

Individual organism probability matching with rats in a two-choice task*

DONALD ROBBINS† and PATRICIA L. WARNER
Emory University, Atlanta, Ga. 30322

On any one trial, 10 rats had reward available for one of two alternatives, where the probability of reward for each alternative was 85:15, 70:30, or 30:70. The side on which reward was available was held constant until the rat responded to the rewarded side. First-choice trials, defined as trials following reward, revealed asymptotic probability matching in individual rats. Analysis of the error-run distributions revealed that Ss may have learned a conditional discrimination, i.e., trials following reward yielded probability matching behavior, while trials following nonreward at asymptote were largely to the side opposite the nonrewarded side.

In this study, we report a procedure that yields individual organism probability matching with rats. It should be noted that, under the standard correction or noncorrection procedure, rats tend to maximize at asymptote, i.e., confine their responses to one or the other side of the maze (e.g., see Bitterman, 1969). The procedure involves rerunning a trial if the rat makes an error, i.e., responds to the side where reward is not placed. In this error rerun procedure, the reward remains on one side for as many trials as necessary for the rat to choose that side. In some unpublished studies,¹ using a T-maze, choice behavior following rewarded trials, called first-choice trials, revealed probability matching in individual organisms. In the present study, a different apparatus from the T-maze was used and yielded relatively rapid probability matching.

METHOD

Subjects

The Ss were 10 male Sprague-Dawley rats obtained from a local supplier; they were 150 days old at the beginning of the experiment. One S was discarded after 60 trials and was replaced at this point. All of the Ss had had experience in a runway with a series of 100% acquisition and extinction trials and thus had been adjusted to a water-deprivation schedule.

Apparatus

The maze was flat white. The startbox, which was 5 in. high, 5 in. wide, and 9 in. long, permitted access to a Y-shaped maze, which was 16 in. high with 16-in.-long side walls, when an aluminum guillotine door was raised. Opposite the startbox door, 7 in. away, was a 6.5-in.-high wall. Two L-shaped goalboxes were located on top of the wall, one on each side of the maze. Each goalbox was 6 x 5 x 12 in. and contained a 3 x 5 in. recessed drinking area. A drinking tube similar to those used in the home cages was mounted on the wall, and a slide could be inserted to prevent access to the tube. The two goalboxes were separated by a 5-in.-high flat white piece of Masonite, which extended 2 in. down the wall to the floor of the

maze. Sliding aluminum doors prevented jumping down into the maze after the S had jumped to one of the ledges. The apparatus was covered with hinge-mounted clear Plexiglas. A 25-W red bulb centered above the apparatus provided illumination, and an air conditioner provided a uniform masking noise.

Procedure

The Ss were given four reinforced goalbox placements in the maze in a LRRL or RLLR order. On all subsequent trials, Ss were placed in the startbox and required to jump to one of the ledges.

During Phase I, four Ss received 470 massed trials, with reward available on one side of the maze for 85% of the trials and on the other side of the maze for the remaining 15% (an 85:15 reward schedule). Four other Ss had a 70:30 reward schedule with massed trials, three had 470 trials, and the other had 410 trials. These eight Ss had a 5-15-sec interval between trials. Two other Ss were given a 70:30 reward schedule for 300 trials with a 10-min intertrial interval (spaced). During Phase II, all Ss received 430 trials under a 30:70 schedule with the same trial spacing as in Phase I. Reward placement for each S was assigned randomly in blocks of 20, with the restriction that in each block reward be available in the appropriate ratio on one or the other side of the maze. Half of the Ss had the right-hand goalbox as the high reward proportion side, while for the remaining half reward was most frequently on the left.

At the beginning of each trial, S was placed in the startbox facing the startbox door. Three seconds later, the startbox door was raised. A response occurred when S jumped onto a ledge. After the response, the sliding aluminum door was closed to prevent retracing. On correct trials, S received 5 sec of water. On error trials, S was detained for 5 sec. Then S was placed in a waiting cage for the intertrial interval. An error rerun procedure was used. If the S chose the side not containing reward, the reward remained on the same side for as many trials as necessary until S made a correct response, i.e., chose the side of reward. As a result, Ss had 10 rewards per day, but more than 10 choices per day.

The Ss were given access to water in their home cages for 90 min each day approximately 30 min after the last S was run.

RESULTS AND DISCUSSION

The data were analyzed in terms of first-choice responses, defined as the choice following a rewarded trial, to the A_1 side, where the A_1 side was defined as the Phase I high reward proportion side.

Figure 1 presents the individual organism data in terms of the proportion of A_1 responses in blocks of first-choice trials for Phases I and II. These data showed that for each S the terminal response proportions were close to the reward proportions (π) of A_1 responses when $\pi = .70, .85,$ and $.30$. Specifically, the proportion of A_1 s during the last 200 trials in Phase I for the two Ss (Nos. 10 and 11) given a 70:30 reward schedule with a 10-min intertrial interval (spaced) was .710 and .715. For the four Ss (Nos. 1, 2, 4, and 9) receiving a 70:30 reward schedule with a 5-15-sec intertrial interval (massed), these proportions were .685, .690, .705, and .710, while for the four Ss (Nos. 5, 6, 7, and 8) given a .85-.15 reward schedule with massed trials, these

*This research was supported in part by a grant from the Emory University Grant Fund.

†Reprint requests should be sent to: Donald Robbins, Department of Psychology, Emory University, Atlanta, Ga. 30322.

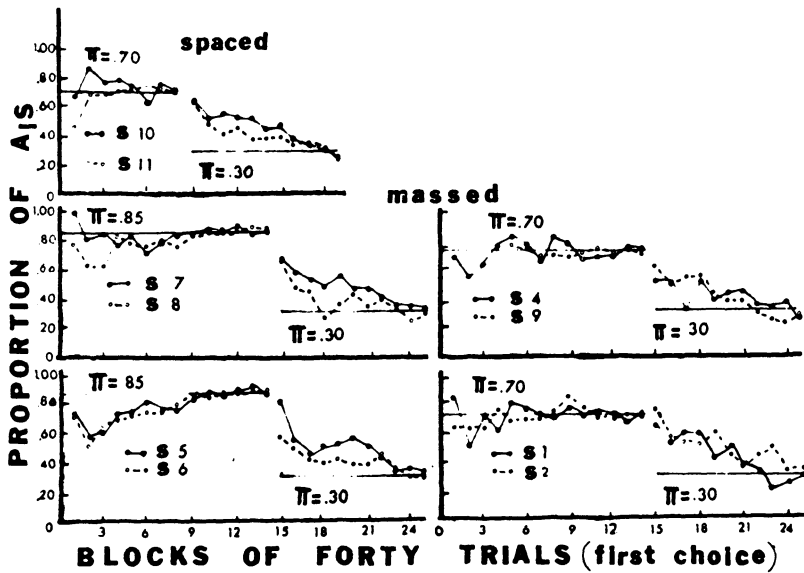


Fig. 1. Proportions of A_1 s in blocks of 40 first-choice trials for individual Ss in Phases I and II.

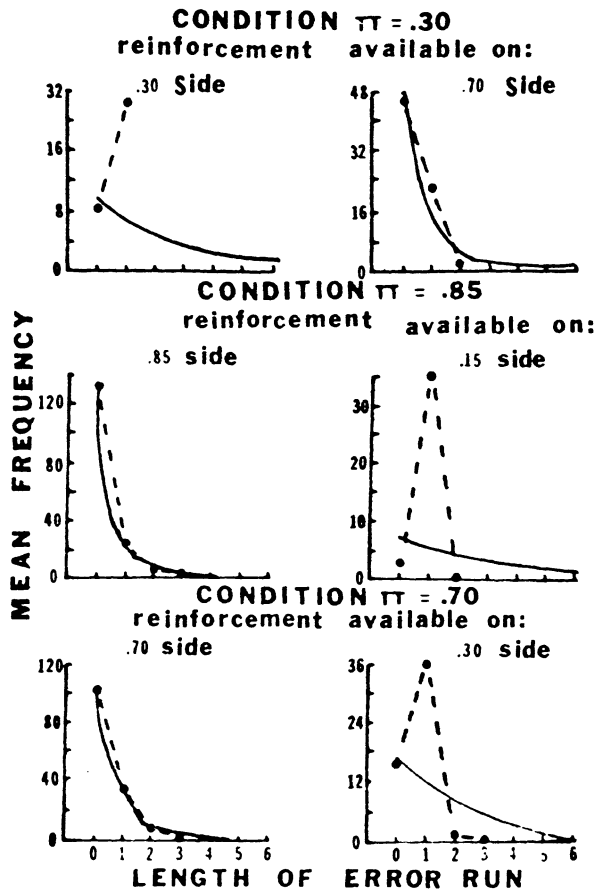


Fig. 2. Mean frequency of error runs for the various reinforcement schedules.

proportions were .840, .850, .855, and .865. During the final 100 trials for the .30-.70 schedule in Phase II, the Ss given a 70:30 reward schedule in Phase I responded with .24 and .31 proportion A_1 s for the two spaced-trial

Ss, and for the four massed-trial Ss, these proportions were .25, .31, .35, and .39. The four Ss given an 85:15 schedule in Phase I (massed) responded with .29, .30, .31, and .39 proportion A_1 s during the final 100 trials under the Phase II 30:70 reward schedule. Thus, all Ss responded at or near the reward proportion in terms of the proportion of A_1 s during the terminal blocks of first-choice trials, i.e., trials immediately following reward.

The major difference between the standard noncorrection and the rerun procedure is that rerunning error trials guarantees that reward will be obtained in a $\pi:1 - \pi$ ratio to one side or the other side of the maze, while for the standard noncorrection procedure the obtained reward schedule is dependent on S's choice behavior. As a result, if we assume, as for the standard noncorrection procedure, that the error trials have no effect, i.e. an identity operator (Weinstock, North, Brody, & LoGuidice, 1965) approximation functions on error trials, the distribution of error runs should be geometric. Since there were no obvious differences between Ss, the individual S data were grouped for these analyses. Figure 2 presents the observed and expected error runs of length 0 to 6 when reward (or reinforcement) was available on the high- and low-proportion side of the maze, and is shown separately for the .70-.30, .85-.15, and .30-.70 schedules. Analysis failed to find an identity operator approximation on these error trials. The critical tests and major departures can be seen to be on the low-proportion side. Specifically, Ss made at least three times as many error runs of Length 1 than were expected when reward was available on the low-proportion side and fewer longer runs than were expected. The inability to find an identity operator operating on error trials suggests that Ss may have learned a conditional discrimination. No water may have been a cue to respond to the other side on the next

Trial; if water was obtained, then S responded according to the ratio of the proportion of conditioned elements to the A₁ and A₂ side. This is plausible in light of the demonstration that rats can learn to reverse successive spatial discriminations with one error (e.g., Dufort, Guttman, & Kimble, 1954). The failure to obtain an identity operator functioning on error trials, however, does not weaken the finding of individual organism matching on first-choice trials.

REFERENCES

Bitterman, M. E. Habit-reversal and probability learning: Rats, birds, and fish. In R. M. Gilbert and N. S. Sutherland (Eds.), *Animal discrimination learning*. New York: Academic Press, 1969.

Dufort, R. H., Guttman, N., & Kimble, G. A. One-trial discrimination reversal in the white rat. *Journal of Comparative & Physiological Psychology*, 1954, 47, 248-249.

Weinstock, S., North, A. J., Brody, A. L., & LoGuidice, J. Probability learning in the I-maze with noncorrection. *Journal of Comparative & Physiological Psychology*, 1965, 60, 76-81.

NOTE

1. Robbins, D., Cohen, S., Levine, M. E., and Olsen, A. Probability learning in the T-maze with an error rerun procedure, unpublished study. Weinstock, S., Robbins, D., and Chen, W. Successive brightness discrimination and stimulus generalization in the T-maze. Unpublished study.

(Received for publication August 31, 1973.)

The effect of two types of baseline training on behavioral contrast and the peak shift*

ABDULAZIZ A. DUKHAYYIL and JOSEPH E. LYONS

University of Arizona, Tucson, Ariz. 85721

Pigeon Ss were given discrimination training between two wavelength stimuli following either a single stimulus or a nondifferential baseline training condition. The birds were then administered a postdiscrimination generalization test. This was followed by a return to the baseline training condition. Two levels of task difficulty were used. The results indicated that neither behavioral contrast nor the peak shift vary systematically as a function of either baseline training condition or task difficulty. The results further show the use of the postdiscrimination baseline training condition to be a weak procedure.

The purpose of establishing a baseline response rate is to provide a stable criterion against which the effects of discrimination training may be judged. Typically, baseline training has been of two types. In the first, single stimulus training to the stimulus correlated with reinforcement (S+) provides the baseline response rate, while, in the second, the Ss are pretrained not only in the presence of S+ but also in the presence of the stimulus to be correlated with extinction (S-). Both baseline training conditions seem to provide a stable response rate provided the baseline training period is 20 to 25 days long (cf. Sadowsky, 1973).

Discrimination training following either baseline condition usually yields positive behavioral contrast, i.e., an increase in the response rate in S+ over the baseline rate as a function of the decrease in the response rate and/or reinforcement density in S-. Unfortunately, little work has been done to determine if, indeed, both the single stimulus and the nondifferential baseline

training conditions have equivalent effects on behavioral contrast. Bloomfield (1969) has compared behavioral contrast following single stimulus baseline training with behavioral contrast following nondifferential baseline training, and concluded that behavioral contrast does not vary as a function of baseline training condition. However, the animals Bloomfield (1969) compared were from two different laboratories, making firm conclusions about the effect of the baseline training conditions tenuous at best. Therefore, the major purpose of the present study was to determine whether behavioral contrast varies as a function of the baseline training condition. A second purpose of this work was to ascertain the effect of discrimination task difficulty (defined here as the physical distance between S+ and S-) on behavioral contrast. Task difficulty has been shown to be related to the position and magnitude of the peak shift (Hanson, 1959), and the peak shift is thought to be related to the occurrence of behavioral contrast (Terrace, 1972). Therefore, another aim of this study was to determine the effect baseline training condition and task difficulty have on the postdiscrimination generalization gradient (i.e., the peak shift).

In some cases, the baseline training condition has followed rather than preceded discrimination training (cf. Terrace, 1964; Weisman & Ramsden, 1973). The present study also attempted to determine the validity of the use of this procedure as a baseline for ascertaining the presence or absence of behavioral contrast during discrimination training.

METHOD

Subjects

The Ss were eight experimentally naive White and Silver King pigeons maintained at 70%-75% of their free feeding weights throughout the experiment.

*Requests for reprints should be sent to Joseph E. Lyons, Department of Psychology, University of Arizona, Tucson, Ariz. 85721.