DOCUMENT RESUME

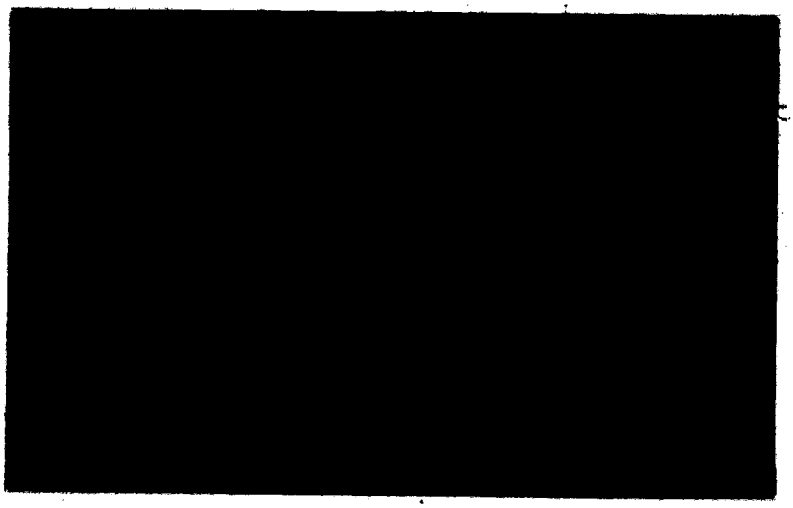
ED 229 446 TM 830 354

AUTHOR Freeman, Donald J.; And Others
TITLE Consequences of Different Styles of Textbook Use in Preparing Students for Standardized Tests.
INSTITUTION Michigan State Univ., East Lansing. Inst. for Research on Teaching.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
REPORT NO IRT-RS-107
PUB DATE Mar 83
CONTRACT 400-81-0014
NOTE 27p.
AVAILABLE FROM Institute for Research on Teaching, College of Education, Michigan State University, 252 Erickson Hall, East Lansing, MI 48824 ($3.00).
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Classroom Techniques; *Content Analysis; Grade 4; Intermediate Grades; Interrater Reliability; Management by Objectives; Mathematics Materials; *Standardized Tests; *Teaching Styles; Test Wiseness; *Textbook Content; *Textbook Research
IDENTIFIERS Holt Rinehart Winston Company

ABSTRACT

Earlier content analyses showed that the match between content covered by textbooks and tests varied as a function of the particular textbook and test a teacher was asked to use. This study tried to determine if the congruity in textbook-test content varied as a function of different styles of textbook use. Using year-long case studies of seven teachers as a guide, five distinct styles of textbook use were identified. These ranged from a page-by-page progression through the book to the selection of only those lessons that conformed to a management-by-objectives system (MBOS). The match between content covered on each of five standardized tests and the lessons covered by each usage style of the Holt fourth-grade mathematics textbook was determined. The results support two important conclusions: (1) when a MBOS serves as the core of an individualized approach to instruction, low achievement students may suffer serious handicaps in their opportunities to learn content covered on standardized tests, and (2) overall levels of student performance on standardized tests of achievement may be relatively insensitive to variation in the content of classroom instruction resulting from differences in how teachers use textbooks. (Author/PN)



Institute for Research on Teaching

College of Education — Michigan State University

Research Series No. 107

CONSEQUENCES OF DIFFERENT STYLES OF TEXTBOOK USE IN PREPARING STUDENTS FOR STANDARDIZED TESTS

Donald J. Freeman, Gabriella M. Belli,
Andrew C. Porter, Robert E. Floden,
William H. Schmidt, and John R. Schwille

Published By

The Institute for Research on Teaching
252 Erickson Hall
Michigan State University
East Lansing, Michigan 48824-1034

March 1983

This work is sponsored in part by the Institute for Research on Teaching, College of Education, Michigan State University. The Institute for Research on Teaching is funded primarily by the Program for Teaching and Instruction of the National Institute of Education, United States Department of Education. The opinions expressed in this publication do not necessarily reflect the position, policy, or endorsement of the National Institute of Education. (Contract No. 400-81-0014)

Institute for Research on Teaching

The **Institute for Research on Teaching** was founded at Michigan State University in 1976 by the National Institute of Education. Following a nationwide competition in 1981, the NIE awarded a second contract to the IRT, extending work through 1984. Funding is also received from other agencies and foundations for individual research projects.

The IRT conducts major research projects aimed at improving classroom teaching, including studies of classroom management strategies, student socialization, the diagnosis and remediation of reading difficulties, and teacher education. IRT researchers are also examining the teaching of specific school subjects such as reading, writing, general mathematics, and science, and are seeking to understand how factors outside the classroom affect teacher decision making.

Researchers from such diverse disciplines as educational psychology, anthropology, sociology, and philosophy cooperate in conducting IRT research. They join forces with public school teachers, who work at the IRT as half-time collaborators in research, helping to design and plan studies, collect data, analyze and interpret results, and disseminate findings.

The IRT publishes research reports, occasional papers, conference proceedings, a newsletter for practitioners, and lists and catalogs of IRT publications. For more information, to receive a list or catalog, and/or to be placed on the IRT mailing list to receive the newsletter, please write to the IRT Editor, Institute for Research on Teaching, 252 Erickson Hall, Michigan State University, East Lansing, Michigan 48824-1034.

Co-Directors: Jere E. Brophy and Andrew C. Porter

Associate Directors: Judith E. Lanier and Richard S. Prawat

Editorial Staff
Editor: Janet Eaton
Assistant Editor: Patricia Nischan

Abstract.

Earlier content analyses conducted by the authors showed that the match between content covered by textbooks and tests varied as a function of the particular textbook and test a teacher was asked to use. The authors also tried to determine if the congruity in textbook-test content varied as a function of different styles of textbook use. Using year-long case studies of seven teachers as a guide, the authors identified five distinct styles of textbook use. These ranged from a page-by-page progression through the book to the selection of only those lessons that conformed to a management-by-objectives system. The authors then determined the match between content covered on each of five standardized tests and the lessons covered by each usage style of the Holt fourth-grade mathematics textbook. Despite clear limits to the generalizability of this study, the results support two important conclusions.

1. When a management-by-objectives system serves as the core of an indivdiualized approach to instruction, low achievement students may suffer serious handicaps in their opportunities to learn content covered on standardized tests.

2. Overall levels of student performance on standardized tests of achievement may be relatively insensitive to variation in the content of classroom instruction resulting from differences in how teachers use textbooks.

CONSEQUENCES OF DIFFERENT STYLES OF TEXTBOOK USE IN PREPARING STUDENTS FOR STANDARDIZED TESTS[1]

Donald J. Freeman, Gabriella M. Belli, Andrew C. Porter, Robert E. Floden, William H. Schmidt, and John R. Schwille[2]

Educators and researchers agree that textbooks greatly influence what is taught. Therefore, it is not surprising that our year-long case studies of seven elementary-school teachers[3] support the results of national surveys: that, for most teachers, textbooks are an important determinant of what is taught (National Committee on Mathematics Education, Note 1). In a series of interview questions focusing on textbook use in fourth-grade mathematics instruction, one case-study teacher remarked, "Most people will teach what is in the book, so what is being taught depends on what is in the math book." In a comparable interview, a principal stated, "The math book is a curriculum, you might say, for the district....The curriculum itself is pretty much oriented to the textbook."

Many believe that in elementary-school mathematics a national curriculum guides the development of textbooks and tests and that when teachers rely on textbooks in deciding what to teach, they provide an opportunity for students to learn what is tested in standardized tests. Similarly, they expect that the commonly used standardized tests of elementary-school mathematics are so similar in content that they may be used interchangeably.

---

1. This paper has been submitted to the Journal of Educational Measurement.

2. Donald Freeman is a professor in the Department of Teacher Education. Gabriella Belli was a research assistant on the Content Determinants Project in the IRT. Andrew Porter is the project's director. Robert Floden is an associate professor in the Department of Teacher Education. William Schmidt is the chairperson of the Department of Counseling, Educational Psychology and Special Education. John Schwille is a professor in the Department of Teacher Education.

3. The results of the case study investigations will be summarized in a forthcoming monograph.

Our previous content analyses of textbooks and tests of fourth-grade mathematics challenge the concept of a national curriculum (Freeman, Kuhs, Porter, Knappen, Floden, Schmidt, & Schwille, Note 2). Using a three-dimensional taxonomy, we found that relatively few topics were consistently emphasized across four textbooks and five tests of fourth-grade mathematics. In fact, only six specific topics were emphasized in all of these sources (e.g., pictorial models of a fraction). All other topics one might attribute to the fourth-grade mathematics curriculum were covered in some textbooks or standardized tests, but not in others.

These analyses also provided evidence that the match in content covered was better for some textbook-test pairs than for others. For example, the proportion of tested topics covered by at least one problem in a book varied from 52.8% to 86.8% for different pairs of textbooks and tests. For even the best matched textbook-test pair, less than 50% of the topics on the test were covered by the equivalent of one lesson in the textbook.

Collectively, these results highlight the need to examine the degree to which teacher reliance on textbooks ensures that students will deal with content that is covered on standardized tests of achievement. The results to date indicate that student level of preparation varies as a function of the particular textbook and test the teacher has been asked to use. The present study investigates the degree to which the match between textbook and test contents varies as a function of a teacher's use of the book.

Models of Textbook Use

David Berliner (1979), suggests that

> Different philosophies of education yield different beliefs about what is important for students to learn. These beliefs, along with the teacher's likes and dislikes for teaching certain areas, result in some interesting differences in the **functional** curriculum of a class. (p. 126)

Differences in beliefs also lead to variations in how textbooks are used. As part of a comprehensive series of interrelated studies of how teachers decide what to teach, the authors conducted year-long case studies of seven elementary-school teachers. These

teachers demonstrated four district styles of textbook use: textbook bound, selective omission, focus on basics, and instruction suggested by a managment-by-objectives system.

Traditional Textbook Bound

The most common image of a textbook-dependent teacher is that of an individual who begins the school year with the lesson page one and progresses page by page through the book over the course of the year. Judging from the two case-study teachers who adopted this style, one likely consequence is that the last few chapters of the textbook will not be presented.

Selective Omission

The selective-omission style is closely related to the traditional textbook-bound style. In this approach, the teacher progresses lesson by lesson through the textbook, but skips some chapters entirely. Although reasons for skipping chapters vary, they frequently reflect beliefs that some general topics are not particularly important for students to learn, at least not in the their present grade level. Lack of confidence in one's ability to teach certain content may also prompt the omission of chapters. From the case-study teacher who followed this style and from similar teachers we have interviewed, it is possible to identify textbook chapters that fourth-grade teachers are most likely to omit. These include geometry, advanced work with fractions (e.g., adding fractions with unlike denominators), and areas that are traditionally emphasized in later grades (e.g., decimals).

The Basics

Interviews of approximately 20 fourth-grade teachers suggest that there are certain general topics teachers have come to associate with the fourth-grade level.[4] For teachers adopting the basics style, the fourth-grade mathematics curriculum is focused almost

---

4. Descriptions of consistencies in beliefs about what should be taught almost always describe content at a very general level of detail. When content is described in more specific terms (e.g., adding fractions with unlike denominators vs. fractions in general), the level of agreement drops dramatically.

entirely on review of addition, subtraction, and place-value concepts and the introduction or refinement of skills in multiplication, division, and fractions. A seventh general area, measurement, may also be included in some teachers' descriptions of the basics of fourth-grade mathematics. The two case-study teachers, who adopted this style, focused all of their textbook instruction on lessons within the seven basic areas. One teacher included measurement; the other did not. Lessons in the textbook that were not directly related to the basics were omitted.

Management by Objectives (MBO)

Three case-study teachers worked in a school district that required teachers to use a management-by-objectives (MBO) system in teaching mathematics. The system was designed to ensure that all students would acquire minimal competencies in mathematics. As an aid to teachers, the district provided a list of textbook exercises that were coordinated with each of the instructional objectives.

Although methods of implementation varied across classrooms, the MBO system served as the core of an individualized approach to mathematics instruction for many of the teachers in the district. One case-study teacher required students to demonstrate mastery of a set of MBO objectives before they were allowed to work on other aspects of the mathematics curriculum. In this and other classrooms like it, some students had still not mastered all required objectives by the end of the school year. For these students, the MBO system, and not the teacher, defined which textbook lessons would be covered.

Applying the Styles to One Book

This investigation used the 1978 edition of Holt School Mathematics, the fourth-grade textbook (Nichols, Anderson, Dwight, Flourney, Kalin, Schluep, & Simon, 1978) to illustrate what effect the different styles of textbook use had on the match between content taught and content tested. Table 1 describes the chapters in the Holt book most likely to be assigned when a teacher adopts the textbook-bound, the selective-omission, or the basics

Table 1

Assigned Chapters in the Holt Textbook
for Four Models of Textbook Use

		Assigned Chapters			
	Chapter Descriptions	Textbook Bound	Selective Omission	Basics w/Meas.	Basics w/o Meas
1.	Numeration (including place value)	Yes	Yes	Yes	Yes
2.	Number sentences	Yes	Yes	No	No
3.	Adding	Yes	Yes	Yes	Yes
4.	Subtracting	Yes	Yes	Yes	Yes
5	Linear Geometry	Yes	No	No	No
6.	Multiplication & Division	Yes	Yes	Yes	Yes
7.	Multiplying	Yes	Yes	Yes	Yes
8.	Division	Yes	Yes	Yes	Yes
9.	Fractions	Yes	Yes	Yes	Yes
10.	Adding & Subtracting Fractions	No	No	Yes	Yes
11.	Plane & Solid Geometry	No	No	No	No
12.	Measurement	No	Yes	Yes	No
13.	Graphs and Probability	No	No	No	No

Note. Meas. stands for measurement; w/ stands for with; w/o stands for without.

style of textbook use. The basics style is described in two ways: with and without measurement.

For the management-by-objectives model, a coordinated list of textbook assignments determines a teacher's selection of lessons. Because the list of assignments in the case-study district that used this system did not consider the 1978 edition of the Holt textbook, a special procedure was used to identify the textbook topics represented by this model. First, objectives that corresponded to the fourth-grade level were determined. Although the district identified 13 fourth-grade objectives, we observed that teachers generally expanded this list. Therefore, the five objectives immediately preceding and the five following the district's list of fourth-grade objectives were also considered. Observations confirm that this set of 23 objectives is generally consistent with what the majority of fourth-grade students would cover under this MBO system.

Next, we determined the content of an objective through content analysis of the mastery-test items for that objective. By considering alternate forms of the mastery tests, it was possible to analyze 20 test items for each of the 23 fourth-grade objectives. The following descriptions of textbook topics covered in the MBO model assume that a coordinated list of assignments for the Holt text would include all problems that focus on each of the test topics identified.

Procedure

Selection of a Textbook and Tests

The fourth-grade Holt School Mathematics textbook was selected for a variety of reasons. First, it is one of several texts widely used in grades four to six (Weiss, Note 3). Second, among a representative sample of textbooks and tests of fourth-grade mathematics, the Holt book generally provided the best overall content match with standardized tests (Freeman, Note 4). Third, the fact that this book was the primary reference for two case-study teachers provided a fairly clear sense of how it might be used in the classroom. For example, case-study data indicated that teachers would cover approximately nine chapters

during the academic year (See Table 1).

It is possible to identify three distinct components in each Holt lesson: instructional activities directed by the teacher, practice exercises assigned to students, and optional enrichment activities such as "brainteasers" that might be offered to some students. Some lessons also include tests of achievement. In this investigation, all problems in the Holt textbook that appeared in teacher-directed activities, student exercises, and achievement tests were classified (N = 6,316 problems for the entire text). Because few, if any, students covered the enrichment activities in each lesson, this material was not considered.

The standardized tests of fourth-grade mathematics selected for analysis are listed below.

CTBS-I - Comprehensive Tests of Basic Skills (Level I/Grades 2.5-4.9) (CTB/McGraw-Hill, 1975)

CTBS II - Comprehensive Tests of Basic Skills (Level II/Grades 4.5-6.9) (CTB/McGraw-Hill, 1975)

IOWA - Iowa Test of Basic Skills (Level 10/Grade 4) (Hieronymus,et al., 1978)

MAT - Metropolitan Achievement Tests (Elementary Level/Grades 3.5-4.9) (Prescott, Balow, Hogan, & Farr, 1978)

STAN - Stanford Achievement Test (Intermediate Level/Grades 4.5-5.6) (Madden, Gardner, Rudman, Karlsen, & Merwin, 1973)

According to the publishers, each of these tests provides a measure of mathematics achievement at the end of fourth grade. Each contains sets of items on mathematical concepts, computations, and applications. Our content analyses considered all items in each of these areas. The number of items classified ranged from 50 items on the Metropolitan (MAT) to 112 items on the Stanford (STAN).

The Taxonomy on Which Analyses Were Based

The content analyses of the Holt textbook and five tests were guided by a classification manual describing the rules for using a three-dimensional taxonomy of elementary school mathematics (Kuhs, Schmidt, Porter, Floden, Freeman, & Schwille, Note 5). The three dimensions of the taxonomy describe the general intent of the items (e.g.,

conceptual understanding or application), the nature of materials presented to students (e.g., fractions or decimals), and the operation the student must perform (e.g., estimate or multiply).

Table 2 illustrates the flexibility of the taxonomy in describing content at different levels of detail. Specific topics covered on the Stanford Achievement Test (STAN) are represented by cells of the classification matrix (e.g., three of the 112 STAN items focus on column addition of multiple digit numbers). More general topics are addressed by summing across cells to obtain margin totals (e.g., seven of the 112 items deal with column addition).

Level of Interrater Agreement

The Holt textbook and the five tests were independently analyzed by two raters. All items in each test and the Holt textbook were described along all three dimensions of the taxonomy. The results for a given pair of raters were then compared and discrepancies resolved by a third rater. Although there was some variance among the five raters who participated in the study, interrater agreements were generally high. Interrater correlation coefficients at the cell level of the taxonomy were consistently above .98 for the five standardized tests. The corresponding figures for the Holt textbook exceeded .94 for both marginal and cell-level totals.[5]

Measures of the Relation Between Content Taught and Content Tested

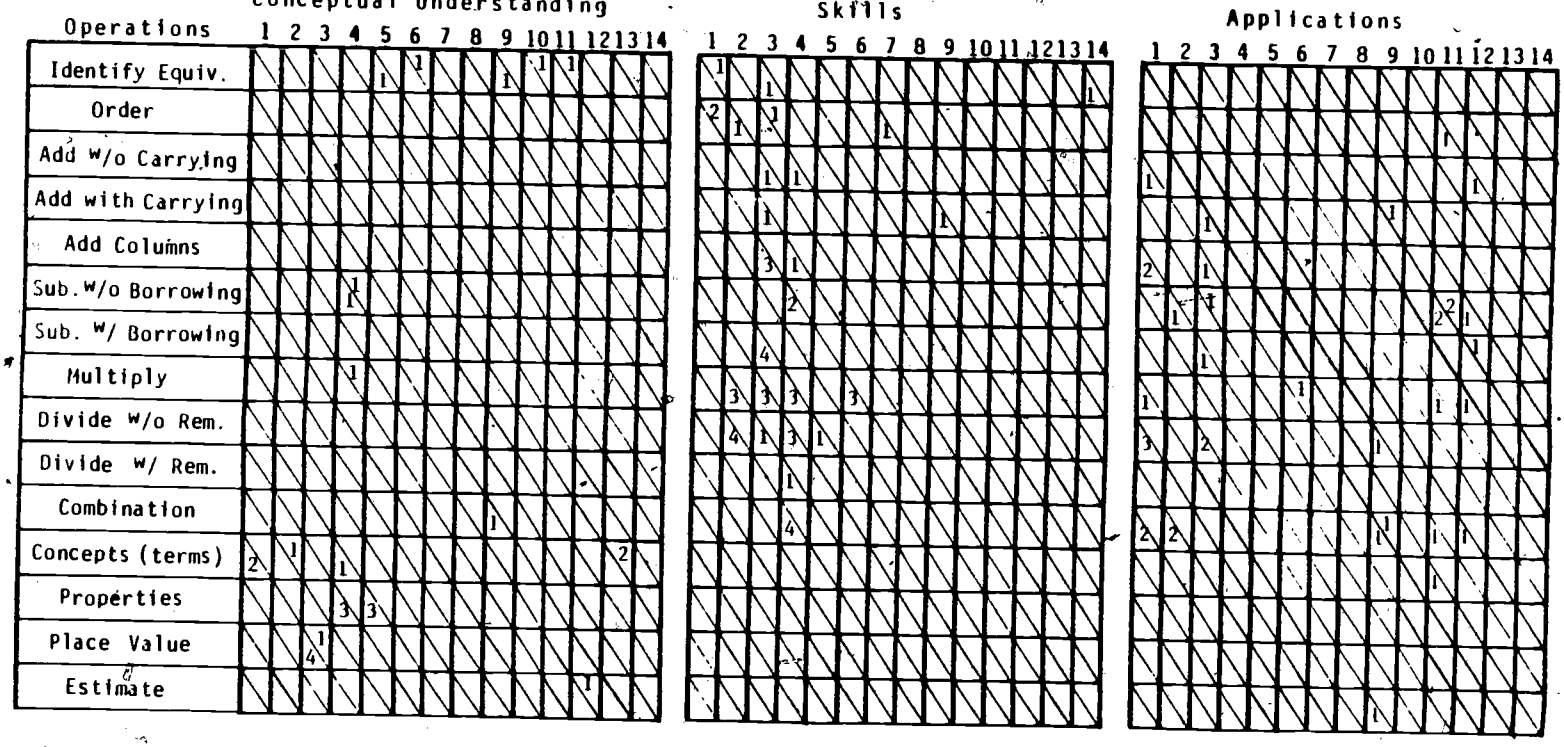
The match between instructional content and tested content, or instructional validity, is usually described in terms of the proportion of the test items that focus on content covered by instruction (Schmidt, Porter, Schwille, Floden, & Freeman, Note 6). The measure of instructional validity selected for this investigation provides an estimate of the proportion of a test that deals with content that students have had an adequate

---

5. Measures of interrater agreement were based on the domain of cell-level topics that were identified for the book or a given test by either of the two raters. The number of items one rater attributed to each of the topics within this domain was correlated with the corresponding information from the second rater.

Table 2

Content Analysis of 1973 Stanford Achievement Test (Intermediate Level I)[a]



Note: Numerals at the top of each column stand for the following:

w/out  with pictures

Nature of Material:

1. single digit/basic # fact
2. single & multiple digit #
3. multiple digit numbers
4. number sentence/phrase
5. algebraic sentence/phrase
6. single/like fractions
7. unlike fractions
8. mixed numbers
9. decimals
11. percents
12. measurements
13. essn, units of measurement
13. geometry
14. other

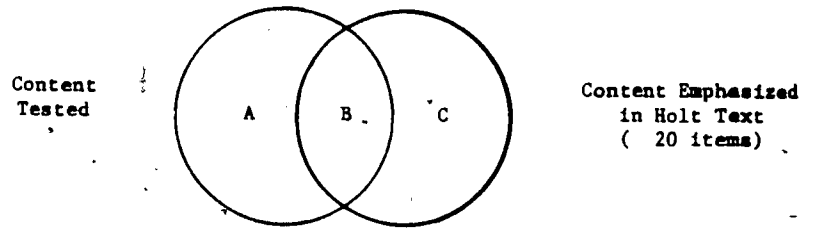
[a]Entries represent the number of items for each cell-level topic (N = 112)

opportunity to learn or to review during the academic year preceding the test's administration. Toward that end, only those cell-level topics that were covered by 20 or more items in the Holt textbook were considered.[6] We will refer to the Holt text covered by 20 or more items as "emphasized" topics throughout this report. Of the more than 6,000 items initially classified, 5,094 (80.7%) dealt with emphasized topics and were therefore considered in the following analyses.

Instructional validity is represented in Figure 1 by the ratio of the number of items on a test that focus on topics emphasized in the textbook (subset B) to the total number of test items. In the analyses that follow, this statistic provides an estimate of the proportion of each standardized test that deals with content students have had an adequate opportunity to learn in each style of textbook use. In order to examine variation in content match from a different perspective, the proportion of instructional content that was tested was also evaluated for each test-textbook style combination. This statistic will be referred to as instructional focus. It provides a measure of the relative attention given to test topics within the total curriculum as defined by a specific style of textbook use. In Figure 1, the instructional focus measure is represented by the ratio of the number of textbook problems that focus on emphasized topics that are also tested (subset B) to the number of textbook items that deal with emphasized topics in each style of textbook use.

---

6. This standard for defining the lower limit of an "adequate opportunity to learn" represents our best estimate of the number of problems a student would need to practice to learn cell-level topics considered in the Holt textbook. Although 20 items represent an average of approximately half of a lesson in the Holt book, the number of items in a given lesson varies as a function of the type of content covered. Twenty items may represent as much as a lesson and a half for content areas such as story problems, or as little as a third of a lesson for areas such as basic number facts.



$$\text{Instructional Validity} = \frac{\text{Number of test items that focus on topics in B}}{\text{Total number of items on test}}$$

$$\text{Instructional Focus} = \frac{\text{Number of text items that focus on topics in B}}{\text{Total number of textbook items dealing with emphasized topics}}$$

Figure 1. A schematic representation of instructional validity and instructional focus.

Results

Proportion of Test Items Covered in Each Model

Assume that a school district administers one of the five standardized tests at the end of fourth-grade year. How many of the items on each test deal with mathematics topics that have been covered by textbook lessons during the previous year? Will this number vary as a function of how the teacher has used the textbook?

Table 3 describes the percent of test items that deal with content covered by 20 or more problems in each style of textbook use. As the numbers suggest, the match in topics covered by a textbook and test varies as a function of the style of textbook use the teacher adopts. For all five standardized tests, the percent of test items that deal with topics emphasized in the Holt text is lowest for the management-by-objectives model. Students who do not move beyond the 23 fourth-grade objectives associated with the MBO model may therefore be handicapped in their performance on tests relative to students in classrooms that conform to one of the other styles of textbook use (or students in MBO classrooms who study content beyond the minimal objectives).

Table 3

Instructional Validity: Percent of Test Items that Deal with Emphasized Instructional Topics*

	Total Book (n=5,094)	Textbook Bound (n=3,786)	Selective Omission (n=3,832)	Basics w/Measurement (n=3,892)	Basics w/o Measurement (n=3,500)	Management by Objectives (n=1,589)
CTBS I (n = 98 items)	62.2	61.2	62.2	61.2	60.2	40.8
CTBS II (n = 98 items)	50.0	39.8	40.8	48.0	46.9	24.5
IOWA (n = 104 items)	53.8	50.0	50.0	52.9	51.9	35.6
METROPOLITAN (n = 50 items)	60.0	48.0	60.0	58.0	46.0	18.0
STANFORD (n = 112 items)	35.7	31.3	31.3	32.1	30.4	17.9

*Emphasized instructional topics are those that serve as the focus of at least 20 items in taught portions of the Holt textbook.

For three of the five tests, differences in instructional validity across the other four styles of textbook use are not likely to result in large discrepancies in the overall level of student performance on a given test. This is particularly true for the CTBS Level I and the Iowa tests, where instructional validity for all four styles is nearly identical to that of the total book. With the Stanford test, measures of instructional validity for the four styles are nearly equal and are only slightly inferior to that for the total text. However, corresponding measures for the other two tests do vary across different styles of textbook use. Whereas, the two basics models cover almost as much CTBS Level II content as the whole book does, instructional validity for the textbook-bound and selective-omission styles is about 10% percentage points lower than that for the entire text. With respect to the Metropolitan test, instructional validity for the selective-omission style is identical to that of the whole book and is only slightly lower for the basics-with-measurement style. However, textbook-bound and basics-without-measurement styles fall below the whole text by 12 and 14 percentage points respectively.

Collectively, these results suggest that differences in the content of classroom instruction resulting from different styles of textbook use may have little, if any, influence on the overall level of student performance on some standardized tests. For example, of the 45 topics emphasized in one or both the selective-omission and basics-with-measurement styles of textbook usage, only 35 were emphasized in both styles. Yet the difference in level of instructional validity of these two styles never varied more than eight percentage points across the five standardized tests. Among the other standardized tests considered in this analysis, the CTBS I, Iowa, and Stanford appear to be relatively insensitive to different styles of using the Holt textbook of fourth-grade mathematics. However, for the CTBS II and Metropolitan, the match in textbook-test content will vary as a function of how the teacher uses the book.

Overall, the basics-with-measurement style provides the highest levels of instructional validity for three of the five tests considered and is only one and two percentage points lower than the best style of usage for the other two tests. In fact, averaged across tests,

measures of instructional validity for the basics-with-measurement style are only two percentage points lower than if the whole textbook were taught.

Proportion of Instructional Items that Are Tested

Considering the portion of a test that focuses on content emphasized in instruction provides information from only one perspective. A more complete picture is provided if one also considers the relative emphasis the tested topics receive in instruction. For example, assume that two curricula both cover half the items of a given test. Further assume that, although the total amount of time allocated to mathematics is the same in both settings, tested topics represent 80% of the instructional items in one curriculum and only 40% in the other. Achievement results might be expected to differ in favor of students in the former curriculum since, presumably, they would have had a greater opportunity to practice the tested material.

Table 4 describes the relative emphasis given to tested topics for each style of textbook usage, as well as for the entire Holt textbook. The results indicate that measures of instructional focus also vary as a function of style of textbook use.[7] The MBO model devotes the highest proportion of instructional time to tested content across all five standardized tests. From 60-95% of all textbook problems that would be assigned in this model deal with content that is tested. However, it is important to consider this finding in conjunction with results from Table 3, which indicate that the MBO system provides the lowest overlap with test content (only 18-40% of the test items deal with content emphasized in instruction). Hence, students who use only the Holt textbook and who do not progress beyond the 23 MBO objectives are exposed to substantially less test content than students who receive instruction under the other models. At the same time, they are spending the majority of their time on that limited domain of content.

7. Because only those topics emphasized in the book were considered, the measures reported in Table 4 provide liberal estimates of instructional focus. However, when the authors recalculated these measures for the total number of problems in the book, the pattern of results did not change. However, each measure did decrease by an average of about 5-10%.

Table 4

Instructional Focus: Percent of Problems Presented During Instruction That Focus on Emphasized Topics That Are Tested

	Total Book (n=5,094)	Textbook Bound (n=3,786)	Selective Omission (n=3,832)	Basics w/Measurement (n=3,892)	Basics w/o Measurement (n=3,500)	Management by Objectives (n=1,589)
CTBS I (n = 98 items)	51.6	58.0	52.5	50.8	53.6	79.5
CTBS II (n = 98 items)	49.3	47.7	42.8	47.0	49.1	82.7
IOWA (n = 104 items)	56.7	63.2	56.7	57.7	61.7	95.0
METROPOLITAN (n = 50 items)	52.6	54.1	54.5	52.6	49.5	60.5
STANFORD (n = 112 items)	44.6	47.8	41.4	40.6	42.8	62.5

In terms of the proportion of the curriculum devoted to tested topics in the other four styles of textbook use, the textbook-bound model ranks highest on three tests (CTBS I, Iowa, and Stanford) and second on the other two. However, differences among the four styles on any one test are typically small and are never larger than seven percentage points.

Level of Practice on Tested Topics

After reviewing the preceding tables, an important question to consider is how much practice each style of textbook use will provide for those topics that are tested. Table 5 presents, by style of text usage and test, the median number of problems in the Holt book that focused on tested topics. It also describes the range for each distribution of item frequencies.[8]

Because the two measures are not independent, it should come as no surprise that the pattern of results portrayed in Table 5 parallel those described in Table 4. For all five tests, the median number of problems per tested topic was higher for the MBO model than for the other four styles of textbook use (despite the MBO covering less than half as many textbook problems as the other styles; see top row of Table 3). Among the other four styles, the textbook-bound model provided the highest level of practice on tested topics for four of the five topics.

In general, for each test topic cosidered, MBO students received an average of about one half to two full lessons of instruction (26 to 91 problems) beyond that provided in the other styles of textbook use. If this style of textbook usage considered as many test topics as the other four styles, one might expect these students to enjoy a competitive advantage on tests. However, even if one assumes that the level of practice provided by the MBO model assures total command of the material, students could be expected to successfully answer only 41% of the items on the CTBS I test (see Table 3). In contrast, students working

8. Because almost all of these distributions were skewed, the median provided the best measure of central tendency. In interpreting the figures in Table 5, it is important to recall that these analyses considered only those topics that were covered by a least 20 items in the Holt text. Thus a frequency of 20 serves as the lower limit of the range for each distribution.

Table 5

Median Number of Textbook Items Per Tested Topic
(Range of Item Frequencies Per Tested Topic)

	Total Book	Textbook Bound	Selective Omission	Basics w/Measurement	Basic w/o Measurement	Management by Objective
CTBS I	97 (20-561)	99 (20-268)	80 (20-228)	79 (25-228)	78 (25-228)	155 (55-234)
CTBS II	60 (20-561)	80 (23-268)	65 (23-223)	59 (22-225)	57 (21-222)	127 (55-234)
IOWA	97 (23-561)	120 (23-268)	100 (23-228)	79 (22-228)	75 (21-228)	146 (27-234)
METROPOLITAN	123 (25-561)	123 (25-268)	99 (38-228)	102 (38-228)	104 (38-228)	214 (124-234)
STANFORD	82 (20-561)	100 (20-268)	80 (20-223)	80 (22-225)	101 (22-222)	146 (27-234)

under one of the other four styles of using textbooks would have somewhat less practice for over 60% of content tested on this exam. Thus, for students to perform as well under the MBO model, the extra practice they receive must increase their probability of success on tested items to about 1½ times that for students working under one of the other four textbook styles. For other tests, this ratio will be even higher. As the data in Table 3 suggest, the differential level of success for MBO students must be about three times higher than that for other students when the Metropolitan test is administered.

The exact number of items or amount of practice needed to ensure a high probability of getting a test item correct is an empirical question, and cannot be answered with these data. A complete assessment of potential test performance would need to consider other variables such as complexity of the topic, massed versus distributed practice, and the sequence of instruction, in addition to the absolute number of items practiced. For example, a relatively small number of instructional items practiced immediately before a test might produce better test performance than many items on that topic presented early in the year. However, given the differential levels of success outlined above, it seems questionable that the additional practice provided by the MBO style of textbook usage on a subset of tested topics would compensate for the large number of tested topics that did not receive any attention at all.

Discussion

A number of factors limit the generalizability of these results. Only one textbook, one grade level, and one subject-matter area were considered. Changes in one or more of these factors are likely to alter the magnitude, and perhaps even the pattern, of results. For these and other reasons, the data summarized in this report should be viewed as illustrative, rather than definitive evidence, of variation in the level of instructional validity of tests that may result from differences in how a textbook is used.

Additionally, these analyses considered only curriculum-based content as presented in the textbook. Analyses of the content of classroom instruction would also include teacher-

presented topics and might produce different results. However, while teachers may use other instructional materials, they rarely teach content not considered in textbooks. The work of Leinhardt and Seewald (1981) supports this contention. They found that analyses of textbooks and of classroom instruction were comparable in estimating posttest results.

Despite these limitatons, the data from the study provide support for two important conclusions regarding the relation between content covered in textbooks and tests.

1. When a management-by-objectives system serves as the core of an individualized approach to instruction, low achievement students may have limited opportunities to learn more than a narrow range of the content covered on standardized tests.

2. Although clear differences in the content of classroom instruction may result from different styles of textbook use, these differences will have little, if any, influence on the overall level of student performance on some standardized tests of achievement.

Even though the management-by-objectives system reviewed in this investigation provides greater depth of coverage of those test topics it considers, it is clearly inferior to the other four styles of textbook use in regard to the match it provides between textbook and test content. However, since the MBO is meant to define the minimum mathematics curriculum, most youngsters will receive instruction in other areas of mathematics. Nevertheless, five of the students in one of the case-study classrooms spent the entire academic year attempting to master the set of 23 objectives. These five students, and others like them, were at a competitive disadvantage in their opportunity to learn many of the topics that are covered on standardized tests. If school districts use an MBO system of this sort and are anxious to maximize achievement test scores, they should therefore make a concerted effort to

1. encourage teachers to use instructional approaches that ensure that students receive instruction in areas of mathematics beyond those defined by the objectives, and

2. focus objectives on content that is emphasized on the standardized test administered in the district.

Districts that choose to ignore either or both of these conditions should anticipate at least

some decline in standardized test scores, particularly among those low-achievement students for whom the system was developed.

In an earlier paper (Porter, Schmidt, Floden, & Freeman, 1978), the authors stressed that overall scores on standardized tests may be insensitive to important differences in curriculum outcomes. The data from this investigation provide a graphic illustration of that point. Although different styles of using the Holt textbook yield substantive differences in the content of classroom instruction, these differences will be masked when looking at total scores on three of the five tests considered. It is clear from these results that teachers and administrators must look beyond total test scores if they wish to understand the strengths and weaknesses of the curriculum.

References Notes

1. National Advisory Committee on Mathematics Education. Overview and analysis of school mathematics: Grade K-12. Washington, D.C.: Conference Board on the Mathematical Sciences, 1975.

2. Freeman, D.J., Kuhs, T.M., Porter, A.C., Knappen, L.B., Floden, R.E., Schmidt, W.H., & Schwille, J.R. The fourth-grade mathematics curriculum as inferred from textbooks and tests (Research Series No. 82). East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1980.

3. Weiss, I.R. Report of the 1977 national survey of science, mathematics and social studies (RTI/1266/06/01F). Research Triangle Park, North Carolina: Research Triangle Institute, 1978.

4. Freeman, D.J. Myths in math: A closer look at the elementary school curriculum. Paper presented at the annual meeting of the Michigan Educational Research Association, Detroit, Michigan, March 1981.

5. Kuhs, T.M., Schmidt, W.H., Porter, A.C., Floden, R.E., Freeman, D.J., & Schwille, J.R. A Taxonomy for classifying elementary school mathematics content (Research Series No. 4). East Lansing, Michigan: Institute for Research on Teaching, Michigan State University, 1979.

6. Schmidt, W.H., Porter, A.C., Schwille, J.R., Floden, R.E., & Freeman, D.J. Validity as a variable: Can the same certification test be valid for all students? Paper presented at the Conference on the Courts and Content Validity, Boston College, Boston, Massachusetts, October 1981.

References

Berliner, D.C. Tempus educare. In P.L. Peterson and H.L. Walberg (Eds.), Research on teaching: Concepts, findings, and implications. California: McCutchan, 1979.

CTB/McGraw-Hill. Comprehensive tests of basic skills. Monterey, California: McGraw-Hill, 1976.

Hieronymus, A.N., et al. Iowa test of basic skills. Boston, Massachusetts: Houghton Miffin, 1978.

Leinhardt, G., & Seewald, A.M. Overlap: What's tested, what's taught? Journal of Educational Measurement, 1981, 18, 85-96.

Madden, R., Gardner, E., Rudman, H., Karlsen, B., & Merwin, J. Stanford achievement test. New York: Harcourt Brace Jovanovich, 1973.

Nichols, E.D., Anderson, P.A., Dwight, L.A., Flournoy, F., Kalin, R., Schluep, J., & Simon, L. Holt school mathematics. New York: Holt, Rinehart and Winston, 1978.

Porter, A.C., Schmidt, W.H., Floden, R.E., & Freeman, D.J. Practical significance in program evaluation. American Educational Research Journal, 1978, 15, 529-539.

Prescott, G.A., Balow, I.H., Hogan, T.P., & Farr, R.C. Metropolitan achievement tests. New York: Harcourt Brace Jovanovich, 1978.