Armbruster, Bonnie B.; And Others

National Inst. of Education (DHEW), Washington, D.C.

TE-26
Mar 77
400-76-0116
22p.

MP-$0.83 BC-$1.67 Plus Postage.

Comparative Analysis; *Course Content; Grade 3; Primary Education; *Reading Comprehension; *Reading Research; *Reading Skills; *Reading Tests; *Textbook Content

*Center for the Study of Reading (Illinois)

Content related to reading comprehension, in three nationally used third grade reading curricula and two common standardized tests, was analyzed in order to discover congruencies and incongruencies in content coverage and emphases. Frequencies of exercises from the curriculum materials and of items from the standardized tests were calculated for 16 subcategories of reading comprehension. Results indicated that the texts differed widely in their relative emphases on reading comprehension and on particular reading comprehension subcategories. The two standardized tests were quite similar in relative emphasis on reading comprehension, but their emphases were quite different from those of the texts. Only a small percentage of the skills emphasized in the curricula had counterparts on the standardized tests, and those tended to reflect factual items entailing locating information in presented materials, whereas two of the three curricula gave heavier emphasis to inferential comprehension skills. (AA)
Technical Report No. 26

Analyzing Content Coverage and Emphasis: A Study of Three Curricula and Two Tests

Bonnie B. Armbruster, Robert J. Stevens, and Barak Rosenshine

University of Illinois at Urbana-Champaign

March 1977

The research reported herein was supported in part by the National Institute of Education under Contract No. HS-11IE-C-400-76-0116.

*Designation as a Center pending approval.
Theoretical Framework

The conclusion emerging from recent studies of the relationship between classroom processes and student achievement is that "major factors in the process of knowledge acquisition in the classroom are the content and emphasis of the curriculum in use" (Berliner & Rosenshine, 1976). For example, studies by Armento (1975), Chang and Raths (1971), Rosenshine (1963), and Shutes (1969) found significant positive correlations between content covered and achievement. The related variable of content emphasis was found to be correlated with achievement in studies reviewed by Walker and Schaffarzick (1974) and in a major study of 166 low SES classrooms (Stallings & Kaskowitz, 1974). Such studies show that content coverage and emphasis are at least as important as teaching method for student achievement. Yet research analyzing and comparing the content of various curricula in the attempt to discover specific variables affecting achievement is rare.

Achievement is a function of content coverage and emphasis of curricula, but the measures of achievement are in turn dependent on the particular content and emphasis of the instrument employed. As Walker and Schaffarzick (1974) show, the demonstrated effectiveness of curricula (measured by "achievement") is a function of the content
of the posttest: each curriculum in the studies reviewed seemed to be most effective when the posttest reflected the content emphasis of that curriculum. Therefore, achievement is most accurately defined with respect to curricula. To make judgments of effectiveness, researchers must know both what the curriculum teaches and whether the tests adequately measure what is taught.

Objectives

The purpose of the present study was to analyze part of the content of three nationally used reading curricula and two common standardized tests in order to discover congruencies and incongruencies of content coverage and emphases among curricula, between tests, and between curricula and tests. The particular content of interest was reading comprehension.

Methods

Three reading curricula designed for the second half of the third grade were chosen for analysis:

2) Ginn and Company: Reading 360: All Sorts of Things (Level 10), 1969.

Two standardized tests judged to be in common use as measures of achievement in educational research were also selected (forms appropriate for the third grade were employed):
Analyzing Content Coverage and Emphasis


In order to have a measure of content coverage and emphasis comparable across curricula and tests, it was decided to obtain frequencies of exercises from the curriculum materials and items from the tests in several subcategories of reading comprehension. Sixteen categories were adapted from SRA's SOBAR Reading Catalog of Objectives 3-9 (1975) and were operationally defined as follows:

**Detail:** The reader answers questions that either have the exact same surface form as single text sentences or that have the same surface form except for pronoun substitutions.

**Paraphrase Level 1:** The reader answers questions that have different surface forms but the same meaning as single text sentences; i.e., questions and text sentences match except for synonym substitutions.

**Paraphrase Level 2:** The reader answers questions that have different surface forms but the same meaning as two or more text sentences.

**Cloze Sentences:** Given a sentence with a word deleted, the reader selects the appropriate word from several alternatives.

**Classifying:** Given a set of reading passages, the reader identifies similar passages according to some criterion.
Following Directions: Given a set of written directions, the reader performs the indicated task.

Sequence: The reader orders presented events into a sequence matching the presented sequence.

Drawing Conclusions: The reader answers questions calling for a conclusion based on the material in the reading selection.

Main Idea: The reader answers questions calling for identification of the main idea and/or an appropriate title.

Supporting Information: The reader identifies or supplies subordinate topics given the main topic in a content outline.

Cause and Effect: The reader answers questions concerning cause-and-effect relationships.

Words in Context: Given a sentence containing context clues to the meaning of an unknown word, the reader selects the appropriate meaning from two or more alternatives.

Figurative Language: The reader identifies similes and metaphors in a reading selection.

Fantasy - Reality: The reader identifies a reading selection as either a representation of fantasy or reality.

Mood - Setting: The reader answers questions concerning the mood or setting of a reading selection.

Character's Emotions and Traits: The reader answers questions concerning the emotions and traits of characters in a reading selection.
The authors classified and recorded frequencies of reading comprehension items in the curriculum materials (teacher's manual and workbook) and tests. For the curricula, only written exercises designed to be completed by all students were coded. Oral exercises, questions meant for class discussion, and individualized activities were not coded because these exercises typically do not involve every student. An interrater reliability (percent agreement) of .81 was established on the basis of three blocks of forty items randomly selected from the three curricula. Since this reliability is acceptably high, the final item counts were taken to be the average of the item counts of the three raters.

Results and Conclusions

The texts differ widely in their relative emphasis of reading comprehension in general (Table 1) and of particular reading comprehension categories (Table 2). Economy stresses cloze sentences, supporting information, and sequence; Ginn gives by far the most emphasis to cloze sentences; and Houghton-Mifflin emphasizes words in context and cloze sentences. Table 3 presents another way of looking at the differences among texts. Detail, paraphrase level 1, paraphrase level 2, cloze sentences, and following directions were judged to be literal comprehension items, while all other categories appeared to involve inferential comprehension. According to this classification,
Economy and Houghton-Mifflin give almost the same relative emphasis to literal and inferential comprehension, but Ginn emphasizes literal items much more, largely because of its heavy emphasis (about half of the items) on cloze sentences.

Correlations were computed (based on rank order of emphasis given to different categories) between the curricula based on all the categories (Table 4) and also on the six categories the curricula had in common with the two tests (detail, paraphrase level 1, paraphrase level 2, drawing conclusions, main idea, and character's emotions and traits) (Table 5). Economy and Houghton-Mifflin were moderately correlated in both tables ($r_s = .4$), but the correlations between Economy and Ginn and Ginn and Houghton-Mifflin vary from low positive to low negative depending on whether one uses all the reading comprehension categories (Table 4) or just the main categories used in the reading comprehension tests (Table 5). Based on these data, the three curricula appear moderately to highly distinct.

The two standardized tests are quite similar in relative emphasis of reading comprehension, yielding a high positive rank-order correlation (+.93). For the six categories common to all texts and tests, the rank
order correlations between texts and tests were low (Table 6). The Ginn series and the MAT correlated only +.10; the highest correlations were between Houghton-Mifflin and the two achievement tests, but even these correlations were in the low .40's. This suggests that what is emphasized on the tests is different from what is emphasized in the texts.

A further finding reflecting differences between texts and tests is the fact that a large percentage of the comprehension items taught are not tested on the standardized tests (Table 2). Out of the 16 reading comprehension categories that are covered in one or more texts, only six are tested on the MAT and seven on the CAT. From another perspective, 64% of the Economy, 65% of the Ginn, and 79% of the Houghton-Mifflin reading comprehension items do not have counterparts on the standardized tests.

Approximately two-thirds of the reading comprehension items on both the MAT and CAT were categorized as detail, paraphrase level 1, and paraphrase level 2. According to our classification, these categories involve literal comprehension. As seen from Table 3, however, two of the three texts (Economy and Houghton-Mifflin) stress inferential categories. Detail, paraphrase level 1, and paraphrase
level 2 comprise only 18% of Economy's, 28% of Ginn's, and 13% of Houghton-Mifflin's items.

Discussion

The most significant finding of this study is the large discrepancy between reading comprehension skills taught and tested. Only a small percentage of skills emphasized in the curricula have counterparts on the standardized tests. Furthermore, the skills in the tests tend to be factual items entailing locating information in the presented text, whereas two out of three curricula give heavier emphasis to comprehension skills that appear to require inference, interpretation, identification of relationships, and synthesis.

The study also revealed high variation among the texts but high consistency between the tests for relative content coverage.

Judging from the correlations between texts and tests, Houghton-Mifflin was more related to the standardized tests than Ginn or Economy. Such a table of correlations might enable one to determine which curriculum is the most appropriate preparation for a particular standardized test, or alternatively, which test is the best measure of the material covered in a particular curriculum. However, such use of a table of correlations is not yet justifiable, for we do not presently know whether the comprehension categories are psychologically distinct. If they are, then the categories which are taught and probably the
sequence of training becomes important. However, it may be that reading comprehension is a global or general ability, with no component subskills. In that case, the choice of particular exercises would be irrelevant; all are equally suited to developing and testing a general reading comprehension ability. In sum, we do not yet know whether a mismatch between what is taught and what is tested in reading comprehension has any practical significance.

Despite unanswered questions, the present study is important in its demonstration of a feasible methodology for addressing a long-neglected research problem—determining content coverage and content emphasis of both curricula and tests. More such studies comparing curricula and tests in different content areas and grade levels are needed.
References


Table 1

Relative Emphasis on Reading and Reading Comprehension
for Three Curricula and Two Tests

<table>
<thead>
<tr>
<th></th>
<th>Curricula</th>
<th>Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Economy Ginn</td>
<td>Houghton-Hiifflin</td>
</tr>
<tr>
<td>Total reading exercises</td>
<td>3,060  2,425  1,134</td>
<td>145     85</td>
</tr>
<tr>
<td>Total reading comprehension exercises</td>
<td>333    892    614</td>
<td>24      35</td>
</tr>
<tr>
<td>Percent reading comprehension exercises</td>
<td>27.2    36.0    54.1</td>
<td>16.6    41.2</td>
</tr>
</tbody>
</table>
Relative Emphasis of Categories of Reading Comprehension for Three Curricula and Two Tests

<table>
<thead>
<tr>
<th>Category</th>
<th>Economy</th>
<th>Ginn</th>
<th>Houghton-Mifflin</th>
<th>NAT</th>
<th>CAT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
<td>Percent</td>
<td>Number</td>
</tr>
<tr>
<td>Number of Items</td>
<td>Items</td>
<td></td>
<td>Items</td>
<td></td>
<td>Items</td>
</tr>
<tr>
<td>Detial</td>
<td>32</td>
<td>4</td>
<td>56</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Paraphrase 1</td>
<td>60</td>
<td>6</td>
<td>60</td>
<td>6</td>
<td>64</td>
</tr>
<tr>
<td>Paraphrase 2</td>
<td>66</td>
<td>8</td>
<td>122</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>Cloze Sentences</td>
<td>300</td>
<td>28</td>
<td>451</td>
<td>51</td>
<td>169</td>
</tr>
<tr>
<td>Classifying</td>
<td>22</td>
<td>3</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Following directions</td>
<td>6</td>
<td>1</td>
<td>13</td>
<td>5</td>
<td>--</td>
</tr>
<tr>
<td>Sequence</td>
<td>37</td>
<td>12</td>
<td>43</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td>Drawing Conclusions</td>
<td>66</td>
<td>16</td>
<td>17</td>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>Main Idea</td>
<td>83</td>
<td>5</td>
<td>12</td>
<td>1</td>
<td>26</td>
</tr>
<tr>
<td>Supporting Information</td>
<td>122</td>
<td>15</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Cause and Effect</td>
<td>--</td>
<td>--</td>
<td>15</td>
<td>2</td>
<td>--</td>
</tr>
<tr>
<td>Words in Context</td>
<td>10</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>220</td>
</tr>
<tr>
<td>Figurative Language</td>
<td>20</td>
<td>3</td>
<td>9</td>
<td>1</td>
<td>56</td>
</tr>
<tr>
<td>Fantasy-Reality</td>
<td>19</td>
<td>2</td>
<td>13</td>
<td>1</td>
<td>--</td>
</tr>
<tr>
<td>Mood-Setting</td>
<td>25</td>
<td>3</td>
<td>--</td>
<td>--</td>
<td>4</td>
</tr>
<tr>
<td>Character's Emotions-Traits</td>
<td>9</td>
<td>1</td>
<td>32</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>333</strong></td>
<td></td>
<td><strong>692</strong></td>
<td></td>
<td><strong>614</strong></td>
</tr>
</tbody>
</table>
Analyzing Content Coverage and Emphasis

Table 3
Percentages of Literal and Inferential Items for Three Curricula

<table>
<thead>
<tr>
<th>Item type</th>
<th>Economy</th>
<th>Ginn</th>
<th>Houghton-Mifflin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literal</td>
<td>42%</td>
<td>33%</td>
<td>41%</td>
</tr>
<tr>
<td>Inferential</td>
<td>58%</td>
<td>17%</td>
<td>59%</td>
</tr>
</tbody>
</table>
Table 4
Correlations ($r_s$) Between Curricula for All Reading Comprehension Categories

<table>
<thead>
<tr>
<th></th>
<th>Economy</th>
<th>Ginn</th>
<th>Houghton-Illiflin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economy</td>
<td>+1.00</td>
<td>+.31</td>
<td>+.40</td>
</tr>
<tr>
<td>Ginn</td>
<td>+.31</td>
<td>+1.00</td>
<td>+.33</td>
</tr>
<tr>
<td>Houghton-Illiflin</td>
<td>+.40</td>
<td>+.33</td>
<td>+1.00</td>
</tr>
</tbody>
</table>
Table 5
Correlations \( (r_s) \) Between Curricula for Six Reading Comprehension Categories in Common with MAT and CAT

<table>
<thead>
<tr>
<th></th>
<th>Economy</th>
<th>Ginn</th>
<th>Houghton-Mifflin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economy</td>
<td>+1.00</td>
<td>-0.08</td>
<td>+0.43</td>
</tr>
<tr>
<td>Ginn</td>
<td>-0.08</td>
<td>+1.00</td>
<td>-0.14</td>
</tr>
<tr>
<td>Houghton-Mifflin</td>
<td>+0.43</td>
<td>-0.14</td>
<td>+1.00</td>
</tr>
</tbody>
</table>
Table 6
Correlations ($r_s$) Between Curricula and Tests
for Six Common Reading Comprehension Categories

<table>
<thead>
<tr>
<th>Curricula</th>
<th>Tests</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NAT</td>
<td>CAT</td>
<td></td>
</tr>
<tr>
<td>Economy</td>
<td>.21</td>
<td>.37</td>
<td></td>
</tr>
<tr>
<td>Ginn</td>
<td>.10</td>
<td>.31</td>
<td></td>
</tr>
<tr>
<td>Houghton-Mifflin</td>
<td>.41</td>
<td>.43</td>
<td></td>
</tr>
</tbody>
</table>

No. 2: Spiro, R. J. Inferential Reconstruction in Memory for Connected Discourse, October 1975.


No. 4: Alessi, S. M., Anderson, T. H., & Biddle, W. B. Hardware and Software Considerations in Computer Based Course Management, November 1975.


No. 7: Ortony, A. Names, Descriptions, and Pragmatics, February 1976.

No. 8: Mason, J. M. Questioning the Notion of Independent Processing Stages in Reading, February 1976.

No. 9: Siegel, M. A. Teacher Behaviors and Curriculum Packages: Implications for Research and Teacher Education, April 1976.


No. 16: Jenkins, J. R., & Pany, D. Curriculum Biases in Reading Achievement Tests, November 1976.


No. 20: Kleiman, G. M. The Effect of Previous Context on Reading Individual Words, February 1977.


No. 30: Goetz, E. T., & Osborn, J. Procedures for Sampling Texts and Tasks in Kindergarten through Eighth Grade, April 1977.


No. 34: Bruce, B. C. Plans and Social Actions, April 1977.

No. 35: Rubin, A. D. Comprehension Processes in Oral and Written Language, April 1977.


