

Same-different judgments of size and of weight in children: Does reward make a difference?

FRANK S. MURRAY and ESTHER R. MORRISON
Randolph-Macon Woman's College, Lynchburg, Virginia 24503

The effect of reward on same-different judgments of size and of weight was investigated. Eighty-one children (ages 4 to 6 years) were (1) rewarded for selecting and inserting a black disk into a box, (2) rewarded for choosing a white disk in the same task, or (3) not rewarded while participating in the task. The children were then required to report "same" or "different" in size judgments of black, white, and gray disks presented for comparison. Same-different judgments of sets of weights of black, white, and gray brightnesses were investigated as well. It was found that reward did not measurably affect the children's judgments of size or weight. A developmental trend was manifested in size judgments but not in estimates of weight.

A number of studies have demonstrated that stimuli with which rewards have been associated tend to be dominant or enhanced in the individual's perceptual field (Lambert & Lambert, 1953; Lambert, Solomon, & Watson, 1949; Proshansky & Murphy, 1942; Schafer & Murphy, 1943; Snyder & Snyder, 1956; Sommer, 1957; Sommer & Ayllon, 1956). The reference experiment on the effects of rewards on size estimation by children was performed by Lambert et al. (1949). Two groups of children, ages 3-5 years, were rewarded for turning a crank handle. The control group received candy directly after turning the crank; for the same number of turns, the experimental group received a white poker chip which could be inserted into a slot for candy. Size of the poker chip was significantly overestimated by the experimental group and not by the control group. Similar results were reported by Lambert and Lambert (1953).

In both of these studies, however, only one size of token was used, and all judgments were made by turning a knob controlling the diameter of a circular patch of light. Additionally, several experiments (Murray & Garrett, 1976) conducted in our laboratory over the last 5 years (using different stimuli, responses, and children of the same age) have failed to obtain any enhancement of response through reward. If such an effect exists, it is very difficult to obtain. It is our impression that many of the findings of the directive-state theorists (advocates of motivational determinants on perception) are attributable to methodological inadequacies, such as demand characteristics (Orne, 1962), experimenter's bias (Rosenthal, 1966), and general lack of sensitivity in the designs of the experiments to rule out alternative rival hypotheses.

It was the purpose of the present investigation to control for some of these methodological difficulties, and to design our experiments to answer some of the following questions: What are the effects of reward on size estimations of objects of different sizes? Does rewarding of a colored object affect subsequent judgments of objects of that color or of different colors? More specifically, are objects, the colors of which have

been associated with reward, judged the same or different in sizes or weights as the nonrewarded ones?

The design of our experiment was such that there were two reward conditions and a control condition. Instead of having only one experimental group, two groups, each rewarded on different brightness values, were used. In this way a more sensitive measure and the differential effects of reward on the different brightness values could be ascertained. The disks used in the present experiment were presented successively so as to augment the response biases to the rewarded brightnesses. Additionally, the reward training was constructed so that the attribute of the disk associated with reward could be verbalized by the child as well as illustrated by him in a free-choice task. Thus, both the effectiveness of the reward training and the child's understanding of its nature could be assessed. A pretest-posttest design with neutral colored disks was used to assay the effects of the reward training on size judgments. In addition, the effect of this training on same-different judgments of weights was also examined. Specifically, would the effect of reward training transfer from size to weight? Finally, to increase the overall sensitivity of the design to detect statistical significance, a large number of children was used. These children were also selected from three age groups to determine if any developmental differences in judgment exist.

SIZE

Method

Subjects. Ninety-two children (43 boys and 49 girls) aged 4, 5, and 6 years participated in this study. Of the total number of children, 11 failed to meet the requirements of the pretraining task and were not selected to continue the study. Thus, there were 81 children (38 boys and 43 girls) who completed the task. The mean ages of the 4-, 5-, and 6-year-olds were 4.7, 5.5, and 6.5 years, respectively. Of the 81 participants, 50 were students at the kindergarten or first-grade level at Villa Maria Academy, and 31 were enrolled at the First Presbyterian Nursery School. Both institutions are located in Lynchburg, Virginia.

Procedure: Pretraining period. The purpose of the pretraining period was to allow the child to become familiar with the procedure and to assess his understanding of the concept of same and different with respect to size.

The materials used in this task were sets of two geometric shapes: a triangle and a square. The sets were composed of three sizes that were classified as small, medium, and large. All the shapes were the same color of blue in order to prevent the

effects of color from being confounded with those of size in the task. The center points of each shape contained a circular hole so the shape could be displayed for inspection and judgment by mounting it on a dowel which protruded from a board. An unpainted plywood board 28 x 28 in. (72.12 x 72.12 cm) with a dowel 3 in. (7.62 cm) in length was used for this purpose.

The child was told by the experimenter that he had been selected to participate in a game that had to do with things that were the same and with things that were different in size. This game was called the "same-size/different-size" game and the experimenter proceeded with an illustration of how it was to be played. Beginning with the triangles, the experimenter placed a small, medium, and a large triangle randomly on the table and asked the child to report, either by picking up or by pointing to the triangle, which one was the largest, the medium, and the smallest in size. Then the experimenter asked him if the three triangles were the same size or different in size. The same procedure was repeated with the set of squares.

The experimenter then informed the child that he would play another "same-size/different-size" game with the triangles and with the squares. The child was requested to sit in a chair placed 1 m from and facing the board. The experimenter explained that she would show the child a triangle by placing it on the board (experimenter illustrated by doing so) for him to inspect for 5 sec. Then she would replace the first triangle with another triangle (experimenter removed the first triangle and placed another triangle on the dowel) and would ask him to report whether the second triangle was the same size or a different size than the first triangle shown. If the child responded "different," he was asked by the experimenter to indicate how the second triangle differed from the first triangle (the direction of the difference). A series of three random presentations followed in which same-different judgments were made between comparisons of three pairs of blue triangles. Of these judgments, two pairs were different in size and one pair was the same size. The same procedure was followed using the set of blue squares. If the child met the criteria of two correct responses out of the total of three presentations on each of the two tasks, it was assumed he understood the concept of same and different with respect to size and he was selected to continue the experiment.

Prejudgment. A series of five gray disks was used. The five disks measured in diameter 2.5 in. (6.35 cm), 2.25 in. (5.715 cm), 2.0 in. (5.08 cm), 1.75 in. (4.445 cm), and 1.5 in. (3.81 cm).

The child was required to make same-different judgments of a standard 2-in. (5.08-cm) disk and comparison disks. There were eight presentations, four in which the standard disk was judged against a comparison disk of the same size and four in which the standard was judged against a comparison disk of a different size. In all trials, the order of presentation of the standard and comparison disks was randomly varied. Of the four presentations in which a standard disk was judged against a comparison disk of the same size, there were two trials in which the standard served as the inspection disk and two trials in which the comparison stimulus was presented for inspection.

As before, the disks were presented for successive judgments on the dowel protruding from the center of the board. The child inspected a disk for 5 sec, at the end of which time the experimenter removed the disk and replaced it with a test disk for him to judge as same or as different in size. If judged as different, the child was asked to indicate whether the comparison disk was smaller or larger than the inspection disk.

Reward training. The materials used in this task were disks of the same sizes as in the prejudgment phase of the experiment. There were three sets of disks of three brightnesses: a gray series, a white series, and a black series. The brightnesses were obtained in the following way. Ten students enrolled at Randolph-Macon Woman's College for the 1976-1977 academic year were asked to match the brightnesses of the painted disks to the chips on the Munsell color chart. This chart displays a continuum of brightness from black (2/) through gray (5/) to white (9/). Eight of the 10 students matched the brightness of the black disks to the 2/ value and two students matched the disks to 3/. Six of the 10 adults judged the brightness of the gray

disks equivalent to a value of 6/. Of the four remaining students, three judged them as 5/ and one judged them as 7/. All 10 of the students rated the white disks a brightness of 9/. It is apparent from these results that the brightnesses of the disks were noticeably discriminable.

After the prejudgment task, the children were assigned at random to each of three conditions, two experimental and one control. In one of the experimental conditions black was rewarded, and in the other white was rewarded. The control condition was classified as gray and the children in this group received no reward. The children in each of the three conditions were then further assigned at random to each of nine groups: black-black, black-white, black-gray, white-black, white-white, white-gray, gray-black, gray-white, and gray-gray. There were nine children in each group. These test groupings indicated the color (brightness) to be rewarded and which color disks (black, white, or gray) would be paired for judgments in the task. The three age categories, 4-, 5-, and 6-year-olds, were also randomly assigned over the nine test groups for the three conditions. Children in the black condition (black-black, black-white, and black-gray) were given an M & M candy for selecting the color black. Children in the white condition (white-black, white-white, and white-gray) were given the candy reward for a selection of a white disk. Children in the gray condition (gray-black, gray-white, and gray-gray) participated in the game situation but received no reward.

In the reward training, the children were given a reward according to the specification of the three conditions. The child was informed by the experimenter that he was going to play another game with the disks. The disks were of two brightnesses, black and white, and of the same sizes as those used in the prejudgment task. The experimenter placed a pair of disks, one white and other black, on a table. Then the experimenter told the child which of the two disks to choose and to insert it in the slot at the top of a large blue box. The child received his rewards from a blue box, 18.5 x 13.0 x 13.5 in. (46.99 x 33.02 x 34.29 cm, height, length, and width, respectively), which contained an M & M dispenser (Model A10052, Behavior Control, Inc.). The child was told that he could insert only that disk specified by the experimenter, who operated a concealed switch to reward the child's insertion of the disk. For instance, the child was shown a black and a white disk on a table and was told by the experimenter to insert the black disk in the box. If the child was in the black condition, he received an M & M candy immediately after insertion of the disk. If the child was in the gray or white condition, he received no candy.

There were 12 presentations of pairs of disks. For each presentation, the experimenter followed a schedule which designated the selection of each brightness at random half of the time. The children in the experimental conditions received a total of six M & Ms. Those in the control condition received no reward upon the insertion of any disk in the box. The order of presentation of the black and of the white disks was randomly varied so that each color was placed on the table on the child's left in 6 of the 12 trials. Of the 12 pairs of 6 disks presented, 3 were of the same size and 9 were of different sizes.

At the conclusion of this training, the children in the two reward conditions were asked to explain how the blue box operated in order to determine if they could verbalize the concept that the brightness rather than the size of the disk was associated with reward. Children in the control condition were not asked how the box worked. To determine if the child would employ the relationship that he had verbalized, a free-choice task was devised. In this task, the child was told that he was to play another game with the disks. As before, he would be presented with a pair of disks, one black and the other white, and would be required to choose one disk and to insert it into the blue box. There were four random presentations of the disks in which the pairs were of different sizes.

Judgment. After reward training, the child had to make eight same-different judgments of the standard and comparison disks in the manner identical to the prejudgment phase of the experiment, except that the judgments were of the three bright-

nesses of disks (black, white, and gray) according to the specifications of the experimental groupings. An example of this procedure using the black reward condition is: children in the black-black group inspected a black disk for 5 sec and were then presented with a black comparison disk of a different size or of the same size for judgment; in the black-white group, either a black or a white disk was presented as a standard for inspection and a white or a black comparison disk followed in such a manner that a black disk was always paired with a white, and not another black, disk; in the black-gray group, either a black or a gray disk was presented as a standard to be inspected, after which a black or a gray comparison disk followed such that a black was always paired with a gray disk and not with another black disk.

Postjudgment. In a situation identical to the prejudgment phase of the experiment, the child made same-different judgments of a series of five gray disks.

Results and Discussion

The number of errors for each child was obtained in the prejudgment and postjudgment tasks. From these data, a 3 by 3 by 2 (conditions by ages by prejudgment and postjudgment) analysis of variance was performed. This analysis indicated that there was no significant difference among the black, white, and gray conditions in the number of errors that occurred in the prejudgment and postjudgment tasks. The mean number of errors for the black, white, and gray conditions in the prejudgment task was 1.4, 1.1, and 1.6, respectively. In the postjudgment task, the mean number of errors for the black, white, and gray conditions was, in order, 1.6, 1.6, and 1.8. These results indicate that, in judgment of gray disks, the performance of each of the conditions was approximately of the same level before reward training and remained so afterward. There was a significant difference among the three age groups in the number of errors made in judging gray disks in the prejudgment and in the postjudgment phase of the experiment. The mean number of errors made by the 4-, 5-, and 6-year-olds in both tasks was 1.2, 1.0, and .6, respectively. In order to examine the age difference more closely, the number of correct responses (reporting of "same" when the comparison disk was the same size as standard and reporting "different" when the test disk differed in size from the standard) was obtained for each child in the prejudgment and postjudgment tasks. These responses were used in calculating a one-way analysis of variance for each of these tasks. In the prejudgment task, there was a significant difference among the age groups [$F(2,72) = 3.72, p < .05$]. A Newman-Keuls test showed that the 6-year-olds made more correct judgments than the 4-year-olds. There was no significant difference between the 4- and 5-year-olds in the number of correct judgments of gray disks in the prejudgment task. In the postjudgment task, this developmental trend was also significant [$F(2,72) = 53.28, p < .01$].

To determine whether there were any differences in the children's responses of "same" in the prejudgment and postjudgment tasks, two measures were obtained: the proportion of "hits" (responses of "same" when the disks were the same size) and the proportion of "false alarms" (responses of "same" when the disks were different sizes). In the analyses of the prejudgment task,

no significant differences were obtained in either hits or false alarms. The mean proportions of false alarms for the 4-, 5-, and 6-year-olds were .26, .22, and .15, respectively. The mean proportions of hits for the 4-, 5-, and 6-year-olds were, in order, .86, .87, and .96. These proportions show that the children had very little difficulty in judging sizes of the disks in the prejudgment task. In the postjudgment task, no significant difference was found in the proportion of false alarms, but a significant difference was obtained in the proportion of hits [$F(2,72) = 4.5, p < .05$]. To establish the origin of this difference, a Newman-Keuls test indicated that the 6-year-olds reported more hits (mean proportion of .89) than the 4-year-olds (mean proportion of .72). The mean proportion of hits reported by the 5-year-olds was .69. This proportion did not differ significantly from the mean proportion reported by the 6-year-olds and by the 4-year-olds.

The effects of reward training showed that the children were able to associate the reward with the brightness of the disks. In the two experimental conditions, 85% of the children in the white rewarded condition verbalized the concept and 100% of the children in the black condition verbalized the relationship. In fact, only 4 out of the 54 children in the two experimental conditions failed to verbalize this relationship (one 6-, two 5-, and one 4-year-old). In the free-choice task, the black disk was chosen 87% of the time by those children rewarded for that brightness. The white disk was chosen 80% of the time by children in the white reward condition, and the children in the control condition chose the black disk 55% of the time. From the large percentage of children who told the experimenter how the box operated and who used this knowledge to gain more candy, it can be assumed that the children were able to use the relationship between reward and brightness in the context of this task.

A second test of the effects of reward on judgment was the frequency with which the disk of the rewarded brightness was overestimated by each child in the judgment task. The total number of overestimations consisted of two types of responses. The first type occurred when the disks were the same size and the child reported that the disk of the rewarded brightness was both "different" from the "larger" than the disk of the nonrewarded brightness. The second type occurred when the disks were different sizes. If the disk of the rewarded brightness was actually smaller than that of the nonrewarded brightness and the child responded "same," then the response was considered an overestimation. Overestimations of both types were used in calculating a 3 by 3 by 2 (conditions by ages by types of overestimations) analysis of variance. The results showed that there was no significant difference among conditions or types in the number of overestimations. However, a significant difference among the three age groups in reporting the number of overestimations for the 4-, 5-, and 6-year-olds was 1.04, .96, and .44, respectively. It was found that the 4-year-olds made significantly more overestimations than the 6-year-olds and that the 5-year-

olds made significantly more than the 6-year-olds.

WEIGHT

Method

Procedure: Pretraining period. The purpose of the pretraining period was to familiarize the child with the procedure and to assess his understanding of the relationship of same and different with respect to weight.

The materials used in this task were two sets of cans, a large and a small set. Each set consisted of two unpainted silver tins. The cans in each set were equal in height but different in weight.

The experimenter informed the child that he would play a game with weights. After inquiring if the child knew what it meant to say that some things were heavy and that some things were light, the experimenter showed him a set of cans. Beginning with the large set, the experimenter asked the child to lift each can one at a time and then to tell her if the two cans weighed the same. He was told to use both hands to lift the can if he wished. If the child responded "different," he was asked which of the two cans was heavier. The same procedure was used with the small set except that the child was asked which can was lighter if he responded "different."

Judgment. Three sets of containers (plastic pill jars), a small, a medium, and a large set were used. Each set contained three jars of three different brightnesses (black, white, and gray), but were identical in size and weight. The small set weighed 100 g and measures 4.5 x 2.0 in. (11.43 x 5.08 cm). The medium set weighed 300 g and measured 6.25 x 3.0 in. (15.88 x 7.62 cm). The weight of each of the containers in the largest set was 600 g.

After the pretraining period, the child was told that he would play a game with some plastic jars. There were three random presentations for each set in which a black jar was paired with a white jar, a black jar was paired with a gray jar, and a white jar was paired with a gray jar for judgment. The order of the colored jars was randomly presented, so that each jar was presented an equal number of times in each position. The child was instructed by the experimenter to lift each of the two jars one at a time and to report if they were the same or different in weight. The child was reminded that he could use both hands to pick up the jar if he wished. If the child responded "different," he was asked to indicate the direction of the difference (which jar was heavier or lighter). A total of nine judgments was made for the three sets of different jars.

Results and Discussion

The number of errors (responses of "different") and the number of correct responses (responses of "same") was obtained for each child. Of the total number of responses, 46% were errors. When the child responded "different," the direction of the difference was reported to be "heavier" 51% of the time and to be "lighter" 49% of the time. The results were not statistically significant for either ages or conditions. The mean number of "same" responses for the 4-, 5-, and 6-year-olds was 4.1, 5.1, and 5.5, respectively. In order, the mean number of responses of "same" for the black, white, and gray conditions was 4.3, 6.0, and 4.4.

A further test of reward on judgments was done using a one-way analysis of variance with the number of overestimations (a response in which the child reported that one jar was both "different" from and "heavier" than another jar). No significant differences were obtained among the three age groups or among the black, white, and gray conditions in the frequency with which the weight of jars was overestimated.

GENERAL DISCUSSION

The results of the present study failed to demonstrate a significant effect of reward training on subsequent same-different size judgments of disks, the brightnesses of which had previously been associated with reward. Neither were the judgments of the weights of jars differentially affected by rewards associated with the colors of the jars.

In view of the negative findings, the present study lends no support to the assumptions of directive-state theory and the argument for the influence of rewards on perceptual judgments. If the predictions of directive-state theory or the postulations of other "new look" advocates had been borne out, one would have expected evidence for the effects of reward. The rewarding of a particular brightness in this study did not significantly bias subjects in reward conditions to respond "same" or "different" any more or less often in size or weight judgments of black, white, and gray objects than was the case in the control group.

The present study sheds considerable light, however, in the area of methodology. The inclusion of two reward conditions in the experimental design is significant. Because the majority of previous experimentation has employed only one reward condition, differential effects of different rewards could not be determined. In this study, an attempt was made to test for the effects of both black and white rewards. If reward influenced perception, then these effects should have been manifested in both reward conditions. This was not the case.

The major differences in judgments of size resulted as a consequence of age rather than of the reward. Discrimination of size seemed to improve with age. This developmental trend was not observed in judgments of weight. The finding that age was a significant factor in the judgment of size, but not of weight, is consistent with the results of previous research dealing with the development of size and weight discrimination in children (Gibson, 1969).

REFERENCES

- GIBSON, E. J. *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts, 1969.
- LAMBERT, W. W., & LAMBERT, E. C. Some indirect effects of reward on children's size estimations. *Journal of Abnormal and Social Psychology*, 1953, 48, 507-510.
- LAMBERT, W. W., SOLOMON, R. L., & WATSON, P. D. Reinforcement and extinction as factors in size estimation. *Journal of Experimental Psychology*, 1949, 39, 637-641.
- MURRAY, F. S., & GARRETT, E. C. *A methodological analysis of reward on same-different judgments in children*. Paper presented at the 84th annual meeting of the American Psychological Association, Washington, D.C., 1976. (ERIC Document Reproduction Service No. ED 130 770).
- ORNE, M. T. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 1962, 17, 776-783.
- PROSHANSKY, H., & MURPHY, G. The effects of reward and punishment on perception. *Journal of Psychology*, 1942, 13, 295-305.
- ROSENTHAL, R. *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts, 1966.
- SCHAFER, R., & MURPHY, G. The role of autism in a visual figure-ground relationship. *Journal of Experimental Psychology*, 1943, 32, 335-343.
- SNYDER, F. W., & SNYDER, C. W. The effects of monetary reward and punishment on auditory perception. *Journal of Psychology*, 1956, 41, 177-184.
- SOMMER, R. The effects of rewards and punishments during perceptual organization. *Journal of Personality*, 1957, 25, 550-558.
- SOMMER, R., & AYLLON, T. Perception and monetary reinforcements: I. The effects of rewards in the tactual modality. *Journal of Psychology*, 1956, 42, 137-141.

(Received for publication November 15, 1977.)