
Minnesota Univ., Minneapolis. Research, Development, and Demonstration Center in Education of Handicapped Children.

Bureau of Education for the Handicapped (DHEW/ OE), Washington, D.C.

BR-332189

Jan 75

OEG-09-332189-4533(032)

55p.

EDR S PRICE

MF-$0.76 HC-$3.32 Plus Postage

EDRS DESCRIPTORS

*Educable Mentally Handicapped; Exceptional Child Research; Mentally Handicapped; *Paired Associate Learning; *Performance Factors; *Questioning Techniques; *Verbal Learning

ABSTRACT

Thirty-six normal 5-year-olds and 60 educable mentally retarded (EMR) students with a mean age of 7 years were tested to determine whether specific formats of interrogatives would be instrumental in inducing the generation of effective verbal mediators. Ten EMR Ss and 6 normal Ss were tested on 21 paired associates in each six experimental conditions: labeling, sentence generation, sentence repetition, response to Why "A" (auxiliary did not function as the main verb) and response to why "B" (auxiliary functioned as the main verb). Analysis of correct response indicated that Ss in the three question conditions performed better than Ss in the nonquestion conditions. No differences existed between the two subject categories. The question conditions were found to induce greater semantic analysis in the Ss than the nonquestion conditions. (Author/CL)
THE PRODUCTION DEFICIENCY MODEL OF VERBAL ELABORATION:
SOME CONTRARY FINDINGS AND CONCEPTUAL COMPLEXITIES

James Turnure, Nissan Builum and Martha-Thurolw
University of Minnesota
Research, Development and Demonstration
Center in Education of Handicapped Children
Minneapolis, Minnesota

January 1975

The research reported herein was performed pursuant to a grant from the Bureau of Education for the Handicapped, U.S. Office of Education, Department of Health, Education, and Welfare to the Center of Research and Development in Education of Handicapped Children, Department of Psychosocial Educational Studies, University of Minnesota. Contractors undertaking such projects under government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official position of the Bureau of Education for the Handicapped.
The University of Minnesota Research, Development and Demonstration Center in Education of Handicapped Children has been established to concentrate on intervention strategies and materials which develop and improve language and communication skills in young handicapped children.

The long term objective of the Center is to improve the language and communication abilities of handicapped children by means of identification of linguistically and potentially linguistically handicapped children, development and evaluation of intervention strategies with young handicapped children and dissemination of findings and products of benefit to young handicapped children.
Abstract

Thirty-six normal and 60 EMR children were tested to determine whether specific formats of interrogatives would be instrumental in inducing the generation of effective verbal mediators. Ten EMRs and six normals were tested on 21 paired-associates in each of six experimental conditions: Labeling, Sentence Generation, Sentence Repetition, Response to What, Response to Why A, and Response to Why B. Analysis of correct responses indicated that children in the three question conditions performed better than children in the non-question conditions. No differences existed between the two subject categories. In addition, analyses of semantic and non-semantic errors suggested that the question conditions induced greater semantic analysis in the children than the non-question conditions. The implications of the findings as a basis for reconsideration of a production deficiency hypothesis, and their relation to the concept of "spontaneous production" are discussed.
Research on verbal learning and language functioning in young normal and mentally retarded children in the first half of the 1960's produced suggestions by Luria (1961), Reese (1962), and Kendler (1963) that children younger than six years of age may be unable to use language to form basic associations between two disparate perceptual items or physical objects. This research was generally based on theoretical contentions regarding the importance of either the acquired distinctiveness or equivalence of cues (cf. Dollard & Miller, 1950). Words often served as cues, and the effects of varying distinctive verbal labels across similar physical stimuli, or the converse, were explored in a variety of increasingly complicated learning situations (cf. DiVesta & Paiermo, 1974). The apparent inability of young children to effectively utilize such verbal mediators become known as the mediational deficiency hypothesis (Flavell, 1970; Maccoby, 1964).

Subsequent research regarding mediational skills of children resulted in the questioning of the mediational deficiency model. Flavell (1970) has suggested that the poor performance of young children in tasks which required them to mediate associations
between two items resulted from a production deficiency rather
than a mediational deficiency; they were not able to spontaneously
produce effective mediators, but were able to use certain ready-made
(i.e., experimenter provided) mediators appropriately. Empirical
findings from the laboratories of Jensen and Rohwer (1963), Martin
(1967), MacMillan (1970), and Turnure (1971) indicated that groups
of both young normal and educable mentally retarded (EMR) children,
when provided with experimenter-designed mediators that included
verbal contexts greater than mere labels, used them effectively in
the recall of the test items. Those of the above researchers who
investigated the subjects' ability to generate their own mediators
(Martin, MacMillan, and Jensen & Rohwer) further concluded that their
young subjects' self-generated verbal elaborations (usually "sentence
fragments" or "conjunctives"; e.g., "The cup and the soap") did not
function as effective verbal mediations. Thus, these investigators
emphasized a production deficiency hypothesis, although Flavell might
refer to the subjects' ineffective efforts as production inefficiencies
(Flavell, 1970, p. 199), but with both formulations being clearly
distinct from the earlier mediational deficiency hypothesis. Investi-
gators also refined the basis of the production deficiency hypothe-
sis, relating it to emerging psycholinguistic formulations promoting
the fundamental significance of syntactic structure and sentential
relations (cf. Blumenthal, 1967; Chomsky, 1957; McNeill, 1966; Suzuki
& Rohwer, 1968).

Examination of a study by Jensen and Rohwer (1965) does in fact
suggest that the kinds of utterances produced by young children are ineffective as verbal mediators. However, before one concludes that the poor performance observed is due to a psycholinguistic related production deficiency, one needs to insure that alternative explanations would not be equally viable. There are other factors, such as instructions, whose role in the mediation task needs to be clarified.

When older children were asked by Jensen and Rohwer (1965) to "make up a sentence" regarding a pair of associates, they were able to meaningfully integrate the items. On the other hand, children younger than six years of age tended to link the items in a conjunction format (e.g., "The cow and the ball") in response to the instructions to "make up a sentence." Perhaps young children's comprehension of the abstract term "sentence" is different than that of older children (i.e., young children do not share the adult meaning of the word "sentence"; Downing, Note 1; Samuels, Note 2) and such instructions may not be the most effective ones to elicit effective verbal elaborations in young children.

All of those who either work with young children or have children of their own know that children often engage in lengthy, meaningful verbal utterances in diverse circumstances (cf. for instance, Chukovsky, 1968; Rosen & Rosen, 1973); and, most pertinent to our concerns, they often do this in response to questions. It appears conceivable to use the interrogative format as a method to induce the young child in a learning situation to produce suitably extended verbal responses. This tactic would be a version of the "promotive usage" of interrogatives in Reichenbach's (1947) analysis of instru-
mental language. Whether the child's response would be an effective verbal elaboration (i.e., one which would meaningfully integrate two items in a paired-associate format and thus enhance the child's future recall of the items) was the experimental question proposed here. Why is the child's own verbal participation in mediational tasks to be so sought after? Jenkins (1974a) has suggested that the primary organization of memory is semantic-based, thus favoring the encoding of "meaning" variables (semantic attributes) over the encoding of structural variables such as form and syntax. The activation of semantic memory is seen as a function of the cognitive ability of the child with respect to the given materials (see also, Craik, 1973). Semantic analysis appears to be best insured when the child is given an active role in the mediation process (Anderson, 1970). When materials are subjected to semantic analysis, they are recalled well whether the child has consciously formulated a strategy to store them in memory or not (Turnure, 1971).

The present study investigated whether specific formats of interrogatives (WH types) would be instrumental in inducing young children and EMRs, presently characterized as "production deficient," to generate verbal responses that function as effective verbal mediators in enhancing the acquisition and recall of paired-associates. Based on previous evidence, it is apparent that a direct comparison of labeling (L), sentence generation (SG), and sentence repetition (SR) conditions among equivalent groups of young children (at least
below 7 years of age, cf. Jensen & Rohwer, 1965) should produce paired-associate performance ordered as L < SG < SR. The implication of the analysis of interrogatives, and their potential for activating satisfactorily "deep" semantic processing (cf. Craik, 1973), would be that performance under interrogative instructional conditions would exceed that under sentence repetition conditions. However, the linguistic category of interrogatives is varied, and can be subdivided along the lines of: (a) formal class restrictions (cf. Miller & Ervin, 1964) which pertain to the structural agreement required by formulating answers to specific questions (e.g., Wh as opposed to, say, Yes-No Questions); (b) transformational rules applying between and within such restrictions (cf. Dale, 1972, Ch. 4); and (c) developmental stages associated with the appropriate utilization of such restrictions (cf. Lee & Cantor, 1971). Since all subjects to be involved in the experiment would theoretically be expected to have developed basic mastery of all such question types, no significant performance differences among the Wh question types involved here (see Method) would necessarily be expected. However, it is an open question whether the cognitive demands of the various Wh question types would induce equally effective verbal mediators in the paired-associate task. Data available indicate, for instance, that Yes-No questions do not surpass sentence repetition as an elaborative technique (Buium & Turnure, Note 4).

The investigation was designed so that the results would also speak to the issue of what the subject does when he is not respond-
ing correctly. Specific questions asked were: (a) What is the extent of semantic errors induced by each condition? (b) Does the extent of semantic errors reflect the extent of semantic analysis induced by the various conditions? It is assumed that the task encountered by the child in the Wh conditions may be as follows: Upon hearing the question "Why is the turtle walking to the house?" the child proceeds (1) to analyze the semantic attributes (Barclay, Bransford, Franks, McCarrell, & Nitsh, 1974) or conceptual features (Reid, 1974) (both forms of terminology refer to the item's semantic properties) of the turtle and house which intercept or interlock within the context of the questions (semantic analysis), and (2) to integrate these conceptual features in his response (semantic integration). For example, in responding to the above question, a subject presumably identified one conceptual feature of the turtle as having a need for food, and a conceptual feature of the house as being capable of food storage. These two features were interlocked in the following way: "Cos he wants to go in to eat something like nuts...."

Evidence for this presumptive analysis may be obtained by including in the test certain items that vary in sharing conceptual features. When items share a large number of conceptual features, they are said to belong to a given paradigmatic set. Unlike the traditional, narrow grammatical definition (Jenkins, 1954; Saporta, 1955) of a paradigmatic set that has largely limited its usefulness to free association tests, Reid (1974) suggests an expanded and more flexible definition in which paradigmatic relations are
conceived of in terms of shared conceptual features that characterize
the members of a given paradigmatic set. When items are analyzed in
terms of their conceptual features or semantic attributes, members
of the same paradigmatic set are more likely to be substituted for
each other, primarily due to their minimal contrast: "A unit is
meaningful only to the extent that it is in contrast with other units
that could have occurred in the same context" (Meld, 1974, p. 327).
Experimental conditions which necessitate an increased amount of
conceptual features analysis would result in a higher proportion of
such substitution class errors (semantic errors) whereas conditions
that require minimal semantic or conceptual analysis (e.g., labeling)
would result in a lower proportion of such semantic errors, although
total errors would be higher due to the general lack of semantic
relations established between item pairs (Turnure, 1971; Turnure &
Walsh, 1971).

Samples of both mentally retarded and normal populations were
included in the experiment to test the generality of effects. The
examination of effects in both populations appeared to be especially
pertinent inasmuch as the retarded are widely imputed to be particularly
prone to production deficiencies (cf. Brown, 1974; Milgram, 1973).

Method

Subjects. Ninety-six children, 36 normals and 60 EMRs, partici-
pated in this study. The two groups were of equivalent mental age
(MA), with MAs ranging from 4.6 to 5.6 years. The normal children
(CA range 5.0 - 5.6 years) were selected from four kindergartens
in the Minneapolis/St. Paul area, while the EMR children came from seven Special Education classes from the same area. The mean CA and IQ of the EMR subjects was 7.5 (SD = 1.3) and 67.5 (SD = 8.7), respectively.

**Conditions.** Ten EMR children and six normal children were randomly assigned to each of the following conditions (examples in parentheses):

1. **Labeling** - subject repeated labels of the items after the tester (Soap - Jacket; Light - Shoes)
2. **Sentence Generation** - subject made up a sentence (Make up a sentence about...)
3. **Sentence Repetition** - subject repeated standard, one-relation sentence after tester (The soap is hiding in the jacket. The light is shining on the shoes)
4. **Response to What** - subject responded to standard What question (What is the soap doing under the jacket? What is the light doing to the shoes?)
5. **Response to Why A** - subject responded to standard Why question in which the auxiliary did not function as the main verb (Why is the soap hiding in the jacket? Why is the light shining on the shoes?)
6. **Response to Why B** - subject responded to standard Why question in which the auxiliary functioned as the main verb (Why is the soap in the jacket? Why is the light on the shoes?)

The experiment therefore had the form of a 2(Subject Category) x 6(Conditions) factorial design.

**Materials.** Forty-eight color pictures of common objects from a pre-primer workbook were used as the stimulus materials. From these 48 pictures, 24 pairs were formed with no common or obvious relationships of meaning existing between the members of any pair.
Twenty-one test pairs were chosen in order to construct semantic categories either of the stimulus or the response items. Seven such categories were constructed: 1) Clothing (hat, mittens, jacket), 2) Furniture (bed, table, chair), 3) Tools (shovel, saw, hammer), 4) Footwear (boots, socks, shoes), 5) Water animals (turtle, fish, duck), 6) Land animals (monkey, cat, dog), and 7) Containers (box, basket, cup). (See Table 1 for a complete list of the test pairs.)

Procedure. Each child was tested individually. First, the child was pretrained with three pairs to insure that the instructions of each experimental condition were clearly understood. Each pair was presented to the child for 15 seconds, during which the child was requested to respond to the items according to the condition directions. When more than 15 seconds were needed for the child to complete his response, an additional 15 seconds were allowed.

In all conditions the tester had pre-established procedures to follow in the event that the child did not respond. Essentially, these procedures consisted of repeating the instructions with an additional encouragement (e.g., in the Sentence Generation condition subjects were told: "Make up a sentence; any sentence that you think is right will do"; in the Response to Why questions subjects were told: "There is no right or wrong answer. Any answer that you think is right will do." In the Sentence Repetition condition subjects were instructed to "please repeat the entire sentence....")

The frequency of such instructional enhancement was relatively minimal, and constant among conditions; some enhancement was necessary for
Table 1
Semantic Categories of the Twenty-one Test Pairs

<table>
<thead>
<tr>
<th>Clothing (Response item)</th>
<th>Footwear (Response item)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doll - Hat</td>
<td>Wheel - Boots</td>
</tr>
<tr>
<td>Pencil - Mittens</td>
<td>Tent - Socks</td>
</tr>
<tr>
<td>Soap - Jacket</td>
<td>Light - Shoes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Furniture (Response item)</th>
<th>Water animals (Stimulus item)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comb - Bed</td>
<td>Turtle - House</td>
</tr>
<tr>
<td>Wagon - Table</td>
<td>Fish - Book</td>
</tr>
<tr>
<td>Ball - Chair</td>
<td>Duck - Toaster</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tools (Response item)</th>
<th>Land animals (Stimulus item)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Telephone - Shovel</td>
<td>Monkey - Kite</td>
</tr>
<tr>
<td>Candle - Saw</td>
<td>Cat - Gun</td>
</tr>
<tr>
<td>Pie - Hammer</td>
<td>Dog - Clock</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Containers (Response item)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gate - Box</td>
</tr>
<tr>
<td>Ball - Basket</td>
</tr>
<tr>
<td>Boat - Cup</td>
</tr>
</tbody>
</table>
some children in each condition including the presumably "simple"
labeling condition.

This same procedure was employed for one or two children in
one or another of the WH conditions, when, due to misinterpretation
of the guidelines for employing this procedure, a tester refused to
accept the answer "cause" as a satisfactory response, and requested
another response.

In the training phase, the child was presented with the 21 pic-
torial item pairs and, according to the condition instructions, was
required to respond to each pair. Each pair was presented for 15
seconds; again, if the child's answer required more time, an additional
15 seconds were allowed after the items were withdrawn from the sub-
ject's view. The procedure of extending the response time most often
occurred with the first few pairs, and it may be noted that the
procedure was utilized in all conditions requiring production of a
verbal mediator, but with only a few children in each condition.

After completion of the training phase, the child was presented
with the recall test, in the form of a standardized paired-associate
anticipation format. In the recall test, each subject was shown
only one picture of the pair (the first item) for up to 20 seconds,
and was asked to identify the picture that "goes with it." The order
of pairs for recall was a constant, randomized variant of the original
list. Subjects' responses on recall were manually recorded.

Error classification and analysis. All errors made by subjects
during the recall test were recorded and later classified as being
either semantic errors, non-semantic errors, or non-responses. Errors
were classified into 5 types (3 semantic, 1 non-semantic, and 1 non-response), according to the following scheme:

A. Semantic errors

Type I - errors within experimental categories. These errors consisted of non-correct responses that were included in the predesigned semantic category of the stimulus or response item. Example: Box instead of basket.

Type II - errors due to the child's categorizations. These errors consisted of non-correct responses that were (1) within the list of items presented to the child, (2) outside the experimentally designed semantic categories of the stimulus or response items, (3) indicative of a meaningful relation belonging to an experimentally unintended semantic category. Example: Candle linked to Light (lightings category).

Type III - errors characterized by an association. These errors consisted of non-correct responses that (1) were outside the list of items presented to the child, and (2) had a high probabilistic value that they would be associated with the stimulus item. Examples: Rabbit associated with carrots.

B. Non-semantic errors

Type IV - non-semantic errors. These errors consisted of non-correct responses that did not lend themselves to any interpretable meaningful relation. Example: Book instead of socks.

C. Non-responses

Type V - non-responses. These consisted of the events in which the child failed to name any response item.

Errors were classified by three individuals in accordance with the above classification scheme. To test the reliability of the categories, error types III and IV (which were open to subjective
judgment) were subjected to further investigation. Ten samples of each error type, defined as such by the experimenters, were presented in a randomized order to 30 judges who (1) were given the definition of the two error types, and (2) asked to sort each pair of items as belonging to either error type III or error type IV.

Three hundred judgments made regarding the error type III samples were in complete agreement with the original classification (30 judges X 10 pairs of items). For error type IV, 289 judgments, or 96.3% of all judgments made were in agreement with the original classification of type IV errors.

Due to a special interest on the part of one of the four testers, in addition to recording recall responses manually, this tester also systematically tape-recorded the subjects' complete performances. Unfortunately these data were insufficient in the present study to reliably analyze the possible relationships between aspects of the subjects' verbal productions in training and their recall. However, several excerpts from these tapes will be utilized as illustrative examples. An examination of the complete records from several subsequent studies utilizing the procedures of this experiment (cf. Buium & Turnure, Note 4) have substantiated the facts and principles addressed in the examples presented here.

Results

Table 2 presents the recall data for both normal and EMR subjects in the six conditions. A 6x2 (Conditions X Subject Category)
Table 2

Means and Standard Deviations of Number Correct in Six Experimental Conditions and Percentage of Recall

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Normals</th>
<th></th>
<th>EMRs</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\bar{X}$ (SD)</td>
<td>$%$</td>
<td>$\bar{X}$ (SD)</td>
<td>$%$</td>
</tr>
<tr>
<td>Labeling</td>
<td>1.3 (1.0)</td>
<td>6.2</td>
<td>2.0 (1.8)</td>
<td>9.5</td>
</tr>
<tr>
<td>Sentence Generation</td>
<td>3.0 (1.6)</td>
<td>14.3</td>
<td>4.0 (3.5)</td>
<td>19.0</td>
</tr>
<tr>
<td>Sentence Repetition</td>
<td>8.3 (3.3)</td>
<td>39.5</td>
<td>8.4 (4.4)</td>
<td>40.0</td>
</tr>
<tr>
<td>Response to &quot;What&quot;</td>
<td>13.5 (5.0)</td>
<td>64.3</td>
<td>13.4 (3.9)</td>
<td>63.8</td>
</tr>
<tr>
<td>Response to &quot;Why&quot; A</td>
<td>16.1 (3.4)</td>
<td>76.7</td>
<td>15.9 (2.2)</td>
<td>75.7</td>
</tr>
<tr>
<td>Response to &quot;Why&quot; B</td>
<td>12.3 (3.7)</td>
<td>58.6</td>
<td>13.9 (3.8)</td>
<td>66.2</td>
</tr>
</tbody>
</table>
analysis of variance of the number correct data revealed that Conditions was the only significant factor, $F(5, 84) = 48.41$, $p < .001$.

Clearly, the normal and EMR children performed at equivalent levels within each condition in the present study (see Figure 1, appended).

Inspection of the data in Table 2 further indicates that the children in the three question conditions ("What", "Why" A, and "Why" B) performed better than children in the other three conditions. This observation was confirmed by a Newman-Keuls test, where it was found that the number correct in each of the three question conditions was higher than in the Labeling, Sentence Generation, or Sentence Repetition Condition (all $p$s < .01). In addition, the Newman-Keuls test indicated that none of the three question conditions led to performances different than in the other question conditions.

Furthermore, the Sentence Repetition Condition did lead to performances which were better than in the Labeling and Sentence Generation conditions ($p$s < .01). The Labeling and Sentence Generation Conditions did not differ in number correct.

The total numbers of non-correct responses (including failures to respond) made by both normal and EMR subjects in the six conditions are presented in Table 3. Also included in this table is a breakdown of these non-correct responses into Semantic errors, Non-semantic errors, and non-responses.

A 6x2 (Conditions X Subject Category) analysis of variance of the percentage of Semantic errors indicated that Conditions was the only significant factor, $F(5, 84) = 5.17$, $p < .01$. Further analysis by means of a Newman-Keuls test for differences suggested that subjects
Table 3

Frequency of Non-Correct Responses and Percentages of Semantic Errors, Non-Semantic Errors, and Non-Responses

NORMAL SUBJECTS

<table>
<thead>
<tr>
<th>Condition</th>
<th>Total Non-Correct</th>
<th>% Semantic Errors</th>
<th>% Non-Semantic Errors</th>
<th>% Non-Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labeling</td>
<td>118</td>
<td>10.2</td>
<td>61.9</td>
<td>28.0</td>
</tr>
<tr>
<td>Sentence Generation</td>
<td>106</td>
<td>16.0</td>
<td>64.2</td>
<td>19.8</td>
</tr>
<tr>
<td>Sentence Repetition</td>
<td>76</td>
<td>26.3</td>
<td>44.7</td>
<td>28.9</td>
</tr>
<tr>
<td>Response to &quot;What&quot;</td>
<td>45</td>
<td>31.1</td>
<td>33.3</td>
<td>35.6</td>
</tr>
<tr>
<td>Response to &quot;Why&quot; A</td>
<td>29</td>
<td>44.8</td>
<td>31.0</td>
<td>24.1</td>
</tr>
<tr>
<td>Response to &quot;Why&quot; B</td>
<td>51</td>
<td>29.4</td>
<td>29.4</td>
<td>41.2</td>
</tr>
</tbody>
</table>

EMR SUBJECTS

<table>
<thead>
<tr>
<th>Condition</th>
<th>Total Non-Correct</th>
<th>% Semantic Errors</th>
<th>% Non-Semantic Errors</th>
<th>% Non-Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labeling</td>
<td>187</td>
<td>20.9</td>
<td>57.2</td>
<td>21.9</td>
</tr>
<tr>
<td>Sentence Generation</td>
<td>170</td>
<td>14.1</td>
<td>34.7</td>
<td>51.2</td>
</tr>
<tr>
<td>Sentence Repetition</td>
<td>125</td>
<td>23.2</td>
<td>48.8</td>
<td>28.0</td>
</tr>
<tr>
<td>Response to &quot;What&quot;</td>
<td>76</td>
<td>35.5</td>
<td>34.2</td>
<td>30.3</td>
</tr>
<tr>
<td>Response to &quot;Why&quot; A</td>
<td>51</td>
<td>43.1</td>
<td>9.8</td>
<td>47.1</td>
</tr>
<tr>
<td>Response to &quot;Why&quot; B</td>
<td>71</td>
<td>36.6</td>
<td>15.5</td>
<td>47.9</td>
</tr>
</tbody>
</table>
in the three question conditions made significantly more Semantic errors than subjects in the Labeling and Sentence Generation conditions \( (p < .05) \). Only the "Why" A condition led to significantly more Semantic errors than the Sentence Repetition condition \( (p < .05) \).

A 6x2 (Conditions X Subject Category) analysis of variance of the percentage of Non-semantic errors again revealed only a significant Conditions effect, \( F(5,84) = 7.62, p < .01 \). A Newman-Keuls test for differences indicated that subjects in all three question conditions made fewer Non-semantic errors than subjects in the Labeling condition \( (p < .05) \). Subjects in the two "Why" conditions also exhibited fewer Non-semantic errors than subjects in the Sentence Generation condition \( (p < .05) \). Only the "Why" A condition led to significantly fewer Non-semantic errors than the Sentence Repetition condition \( (p < .05) \).

Further processing of the differences between the percentages of Semantic and Non-semantic errors within conditions was done by means of related t tests. These tests indicated that significant differences existed between the proportion of Semantic and Non-semantic errors in the Labeling, Sentence Generation and "Why" A conditions. In both the Labeling, \( t(15) = 7.71 \), and the Sentence Generation, \( t(15) = 4.80 \), conditions, significantly more Non-semantic than Semantic errors were made \( (both ps < .001) \). In the "Why" A condition, subjects made significantly more Semantic than Non-semantic errors, \( t(15) = 3.16, p < .01 \). Related t tests did not reveal any significant differences in the Sentence Repetition, \( t(15) = 1.79 \); What, \( t(15) = .89 \); or Why B, \( t(15) = 1.70 \), conditions.
Discussion

The data reported above have clear implications regarding the production deficiency hypothesis as applied to young children's verbal learning. To the extent that the findings of Jensen and Rohwer (1963, 1965) can be considered representative of the empirical basis for production deficiency arguments in the area of verbal learning, the present findings constitute clear evidence of the insufficiency of that hypothesis. While the results replicated the pattern of their findings for the Labeling, Sentence Repetition, and Sentence Generation conditions, the overwhelming superiority of the performances exhibited by subjects in the interrogative conditions of this study, even compared to Sentence Repetition (i.e., the previous basis for rejecting the mediation deficiency hypothesis), clearly demonstrates that certain data used to infer and support a production deficiency hypothesis are task or situation specific. It should be pointed out that by replicating the Jensen and Rohwer (1963, 1965) patterns of results with the present samples of subjects and with comparable conditions, the results provide empirical justification for arguing against a production deficiency hypothesis on standards of performance grounds, and for generalizing our findings to populations of children previously characterized as "production deficient" on the basis of findings similar to those of Jensen and Rohwer.

Clearly, had our recall data been based solely on the Labeling, Sentence Generation and Sentence Repetition conditions, the present results would also have supported a production deficiency type argu-
ment, namely that young children and EMRs cannot produce effective verbal mediators. In fact, the type of sentences the present subjects produced in response to the "make up a sentence" instructions were of a conjunctive nature (e.g., the soap and the jacket). It is the results from the Sentence Generation condition and others like it that have led investigators to characterize young children and EMRs as production deficient, and so not to be expected to produce and employ effective verbal mediators without proper training (Flavell, 1970; Kellas, Ashcraft, & Johnson, 1973; MacMillan, 1970; Martin, 1967; Rohwer, 1973). The present results suggest either that young children in such conditions do not understand the instructions, or that such instructions inhibit their use of imagination or their creative generative use of language (cf. Johnson, Note 3). Thus, the kinds of verbal mediators they produce appear to be insufficient given the requirements of the semantic-based organization of memory model (Jenkins, 1974a). Within the semantic-based (Jenkins, 1974a) or relations-based (Asch, 1969) organization of memory model, the key issue in verbal elaboration research appears to be the extent to which the subject is induced by the experimental conditions to integrate the two paired items via "shared meaning" (Rohwer, 1973) or in a semantic or meaningful relationship (Turnure, 1971). Such relations then function as effective verbal mediators in enhancing correct recall (Turnure & Thurlow, 1973).

Further inspection of the kinds of mediators produced in this study suggests that the three conditions which required children to respond to WH interrogatives induced the subjects (a) to perform
semantic analysis on the paired-associates in the context of the
question associating them, (b) to integrate the items semantically,
and (c) to express this semantic integration verbally through a
sensible implication, as in the following examples:

Experimenter: (Holding pictures of soap, jacket)
"What is the soap doing under the
jacket?"
EMR Subject: "A lady's back there washing um
washing um the coat in the back."

Experimenter: (Holding pictures of light, shoes)
"What is the light doing to the shoes?"
EMR Subject: "Putting a light inside the shoes to
see where a where a tag is what size
shoes they were."

Our records (see also Buium & Turnure, Note 4) of children's responses
to the WH interrogatives clearly demonstrate that young children and
EMRs can produce the kind of verbal responses that constitute effective
verbal mediators. Indeed, one of the strengths of the questioning
techniques employed in this study is that they lend themselves to a
very natural induction of overt responding, with consequent records
of young children's "semantic processing," but do so within the con-
trolled confines of the paired-associate task. It should be empha-
sized here that these records of the children's well-formed responses
are prima facie, logical evidence of production capability. The
present findings, then, show that conditions differ in the extent
to which they induce young children to perform semantic analyses and
to relate these semantic linkages in verbal mediators. Thus, the
results suggest that certain of the putative "production deficiencies"
of young normal and EMR children may be more appropriately character-
ized as "instructional deficiencies."
The effectiveness of the "promotive usage" of interrogatives appears to raise serious questions regarding the adequacy of Rohwer's recent (1973) "prompt" scheme (cf. Flavell, 1970), whereby he has attempted to organize the results of much of the previous research on verbal mediation and elaboration along a dimension reflecting the explicitness to which task conditions orient subjects to generate an event serving as a common referent for any two item pairs (see also Milgram, 1968a). There are two interrelated reasons for this. The first is that his scheme does not provide a category that would encompass interrogatives, although it appears they would fit closest to the augmented explicit prompt. The augmented prompt provides the subject with an "event" that can serve as a common referent for every set of items he is to couple. This prompt type is the basis for the Sentence Repetition condition of the present study. It is obvious that the Sentence Repetition condition and the various interrogative conditions vary in diverse ways, in terms of both their linguistic and psychological demands. They also differ quite markedly in their response demands, and this observation leads directly to the second reason for questioning the applicability of Rohwer's scheme in relation to the effectiveness of interrogatives in enhancing paired-associate learning.

In Rohwer's formulation, the various instructional manipulations defining the prompt types appear to be necessarily bound to the nature of the children's responses. That is, antagonistic prompts require "irrelevant" responses which by definition are inimical to learning.
Minimally explicit prompts are as nonconstraining as is possible (see below regarding the "cue" effects of formal situations), but as such are totally uncontrolled in substance, and are therefore uninterpretable. (Our colleague Arthur Taylor, Note 5, refers to this situation as the "Do your own thing" condition.) Explicit prompts specifically direct the subject to create a referential event encompassing the item pairs: The Sentence Generation condition of this study exemplifies this prompt instruction. Responses to such instruction are necessarily restricted to precise expressions of the demanded behavior. Augmented explicit prompts have been described above (Sentence Repetition is an example), and responses to this prompt are either active overt repetition of the prompt or passive-covert "apprehension" of same. Maximally explicit prompts entail involving the subject in an enactment of an event and the response required is, therefore, the "experience" of the subject. Nowhere in Rohwer's scheme is there room for a "sort-of-augmented prompt," which requires as a response a sensible continuation, extension, or explanation of an event, as is the case for the Wh questions.

The rather straightforward interpretation of the effective "promotive usage" of interrogatives advanced so far may appear to founder when scrutinized in the light of various aspects of an obvious question concerning the extent to which the semantic contents of our interrogatives themselves "cue" the responses of the children. In terms usually employed in discussing mediational and production deficiency models, this question would pertain to the degree to which
any "cognitive strategy" or "mnemonic mediation" employed by a subject in the service of his performance was a "spontaneous" gesture on his part. Flavell (1970) has put the case clearly: "A production deficiency versus some other characterization of the child's performance seems warranted in direct proportion to one's conviction that this particular child could have produced the target mediator spontaneously, and that his failure to emit it is virtually all that one can find to differentiate him from a producing child" (p. 198, italics in original). While Flavell's analysis then proceeds to discuss numerous interesting developmental complexities regarding the transition of such a child to a "production sufficient" status, it would appear more directly to the point here to grapple, however inconclusively, with the basically undefined notion of "spontaneous production."

The term "spontaneous" is usually applied to such manifestations of children's behaviors as are emitted in the absence of explicit and direct demands or inducements for the appearance of those behaviors. In "mediated memory" tasks, this generally means that some particular technique known to be efficacious for the acquisition or retention of task materials is withheld or not revealed to certain subjects, with the aim of determining if such subjects will, nevertheless, give evidence of utilizing said technique, or similar, sans "cue." However, it is readily apparent that in "appropriate memory tasks" it is not the case that there is no stimulus, since the experimenter has contrived to confront the child with some sort
of task with its attendant materials, procedures, general inducements, and constraints. Thus, in some sense there is always some "cue" influencing the subject.

Without going to metaphysical lengths about it, the foregoing observations conform to Donaldson's (1970) definition of a "formal situation," and support the validity of her observation that in an experimental (as opposed to naturalistic) approach to the study of cognitive development "an inescapable complication arises, for [such studies] demand not only competence in respect of the behaviour which the instruction is meant to elicit but competence in the very business of responding 'to order'" (p. 397). Donaldson's (1970) comments on the complexities abounding in the experimental analysis of language competence had been presaged by the somewhat more specific observations of Jenkins (1967), which he addressed explicitly to the paired-associate learning task, as actually "a series of tasks which must be accomplished by the subject" (p. 48), with the initial task facing the subject as one of understanding the requirements of the task and getting a "feel" for the procedures (see also Turnure, 1971, p. 311; Note 6, p. 7). This analysis of the ambiguity facing the young child during his first exposures to "formal situations" appears to be quite congruent with Flavell's (1970) distinction regarding the development of "general cognitive factors," such as appreciation of the need for planful and preparatory activities, as compared with specific factors, such as the actual mechanisms of rehearsal, clustering and the like. In essence then, the argument here is that by
engaging the child's propensity for responding appropriately to effective questioning we have circumvented the constraining effects entailed in such "formal testing situations," and thereby illuminated the child's competence in "naturally elaborating" on informational input, including all the cognitive processing that that implies. Thus, one aspect of the production deficiency argument based on an artifactual and conceptually confusing constraint masking children's capability to "spontaneously" elaborate may have been nullified.

In concluding this section on the nature of "spontaneous production," it appears clearly relevant to at least allude to the compelling arguments regarding the novel, creative, or generative characteristics of most of the language children produce. Inasmuch as our subjects did not merely repeat the "prompts" introducing the item pairs to them, but created reasonable new (and often very novel) contexts wherein to perceive additional relations among the pairs, their responses can be fairly characterized as generative, and so spontaneous, at least to some degree. In effect, these sensible extensions of the implications of a question pertaining to an item pair demonstrate that the items in context have been actively processed at the level of understanding (Craik, 1973). Therefore, at a minimum, the performances of the children in the interrogative conditions constitute intriguing indicators of further processes that may be related to cognitive growth and memory functioning in ways that are more "natural" than those investigated in studies of rehearsal, clustering, "pegging," or other formal mnemonic devices.

The point here is that it appears unlikely that the children in
this study were deliberately or voluntarily intending to remember the items and so strategically answered the questions in a way which would serve as a mnemonic mediator enabling them to do so (cf. Brown, 1974). In its broad implications, this argument appears closely related to Jenkins' (1974b) formulation of "contextualism," particularly in regard to emphasizing the "event" as interpreted by the subject, as opposed to the "machinery" utilized merely to remember the items. The virtually universal custom of questioning children (Hesse, Turnure & Buium, Note 7) as a form of general social discourse, or more formally and didactically during tuition, suggests that the pervasive and cumulative impact of such activities during the child's experiential history may well be a major developmental impetus to progressive changes in general and specific cognitive factors (cf. Flavell, 1970), and, in broad terms, to increases in the child's "processing space" (cf. Pascual-Leone, 1970; Rohwer, 1973). One implication from these speculations is to remind psychological theorists that making more than descriptive reference to "spontaneous" functioning of any kind is a dubious explanatory device. The weakness of such explanations has been discovered by workers in chemistry and physics who relied on a conception of "spontaneous combustion" to "explain" the unprovoked igniting of diverse materials, and in biology where Pasteur effectively refuted the doctrine of "spontaneous generation" just over 100 years ago (cf. Moulton & Schifferes, 1945, especially p. 429).

The recent resurgence of interest in applying transfer-of-training to memory processes (Campione & Brown, 1974; Goulet, 1970, 1973),
particularly in regard to general or nonspecific transfer effects, indicates at least a general awareness of the problem. It only needs mentioning that the study of transfer poses complicated methodological and logical problems (cf. Woodworth & Schlosberg, 1954, especially Chaps. 24 & 26), not the least of which is establishing a realistic relationship between the "transfer effect" measured and the "transfer" that produces it (Woodworth & Schlosberg, 1954, p. 734). In other words, tracing the process pertaining to the performance (Woodworth & Schlosberg, 1954) will, at a minimum, require demonstrating, for instance, the operation of cognitive structures, mechanisms, or strategies that are generally only vaguely defined, and, possibly, inherently unobservable. The appalling lack of empirical evidence pertaining even to the transfer of general principles applicable to the solution of some whole class of problems (compare, for instance, Travers, 1967, ch. 8 "Transfer of Training" on the studies of Bagley, 1905, Hendrikson & Schroeder, 1941, and Judd, 1908), which would appear to be a more specific and readily operationalizable research problem than studies of strategies, seems to reflect the diffidence with which researchers have treaded in the domain of transfer of other than simple S-R relationships.

The other condition effects in the present study appear to be congruent with the semantic analysis interpretation. We may note that the Sentence Repetition condition is theoretically significant in the sense that it provides the children with a semantic relation between the two paired-associates, yet the child's repetition of the
mediator may or may not reflect satisfactory semantic integration on his part. In another study (Buium & Turnure, Note 4), it was found that a "response to yes-no interrogatives" condition produced performance similar to a Sentence Repetition condition, and on the grounds that yes-no questions shift the "cognitive load" from responder to speaker, advanced a similar lack of semantic integration argument. If the subjects had been involved in constructing the relation in a way that was meaningful to them, or had been provided more relations (Turnure & Thurlow, 1973), or it had been insured otherwise that the children comprehended the meaning of the elaborations encompassing the paired items, they might have been at a lesser disadvantage.

The Wh type interrogative conditions appear to have induced the subjects to search, relate, analyze and integrate the items' most appropriate conceptual features. It is in this light that the high proportion of the semantic errors in an otherwise relatively low number of total errors (Table 2) is understood. Conversely, the Sentence Repetition and Labeling conditions, respectively, necessitate less in the way of semantic analyses, thus resulting in respectively fewer semantic errors (see Table 2). From a semantic (Jenkins, 1974a) or relation (Asch, 1969) based organization of memory model, the Labeling condition appears to impose no semantic relation between the paired-associates, and thus the low correct score (Turnure, 1971).

In the present study, the performance levels of the normal and the EMR children in each of the six conditions were quite similar.
Other recent studies comparing the performances of MA-matched normal and EMR children on verbal elaboration tasks have also found performance levels to be quite similar, at least under certain conditions. For example, Turnure, Thurlow and Larson (Note 8) compared the performances of normal and EMR subjects (MAs of about 7 years) on a 4-pair task under labeling and elaboration conditions. They found no significant differences between the number of trials taken by each subject category to reach a criterion of two errorless trials. In an 8-pair task, however, the comparison of normals and EMRs of MA = 7 years indicated that the difference between the performance levels was statistically significant under elaboration conditions, with the normal subjects requiring an average of one less trial to reach criterion (X̄ = 3.17) compared to the retarded subjects (X̄ = 4.17). The practical significance of the difference was questionable, however, since it reflected only a difference of one trial in learning to criterion by the two groups.

Subsequent studies have looked at the relative performance of CA-matched normals and EMRs. In one study (CA = 9), the effects of Compound sentence, Complex sentence, and Paragraph elaborations on a 12-pair list were compared (Turnure & Thurlow, in press). Analysis of the number of first trial errors made by each subject category (Normals: X̄ = 3.00; EMRs: X̄ = 4.17) did not produce any significant differences between the groups. Trials to criterion data in the same study did reveal a difference, with normals requiring fewer mean trials to reach criterion than EMRs (Normals: X̄ = 4.07; EMRs: X̄ = 5.57). The difference, however, was directly attributable to
the inadequate performances of three EMR children, children who possibly were brain-damaged and therefore not to be expected to perform as well as the other EMR children (cf. Turnure, Larsen & Thurlow, 1973). In a second study (CA = 8.5), the effects of four types of elaborations were studied (Declarative sentences, Interrogative sentences, Declarative paragraphs, Interrogative paragraphs). In this study (Turnure & Thurlow, Note 9), significant differences did emerge between the first trial error performances of normals and EMRs. This significant subject category effect, however, was clearly attributable to the poor performances of the EMR subjects in the Interrogative Sentence Condition, the condition in which these children were expected to do less well.

Other studies have, similarly, found (Jensen, 1965; Milgram, 1968b; Ring, 1965) and not found (Baumeister & Campbell, 1971; Milgram, 1968b; Ferguson, 1964) differences between EMRs and normals in verbal learning (cf. Goulet, 1968, for a more extensive review). One probably must conclude from the studies summarized here that the emergence of a "subject category effect" depends upon several situational variables, and vagaries in sampling from the heterogeneous population of mentally retarded children, as well as on the interaction of these with the effects or different treatments. The equivalent performances of MA-matched normal and retarded children tend to support a developmental as opposed to a defect interpretation of mental retardation (cf. Zigler, 1973), and conform to Milgram's (1973) observation that "a moderately retarded individual will display..."
adequate linguistic competence, although he may acquire it at a
retarded rate" (p. 165). Beyond this, the interpretation of the
condition effects offered here would appear to hold the same impli-
cations for students of mental retardation as they do for child
development in general, and may be even more pertinent given the
contemporary tendency to attribute all manner of retardate
inadequacies to "production deficiencies" (cf. Brown, 1974;
Reference Notes


2. Samuels, S. J. Personal communication.

   (Available from Michael G. Johnson, Department of Psychology, University of Tennessee, Knoxville, Tennessee 37916.)


5. Taylor, A. M. Personal communication.


References


Figure 1: Mean Number Correct by Normals and EMRs in Six Experimental Conditions.
Labeling Sentence Repetition Generation

What Why A Why B

CONDITION

* Normals exceed EMR's

MEAN NUMBER CORRECT


P. Krus. Use of family history data to predict intellectual and educational functioning longitudinally from ages four to seven. Research Report #64. June 1974.


D. Moores. Recent research on manual communication. Occasional Paper #7. April 1971. (Keynote Address, Division of Communication Disorders, Council for Exceptional Children Annual Convention, Miami Beach, April 1971.)


