Previous experiments with nursery school children have suggested that (1) subjects of preschool age do not verbalize during transfer learning or that (2) for these subjects, self-produced verbal cues have little influence on the learning process. To investigate the relative merits of these alternative positions, research was conducted among 80 nursery school children in Champaign and Urbana, Illinois. The subjects were assigned discrimination tasks with stimuli varying in either brightness or size. Half of the subjects were required to verbalize their choices. The others responded throughout the experiment without speaking. Confirmation was found for the preliminary postulates that verbalization would cause most subjects to use brightness labels to describe the stimulus regardless of the dimension and that verbalization would significantly facilitate performance on the brightness dimension. Contrary to prediction, however, verbalization did not interfere with performance on the size dimension. The results of this research suggest that, for nursery school children, size discrimination is determined by proprioceptive feedback and is not, therefore, greatly influenced by verbalization. Detailed methodological and theoretical discussions are included in this research report and statistics are reported in five tables. (JS)
The Effect of Subject-Determined Verbalization on Discrimination Learning in Preschoolers

Joseph L. Wolff
Department of Educational Psychology
University of Illinois

In a previous experiment (Wolff, 1967) it was found that a transfer problem involving object stimuli varying simultaneously in the dimensions of size and brightness is learned more easily by nursery-school children when size is the relevant dimension than when the relevant dimension is brightness. This is true despite the fact that such children normally use brightness rather than size labels in describing the positive stimulus at the termination of the transfer task. On the assumption that posttask descriptions reflect the relative dominance of size and brightness labels in the S's operative verbal hierarchy, such a discrepancy suggests either that (a) nursery-school Ss do not verbalize during the transfer task or that (b) for these Ss, self-produced verbal cues have little influence on the learning process. The latter position corresponds to the Kendlers' conjecture, termed by Reese (1962) the "mediational-deficiency hypothesis," that nursery-schoolers are at "a stage in human development in which verbal responses, though available, do not readily mediate between external stimuli and overt responses (Kendler, Kendler, & Wells, 1960, p. 87)." The former position corresponds to Kendler's (1963) subsequent conclusion that nursery-school children do not mediate simply because they normally do not make verbal responses while learning transfer problems. Although a number of studies have investigated the relative merits of these alternative positions, definitive evidence favoring one position over the other has not been forthcoming. It is the aim of the present experiment to throw further light on this issue within the context of the experimental conditions employed by Wolff (1967).
All Ss performed an initial form discrimination with stimuli varying in a single dimension (i.e., form) and were then nonreversal-shifted to either a brightness or a size discrimination with stimuli varying simultaneously in the dimensions of brightness and size (form held constant). Half of the Ss verbalized their choices during both the original-learning (OL) and transfer-learning (TL) periods, while the remaining half responded throughout the entire task without speaking. On the basis of Wolff's (1967) results and previous data, it was predicted that requiring subjects to "name" the stimulus to which they wished to respond before making their choice would cause most Ss to use brightness labels to describe the stimuli regardless of the dimension relevant. Contrary to the mediational-deficiency hypothesis, it was further predicted that verbalization would retard learning of a size discrimination and facilitate learning of a brightness discrimination as compared to a nonverbalization control.

Method

Subjects.--Subjects were eighty nursery-school children drawn from two nursery schools and two day-care centers in the Champaign-Urbana area. In addition to these Ss, twelve Ss were tested whose data were not included. Three of these Ss failed the OL task (two in the verbalization condition and one in the nonverbalization condition), three refused to cooperate during the task, and six had IQ's below the predetermined minimum level--38. Further data on the four experimental groups are shown in Table 1.

Apparatus.--Stimuli were wooden blocks of various shapes, colors, and heights. Two of the stimuli (one diamond-shaped and the other spade-shaped) were red and stood 9/8 in. high. These stimuli were used during OL for all Ss. Of the remaining four stimuli (each of which was spade-shaped), two were black and two were white, one of each shade being 13/8 in. high (tall) and the other 7/8 in. high (short). These stimuli were used during TL. Each stimulus had a circular indentation drilled in
its base to permit it to fit over a marble reward. Other apparatus included the same turntable, box of marbles, and mirror as were used by Wolff (1967).

Procedure.--All Ss were individually administered the Peabody Picture Vocabulary Test and the nonreversal-learning task in a single experimental session. Separating the administrations of the tasks--which were always given in the above order--was a 40 to 60 second interval. This time was used by E to check S's IQ and prepare the apparatus for the second task. Prior to this task Ss were randomly assigned to OL positive cue, TL relevant dimension, verbalization condition, and TL positive cue. The single restriction that was placed upon the randomization was that 20 Ss successfully completing OL be eventually assigned to each of the four experimental conditions created by the $2 \times 2 \times 2$ (verbalization vs. nonverbalization) design.

To begin the nonreversal task, Ss assigned to the verbalization condition were read the following instructions (Ss assigned to the nonverbalization condition were read only the first and last paragraph):

Now __________ we are going to play a game with these things. The game is to win as many marbles as you can. Here is how we play the game. See these two things? One of them is a winner and the other one is a loser. Under the winner, there will always be a marble, but under the loser, there will never be a marble. If you pick the winner, you may keep the marble that is under it and put it in that box (E points to box), but if you pick the loser, you won't get anything. Each time you may choose only one. Then I will turn it around like this and you will have another turn. But on each turn you may choose only one.

Now, one last thing. I want you to tell me which one you want to pick before you choose it. For instance, if you thought this one
was the winner (E points to club), you would say, "the club," and then pick it up. If you thought this one was the winner (E points to diamond), you would say "the diamond," and then pick it up. Do you understand? Now, I called this one the club and this one the diamond, because these are the names that I like to use. But you can use any names that you want.

See if you can get a marble every time--OK?

In the OL portion of the task all Ss made an original discrimination between two stimulus blocks--one medium-size, red, and club-shaped and the other one medium-size, red, and diamond-shaped. Positioning of stimuli during this period was determined randomly with two constraints: (1) stimuli were not allowed to remain in the same position over more than three successful trials and (2) stimuli maintained their position over trials until S made a correct response. Marbles baited under the discriminanda served as rewards and Ss worked to a criterion of 9 of 10 successive correct responses or to an elimination point of 70 trials. Approximately every tenth trial E said, "try to get a marble every time you choose."

A noncorrection procedure was used at all times.

Following OL, all Ss not eliminated received a nonreversal shift (half to size as the relevant dimension and half to brightness as the relevant dimension) with no other break in procedure. Stimuli for the shift period consisted of four spade-shaped discriminanda varying in height and brightness. During this period, stimuli were presented tall-white, short-black on half the trials and tall-black, short-white on the remaining half. Pair presentations and positioning of stimuli within pairs were determined randomly with two exceptions: (1) the same pair or positioning within pairs never appeared over three successive successful trials and (2) stimuli remained constant over trials until S made one correct response.
Criterion remained the same as in the OL period and cut off was set at 50 trials. As in OL, Ss were periodically instructed to try to get a marble every time they chose.

Subjects assigned to the verbalization condition vocalized their choices in both OL and TL; Ss in the nonverbalization condition responded silently throughout the task. During the OL period, Ss in the verbalization condition were allowed to use any label for the stimulus that they desired. However, in the event that an S failed to supply a label within approximately five seconds, he was told, "Tell me which one you want to choose and then pick it up," and this instruction was repeated until S either gave a label or stated that he did not know what to call the discriminanda, in which case the labels "club" and "diamond" were supplied to him. In the event that S responded to a stimulus without speaking, he was told, "No, you must tell me first and then pick it up. So tell me which one you chose and then pick it up again." This extra trial was not counted.

During the TL period, verbalization-group Ss were restricted to the use of three types of labels: size labels (e.g., big, little), brightness labels (e.g., white, dark, yellow), or compound labels involving size or brightness (e.g., white-heart, big-black one). To insure that these labels were used, one of two correction procedures was employed whenever S's description of the stimulus fell outside these three categories: (a) If S attempted to use the previously relevant form words, he was asked, "Which one is the club? (S pointed.) Which one is the diamond? (S pointed.)" S then said, "See, I can't tell which one you want if you use those words; tell me so that I can tell without looking." (b) If S used any other nonpermissible label he was told, "No, I can't tell which one you want if you use that word; tell me so that I can tell without looking." Following the occurrence of the first permissible label elicited by either correction procedure, S said, "Good." In both correction procedures pertinent parts of the instructions
were repeated until S responded with a permissible label, and although all motor choices were scored, recording of labels commenced only after the emission of the first permissible label. Subjects failing to verbalize their choices or responding without speaking were treated as an OL. Subjects claiming not to "know" what to call the discriminanda were told, "Just tell me so that I can tell which one you want without looking."

Following the completion of TL, all Ss were shown the last positive and negative stimulus used in the shift period and were asked to describe the positive stimulus by means of the following ordered set of questions: (1) Which one is the winner. (2) Tell me, I can't see. (3) What does it look like? Questioning was stopped at any point at which S made a verbalization.

**Results**

The number of Ss in the verbalization condition emitting various stimulus labels **during TL** as a function of the dimension relevant is shown in Table 2. To test the effect of the dimensional variable on verbalization, the percentage of brightness labels occurring among all labels recorded in the TL period was determined for each S individually. For the brightness-relevant group the mean of these percentages was 99.6; for the size-relevant group the mean was 91.4. According to a Mann-Whitney test, the distributions of these percentages within the two groups were not significantly different ($z < 1, p > .10$).

Table 3 shows the post-TL verbalization data broken down into four categories. From this table it is apparent without statistical analysis that the incidence of brightness labels given was independent of the dimension to which S was assigned. Both this finding and the results of the Mann-Whitney test above are in accord with prediction 1.
Trials to criterion in OL for the four experimental groups are shown in Table 4. In interpreting this table, it is important to note that the dimensional variable represented by the two rows did not become a true variable until TL, and that the OL problems of the groups labeled size and brightness were, in fact, the same. Thus, while a verbalization effect was considered probable, a dimensional effect was not anticipated. Contrary to expectation, analysis of variance indicated no effect due to verbalization or to interaction and a significant effect \( (F(1,76) = 6.16, p < .05) \) due to dimensions.

This finding may be given one of three interpretations: (a) by chance, better problem solvers were assigned to the size-relevant condition than to the brightness-relevant condition. (b) By accident, experimental arrangements during OL happened to favor the size-relevant group (e.g., more Ss in this group had their preferred cue as S+). (c) The experimenter inadvertently radiated a systematic bias favoring the size-relevant condition. Because OL and TL scores were uncorrelated, the first two possibilities pose no problem for the analysis of the TL results. Since the third possibility cannot be eliminated, however, no attempt can be made to interpret any main effect for dimensions obtained in TL. Aside from this, the interpretation of the TL results would not appear to be affected by the unexpected dimensional effect in the OL data.

Trials to criterion in TL for the four experimental groups are shown in Table 5. Since an \( F_{\text{max}} \) test showed significant heterogeneity of variance among the four groups \( (F_{\text{max}}(k,19) = 11.23, p < .05) \), scores were subjected to a square-root transformation to restore homogeneity \( (F_{\text{max}}(k,19) = 2.22, p > .05, \text{for the transformed scores}) \). A \( 2 \times 2 \) analysis of variance of the resulting data indicated significant main effects for dimensions \( (F(1,76) = 29.88, p < .001) \) and verbalization \( (F(1,76) = 10.29, p < .01) \) and a marginally significant dimensions \( \times \) verbalization...
interaction \((F(1, 76) = 3.37, p < .10)\). Individual comparisons further suggested that verbalization facilitated performance on the brightness dimension \((t(33) = 3.12, p < .01)\) but had no effect on the size dimension \((t(33) = 1.16, p > .10)\).

Discussion

It was hypothesized at the beginning of this paper that requiring subjects to verbalize their choice during original and transfer learning would (1) cause most Ss to use brightness labels to describe the stimulus regardless of the dimension relevant and, as a consequence, (2) retard learning of a size discrimination and facilitate learning of a brightness discrimination as compared to a non-verbalization control. The first part of this hypothesis was strikingly confirmed. All but four of the forty Ss in the two verbalization conditions used brightness labels exclusively in describing their TL choices, and the dimension relevant had almost no effect on the incidence of such labels emitted. Similar effects were observed in the posttask verbalization period.

The second part of the hypothesis was confirmed only in part. In line with prediction, verbalization significantly facilitated performance on the brightness dimension. Contrary to prediction, it did not interfere with performance on the size dimension.

According to the mediational-deficiency hypothesis, nursery schoolers are at "a stage in human development in which verbal responses, though available, do not readily mediate between external stimuli and overt responses." According to the second hypothesis considered at the first of this paper, nursery schoolers do not mediate because they normally do not make verbal responses while learning transfer problems. Supporting the former hypothesis over the latter are the findings that verbalizing the dimension of brightness fails to inhibit the learning of a size discrimination and that the type of label used to describe the TL discriminanda is
not a function of the type of discrimination (size or brightness) that is being undergone. Supporting the latter hypothesis over the former is the finding that learning a brightness discrimination is facilitated by the overt use of brightness labels.

Clearly all the present results are explained by neither the meditational-deficiency hypothesis nor its alternative as proposed above. The presumption is, therefore, that both hypotheses, as they now stand, fail to afford an adequate description of the mechanisms governing transfer behavior in the nursery-school child. Is there some other theory or hypothesis with which the present results are more congruent?

In an attempt to explain the data of a previous experiment, Wolff (1967) hypothesized that kinesthetic stimulation arising from the differential strain of lifting bigger (heavier) and smaller (lighter) discriminanda provides highly salient cues for nursery-school children in learning simple size discriminations, and that the salience of such cues diminishes with age. A recent study by Milgram and Furth (1964) suggests that this hypothesis, originally suggested by a more general position proposed by White (1965), may also be of use in explaining the present data.

According to Milgram and Furth, position discriminations are less susceptible than visual discriminations to the facilitory and inhibitory effects of verbal mediation (see also House, 1964). On the commonly held assumption that the principle stimulation directing a position discrimination is the differential proprioceptive feedback arising from movements from one side to the other (see Sperling, 1967, for instance), one plausible inference from this hypothesis is that all response under the control of proprioceptive stimuli are generally less amenable to verbal influence than are responses controlled by other (e.g., exteroceptive) sources of stimulation. Thus, if the hypothesis be accepted that the apparent
discrimination of size may in some cases involve the actual discrimination of proprioceptive stimuli, it follows from this conjecture that, in these cases, verbalization might well have a markedly different effect on size and brightness discrimination in just the manner that was in fact observed.

In view of this argument, it is possible to consider the results of both the present and a previous experiment (Wolff, 1967) as pointing to the same tentative conclusion: that at early ages, and for certain types of tasks, children tend to rely to a greater extent on proprioceptive feedback than on exteroceptive stimulation in making simple transfer discriminations. The implications of this notion for several theories of development and cognition (e.g., Bruner, 1966; Piaget, 1950) are deserving of consideration.
Footnotes

1. This research was supported by a contract with the United States Office of Education (OE-6-8934).

2. These two positions have been designated respectively the "mediational-deficiency hypothesis" and the "production-deficiency hypothesis" by Flavell, Beach and Chinsky (1966).

3. These instructions are a modified version of instructions used by Kendler and Kendler (1959).

4. From previous work it was anticipated that a small number of Ss would use incorrect color labels in designating the stimuli (e.g., "yellow" to designate white). Since eliminating such Ss might create a bias, any color label emitted during the TL period was treated as if it were a brightness label.

5. The within-cell correlation as determined from an analysis of covariance was nonsignificant +.13. This covariance analysis, incidentally, yielded almost identical results as the analysis of variance reported below. Analyses of covariance using MA as a covariate were also performed on the OL and TL data since MA differences favoring the size groups were indicated in Table 1. In both cases, MA failed to correlate significantly with trials-to-criterion scores and the analysis of covariance yielded the same significance levels for all effects as the analysis of variance.

The lack of correlation between OL and TL scores removes the necessity for considering possibility (a) because it indicates that the ability to solve the original problem is unrelated to whatever ability may be involved in solving the nonreversal shift; hence, there is no reason to believe that the four experimental groups differed in this latter intellectual skill. Possibility (b) is also contraindicated by this result because if the experimental arrangements favoring the Ss of the size condition operated during both OL and TL, those Ss
favored in OL would also have been favored in TL with the result that OL and TL scores would surely have been correlated. The only alternative to this conclusion is the assumption that each S in the size-relevant condition received the same amount of chance facilitation or that each S in the brightness-relevant condition received the same amount of chance interference. Such a state of affairs is too unlikely to warrant consideration.
References


Table 1
Age and Intelligence Characteristics of the Sample

<table>
<thead>
<tr>
<th>Verbalization Condition</th>
<th>Dimension</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Size</td>
<td>Brightness</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>IQ</td>
<td></td>
</tr>
<tr>
<td></td>
<td>57.5 (46-73)</td>
<td>55.5 (38-71)</td>
<td></td>
</tr>
<tr>
<td>Verbalization</td>
<td>IQ 110.5 (89-125)</td>
<td>109 (98-140)</td>
<td></td>
</tr>
<tr>
<td>Non-verbalization</td>
<td>Age 56 (45-76)</td>
<td>53 (41-65)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IQ 113 (103-130)</td>
<td>110 (89-124)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Ages are reported in months. The first statistic in each cell is a median. The figures in parentheses denote ranges.
Table 2

Frequency of Subjects Using Various Types of Labels to Describe the Stimuli During TL

<table>
<thead>
<tr>
<th>Relevant Dimension</th>
<th>Type of Verbalization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Brightness^a</td>
</tr>
<tr>
<td>Size</td>
<td>18</td>
</tr>
<tr>
<td>Brightness</td>
<td>18^e</td>
</tr>
</tbody>
</table>

^a. Includes only those Ss who used no size, size-compound, or irrelevant labels.

^b. Includes Ss who used at least one size label.

c. Includes Ss who used at least one size-brightness-compound label (e.g., big-black one).

d. Includes Ss who used at least one inappropriate label after the emission of an appropriate label.

e. Two of these Ss used color labels (yellow, green, purple) to designate the stimuli on certain trials. Two used brightness-form compounds on one or more trials.
Table 3

Frequency of Subjects Verbalizing Various Aspects of the Stimuli During Post-TL Questioning

<table>
<thead>
<tr>
<th>Condition</th>
<th>Type of Verbalization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Brightness</td>
</tr>
<tr>
<td>Verbalization Brightness</td>
<td>19</td>
</tr>
<tr>
<td>Verbalization Size</td>
<td>17</td>
</tr>
<tr>
<td>Nonverbalization Brightness</td>
<td>13</td>
</tr>
<tr>
<td>Nonverbalization Size</td>
<td>15</td>
</tr>
</tbody>
</table>

<sup>a</sup> This category includes those Ss verbalizing both dimensions.

<sup>b</sup> This category includes those Ss giving irrelevant verbalizations (e.g., "me," "winner," "heart," "fire"), Ss claiming not to know, and Ss refusing to verbalize.
Table 4: Means and Standard Deviations of Trials to Criterion Scores in Original Learning for Four Groups

<table>
<thead>
<tr>
<th>Relevant TL Dimension</th>
<th>Condition</th>
<th>Mean</th>
<th>S.D.</th>
<th>Combined Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>Verbal</td>
<td>3.95</td>
<td>7.37</td>
<td>5.66</td>
</tr>
<tr>
<td></td>
<td>Nonverbal</td>
<td>5.60</td>
<td>7.57</td>
<td></td>
</tr>
<tr>
<td>Brightness</td>
<td>Verbal</td>
<td>9.25</td>
<td>12.94</td>
<td>11.10</td>
</tr>
<tr>
<td></td>
<td>Nonverbal</td>
<td>12.40</td>
<td>14.60</td>
<td></td>
</tr>
</tbody>
</table>

Note: n=20.
<table>
<thead>
<tr>
<th>Relevant TL Dimension</th>
<th>Condition</th>
<th>Mean</th>
<th>S.D.</th>
<th>Combined Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>Verbal.</td>
<td>3.50</td>
<td>6.13</td>
<td>5.68</td>
</tr>
<tr>
<td></td>
<td>Nonverbal.</td>
<td>6.85</td>
<td>12.39</td>
<td>5.68</td>
</tr>
<tr>
<td>Brightness</td>
<td>Verbal.</td>
<td>13.85</td>
<td>15.60</td>
<td>22.65</td>
</tr>
<tr>
<td></td>
<td>Nonverbal.</td>
<td>31.65</td>
<td>20.56</td>
<td>22.65</td>
</tr>
</tbody>
</table>

Note: n=20.