

NATIONAL CENTER FOR EDUCATION STATISTICS

Working Paper Series

The Working Paper Series was initiated to promote the sharing of the valuable work experience and knowledge reflected in these preliminary reports. These reports are viewed as works in progress, and have not undergone a rigorous review for consistency with NCES Statistical Standards prior to inclusion in the Working Paper Series.

U. S. Department of Education
Institute of Education Sciences

This page intentionally left blank.

NATIONAL CENTER FOR EDUCATION STATISTICS

Working Paper Series

PIRLS-IEA Reading Literacy Framework: Comparative Analysis of the 1991 IEA Reading Study and the Progress in International Reading Literacy Study

Working Paper No. 2003-05

April 2003

Contact: Laurence Ogle
International Activities Program Division
Laurence.Ogle@ed.gov

**U. S. Department of Education
Institute of Education Sciences**

U.S. Department of Education

Rod Paige
Secretary

Institute of Education Sciences

Grover J. Whitehurst
Director

National Center for Education Statistics

Valena Plisko
Associate Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

National Center for Education Statistics
Institute of Education Sciences
U.S. Department of Education
1990 K Street NW
Washington, DC 20006–5651

April 2003

The NCES World Wide Web Home Page address is <http://nces.ed.gov>

The NCES World Wide Web Electronic Catalog is: <http://nces.ed.gov/pubsearch>

Suggested Citation

U.S. Department of Education, National Center for Education Statistics. PIRLS-IEA Reading Literacy Framework: Comparative Analysis of the 1991 Reading Study and Progress in International Reading Literacy Study, NCES 2003–05, by Barbara Kapinus. Project Officer: Laurence Ogle. Washington, DC: 2003

For ordering information on this report, write:

U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794–1398

Or call toll free 1–877–4ED–Pubs

Content Contact:

Laurence Ogle
(202) 502–7426
Laurence.Ogle@ed.gov

Foreword

In addition to official NCES publications, NCES staff and individuals commissioned by NCES produce preliminary research reports that include analyses of survey results, and presentations of technical, methodological, and statistical evaluation issues.

The *Working Paper Series* was initiated to promote the sharing of the valuable work experience and knowledge reflected in these preliminary reports. These reports are viewed as works in progress, and have not undergone a rigorous review for consistency with NCES Statistical Standards prior to inclusion in the Working Paper Series.

Copies of Working Papers can be downloaded as pdf files from the NCES Electronic Catalog (<http://nces.ed.gov/pubsearch/>), or contact Sheilah Jupiter at (202) 502-7444, e-mail: sheilah.jupiter@ed.gov, or mail: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 1990 K Street NW, Room 9048, Washington DC 20006.

Marilyn M. Seastrom
Chief Mathematical Statistician
Statistical Standards Program

Ralph Lee
Mathematical Statistician
Statistical Standards Program

This page intentionally left blank.

PIRLS-IEA Reading Literacy Framework: Comparative Analysis of the 1991 IEA Reading Study and the Progress in International Reading Literacy Study

Prepared by:

Barbara Kapinus

Prepared for:

U.S. Department of Education
Institute of Education Statistics
National Center for Education Statistics

April 2003

This page intentionally left blank.

Background

In 1991, the United States participated in the International Association for the Evaluation of Educational Achievement (IEA) Reading Literacy Study that assessed the reading literacy of 4th and 9th grade students in 32 countries. Fourth grade students in the U.S. did very well on the study - “With the exception of Finland, no country consistently outperforms the United States “ - (Binkley and Williams, 1996, p. 12).

This result was very different from the results of the National Assessment of Educational Progress (NAEP) in 1992. That assessment of 4th, 8th, and 12th grade students, indicated that less than one third (29 percent) of the nation’s fourth graders were at the Proficient Level or above and only 6 percent achieved the Advanced level. It is not surprising that the IEA and NAEP provided different pictures of U. S. students’ proficiency since the NAEP was designed to indicate how well U.S. students were performing relative to our own high standards and the IEA was designed to compare the performance of students populations across several countries. Nevertheless, the differences in the picture of student literacy provided by the two assessments raised questions about the comparability of the two studies. As a result, a study of the two assessments was conducted (Binkley and Williams, 1996). This study found important differences in the items and passages used on the two assessments with NAEP appearing to be a substantially more difficult assessment. Now that a new international assessment to compare fourth-grade reading achievement across countries has been developed and administered, it is important to note differences in the former international assessment and the new one. That comparison is the subject of this paper.

When a second study of 4th grade literacy was being planned for 2001, the IEA decided to create a new assessment: The Progress in International Reading Literacy Study (PIRLS). A framework was developed with the intent to “incorporate in PIRLS the latest approaches to measuring reading literacy.” (Campbell et al., 2001). In the United States alone some important changes have taken place in reading assessment in the decade since the IEA. For example, the National Assessment of Educational Progress (NAEP), as well as several state assessments (e.g., Maryland, Connecticut) and commercial tests (e.g., the Stanford 9), have incorporated more extensive use of constructed response items and taxonomies that focus on responding to text as well as getting ideas from the text. Thus, it is logical to assume that the PIRLS is a very different assessment from the 1991 IEA.

In the meantime, high standards of performance, especially in reading and math, were called for by policy makers and education leaders across the United States. In addition to state and national benchmarks for performance, many policy makers called for international studies to determine not only the achievement of U.S. students but also to indicate how rigorous our standards of performance were. Given that context, which continues and is more compelling than ever, it is important to take a close look at how the two international studies compare to each other in the aspects of reading literacy each assessed, the types of texts they used in the assessments, and the types and difficulty of the questions they used.

In order to carry out this inspection and comparison, three experts reviewed the taxonomies, texts, and questions of each assessment. These reading specialists had extensive experience with the National Assessment of Education Progress, and also with a state assessment using the same taxonomy and question types as NAEP. One had also worked on the development of both the IEA and the PIRLS.

Frameworks and Aspects of Reading

The IEA Reading Literacy Study Framework

The IEA Reading Literacy Study defined literacy as

...the ability to understand and use those written language forms required by society and/or valued by the individual. Literacy occurs in a variety of language contexts (e.g., school, home, work, and religious or civic institutions) and involves both a range of competencies and a set of habits and/or practices, arrayed along various dimensions (Binkley and Rust, August, 1994, p. 8.)

The fourth grade assessment for the IEA used three types of texts:

- narrative prose – continuous text written to tell a story, either factual or fictional;
- expository text - continuous text written to describe or explain things;
- documents – structured tabular texts such as forms, charts, labels, graphs, lists, and directions.

The questions on the IEA assessment focused on six types of reading processes: verbatim, paraphrase, inference, main theme, locating information, and following directions. Figure 1 below summarizes the types of text and questions.

Figure 1: Types of texts and skills assessed on the IEA

Types of Text	Skills Assessed					
	Verbatim	Paraphrase	Main Theme	Inference	Locate Information	Follow Directions
Narrative Prose						
Expository Prose						
Document						

Based on Binkley, 1994, p. 106

The PIRLS Assessment

For PIRLS, reading literacy was defined as

the ability to understand and use those written language forms required by society and/or valued by the individual. Young readers can construct meaning from a

variety of texts. They read to learn, to participate in communities of readers, and for enjoyment (Campbell, J. et al., March, 2000, p. 3).

PIRLS employed three aspects of reading in order to construct the assessment:

- processes of comprehension;
- purposes for reading; and
- reading behaviors and attitudes.

The latter was addressed through a student questionnaire. The first two aspects framed the selection of passages and the development of items.

The following chart lists the purposes for reading and the processes of comprehension that guided passage selection and item development.

Figure 2: Reading Purposes and Processes on the PIRLS Assessment

Processes of Comprehension	Purposes for Reading	
	Literary Experience	Acquire and Use Information
Focus on and Retrieve Explicitly Stated Information		
Make Straightforward Inferences		
Interpret and Integrate Ideas and Information		
Examine and Evaluate Content, Language, and Textual Elements		

Based on Campbell et al., August 2001

Comparison of the Frameworks

While the two definitions of reading literacy appear to be similar, they do seem to differ in scope. PIRLS specifically focuses on fourth graders compared to the focus of the IEA that took into account ninth graders as well as fourth graders. Consequently, the PIRLS definition speaks to the more limited applications of reading that would be asked of fourth-grade students. It does not name contexts such as school, home, work, religious and civic institutions, but rather names purposes of reading for the fourth-grade population.

While the types of texts used in the assessments are delineated differently in the two frameworks, they essentially encompass the same bodies of texts. The PIRLS identifies the types of texts used on the assessment within the categories of purposes for reading while IEA directly lists the types of texts without referring to different purposes for reading. Narrative prose on the IEA includes both fictional and factual texts while the texts included under reading for literary experience on the PIRLS are mainly narrative fiction. On the IEA, expository prose and documents are two separate categories while these are both included under reading to acquire and use information on the PIRLS. The general types of texts used on the two assessments were the same. However, there were

major differences in the characteristics of the texts used. These differences are described in the following section.

Analysis of Texts

The IEA used fifteen separate texts with two to seven questions on each text. Passage length ranged from 56 words to 706 words. All of the passages were under 500 words with the exception of one narrative text. Six of the texts were charts or graphics such as a map and the word count does not really capture everything students had to process in order to answer questions on those prompts. Usually, these prompts did not have much surrounding text. The passages were presented in block type with basic, line drawings for illustrations. All of the passages were administered to each student taking the assessment.

Table 1: IEA number of words per passage by text type

Narrative		Expository	
The Bird and the Elephant	292	Postcard	56
Grandpa	310	Quicksand	141
A Shark Makes Friends	452	Walrus	207
No Dogs Is Not Enough	706	Marmots	222
		Trees	383

(Binkley, 1994, p. 104)

The PIRLS used seven texts, four for measuring reading to acquire information and three for measuring reading for literary experience. Passage length ranged from 405 words to 801 words. Three of the passages were over 500 words. Only one passage dealt with reading graphics. The brochure on the river trail involved reading both a chart and a map as well as surrounding text. In order to present passages in such a way as to approximate authentic reading experiences of fourth grade students, the passages on PIRLS were accompanied by their original illustrations or by illustrations specially commissioned for the assessment. The font was similar to that in children’s books and magazines. Two passages were presented in a magazine format, the *PIRLS Reader*, with color pictures.

Because it would take a student approximately four hours to read and respond to all the texts on the PIRLS, the passages were paired in blocks that required only an hour and twenty minutes of testing time per student. The blocks were then randomly assigned to students in the sample. Consequently any one student would only respond to two of the reading passages.

Table 2:PIRLS number of words per passage by text type

Reading for Literary Experience		Reading to Acquire and Use Information	
Flowers on the Roof	801	Nights of the Pufflings	683
The Upside-Down Mice	405	Antarctica: Land of Ice	392
Rabbit Raises the earthquake Alarm	444	River Trail	378
		Leonardo da Vinci	545

In considering the passages on the two assessments, the three experts reviewing the items and texts agreed that the PIRLS passages appeared more engaging and challenging. Not only were the PIRLS texts longer, they were generally more complex. The format of the PIRLS passages, more like texts that fourth graders are likely to encounter daily, was more engaging. However, while this might be more encouraging to students to read the PIRLS passages than the IEA passages, the longer length of the PIRLS passages could have countered that effect since students are often more willing to read shorter passages than longer ones.

Analysis of Items

Procedure

The three experts read the passages of one assessment at a time starting with the PIRLS assessment. After they read each passage on the PIRLS and categorized the items on a chart indicating the taxonomy categories for IEA (See Appendix A), they discussed the classification of each item until they reached consensus. They also categorized each passage according to the IEA passage types. The process was then repeated categorizing IEA items on a chart of the PIRLS taxonomy and the passages according to PIRLS passage types. Tables 1 and 2 show the results of the item analyses.

Table 3: Results of classification of PIRLS items using the IEA taxonomy

Item Type	Verbatim	Paraphrase	Follow Direction	Main Theme	Inference	Locate Information in a Document
Number of Items in this Category (Percent of Total)	21 (25%)	12 (14%)	1 (1%)	5 (6%)	42 (49%)	4 (5%)

Table 4. Results of the classification of IEA items using the PIRLS taxonomy

Item Type	Focus on and retrieve Explicitly Stated Information	Make Straightforward Inferences	Interpret and Integrate Ideas and Information	Examine and Evaluate Content, Language, and Textual Elements
Number of Items in this Category (Percent of Total Items)	37 (54%)	21 (31%)	9 (13%)	1 (1%)

While all of the PIRLS items could be put in a category in the IEA taxonomy, not all of the items on the IEA fit into the PIRLS categories. Specifically, the first 40 items on the IEA were *sight vocabulary* items where a word was presented and students were asked to select the picture that matched the word. This type of item was not included on the PIRLS or even considered in its taxonomy. It was possible to fit all of the other items into the PIRLS categories. As a result, those first forty items were not considered in the total items for indicating percents of items in each category of the taxonomies. Most of the PIRLS items (49 percent) fit in the IEA category of “inferences.” Most of the items that appeared to be written for the PIRLS category of “examine and evaluate content, language, and textual elements” were placed in this category during the analysis because that there was no IEA category analogous to that particular PIRLS category. PIRLS items categorized as verbatim on the IEA taxonomy were most likely categorized as “focus on and retrieve explicitly stated information” on the PIRLS taxonomy.

In order to consider how the percent of items in each category for the taxonomies compare to the percentages planned or reported for the each of the assessments, Tables 5 and 6 summarize that information.

Table 5: Percent of IEA items in each category for IEA questions

Item Type	Verbatim	Paraphrase	Follow Direction	Main Theme	Inference	Locate Information in a Document
Percent of Total	16	24	-	6	34	-

Table 6: Percent of PIRLS items planned for each category for PIRLS questions

Item Type	Focus on and retrieve Explicitly Stated Information	Make Straightforward Inferences	Interpret and Integrate Ideas and Information	Examine and Evaluate Content, Language, and Textual Elements
Planned Percent of Total	20	30	30	20

An important difference in the items on the two tests was that the IEA had two constructed response items on the entire assessment and it was optional for countries to count those items in scores. The PIRLS, however, had 43 constructed response items. These ranged in value from one point to three points, depending on the length and difficulty of the expected response. This difference alone suggests that the PIRLS might have been more difficult than the IEA.

In order to explore the question of the difficulty level of the items, seven items were randomly selected from each assessment and rated using a short form of the system for determining levels of processing developed by Wixson, 1999. See Appendix B. This system consists of four levels of processing for reading activities:

- Level 1 requires only a shallow understanding of the text. Oral reading, verbatim recall, and simple understanding of a single phrase are some of the activities in this category.
- Level 2 requires some additional mental processing; comprehension plus some interpretation. Inter sentence analysis, making inferences, classifying, comparing, and indicating literal main ideas are some of the tasks at this level.
- Level 3 involves deep knowledge and going beyond the text. Activities at this level include generalizing, explaining, connecting ideas, making inferences across entire passages, and identifying abstract themes.
- Level 4 requires higher order thinking such as developing hypotheses, complex analyses, and connecting ideas across passages.

The results of the analysis of levels of process are summarized in Table 7 below.

Table 7: Number of items in random samples from each assessment at each level of processing

Assessment	Level 1	Level 2	Level 3	Level 4
IEA	4	3		
PIRLS	1	3	2	1

Although the random sample of items was small, it suggests that the items on the PIRLS required deeper reading knowledge and cognitive skills than those on the IEA. This is

not surprising given the fact that almost half of the PIRLS items were constructed response while there were only two constructed-response items on the IEA. In addition, as discussed earlier, the passages on PIRLS might be considered richer and more likely to elicit higher-level thinking. Finally, the small number of items per passage on the IEA, 2–7, compared to 11-13 items per passage on PIRLS implies that the PIRLS items required students to delve more deeply into the texts and to employ a wider range of cognitive skills.

Summary

In comparing the frameworks, passages, and items of the IEA administered in 1991 and the PIRLS administered in 2001, there were both differences and similarities. The definitions of reading literacy were very similar. The types of passages used were similar but in actually choosing and categorizing passages, the IEA emphasized the types of texts while the PIRLS focused on purposes for reading. The passages/texts on the PIRLS were longer, more engaging, and more complex in most cases. The question taxonomies that were generated to frame the tasks on the assessments were very different. The IEA taxonomy had a text focus with activities such as verbatim responses, main theme, and locating information. The PIRLS taxonomy suggested more consideration of the readers' interaction with the passage especially in the categories of "interpret and integrate ideas and information" and "examine and evaluate content, language, and textual elements." The use of a high number of constructed response items permitted the PIRLS questions to tap a wider range of reading responses and this is supported by the limited analysis of a sample of questions using Wixson's Levels of Depth of Knowledge. The PIRLS passages were presented in an engaging and authentic manner that might have improved students' motivation to read and respond to the texts. This is one area where the form of the PIRLS might have contributed to students' level of performance. However, if students lacked the skills necessary to respond to the items, engaging texts would not have helped much. With a wider range of skills being called for on the PIRLS than on the IEA, especially skills requiring deeper thinking, it is possible that the results of the 2001 PIRLS will show U.S. students spread out more at the top and a wider distribution of their scores than on the IEA in 1991.

References

- Binkley, M. (1994). The IEA reading literacy test. In Binkley, M. and Rust, K. Eds. (1994) *Reading Literacy in the United States: Technical Report of the U.S. Component of the IEA Reading Literacy Study*. Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.
- Binkley, M. and Rust, K. Eds. (1994). *Reading Literacy in the United States: Technical Report of the U.S. Component of the IEA Reading Literacy Study*. Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.
- Binkley, M. and Williams, T. (1966). *Reading Literacy in the United States Findings from the IEA Reading Literacy Study*. Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.
- Campbell, J.R.; Kelley, D.L.; Mullis, I.V.S.; Martin, M.O.; Sainsbury, M. (March, 2001). *Framework and Specifications for PIRLS Assessment 2001*, Chestnut Hill, MA: PIRLS International Study Center, Lynch School of Education, Boston College.
- Wixson, K., Fisk, M.C., Dutro, E., and McDaniel, J.(February 4, 1999). The alignment of state standards and assessments in elementary reading. A report commissioned by the National Research Council's Committee on Title I Testing and Assessment. Ann Arbor, Michigan: The University of Michigan.

Appendix A
Item Rating Sheets

Appendix B
Levels of Processing

Levels of Depth of Knowledge for Language Arts Reading (based on Wixson, 1999)

Interpreting and assigning depth-of-knowledge levels of both objectives within standards and assessment is an essential requirement of alignment analysis. Four levels of depth of knowledge are used for this analysis.

Level 1

Level 1 requires students to receive or recite facts or to use simple skills or abilities. Oral reading that does not include analysis of the text as well as basic comprehension of a text is included. Items require only a shallow understanding of text presented and often consist of verbatim recall from text or simple understanding of a single word or phrase. Some examples that represent but do not constitute all of level 1 performance are:

- Support ideas by reference to details in the text.
- Use a dictionary to find the meaning of words.
- Identify figurative language in a reading passage.

Level 2

Level 2 includes the engagement of some mental processing beyond recalling or reproducing a response; it requires both comprehension and subsequent processing of text or portions of text. Intersentence analysis of inference is required. Some important concepts are covered but not in a complex way. Standards and items at this level may include words such as summarize, interpret, infer, classify, organize, collect, display, compare, and determine whether fact or opinion. Literal main ideas are stressed. A level 2-assessment item may require students to apply some of the skills and concepts that are covered in level 1. Some examples that represent but do not constitute all of level 2 performance are:

- Use context cues to identify the meaning of unfamiliar words.
- Predict a logical outcome based on information in a reading selection.
- Identify and summarize the major events in a narrative.

Level 3

Deep knowledge becomes more of a focus at level 3. Students are encouraged to go beyond the text; however, they are still required to show understanding of the ideas in the text. Students may be encouraged to explain, generalize, or connect ideas. Standard and items at level 3 involve reasoning and planning. Students must be able to support their thinking. Items may involve abstract theme identification, inference across an entire passage, or students' application of prior knowledge. Items may also involve more superficial connections between texts. Some examples that represent but do not constitute all of level 3 performance are:

- Determine the author's purpose and describe how it affects the interpretation of a reading selection.
- Summarize information from multiple sources to address a specific topic.
- Analyze and describe the characteristics of various types of literature.

Level 4

Higher order thinking is central and knowledge is deep at level 4. The standard or assessment item at this level will probably be an extended activity, with extended time provided. The extended time period is not a distinguishing factor if the required work is only repetitive and does not require applying significant conceptual understanding and higher-order thinking. Students take information from at least one passage and are asked to apply this information to a new task. They may also be asked to develop hypotheses and perform complex analyses of the connections among texts. Some examples that represent but do not constitute all of level 4 performance are:

- Analyze and synthesize information from multiple sources.
- Examine and explain alternative perspectives across a variety of sources.
- Describe and illustrate how common themes are found across texts from different cultures.