
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.

This page intentionally left blank.

NATIONAL CENTER FOR EDUCATION STATISTICS

Working Paper Series

**NAEP VALIDITY STUDIES:
IMPLICATIONS OF ELECTRONIC
TECHNOLOGY FOR THE
NAEP ASSESSMENT**

Working Paper No. 2003-16

April 2003

Contact: Patricia Dabbs
Field Operations and Publication Management
Patricia.Dabbs@ed.gov

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

U.S. Department of Education

Rod Paige
Secretary

Institute of Education Sciences

Grover J. Whitehurst
Director

National Center for Education Statistics

Val 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

March 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. *NAEP Validity Studies: Implications of Electronic Technology for the NAEP Assessment*. NCES 2003-16, by Richard P. Durán. Project Officer: Patricia Dabbs. 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:

Patricia Dabbs
(202) 502-7332
Patricia.Dabbs@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) 502-7363, e-mail: sheilah.jupiter@ed.gov, or mail: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 1990 K Street NW, Room 9048, Washington, DC 20006.

Marilyn M. Seastrom
Chief Mathematical Statistician
Statistical Standards Program

Ralph Lee
Mathematical Statistician
Statistical Standards Program

This page intentionally left blank.

Implications of Electronic Technology for the NAEP Assessment

Richard P. Durán
University of California, Santa Barbara

Commissioned by the NAEP Validity Studies (NVS) Panel
August 2000

George W. Bohrnstedt, Panel Chair
Frances B. Stancavage, Project Director

The NAEP Validity Studies Panel was formed by the American Institutes for Research under contract with the National Center for Education Statistics. Points of view or opinions expressed in this paper do not necessarily represent the official positions of the U.S. Department of Education or the American Institutes for Research.

The NAEP Validity Studies (NVS) Panel was formed in 1995 to provide a technical review of NAEP plans and products and to identify technical concerns and promising techniques worthy of further study and research. The members of the panel have been charged with writing focused studies and issue papers on the most salient of the identified issues.

Panel Members:

Albert E. Beaton
Boston College

John A. Dossey
Illinois State University

Don McLaughlin
American Institutes for Research

R. Darrell Bock
University of Chicago

Richard P. Durán
University of California

Ina V.S. Mullis
Boston College

George W. Bohrnstedt, Chair
American Institutes for Research

David Grissmer
RAND

Harry O'Neil
University of Southern California

Audrey Champagne
University at Albany, SUNY

Larry Hedges
University of Chicago

P. David Pearson
Michigan State University

James R. Chromy
Research Triangle Institute

Gerunda Hughes
Howard University

Lorrie Shepard
University of Colorado

Pasquale DeVito
Rhode Island Dept. of Education

Richard Jaeger
University of North Carolina

Project Director:

Frances B. Stancavage
American Institutes for Research

Project Officer:

Patricia Dabbs
National Center for Education Statistics

For Information:

NAEP Validity Studies (NVS)
American Institutes for Research
1791 Arastradero Road
Palo Alto, CA 94304-1337
Phone: 650/ 493-3550
Fax: 650/ 858-0958

Acknowledgments

Appreciation is expressed to NAEP Validity Studies Panel members for their guidance and review. Special thanks are extended to Albert Beaton, Darrell Bock, James Chromy, Robert Linn, and Donald McLaughlin for their editorial help.

Table of Contents

Overview	4
NAEP and Electronic Technology: A Public Concern	5
Changes in Use of Technology in Schooling: Implications for NAEP.....	6
Shaping and Sharpening NAEP’s Technology Agenda: A Construct Validity Centered Perspective	7
Validity Issues for Technological Assessment Innovations.....	8
Shaping and Sharpening NAEP’s Technology Agenda: Identifying Priorities	9
Assessment of Existing NAEP Constructs, Using Computers for Item Presentation and Response Recording	10
Recommendation of the ETS/NCS/Westat Redesign Team: CBT Pilot in Math and Science	14
Validity Issues for Computerized Assessment of Existing Constructs	15
Extension of NAEP to the Assessment of New Constructs, Using Computers	16
Priority of Computer Administered Writing Test	16
Potential for Assessing New Modes of Learning and Problem-Solving.....	17
Validity Issues for New Constructs.....	19
Technological Enhancements of Assessment Processes Other than Computer Presentation of Items and Recording of Responses	20
Redesign of NAEP Assessments as an Integrated Electronic Information System	21
Validity Issues	23
Recommendations and Summary	23
Design of a Computerized NAEP Field Trial	25
Selection of an Assessment	25
Choice of a Sample	25
Computer Delivery	26

Internet Delivery	26
Computer Experience and Accommodations.....	26
References	27

List of Figures

Figure 1. Possible NAEP Priority Areas for the Introduction of New Technologies into Assessments	12
---	----

Overview

The National Assessment of Educational Progress (NAEP) lives in a changing assessment world. Its credibility and utility as an assessment system requires that NAEP maintain the stability of its assessments so that they capture change in students' learning across years. In addition, in its role as a model for the state of the art in the assessment of educational achievement, NAEP must embrace new forms of assessment that are consistent with changing educational needs of students. Electronic technologies have relevance to both goals. The rapid spread and infusion of technologies such as computers, digital media, the Internet, video and audio recorders, and playback devices into every phase of contemporary life is affecting both the methods of learning and assessment and the content of what needs to be learned in schools.

However, embracing new technologies does not mean NAEP should rush to use every new technology in operational assessments. On the contrary, in its position of leadership, NAEP must thoroughly evaluate new technologies to address both validity and cost issues and introduce them to operational NAEP only when these issues have been addressed. Thorough evaluation studies are, themselves, costly in time and money, so it is important to identify and concentrate on technological innovations in NAEP assessments that build coherently on NAEP's existing mission and on the planning process for the future of the NAEP. These concerns have been acknowledged in deliberations on the redesign of NAEP commissioned by the National Assessment Governing Board (NAGB) (Forysth, Hambleton, Linn, Mislavy, and Yen, 1996), the National Academy of Education (NAE) Capstone Report on the evaluation of the NAEP state assessment system, *Assessment in Transition: Monitoring the Nation's Educational Progress* (Glaser, Linn, and Bohrnstedt, 1997), and the National Academy of Science (NAS) evaluation of NAEP, *Grading the Nation's Report Card*, (Pellegrino, Jones, and Mitchell, 1999).

The focus needs to be on how computer-administered NAEP assessments and electronic technology can help us better assess what students know and are capable of doing, and better help us inform the public on students' educational achievement. It is important that NAEP proceed in a timely manner to evaluate and recommend priorities on major electronic technology innovations in the near, intermediate, and far-term future. As part of this process, it will also be valuable to point out ways in which technological developments are having a broader impact on the design and implementation of assessment systems, and on the facilitation of classroom learning and instruction.

The purpose of this paper is to review major options NAEP faces regarding introduction of technology into the assessment and to review priorities that can guide this introduction.

Technological innovations for NAEP currently being considered fall into three broad categories:

-
1. Computer-based presentation of items and recording of responses on assessments of existing NAEP constructs
 2. Extension of NAEP to assessment of new constructs, using technology
 3. Computer enhancement of assessment processes other than presenting items to students and recording their responses

These categories lead to different requirements for research on the impact of innovations on NAEP validity. Thus, early identification of the highest priority innovations is important, so that validity research can be carried out prior to their introduction into NAEP operations.

NAEP and Electronic Technology: A Public Concern

In a background paper for the NAE Capstone Report, Eva Baker noted that the continuing public credibility of NAEP must build upon NAEP's incorporation of new technologies as integral components of the testing process (Baker, 1997). Arguably, NAEP is the nation's most visible assessment. As such it is a showcase for the nation's capability in the field of assessment. Baker posits that the public expects NAEP to be at the forefront of technology and to be a leader in the application of technology to assessment in all its systems. NAEP's use of technology needs to reflect the technology that surrounds and makes possible everyday life, but beyond this, NAEP should consider being a leader in the application of technology to testing in fulfilling its public mission. Nowhere would this visibility be more prominent than in the delivery of computerized assessments.

Computer-delivered assessments are already being implemented by the U.S. Department of Defense (Sands, Waters, and McBride, 1997) and major college and professional school admissions and placement programs (Cole, 1997). Their implementation in K–12 schooling assessment programs, e.g., at the state level, is close on the horizon in states such as Georgia and North Carolina. K–12 test publishers are actively pursuing computer delivered assessments as well. In this climate, it is essential that NAEP keep pace and demonstrate to the public the widespread impact of computers and technology on delivery of its own assessments.

Technological innovations in educational assessment are needed. There is a widespread belief in the educational community that current multiple-choice forms of assessment are not representative of authentic tasks encountered by students as they learn. In particular, multiple-choice item types involving minimal contextualization of text and problem questions are viewed as inherently inadequate for assessing the complex reasoning and literacy skills specified in content and performance standards advocated by reformers. There is a belief that assessment exercises should include complex problems and learning materials that students encounter in effective instruction and that the performance of students on assessments should be measured on constructs appropriate to standards.

NAEP's decisions and strategies on how to incorporate computers and related technologies into its assessments should help the public better understand the implications of the reform movement. In particular, it should help shape the public's responsiveness to the costs and benefits of implementing effective reform, given psychometrically bounded standards of evidence regarding outcomes for students. Presently, NAEP is the only K–12 assessment system capable of addressing these questions at a nationally representative and cross-state level. NAEP's development of computer-delivered tests, new item types utilizing electronic technologies, and new scoring and reporting strategies aligned with reform goals will have a direct and major impact on the public's view of educational reform and its progress.

An important public policy issue that NAEP's use of technology can and must address is the fair testing of all students. Access to computers is greatly affected by the income level of families (Coley, Cradler, and Engel, 1997). Introduction of computer delivered NAEP assessments may aggravate public concern about access to computers across socioeconomic and racial/ethnic groups, and the fairness of computer-delivered tests. In addition, federal Civil Rights policies and education code specify that NAEP and other federally sponsored assessments must include all students in assessments (Olson and Goldstein, 1997).

The introduction of computer delivered NAEP assessments, along with associated multimedia tools, has the potential to address fairness concerns tied to inclusion. Students with disabilities or from non-English backgrounds who have previously been excluded from assessments may be able to participate meaningfully in a NAEP computerized assessment that offers accommodations tailored to overcome particular disabilities or lack of knowledge of English, when that knowledge is not central to constructs under assessment. NAEP is actively engaging in research on accommodations (Olson and Goldstein, 1997), and it is foreseeable that further introduction of technology into its assessments will continue and expand investigations of these concerns.

Changes in Use of Technology in Schooling: Implications for NAEP

Access to computers and related electronic technologies is accelerating in schools, but the implications of the use of these technologies in schools for the NAEP assessment system in the near term are inchoate and need careful monitoring. Some of these trends are cited by Viadero (1997) in a special issue of *Education Week* entitled "Technology Counts, Schools and Reform in the Information Age." For example, NCES data cited by *Education Week* indicate that the percentage of schools with access to the Internet jumped from 35 to 65 percent from 1994 to 1996, while access in individual classrooms jumped from 3 to 14 percent. NCES also found that frequency of computer use is rising. In 1994, for example, 24 percent of eleventh graders used a computer at least once a week, and these uses included math games, simulation and applications, and demonstrations of new topics in math.

Teachers' use of computers in instruction is currently very uneven. For example, NAEP 1996 teacher survey data indicated that 52 percent of eighth grade teachers

didn't use computers for math at all, while 18 percent used computers for drill and practice. Interestingly, NAEP 1994 teacher survey data indicate that fourth graders were more likely to use computers in subject matter learning of reading, history, and geography than eighth graders.

There is also evidence of racial/ethnic differences in access to computers. For example, among 3 to 17-year-olds, Black and Hispanic students have had less access to computers than white students. NAEP 1994 data indicated that 53 percent of Hispanics and 51 percent of Blacks had access to computers at school, as compared to 63 percent of whites. The gap was wider in home settings; 12 percent of Hispanics and 13 percent of Blacks had access to computers at home as compared to 36 percent of whites (U.S. Bureau of the Census, 1993).

These findings (see also U.S. Department of Commerce, 1999) suggest that despite the overall trend of increasing access to computers, immense variability in access remains. These survey data have value as national-level indicators of access to and use of computers. These and future data regarding differences in access offer important contextual information that should be carefully scrutinized as NAEP pursues new measures and item types responsive to changes in the role of electronic technology in the education of students from different backgrounds.

Shaping and Sharpening NAEP's Technology Agenda: A Construct Validity Centered Perspective

The NAEP assessment has contributed in an important manner to national awareness of what students know and should be able to do (Glaser et al., 1997). Measurement of basic skills and content knowledge in critical domains will remain central to NAEP for the foreseeable future. Assessment of these skills could be enhanced considerably by computer delivered NAEP through computerized display of richer textual and figural item content, or through display of dynamically changing information relevant to understanding an item.

But the possibilities of computer delivered assessments, as well as other technological enhancements to NAEP, can go beyond this and make possible improved assessments of more complex thinking skills and assessments of complex skills that have yet to be defined and measured. The NAE Capstone Report mentions a number of possible new assessment areas for NAEP that have this flavor and that could become extensions of existing assessments. Areas include enhanced assessment of:

4. Students' active demonstrations and applications of skill and knowledge
5. Problem-representation skills using student constructed graphs and figures
6. Problem solving strategies, based on analysis of students' response patterns while solving problems on a computer

-
7. Self-monitoring skills, such as checking of work while problem solving
 8. Skill in explaining one's problem solving, such as through automated analysis of written explanations or terms entered into a computer
 9. Skills in interpreting and synthesizing complex pieces of textual and figural information displayed on a computer

While students' constructed responses to new computer delivered item types will be an important source of assessment information, measures of reaction time during different portions of item presentation and student response activity may become an important new source of assessment information, which could help measure constructs associated with a range of cognitive skills and motivation.

Regardless of whether technological innovations are targeted at better assessing existing NAEP skill areas or new skill areas altogether, an assessment validity framework and associated validity research are essential. A central question in this endeavor is "Can we draw accurate inferences about students' performance on a computer administered assessment in a manner that reflects skills, problem solving ability, and knowledge of a target domain of learning identified by NAEP?"

Validity Issues for Technological Assessment Innovations

Regardless of how introduction of new technologies in NAEP assessments arises, there are a number of prominent validity research questions that will surface. These are questions that have been of increasing prominence in the field of computerized testing as a whole (Bennett, Steffen, Singley, Morley, and Jacquemin, 1997; Green, 1998; Segall, 1998). While questions have arisen in the context of computerized testing, they apply more broadly to introduction of electronic technologies into testing. Questions relevant to a broad range of technology-based changes to NAEP might include:

- Comparability across forms: Do different forms of the same computerized (or technologically mediated tests) measure the same constructs and have the same psychometric properties?
- Standardization: What are the critical computer interface components that need standardization across assessments and platforms?
- Differential familiarity: How does examinee familiarity and comfort with technologies used in new assessments affect performance and measurement characteristics of assessments?
- Practice effects: How do instructions and practice affect examinees' readiness for technology based assessments?

-
- Effects on test-taking strategies: Does constraining examinee options to review or omit items in technology based assessments affect performance and the measurement characteristics of assessments?

Each of the foregoing question areas has relevance for examinees regardless of their background, yet each question area has special value in examining the fairness of new technology-based assessments for students from backgrounds that are associated with low school performance.

Shaping and Sharpening NAEP's Technology Agenda: Identifying Priorities

Because the validity of assessment results has been shown to be sensitive to a wide variety of factors, it is essential to ensure the continuing validity of NAEP after introduction of each technological innovation by completing research studies *before* the innovation is introduced. The program of research that would be required to explore all of the innovations that have recently been proposed would be prohibitively expensive, so priorities are needed. Studies can then focus on high priority options before others.

There are three major considerations in determining priorities:

10. Does the innovation address a critical need?
11. How great is the probability of success of the innovation?
12. What are the operational cost implications?

For example, among the most critical needs for NAEP is the development of methods for incorporating testing accommodations that will allow increased participation of special student populations such as students with disabilities and English language learners into NAEP. Technological innovations which might solve accommodation problems should be considered a high priority.

Attending to the second consideration will help to avoid proceeding down dead-end paths. For example, among alternative innovations, those that are based on use of a technology that is stable across assessment contexts should be given higher priority, when other factors are equal. The third consideration is necessary because there are severe constraints on the level of investment that can be made in a national assessment. Organizations which have attempted to move into computerized testing have found that not all operational costs are readily apparent in the initial design phase.

Determining the critical needs of NAEP is a complex task involving many players. In particular, priorities will be affected by measurement policy objectives set by the National Assessment Governing Board (NAGB) in consultation with the National Center for Education Statistics (NCES) and relevant constituencies. Inputs from a

variety of perspectives will help to ensure that the priorities reflect analysis of costs, system redesign implications, and time implementation goals.

Reports from the NAGB Design/Feasibility Team, the NAE Capstone Report, and the NAS NAEP evaluation concur that there are numerous ways that technological innovations might improve NAEP and the achievement constructs it measures. Implementation of computer administered NAEP assessments for the purposes of improving assessment of important learning achievement constructs is a high priority and includes the possibility that introduction of computerized NAEP assessments would be coupled with other uses of technology to develop and score items presented via a computer.

In 1996 NCES awarded a series of small grants to individuals and organizations interested in contributing to the dialogue on the NAEP redesign. Working under one of these grants, the current NAEP contractors—the Educational Testing Service (ETS), National Computer Systems (NCS), and Westat—collaborated to produce a proposal for an integrated redesign of NAEP (Johnson, Lazar, and O’Sullivan, 1997). Their report included several specific suggestions for the near-term introduction of technology into the NAEP program, thereby moving the dialogue forward beyond any previous published position. Figure 1 shows the priorities for introducing technology into NAEP assessments that are reflected in the ETS/NCS/Westat redesign report, the NAE Capstone Report, and NCES priorities in assessment accommodations research. Innovations are divided into “highest priority” and “other high priorities,” within the broad categories of actually conducting assessments of (1) existing, (2) new constructs, and (3) supporting other technological enhancements of the assessment process. While research priorities need to be set, it is not necessarily the case that only priorities that are the highest for the existing NAEP ought to determine what immediate research and development occurs. Designing and planning extensions to the existing NAEP that utilizes technology also need prioritization in the near term if NAEP is to expand its coverage of constructs in tune with students’ increasing reliance on computers in day-to-day instruction (Pellegrino, et al., 1999).

Assessment of Existing NAEP Constructs, Using Computers for Item Presentation and Response Recording

NAEP’s introduction of computer administered assessments should stem from decisions about which target assessment areas are in need most of improved coverage or which assessment areas would demonstrate the advantages of computer administration in ensuring score accuracy and quality of information derived from an assessment. NAEP has multiple measurement targets, and computer administered assessments could help improve inferences about student achievement in either broad domains of knowledge or with regard to specific skills and knowledge. The two prominent options examined by the ETS/NCS/Westat redesign team (Johnson et al., 1997) include:

- Computer administered, non adaptive assessments that assess skills not amenable to pencil-and-paper tests or introduce efficiencies

-
- Computer adaptive tests that produce NAEP scale scores

Under the first option, a computer presents items to students in a single predetermined order, but with potential improvements in presenting textual or figural information and recording responses. These enhancements could improve an assessment's coverage of existing achievement constructs, if research confirms that they do not compromise those constructs.

**Figure 1—
Possible NAEP Priority Areas for the Introduction of New
Technologies into Assessments**

Use of Technology	Highest Priorities for NAEP	Other High Priorities for Extending NAEP
Assessment of existing NAEP constructs by computer	<p>Implementation of computer based testing (CBT) in a core area, such as math or science</p> <p>Computer- and multimedia-based accommodations for students with special sensory, response, and timing needs</p>	<p>Implementation of computer adaptive testing (CAT)</p> <p>Accommodation by broadening the skill range of test items.</p>
Extension of NAEP to the assessment of new constructs, using computers	<p>A demonstration of innovative computerized assessment of a skill area in which computer use is a significant component</p> <p>For example, assessment of writing using word-processing</p>	<p>CBT in new domains of school learning, such as the use of the Internet for research</p> <p>New assessment strategies that make use of technology:</p> <ul style="list-style-type: none"> • conditional test items, to probe proficiency in content, skill, and problem solving areas, such as scientific inquiry • integrated assessment and instruction • audio computer assisted self-interviewing in new and existing construct areas

Use of Technology	Highest Priorities for NAEP	Other High Priorities for Extending NAEP
Other technological enhancements of assessment processes	<p>Computerized item development</p> <p>Computerized scoring of responses on existing assessments against new constructs</p> <p>Use of Internet and multimedia technologies to disseminate reports</p>	<p>Computerized delivery of assessment materials to schools with possible further local processing of response data</p> <p>Internet and wireless item presentation</p> <p>Use of technology and the Internet to permit on-line construction of custom data reports</p> <p>On line electronic forums for the discussion and interpretation of NAEP results</p> <p>Redesign of NAEP assessments as integrated electronic information systems</p>

The second option computer adaptive testing (CAT) has the same potential for improvements, but differs from the first in that each examinee's ability in a construct area is provisionally estimated at several points during the test, based on computerized scoring of responses to items already presented. The record of items and estimates of ability based on previous responses are used to select items to be presented subsequently, improving the accuracy of the estimate of an examinee's underlying achievement level. This procedure is both statistically and practically efficient, in that examinees' achievement levels are assessed with greater accuracy across levels and with fewer items than is possible on a linear test. In the extreme case, provisional estimates and item selections are made after each item; however, a less radical departure from traditional (linear) testing might involve scoring a "pretest" of several items to decide which of several alternative forms to administer during the remainder of the test session.

A NAEP CAT assessment in a target assessment area could yield individual examinee scores with acceptable accuracy for reporting, something that is not possible now given the matrix sampling design of NAEP forms and the length of NAEP tests. Further, NAEP CAT assessments would not need to rely on plausible values imputation of NAEP scores, obviating the misperception that examinee background variables are inferred by NAEP to be sufficient to accurately explain students' achievement level.

Recommendation of the ETS/NCS/Westat Redesign Team: CBT Pilot in Math and Science

After reviewing various arguments for and against implementation of a computer adaptive NAEP versus a linear computer delivered NAEP, the ETS/NCS/Westat redesign team recommended against immediate implementation of a NAEP CAT (Johnson et al., 1997).

The arguments against immediate implementation of a NAEP CAT are substantial. First and foremost, NAEP is not designed to be a precise individual student assessment, but rather an assessment of how well students in aggregate perform on tasks systematically sampled from target achievement domains. Because the greatest advantage of CAT is in the efficient diagnosis of individual abilities, a CAT assessment is not particularly well-suited for NAEP. A second concern relates to the size of the item pool needed for CAT. Because the focus of NAEP is more on representative achievement in curricular domains than on individual student differences, there is always continuing pressure for assessments to include more items representative of the breadth and depth of domains and to include sufficient numbers of items to ensure reliability of average proficiency estimates. The feasibility of rapidly developing item pools with known psychometric parameters of sufficient size for use in adaptive tests is not clear.

The ETS/NCS/Westat redesign team was more sanguine regarding development of a NAEP computer-based test involving a fixed linear presentation of test items on a computer. They recommended that such a test be developed in the near term (within the next two years) and

...that as a first step NAEP develop an explicit set of criteria for evaluating the potential uses of computerized testing in NAEP and then use these criteria to identify a small number of promising opportunities (Johnson et al., 1997, p. 4-41).

The team concluded by recommending that (pp. 4-41–4-42):

...NAEP commit to use one CBT pilot module with either science or mathematics assessment by the year 2000. This module should be experimental, and would likely involve the use of response protocols with students. The results of this study would give NAEP valuable information on equity and feasibility issues, and would help point the direction for future efforts. We strongly believe that the initial uses of CBT in NAEP should be to measure outcomes not easily accessed through pencil-and-paper testing, and that the program's short-term goal should not be to use CATs to generate the domain scales used in primary NAEP reporting.

Perhaps the most urgently needed enhancements of NAEP which computer administration of the assessment could provide are accommodations for students with special assessment needs. Both presentation of items and recording of responses could potentially be implemented more efficiently by computer rather than one-on-one administrations by trained administrators. The flexibility added when

paper and pencil are left behind offers many new avenues for accommodation. In particular, whether by computer or in some other manner, NAEP needs to expand the range of performance it measures, challenging the most proficient students while not frustrating less prepared students. If choices between easier and harder versions of the assessment can be made based on performance on responses to initial items, then precision of estimates need not be sacrificed when accommodating the broader range of proficiency.

Validity Issues for Computerized Assessment of Existing Constructs

In addition to the overall validity issues for all technological enhancements of NAEP, four validity questions are particularly important when computerized assessments are implemented for existing NAEP constructs.

1. Does computer-based presentation of items and recording of responses assess the constructs that are specified in the framework?

Currently, NAEP content domain frameworks are specified without reference to the mode of assessment. If IRT scaling of computer-based items leads to different scale constructions than the paper-and-pencil presentation, then a careful study of the differences in metacognitive (test-taking) skills involved in different testing modes is essential in order to relate the results to existing frameworks.

2. Will the trend be accurate?

Continued assessment of the same constructs provides the basis for measuring the educational progress of the nation over time. Changing the mode of assessment cannot be allowed to undermine that capability. This applies both to mean proficiency levels and to correlations with background measures. For example, if the minority gap widens, we want to be sure that that is not an artifact of changing from one form of test administration to another. There may be fundamental differences that cannot be directly overcome; however, based on proper research, it may be possible to determine a score adjustment that validates the trend.

3. Are differences between groups affected by computer-based administration?

Even if the trend can be approximately validated only, it is important that differences between groups, for example in mathematics performance, reflect differences in knowledge and skills in the curricular content area, not differences in interactions with a computer-based test administration.

4. Does a computer-based accommodation for an assessment disability alter the construct being assessed?

Like all accommodations, computer-based accommodations raise the

question of whether as a result of the accommodation, an essential part of the skill being assessed is omitted. For example, a “clickable” glossary may be fine for some assessments but not for others.

Extension of NAEP to the Assessment of New Constructs, Using Computers

While the ETS/NCS/Westat redesign team was clear in recommending science or mathematics as target achievement areas for a CBT pilot, no advice was generated regarding ways in which the construct coverage of a CBT might merge with strategies to improve NAEP score reporting or modularization to hone construct coverage— areas that were treated in other sections of the redesign proposal. This issue deserves systematic examination in laying out a research agenda for computer delivered NAEP assessments.

The rapid evolution of technology, coupled with effective innovations in what is taught in schools, opens a wide variety of needs and opportunities for new NAEP assessments; and although new assessments need not be concerned with comparisons of performance levels to prior assessments, they require careful development. Although a variety of possibilities call for study, one area currently stands out as a priority: the assessment of proficiency in the new skills of writing that arise when students use computers for word-processing.

Priority of a Computer Administered Writing Test

While not discussed in the ETS/NCS/Westat redesign report or the NAE Capstone Report, the delivery of a computerized writing assessment has emerged as a high priority issue. Recent data suggest there is growing, widespread use of computers in schools for word processing. 1994 NAEP data (cited by Coley, et al., 1997) indicated that the percentage of students using computers at home or at school to write stories or papers grew in importance over the fourth (68 percent), eighth (82 percent), and eleventh grade (87 percent). The eleventh grade figure represented the most frequent use for computers identified by students. College Board data for 1996 (cited by Coley, et al.) indicated that 72 percent of SAT Program test takers reported previous experience in using computers for word processing. NAEP’s introduction of a computer delivered writing assessment would resonate with this changeover and could be coupled with introduction of automated or partially-automated essay scoring technologies (Burstein, Kukich, Wolff, Lu, and Chodorow, 1998).

While there is a dearth of research, evidence has begun to emerge that traditional pencil-and-paper assessment of writing may underestimate students’ writing ability compared to a computer administered writing assessment. Russell and Haney (1997) found that sixth- to eighth-grade students randomly assigned to a pencil-and-paper writing assessment or to a computer writing assessment performed significantly better on the latter using a 4-point wholistic scoring rubric. (Prior to scoring, the handwritten essays had been transferred onto a computer and randomly intermixed

with essays that were originally entered onto a computer so as to ensure that scoring was blind.) Russell and Haney (1997) conclude:

Increasingly, schools are encouraging students to use computers in their writing. As result, it is likely that increasing numbers of students are growing accustomed to writing on computers. Nevertheless, large scale assessments of writing, at state, national, and even international levels, are attempting to estimate students' writing skills by having them use pencil-and-paper. Our results, if generalizable, suggest that for students accustomed to writing on a computer for only a year or two, such estimates of student writing abilities based on responses written by hand may be substantial underestimates of their abilities to write when using a computer.

Thus, while NAEP has assessed writing in the past, the skill set in writing extended text with paper and pencil may gradually become obsolete, to be replaced by a skill set which makes use of technology to produce higher quality documents than would previously have been expected of students. To maintain its relevance to educational progress, NAEP must develop an assessment of the new construct of writing.

Potential for Assessing New Modes of Learning and Problem-Solving

In addition to developing measures of new problem solving skills tied to existing assessments, NAEP should monitor research progress on altogether new forms of assessment that pertain to 1) new domains of school learning, 2) on-line computerized diagnostic assessment of student learning, and 3) use of computers to simultaneously assess and tutor learning skills and subject matter content.

While computers and multi-media electronic technologies have been available in schools for more than four decades, they generally have not had a visible and dramatic impact on what students are expected to know and do in school. More recently, constructivist-based explorations have arisen of how technology might transform schooling in a manner attuned to the most lofty goals of the current education reform (Viadero, 1997).

Dede (1996) foregrounds the importance of collaboration through electronic media and characterizes these practices as "distributed learning." He notes three forms of distributed learning that are emerging (p. 3):

- Knowledge webs that complement teachers, texts, libraries, and archives as sources of information
- Interactions in virtual communities that complement face-to-face relationships in classrooms
- Immersive experiences in shared synthetic environments that extend learning-by-doing in real world settings

A number of efforts are underway nationally to explore radical transformations of education involving one or more of the foregoing themes (see Dede, 1996 for a collection of non-technical papers). One of the most cited and best researched programs of this sort, known as *Schools for Thought*, has been implemented by The Cognition and Technology Group at Vanderbilt (Williams, et al., 1998). *Schools for Thought* features four elements:

- Use of video, computer, and internet technology to bring information about complex problems into the classroom and to provide resources for problem solving
- Interaction among students and with others through video and interactive computer resources
- Use of electronic tools such as computer-based simulations to probe students' understanding and to assess students' learning while also providing feedback to students
- Use of technology to support collaborations among students, teachers, and community

Evaluation studies of the "Jasper Series," one of the components of *Schools for Thought*, have shown statistically significant cognitive outcomes and improvements in student attitudes towards math and science relative to control group students not exposed to the program (Goldman, Pellegrino, and Bransford, 1994).

One of the most common features among the constructivist approaches that utilize multi-media technologies and computers is student research projects. Investigators such as Guzdial (1998) assert that student research projects are effective for learning when they follow a process model for inquiry, guided by the following steps: initial statement of a problem and solution; decomposition into subproblems; composition of solution elements; debugging of a full solution; and final review of learning. In a similar manner, Krajcik, Soloway, Blumenfeld, and Marx (1998) emphasize students' active awareness of how they undertake a project and the role of collaboration among students in carrying out projects.

Windschitl (1998) sees three different areas of learning, centered on the World Wide Web, which complement the foregoing. These include: students' learning through interaction with the Web; students' inquiry processes and searching of the Web; and students' communication with others through the Internet. Windschitl argues that each of these areas is in direct need of research and that we lack a coherent knowledge base by which to evaluate the impact of technology on learning. He asserts that qualitative research will play an important role in this undertaking.

Some possible themes that are emerging from the cognitive science research approaches relevant to NAEP's consideration of new assessments include:

-
- Interpersonal cooperation in problem solving
 - Screening and prioritization of information obtained from the Internet that is needed to solve problems
 - Use of information-rich multi-media presentation of problems where students can review, replay, and search problem representations to support problem solving

Areas such as these (see O'Neil, Chung, and Brown, 1995; O'Neil, Klein, and Baker, 1997) will require much research before they can become meaningful targets for an assessment such as NAEP. This will take time, and NAEP needs to act now to prepare for assessments of the future.

One suggestion toward pursuing these goals is for NAEP to re-invent itself by adding "Multiple-Methods NAEP" to its existing core assessments (Pellegrino, et al., 1999). New assessment methods associated with the "Multiple-Methods NAEP" could be used to explore and extend the constructs assessed by NAEP through innovative uses of technology, supplementing and enriching the existing core NAEP. New assessment strategies might include integrating assessment with instruction, as well as conditional probes which track a student's search process to assess metacognition and problem-solving strategies.

Validity Issues for New Constructs

Three validity issues arise when NAEP introduces assessment of new constructs. These apply particularly in the case of new assessment areas involving technology-related skills, because the nature of what is to be learned for these constructs has not had the time to undergo thorough study.

1. Does the new construct have a coherent assessment framework?

Existing NAEP constructs are based on well-developed (although not unchanging) curricular content standards. Developing similar standards in emerging areas with which most of today's teachers are inexperienced presents an important challenge.

2. Does the new construct have valid performance expectations (achievement levels)?

The development of achievement levels, or performance standards, requires complex judgments; and methods for establishing these standards have been developed for traditional paper-and-pencil tests. Development of standards for new types of exercises, in areas in which performance expectations have not had time to develop, will require substantial efforts.

-
3. Is there a consensus that the new construct is central to American education, in the sense that information about student population performance on the construct is useful for educational policymaking?

The lack of formal curriculum frameworks for teaching elementary and secondary students the skills of cooperative problem-solving, complex problem simulations, and Internet use suggests the need for developing a national consensus that such areas should be included in the Nation's Report Card.

Technological Enhancements of Assessment Processes Other than Computer Presentation of Items and Recording of Responses

Although computerizing the interface between the assessment and individual student participants offers exciting opportunities and complex challenges, other technological innovations also warrant attention. These include 1) both non-computer enhancements to the interface with students and 2) computer enhancements of components of the assessment other than the actual presentation of items and recording of responses. In particular, the following six topics warrant study. The first four of these stand out as high priorities consistent with advice offered by Baker (1997) in her Capstone Report working paper on NAEP and technology and by others such as Bejar (1991). These four, it should be pointed out, do not require that a NAEP assessment be delivered on a computer.

1. *Assessment presentation accommodations.* For example, use of audio tapes to eliminate extraneous English reading factors for students with dyslexia or limited English reading proficiency does not require computers.
2. *Computer automated item/exercise generation.* This might be possible based on algorithms targeting desired item/exercise construct characteristics—Bejar (1991) and others have explored strategies to “recombine” existing items so as to create new items observing test specifications for item types based on analogical pattern matching rules. Singley and Bennett (1998) describe implementation of a Math Test Creation Assistant that utilizes cognitive schema theory to automatically generate alternative mathematics word problems that assess the same mathematical concept and performance skills.

-
3. *Computer scoring of extended open-ended responses.* Systematic classification schemes for extended open-ended responses, based on cognitive theories of conceptual development and problem solving, could be computerized for use in item scoring, score reporting, and score interpretation. Tatsuoka (1993), describes “rule space” models which automate interpretation of patterns of correct-incorrect responses to mathematics items in terms of cognitive models of the mathematical competencies of examinees. Bennett et al. (1997) evaluated a model for automated scoring of open-ended mathematical reasoning exercises.
 4. *Multimedia presentation of NAEP results.* The NAEP 1997 Arts Report Card for Eighth Graders (Persky, Sandene, and Askew, 1998) is available on the World Wide Web as well as in printed paper format. A CD-ROM version is available that includes the complete text of the report as well as many multimedia clips of student responses to assessment exercises. The multi-media format of this report represents a major advance in NCES capacity to utilize technology in dissemination of NAEP results.
 5. *Web delivered and wireless NAEP assessments.* Progress is being made in the delivery of assessments via the World Wide Web (McNichols, 1998). Use of programming languages, such as Java, permits uniform graphic representation of testing exercises across computer platforms. A centralized or regional computer server could be used to both administer new or existing NAEP exercises and store assessment data for scoring and further processing. Existing wireless Internet technologies could also be used to communicate assessment information to examinees’ computer stations by means of radio signals. Maintaining security and integrity of assessment information delivered via the Web and wireless technologies would be a central concern. However, current computer encryption technologies could ameliorate this concern.
 6. *Electronic transmission of traditional assessment materials.* Even without a computer interface with each student, NAEP might opt for secure electronic transmission of assessment materials to local schools for paper reproduction, or for teacher background surveys. Completed booklets might also be scanned at the school for digital transmission of student responses.

Redesign of NAEP Assessments as an Integrated Electronic Information System

A coherent approach to technological innovation will ultimately be more effective if built on an integrated plan. Rapid progress is being made by the Defense Department in its Armed Service Vocational Aptitude Battery (ASVAB) re-norming work and by

test developers such as Educational Testing Service in designing assessments as integrated electronic information systems. Bennett (1994) provides a helpful overview of key elements and their interaction in constructing an “electronic infrastructure” for new generations of computerized tests. He envisions a complex multi-organizational testing system connected and sharing information electronically. Each subsystem in the network performs essential tasks, much as is the case now with major testing programs, but the capacity to share and disseminate information is made immediate via electronic networking. Further, with planning, the testing system builds new capacities to interact with examinees and provides potential connections to information resources aiding examinees and schools in better preparing students.

The NAEP item and exercise development process might use technology strategically, and in a widespread manner, in devising certain assessments. Within ETS this strategy is exemplified by systems such as the “Test Creation Assistant” (Singley and Bennett, 1998). A more ambitious strategy would be to design and implement a system of integrated, computerized assessment development, delivery, and skill estimation tools based on cognitive models of competence in a problem solving domain. This approach is exemplified by the “Portal” project undertaken by Mislavy, Steinberg, Almond, and Johnson, 1998. A major strength of this approach is that it develops assessment task models based on expert judgments of competence in a problem solving domain. The task models lead to careful specification of the forms of evidence in students’ performance on tasks that support inferences about students’ competence in the target problem solving domain as represented by a separate student model.

While the complete redesign of the NAEP assessment as an integrated electronic information system would involve assessment delivery, as well as the scoring and score interpretation processes, it would also affect many other areas. These include development of the assessment frameworks, standards, and exercises; specification of the population and samples; collection of the data; analysis of student, school, and teacher surveys; preparation and dissemination of NAEP reports; dissemination of NAEP released exercises; and dissemination of public use data.

Important technology and infrastructure issues need examination across all of these possibilities, but these issues need to be addressed in the context of concrete implementation plans. For example, what kinds of electronic interfaces with NAEP public use data will be optimal in the future for educational policymakers and researchers? NAEP and the U.S. Department of Education have begun to make pre-compiled data and many summary reports and tables available on the World Wide Web, but they have yet to introduce on-line data analysis engines to produce summaries on demand. This example illustrates the promise and hazards of technology, and the complexity of problems to be faced in making NAEP a leader in use of technology. Thus, how will NAEP ensure that users understand the structure and limitations of NAEP data as users interact with on-line data directly? Availability of public use guidelines on the Internet may help, but what additional safeguards might be in order?

While the foregoing are important and fundamental potential enhancements of NAEP to its technology infrastructure for the long term, they call for careful research to address validity issues.

Validity Issues

Each of the proposed enhancements involves changes to NAEP, and it is important for the integrity of NAEP that the validity of NAEP not be reduced by those changes. The following exemplify the kinds of validity research questions that arise.

Media accommodations. Do the accommodations change the construct assessed?

Although audiotape has been used in NAEP for thirty years, their use as a method of accommodation has not been well-studied. The range of issues with respect to the validity of scores on accommodated tests must be addressed for this method of accommodation as well.

Automated item development. Are items equivalent?

Generating replicates of items that superficially assess the same skill domain does not provide any assurance that the replicates have the same difficulty or the same discriminability.

Internet background questionnaire administration. Are responses to computer/Internet administered surveys equivalent to paper-and-pencil surveys?

Teachers and principals may have a different tendency to omit items or to select particular response options when items are presented in a different format.

Automated scoring of open-ended responses. Do rubrics used in automated scoring assess the skills specified in the framework?

There may be a tendency to specify rubrics with simpler scoring rules, and there may be a tendency to misinterpret unusual correct responses.

Recommendations and Summary

Based on the review provided here, NAEP should place immediate priority on introducing computerized assessments that enhance its existing assessment priorities. Three high priority options stand out that deserve careful consideration for immediate research and implementation:

- Implementation of a linear computer administered NAEP in a target subject matter area such as mathematics or science
- Development and implementation of a computer administered writing assessment
- Continued introduction and evaluation of technology-based testing accommodations to include students with disabilities and students who are English learners.

Presumably, pursuit of these three alternatives would be benchmarked by a full field trial as a first initiative that would inform and guide the feasibility of a full implementation. Some of the issues for such a field trial are discussed later in this section.

Choosing among these priorities (or others that may arise) is inherently a policy decision informed by an analysis of cost-benefits and validity concerns. Successful implementation of any of the three priorities outlined above would accomplish the end of making NAEP a national leader in the application of technology to improved assessment. Each option would offer NAEP the opportunity to enhance its coverage of assessment constructs that are at the center of education reform.

While each option could be pursued without the other, it may be possible that implementation of a linear computer administered (CBT) NAEP could be yoked to eventual introduction of a CAT NAEP in the same assessment area. Implementation of a CBT NAEP first would give NAEP more time in which to develop a larger item pool for a CAT NAEP assessment and to address the research questions noted below.

Operational implementation of a computer-based assessment would need to grapple with the problem of providing examinees with adequate and timely access to computer testing stations. More important, introduction of computerized NAEP assessments must be preceded by a carefully planned set of research studies to investigate the scaling of scores on computerized versus non-computerized assessments, to evaluate the effects of computerized versus pencil-and-paper assessment administration on the performance of different groups of students, and to record the effects of prior computer experience on performance.

The introduction of a computerized NAEP writing assessment which assesses competence in using computer word processing tools is attractive because it resonates with the increasing importance of word processing in students' academic and non-academic lives. A computerized NAEP writing assessment could also be coupled with a research program to automate or semi-automate the scoring of writing samples. In addition, it could be coupled with research on new ways to assess complex reasoning skills such as those identified in the NAE Capstone Report (Glaser et al., 1997). A computerized writing assessment would entail some of the same logistical problems as a computerized assessment of an existing construct in providing students with appropriate access to computers. It would also entail validity research to ensure comparability of performances across key groups of students with differential exposure to computers; and it would require development of coherent content and performance standards for the new construct.

Pursuit of computer-based testing priorities needs to be coupled with careful examination of the fuller implications of technology for NAEP. Many important technological innovations are possible within the existing NAEP pencil-and-paper system, including computerized test development, scoring, and score reporting of NAEP results. As part of its deliberations on redesign, NAEP should begin planned initiatives to examine how technology should be infused throughout its infrastructure.

Design of a Computerized NAEP Field Trial

It is easy to recommend that NAEP should immediately undertake design and implementation of a field trial of a computer delivered linear assessment in reading, mathematics, or writing; however, the design of that field trial is complex and must be carefully developed to avoid failure at various points. The expectations for outcomes of an initial field trial should be fairly limited, as the field trial should be intended primarily to give NAEP an experience base appropriate for guiding a more intensive effort in implementing computerized assessments. Conceivably, more than one field trial would be helpful. An initial field trial would be a “proof of concept.” It would enable NAEP to assess the logistic capabilities needed for a computerized assessment and to evaluate the requisite test design and psychometric procedures for a computerized assessment. There are many decisions to be made and procedures to be developed for an initial field trial of a computerized assessment. The four listed below are particularly important.

Selection of an Assessment

The selection of an assessment for a computerized field trial will need to take into account at least two issues. One is the pay-off that development of a computerized assessment in the initial assessment area would have for implementation of computerized assessments in other areas. In this regard, initial implementation of a computerized assessment in reading or mathematics would be preferable as these assessments are more like each other and other NAEP assessments than would be the case for writing—the latter being limited to constructed responses to one or more writing prompts, while most other NAEP assessments involve a mixture of multiple-choice and short-answer questions.

It is possible that the initial field trial of a linear computer based assessment in reading or mathematics would be part of a process leading to a computer adaptive assessment. If one of these areas was selected for a field trial, a second factor would be the feasibility of rapidly developing an adequately sized item pool for a computerized adaptive test. It is not clear, however, how such a pool would need to grow in size in order to accommodate an adaptive test, even if we presume that most computerized versions of pencil-and-paper items would retain desirable psychometric characteristics.

Choice of a Sample

Because so many new variables are present, investment in a particular strategy should not be sufficiently large that needs for starting over would be disastrous for the program. Therefore, an initial computerized NAEP field trial should involve a limited number of students, perhaps no more than several hundred students. The sample should represent students from diverse backgrounds and demographic settings, but it also needs to be drawn as part of a formal stratified sampling plan in order to support meaningful comparisons.

Computer Delivery

Computers used in delivery of an assessment should be a standardized laptop or desktop platform with standardized monitors and keyboards. Nevertheless, variations should be anticipated and steps should be taken to learn from the impact of variations on assessment outcomes.

Internet Delivery

Use of the Internet to deliver assessments would be valuable to investigate early on. Internet presentation of assessments would ensure considerable standardization in the appearance of assessment items and response formats akin to administration via an assessment stored on local computer or network servers. Such a strategy would improve access of students to assessments on a more flexible schedule and would enable immediate centralized recording of performance data at a remote site.

Computer Experience and Accommodations

Implementation of a field trial should be coupled with a survey of participants' prior exposure to computers and their reactions to participation in a computerized NAEP assessment. This information will be valuable in exploring whether a computerized NAEP assessment will be accepted by students, and it will provide initial evidence of potential biases in assessment performance tied to student characteristics. In addition, the field trial could explore the use of selected assessment accommodations for students with disabilities and limited English familiarity. These accommodations, as appropriate to a population, might include increasing length of assessment time, use of primary language dictionaries for English language terms, etc.

A clear momentum has evolved for NAEP to undertake careful steps leading to systematic integration of technology into its activities. The work cited in this report suggests that NAEP needs to act in a timely manner in implementing a field trial of a computer delivered NAEP. Furthermore, NAEP needs to couple this effort with a continued effort to examine ways that technology can impact its functioning across all of its systems. Finally, NAEP needs to carry out research on validity issues to ensure that technological innovations do not threaten the validity of the Nation's Report Card.

References

- Baker, E. (1997). Readyng the National Assessment of Educational Progress to meet the future. In R. Glaser, R. Linn, and G. Bohrnstedt (Eds.), *Assessment in transition: Monitoring the nation's educational progress, background studies* (pp. 123–150). Stanford, CA: National Academy of Education.
- Bejar, I. (1991). *A generative approach to psychological and educational measurement*. RR-91-20. Princeton: Educational Testing Service.
- Bellamy, R.K.E. (1996). Designing educational technology: Computer-mediated change. In B. Nardi. (Ed.), *Context and consciousness. Activity theory and human-computer interaction*. Cambridge, MA: MIT Press.
- Bennett, R.E. (1994). *An electronic infrastructure for a future generation of tests*. Research Report RR-94-61. Princeton: Educational Testing Service.
- Bennett, R., Steffen, M., Singley, M., Morley, M., and Jacquemin, D. (1997). Evaluating an automatically scoreable, open-ended response type for measuring mathematical reasoning in computer-adaptive tests. *Journal of Educational Measurement*, 34, 163–177.
- Bennett, R.E., Goodman, M., Hessinger, J., Ligget, J., Marshall, G., Kahn, H., Zack, J. (In Press). Using multimedia in large-scale computer-based testing programs. *Computers in Human Behavior*.
- Burstein, J., Kukich, K., Wolff, S., Lu, C., and Chodorow, M. (1998). *Computer analysis of essays*. Paper presented at the Annual Meeting of the National Council on Measurement and Education.
- Carnavale, A., and Rose, S. (1998). *Education for what? The new office economy. executive summary*. Princeton: Educational Testing Service.
- Cole, N. (December 1997). *Why computerize assessment? An Issue Brief for the Colleagues of Educational Testing Service*. Vol. 1
- Coley, R., Cradler, J., and Engel, P. (May 1997). *Computers and classrooms: The status of technology in U.S. schools*. Policy Information Report. Princeton, NJ: Educational Testing Service.
- Dede, C. (1996). The evolution of distance education: Emerging technologies and distributed learning. *American Journal of Distance Education*, 10, (2) 4–36.
- Feldman, A. (1997). Digital kids. There's no turning back. *Hands On!* 20 (2).
- Forsyth, R., Hambleton, R.K., Linn, R., Mislevy, R., and Yen, W. *Design/feasibility team: Report to the National Governing Board*. July 1, 1996.

-
- Gilster, P. (1997). *Digital literacy*. New York: John Wiley.
- Glaser, R., Linn, R., and Bohrnstedt, G. (Eds.), (1997). *Assessment in transition: Monitoring the nation's educational progress*. Stanford, CA: National Academy of Education.
- Goldman, S., Pellegrino, J., and Bransford, J. (1994). Assessing programs that invite thinking. In E. Baker and H. O'Neil (Eds.), *Technology assessment in education and training*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Green, B. (1998). Panel presentation on "Trends in Computerized Testing-A synthesis of This Year's Sessions" Annual NCME Meeting. San Diego.
- Guzdial, M. (1998). Technological support for project-based learning, In C. Dede (Ed.), *ASCD 1998 year book. Learning with technology*. Alexandria, VA: Association for Supervision and Curriculum Development
- Heaviside, S., Riggins, T., and Farris, E. (1997). *Advanced telecommunications in U.S. public elementary and secondary schools, fall 1996*. NCES Statistics in Brief (NCES 97-944). Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Krajcik, J., Soloway, E., Blumenfeld, P., and Marx, R. (1998). Scaffolded technology tools to promote teaching and learning in science. In C. Dede (Ed.), *ASCD 1998 year book. Learning with technology*. Alexandria, VA: Association for Supervision and Curriculum Development
- Kulik, J. (1994). Meta-analytic studies of findings on computer-based instruction. In E. Baker and H. O'Neil, Jr. (Eds.), *Technology assessment in education and training*. Hillsdale, NJ: Erlbaum.
- McNichols, C. (1998) Web-based testing with active server pages. *Web Techniques* 3, (12), 69-75.
- Mislevy, R., Steinberg, L., Breyer, F., Almond, R., and Johnson, L. (1998). *A cognitive task analysis, with implications for designing a simulation-based performance assessment*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.
- Nardi, B. (Ed.), (1996). *Context and consciousness. Activity theory and human-computer interaction*. Cambridge, MA: MIT Press.
- Persky, H.R., Sandene, B.A., and Askew, J.M. (1998). *The NAEP 1997 arts report card*. (NCES 1999-486). Washington, DC: National Center for Education Statistics.
- Johnson, E.G., Lazar, S., and O'Sullivan, C.Y. (1997). *NAEP reconfigured: An integrated redesign of the National Assessment of Educational Progress*. (NCES

-
- 97-31). Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Norman, D. (1988). *The psychology of everyday things*. New York: Basic Books.
- Olson, J. and Goldstein, A. (1997). *The inclusion of students with disabilities and limited english proficient students in large-scale assessments. A summary of recent progress*. (NCES 97-482). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.
- O'Neil, H. Jr., Klein, D., and Baker, E. (1997). *A cognitive demands analysis of innovative technologies*. CSE Technical Report 460, Los Angeles: UCLA, CRESST.
- O'Neil, H. Jr., Chung, G., and Brown, R. (1995). *Measurement of teamwork processes using computer simulation*. CSE Technical Report 399. Los Angeles: UCLA, CRESST.
- Pellegrino, J., Jones, L., and Mitchell, K. (Eds.), (1999). *Grading the Nation's report card: Evaluating NAEP and transforming the assessment of educational progress*. Washington DC: National Academy Press.
- President's Committee of Advisors on Science and Technology (March 1997). *Report to the President on the use of technology to strengthen K-12 education in the United States*. Panel on Educational Technology.
- Russell, M. and Haney, W. (1997). Testing writing on computers: An experiment comparing student performance on tests conducted via computer and via paper-and-pencil. *Education Policy Analysis Archives*, 5 (3).
- Sands, W., Waters, B., and McBride, J. (1997). *Computer adaptive testing: From inquiry to operation*. Washington, DC: American Psychological Association.
- Segall, D. (1998). Panel presentation on "Trends in Computerized Testing-A Synthesis of This Year's Sessions" Annual NCME Meeting. San Diego.
- Singley, M., and Bennett, R. (1998). *Item generation and beyond: Applications of schema theory to mathematics assessment*. Paper presented at the "Generating Items for Cognitive Tests: Theory and Practice Conference", Princeton, Educational Testing Service.
- Tatsuoka, K. (1993). Item construction and psychometric models appropriate for constructed responses. In R. Bennett and W. Ward (Eds.), *Constructed response versus choice in cognitive measurement*. Hillsdale, NJ: Lawrence Erlbaum, 107-133.
- U.S. Bureau of the Census, 1993 Current Population Survey, Educational and Social Stratification Branch.

-
- U.S. Department of Commerce, National Telecommunications and Information Administration. (1999). *Falling through the net: Defining the digital divide*. A Report on the Telecommunications and Information Gap in America. Washington D.C.
- Viadero, D. (Nov. 10, 1997). A tool for learning. *Education Week. Technology Counts. Schools and Reform in the Information Age*.
- Williams, S., Burgess, K., Bray, M., Bransford, J., Goldman, S., and The Cognition and Technology Group at Vanderbilt (1998). Technology and learning in schools for thought classrooms. In C. Dede (Ed.), *ASCD 1998 year book. Learning with technology*. Alexandria, VA: Association for Supervision and Curriculum Development, pp. 97-119.
- Williams, V. and Sweeny, S. (1997). *A prototype audio-CAT for the national assessment*. NAEP Redesign CFDA No. 84.902A. Research Triangle Park, NC: Research Triangle Institute.
- Windschitl, M. (1998). The WWW and classroom research: What paths should we take? *Educational Researcher*, 28 (1), pp. 28–33.

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) 502-7363 (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

No.	Title	NCES contact
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
2003-06	NAEP Validity Studies: The Validity of Oral Accommodation in Testing	Patricia Dabbs
2003-07	NAEP Validity Studies: An Agenda for NAEP Validity Research	Patricia Dabbs
2003-08	NAEP Validity Studies: Improving the Information Value of Performance Items in Large Scale Assessments	Patricia Dabbs
2003-09	NAEP Validity Studies: Optimizing State NAEP: Issues and Possible Improvements	Patricia Dabbs
2003-10	A Content Comparison of the NAEP and PIRLS Fourth-Grade Reading Assessments	Marilyn Binkley
2003-11	NAEP Validity Studies: Reporting the Results of the National Assessment of Educational Progress	Patricia Dabbs
2003-12	NAEP Validity Studies: An Investigation of Why Students Do Not Respond to Questions	Patricia Dabbs
2003-13	NAEP Validity Studies: A Study of Equating in NAEP	Patricia Dabbs
2003-14	NAEP Validity Studies: Feasibility Studies of Two-Stage Testing in Large-Scale Educational Assessment: Implications for NAEP	Patricia Dabbs
2003-15	NAEP Validity Studies: Computer Use and Its Relation to Academic Achievement in Mathematics, Reading, and Writing	Patricia Dabbs
2003-16	NAEP Validity Studies: Implications of Electronic Technology for the NAEP Assessment	Patricia Dabbs
2003-17	NAEP Validity Studies: The Effects of Finite Sampling on State Assessment Sample Requirements	Patricia Dabbs
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
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

No.	Title	NCES contact
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
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

No.	Title	NCES contact
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
Progress in International Reading Literacy Study (PIRLS)		
2003-05	PIRLS-IEA Reading Literacy Framework: Comparative Analysis of the 1991 IEA Reading Study and the Progress in International Reading Literacy Study	Laurence Ogle
2003-10	A Content Comparison of the NAEP and PIRLS Fourth-Grade Reading Assessments	Marilyn Binkley
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
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

No.	Title	NCES contact
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
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

No.	Title	NCES contact
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
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
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		

No.	Title	NCES contact
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		
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		

No.	Title	NCES contact
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
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

No.	Title	NCES contact
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
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

No.	Title	NCES contact
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
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

No.	Title	NCES contact
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
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

No.	Title	NCES contact
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
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

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
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
1999-06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson