DOCUMENT RESUME

ED 475 855 CS 511 951

AUTHOR Binkley, Marilyn; Kelly, Dana L.

TITLE A Content Comparison of the NAEP and PIRLS Fourth-Grade Reading Assessments. Working Paper Series.

INSTITUTION National Center for Education Statistics (ED), Washington, DC.

REPORT NO NCES-WP-2003-10

PUB DATE 2003-04-00

NOTE 58p.

AVAILABLE FROM ED Pubs, P.O. Box 1398, Jessup, MD 20794-1398. Tel: 877-433-7827 (Toll Free); Fax: 301-470-1244; e-mail: edpubs@inet.ed.gov; Web site: http://www.ed.gov/pubs/edpubs.html. For full text: http://nces.ed.gov/pubs2003/200310.pdf.

PUB TYPE Reports - Research (143)

EDRS PRICE EDRS Price MF01/PC03 Plus Postage.

DESCRIPTORS Comparative Analysis; *Grade 4; Intermediate Grades; *Literacy; Measures (Individuals); *Reading Achievement; *Reading Skills

IDENTIFIERS IEA Reading Literacy Study; International Assn Evaluation Educ Achievement; International Surveys; *National Assessment of Educational Progress

ABSTRACT

In 1991, the United States participated in the International Evaluation of Educational Achievement (IEA) Reading Literacy Study that assessed the reading literacy of fourth-grade students in 32 countries. When a new study of fourth-grade reading literacy was being planned for 2001, the IEA decided to create a new assessment: Progress in International Reading Literacy Study (PIRLS). This Working Paper compares the frameworks, texts, and items of these two international studies. The paper is divided into the following sections: Executive Summary; Introduction; Definitions of Reading; Purposes for Reading; Reading Passages; Distribution of Item Types in NAEP and PIRLS; Reading Processes Assessed by NAEP and PIRLS; To What Extent Do NAEP and PIRLS Measure Similar Skills?; Detailed View of Cross-Classification; and Conclusion. Appended are: Expert Panel Members; Example Passages; Readability and Lexile Analysis; and Classification of Items. (Contains 7 references, 14 tables, and 4 figures.) (NKA)

NATIONAL CENTER FOR EDUCATION STATISTICS

Working Paper Series

A Content Comparison of the NAEP and PIRLS Fourth-Grade Reading Assessments

Working Paper No. 2003-10 April 2003

Contact: Marilyn Binkley
Assessment Division
Marilyn.Binkley@ed.gov

U. S. Department of Education
Office of Educational Research and Improvement

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
Office of Educational Research and Improvement

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. A Content Comparison of the NAEP and PIRLS Fourth-Grade Reading Assessments, NCES 2003–10, by Marilyn Binkley and Dana Kelly. Project Officer: Marilyn Binkley. 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:
Marilyn Binkley
(202) 502–7492
Marilyn.Binkley@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

A Content Comparison of the NAEP and PIRLS Fourth-Grade Reading Assessments

Prepared by:

Marilyn Binkley
National Center for Education Statistics
Dana L. Kelly
Education Statistics Services Institute

Prepared for:

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

April 2003

Contents

Executive Summary

In 2003, NCES will be releasing results for both the 2001 PIRLS fourth-grade assessment and the 2002 NAEP fourth grade reading assessment. In anticipation of questions about how these two assessments compare, NCES convened an expert panel to compare the content of the PIRLS and NAEP assessments and determine if they are measuring the same construct. This involved a close examination of how PIRLS and NAEP define reading, the texts used as the basis for the assessments, and the reading processes required of students in each.

The comparison of the NAEP and PIRLS fourth-grade reading assessments suggests that there is a great deal of overlap in what the two assessments are measuring. While they do seem to defining and measuring the same kind of "reading," PIRLS is an easier assessment than NAEP, with more text-based tasks and shorter, less complex reading passages. The similarities and differences between the two are listed below.

Similarities

- PIRLS and NAEP define "reading" similarly, as a constructive process.
- PIRLS and NAEP assess reading for a literary experience and reading to be informed.
- PIRLS and NAEP call for students to develop interpretations, make connections across text, and evaluate aspects of what they have read.
- PIRLS and NAEP use authentic texts as the basis for the reading assessment.
- PIRLS and NAEP use multiple-choice and constructed response questions with similar distributions of these types of questions. In both, about half of the items are constructed-response format.

Differences

- PIRLS calls for more text-based interpretation than NAEP. NAEP places more emphasis on having students take what they have read and connect to other readings or knowledge and to critically evaluate what they have read.
- Close to 20% (18%) of the items in PIRLS require students to locate information in the text that is virtually an identical match to what is in the stem of the item. NAEP does not have any items requiring a verbatim match.
- PIRLS reading passages are, on average, about half the length of the NAEP reading passages. PIRLS passages are, on average, about 547 words, while NAEP passages are, on average, about 1000 words.

- Results of readability analyses suggest that the PIRLS reading passages are easier than the NAEP passages (one to two grade levels lower, on average)

A Content Comparison of the NAEP and PIRLS Fourth-Grade Reading Assessments

Introduction

In April 2003, NCES will release results of fourth-grade students' reading achievement in the United States from two assessments: the 2002 National Assessment of Educational Progress (NAEP) reading assessment and the 2001 Progress in International Reading Literacy Study (PIRLS) reading literacy assessment. PIRLS is a study of the International Association for the Evaluation of Educational Achievement (IEA) and was administered in 2001 in 35 countries. It is anticipated that when the results for NAEP and PIRLS are released there will be questions about how the two assessments compare. Policymakers, educators, and the general public are likely to wonder whether NAEP and PIRLS are assessing the same thing. In anticipation of this question, a comparison of the content of the two fourth-grade assessments was carried out. The frameworks, reading passages, and assessment items for each assessment were examined and systematically compared. This paper describes the analysis and how the content of the two assessments compare.

The NAEP reading assessment is based on a framework first developed for the 1992 assessment through a widely deliberative process involving teacher, curriculum specialists, reading researchers, policymakers, and the general public representing a broad cross-section of the United States (National Assessment Governing Board (NAGB), 2001*)*. The PIRLS framework and assessment were developed through a collaborative process involving reading and assessment specialists from the participating countries (Campbell et al. 2001). Both NAEP and PIRLS assessed nationally-representative samples of fourth-graders in the United States. The international target population for PIRLS was the "upper of the two adjacent grades with the most nine-year-olds"; this corresponds to the fourth grade in most countries but the third and fifth grades in some countries in the study.

The NAEP-PIRLS comparison study reported on in this paper is based on a similar comparison conducted by NCES of the IEA's 1991 reading literacy study and NAEP (Binkley and Rust, 1994). In the 1991 IEA reading literacy study, United States fourth-graders performed very well compared to their counterparts in other countries, ranking second out of 23 countries. This was incongruous with how fourth-grade reading achievement was being reported by NAEP-- only 29 percent of fourth-graders in 1992 met the Proficient Achievement Level set by the National Assessment Governing Board (Mullis, Campbell, and Farstrup 1993). In the analysis of the 1991 IEA study and NAEP, the two assessments were compared in terms of how they defined reading, the aspects of reading they measured, and the kinds of texts they used. The analysis revealed that the 1991 IEA reading literacy study assessed a small subset of what NAEP assessed, and contained less challenging texts and tasks (Binkley and Williams 1996).

The 2001 PIRLS assessment is based on a new framework that aims to assess a broader and higher level of reading than the 1991 study. The passages in PIRLS are longer and more

developed than those in the 1991 IEA study, and the kinds of questions that students are asked are more probing and cover a broader range of reading processes. A comparison of the 1991 IEA study and PIRLS showed that these two assessments are markedly different with respect to the texts used and the extent to which students are asked to interpret and think critically about what they have read (Kapinus, 2002).

This paper attempts to answer the inevitable question of whether NAEP and PIRLS are measuring the same thing. Using the comparison of NAEP and the 1991 IEA study as a model, the NAEP and PIRLS fourth-grade assessments were compared by a group of individuals who, collectively, have extensive experience with both assessments, including developing the frameworks and items (this group, listed in Appendix A, is hereafter referred to as the expert panel). A subset of the expert panel conducted a systematic comparison of the items in both assessments.

While content is an important consideration in how the two assessments compare, and is the focus of this paper, other aspects of the assessments should be examined as well, including how the achievement results are scaled and reported. In this paper, we address the content of the assessment, and leave the comparison with respect to other aspects of assessment to other researchers to carry out.

We begin by describing how each assessment defines reading. We examine their formal definitions and the aspects of reading they assessed. We then compare the passages in each assessment with respect to the kinds of texts as well as their lengths and difficulty. We conclude by describing how the assessments compare with respect to types of reading processes and skills evaluated in each assessment. These comparisons are based on a systematic classification of items by categories in the two frameworks.

Definitions of Reading

NAEP's definition of reading literacy is reflected in the following excerpts from the framework (NAGB, 2001).[1]

> The term reading literacy is not intended to imply only basic or functional literacy. Rather the term connotes a broader sense of reading, including knowing when to read, how to read, and how to reflect on what has been read. Contemporary research indicates that reading is a complex process that involves an interaction among the reader, text, and the context in which something is read. (p. 8)
>
> Reading for meaning involves a dynamic, complex interaction among three elements: the reader, the text, and the context. The context of a reading situation includes the purposes for reading that the reader might use in building a meaning of the text...Good readers bring to this interaction their prior knowledge about the topic of the text and their purposes for reading it, as well as their skill in reading, which includes their knowledge about the reading process and the structure of texts. (p. 12)

[1] Although the NAEP reading framework was "revisited" and updated for the 2003 assessment, resulting in changes to the names of categories, because this analysis focuses on the 2002 assessment, the framework categories used for the 2002 and previously are used.

> Readers respond to a given text in a variety of ways as they use background knowledge and information from the text to construct an initial understanding, develop an interpretation to extend the text's meaning, and examine the meaning so as to respond personally and critically to the text. (p. 15)

The PIRLS framework (Campbell et al., 2001) defines reading literacy for PIRLS as

> the ability to understand and use those written 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. (p. 3)

The PIRLS framework continues:

> Readers are regarded as actively constructing meaning and as knowing effective reading strategies and how to reflect on reading. They have positive attitudes toward reading and read both for recreation and to acquire information. Meaning is constructed in the interaction between reader and text in the context of a particular reading experience. The reader brings a repertoire of skills, cognitive and metacognitive strategies, and background knowledge. The text contains certain language and structural elements and focuses on a particular topic. The context of the reading situation promotes engagement and motivation to read, and often places specific demands on the reader (p. 3).

The expert panel noted that there is considerable overlap in how NAEP and PIRLS define reading. Both frameworks acknowledge that reading is a constructive and interactive process involving interaction between the reader and the text. Both address that the context for reading is an important element in how readers make meaning and in the skills and strategies they use. Both acknowledge that the structural elements of a text will influence a reader's strategies. Authors of the updated NAEP reading framework (2002) also noted that the two assessments had similarities. They wrote that the definitions of reading literacy "convey the notion that reading involves developing an understanding of text, thinking about it, and using various texts for many different purposes…The congruence in framework definitions clearly represents a growing international agreement on the important dimensions of reading literacy" (NAGB 2002, p. 8).

The panel comparing NAEP and PIRLS also noted differences in what is emphasized in the NAEP and PIRLS definitions of reading. The PIRLS definition is more explicitly targeted to younger readers and the reading tasks and processes in which children engage, noting in the definition the purposes for which young readers read (to learn, to participate in communities of readers, and for enjoyment). Of course, PIRLS assessed one population, fourth-graders, while NAEP assesses three populations across a broad age and grade span. Another difference is the emphasis on readers' response to text: in its definition, NAEP appears to place more emphasis on students' personal response to a text than does PIRLS. NAEP devotes an entire category of items to reader-text connection (a target of 15% of student assessment time on items classified this way), while PIRLS does not have a separate category for this type of item.

The NAEP and PIRLS reading assessments are each based on a two-dimensional matrix, with the *purposes of reading* on one dimension and the *processes of reading*[2] on the other. In both assessments, each process is assessed within each purpose for reading. The purposes for reading and processes of comprehension assessed by each assessment are shown in Figures 1 and 2.

Figure 1: NAEP Framework Dimensions[3]

	Purposes for Reading[4]	
Processes	Reading for Literary Experience	Reading for Information
Forming an Initial Understanding		
Developing an Interpretation		
Personal Reflection and Response		
Demonstrating a Critical Stance		

Source: National Assessment of Educational Progress, 2002.

Figure 2: PIRLS Framework Dimensions

	Purposes for Reading	
Processes	Reading for Literary Experience	Reading to 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		

Source: IEA's Progress in International Reading Literacy Study, 2001

[2] Although NAEP refers to these as "reading stances" or aspects of "constructing, extending, and examining meaning," they are referred to as "processes" in this document for ease of comparison.

[3] The updated framework (NAGB 2002) uses the following labels for the four process categories: (1) forming a general understanding; (2) developing interpretation; (3) making reader-text connections; and (4) examining content and structure.

[4] At fourth-grade, NAEP assesses two purposes for reading; at eighth- and twelfth grades NAEP assesses a third purpose—Reading to Perform a Task.

Purposes for Reading

Both NAEP and PIRLS assess and report on two purposes for which young readers read,[5] reading for a literary experience and reading for information. Both call the former purpose as "reading for literacy experience." NAEP calls its informational category "reading for information," while PIRLS calls it "reading to acquire and use information." Each assessment devotes about half of the assessment to each reading purpose. NAEP devotes 55% of the assessment to the literary purpose and 45% to the informational purpose and PIRLS devotes 50% to each purpose.

The two literary purposes are defined in very similar ways. The NAEP framework says:

> Reading for literary experience usually involves the reading of novels, short stories, poems, plays, and essays. In these reading situations, readers explore the human condition and consider interplays among events, emotions and possibilities. In reading for literary experience, readers are guided by what and how an author might write in a specific genre and by their expectations of how the text will be organized. The readers' orientation when reading for literary experience usually involves looking for how the author explores or uncovers experiences and engaging in vicarious experiences through the text. (NAGB, 2001, p. 13)

The PIRLS framework describes reading for literary experience as follows.

> In literary reading, the reader engages with the text to become involved in imagined events, settings, actions, consequences, characters, atmosphere, feelings and ideas, and to enjoy language itself. To understand and appreciate literature, the reader must bring to the text his or her own experiences, feelings, appreciation of language and knowledge of literary forms. For young readers, literature offers the opportunity to explore situations and feelings they have not yet encountered, and to experience imaginatively an autonomy not yet available to them....Events, actions, and consequences depicted in narrative fiction allow the reader to experience vicariously and reflect upon situations that, although they may be fantasy, illuminate those of real life. The text may present the perspective of the narrator or a principle character, or there may be several such viewpoints in a more complex text. Information and ideas may be described directly or through dialogue and events. Short stories or novels sometimes narrate events chronologically, or sometimes make more complex use of time with flashbacks or time shifts. (Campbell et al. 2001, p. 17)

The two assessments both address reading as a means of acquiring information. NAEP emphasizes the reading of information texts in order to read and understand informational texts for a variety of purposes:

> Reading to be informed usually involves the reading of articles in magazines and newspapers, chapters in textbooks, entries in encyclopedias and catalogues, and books on particular topics. The type of prose found in such texts has its own features and readers need to be aware of those features to understand it...Readers read to be informed for different purposes; for example, to find specific pieces of

[5] NAEP also calls the "purposes" "reading situations."

> information when preparing a research project or to get some general information when glancing through a magazine article. (NAGB, 2001, p. 13)

In addition to addressing reading to *acquire* information, PIRLS also addresses reading to *use* information and in that way goes beyond NAEP's informational category (NAEP assesses "reading to perform a task" at eighth- and twelfth-grades). The PIRLS framework says the following:

> In reading for information, the reader engages not with imagined worlds, but with aspects of the real universe. Through information texts, one can understand how the world is and has been, and why things work as they do. Readers can go beyond the acquisition of information and use it in reasoning and in action. Information texts need not be read from beginning to end; readers may select the parts they need. (Campbell et al. 2001, p. 17)

PIRLS does not include a separate reading purpose to address the use of reading to perform a task, such as the use of documents like maps or charts or the use of instructions on how to assemble something because generally young children are not involved in that type of reading very often. However, PIRLS does acknowledge that young children do engage in this type of reading sometimes by including it in the informational category. The PIRLS framework points out that information texts can be organized chronologically or non-chronologically (or may include both), which may require the reader to draw upon different skills.

Reading Passages

The types of reading passages included in NAEP and PIRLS reflect the purposes for reading that are assessed. In NAEP, students are presented with short stories, legends, biographies, and folktales in the assessment of reading for literary experience. In PIRLS, similarly, narrative fiction in the form of short stories serves as the basis for the PIRLS assessment of reading for literary experience. NAEP's assessment of reading to be informed includes magazine articles that focus on people, places, and events of interest to children. Similarly, in PIRLS, students are presented with articles about people, places, and events, as well as an informational brochure.

Altogether, the NAEP and PIRLS fourth-grade assessments each comprise eight reading passages, four for each purpose. Each student, however, reads and answers questions about two passages. Appendix B contains a passage from each assessment; one informational text from NAEP (administered in 1998) and one literary passage for PIRLS to illustrate the nature of the texts used in the two assessments. These passages will be referenced in discussions about the assessments to follow.

Both NAEP and PIRLS strive to have the assessment be an "authentic" reading experience for students. That is, both try to present tasks that are similar to what students would actually encounter when reading in and out of school. The NAEP framework calls for the use of "authentic" texts. All passages were previously published in children's magazines or other publications and generally are not edited at all for use in NAEP. Also, the format is kept as close to the original as possible. In selecting passages, NAEP uses passages that are

like those that students encounter in and out of school, but that were published long enough ago that the students taking the assessment probably have not read the particular passages.

PIRLS also strives to use authentic, previously published texts. However, PIRLS has a more liberal policy on editing and changing the format of the texts that are used. Largely because of the constraints inherent in an assessment that is administered across languages and cultures, PIRLS did change some of the original language in the passages so that it would be acceptable to a more diverse audience, could be more easily translated and retain the same meaning, and could meet requirements for length.

One feature of the PIRLS assessment that speaks to authenticity is the use of the "PIRLS Reader." While most students in the assessment get a test booklet that contains two sections, each with one reading passage and 11-14 questions about the passage, some students receive a color booklet with two passages (one literary and one informational) and the accompanying items in a separate test booklet. The purpose of this layout was to evoke a more authentic reading experience for students. About a quarter of the students in the PIRLS assessment were administered the PIRLS Reader.

How Interchangeable are the NAEP and PIRLS Reading Passages?

After reading all of the passages in both assessments, the expert panel concluded that most of the NAEP passages could appear on PIRLS and most of the PIRLS passages could appear on NAEP, although overall the PIRLS passages were shorter and appeared less complex than the NAEP passages. This judgment was made without the benefit of word counts or results of the empirical analyses.

For both NAEP and PIRLS, there were two passages that the panel felt would not appear on the other assessment. The panel felt that two of the eight NAEP passages would be inappropriate for PIRLS because of their length and complexity. Both of these texts (one a short story and one an informational article) were included in NAEP at both fourth- and eighth grades and were thus the more difficult of the NAEP fourth-grade passages.

Of the eight PIRLS passages, the panel felt that all but two would be appropriate for NAEP. One of the two that the panel felt would not appear on NAEP was an informational brochure that the group felt was more appropriate for assessing "reading to perform a task" than "reading to be informed," and therefore might appear on NAEP at eighth grade where "reading to perform a task" is assessed. The other passage was considered inappropriate because it had been excerpted from a book and revised substantially and therefore lacked the authenticity required by NAEP.

Length

The panel noted in their review of the NAEP and PIRLS passages that the PIRLS passages were, on average, shorter. Indeed, the NAEP reading passages are, on average, twice as long as the PIRLS reading passages. As shown in Table 1, NAEP passages range in length from 691 to 1365 words, with an average word length of 1000. In contrast, PIRLS passages range in length from 293 (a brochure) and 322 (mainly continuous text) to 804 words, with an average of 547 words. The passages included in Appendix B illustrate the difference in length. The PIRLS story, *Upside-Down Mice*, has 521 words, close to the average for the

PIRLS passages. The NAEP passage, *Blue Crabs* (administered in 1998 and not included in the counts in Table 1), is an informational article that is 880 words, a typical length for the NAEP fourth-grade passages.

Table 1: Number of Words in NAEP and PIRLS Fourth-Grade Reading Passages

	Number of Words in Each Passage	Average
NAEP Passages	691, 790, 851, 946, 1038, 1120, 1196, 1365	**1000**
PIRLS Passages	293, 322, 521, 537, 546, 621, 730, 804	**547**

Source: National Assessment of Educational Progress, 2002 Assessment; IEA's Progress in International Reading Literacy Assessment, 2001 Assessment.

Passage Difficulty

While the panel of experts recognized differences in difficulty between the NAEP and PIRLS passages, and noted the differences in length, they wanted to provide a more empirically-based measure of difficulty. Two widely used readability formulas, the Fry and the Flesch, were used to obtain measures for the NAEP and PIRLS fourth-grade reading passages. In addition, a Lexile analysis was conducted to obtain an additional measure that indicates the reading demand of the text in terms of the semantic difficulty (vocabulary) and syntactic complexity (sentence length).

Readability

For the Fry and Flesh readability formulas, the number of words and syllables in each of three 100-words samples (per text) are counted and used as the basis for determining an age and grade level for the text. See Appendix C for a description of the formulas and procedures used. Counting the number of sentences or syllables per 100 words is quite straightforward and easily replicable across raters and across time. However, readability formulae go one step further. They associate length of words and sentences with increasing difficulty that can then again be associated with the increasing skill that is accrued with age or grade level. This connection is derived from materials used in classrooms at a particular point in time. The strength of this association was used as a predictor of the probability that other materials with the same degree of characteristics would then be appropriate at the age or grade level as well.

Counts of the number of syllables and sentences per 100 words provide some evidence that the NAEP passages are, overall, more complex. As shown in Table 2, NAEP has, on average, more syllables per 100 words than PIRLS (138.2 compared with 133.6), suggesting the use of longer and perhaps more unfamiliar words in the NAEP passages. NAEP has an average of 6.8 sentences per 100 words, compared with 8.6 sentences for PIRLS, indicating that the NAEP sentences are longer (also in Table 2). The difference in the average number of sentences suggests that the NAEP texts are more complex, with more embedded clauses and a more complex syntactical structure.

Table 2: Average Number of Sentences and Syllables per 100 Words, NAEP and PIRLS Reading Passages

	Sentences	Syllables
NAEP	6.8	138.2
PIRLS	8.6	133.6

Note: See Appendix C for details on analyses

Based on the Fry analysis, as shown in Table 3, predictions could be made that on average the NAEP passages would be appropriate for grade 7 (and average age of 12.3) and PIRLS for grade 5 (and average age of 10.4). Similarly, the Flesch analysis indicated that the NAEP passages would be appropriate for grade 7 and PIRLS for grade 5-6. The Flesch analysis provides a "reading ease" score rather than an age, and it too shows that NAEP is more difficult; NAEP was rated "fairly easy" while PIRLS was rated "easy." Appendix C presents results of the Fry and Flesch analyses on each of the NAEP and PIRLS reading passages.

Overall, both formulae indicate about a two grade-level spread between NAEP, which is more difficult, and PIRLS. Therefore, the panelists' impression was confirmed by a less subjective measure.

Table 3: Average Age and Grade Level Determined by Fry and Flesch Readability Analyses: PIRLS and NAEP Fourth Grade Reading Passages

	Fry Analysis		Flesch Analysis	
	Average Age	Average Grade Level	Average Reading Ease	Average Grade Level
NAEP	12.3	6.9	Fairly Easy (74.5)	7th
PIRLS	10.4	5.0	Easy (81.7)	5th to 6th

Note: See Appendix for details on analyses

Although the results of the readability analyses support the panel's observation that the NAEP passages were more difficult than those in PIRLS, there are some caveats associated with readability formulae. Historically, readability formulae have been used to provide a quick estimate of the difficulty and grade appropriateness of a text. In general, as a class of tools they were good predictors for these purposes. However, as more and more states moved towards stringent requirements related to the evaluation of grade appropriateness for textbook adoption, these indicators were found to be insufficient for the expanded purpose and fell into disfavor. The reading formulae did not account for a sufficient number of salient text features that figure prominently in comprehension, such as variation in vocabulary, topic, and appeal.

Lexile Analysis

Whether the NAEP and PIRLS passages are or are not appropriate for fourth-graders could not be concluded based on older readability formulae not only because of their more limited measurement, but also because the norming information that is the basis of the two readability formulae was established over twenty-five years ago. To address this issue a more current measure of passage difficulty was obtained using the Lexile system (see Appendix C for description of procedure). The Lexile provides a measure of difficulty that takes into account semantic difficulty (vocabulary) and syntactic complexity (sentence length) and assigns an appropriate grade level. Developed and normed in recent years, Lexiles provide a more accurate measure of grade appropriateness. As shown in Table 4, the NAEP passages were determined to be appropriate for 4th to 5th grade and the PIRLS passages appropriate for 3rd to 4th grade. This is consistent with the judgment of the expert panel.

Table 4: Lexile Score and Corresponding Grade Levels for PIRLS and NAEP Fourth Grade Reading Passages

	Lexile Score	Corresponding Grade Level
NAEP Passages	818.8	4th or 5th
PIRLS Passages	695.7	3rd or 4th

Note: See Appendix C for details on analyses

Distribution of Item Types in NAEP and PIRLS

In both NAEP and PIRLS, students respond to multiple-choice and constructed-response questions (short answer and extended response) based on the passages, and the distribution of these item types are similar in both assessment (Table 4). In each assessment, close to half of the items are in multiple-choice format, with four response options (45% of NAEP items and 47% of PIRLS items). Both NAEP and PIRLS include constructed-response items that are scored for two, three, or four levels of correctness (including incorrect), with "short constructed-response" items scored for two or three levels and "extended constructed response" items scored for four levels. Each item has a scoring guide that delineates the requirements for different levels. Scoring guides are used by scorers to make reliable scoring decisions. The percentages of short and extended constructed-response items in each assessment are comparable: In NAEP, 45% of the items are short-constructed response and 10% are extended; in PIRLS, 44% of the items are short-constructed response and 8% are extended.

Table 4: Distribution of Item Types in NAEP and PIRLS, based on Percentage of the Number of Items in Each Assessment

	Multiple-choice	Short Constructed Response	Extended Constructed Response
NAEP (Fourth-grade)	45%	45%	10%
PIRLS	47%	44%	8%

Source: National Assessment of Educational Progress, 2002 Assessment; IEA's Progress in International Reading Literacy Assessment, 2001 Assessment.

Reading Processes Assessed by NAEP and PIRLS

While the purposes for reading and the types of texts used form the content of the assessment, the *reading processes* define the skills and abilities that students must draw on in response to the texts. Each item is written to address one process. Both NAEP and PIRLS have four categories of reading processes; each is described briefly below.

NAEP Processes Categories

Forming an Initial Understanding: "requires the reader to provide an initial impression or global understanding of what was read. It involves considering the text as a whole or in a broad perspective..."

Developing an Interpretation: "...requires the reader to go beyond the initial impression to develop a more complex understanding of what was read. It involves linking information across parts of a text as well as focusing on specific information..."

Personal Reflection and Response: "...require[s] the reader to connect knowledge from the text with his or her personal background knowledge. The focus here is on how the text relates to personal knowledge."

Developing a Critical Stance: "...requires the reader to stand apart from the text and consider it objectively. It involves a range of tasks including such behaviors as critical evaluation, comparing and contrasting, application to practical tasks, and understanding the impact of such text features as irony, humor, and organization." (Reading Framework 2001, pp 15–17)

PIRLS Process Categories

Focus on and Retrieve Explicitly Stated Information: "Successful retrieval requires a fairly immediate or automatic understanding of the text. This process requires little or no inferring or interpreting. There are no "gaps" in meaning to be filled – the

meaning is evident and stated in the text....Focus on the text typically remains at the sentence or phrase level in this type of processing."

Make Straightforward Inferences: "Making inferences allows the reader to move beyond the surface of texts and to fill in the "gaps" in meaning that often occur in texts. ... Although the ideas may be explicitly stated, the connection between them is not, and thus must be inferred. Straightforward inferences are very much text-based."

Interpret and Integrate Ideas and Information: "...the reader is processing text beyond the phrase or sentence level. ...When they interpret and integrate text information and ideas, readers may need to draw on their background knowledge and experience more than they do for straightforward inferences. ...By engaging in this interpretive process, readers are attempting to construct a more specific or more complete understanding of the text by integrating personal knowledge and experience with meaning that resides in the text."

Examine and Evaluate Content, Language, and Textual Elements: "The reader engaged in this process is standing apart from the text and examining or evaluating it. The text content, or meaning, may be examined from a very personal perspective or with a critical or objective view...In examining or evaluating elements of text structure and language, readers draw upon their knowledge of language usage and general or genre-specific features of texts." (Campbell et al. 2001, pp 10–15)

The expert panel examined the descriptions of the categories and discussed the apparent similarities and differences among them in an attempt to determine where categories might overlap in what they assessed. Their initial impressions of how the categories compared, prior to systematically classifying the items, suggested that overall the assessments were calling on similar reading skills, but that they do not break out and describe the skills in the same manner in the framework. The panel noted similarities between NAEP's "Developing an interpretation" and PIRLS's "Make straightforward inferences" and "interpret and integrate information and ideas" and some overlap between NAEP's "demonstrating a critical stance" and PIRLS's "Examine and evaluate content, language, and textual elements."

The panel's initial impression that the two assessments are, overall, assessing the same domain was not upheld entirely when it delved more deeply into the content by examining and classifying the items. Through this exercise, described below, the panel determined that while there are many similarities in what is measured on each assessment, there also are distinct differences.

First, in Figures 1 and 2 we present items from NAEP and PIRLS, respectively, to illustrate how the different reading processes are assessed (passages are included in Appendix B). The example NAEP passage was administered in 2000.

Figure 1: Example NAEP Fourth-grade Items

Students read an informational passage, A Brick to Cuddle Up To, *about how people kept warm during colonial times. See Appendix B.*

Developing an initial understanding

1) You would probably read this article if you wanted to know how the colonists
   - A) cooked their food
   - B) traveled in the winter
   - C) washed their clothes
   - *D) kept warm in cold weather.

Developing an interpretation

2) A colonist would probably have used a foot stove when
   - *A) going on a trip
   - B) sleeping in a bed
   - C) sitting by the fireplace
   - D) working around the house

Source: National Assessment of Educational Progress, 2000 Assessment.

Students read an informational passage, Blue Crabs, *describing the experiences of hunting for and catching blue crabs. It includes information about the crabs' appearance, habits, habitats, and survival techniques, and what it is like to hunt for them. See Appendix B*

Developing an interpretation

3) Why does a blue crab hide after molting?

So that a predator does not eat him while he is growing back a new shell.

Personal reflection and response

4) What is the most interesting thing you learned from the passage about blue crabs?

that they can loose a leg and grow it back

Demonstrating a critical stance

The author of the article helps you to learn about blue crabs by

A) explaining why they are an endangered species
B) comparing them to other arthropods
C) discussing their place in the food chain
*D) providing details about their unique characteristics

Source: National Assessment of Educational Progress, 1998 Assessment.

Figure 2: Example PIRLS Fourth-grade Items

Students read a short story by Roald Dahl called "Upside-down Mice." In the story, an elderly man with mice in his house carries out a plan to trick the mice and get rid of them. The mice observe the man as he carries out his plan, unaware of their fate. See Appendix B.

Focus on and retrieve explicitly stated information

1) Where did Labon put the mousetraps?

A) in a basket

B) near the mouseholes

C) under the chairs

*D) on the ceiling

Make straightforward inferences

2) Why did Labon want to get rid of the mice?

A) He had always hated mice.

*B) There were too many of them.

C) They laughed too loudly

D) They ate all his cheese.

Interpret and integrate ideas and information

3) You learn what Labon is like from the things he does. Describe what he is like, and give **two** examples

Labon is tricky because he glued the mousetraps to the ceiling on the first night just to make them think he was silly. Then he glued the furniture to the ceiling to make them think they were upside down

Examine and evaluate content, language, and textual elements

4) Think about what Labon and the mice did in the story. Explain what makes the story unbelievable.

No one would be able to glue furniture to the ceiling like Labon did

Source: IEA's Progress in International Reading Literacy Assessment, 2001 Assessment.

To What Extent Do NAEP and PIRLS Measure Similar Skills?

To understand the extent to which each assessment measured aspects of reading comprehension measured by the other, the 82 NAEP and 98 PIRLS items were classified by the panel by the process categories in each other's frameworks. That is, each NAEP item was classified by one of the four PIRLS process categories and each PIRLS item was classified by one of the four NAEP process categories. The results of this cross-classification reveal considerable overlap in what the two assessments are assessing, yet also some striking differences in what each emphasizes.

Table 5 shows the distribution of items in each assessment by the PIRLS classification scheme. The first row, NAEP items, shows the percentage of NAEP items in each PIRLS category, according to the expert panel. The second row, PIRLS items, shows the percentage of items in each category according to the PIRLS assessment developers.[6]

Cross-Classification of Items by PIRLS Framework

Both assessments have similar levels of emphasis on students' ability to make straightforward inferences. In NAEP and PIRLS, a little more than a quarter of the items (27% for NAEP and 28% for PIRLS) assess this skill. The two assessments do not place similar levels of emphasis on the other PIRLS process categories. For example, NAEP places more weight on complex interpretation than does PIRLS. Nearly half of NAEP items (46%) were classified as "Interpret and integrate ideas and information," while about one-third (32%) of PIRLS items were. Compared to NAEP, PIRLS places greater emphasis on examining and evaluating content, language, and textual elements (15% in PIRLS compared with 9% in NAEP). PIRLS also places more emphasis on focusing on and retrieving information than NAEP—26% of the PIRLS items were classified "Focus on and retrieve explicitly stated information" compared with 18% of NAEP items.

Table 5: NAEP and PIRLS Items Classified by PIRLS Framework

	Focus on and retrieve explicitly stated information	Make straightforward inferences	Interpret and integrate ideas and information	Examine and evaluate content, language, and textual elements
NAEP Items*	18%	27%	46%	9%
PIRLS Items**	26%	28%	32%	15%

*As classified by expert panel.
**As classified by PIRLS developers.

The cross-classification exercise involved having the panel members classify the PIRLS items by the PIRLS categories, although the PIRLS classifications reported in Table 5 are the PIRLS item classifications according to the developers of the PIRLS assessment (the "actual" item classifications). But, to what extent did the panel agree with the classifications of the PIRLS items? Table D.1 in the appendix shows panel had a nearly identical

[6] Classification of PIRLS items was provided by the PIRLS International Study Center, Boston College.

distribution of items across the PIRLS categories when it classified the PIRLS items by the PIRLS framework. This suggests that each category in PIRLS is distinct from every other category, with little overlap and also that the PIRLS framework was explicit enough to communicate the types of skills and processes subsumed under each category.

Cross-Classification of Items by NAEP Framework

Table 6 shows the distribution of items in each assessment according to the NAEP classification scheme. Again, we see differences in what each assessment emphasizes. Few PIRLS items fit into "Forming an initial understanding" (4%) and "Personal reflection and response" (3%), while 10% and 15% of NAEP items were so classified by NAEP developers. While 12% of PIRLS items were classified as "Demonstrating a critical stance," 21% of NAEP items were classified this way. The area in which more PIRLS items than NAEP items were classified was "Developing an interpretation" (62% compared with 55%).

The most striking outcome of this exercise was that it was not possible to fit 18% of the PIRLS items into the NAEP framework. These were items that asked students to locate information in the text that was virtually an identical match between the item and the text. These items are shown in the table in the category labeled "other." It is expected that such items would be easier for students, although other factors beyond the match between the item stem and the text may influence difficulty, and these features were not examined. We can conclude, however, that 18% of PIRLS items are unlike those that appear on NAEP since NAEP does not include items of this nature.

An example of this type of item is shown in Figure 2 (page 17) as example item 1. There, students are asked to locate a specifically stated piece of information and although they need to identify the appropriate area of the text and determine which piece of information is relevant, the information is explicitly stated and there is no need for interpretation.

Table 6: NAEP and PIRLS Items Classified by NAEP Framework

	Forming an initial understanding	Developing an interpretation	Personal reflection and response	Demonstrating a critical stance	Other
NAEP Items*	10%	55%	15%	21%	0%
PIRLS Items**	4%	62%	3%	12%	18%

*As classified by NAEP assessment developers.
**As classified by panel.

The cross-classification exercise involved having the panel members classify the NAEP items by the NAEP categories, although the classifications reported in Table 5 are the NAEP item classifications according to NAEP developers (the "actual" item classifications). But, to what extent did the panel agree with the classifications of the NAEP items? Unlike the close agreement between the panel and PIRLS assessment developers (Table D.1 in Appendix D), the panel was not in full agreement with the "actual" NAEP item classifications. As shown in Table D.1, the panel classified more items "Developing an interpretation" and fewer items "Demonstrating a critical stance" than the NAEP developers. This suggests that there is some overlap in what these categories are assessing

and perhaps some ambiguity in how they are described and communicated in the framework. It should be noted that in coming to consensus on how each item should be classified, the panel found it more difficult to agree with each other on the NAEP items, than on the PIRLS item classifications, and that could be contributing to the disagreement with the "actual" classifications.

Detailed View of Cross-Classification

In order to more fully understand how the NAEP and PIRLS assessments compare, the following tables show more detailed information about how the expert panel classified the items in each. Taken together, the distributions suggest that the NAEP process categories are broader than those in PIRLS. Items from a category in the NAEP framework tend to be dispersed across the PIRLS categories, while items from a category in the PIRLS framework tend to fall fairly cleanly into one NAEP category.

Cross-Classification of Items by PIRLS Framework

Table 7 shows the items classified as NAEP's "Forming an initial understanding" distributed across PIRLS categories according to how the committee classified the items during the cross-classification exercise: 13% were classified by the panel as "Focus on and retrieve explicitly stated information," 25% were classified "Make straightforward inferences," 63% were classified "Interpret and integrate ideas and information," and 0% were classified "Examine and evaluate content, language, and textual elements." This suggests that the "Forming an initial understanding" category in NAEP includes items assessing retrieval, inference, and interpretation skills.

Table 7: NAEP "Forming an initial understanding" Items Classified by PIRLS Framework Categories

PIRLS Process Category	Percent of Items According to Panel
Focus on and retrieve explicitly stated information	13%
Make straightforward inferences	25%
Interpret and integrate ideas and information	63%
Examine and evaluate content, language, and textual elements	0%
Total	100%

Similar to items in the "Forming an initial understanding category," NAEP items classified as "Developing an interpretation" are dispersed across the PIRLS process categories (Table 8). About one-third of NAEP items fell into each of three PIRLS categories, and 2% in "Examine and evaluate content, language, and textual elements," suggesting that this category includes items tapping a range of skills, including retrieval of information.

Table 8: NAEP "Developing an interpretation" Items Classified by PIRLS Framework

PIRLS Process Category	Percent of Items According to Panel
Focus on and retrieve explicitly stated information	29%
Make straightforward inferences	36%
Interpret and integrate ideas and information	33%
Examine and evaluate content, language, and textual elements	2%
Total	100%

NAEP items classified as "Personal reflection and response" are not as dispersed across the PIRLS categories to the extent that items in the previous two categories are, suggesting that this is a more tightly defined category: three-quarters (75%) of these items were classified as "Interpret and integrate ideas and information." Close to 20% of these items were deemed to fit into "Examine and evaluate content, language, and textual elements." None of the items were classified as "Focus on and retrieve explicitly stated information" reflecting the fact that these items are not asking students to retrieve information from the text but rather to think about what they have read in reference to their own experiences or points of view. See Table 9.

Table 9: NAEP "Personal reflection and response" Items Classified by PIRLS Framework

PIRLS Process Category	Percent of Items According to Panel
focus on and retrieve explicitly stated information	0%
make straightforward inferences	8%
interpret and integrate ideas and information	75%
examine and evaluate content, language, and textual elements	17%
Total	100%

NAEP "Demonstrating a critical stance" items (Table 10) fall across the four PIRLS categories, suggesting that it too is a broad category of skills, including retrieving, making inferences, and examining and evaluating the text, and an emphasis (about half of the items) on requiring students to interpret and integrate ideas and information.

Table 10: NAEP "Demonstrating a critical stance" Items Classified by PIRLS Framework

PIRLS Process Category	Percent of Items According to Panel
Focus on and retrieve explicitly stated information	6%
Make straightforward inferences	18%
Interpret and integrate ideas and information	53%
Examine and evaluate content, language, and textual elements	24%
Total	100%

Cross-Classification of Items by NAEP Framework

In contrast to the way NAEP items from a single NAEP category were distributed across the PIRLS categories for three of the four NAEP categories, the PIRLS items tend to fit clearly into one or two NAEP categories (or not in any). As shown in Table 11, about two-thirds of the PIRLS "Focus on and retrieve explicitly stated information and ideas" items did not fit a NAEP category and were therefore classified as "Other." These were items that required students to make virtually an exact match between the item stem and text. Most of the remaining items were classified as "Developing an interpretation," suggesting that there is some overlap between this NAEP category and PIRLS' "Focus on and retrieve."

Table 11: PIRLS "Focus on and retrieve explicitly stated information and ideas" Items Classified by NAEP Framework

PIRLS Process Category	Percent of Items According to Panel
Forming an initial understanding	4%
Developing an interpretation	28%
Personal reflection and response	0%
Demonstrating a critical stance	0%
Other (not a NAEP category)	68%
Total	100%

Almost 90% of PIRLS "Make straightforward inferences" items were classified as "Developing an interpretation," indicating an overlap between the skills assessed by these two categories. Some of these items (11%) tap skills assessed by the "Demonstrating a critical stance" category.

Table 12: PIRLS "Make straightforward inferences" Items Classified by NAEP Framework

PIRLS Process Category	Percent of Items According to Panel
Forming an initial understanding	0%
Developing an interpretation	89%
Personal reflection and response	0%
Demonstrating a critical stance	11%
Total	100%

Nearly all PIRLS "Interpret and integrate" items were classified as NAEP's "Developing an interpretation," indicating an overlap in the skills assessed by items in these two categories (Table 13).

Table 13: PIRLS "Interpret and integrate ideas and information" Items Classified by NAEP Framework

PIRLS Process Category	Percent of Items According to Panel
Forming an initial understanding	0%
Developing an interpretation	94%
Personal reflection and response	3%
Demonstrating a critical stance	3%
Total	100%

PIRLS items classified as "Examine and evaluate content, language, and textual elements" cut across the NAEP process categories, with some items in each, although 53% were classified as "Demonstrating a critical stance" (Table 14).

Table 14: PIRLS "Examine and evaluate content, language, and textual elements" Items Classified by NAEP Framework

PIRLS Process Category	Percent of Items According to Panel
forming an initial understanding	20%
developing an interpretation	7%
personal reflection and response	13%
demonstrating a critical stance	53%
Other (not a NAEP category)	7%
Total	100%

Conclusion

The comparison of the NAEP and PIRLS fourth-grade reading assessments suggests that there is a great deal of overlap in what the two assessments are measuring. However, while they do seem to define reading in similar ways. PIRLS has more text-based tasks and shorter, less complex reading passages than NAEP. The similarities and differences between the two are described below.

This comparison revealed that overall, the NAEP and PIRLS reading assessments are quite similar. Both define "reading" similarly, as a constructive process. Both use high-quality reading passages and address similar purposes for which young children read—for literary experience and information. Both call for students to develop interpretations, make connections across text, and evaluate aspects of what they have read. Finally, both have a similar distribution of multiple-choice and constructed-response items: in each, about half of the items are constructed-response items.

While the two assessments have similar definitions of reading and assess many of the same aspects of reading, a closer look at how the domain is operationalized by each revealed that there are some important differences. NAEP places more emphasis than PIRLS on having students take what they have read and connect to other readings or knowledge. This is consistent with the value placed on readers' response to text when the NAEP reading framework was developed.

PIRLS places a greater emphasis than NAEP on text-based reading skills and interactions, including items that ask students to locate information in the text, make text-based inferences and interpretations, and evaluate aspects of the text. Perhaps the most striking difference is that close to 20% (18%) of the items in PIRLS require students to locate information in the text that is virtually an identical match to what is in the stem of the item. NAEP does not have any items requiring a verbatim match. NAEP does have items asking students to locate specific information or ideas in the text, but the information is not a verbatim match.

The PIRLS reading passages are, on average, about half the length of the NAEP reading passages. PIRLS passages are, on average, about 547 words, while NAEP passages are, on average, about 1000 words. Readability formulas indicate that the passages used in PIRLS are less complex than those used in NAEP (one to two grade levels lower, on average).

The classification of items also revealed differences in how the two frameworks function. The panel had an easier time classifying PIRLS and NAEP items by the PIRLS framework categories than by the NAEP framework categories. When classifying PIRLS items by the PIRLS framework the panel had almost 100% agreement with the IEA assessment developers. When classifying by the NAEP framework, however, there was more disagreement as to the categories in which NAEP and PIRLS items should be classified.

References

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. (1996). *Reading Literacy in the United States: Findings from the IEA Reading Literacy Study* (NCES 96–258). Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

Campbell, J.R., Kelly, 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.

Kapinus, B. (2002). *Comparative Analysis of the 1991 IEA Reading Study and PIRLS.* Unpublished manuscript.

Mullis, I.V.S., Campbell, J.R., and Farstrup, A.E. (1993). *NAEP 1992 Reading Report Card for the Nation and the States.* Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

National Assessment Governing Board. *Reading Framework for the 2003 National Assessment of Educational Progress.* (September 2002). Washington DC: Author.

National Assessment Governing Board. Reading Framework for the National Assessment of Educational Progress: 1992–2000, Washington, DC: Author.

Appendix A: Expert Panel Members

Peter Afflerbach
University of Maryland

Patricia Donahue
Educational Testing Service

Robert Finnegan
Educational Testing Service

Janet Jones
Berry Elementary School

Barbara Kapinus
Consultant

Irwin Kirsh
Educational Testing Service

Jesse Miller
Educational Testing Service

Jean Osborn
Consultant

Terry Salinger
American Institutes for Research

Staff

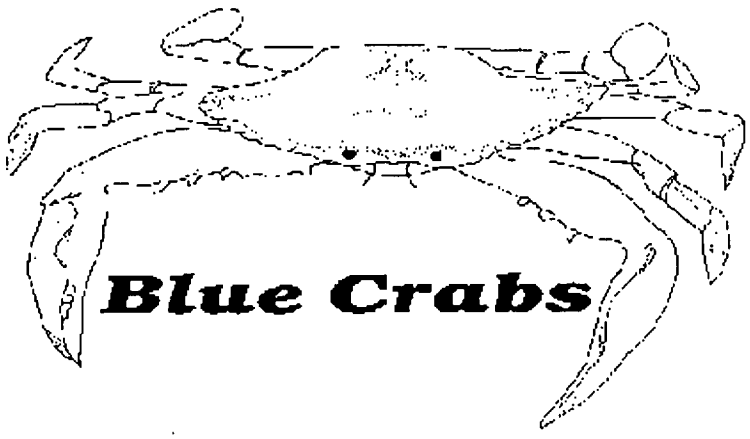
Marilyn Binkley
National Center for Education Statistics

Dana Kelly
Education Statistics Services Institute

Kim Gattis
Education Statistics Services Institute

Appendix B: Example Passages

NAEP Example Passages



By George W. Frame

Nearly every day last summer my nephew Keith and I went crabbing in a creek on the New Jersey coast. We used a wire trap baited with scraps of fish and meat. Each time a crab entered the trap to eat, we pulled the doors closed. We cooked and ate the crabs we caught.

Blue crabs are very strong. Their big claws can make a painful pinch. When cornered, the crabs boldly defend themselves. They wave their outstretched claws and are fast and ready to fight. Keith and I had to be very careful to avoid having our fingers pinched.

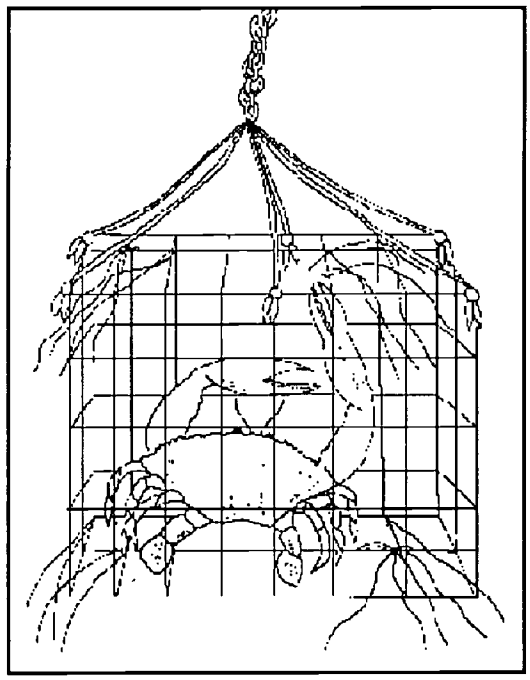
Crabs are **arthropods**, a very large group of animals that have an external skeleton and jointed legs. Other kinds of arthropods are insects, spiders, and centipedes. Blue crabs belong to a particular arthropod group called **crustaceans**. Crustaceans are abundant in the ocean, just as insects are on land.

The blue crab's hard shell is a strong armor. But the armor must be cast off from time to time so the crab can grow bigger. Getting rid of its shell is called **molting**.

Each blue crab molts about twenty times during its life. Just before molting, a new soft shell forms under the hard outer shell. Then the outer shell splits apart, and the crab backs out. This leaves the crab with a soft, wrinkled, outer covering. The body increases in size by absorbing water, stretching the soft shell to a much larger size. The crab hides for a few hours until its new shell has hardened.

Keith and I sometimes found these soft-shell crabs clinging to pilings and hiding beneath seaweed.

Blue crabs mate when the female undergoes her last molt and still has a soft shell. The male courts her by dancing from side to side while holding his claws



outstretched. He then transfers sperm to the female, where they are stored until egg laying begins several months later. The female blue crab mates only once but receives enough sperm to fertilize all the eggs that she will lay in her lifetime. Usually she lays eggs two or three times during the summer, and then she dies.

When the eggs are fertilized and laid, they become glued to long hairs on the underside of the female's abdomen. The egg mass sometimes looks like an orange-brown sponge and contains up to two million eggs until they hatch --- about nine to fourteen days later. Only one of the blue crabs that we caught last summer was carrying eggs, and we returned her to the water so her eggs could hatch. Most females with eggs stay in the deeper, saltier water at the ocean's edge rather than in the marshes.

The young blue crabs, and most other young crustaceans, hatch into larvae that look very different from their parents. The tiny blue crab babies are hardly bigger than a speck of dust. They are transparent and look like they are all head and tail. These larvae swim near the surface of the sea, and grow a new and bigger shell every few days. They soon change in shape so that they can either swim or crawl around on the bottom. Then they molt again and look like tiny adult crabs. After that their appearance does not change, but they continue to molt every twenty or thirty days as they grow.

As blue crabs become older, some move into shallower waters. The males in particular go into creeks and marshes, sometimes all the way to the freshwater streams and rivers. Keith and I caught ninety-two blue crabs in the shallow creek of the tide marsh last summer. Eighty-seven of those crabs were males, and only five were females

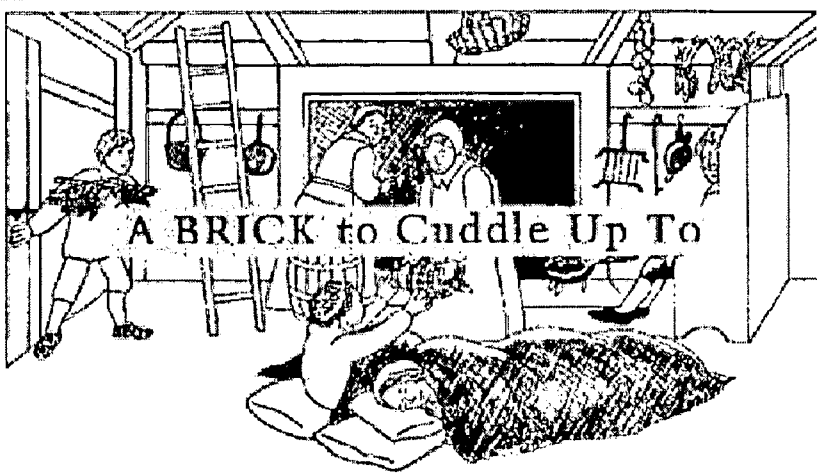
Gulls find and eat many blue crabs. They easily catch crabs that hide in puddles at low tide. Other predators are raccoons, alligators, and people. If caught, the crabs sometimes drop off a leg or claw to escape. Seven of the blue crabs that Keith and I caught were missing a claw.

Crabs are able to replace their lost limbs. If a leg or claw is seriously injured, the crab drops it off. The opening that is left near the body closes to prevent the loss of blood. Soon a new limb begins growing at the break. The next time the crab molts, the tiny limb's covering is cast off, too, and the crab then has a new usable leg or claw. The new limb is smaller that the lost one. But by the time the crab molts two or three more times, the new leg or claw will be normal size.

Many fishermen catch crabs to sell. Most are caught in wire traps or with baited lines during the summer while the crabs are active. In the winter, the fishermen drag big nets through the mud for the dormant crabs. Commercial fishermen catch a lot of crabs, sometimes more than 50 million pounds in a year. And many other crabs are caught by weekend fishermen who crab for fun and food.

The blue crab has a scientific name, just like all other living things. Its name is *Callinectes sapidus*. In the Latin language *Callinectes* means "beautiful swimmer," and *sapidus* means "delicious." I think that scientists gave the blue crab a very appropriate name.

WO000016



Imagine shivering on a cold winter's night. The tip of your nose tingles in the frosty air. Finally, you climb into bed and find the toasty treat you have been waiting for--your very own hot brick.

If you had lived in colonial days, that would not sound as strange as it does today. Winters were hard in this New World, and the colonists had to think of clever ways to fight the cold. At bedtime, they heated soapstones, or bricks, in the fireplace. They wrapped the bricks in cloths and tucked them into their beds. The brick kept them warm at night, at least for as long as its heat lasted.

at night, at least for as long as its heat lasted.

Before the colonists slipped into bed, they rubbed their icy sheets with a bed warmer. This was a metal pan with a long wooden handle. The pan held hot embers from the fireplace. It warmed the bedding so well that sleepy bodies had to wait until the sheets cooled before climbing in.

Staying warm wasn't just a bedtime problem. On winter rides, colonial travelers covered themselves with animal skins and warm blankets. Tucked under the blankets, near their feet, were small tin boxes called foot stoves. A foot stove held burning coals. Hot smoke puffed from small holes in the stove's lid, soothing freezing feet and legs. When the colonists went to Sunday services, their foot stoves, furs, and blankets went with them. The meeting houses had no heat of their own until the 1800s.

At home, colonial families huddled close to the fireplace, or hearth. The fireplace was wide and high enough to hold a large fire, but its chimney was large, too. That caused a problem: Gusts of cold air blew into the house. The area near the fire was warm, but in the rest of the room it might still be cold enough to see your breath.

Reading or needlework was done by candlelight, or by the light of the fire. During the winter, animal skins sealed the drafty windows of some cabins and blocked out the daylight. The living area inside was gloomy, except in the circle of light at the hearth.

Early Americans did not bathe as often as we do. When they did, their "bathroom" was the kitchen, in that toasty space by the hearth. They partially filled a tub with cold water, then warmed it up with water heated in the fireplace. A blanket draped from chairs for privacy also let the fire's warmth surround the bather.

The household cooks spent hours at the hearth. They stirred the kettle of corn pudding or checked the baking bread while the rest of the family carried on their own fireside activities. So you can see why the fireplace was the center of a colonial home.

The only time the fire was allowed to die down was at bedtime. Ashes would be piled over the fire, reducing it to embers that might glow until morning.

By sunrise, the hot brick had become a cold stone once more. An early riser might get dressed under the covers, then hurry to the hearth to warm up.

Maybe you'd enjoy hearing someone who kept warm in these ways tell you what it was like. You wouldn't need to look for someone who has been living for two hundred years. In many parts of the country the modern ways didn't take over from the old ones until recently. Your own grandparents or other older people might remember the warmth of a hearthside and the joy of having a brick to cuddle up to.

PIRLS Example Passage

The Upside-Down Mice

by Roald Dahl

Once upon a time there lived an old man of 87 whose name was Labon. All his life he had been a quiet and peaceful person. He was very poor and very happy.

When Labon discovered that he had mice in his house, it did not bother him much at first. But the mice multiplied. They began to bother him. They kept on multiplying and finally there came a time when even he could stand it no longer.

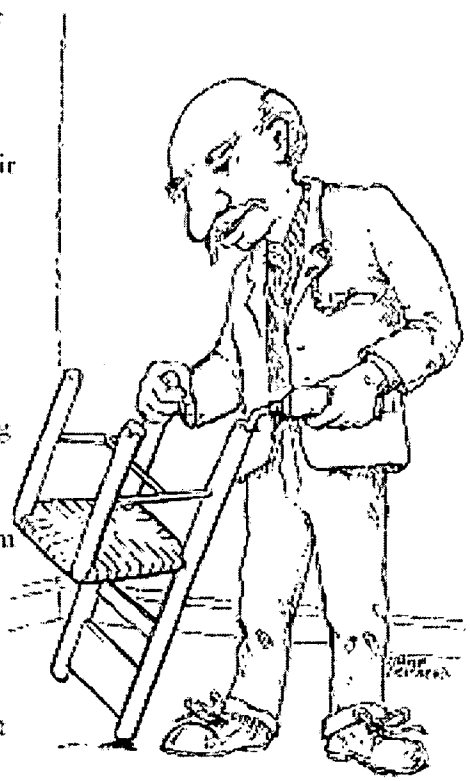
"This is too much," he said. "This really is going a bit too far." He hobbled out of the house down the road to a shop where he bought some mousetraps, a piece of cheese and some glue.

When he got home, he put the glue on the underneath of the mousetraps and stuck them to the ceiling. Then he baited them carefully with pieces of cheese and set them to go off.

That night when the mice came out of their holes and saw the mousetraps on the ceiling, they thought it was a tremendous joke. They walked around on the floor, nudging each other and pointing up with their front paws and roaring with laughter. After all, it was pretty silly, mousetraps on the ceiling.

When Labon came down the next morning and saw that there were no mice caught in the traps, he smiled but said nothing.

He took a chair and put glue on the bottom of its legs and stuck it upside-down to the ceiling, near the mousetraps. He did the same with the table, the television set and the lamp. He took everything that was on the floor and stuck it upside-down on the ceiling. He even put a little carpet up there.



The next night when the mice came out of their holes they were still joking and laughing about what they had seen the night before. But now, when they looked up at the ceiling, they stopped laughing very suddenly.

"Good gracious me!" cried one. "Look up there! There's the floor!"

"Heavens above!" shouted another. "We must be standing on the ceiling!"

"I'm beginning to feel a little giddy," said another.

"All the blood's going to my head," said another.

"This is terrible!" said a very senior mouse with long whiskers. "This is really terrible! We must do something about it at once!"

"I shall faint if I have to stand on my head any longer!" shouted a young mouse.

"Me too!"

"I can't stand it!"

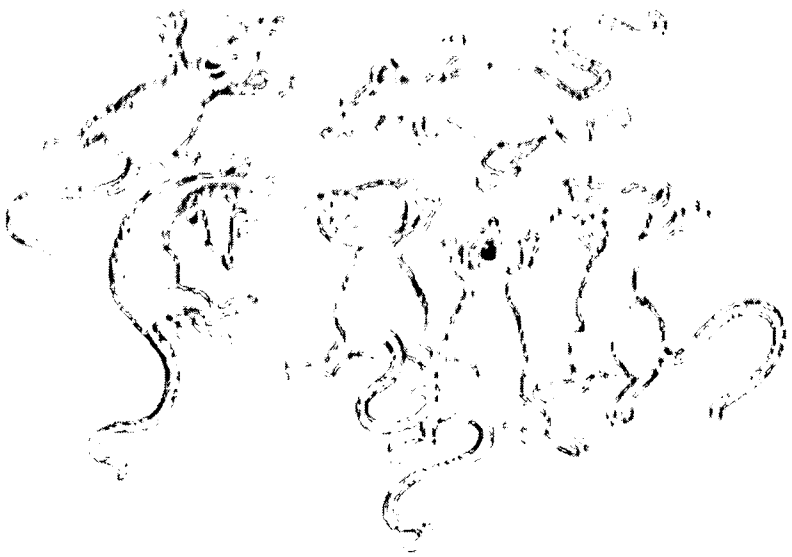
"Save us! Do something, somebody, quick!"

They were getting hysterical now. "I know what we'll do," said the very senior mouse. "We'll all stand on our heads, then we'll be the right way up."

Obediently, they all stood on their heads, and after a long time, one by one they fainted from a rush of blood to their brains.

When Labon came down the next morning the floor was littered with mice. Quickly he gathered them up and popped them all in a basket.

So the thing to remember is this: whenever the world seems to be terribly upside-down, make sure you keep your feet firmly on the ground.



Appendix C: Readability and Lexile Analyses

For each passage in NAEP fourth-grade and PIRLS, three 100-word samples were randomly chosen and for each 100-word sample, the number of sentences and syllables were counted. For one PIRLS passage--a brochure--with fewer than 300 words only two 100-word samples were chosen. Sentence and syllable tabulation was accomplished with the help of a program found at: http://www.educational-psychologist.co.uk/fry_readability_program.htm.

The above-mentioned program also automatically calculated the "Fry Readability Age." The "Average Age" for each passage was calculated by averaging the "Fry Readability Age" from the three samples. The "Average Age" for each assessment was calculated by averaging the "Average Age" from the eight passages of each respective assessment. Using the average number of sentences and syllables for each passage, the "Average Grade Level" for each passage was found using the Fry Graph. To find the "Average Grade Level" for each assessment, the Average Grade Level" from the eight passages of each respective assessment were averaged.

The same average sentence and syllable counts used in the Fry Analyses were used to compute the "Average Reading Ease" for each passage (Appendix, Tables A3 and A4). The formula used: ([(Words/Sentences)*1.015]+[(Syllables/Words)*84.6]). The resulting score is on a scale of 1-100, with higher numbers indicating easier reading levels. To find the "Average Reading Ease" for each assessment, the "Average Reading Ease" scores from the eight passages of each respective assessment were averaged. The qualitative labels (e.g. "Easy," "Fairly Easy") attached to these Reading Ease scores were assigned based on 10-point intervals, reprinted below. Also assigned based on the table's translation of Reading Ease scores is the "Average Grade Level".

Reading Ease Score	*Difficulty*	*Flesch Grade Level*
0-29	Very Difficult	Post Graduate
30-49	Difficult	College
50-59	Fairly Difficult	High School
60-69	Standard	8th to 9th Grade
70-79	Fairly Easy	7th Grade
80-89	Easy	5th to 6th Grade
90-100	Very Easy	4th to 5th Grade

The Lexile analysis was conducted by using the software provided by MetaMetrics and accessed on-line. Each reading passage was formatted according to MetaMetrics guidelines and run through the software program to obtain the Lexile score.

Appendix D: Classification of Items

Did the Expert Panel Agree with PIRLS Developers on the Processes Assessed by Each PIRLS Item?

Table D.1: Comparison between Classification by the PIRLS Framework Categories by the PIRLS Developers and by the Expert Panel

	Focus on and Retrieve Explicitly Stated Information	Make Straightforward Inferences	Interpret and Integrate Ideas and Information	Examine and Evaluate Content, Language, and Textual Elements
Classified by PIRLS Developers	26%	28%	32%	15%
Classified by Panel	24%	32%	28%	16%

Did the Expert Panel Agree with NAEP Developers on the Processes Assessed by Each NAEP Item?

Table D.2: Comparison between Classification by the NAEP Framework Categories by NAEP Developers and by the Expert Panel

	Forming an Initial Understanding	Developing an Interpretation	Personal Reflection and Response	Demonstrating a Critical Stance
Classified by NAEP Developers	10%	55%	15%	21%
Classified by Panel	10%	72%	11%	7%

Listing of NCES Working Papers to Date

Working papers can be downloaded as .pdf files from the NCES Electronic Catalog (http://nces.ed.gov/pubsearch/). You can also contact Sheilah Jupiter at (202) 502–7444 (sheilah.jupiter@ed.gov) if you are interested in any of the following papers.

Listing of NCES Working Papers by Program Area

No.	Title	NCES contact
Baccalaureate and Beyond (B&B)		
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
2001–15	Baccalaureate and Beyond Longitudinal Study: 2000/01 Follow-Up Field Test Methodology Report	Andrew G. Malizio
2002–04	Improving Consistency of Response Categories Across NCES Surveys	Marilyn Seastrom
Beginning Postsecondary Students (BPS) Longitudinal Study		
98–11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96–98) Field Test Report	Aurora D'Amico
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
1999–15	Projected Postsecondary Outcomes of 1992 High School Graduates	Aurora D'Amico
2001–04	Beginning Postsecondary Students Longitudinal Study: 1996–2001 (BPS:1996/2001) Field Test Methodology Report	Paula Knepper
2002–04	Improving Consistency of Response Categories Across NCES Surveys	Marilyn Seastrom
Common Core of Data (CCD)		
95–12	Rural Education Data User's Guide	Samuel Peng
96–19	Assessment and Analysis of School-Level Expenditures	William J. Fowler, Jr.
97–15	Customer Service Survey: Common Core of Data Coordinators	Lee Hoffman
97–43	Measuring Inflation in Public School Costs	William J. Fowler, Jr.
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
1999–03	Evaluation of the 1996–97 Nonfiscal Common Core of Data Surveys Data Collection, Processing, and Editing Cycle	Beth Young
2000–12	Coverage Evaluation of the 1994–95 Common Core of Data: Public Elementary/Secondary School Universe Survey	Beth Young
2000–13	Non-professional Staff in the Schools and Staffing Survey (SASS) and Common Core of Data (CCD)	Kerry Gruber
2002–02	School Locale Codes 1987 - 2000	Frank Johnson
Data Development		
2000–16a	Lifelong Learning NCES Task Force: Final Report Volume I	Lisa Hudson
2000–16b	Lifelong Learning NCES Task Force: Final Report Volume II	Lisa Hudson
Decennial Census School District Project		
95–12	Rural Education Data User's Guide	Samuel Peng
96–04	Census Mapping Project/School District Data Book	Tai Phan
98–07	Decennial Census School District Project Planning Report	Tai Phan
Early Childhood Longitudinal Study (ECLS)		
96–08	How Accurate are Teacher Judgments of Students' Academic Performance?	Jerry West
96–18	Assessment of Social Competence, Adaptive Behaviors, and Approaches to Learning with Young Children	Jerry West
97–24	Formulating a Design for the ECLS: A Review of Longitudinal Studies	Jerry West
97–36	Measuring the Quality of Program Environments in Head Start and Other Early Childhood Programs: A Review and Recommendations for Future Research	Jerry West
1999–01	A Birth Cohort Study: Conceptual and Design Considerations and Rationale	Jerry West
2000–04	Selected Papers on Education Surveys: Papers Presented at the 1998 and 1999 ASA and 1999 AAPOR Meetings	Dan Kasprzyk
2001–02	Measuring Father Involvement in Young Children's Lives: Recommendations for a Fatherhood Module for the ECLS-B	Jerry West
2001–03	Measures of Socio-Emotional Development in Middle Childhood	Elvira Hausken

No.	Title	NCES contact
2001–06	Papers from the Early Childhood Longitudinal Studies Program: Presented at the 2001 AERA and SRCD Meetings	Jerry West
2002-05	Early Childhood Longitudinal Study-Kindergarten Class of 1998–99 (ECLS–K), Psychometric Report for Kindergarten Through First Grade	Elvira Hausken
Education Finance Statistics Center (EDFIN)		
94–05	Cost-of-Education Differentials Across the States	William J. Fowler, Jr.
96–19	Assessment and Analysis of School-Level Expenditures	William J. Fowler, Jr.
97–43	Measuring Inflation in Public School Costs	William J. Fowler, Jr.
98–04	Geographic Variations in Public Schools' Costs	William J. Fowler, Jr.
1999–16	Measuring Resources in Education: From Accounting to the Resource Cost Model Approach	William J. Fowler, Jr.
Education Longitudinal Study: 2002 (ELS:2002)		
2003-03	Education Longitudinal Study: 2002 (ELS: 2002) Field Test Report	Jeffrey Owings
High School and Beyond (HS&B)		
95–12	Rural Education Data User's Guide	Samuel Peng
1999–05	Procedures Guide for Transcript Studies	Dawn Nelson
1999–06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson
2002–04	Improving Consistency of Response Categories Across NCES Surveys	Marilyn Seastrom
HS Transcript Studies		
1999–05	Procedures Guide for Transcript Studies	Dawn Nelson
1999–06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson
2003–01	Mathematics, Foreign Language, and Science Coursetaking and the NELS:88 Transcript Data	Jeffrey Owings
2003–02	English Coursetaking and the NELS:88 Transcript Data	Jeffrey Owings
International Adult Literacy Survey (IALS)		
97–33	Adult Literacy: An International Perspective	Marilyn Binkley
Integrated Postsecondary Education Data System (IPEDS)		
97–27	Pilot Test of IPEDS Finance Survey	Peter Stowe
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
2000–14	IPEDS Finance Data Comparisons Under the 1997 Financial Accounting Standards for Private, Not-for-Profit Institutes: A Concept Paper	Peter Stowe
National Assessment of Adult Literacy (NAAL)		
98–17	Developing the National Assessment of Adult Literacy: Recommendations from Stakeholders	Sheida White
1999–09a	1992 National Adult Literacy Survey: An Overview	Alex Sedlacek
1999–09b	1992 National Adult Literacy Survey: Sample Design	Alex Sedlacek
1999–09c	1992 National Adult Literacy Survey: Weighting and Population Estimates	Alex Sedlacek
1999–09d	1992 National Adult Literacy Survey: Development of the Survey Instruments	Alex Sedlacek
1999–09e	1992 National Adult Literacy Survey: Scaling and Proficiency Estimates	Alex Sedlacek
1999–09f	1992 National Adult Literacy Survey: Interpreting the Adult Literacy Scales and Literacy Levels	Alex Sedlacek
1999–09g	1992 National Adult Literacy Survey: Literacy Levels and the Response Probability Convention	Alex Sedlacek
2000–05	Secondary Statistical Modeling With the National Assessment of Adult Literacy: Implications for the Design of the Background Questionnaire	Sheida White
2000–06	Using Telephone and Mail Surveys as a Supplement or Alternative to Door-to-Door Surveys in the Assessment of Adult Literacy	Sheida White
2000–07	"How Much Literacy is Enough?" Issues in Defining and Reporting Performance Standards for the National Assessment of Adult Literacy	Sheida White
2000–08	Evaluation of the 1992 NALS Background Survey Questionnaire: An Analysis of Uses with Recommendations for Revisions	Sheida White
2000–09	Demographic Changes and Literacy Development in a Decade	Sheida White
2001–08	Assessing the Lexile Framework: Results of a Panel Meeting	Sheida White

No.	Title	NCES contact
2002–04	Improving Consistency of Response Categories Across NCES Surveys	Marilyn Seastrom
National Assessment of Educational Progress (NAEP)		
95–12	Rural Education Data User's Guide	Samuel Peng
97–29	Can State Assessment Data be Used to Reduce State NAEP Sample Sizes?	Steven Gorman
97–30	ACT's NAEP Redesign Project: Assessment Design is the Key to Useful and Stable Assessment Results	Steven Gorman
97–31	NAEP Reconfigured: An Integrated Redesign of the National Assessment of Educational Progress	Steven Gorman
97–32	Innovative Solutions to Intractable Large Scale Assessment (Problem 2: Background Questionnaires)	Steven Gorman
97–37	Optimal Rating Procedures and Methodology for NAEP Open-ended Items	Steven Gorman
97–44	Development of a SASS 1993–94 School-Level Student Achievement Subfile: Using State Assessments and State NAEP, Feasibility Study	Michael Ross
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
1999–05	Procedures Guide for Transcript Studies	Dawn Nelson
1999–06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson
2001–07	A Comparison of the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study Repeat (TIMSS-R), and the Programme for International Student Assessment (PISA)	Arnold Goldstein
2001–08	Assessing the Lexile Framework: Results of a Panel Meeting	Sheida White
2001–11	Impact of Selected Background Variables on Students' NAEP Math Performance	Arnold Goldstein
2001–13	The Effects of Accommodations on the Assessment of LEP Students in NAEP	Arnold Goldstein
2001–19	The Measurement of Home Background Indicators: Cognitive Laboratory Investigations of the Responses of Fourth and Eighth Graders to Questionnaire Items and Parental Assessment of the Invasiveness of These Items	Arnold Goldstein
2002–04	Improving Consistency of Response Categories Across NCES Surveys	Marilyn Seastrom
2002-06	The Measurement of Instructional Background Indicators: Cognitive Laboratory Investigations of the Responses of Fourth and Eighth Grade Students and Teachers to Questionnaire Items	Arnold Goldstein
2002–07	Teacher Quality, School Context, and Student Race/Ethnicity: Findings from the Eighth Grade National Assessment of Educational Progress 2000 Mathematics Assessment	Janis Brown
2003–10	A Content Comparison of NAEP and PIRLS Fourth-Grade Reading Assessments	Marilyn Binkley
National Education Longitudinal Study of 1988 (NELS:88)		
95–04	National Education Longitudinal Study of 1988: Second Follow-up Questionnaire Content Areas and Research Issues	Jeffrey Owings
95–05	National Education Longitudinal Study of 1988: Conducting Trend Analyses of NLS-72, HS&B, and NELS:88 Seniors	Jeffrey Owings
95–06	National Education Longitudinal Study of 1988: Conducting Cross-Cohort Comparisons Using HS&B, NAEP, and NELS:88 Academic Transcript Data	Jeffrey Owings
95–07	National Education Longitudinal Study of 1988: Conducting Trend Analyses HS&B and NELS:88 Sophomore Cohort Dropouts	Jeffrey Owings
95–12	Rural Education Data User's Guide	Samuel Peng
95–14	Empirical Evaluation of Social, Psychological, & Educational Construct Variables Used in NCES Surveys	Samuel Peng
96–03	National Education Longitudinal Study of 1988 (NELS:88) Research Framework and Issues	Jeffrey Owings
98–06	National Education Longitudinal Study of 1988 (NELS:88) Base Year through Second Follow-Up: Final Methodology Report	Ralph Lee
98–09	High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988	Jeffrey Owings
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
1999–05	Procedures Guide for Transcript Studies	Dawn Nelson
1999–06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson
1999–15	Projected Postsecondary Outcomes of 1992 High School Graduates	Aurora D'Amico
2001–16	Imputation of Test Scores in the National Education Longitudinal Study of 1988	Ralph Lee

No.	Title	NCES contact
2002–04	Improving Consistency of Response Categories Across NCES Surveys	Marilyn Seastrom
2003–01	Mathematics, Foreign Language, and Science Coursetaking and the NELS:88 Transcript Data	Jeffrey Owings
2003–02	English Coursetaking and the NELS:88 Transcript Data	Jeffrey Owings
National Household Education Survey (NHES)		
95–12	Rural Education Data User's Guide	Samuel Peng
96–13	Estimation of Response Bias in the NHES:95 Adult Education Survey	Steven Kaufman
96–14	The 1995 National Household Education Survey: Reinterview Results for the Adult Education Component	Steven Kaufman
96–20	1991 National Household Education Survey (NHES:91) Questionnaires: Screener, Early Childhood Education, and Adult Education	Kathryn Chandler
96–21	1993 National Household Education Survey (NHES:93) Questionnaires: Screener, School Readiness, and School Safety and Discipline	Kathryn Chandler
96–22	1995 National Household Education Survey (NHES:95) Questionnaires: Screener, Early Childhood Program Participation, and Adult Education	Kathryn Chandler
96–29	Undercoverage Bias in Estimates of Characteristics of Adults and 0- to 2-Year-Olds in the 1995 National Household Education Survey (NHES:95)	Kathryn Chandler
96–30	Comparison of Estimates from the 1995 National Household Education Survey (NHES:95)	Kathryn Chandler
97–02	Telephone Coverage Bias and Recorded Interviews in the 1993 National Household Education Survey (NHES:93)	Kathryn Chandler
97–03	1991 and 1995 National Household Education Survey Questionnaires: NHES:91 Screener, NHES:91 Adult Education, NHES:95 Basic Screener, and NHES:95 Adult Education	Kathryn Chandler
97–04	Design, Data Collection, Monitoring, Interview Administration Time, and Data Editing in the 1993 National Household Education Survey (NHES:93)	Kathryn Chandler
97–05	Unit and Item Response, Weighting, and Imputation Procedures in the 1993 National Household Education Survey (NHES:93)	Kathryn Chandler
97–06	Unit and Item Response, Weighting, and Imputation Procedures in the 1995 National Household Education Survey (NHES:95)	Kathryn Chandler
97–08	Design, Data Collection, Interview Timing, and Data Editing in the 1995 National Household Education Survey	Kathryn Chandler
97–19	National Household Education Survey of 1995: Adult Education Course Coding Manual	Peter Stowe
97–20	National Household Education Survey of 1995: Adult Education Course Code Merge Files User's Guide	Peter Stowe
97–25	1996 National Household Education Survey (NHES:96) Questionnaires: Screener/Household and Library, Parent and Family Involvement in Education and Civic Involvement, Youth Civic Involvement, and Adult Civic Involvement	Kathryn Chandler
97–28	Comparison of Estimates in the 1996 National Household Education Survey	Kathryn Chandler
97–34	Comparison of Estimates from the 1993 National Household Education Survey	Kathryn Chandler
97–35	Design, Data Collection, Interview Administration Time, and Data Editing in the 1996 National Household Education Survey	Kathryn Chandler
97–38	Reinterview Results for the Parent and Youth Components of the 1996 National Household Education Survey	Kathryn Chandler
97–39	Undercoverage Bias in Estimates of Characteristics of Households and Adults in the 1996 National Household Education Survey	Kathryn Chandler
97–40	Unit and Item Response Rates, Weighting, and Imputation Procedures in the 1996 National Household Education Survey	Kathryn Chandler
98–03	Adult Education in the 1990s: A Report on the 1991 National Household Education Survey	Peter Stowe
98–10	Adult Education Participation Decisions and Barriers: Review of Conceptual Frameworks and Empirical Studies	Peter Stowe
2002–04	Improving Consistency of Response Categories Across NCES Surveys	Marilyn Seastrom
National Longitudinal Study of the High School Class of 1972 (NLS-72)		
95–12	Rural Education Data User's Guide	Samuel Peng
2002–04	Improving Consistency of Response Categories Across NCES Surveys	Marilyn Seastrom
National Postsecondary Student Aid Study (NPSAS)		
96–17	National Postsecondary Student Aid Study: 1996 Field Test Methodology Report	Andrew G. Malizio
2000–17	National Postsecondary Student Aid Study:2000 Field Test Methodology Report	Andrew G. Malizio

No.	Title	NCES contact
2002–03	National Postsecondary Student Aid Study, 1999–2000 (NPSAS:2000), CATI Nonresponse Bias Analysis Report.	Andrew Malizio
2002–04	Improving Consistency of Response Categories Across NCES Surveys	Marilyn Seastrom
National Study of Postsecondary Faculty (NSOPF)		
97–26	Strategies for Improving Accuracy of Postsecondary Faculty Lists	Linda Zimbler
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
2000–01	1999 National Study of Postsecondary Faculty (NSOPF:99) Field Test Report	Linda Zimbler
2002–04	Improving Consistency of Response Categories Across NCES Surveys	Marilyn Seastrom
2002–08	A Profile of Part-time Faculty: Fall 1998	Linda Zimbler
Postsecondary Education Descriptive Analysis Reports (PEDAR)		
2000–11	Financial Aid Profile of Graduate Students in Science and Engineering	Aurora D'Amico
Private School Universe Survey (PSS)		
95–16	Intersurvey Consistency in NCES Private School Surveys	Steven Kaufman
95–17	Estimates of Expenditures for Private K–12 Schools	Stephen Broughman
96–16	Strategies for Collecting Finance Data from Private Schools	Stephen Broughman
96–26	Improving the Coverage of Private Elementary-Secondary Schools	Steven Kaufman
96–27	Intersurvey Consistency in NCES Private School Surveys for 1993–94	Steven Kaufman
97–07	The Determinants of Per-Pupil Expenditures in Private Elementary and Secondary Schools: An Exploratory Analysis	Stephen Broughman
97–22	Collection of Private School Finance Data: Development of a Questionnaire	Stephen Broughman
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
2000–04	Selected Papers on Education Surveys: Papers Presented at the 1998 and 1999 ASA and 1999 AAPOR Meetings	Dan Kasprzyk
2000–15	Feasibility Report: School-Level Finance Pretest, Private School Questionnaire	Stephen Broughman
Recent College Graduates (RCG)		
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
2002–04	Improving Consistency of Response Categories Across NCES Surveys	Marilyn Seastrom
Schools and Staffing Survey (SASS)		
94–01	Schools and Staffing Survey (SASS) Papers Presented at Meetings of the American Statistical Association	Dan Kasprzyk
94–02	Generalized Variance Estimate for Schools and Staffing Survey (SASS)	Dan Kasprzyk
94–03	1991 Schools and Staffing Survey (SASS) Reinterview Response Variance Report	Dan Kasprzyk
94–04	The Accuracy of Teachers' Self-reports on their Postsecondary Education: Teacher Transcript Study, Schools and Staffing Survey	Dan Kasprzyk
94–06	Six Papers on Teachers from the 1990–91 Schools and Staffing Survey and Other Related Surveys	Dan Kasprzyk
95–01	Schools and Staffing Survey: 1994 Papers Presented at the 1994 Meeting of the American Statistical Association	Dan Kasprzyk
95–02	QED Estimates of the 1990–91 Schools and Staffing Survey: Deriving and Comparing QED School Estimates with CCD Estimates	Dan Kasprzyk
95–03	Schools and Staffing Survey: 1990–91 SASS Cross-Questionnaire Analysis	Dan Kasprzyk
95–08	CCD Adjustment to the 1990–91 SASS: A Comparison of Estimates	Dan Kasprzyk
95–09	The Results of the 1993 Teacher List Validation Study (TLVS)	Dan Kasprzyk
95–10	The Results of the 1991–92 Teacher Follow-up Survey (TFS) Reinterview and Extensive Reconciliation	Dan Kasprzyk
95–11	Measuring Instruction, Curriculum Content, and Instructional Resources: The Status of Recent Work	Sharon Bobbitt & John Ralph
95–12	Rural Education Data User's Guide	Samuel Peng
95–14	Empirical Evaluation of Social, Psychological, & Educational Construct Variables Used in NCES Surveys	Samuel Peng
95–15	Classroom Instructional Processes: A Review of Existing Measurement Approaches and Their Applicability for the Teacher Follow-up Survey	Sharon Bobbitt
95–16	Intersurvey Consistency in NCES Private School Surveys	Steven Kaufman
95–18	An Agenda for Research on Teachers and Schools: Revisiting NCES' Schools and Staffing Survey	Dan Kasprzyk

No.	Title	NCES contact
96–01	Methodological Issues in the Study of Teachers' Careers: Critical Features of a Truly Longitudinal Study	Dan Kasprzyk
96–02	Schools and Staffing Survey (SASS): 1995 Selected papers presented at the 1995 Meeting of the American Statistical Association	Dan Kasprzyk
96–05	Cognitive Research on the Teacher Listing Form for the Schools and Staffing Survey	Dan Kasprzyk
96–06	The Schools and Staffing Survey (SASS) for 1998–99: Design Recommendations to Inform Broad Education Policy	Dan Kasprzyk
96–07	Should SASS Measure Instructional Processes and Teacher Effectiveness?	Dan Kasprzyk
96–09	Making Data Relevant for Policy Discussions: Redesigning the School Administrator Questionnaire for the 1998–99 SASS	Dan Kasprzyk
96–10	1998–99 Schools and Staffing Survey: Issues Related to Survey Depth	Dan Kasprzyk
96–11	Towards an Organizational Database on America's Schools: A Proposal for the Future of SASS, with comments on School Reform, Governance, and Finance	Dan Kasprzyk
96–12	Predictors of Retention, Transfer, and Attrition of Special and General Education Teachers: Data from the 1989 Teacher Followup Survey	Dan Kasprzyk
96–15	Nested Structures: District-Level Data in the Schools and Staffing Survey	Dan Kasprzyk
96–23	Linking Student Data to SASS: Why, When, How	Dan Kasprzyk
96–24	National Assessments of Teacher Quality	Dan Kasprzyk
96–25	Measures of Inservice Professional Development: Suggested Items for the 1998–1999 Schools and Staffing Survey	Dan Kasprzyk
96–28	Student Learning, Teaching Quality, and Professional Development: Theoretical Linkages, Current Measurement, and Recommendations for Future Data Collection	Mary Rollefson
97–01	Selected Papers on Education Surveys: Papers Presented at the 1996 Meeting of the American Statistical Association	Dan Kasprzyk
97–07	The Determinants of Per-Pupil Expenditures in Private Elementary and Secondary Schools: An Exploratory Analysis	Stephen Broughman
97–09	Status of Data on Crime and Violence in Schools: Final Report	Lee Hoffman
97–10	Report of Cognitive Research on the Public and Private School Teacher Questionnaires for the Schools and Staffing Survey 1993–94 School Year	Dan Kasprzyk
97–11	International Comparisons of Inservice Professional Development	Dan Kasprzyk
97–12	Measuring School Reform: Recommendations for Future SASS Data Collection	Mary Rollefson
97–14	Optimal Choice of Periodicities for the Schools and Staffing Survey: Modeling and Analysis	Steven Kaufman
97–18	Improving the Mail Return Rates of SASS Surveys: A Review of the Literature	Steven Kaufman
97–22	Collection of Private School Finance Data: Development of a Questionnaire	Stephen Broughman
97–23	Further Cognitive Research on the Schools and Staffing Survey (SASS) Teacher Listing Form	Dan Kasprzyk
97–41	Selected Papers on the Schools and Staffing Survey: Papers Presented at the 1997 Meeting of the American Statistical Association	Steve Kaufman
97–42	Improving the Measurement of Staffing Resources at the School Level: The Development of Recommendations for NCES for the Schools and Staffing Survey (SASS)	Mary Rollefson
97–44	Development of a SASS 1993–94 School-Level Student Achievement Subfile: Using State Assessments and State NAEP, Feasibility Study	Michael Ross
98–01	Collection of Public School Expenditure Data: Development of a Questionnaire	Stephen Broughman
98–02	Response Variance in the 1993–94 Schools and Staffing Survey: A Reinterview Report	Steven Kaufman
98–04	Geographic Variations in Public Schools' Costs	William J. Fowler, Jr.
98–05	SASS Documentation: 1993–94 SASS Student Sampling Problems; Solutions for Determining the Numerators for the SASS Private School (3B) Second-Stage Factors	Steven Kaufman
98–08	The Redesign of the Schools and Staffing Survey for 1999–2000: A Position Paper	Dan Kasprzyk
98–12	A Bootstrap Variance Estimator for Systematic PPS Sampling	Steven Kaufman
98–13	Response Variance in the 1994–95 Teacher Follow-up Survey	Steven Kaufman
98–14	Variance Estimation of Imputed Survey Data	Steven Kaufman
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
98–16	A Feasibility Study of Longitudinal Design for Schools and Staffing Survey	Stephen Broughman
1999–02	Tracking Secondary Use of the Schools and Staffing Survey Data: Preliminary Results	Dan Kasprzyk
1999–04	Measuring Teacher Qualifications	Dan Kasprzyk
1999–07	Collection of Resource and Expenditure Data on the Schools and Staffing Survey	Stephen Broughman
1999–08	Measuring Classroom Instructional Processes: Using Survey and Case Study Fieldtest Results to Improve Item Construction	Dan Kasprzyk
1999–10	What Users Say About Schools and Staffing Survey Publications	Dan Kasprzyk

No.	Title	NCES contact
1999–12	1993–94 Schools and Staffing Survey: Data File User's Manual, Volume III: Public-Use Codebook	Kerry Gruber
1999–13	1993–94 Schools and Staffing Survey: Data File User's Manual, Volume IV: Bureau of Indian Affairs (BIA) Restricted-Use Codebook	Kerry Gruber
1999–14	1994–95 Teacher Followup Survey: Data File User's Manual, Restricted-Use Codebook	Kerry Gruber
1999–17	Secondary Use of the Schools and Staffing Survey Data	Susan Wiley
2000–04	Selected Papers on Education Surveys: Papers Presented at the 1998 and 1999 ASA and 1999 AAPOR Meetings	Dan Kasprzyk
2000–10	A Research Agenda for the 1999–2000 Schools and Staffing Survey	Dan Kasprzyk
2000–13	Non-professional Staff in the Schools and Staffing Survey (SASS) and Common Core of Data (CCD)	Kerry Gruber
2000–18	Feasibility Report: School-Level Finance Pretest, Public School District Questionnaire	Stephen Broughman
2002–04	Improving Consistency of Response Categories Across NCES Surveys	Marilyn Seastrom
Third International Mathematics and Science Study (TIMSS)		
2001–01	Cross-National Variation in Educational Preparation for Adulthood: From Early Adolescence to Young Adulthood	Elvira Hausken
2001–05	Using TIMSS to Analyze Correlates of Performance Variation in Mathematics	Patrick Gonzales
2001–07	A Comparison of the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study Repeat (TIMSS-R), and the Programme for International Student Assessment (PISA)	Arnold Goldstein
2002–01	Legal and Ethical Issues in the Use of Video in Education Research	Patrick Gonzales

Listing of NCES Working Papers by Subject

No.	Title	NCES contact
Achievement (student) - mathematics		
2001–05	Using TIMSS to Analyze Correlates of Performance Variation in Mathematics	Patrick Gonzales
Adult education		
96–14	The 1995 National Household Education Survey: Reinterview Results for the Adult Education Component	Steven Kaufman
96–20	1991 National Household Education Survey (NHES:91) Questionnaires: Screener, Early Childhood Education, and Adult Education	Kathryn Chandler
96–22	1995 National Household Education Survey (NHES:95) Questionnaires: Screener, Early Childhood Program Participation, and Adult Education	Kathryn Chandler
98–03	Adult Education in the 1990s: A Report on the 1991 National Household Education Survey	Peter Stowe
98–10	Adult Education Participation Decisions and Barriers: Review of Conceptual Frameworks and Empirical Studies	Peter Stowe
1999–11	Data Sources on Lifelong Learning Available from the National Center for Education Statistics	Lisa Hudson
2000–16a	Lifelong Learning NCES Task Force: Final Report Volume I	Lisa Hudson
2000–16b	Lifelong Learning NCES Task Force: Final Report Volume II	Lisa Hudson
Adult literacy—see Literacy of adults		
American Indian – education		
1999–13	1993–94 Schools and Staffing Survey: Data File User's Manual, Volume IV: Bureau of Indian Affairs (BIA) Restricted-Use Codebook	Kerry Gruber
Assessment/achievement		
95–12	Rural Education Data User's Guide	Samuel Peng
95–13	Assessing Students with Disabilities and Limited English Proficiency	James Houser
97–29	Can State Assessment Data be Used to Reduce State NAEP Sample Sizes?	Larry Ogle
97–30	ACT's NAEP Redesign Project: Assessment Design is the Key to Useful and Stable Assessment Results	Larry Ogle
97–31	NAEP Reconfigured: An Integrated Redesign of the National Assessment of Educational Progress	Larry Ogle
97–32	Innovative Solutions to Intractable Large Scale Assessment (Problem 2: Background Questions)	Larry Ogle
97–37	Optimal Rating Procedures and Methodology for NAEP Open-ended Items	Larry Ogle
97–44	Development of a SASS 1993–94 School-Level Student Achievement Subfile: Using State Assessments and State NAEP, Feasibility Study	Michael Ross
98–09	High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988	Jeffrey Owings
2001–07	A Comparison of the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study Repeat (TIMSS-R), and the Programme for International Student Assessment (PISA)	Arnold Goldstein
2001–11	Impact of Selected Background Variables on Students' NAEP Math Performance	Arnold Goldstein
2001–13	The Effects of Accommodations on the Assessment of LEP Students in NAEP	Arnold Goldstein
2001–19	The Measurement of Home Background Indicators: Cognitive Laboratory Investigations of the Responses of Fourth and Eighth Graders to Questionnaire Items and Parental Assessment of the Invasiveness of These Items	Arnold Goldstein
2002-05	Early Childhood Longitudinal Study-Kindergarten Class of 1998–99 (ECLS–K), Psychometric Report for Kindergarten Through First Grade	Elvira Hausken

No.	Title	NCES contact
2002-06	The Measurement of Instructional Background Indicators: Cognitive Laboratory Investigations of the Responses of Fourth and Eighth Grade Students and Teachers to Questionnaire Items	Arnold Goldstein
2002-07	Teacher Quality, School Context, and Student Race/Ethnicity: Findings from the Eighth Grade National Assessment of Educational Progress 2000 Mathematics Assessment	Janis Brown
2003-10	A Content Comparison of the NAEP and PIRLS Fourth-Grade Reading Assessments	Marilyn Binkley
Beginning students in postsecondary education		
98–11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96–98) Field Test Report	Aurora D'Amico
2001–04	Beginning Postsecondary Students Longitudinal Study: 1996–2001 (BPS:1996/2001) Field Test Methodology Report	Paula Knepper
Civic participation		
97–25	1996 National Household Education Survey (NHES:96) Questionnaires: Screener/Household and Library, Parent and Family Involvement in Education and Civic Involvement, Youth Civic Involvement, and Adult Civic Involvement	Kathryn Chandler
Climate of schools		
95–14	Empirical Evaluation of Social, Psychological, & Educational Construct Variables Used in NCES Surveys	Samuel Peng
Cost of education indices		
94–05	Cost-of-Education Differentials Across the States	William J. Fowler, Jr.
Course-taking		
95–12	Rural Education Data User's Guide	Samuel Peng
98–09	High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988	Jeffrey Owings
1999–05	Procedures Guide for Transcript Studies	Dawn Nelson
1999–06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson
2003–01	Mathematics, Foreign Language, and Science Coursetaking and the NELS:88 Transcript Data	Jeffrey Owings
2003–02	English Coursetaking and the NELS:88 Transcript Data	Jeffrey Owings
Crime		
97–09	Status of Data on Crime and Violence in Schools: Final Report	Lee Hoffman
Curriculum		
95–11	Measuring Instruction, Curriculum Content, and Instructional Resources: The Status of Recent Work	Sharon Bobbitt & John Ralph
98–09	High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988	Jeffrey Owings
Customer service		
1999–10	What Users Say About Schools and Staffing Survey Publications	Dan Kasprzyk
2000–02	Coordinating NCES Surveys: Options, Issues, Challenges, and Next Steps	Valena Plisko
2000–04	Selected Papers on Education Surveys: Papers Presented at the 1998 and 1999 ASA and 1999 AAPOR Meetings	Dan Kasprzyk
Data quality		
97–13	Improving Data Quality in NCES: Database-to-Report Process	Susan Ahmed
2001–11	Impact of Selected Background Variables on Students' NAEP Math Performance	Arnold Goldstein
2001–13	The Effects of Accommodations on the Assessment of LEP Students in NAEP	Arnold Goldstein
2001–19	The Measurement of Home Background Indicators: Cognitive Laboratory Investigations of the Responses of Fourth and Eighth Graders to Questionnaire Items and Parental Assessment of the Invasiveness of These Items	Arnold Goldstein

No.	Title	NCES contact
2002-06	The Measurement of Instructional Background Indicators: Cognitive Laboratory Investigations of the Responses of Fourth and Eighth Grade Students and Teachers to Questionnaire Items	Arnold Goldstein
Data warehouse		
2000–04	Selected Papers on Education Surveys: Papers Presented at the 1998 and 1999 ASA and 1999 AAPOR Meetings	Dan Kasprzyk
Design effects		
2000–03	Strengths and Limitations of Using SUDAAN, Stata, and WesVarPC for Computing Variances from NCES Data Sets	Ralph Lee
Dropout rates, high school		
95–07	National Education Longitudinal Study of 1988: Conducting Trend Analyses HS&B and NELS:88 Sophomore Cohort Dropouts	Jeffrey Owings
Early childhood education		
96–20	1991 National Household Education Survey (NHES:91) Questionnaires: Screener, Early Childhood Education, and Adult Education	Kathryn Chandler
96–22	1995 National Household Education Survey (NHES:95) Questionnaires: Screener, Early Childhood Program Participation, and Adult Education	Kathryn Chandler
97–24	Formulating a Design for the ECLS: A Review of Longitudinal Studies	Jerry West
97–36	Measuring the Quality of Program Environments in Head Start and Other Early Childhood Programs: A Review and Recommendations for Future Research	Jerry West
1999–01	A Birth Cohort Study: Conceptual and Design Considerations and Rationale	Jerry West
2001–02	Measuring Father Involvement in Young Children's Lives: Recommendations for a Fatherhood Module for the ECLS-B	Jerry West
2001–03	Measures of Socio-Emotional Development in Middle School	Elvira Hausken
2001–06	Papers from the Early Childhood Longitudinal Studies Program: Presented at the 2001 AERA and SRCD Meetings	Jerry West
2002-05	Early Childhood Longitudinal Study-Kindergarten Class of 1998–99 (ECLS–K), Psychometric Report for Kindergarten Through First Grade	Elvira Hausken
Educational attainment		
98–11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96–98) Field Test Report	Aurora D'Amico
2001–15	Baccalaureate and Beyond Longitudinal Study: 2000/01 Follow-Up Field Test Methodology Report	Andrew G. Malizio
Educational research		
2000–02	Coordinating NCES Surveys: Options, Issues, Challenges, and Next Steps	Valena Plisko
2002–01	Legal and Ethical Issues in the Use of Video in Education Research	Patrick Gonzales
Eighth-graders		
2001–05	Using TIMSS to Analyze Correlates of Performance Variation in Mathematics	Patrick Gonzales
2002-07	Teacher Quality, School Context, and Student Race/Ethnicity: Findings from the Eighth Grade National Assessment of Educational Progress 2000 Mathematics Assessment	Janis Brown
Employment		
96–03	National Education Longitudinal Study of 1988 (NELS:88) Research Framework and Issues	Jeffrey Owings
98–11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96–98) Field Test Report	Aurora D'Amico
2000–16a	Lifelong Learning NCES Task Force: Final Report Volume I	Lisa Hudson
2000–16b	Lifelong Learning NCES Task Force: Final Report Volume II	Lisa Hudson
2001–01	Cross-National Variation in Educational Preparation for Adulthood: From Early Adolescence to Young Adulthood	Elvira Hausken
Employment – after college		

No.	Title	NCES contact
2001–15	Baccalaureate and Beyond Longitudinal Study: 2000/01 Follow-Up Field Test Methodology Report	Andrew G. Malizio
Engineering		
2000–11	Financial Aid Profile of Graduate Students in Science and Engineering	Aurora D'Amico
Enrollment – after college		
2001–15	Baccalaureate and Beyond Longitudinal Study: 2000/01 Follow-Up Field Test Methodology Report	Andrew G. Malizio
Faculty – higher education		
97–26	Strategies for Improving Accuracy of Postsecondary Faculty Lists	Linda Zimbler
2000–01	1999 National Study of Postsecondary Faculty (NSOPF:99) Field Test Report	Linda Zimbler
2002–08	A Profile of Part-time Faculty: Fall 1998	Linda Zimbler
Fathers – role in education		
2001–02	Measuring Father Involvement in Young Children's Lives: Recommendations for a Fatherhood Module for the ECLS-B	Jerry West
Finance – elementary and secondary schools		
94–05	Cost-of-Education Differentials Across the States	William J. Fowler, Jr.
96–19	Assessment and Analysis of School-Level Expenditures	William J. Fowler, Jr.
98–01	Collection of Public School Expenditure Data: Development of a Questionnaire	Stephen Broughman
1999–07	Collection of Resource and Expenditure Data on the Schools and Staffing Survey	Stephen Broughman
1999–16	Measuring Resources in Education: From Accounting to the Resource Cost Model Approach	William J. Fowler, Jr.
2000–18	Feasibility Report: School-Level Finance Pretest, Public School District Questionnaire	Stephen Broughman
Finance – postsecondary		
97–27	Pilot Test of IPEDS Finance Survey	Peter Stowe
2000–14	IPEDS Finance Data Comparisons Under the 1997 Financial Accounting Standards for Private, Not-for-Profit Institutes: A Concept Paper	Peter Stowe
Finance – private schools		
95–17	Estimates of Expenditures for Private K–12 Schools	Stephen Broughman
96–16	Strategies for Collecting Finance Data from Private Schools	Stephen Broughman
97–07	The Determinants of Per-Pupil Expenditures in Private Elementary and Secondary Schools: An Exploratory Analysis	Stephen Broughman
97–22	Collection of Private School Finance Data: Development of a Questionnaire	Stephen Broughman
1999–07	Collection of Resource and Expenditure Data on the Schools and Staffing Survey	Stephen Broughman
2000–15	Feasibility Report: School-Level Finance Pretest, Private School Questionnaire	Stephen Broughman
Geography		
98–04	Geographic Variations in Public Schools' Costs	William J. Fowler, Jr.
Graduate students		
2000–11	Financial Aid Profile of Graduate Students in Science and Engineering	Aurora D'Amico
Graduates of postsecondary education		
2001–15	Baccalaureate and Beyond Longitudinal Study: 2000/01 Follow-Up Field Test Methodology Report	Andrew G. Malizio
Imputation		
2000–04	Selected Papers on Education Surveys: Papers Presented at the 1998 and 1999 ASA and 1999 AAPOR Meeting	Dan Kasprzyk
2001–10	Comparison of Proc Impute and Schafer's Multiple Imputation Software	Sam Peng
2001–16	Imputation of Test Scores in the National Education Longitudinal Study of 1988	Ralph Lee
2001–17	A Study of Imputation Algorithms	Ralph Lee
2001–18	A Study of Variance Estimation Methods	Ralph Lee

No.	Title	NCES contact
Inflation		
97–43	Measuring Inflation in Public School Costs	William J. Fowler, Jr.
Institution data		
2000–01	1999 National Study of Postsecondary Faculty (NSOPF:99) Field Test Report	Linda Zimbler
Instructional resources and practices		
95–11	Measuring Instruction, Curriculum Content, and Instructional Resources: The Status of Recent Work	Sharon Bobbitt & John Ralph
1999–08	Measuring Classroom Instructional Processes: Using Survey and Case Study Field Test Results to Improve Item Construction	Dan Kasprzyk
International comparisons		
97–11	International Comparisons of Inservice Professional Development	Dan Kasprzyk
97–16	International Education Expenditure Comparability Study: Final Report, Volume I	Shelley Burns
97–17	International Education Expenditure Comparability Study: Final Report, Volume II, Quantitative Analysis of Expenditure Comparability	Shelley Burns
2001–01	Cross-National Variation in Educational Preparation for Adulthood: From Early Adolescence to Young Adulthood	Elvira Hausken
2001–07	A Comparison of the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study Repeat (TIMSS-R), and the Programme for International Student Assessment (PISA)	Arnold Goldstein
International comparisons – math and science achievement		
2001–05	Using TIMSS to Analyze Correlates of Performance Variation in Mathematics	Patrick Gonzales
Libraries		
94–07	Data Comparability and Public Policy: New Interest in Public Library Data Papers Presented at Meetings of the American Statistical Association	Carrol Kindel
97–25	1996 National Household Education Survey (NHES:96) Questionnaires: Screener/Household and Library, Parent and Family Involvement in Education and Civic Involvement, Youth Civic Involvement, and Adult Civic Involvement	Kathryn Chandler
Limited English Proficiency		
95–13	Assessing Students with Disabilities and Limited English Proficiency	James Houser
2001–11	Impact of Selected Background Variables on Students' NAEP Math Performance	Arnold Goldstein
2001–13	The Effects of Accommodations on the Assessment of LEP Students in NAEP	Arnold Goldstein
Literacy of adults		
98–17	Developing the National Assessment of Adult Literacy: Recommendations from Stakeholders	Sheida White
1999–09a	1992 National Adult Literacy Survey: An Overview	Alex Sedlacek
1999–09b	1992 National Adult Literacy Survey: Sample Design	Alex Sedlacek
1999–09c	1992 National Adult Literacy Survey: Weighting and Population Estimates	Alex Sedlacek
1999–09d	1992 National Adult Literacy Survey: Development of the Survey Instruments	Alex Sedlacek
1999–09e	1992 National Adult Literacy Survey: Scaling and Proficiency Estimates	Alex Sedlacek
1999–09f	1992 National Adult Literacy Survey: Interpreting the Adult Literacy Scales and Literacy Levels	Alex Sedlacek
1999–09g	1992 National Adult Literacy Survey: Literacy Levels and the Response Probability Convention	Alex Sedlacek
1999–11	Data Sources on Lifelong Learning Available from the National Center for Education Statistics	Lisa Hudson
2000–05	Secondary Statistical Modeling With the National Assessment of Adult Literacy: Implications for the Design of the Background Questionnaire	Sheida White
2000–06	Using Telephone and Mail Surveys as a Supplement or Alternative to Door-to-Door Surveys in the Assessment of Adult Literacy	Sheida White
2000–07	"How Much Literacy is Enough?" Issues in Defining and Reporting Performance Standards for the National Assessment of Adult Literacy	Sheida White
2000–08	Evaluation of the 1992 NALS Background Survey Questionnaire: An Analysis of Uses with Recommendations for Revisions	Sheida White

No.	Title	NCES contact
2000–09	Demographic Changes and Literacy Development in a Decade	Sheida White
2001–08	Assessing the Lexile Framework: Results of a Panel Meeting	Sheida White
Literacy of adults – international		
97–33	Adult Literacy: An International Perspective	Marilyn Binkley
Mathematics		
98–09	High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988	Jeffrey Owings
1999–08	Measuring Classroom Instructional Processes: Using Survey and Case Study Field Test Results to Improve Item Construction	Dan Kasprzyk
2001–05	Using TIMSS to Analyze Correlates of Performance Variation in Mathematics	Patrick Gonzales
2001–07	A Comparison of the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study Repeat (TIMSS-R), and the Programme for International Student Assessment (PISA)	Arnold Goldstein
2001–11	Impact of Selected Background Variables on Students' NAEP Math Performance	Arnold Goldstein
2002-06	The Measurement of Instructional Background Indicators: Cognitive Laboratory Investigations of the Responses of Fourth and Eighth Grade Students and Teachers to Questionnaire Items	Arnold Goldstein
2002-07	Teacher Quality, School Context, and Student Race/Ethnicity: Findings from the Eighth Grade National Assessment of Educational Progress 2000 Mathematics Assessment	Janis Brown
Parental involvement in education		
96–03	National Education Longitudinal Study of 1988 (NELS:88) Research Framework and Issues	Jeffrey Owings
97–25	1996 National Household Education Survey (NHES:96) Questionnaires: Screener/Household and Library, Parent and Family Involvement in Education and Civic Involvement, Youth Civic Involvement, and Adult Civic Involvement	Kathryn Chandler
1999–01	A Birth Cohort Study: Conceptual and Design Considerations and Rationale	Jerry West
2001–06	Papers from the Early Childhood Longitudinal Studies Program: Presented at the 2001 AERA and SRCD Meetings	Jerry West
2001–19	The Measurement of Home Background Indicators: Cognitive Laboratory Investigations of the Responses of Fourth and Eighth Graders to Questionnaire Items and Parental Assessment of the Invasiveness of These Items	Arnold Goldstein
Participation rates		
98–10	Adult Education Participation Decisions and Barriers: Review of Conceptual Frameworks and Empirical Studies	Peter Stowe
Postsecondary education		
1999–11	Data Sources on Lifelong Learning Available from the National Center for Education Statistics	Lisa Hudson
2000–16a	Lifelong Learning NCES Task Force: Final Report Volume I	Lisa Hudson
2000–16b	Lifelong Learning NCES Task Force: Final Report Volume II	Lisa Hudson
Postsecondary education – persistence and attainment		
98–11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96–98) Field Test Report	Aurora D'Amico
1999–15	Projected Postsecondary Outcomes of 1992 High School Graduates	Aurora D'Amico
Postsecondary education – staff		
97–26	Strategies for Improving Accuracy of Postsecondary Faculty Lists	Linda Zimbler
2000–01	1999 National Study of Postsecondary Faculty (NSOPF:99) Field Test Report	Linda Zimbler
2002–08	A Profile of Part-time Faculty: Fall 1998	Linda Zimbler
Principals		
2000–10	A Research Agenda for the 1999–2000 Schools and Staffing Survey	Dan Kasprzyk
Private schools		
96–16	Strategies for Collecting Finance Data from Private Schools	Stephen Broughman

No.	Title	NCES contact
97–07	The Determinants of Per-Pupil Expenditures in Private Elementary and Secondary Schools: An Exploratory Analysis	Stephen Broughman
97–22	Collection of Private School Finance Data: Development of a Questionnaire	Stephen Broughman
2000–13	Non-professional Staff in the Schools and Staffing Survey (SASS) and Common Core of Data (CCD)	Kerry Gruber
2000–15	Feasibility Report: School-Level Finance Pretest, Private School Questionnaire	Stephen Broughman
Projections of education statistics		
1999–15	Projected Postsecondary Outcomes of 1992 High School Graduates	Aurora D'Amico
Public school finance		
1999–16	Measuring Resources in Education: From Accounting to the Resource Cost Model Approach	William J. Fowler, Jr.
2000–18	Feasibility Report: School-Level Finance Pretest, Public School District Questionnaire	Stephen Broughman
Public schools		
97–43	Measuring Inflation in Public School Costs	William J. Fowler, Jr.
98–01	Collection of Public School Expenditure Data: Development of a Questionnaire	Stephen Broughman
98–04	Geographic Variations in Public Schools' Costs	William J. Fowler, Jr.
1999–02	Tracking Secondary Use of the Schools and Staffing Survey Data: Preliminary Results	Dan Kasprzyk
2000–12	Coverage Evaluation of the 1994–95 Public Elementary/Secondary School Universe Survey	Beth Young
2000–13	Non-professional Staff in the Schools and Staffing Survey (SASS) and Common Core of Data (CCD)	Kerry Gruber
2002–02	Locale Codes 1987 - 2000	Frank Johnson
Public schools – secondary		
98–09	High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988	Jeffrey Owings
Reform, educational		
96–03	National Education Longitudinal Study of 1988 (NELS:88) Research Framework and Issues	Jeffrey Owings
Response rates		
98–02	Response Variance in the 1993–94 Schools and Staffing Survey: A Reinterview Report	Steven Kaufman
School districts		
2000–10	A Research Agenda for the 1999–2000 Schools and Staffing Survey	Dan Kasprzyk
School districts, public		
98–07	Decennial Census School District Project Planning Report	Tai Phan
1999–03	Evaluation of the 1996–97 Nonfiscal Common Core of Data Surveys Data Collection, Processing, and Editing Cycle	Beth Young
School districts, public – demographics of		
96–04	Census Mapping Project/School District Data Book	Tai Phan
Schools		
97–42	Improving the Measurement of Staffing Resources at the School Level: The Development of Recommendations for NCES for the Schools and Staffing Survey (SASS)	Mary Rollefson
98–08	The Redesign of the Schools and Staffing Survey for 1999–2000: A Position Paper	Dan Kasprzyk
1999–03	Evaluation of the 1996–97 Nonfiscal Common Core of Data Surveys Data Collection, Processing, and Editing Cycle	Beth Young
2000–10	A Research Agenda for the 1999–2000 Schools and Staffing Survey	Dan Kasprzyk
2002–02	Locale Codes 1987 – 2000	Frank Johnson
2002-07	Teacher Quality, School Context, and Student Race/Ethnicity: Findings from the Eighth Grade National Assessment of Educational Progress 2000 Mathematics Assessment	Janis Brown

No.	Title	NCES contact
Schools – safety and discipline		
97–09	Status of Data on Crime and Violence in Schools: Final Report	Lee Hoffman
Science		
2000–11	Financial Aid Profile of Graduate Students in Science and Engineering	Aurora D'Amico
2001–07	A Comparison of the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study Repeat (TIMSS-R), and the Programme for International Student Assessment (PISA)	Arnold Goldstein
Software evaluation		
2000–03	Strengths and Limitations of Using SUDAAN, Stata, and WesVarPC for Computing Variances from NCES Data Sets	Ralph Lee
Staff		
97–42	Improving the Measurement of Staffing Resources at the School Level: The Development of Recommendations for NCES for the Schools and Staffing Survey (SASS)	Mary Rollefson
98–08	The Redesign of the Schools and Staffing Survey for 1999–2000: A Position Paper	Dan Kasprzyk
Staff – higher education institutions		
97–26	Strategies for Improving Accuracy of Postsecondary Faculty Lists	Linda Zimbler
2002–08	A Profile of Part-time Faculty: Fall 1998	Linda Zimbler
Staff – nonprofessional		
2000–13	Non-professional Staff in the Schools and Staffing Survey (SASS) and Common Core of Data (CCD)	Kerry Gruber
State		
1999–03	Evaluation of the 1996–97 Nonfiscal Common Core of Data Surveys Data Collection, Processing, and Editing Cycle	Beth Young
Statistical methodology		
97–21	Statistics for Policymakers or Everything You Wanted to Know About Statistics But Thought You Could Never Understand	Susan Ahmed
Statistical standards and methodology		
2001–05	Using TIMSS to Analyze Correlates of Performance Variation in Mathematics	Patrick Gonzales
2002–04	Improving Consistency of Response Categories Across NCES Surveys	Marilyn Seastrom
Students with disabilities		
95–13	Assessing Students with Disabilities and Limited English Proficiency	James Houser
2001–13	The Effects of Accommodations on the Assessment of LEP Students in NAEP	Arnold Goldstein
Survey methodology		
96–17	National Postsecondary Student Aid Study: 1996 Field Test Methodology Report	Andrew G. Malizio
97–15	Customer Service Survey: Common Core of Data Coordinators	Lee Hoffman
97–35	Design, Data Collection, Interview Administration Time, and Data Editing in the 1996 National Household Education Survey	Kathryn Chandler
98–06	National Education Longitudinal Study of 1988 (NELS:88) Base Year through Second Follow-Up: Final Methodology Report	Ralph Lee
98–11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96–98) Field Test Report	Aurora D'Amico
98–16	A Feasibility Study of Longitudinal Design for Schools and Staffing Survey	Stephen Broughman
1999–07	Collection of Resource and Expenditure Data on the Schools and Staffing Survey	Stephen Broughman
1999–17	Secondary Use of the Schools and Staffing Survey Data	Susan Wiley
2000–01	1999 National Study of Postsecondary Faculty (NSOPF:99) Field Test Report	Linda Zimbler
2000–02	Coordinating NCES Surveys: Options, Issues, Challenges, and Next Steps	Valena Plisko
2000–04	Selected Papers on Education Surveys: Papers Presented at the 1998 and 1999 ASA and 1999 AAPOR Meetings	Dan Kasprzyk
2000–12	Coverage Evaluation of the 1994–95 Public Elementary/Secondary School Universe Survey	Beth Young
2000–17	National Postsecondary Student Aid Study:2000 Field Test Methodology Report	Andrew G. Malizio

No.	Title	NCES contact
2001–04	Beginning Postsecondary Students Longitudinal Study: 1996–2001 (BPS:1996/2001) Field Test Methodology Report	Paula Knepper
2001–07	A Comparison of the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study Repeat (TIMSS-R), and the Programme for International Student Assessment (PISA)	Arnold Goldstein
2001–11	Impact of Selected Background Variables on Students' NAEP Math Performance	Arnold Goldstein
2001–13	The Effects of Accommodations on the Assessment of LEP Students in NAEP	Arnold Goldstein
2001–19	The Measurement of Home Background Indicators: Cognitive Laboratory Investigations of the Responses of Fourth and Eighth Graders to Questionnaire Items and Parental Assessment of the Invasiveness of These Items	Arnold Goldstein
2002–01	Legal and Ethical Issues in the Use of Video in Education Research	Patrick Gonzales
2002–02	Locale Codes 1987 - 2000	Frank Johnson
2002–03	National Postsecondary Student Aid Study, 1999–2000 (NPSAS:2000), CATI Nonresponse Bias Analysis Report.	Andrew Malizio
2002-06	The Measurement of Instructional Background Indicators: Cognitive Laboratory Investigations of the Responses of Fourth and Eighth Grade Students and Teachers to Questionnaire Items	Arnold Goldstein
2003-03	Education Longitudinal Study: 2002 (ELS: 2002) Field Test Report	Jeffrey Owings
Teachers		
98–13	Response Variance in the 1994–95 Teacher Follow-up Survey	Steven Kaufman
1999–14	1994–95 Teacher Followup Survey: Data File User's Manual, Restricted-Use Codebook	Kerry Gruber
2000–10	A Research Agenda for the 1999–2000 Schools and Staffing Survey	Dan Kasprzyk
2002-07	Teacher Quality, School Context, and Student Race/Ethnicity: Findings from the Eighth Grade National Assessment of Educational Progress 2000 Mathematics Assessment	Janis Brown
Teachers – instructional practices of		
98–08	The Redesign of the Schools and Staffing Survey for 1999–2000: A Position Paper	Dan Kasprzyk
2002-06	The Measurement of Instructional Background Indicators: Cognitive Laboratory Investigations of the Responses of Fourth and Eighth Grade Students and Teachers to Questionnaire Items	Arnold Goldstein
Teachers – opinions regarding safety		
98–08	The Redesign of the Schools and Staffing Survey for 1999–2000: A Position Paper	Dan Kasprzyk
Teachers – performance evaluations		
1999–04	Measuring Teacher Qualifications	Dan Kasprzyk
Teachers – qualifications of		
1999–04	Measuring Teacher Qualifications	Dan Kasprzyk
Teachers – salaries of		
94–05	Cost-of-Education Differentials Across the States	William J. Fowler, Jr.
Training		
2000–16a	Lifelong Learning NCES Task Force: Final Report Volume I	Lisa Hudson
2000–16b	Lifelong Learning NCES Task Force: Final Report Volume II	Lisa Hudson
Variance estimation		
2000–03	Strengths and Limitations of Using SUDAAN, Stata, and WesVarPC for Computing Variances from NCES Data Sets	Ralph Lee
2000–04	Selected Papers on Education Surveys: Papers Presented at the 1998 and 1999 ASA and 1999 AAPOR Meetings	Dan Kasprzyk
2001–18	A Study of Variance Estimation Methods	Ralph Lee
Violence		
97–09	Status of Data on Crime and Violence in Schools: Final Report	Lee Hoffman
Vocational education		
95–12	Rural Education Data User's Guide	Samuel Peng
1999–05	Procedures Guide for Transcript Studies	Dawn Nelson

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

	This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.
X	This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").