concave utility curve accounts for all the results. First, the informative stimulus is preferred. Second, in the pair comparisons, the short-delay stimulus is preferred to the others, and the mixed stimulus is preferred to the long-delay stimulus. Third, assuming that rate of responding to a stimulus is monotonically related to its value, the results on response rate fit this picture fairly well; i.e., response rate to the short-delay stimulus was much higher than to the other stimuli, and rate to the mixed stimulus was slightly higher than to the long-delay stimulus. Fourth, consistent with this view, the differential value of the informative sequence is retained despite the addition of a reliable cue, the cross, at the moment of reinforcement.

Besides accounting for our results, the function in Figure 2 suggests an interpretation of several conditions reported in the literature in which a significant $R_o$ preference did not develop within the allotted training trials. The level of preference and its rate of development will depend on the value differential favoring I over N, which in turn depends on the value differential between the better vs. the poorer reward conditions in Figure 2. Therefore, one may expect an $R_o$ preference to develop slowly or not at all when the values of the two reward conditions differ only slightly. Most of the failures to find $R_o$ learning may be explained in this way. Thus, Mitchell et al. (1965) found no $R_o$ learning when a very small reward (vs. nothing) was used. Also, Kendall (1965), using Wyckoff's paradigm in which an $R_o$ produced brief colors on a response key correlated with different fixed-ratio schedules of reinforcement, found that an $R_o$ preference disappeared when the two fixed ratios were made sufficiently similar. Furthermore, assuming that net value varies with the product of drive level and reward magnitude, as in Hull's (1952) $D \times K$ formula, then the principle explains the poor learning of $R_o$ under low drive found by Wehling and Prokasy (1962). To repeat, the principle is that the values of the two reward conditions must be sufficiently different before it is worth getting advance information about which will occur. The indeterminacy in this statement is the current imprecision surrounding the phrase "sufficiently different."

Figure 2 also accounts for results obtained by Kendall and Gibson (1965). $R_o$ was initially acquired in a condition in which it briefly changed the color on the response key from white (the mixed stimulus) either to blue, associated with a highly valued fixed-ratio schedule of reinforcement, or to red, associated with a poorly valued fixed-interval schedule. Then one of the colors following $R_o$ was deleted, replacing its occurrence by white. When the low value (red) stimulus was deleted, the $R_o$ output was maintained at a high level; when the high value (blue) stimulus was deleted, $R_o$ extinguished to a very

[3] This value of the mixed stimulus is not inconsistent with the results of Logan (1960) or Herrnstein (1964), who found that a mixed-delay (VI) stimulus was preferred to a stimulus having a constant delay (FI) at the average of the mixed delays. In terms of Figure 2, their data would tell us to locate the value of a constant FI 25 below the value of the mixed stimulus involving half 10-sec. and half 40-sec. intervals.

low level. In terms of Figure 2, $R_o$ was maintained in the first condition because the average value of the blue and white stimuli following $R_o$ still exceeded the value of white alone following $R_n$; $R_o$ extinguished in the second condition because the average value of the red and white stimuli following $R_o$ was less than the value of white alone following $R_n$.

In sum, Wyckoff's utility analysis appears to account parsimoniously for most of the results. An annoying feature is that the concave utility curve in Figure 2 is inferred from the data in a post hoc manner. An important question is whether more elementary considerations (theoretical assumptions) can be found which either imply the concave utility curve or else imply $R_o$ learning and which admit of some independent support. Of several possibilities that may be proposed—including frustration, contrast effects, distortions of subjective probabilities—we choose for discussion a proposal of Perkins' (1955) which has been claimed to have the requisite properties.

Perkins proposes that the reinforcing value of a given reward will vary with the level of some appropriate "preparatory response" (PR) occurring just before the reward. The prototypic illustration of a PR is salivation just before dry food is introduced into the mouth. For each amount or quality of reward, an optimal level and type of PR is assumed to exist; it is optimal in the sense that occurrence of this level maximizes the reinforcing value of that particular reward. It is assumed that when possible *S* will learn to make this optimal level of PR due to built-in differential reinforcement of variations in the level. Thus, in classical conditioning *S*s learn to salivate because food-plus-saliva is a more reinforcing state of affairs than is food-plus-no-saliva. With large food rewards, CR amplitude is larger because much-saliva-plus-much-food is more reinforcing than little-saliva-plus-much-food. Similarly, *S*s extinguish salivation because no-saliva-plus-no-food is more reinforcing than is saliva-plus-no-food. This all appears to be post hoc circular reasoning, i.e., the only index of the optimal level is that level of PR which *S* eventually learns. However, Perkins argues that it is not always circular, the main demonstration of this apparently being that the PR proposal predicts that otherwise "useless" observing responses will be learned. And the proposal clearly does predict $R_o$ learning. Following $R_o$, the discriminative cues enable *S* to make differential PRs to optimize the value from whichever reward is signaled; following $R_n$, PRs will usually be less favorable for optimizing the value from whichever reward occurs. For our experiments, the small "scallops" in response rate during the green FI 10 and red FI 40 intervals were presumably correlated with the temporal development of the appropriate PRs (cf. Shapiro, 1961).

Perkins thus accounts for $R_o$ learning by supposing an enhancement of the reward values following $R_o$. It is not that *S* is using a different averaging process on the mixed vs. the multiple schedules (as in Figure 2), but rather that the values being averaged are higher in the latter case. By judicious selection of the hypothetical values, the PR analysis can account too for the Kendall and Gibson (1965) findings, so it does not imply that *S* always wants to know the prevailing condition of reinforcement.

An unfortunate property of the PR theory is that it appears incapable of independent support from any experiments except those on observing responses; by the same token, it appears to have no determinate implications for any experiments except those on observing responses. This appearance is perhaps only temporary, and appropriate ingenuity may discover other testable consequences of the theory. The difficulty is that the utility of a given reward clearly depends on what it is as well as whether it is precisely expected. The former component of utility is almost never known—except for "equality" of two conditions, as in the $R_o$ studies—whereas the PR theory refers only to the excess utility added by the second factor. Thus, for example, the PR theory cannot predict that, say, a constant reward will be preferred to random variations between smaller and larger rewards averaging to the same amount, because the first component of utility for the three rewards is unknown. The theory simply has no determinate predictions for such studies.

## REFERENCES


HERRNSTEIN, R. J. Aperiodicity as a factor in choice. *J. exp. Anal. Behav.*, 1964, **7,** 179–182.

HULL, C. L. *A behavior system.* New Haven: Yale University Press, 1952.

KENDALL, S. B. An observing response analysis of fixed ratio discrimination. *Psychon. Sci.*, 1965, **3,** 281–282.

KENDALL, S. B., & GIBSON, D. A. Effects of discriminative stimulus removal on observing behavior. *Psychol. Rep.*, 1965, **15,** 545–551.

KNAPP, R. K., KAUSE, R. H., & PERKINS, C. C., JR. Immediate versus delayed shock in T-maze performance. *J. exp. Psychol.*, 1959, **58,** 357–362.

LOCKARD, J. S. Choice of a warning signal or no warning signal in an unavoidable shock situation. *J. comp. physiol. Psychol.*, 1963, **56,** 526–530.

LOGAN, F. A. *Incentive.* New Haven: Yale University Press, 1960.

LUTZ, R. F., & PERKINS, C. C., JR. A time variable in the acquisition of observing responses. *J. comp. physiol. Psychol.*, 1960, **53,** 180–182.

MITCHELL, K. M., PERKINS, N. P., & PERKINS, C. C., JR. Conditions affecting acquisition of observing responses in the absence of differential reward. *J. comp. physiol. Psychol.*, 1965, **60,** 435–437.

PERKINS, C. C., JR. The stimulus conditions which follow learned responses. *Psychol. Rev.*, 1955, **62,** 341–348.

PERKINS, C. C., JR., LEVIS, D. J., & SEYMANN, R. Preference for signal-shock vs. shock-signal. *Psychol. Rep.*, 1963, **13,** 735–738.

PROKASY, W. F. The acquisition of observing responses in the absence of differential external reinforcement. *J. comp. physiol. Psychol.*, 1956, **49,** 131–134.

SHAPIRO, M. M. Salivary conditioning in dogs during fixed-interval reinforcement contingent upon lever pressing. *J. exp. Anal. Behav.*, 1961, **4,** 361–364.

WEHLING, H. E., & PROKASY, W. F. Role of food deprivation in the acquisition of the observing response. *Psychol. Rep.*, 1962, **10,** 399–407.

WEYANT, R. G. The relation between preparatory responses and the effectiveness of delayed reward. Unpublished Master's thesis, Kent State University, 1957.

WYCKOFF, L. B. The role of observing responses in discrimination learning. Part I. *Psychol. Rev.*, 1952, **59,** 431–442.

WYCKOFF, L. B. Toward a quantitative theory of secondary reinforcement. *Psychol. Rev.*, 1959, **66,** 68–78.


(Received January 10, 1966)