

Facilitating stimulus effects of reward and punishment in discrimination learning

H. FOWLER, M. HOCHHAUSER, and G. J. WISCHNER
University of Pittsburgh, Pittsburgh, Pennsylvania 15260

To demonstrate a facilitating stimulus effect, as opposed to an incentive effect, of food reward, rats were trained on an easy, light-dark discrimination with different amounts of reward for correct and incorrect responses (1-0, 2-0, 3-1, and 5-1 pellets, respectively), and with shock or no shock administered in the correct goalbox. Both errors and trials to criterion were fewer with a large reward differential (LRD: 2-0 and 5-1), as compared with a small reward differential (SRD: 1-0 and 3-1), but were not affected by the "base" reinforcement condition of either 1 or 0 pellets for the incorrect response. In addition, choice and arm speeds during early training were positively related to the combined, or average, number of pellets contingent upon both correct and incorrect responses, indicating a generalization of reward expectancies. Although shock uniformly suppressed arm speeds under all reward conditions, it facilitated discrimination learning in the SRD conditions. That such facilitation occurred only when the conditions of reward for correct and incorrect responses were relatively similar indicates that not only shock, but also food can function as a distinctive cue: As a stimulus selectively applied to one response, it can decrease the similarity of the alternatives, and, in this manner, it can facilitate performance.

Studies of the facilitating effect of mild shock punishment for the food-rewarded response in visual discrimination learning (see Fowler, 1971; Fowler & Wischner, 1969) have indicated that shock functions as a highly discernible or "distinctive" cue: It reduces the similarity of the discriminative-stimulus compounds constituting the response alternatives, and thus it facilitates performance by reducing the generalization of reward and nonreward expectancies between the alternatives. Consistent with this interpretation, the "shock-right" facilitation effect is typically absent in an easy discrimination in which the discriminative stimuli (e.g., light-dark) are highly dissimilar and hence preclude a distinctive-cue function of the shock (e.g., Wischner & Fowler, 1964; Wischner, Fowler, & Kushnick, 1963); in contrast, the facilitation effect is consistently observed in more difficult (e.g., bright-dim) discriminations in which the discriminative stimuli are similar and thus potentiate the shock's cue function (e.g., Fowler, Spelt, & Wischner, 1967; Fowler & Wischner, 1965). Furthermore, if the aversiveness of the shock is reduced through the administration of sodium amytal (Fowler, Goldman, & Wischner, 1968), the facilitation occurring in a difficult discrimination is, within limits, an

increasing S-shaped function of shock intensity, consistent with the Weber principle relating performance to the discriminable cue properties (e.g., the intensity) of a stimulus. The same relationship between a facilitating effect and the intensity of a stimulus has also been observed for a neutral white-noise cue used in place of shock for the correct response (Fago & Fowler, 1972).

The fact that either an aversive or a neutral stimulus for one response can facilitate discrimination learning suggests that the same effect and principle of operation should occur for any stimulus. As such, the facilitating (cf. "reinforcing") effect of food reward can be reinterpreted as due, in part, to the operation of food as a distinctive cue. Like shock or noise, food reward for one response should decrease the similarity of the stimulus alternatives, especially as the magnitude (cf. intensity) of the reward increases. Inasmuch as this interpretation argues that the stimulus parameters of reward influence the similarity of the alternatives, it follows that the effective difficulty of a discrimination can be regulated both by the similarity of the discriminative stimuli and by the similarity of the conditions of reward for correct and incorrect responses. In other words, a more difficult discrimination task should result either when the discriminative stimuli are similar and the conditions of reward for correct and incorrect responses are dissimilar (e.g., moderate vs. no reward, as is typically the case) or when the discriminative stimuli are dissimilar and the conditions of reward are similar (e.g., moderate vs. small reward). Furthermore, if the latter operation increases task difficulty by increasing

This study was supported in part by Grant MH-08482 from the National Institute of Mental Health, United States Public Health Service. Requests for reprints should be sent to Harry Fowler, Department of Psychology, University of Pittsburgh, Pittsburgh, Pennsylvania 15260. M. Hochhauser is now at the Division of School Health Education, University of Minnesota, Minneapolis, Minnesota 55455. G. J. Wischner is now at the Department of Psychology, Cleveland State University, Cleveland, Ohio 44115.

generalization effects between the alternatives (rather than by reducing their incentive difference), then it should be possible to demonstrate a facilitating cue effect of shock-right training in this "easy" (i.e., light-dark) discrimination in which the effect has typically been absent. In turn, such an outcome would highlight the facilitating cue property of selective reward (i.e., administered for only one response) because the functioning of shock as a distinctive cue in an "easy," nonselective-reward discrimination can only derive from the similarity of the conditions of reward. (Conversely, selective reward in an "easy" discrimination would render the alternatives more discriminable and would therefore offset a distinctive-cue effect of the shock, as has been observed.)

Following the above rationale, Fowler, Fago, and Wischner (1971) manipulated the effective difficulty of a light-dark discrimination by varying the delay of reward for the correct response in conjunction with no reward (cf. an "infinite" delay) for the incorrect response. With reward considerably delayed (i.e., 16 or 24 sec), shock administered immediately in the correct goalbox facilitated discrimination learning, but, with food reward administered immediately or at a short (8-sec) delay, immediate shock had a retarding effect on discrimination performance. Although these results are consistent with a cue interpretation of the effects of reward, they are open to an alternative interpretation based on secondary-reinforcing effects of shock (cf. Fowler, 1971; Fowler & Wischner, 1969). It is possible that shock at the goal elicited stereotyped reactions, such as cringing, which mediated the effects of food reward across the delay and, in this manner, facilitated performance. To eliminate this possibility, the present study administered shock in conjunction with immediate food reward in the correct goalbox, but manipulated the effective difficulty of the light-dark discrimination by varying the magnitude of food reward for both correct and incorrect responses. For different groups of rats, including no-shock controls, food reward in the correct and incorrect goals was set at 1-0, 2-0, 3-1, and 5-1 pellets, respectively. These treatments afforded a comparison of both selective (1-0, 2-0) and nonselective (3-1, 5-1) reward conditions and of small (1-0, 3-1) and large (2-1, 5-1) reward differentials within each set.

METHOD

Subjects

The subjects were 64 naive male albino rats of the Sprague-Dawley strain, 80 to 90 days old at the start of the experiment. They were caged individually in the experimental room under controlled temperature (20°-22°C) and a reversed day-night (12-h) cycle.

Apparatus

The discrimination-training apparatus was an enclosed T-maze made of opposing L-shaped strips of galvanized sheet metal, sup-

ported by an exterior wooden shell (for details, see Fowler, Fago, & Wischner, 1971). Guillotine doors were located 20.3 cm from the end of the stem to form a start compartment, at the entrance to each arm, in order to permit forcing and to prevent retracing, and 30.5 cm from the end of each arm to form goalboxes (total arm length was 91.4 cm). Each goalbox had a food well, which was floor-recessed and not visible from the arm, and an end wall of frosted Plexiglas. The end walls were differentially illuminated from behind by a 5-W incandescent bulb (0 or 120 V) to provide the light-dark discriminative stimuli.

The two L-shaped strips of sheet metal forming each maze section were connected in series with a .2-M Ω resistor across the output of a transformer (60 Hz ac). A single 60-V .2-sec shock (.3 mA) could be delivered to the subject when it interrupted an infrared photoelectric beam crossing the goal at a point 5.1 cm in front of the food cup. Another infrared photobeam, located 33.0 cm from the center of the choice point in each arm, was used to measure the subject's latency of choice from the opening of the startbox door and, in conjunction with the goalbox photobeam, to measure running latency in each arm.

Procedure

The procedure included both nondifferential-pretraining and discrimination-training phases. One week prior to pretraining, the subjects were started on a daily diet of 11 g of Wagne Lab Blox, with water available ad lib. (This diet reduced the subject's weight to about 80% of predeprivation weight.) Pretraining was designed to habituate the subjects to the apparatus and to reduce possible position and brightness preferences. Each subject received a total of 16 forced-choice food-reinforced trials, administered 4/day at an intertrial interval of about 7 min. Daily forced-choice trials were distributed randomly with the restriction that they were balanced over left and right positions and light and dark goalbox cues. Forcing was accomplished by lowering the door to the inappetent goalbox. The pellets consisted of Noyes Formula A rat chow (1.25 g). The pellets were spread liberally throughout a goalbox on Day 1 of pretraining and then were reduced systematically until, on Day 4, only 1 pellet/trial was provided in the food cup. For this and the succeeding phase of training, the subject's daily diet was corrected for the amount of food received during a day's trials.

Discrimination training began on the day following pretraining and consisted of free-choice trials. The subjects were assigned randomly to eight groups of eight subjects each, comprising a 2 by 2 by 2 factorial design: no-shock (NS) or shock-right (SR) training, that is, shock administered in the correct, large-reward goalbox; a "base" reinforcement condition of either 1 or 0 pellets for each incorrect response; and a small or large reward differential between correct and incorrect responses. Specifically, the differential-reward conditions were 1-0, 2-0, 3-1, and 5-1 pellets for each correct and incorrect response, respectively. Thus, within each base-reinforcement condition, there was a small reward differential for correct and incorrect responses (SRD: 1-0 and 3-1) and a relatively large reward differential for the responses (LRD: 2-0 and 5-1). With regard to the stimulus properties of reward, LRD represented an easy (highly discriminable) reward condition for which SR training would not be expected to facilitate performance; in comparison, SRD represented a more difficult (less discriminable) reward condition for which SR training could facilitate performance via the shock's function as a distinctive cue.

All subjects received 4 trials/day for the first 6 days of discrimination training and then 8 trials/day thereafter. Subjects were run until they met a criterion of 15 correct responses out of 16 consecutive choices, with the last 8 correct. To avoid the rapid learning and possible "floor" effects associated with training to a dark discriminative stimulus, the larger food reward was administered for all groups in the lighted goalbox, the right-left positioning of which varied according to a Gellerman sequence. On any training trial, time in the correct or incorrect goal was 10 sec, and the intertrial interval was about 7 min.

RESULTS

Errors and Trials to Criterion

Figure 1 presents group mean errors in blocks of 20 trials for each of the four reward conditions. (NS and SR subgroups have been pooled to reduce the cluster of curves.) As indicated, the LRD groups (2-0 and 5-1) exhibited fewer errors [$F(1,56) = 9.14$, $p < .005$] and trials to criterion [$F(1,56) = 10.10$, $p < .005$] than did the SRD groups (1-0 and 3-1). Figure 1 also suggests that errors and trials to criterion were somewhat fewer for the 0-pellet base groups (2-0 and 1-0) as compared with the 1-pellet base groups (5-1 and 3-1); however, for both measures, the differences were nonsignificant ($ps > .10$). Also, for both measures, the interaction of the differential and base-reinforcement variables was nonsignificant ($ps > .20$). In terms of errors and trials to criterion, then, the SRD conditions (1-0 and 3-1) produced comparably difficult discrimination tasks, and the LRD conditions (2-0 and 5-1) produced relatively easy tasks.

Mean errors to criterion for the NS and SR subgroups are presented as a function of reward differential in Figure 2; the left panel shows the data for the 0-pellet base conditions, and the right panel shows the data for the 1-pellet base conditions. Figure 2 shows that all SR groups had fewer errors than their respective NS controls, but that the differences were more pronounced in the SRD (1-0 and 3-1) conditions. However, the F-test results showed only an overall SR facilitation effect [$F(1,56) = 4.31$, $p < .05$]; the interaction of shock with either the differential or base-reinforcement variable, or both, was unreliable ($ps > .10$). To further assess the relationship between shock and reward differential, NS and SR differences were analyzed separately for the SRD conditions and

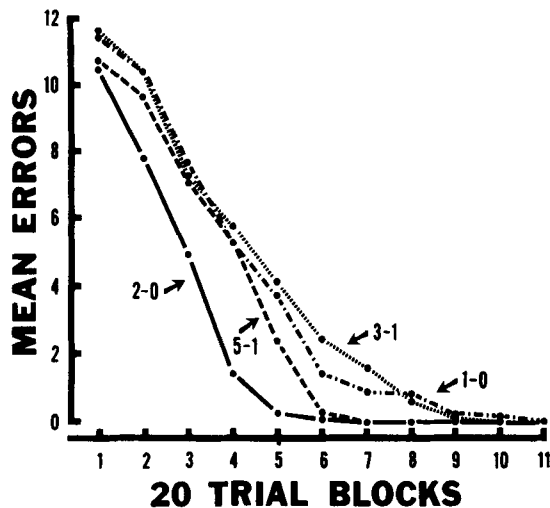


Figure 1. Mean errors in blocks of 20 trials for the differential reward groups. (NS and SR subgroups have been pooled to reduce the cluster of curves.)

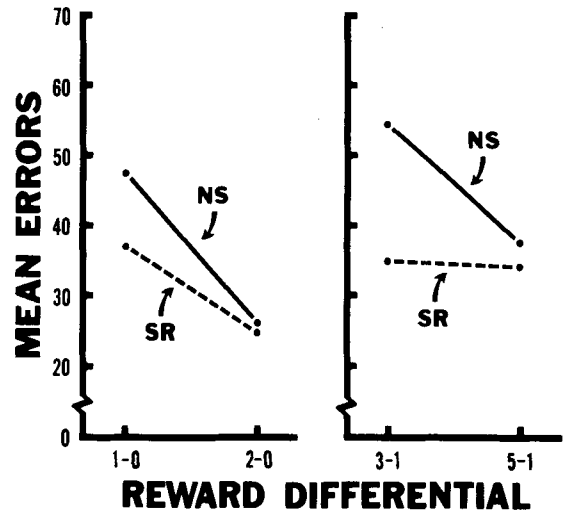


Figure 2. Mean errors to criterion for NS and SR subgroups as a function of reward differential; the left panel presents the data for the 0-pellet base conditions, and the right panel presents those for the 1-pellet base conditions.

for the LRD conditions. Evaluated on the basis of the overall error term, these comparisons showed that the SR facilitation effect was reliable for the SRD conditions [$F(1,56) = 6.16$, $p < .025$], but was clearly not reliable for the LRD conditions ($F < 1$). Furthermore, a Duncan range test assessing differences between respective NS and SR groups indicated that SR facilitation was significant ($p < .05$) only in the 3-1 condition; in the 1-0 condition, the effect approached significance ($p < .10$). Virtually identical results obtained with a measure of trials to criterion.

Arm and Choice Speeds

Median arm and choice latencies were tabulated for each subject over successive blocks of 5 trials and were then transformed to speed scores (10/latency in sec). Mean arm speeds for the NS and SR subgroups are presented as a function of reward differential in Figure 3; the left panel shows the data for the 0-pellet base conditions, and the right panel shows the data for the 1-pellet base conditions. These speed data were restricted to the initial 25 discrimination trials in order to illustrate early training effects and to avoid empty data cells for animals that quickly met the learning criterion and were subsequently removed from training. Also, the data are based on both correct and incorrect responses because analyses showed that neither arm nor choice speeds were reliably differentiated by the correctness of the subjects' responses during these early trials (cf. Figure 1, which shows chance performance during these trials).

Figure 3 shows that SR training uniformly suppressed arm speeds under all reward conditions [$F(1,56) = 4.46$, $p < .05$] and did not interact with any of the reinforcement variables ($Fs < 1$). In addition, the 1-pellet base groups were faster than the 0-pellet

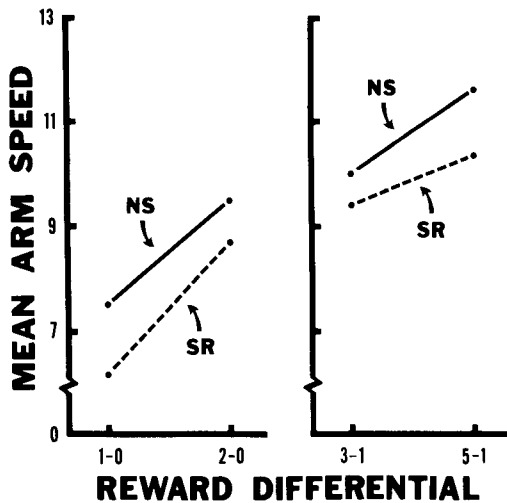


Figure 3. Mean arm speeds over the initial 25 discrimination trials for NS and SR subgroups as a function of reward differential; the left panel presents the data for the 0-pellet base conditions, and the right panel presents those for the 1-pellet base conditions.

base groups [$F(1,56) = 26.77, p < .001$], and the LRD (5-1 and 2-0) groups were faster than the SRD (3-1 and 1-0) groups [$F(1,56) = 13.88, p < .001$]. However, the interaction of these variables was negligible ($F < 1$). As such, the arm-speed data were systematically ordered from 1-0 through 2-0 and 3-1 to 5-1, and could best be described as a positive function of the combined (or average) number of pellets contingent upon the subject's running response to both correct and incorrect goals. Trend analysis by polynomial contrasts showed that the linear component of this relationship was highly reliable [$F(1,56) = 38.96, p < .001$] and that residual-trend variance was nonsignificant ($p > .20$).

Analysis of choice speeds showed virtually the same results as arm speeds, with the exception that there were no significant shock effects, as would be expected given the greater temporal disparity between choice and shock at the goal.

DISCUSSION

The speed data are informative because, in line with a stimulus interpretation of reward effects, they indicate that running speed during early discrimination training was largely regulated by a generalization of incentives, or reward expectancies, between the alternatives (cf. Bower, 1961; Logan & Wagner, 1965). There are three aspects of the data that highlight this effect. First, arm speeds during the early trials were reliably faster in ascending order for the 1-0, 2-0, 3-1, and 5-1 groups, and could best be described as a positive function of the average number of pellets contingent upon both correct and incorrect responses. Second, during these trials, arm speeds were not reliably differentiated by the correctness of

the subjects' responses and indicated, therefore, that such performance was a product of the combined effects of correct and incorrect-response outcomes. Third, choice speeds, which could *only* be influenced by the combined and, for the most part, equally frequent effects of correct and incorrect responses during the early trials (i.e., when choice was still within the bounds of chance; see Figure 1), showed virtually the same outcome as arm speeds.

With the easy light-dark discrimination that was used, a generalization of reward-expectancies would not be mediated by the highly dissimilar discriminative stimuli, but rather by "background" cues that are common to the alternatives, that is, physical features of the arms, interoceptive or "drive" stimuli, feedback from the instrumental running response, and, as argued by the present study, stimulus similarities of the reward conditions for correct and incorrect responses. When the reward conditions are relatively similar (as in the SRD conditions), there should be enhanced generalization effects between the alternatives, with the result that correct and incorrect responses are less subject to the selective effects of reward in each alternative, and hence are less well differentiated over the course of training. Thus, in line with the associative data on errors and trials to criterion, the present interpretation argues that discrimination learning will be slower under the SRD (1-0 and 3-1) conditions than under the LRD (2-0 and 5-1) conditions, but not as a result of a smaller difference in reward magnitude between the alternatives. The fact that the difference in reward magnitude between the alternatives was identical in the 2-0 and 3-1 conditions, and yet learning was far superior in the 2-0 condition (cf. Figures 1 and 2), strongly argues that discrimination performance was regulated by the extent to which the reward conditions were *discriminably* different. Also in line with the present data, a stimulus interpretation of reward effects argues that the rate of discrimination learning will be independent of the base reinforcement condition employed, provided there are equal psychological differences in reward magnitude between the alternatives. This was essentially the case in the present study: By comparison with respective SRD and LRD groups of the 0-pellet base condition, those of the 1-pellet base condition represented a log increase in the *difference* in reward magnitude between the alternatives (i.e., from 1 to 2 pellets for the 1-0 and 3-1 groups, and from 2 to 4 pellets for the 2-0 and 5-1 groups). In short, the progression in reward differential across the 0- and 1-pellet base conditions was consistent with the Weber function on the discriminability of differences.

In discrimination-training contexts in which a generalization of reward-expectancy is enhanced, shock for the correct response should facilitate learning because, despite its aversive property (as shown

by the suppression of arm speeds; cf. Figure 3), it can function as a distinctive cue to reduce the similarity of the stimulus compounds comprising the correct and incorrect alternatives. Consequently, when similarity of the alternatives is promoted by a small reward differential (particularly one entailing consummatory activity and related feedback in both alternatives, e.g., as in the 3-1 condition), shock should render the correct alternative more discriminable from the incorrect alternative, with the result that generalization effects between the alternatives are reduced and learning is facilitated. On the other hand, when similarity of the alternatives is reduced through the use of a relatively large reward differential (e.g., 5-1 and 2-0), a facilitating cue effect of the shock should also be reduced, and now possibly offset, by the shock's aversive or avoidance-producing property. Aversiveness of the shock can also reduce its facilitating cue effect in contexts in which low discriminability of the alternatives is produced by a small reward differential entailing minimal reward for the correct response (e.g., 1-0). Pitted against 1-pellet reward, shock should more readily promote avoidance of the correct alternative, and thus it should generate less facilitation in this condition, as compared with one entailing a larger reward for the correct alternative and yet the same psychological difference between the alternatives (i.e., 3-1). The present interpretation accords well with the obtained findings: SR facilitation was pronounced in the 3-1 condition, marginally present in the 1-0 condition, and absent in the LRD conditions (2-0 and 5-1).

In prior research, a facilitating effect of SR training has typically been restricted to a difficult bright-dim discrimination, that is, when similarity of the stimulus alternatives is accomplished by manipulating the discriminative stimuli (see Fowler, 1971; Fowler & Wischner, 1969). By showing that SR facilitation can occur in an easy, light-dark discrimination under conditions entailing a small, but not a large, reward differential between the alternatives, the present findings strongly argue that food reward itself can function as a distinctive cue. That is to say, for example, in selective-reward training, a moderate or large reward will itself render the alternatives more discriminable (thereby precluding a facilitating cue effect of shock, as in the 2-0 condition). This conclusion holds particular significance because it argues that much of what has been attributed to the effect of selective reward may well relate to its value as a discernible stimulus rather than to its value as an in-

centive. Indeed, based on its discriminable stimulus properties, not only larger amounts, but more immediate and more frequent presentations of food reward for one response should better serve to reduce the similarity of the alternatives and, in this manner, facilitate (cf. "reinforce") discrimination performance. With this view, the effects of the parameters of reward (and of punishment) can be cast as part of the broad context in which stimulus discrimination operates (cf. Fowler, in press; Hulse, 1978).

REFERENCES

- BOWER, G. H. A contrast effect in differential conditioning. *Journal of Experimental Psychology*, 1961, **62**, 196-199.
- FAGO, G. C., & FOWLER, H. Facilitated discrimination learning as effected by response-contingent neutral and aversive stimuli. *Learning and Motivation*, 1972, **3**, 20-30.
- FOWLER, H. Suppression and facilitation by response contingent shock. In F. R. Brush (Ed.), *Aversive conditioning and learning*. New York: Academic Press, 1971.
- FOWLER, H. Facilitating stimulus effects of reward and punishment: Discriminability as a general principle. In D. A. Routh (Ed.), *Learning, speech, and the complex effects of punishment*. New York: Plenum Press, 1981, in press.
- FOWLER, H., FAGO, G. C., & WISCHNER, G. J. Shock-right facilitation in an easy noncorrection problem as effected by delay of reward. *Learning and Motivation*, 1971, **2**, 235-245.
- FOWLER, H., GOLDMAN, L., & WISCHNER, G. J. Sodium amygdala and the shock-right intensity function for visual discrimination learning. *Journal of Comparative and Physiological Psychology*, 1968, **65**, 155-159.
- FOWLER, H., SPELT, P. F., & WISCHNER, G. J. Discrimination performance as affected by training procedure, problem difficulty, and shock for the correct response. *Journal of Experimental Psychology*, 1967, **75**, 432-436.
- FOWLER, H., & WISCHNER, G. J. Discrimination performance as affected by problem difficulty and shock for either the correct or incorrect response. *Journal of Experimental Psychology*, 1965, **69**, 413-418.
- FOWLER, H., & WISCHNER, G. J. The varied functions of punishment in discrimination learning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior*. New York: Appleton-Century-Crofts, 1969.
- HULSE, S. H. Cognitive structure and serial pattern learning by animals. In S. H. Hulse, H. Fowler, & W. K. Honig (Eds.), *Cognitive processes in animal behavior*. Hillsdale, N.J.: Erlbaum, 1978.
- LOGAN, F. A., & WAGNER, A. R. *Reward and punishment*. Boston: Allyn & Bacon, 1965.
- WISCHNER, G. J., & FOWLER, H. Discrimination performance as affected by duration of shock for either the correct or incorrect response. *Psychonomic Science*, 1964, **1**, 239-240.
- WISCHNER, G. J., FOWLER, H., & KUSHNICK, S. A. Effect of strength of punishment for "correct" and "incorrect" responses on visual discrimination performance. *Journal of Experimental Psychology*, 1963, **65**, 131-138.

(Received for publication March 31, 1980;
revision accepted August 19, 1980.)