Founded in 1887 by G. Stanley Hall

OFFPRINTED FROM

THE AMERICAN JOURNAL OF PSYCHOLOGY

EDITED BY

KARL M. DALLENBACH
University of Texas

AND

M. E. BITTERMAN
Bryn Mawr College

E. B. NEWMAN
Harvard University

THE ROLE OF LEARNING, PERCEPTION, AND REWARD IN MONKEYS' CHOICE OF FOOD

By William A. Wilson, Jr., University of Colorado

December, 1959, Vol. LXXII
pp. 560-565

Published by The American Journal of Psychology, Department of Psychology, University of Texas, Austin, Tex.

THE ROLE OF LEARNING, PERCEPTION, AND REWARD IN MONKEYS' CHOICE OF FOOD

By William A. Wilson, Jr., University of Colorado

Harlow and Meyer have reported that monkeys given a choice between two different amounts of food usually choose the larger amount, although 'errors' (*i.e.* choices of the smaller amount) sometimes occur.[1] The authors recognize that the ability of the animals to discriminate differences in amount of food may enter into such choices—at one point in their paper they account for an aberrant finding in terms of "the ease with which a half-peanut may be confused with a whole peanut"—but they give the impression that their results are to be understood primarily in terms of the reinforcing properties of reward. Whether there was any evidence of learning during the early stages of testing, they do not say. It may be significant, however, that the animals used had earlier been trained in paired comparisons.

In the experiment reported here an attempt was made to analyze some of the factors entering into choices of the kind studied by Harlow and Meyer. Different amounts of food were presented to monkeys, and choices were recorded until a stable level of performance was achieved. Then the situation was so changed that the amounts of food presented continued to function as cues but not as rewards for choice; that is, discriminative and reinforcing factors were unconfounded.

Subjects. The *S*s were six immature rhesus monkeys. All had extensive previous training in a probability-learning situation for peanut-reward.[2] They also had some previous experience with an automatic apparatus which yielded a single pellet of glucose for correct response.

Apparatus. Within a Wisconsin General Test Apparatus a special testing board was presented to *S*. This board was painted flat black except for two gray squares, 2.5 in. on a side and 5 in. apart. A piece of 1/8-in. plywood, also black, raised the surface of the 2.5 × 5-in. space between the squares. For Series I, each of the squares was covered by a transparent 2.5-in. square lucite box, 1 in. deep, placed with its open top down and hinged at the side away from the monkey. The food-reward, placed on the gray square, could easily be seen through the bottom of the box and obtained by tilting the box. For Series II, a special set of five boxes was prepared. Transparent tops were placed on the boxes and on the inner sides of the tops appropriate numbers of glucose-pellets were glued. The last 0.5 in. of the top farthest away from the monkey was cut out, and a strip of aluminum was

* Received for publication November 12, 1958. This research was done at the Institute of Living, Hartford, Connecticut. It was supported in part by the Office of the Surgeon General (Department of the Army), Contract DA-49-007-MD-763.

[1] H. F. Harlow and D. F. Meyer, Paired-comparisons scales for monkey rewards, *J. compar. & physiol. Psychol.*, 45, 1952, 73-79.

[2] W. A. Wilson, Jr., and A. R. Rollin, Two-choice behavior of rhesus monkeys in a noncontingent situation, *J. exp. Psychol.*, in press.

so fixed to it that *S* could not see into this part of the box. These boxes were placed topside-down upon the squares, as described below. A microswitch was attached to the testing board next to each gray square, and one or the other was tripped when *S* responded by lifting a box during either series.

Procedure. Since the *Ss* had previous experience with similar boxes in a similar apparatus, no special preliminary training was given. There were two experimental procedures (Series I and II).

Series I. Glucose-pellets, 0.05 gm. in weight and 4 mm. in diameter, were used, in the five amounts of one, two, three, four, and five pellets. On each trial in Series I, one of these amounts of reward was placed under one box and another amount was placed under the second box. (In both series, the pellets visible in each box were in a straight line, parallel to the front of the box, and approximately one pellet-diameter from each other.) Then the one-way screen between the rewards and *S* was opened and *S* was allowed to lift one box and obtain the corresponding reward. Opening the one-way screen started a clock which was stopped by opening either of the boxes. The latency of response, measured in units of 0.25 sec., was recorded throughout the experiment.

There were 40 scored trials given each day for 10 days. Within each block of 10 trials, each possible comparison appeared once, in random order. On any day, two of the four different presentations of a given comparison were made with the larger amount on the right side, and two with the larger amount on the left. Two warm-up trials were presented before the 40 trials of a daily session; for these, the same rewards as were to be used on the last two trials of the day were employed.

Series II. In Series II, the boxes with pellets glued inside were used. These boxes looked very much like those of Series I, but, when *S* lifted a box, the apparent reward came away also and was unobtainable. Instead either four pellets or no pellets (placed beforehand by *E* on the portion of the gray square screened from *S* by the aluminum strips) became available. The random schedule used in Series I was used to determine which boxes (in terms of numbers of pellets) would be used, and on which sides of the testing board they would be placed; four pellets were always put under the box with the larger number of stimulus-pellets, and no pellets under the other box. Again, 40 trials (plus the warm-up trials) were given each day for 10 days.

Results. On early trials of Series I, the choices made were random with respect to the number of pellets in the boxes, but choices of the larger number of pellets increased to an asymptote in approximately 300 trials. Results for individual comparisons (*e.g.* five vs. three pellets) yielded negatively accelerated learning curves similar in form to that of the over-all results shown in the left part of Fig. 1. On the first trials of Series II, the *Ss* appeared to be surprised that they could not get the pellets which now served only as cues, but immediately found and consumed the reward-pellets when they chose the correct box. The second half of Fig. 1 shows the mean and range of accuracy of the tests in Series II. There is no evidence that the procedure in Series II led to a higher percentage of choices of the box with the larger number of pellets.

The gradual increase in mean percentage of choices of the larger reward during the course of Series I presumably reflects a gradual increase in *S*'s knowledge of the situation and of the results of comparison and choice. The mean curve masks the fact that some *Ss* exhibited more rapid increases from a near-chance to a near-perfect mode of behavior, but even in the data for these animals the transition is not abrupt. The 'errors' that remain after the animals have reached asymptotic performance in Series I are due, presumably, to a certain lack of distinctiveness between the groups of pellets being compared. The evidence from Series II leads to the conclusion that the asymptotic number and pattern of errors

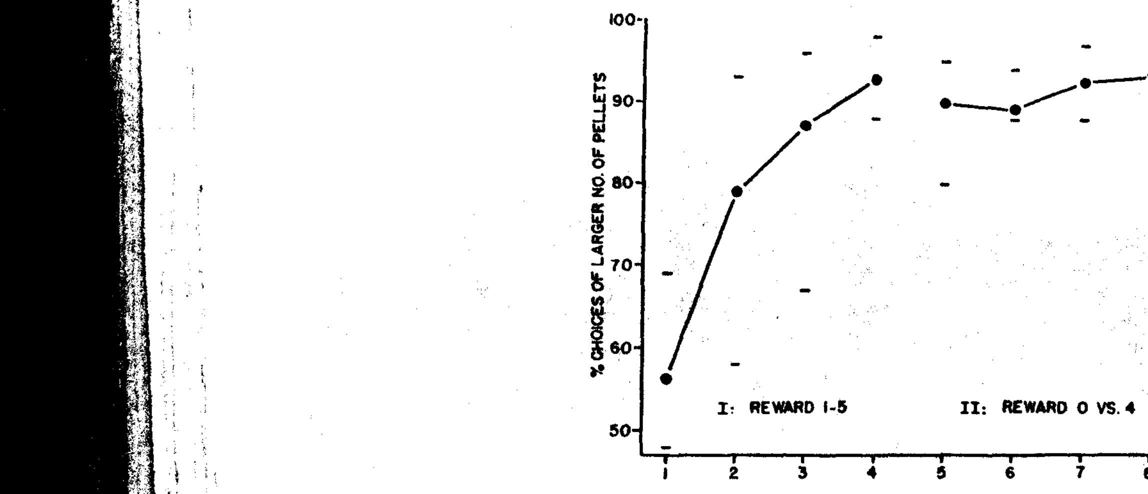

FIG. 1. CHOICES OF BOX WITH LARGER NUMBER OF PELLETS UNDER TWO REWARD-CONDITIONS

(Mean and range are plotted for successive blocks of 100 trials.)

in both series are due primarily to a lack of *perceptual* distinctiveness rather than a lack of distinctiveness in reinforcement-value. In Series II, the difference in reward between a correct and an incorrect choice always is equivalent to the maximal difference in reward between the two choices in the comparisons of Series I, yet the asymptotic performance is no higher.[3] These results cast doubt upon the

[3] The monkeys later were retested after half of them had been subjected to bilateral lesions of the inferotemporal cortex. The *Ss* received 200 trials under the procedure of Series I, followed by 200 trials under the procedure of Series II and then a final set of 100 trials under Series I. The operated animals displayed a temporary reduction in percentage of choice of the larger number of pellets in both situations; on the second and third blocks of 100 postoperative trials the inferotemporal animals were uniformly inferior to the normals. Ablation of the inferotemporal area has been shown to affect the performance of monkeys in problems of visual discrimination, but does not appear to affect (as do lesions of certain rhinencephalic areas) reactions to alterations in deprivation or magnitude of reward (W. A. Wilson, Jr., and Mortimer Mishkin, Comparison of the effects of inferotemporal and lateral

immediate applicability to animals of simple choice-methods of measuring utility.[4]

Scale-values were computed for the different numbers of pellets based upon the choices of all *Ss*, for each block of 100 trials.[5] With arbitrary values of 0 and 10 assigned to the extreme amounts (zero and five pellets, respectively), the values shown in Table I were obtained for two, three, and four pellets. On the first block of trials of Series I, the scale-values for three pellets and for four pellets are about as high or higher than the value for five pellets. After the third block, there is but little evidence for further change in scale-value. The values obtained during the final block of Series I and the final block of Series II were compared by computing

TABLE I

SCALE-VALUES OF DIFFERENT NUMBERS OF PELLETS

(Value of 0.0 assigned to one pellet, 10.0 to five pellets.)

No. of pellets	Blocks of 100 trials (Series I)				(Series II)			
	1	2	3	4	5	6	7	8
2	3.1	3.3	3.4	2.8	3.1	3.3	3.8	3.2
3	9.7	6.5	6.6	5.6	6.1	5.5	5.7	5.8
4	12.2	8.3	9.1	8.3	8.8	8.1	8.0	7.9

scale-values for individuals, and comparing the means then obtained. The differences between the values obtained in the two series were not significant for any amount of reward; nor did the difference between the scale-values for two and four pellets vary significantly from the last block of Series I to the last block of Series II. The results for each *S* on the last block of each of the two series were pooled to give asymptotic scale-values. The relation of these values to number of pellets is neither arithmetic nor logarithmic. Plotted, they seem to lie almost midway between $y = a \log x$ and $y = a\ (x\text{-}1)$, and the experimentally determined points differ significantly from both curves. The relation found by Harlow and Meyer, who used amounts of ¼, ½, 1, 2, and 4 peanuts, was logarithmic. The discrepancy may be due to differences in kind of food or to differences in the pattern of amounts.[6]

---

occipital lesions on visually-guided behavior in monkeys, *J. compar. & physiol. Psychol.*, 52, 1959, 10-17; L. Weiskrantz, Behavioral changes associated with ablation of the amygdaloid complex, unpublished doctoral dissertation, Harvard University, 1953; J. S. Schwartzbaum, Food-maintained behavior in monkeys following bilateral ablation of the amygdaloid complex, unpublished doctoral dissertation, Stanford University, 1958). Thus the deficit displayed by the operated *Ss* in this study is further evidence that the behavior of normal monkeys in such a situation is limited by perceptual or learning factors and not by reinforcement-value distinctiveness. (Beverly D. Overstreet assisted with this postoperative testing, which was done in the Psychological Laboratories of the University of California, Berkeley.)

[4] R. D. Luce and Howard Raiffa, *Games and Decisions*, 1957, 12-38.

[5] J. P. Guilford, *Psychometric Methods*, 1936, 236.

[6] This experiment does not answer the question of exactly what aspect of the stimuli was being discriminated; it may, for example, have been either numerosity, length, or volume. The situation did allow for a simpler discrimination than did that of Harlow and Meyer, who used a mixture of different portions of peanuts and different numbers of peanuts in their series.

The median latency was computed for each of the 10 comparisons, separately for each *S* in each block of 100 trials. The mean of the median latencies for each comparison in the last block of trials of Series I is plotted in Fig. 2. From *a priori* considerations, there exist several reasonable bases upon which to predict the relative latencies of the various choice-combinations as a function of the alternatives presented. Some interesting suggestions are provided by rank-difference correlations computed between the observed latencies and the ranking of latencies predicted from consideration of the total number of pellets presented, the size of

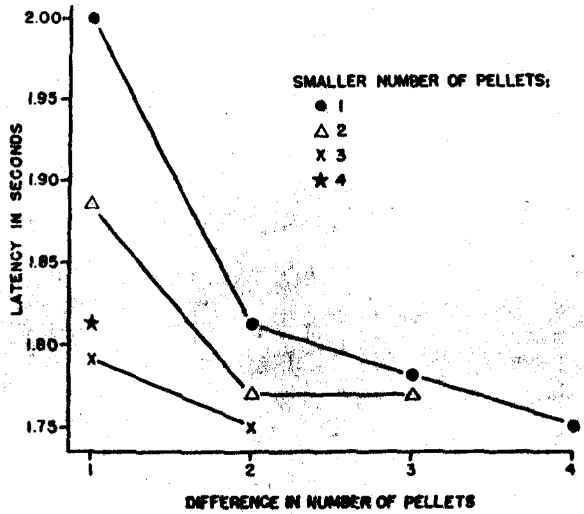

FIG. 2. LATENCY OF RESPONSE FOR EACH OF THE CHOICES IN FINAL BLOCK OF SERIES-I TRIALS

the larger number of pellets presented, the difference in size between the two groups of pellets, and combinations of these factors.

Although more parsimonious explanations might be offered, the latency-data for Series I seem to be most completely explained by invoking three factors. At asymptote, response is more rapid if the total number of pellets presented is large, or if the number of pellets in the larger reward is large. These two factors are themselves highly correlated, and exact analysis of their separate contributions is not possible, but the data are consistent with the suggestion that the total number is more important in the early trials, when responses and latencies are primarily random and the choice-situation is not well understood, and that the larger reward becomes more important at asymptote. The difference in size of the two rewards also affects latency. At asymptote, a larger difference leads to a faster response; it may be that very early in learning an opposite relation holds, since only a very large difference in reward makes the animal pause and choose carefully. Thus, after the choice-situation is well learned, latency of response can be best predicted by some negative function of the maximal reward possible, and the difference in the size of the stimuli. In Series II, the maximal reward is constant from trial to trial; thus we should not expect the larger number of pellets presented (or the total number presented) to contribute directly to determination of the latency of response. The correlational results are generally in agreement with this prediction. The difference in the number of pellets in the two boxes continues to operate, as a stimulus-factor, correlating positively with the speed of response.

It may be concluded, then, that behavior of the monkeys in this experiment may be explained in the following fashion. In the early trials, *S* learns about the nature of the choice-situation, the responses that may be made, and their various outcomes. Only when the stimuli (and the rewards) are greatly different, does he choose carefully. When the situation is well learned, all choices (except for random errors) are made of the larger reward, to the limit of *S*'s perceptual ability; decision-time is greater when the difference between the stimuli is smaller. The difference in reward does not limit the accuracy of response at asymptote; the total amount of reward in the situation (primarily in the early trials) and the larger reward in the situation (primarily after it is learned that only one reward may be obtained) do positively influence the speed of response.

Summary. Six rhesus monkeys were allowed to choose between two transparent boxes, in each of which one, two, three, four or five small glucose pellets could be seen. For the first 400 trials, *S* lifted one of the boxes and was allowed to retrieve and eat the pellets he had seen there. For the next 400 trials, the pellets that were visible in the box could not actually be obtained; instead *S* got four pellets for choosing the box with the larger number of 'stimulus-pellets' and nothing for choosing the alternative box.

The animals learned gradually to make the response that led to the larger of the two possible rewards on each trial. The choices of the larger reward were limited by the perceptual ability of the monkeys, and not by the lack of distinctiveness in value of the rewards. An effect of reward-value which *was* differential between trials may be noted: the amount of reward in a trial interacted with the perceptual factor of the relative size of the stimuli to influence the latency of response.