RESULTS AND DISCUSSION

In order to test statistically for sensitivity, the binomial data point in the unit square was calculated for each test stimulus and two 95% confidence bands (one for the proportions of hits, the other for the proportions of false alarms) were calculated for each data point. These results for the four Ss are shown in Fig. 1. It can be seen that K. J. showed significant sensitivity with all test stimuli but that there was a statistically significant increase in sensitivity only between the 50- and the 56-mm stimuli. T. D. was not reliably sensitive at 50 mm but did show sensitivity at 53 and 56 mm, even though these latter values were not different. K. D. demonstrated significant sensitivity only with the 56-mm test stimulus. C. S. exhibited sensitivity minimally at 53 mm, but this broke down at 56 mm for the only inversion of results among the four Ss. The results show considerable individual differences among Ss, but, taken as a whole, it appears that a test stimulus must exceed the standard by more than 6 mm to enable S to reliably report "different." In short, the DL for supraliminal two-point stimuli under the conditions of this study is in excess of 6 mm. The lack of separation of the two larger stimuli for two Ss, T. D. and C. S., as well as the reversal of these two stimuli for C. S., raises a question about what happens as the magnitude of stimuli are further increased. Does the size of the DL change as the magnitude of the test stimuli increases? There is also a need to investigate with supraliminal stimuli that are progressively smaller than a standard. In addition, it would be interesting to look at other body areas where the initial two-point threshold is larger or smaller. Would a comparable DL for such an area reflect the initial difference in the two-point threshold?

REFERENCES

GREEN, D. M., & SWETS, J. A. *Signal detection theory and psychophysics.* New York: Wiley, 1966.

HILGARD, E. R. *Introduction to psychology.* New York: Harcourt, Brace & World, 1953.

UNDERWOOD, B. J. *Experimental psychology.* New York: Appleton-Century-Crofts, 1966.

Observation vs performance in learning over the fourth to sixth grades*

MELVIN H. MARX and KATHLEEN MARX
University of Missouri, Columbia, Mo. 65201

Fourth-, fifth-, and sixth-grade children were trained and tested in a multiple-choice learning situation under both performance (guessing with knowledge of results) and observation (watching the performer) conditions. Although the observer superiority previously found for college students and grade-school children did not occur, there was a reliable trend in that direction from the fourth to the sixth grade.

The experimental question examined was whether there is any difference in learning by observation, as contrasted with learning by performance (trial-and-error guessing with knowledge of results) in school children of the fourth, fifth, and sixth grades. Superior multiple-choice maze learning by observation has been found in college students (Hillix & Marx, 1960; Rosenbaum & Schutz, 1967) and grade-school children (Rosenbaum, 1967).

SUBJECTS

A total of 140 school children from the fourth, fifth, and sixth grades of a Columbia, Missouri, public school were tested. Two classes from each of the grades were used.

APPARATUS

Stimulus presentation was by means of a Kodak Ektagraphic 2 x 2 in. slide projector (Model AF) and a Radiant Super Champion 40-in. screen. The stimuli consisted of 10 sets of four animals, labeled A, B, C, and D, used as described below. Responses in training trials utilized self-scoring devices called Rapid Raters, manufactured by Research Media, Inc. These small pressboard devices had two columns of 20 items each, with response alternatives A, B, C, and D. Responses were made by inserting a metal stylus into the appropriate hole; the device was modified so that correct answers allowed the stylus to proceed to its hilt, whereas incorrect answers stopped the stylus (by means of a template inserted into the device). Paper answer sheets were perforated by each stylus entry, leaving a permanent record of training responses. Styluses were made by inserting a small nail into the rubber eraser of a standard No. 2 lead pencil. Test answer sheets were mimeographed and required S to write his answer (A, B, C, or D) with the pencil, with no knowledge of results.

PROCEDURE

Each class was tested in the usual classroom. The children were told that they were to participate in an experiment whose purpose was to compare different methods of learning. Their task was to guess which animal from a set of four had been selected as the "correct" one of that set; if their guess was correct, as shown by the full entrance of the stylus into the Rapid Rater, they were to remember that animal, because it would be correct on all future trials. If their guess was wrong, as indicated by the blocking of the stylus, they were to try some other animal on future trials. They were also told not to try to remember the letters, since these identifications for particular animals changed from trial to trial as the positions of the animals changed (randomly determined), although the composition of the sets did not vary (that is, the same four animals always appeared together).

The class was divided into pairs of Ss of the same sex and approximately equal levels of achievement, as rated by the teacher, and one member of each pair was assigned to either the "red team" (left-side S) or the "green team" (right-side S). Red team members performed for the first five sets of animals; that is, the performer did the guessing and used the Rapid Rater, while his partner simply observed his response and its consequence. Then the roles were reversed for the next five sets, which completed the first training trial. Four such training trials were given (except for one sixth-grade class, where time permitted only three trials), with varying orders of sets as well as varied positions of the animals within the sets.

After each training trial a retention test was administered, so that learning under both performance and observation conditions could be compared (there being, of course, no overt record available of observational learning in training). These trials utilized the same kind of stimulus presentation, but with new orders of animal sets and of positions of animals in the slides. The Ss answered on mimeographed test sheets. Pairs of Ss were

*The project reported herein was supported by a grant from the U.S. Office of Education, Department of Health, Education and Welfare. The opinions expressed herein, however, do not necessarily reflect the position or policy of the U.S. Office of Education, and no official endorsement by the Office of Education should be inferred. We wish to express our gratitude to Dr. Robert C. Shaw, Superintendent, and Dr. Harold E. Steere, Assistant Superintendent, of the Columbia Public School System; to Mr. Bob Lincoln, Principal of the Russell Boulevard School; and to the teachers, Mrs. Moen, Mrs. Scott, Mrs. Wilcox, Mrs. Clayton, Mrs. Chastain, and Mrs. Rogers.

Table 1
Number of Ss in Each Grade with More Observation Items Learned (O > P), More Performance Items Learned (P > O), and Equal Numbers of Items Learned (O = P)

	Grade		
	4th	5th	6th
O>P	15	9	24
P>O	29	22	19
O = P	3	11	8
N	47	42	51

separated during testing in order to ensure independence of response and were cautioned not to cooperate.

Stimulus slides were presented automatically at a rate of 15 sec in training trials and 8 sec in test trials by means of the projector timer.

RESULTS

Table 1 summarizes the main data from the experiment. It provides the number of Ss from each grade (two classes combined) who learned more observation items, more performance items, or equal numbers of observation and performance items. This kind of comparison, in which each S serves as his own control, is possible because all of the Ss learned under both conditions. Any biases associated with differences in the two sets of learning materials were controlled, because each set was used for both performance and observation conditions with each pair of Ss.

It is apparent from inspection of these data that the fourth and fifth grades gave very similar results and that they each had a much higher proportion of "performance" learners (approximately 2:1) than did the sixth grade, which showed a reversal of this relationship. The two terminal grades, fourth and sixth, then were compared directly to determine if the apparent shift was a reliable one.

Two statistical analyses were performed. First, four of the tied scores from the sixth grade were eliminated, in order to equalize the Ns at 47. A chi-square test was applied, with the remaining ties divided evenly between the two groups in each grade. The difference in proportion was found to be reliable at just beyond the .05 level of confidence (chi square = 3.87, df = 1).

As a check on this test, a simple t test was performed on the two arrays of total difference scores. The difference scores were obtained by simply subtracting each S's performance total from his observational total. The resulting means were −1.4 for the fourth grade, indicating that these Ss learned almost 150% more performance items on the average, and .1 for the sixth grade, suggesting no real difference between performance and observation learning in this case. The two means were reliably different by the t test ($t = 2.069$, $p < .05$, $df = 96$).

DISCUSSION

These results demonstrate a shift in relative efficiency of learning from the performance to the observation condition between the fourth and the sixth grades, with the fifth grade apparently no different from the fourth. However, although the sixth-grade students surpassed the fourth-grade students on observation learning scores, they did not show an "observer effect" (that is, a real superiority of observation to performance condition). A similar result was found for the other two classes of sixth-grade children at the same school, so that we may tentatively conclude that observational and performance learning processes are equally efficient, at least for these kinds of tests and Ss.

This restriction on the generality of the observer effect is inconsistent with the data on children from the first through the sixth grades reported by Rosenbaum (1967). His Ss learned a multiple-choice maze either as performers or observers. When the grades were paired as lower (first and second), middle (third and fourth), and upper (fifth and sixth), all of his observation groups retained more correct responses than did the performer groups. The observer superiority was statistically reliable ($p < .01$) but there were no reliable interactions. Moreover, calculation of the differences between observation and performance scores by paired grades revealed no clear trend (means of 3.08, 3.71, and 3.50 for the lower, middle, and upper grades, respectively, calculated from data in Table 2 of Rosenbaum, 1967, p. 619).

The discrepancies between the present results and those of Rosenbaum (1967) may be attributed to the many procedural differences and particularly, perhaps, to the fact that the decision-making process in Rosenbaum's performance condition was complicated by a response requirement (inserting a stylus into an 8-pin radio-tube socket) that was presumably more difficult than was the present one (inserting a stylus into a simple hole). Also, Rosenbaum used a correction procedure, allowing up to three errors for each row of sockets, while in the present experiment the noncorrection procedure allowed only one response per stimulus; it is possible that learning under performance conditions is differentially restricted by multiple errors or, in other words, that observers are less distracted by errors and concentrate more on correct responses. Some support for this proposition is provided in an experiment by Chalmers (1964), who reported fewer perserverative errors in observers in a concept-learning task. These suggestions may be considered to be representative of the type of differentiating conditions that will ultimately be identified as critical in the observation/performance dichotomy, but further speculation concerning the differences between the two experimental results does not seem worthwhile until a more thorough experimental analysis is made.

Two provisional explanatory interpretations may be suggested to account for the reliable age difference observed in the present study. The simplest and perhaps most feasible interpretation is that students in the lower grades did not pay sufficient attention during the observation condition; in other words, they were relatively poor observers. This view is supported by some qualitative observations made by the Es, who noticed frequent failures of this kind in the lower-grade classes. Increased maturity in the sixth-grade students apparently enabled them to cope more effectively with the somewhat unusual observation procedure.

A second, somewhat more interesting but at the same time more speculative, interpretation is that the disposition to incorporate someone else's work within one's own personal frame of reference increases substantially over the fourth to sixth grades. The sixth-grade children seemed to understand more readily that they were to be tested on *both* sets of materials. This view was suggested to the Es by the occasional failure of lower-grade children even to attempt to respond on the test trials to the stimuli that they had observed. Unfortunately, there was no systematic recording of this kind of casual observation. Such behavior may be considered symptomatic, nevertheless, of a less prevalent ability in the younger children to understand and cope with the dual learning task.

Each of these two interpretations reflects the growing maturity of the older children, and a decision between them must await the collection of additional, more definitive data.

REFERENCES

CHALMERS, D. K. The effects of direct and observational learning on reversal and nonreversal shifts in concept formation. Unpublished Master's thesis, University of Iowa, 1964.

HILLIX, W. A., & MARX, M. H. Response strengthening by information and effect in human learning. Journal of Experimental Psychology, 1960, 60, 97-102.

ROSENBAUM, M. E. The effect of verbalization of correct responses by performers and observers on retention. Child Development, 1967, 38, 615-622.

ROSENBAUM, M. E., & SCHUTZ, L. J. The effects of extraneous response requirements on learning by performers and observers. Psychonomic Science, 1967, 8, 51-52.