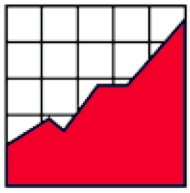


NATIONAL CENTER ON EDUCATIONAL OUTCOMES

Technical Report 2

Inclusion of Students with Disabilities in National and State Data Collection Programs

National Center on Educational Outcomes

University of Minnesota

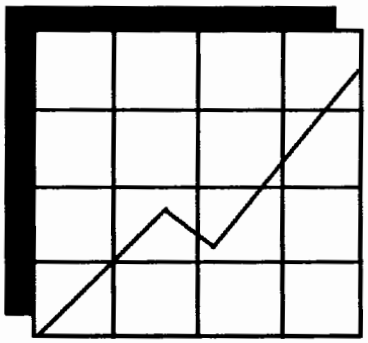
in collaboration with

St. Cloud State University

and

National Association of State Directors of Special Education

Technical Report 2



Inclusion of Students with Disabilities in National and State Data Collection Programs

Kevin S. McGrew
St. Cloud State University

Martha L. Thurlow, James G. Shriner and Amy N. Spiegel
University of Minnesota

National Center on Educational Outcomes
UNIVERSITY OF MINNESOTA

March, 1992

The National Center on Educational Outcomes (NCEO) was established in October, 1990 to work with state departments of education, national policy-making groups, and others to facilitate and enrich the development and use of indicators of educational outcomes for students with disabilities. It is believed that responsible use of such indicators will enable students with disabilities to achieve better results from their educational experiences. The Center represents a collaborative effort of the University of Minnesota, the National Association of State Directors of Special Education, and St. Cloud State University.

The Center is supported through a Cooperative Agreement (H159C00004) with the U.S. Department of Education, Office of Special Education Programs. Opinions or points of view do not necessarily represent those of the U.S. Department of Education or Offices within it.

NCEO Core Staff:

Robert H. Bruininks
Stanley L. Deno
Cheri Gilman
Kevin S. McGrew
James G. Shriner, **State Survey Interview Coordinator**
Martha L. Thurlow, **Assistant Director**
James E. Ysseldyke, **Director**

NASDSE Contributing Staff:

Eileen Ahearn
Linda Lewis, **Associate Director**

Report Design:

Jill T. Jackson

Additional copies of this report can be obtained for $10.00 from:

NCEO Publications Office
350 Elliott Hall
75 E. River Road
Minneapolis, MN 55455

Table of Contents

List of Tables

List of Figures

Inclusion of Students with Disabilities in National and State Data Collection Programs

Executive Summary

Calls for reform in American education during the past decade have resulted in raised expectations, attempts to develop uniform and "world class" standards, and increased emphasis on school accountability and the measurement of educational outcomes. Reform initiatives at both national and state levels are focusing more frequently on outcomes and quantifiable data. The measurement of educational indicators is playing a central role in the current wave of educational reform as various groups seek to produce policy-relevant information on the educational performance and status of children in our nation's schools. With increasing frequency, the data needed to monitor and evaluate these reform initiatives are being drawn from national and state data bases.

This report summarizes activities of the National Center on Educational Outcomes (NCEO) that are directed at determining the extent to which individuals with disabilities are involved in national and state data collection programs that are playing a pivotal role in the current measurement-driven reform movement. Inclusion of individuals with disabilities in the major national data collection programs was evaluated by identifying 30 different data collection programs that are receiving significant attention in the current reform initiatives. For this report, 9 of the 30 targeted national data collection programs were reviewed. The extent to which students with disabilities are included in state outcomes assessment activities was evaluated through a survey of state directors of special education. Together these activities produced the following results and conclusions:

- Most existing national and state data collection programs exclude large portions of the student population with disabilities.

- At the national level, it is estimated that approximately 40% to 50% of school-age students with disabilities are excluded from prominent national data collection programs.

- Exclusion appears most prevalent in data collection programs that require students to complete surveys or tests independently. Data collection programs that are based solely on survey research methods typically exclude few individuals with disabilities when they use third party informants.

- State-level data documenting the extent of exclusion of students with disabilities is, with a few exceptions, largely unavailable at this time.

- Exclusion of students with disabilities from national and state data collection programs occurs at many different points, from the development of assessment instruments, to the reporting of results.

- Exclusion criteria typically are implemented by local school personnel who understandably are most concerned about their immediate setting and who may not appreciate the potential usefulness of such information for school improvement and the development of educational policy.

- The use of different exclusion guidelines across data sets can cause problems in comparing results obtained from different data collection programs. Reasons typically given for exclusion of students with disabilities range from concerns about providing proper accommodations (e.g., in test administration mode, in response mode, in flexible time or setting) to concerns about the potential aversiveness of the assessment situation for the student.

- A sizable portion of excluded students should not have been excluded from data collection programs, and could readily participate (some with testing accommodations, others without) in such data collection programs.

- The ability to extract useful national and state policy-relevant information on the outcomes of students with disabilities from national and state data collection programs is seriously hampered by the extensive exclusion of portions of this population. The exclusion of students with disabilities results in significant problems in obtaining representative samples.

This, in turn, creates difficulties in estimating national and state level statistics (e.g., dropout rates) and in completing accurate policy studies.

In conclusion, large numbers of students with disabilities are currently treated as "outliers" in our national and state data collection programs. This categorical exclusion of students with disabilities perpetuates the myth of inherent differences. Given the magnitude of federal and state support for educational programs for students with disabilities, support that reflects the valuing of this population in our society, it is time that this implied value is matched by the commitment of resources to address the numerous political and technical hurdles that must be overcome in order for these students to participate more fully in our national and state data collection programs. Current and future activities of the NCEO are being focused in this direction.

A number of recommendations for increasing the participation of students with disabilities in national and state data collection programs are offered. These recommendations focus on:

- Developing broader and more uniform definitions of sample eligibility
- Increasing adherence to inclusion guidelines, particularly the "if in doubt, include" component of inclusion guidelines
- Developing sampling frames for data collection programs that are more inclusive
- Routinely conducting follow-up studies of ineligible students as part of data collection programs
- Increasing the partial participation of students with disabilities during instrument development
- Researching and developing assessment modifications, accommodations, or alternatives that allow more students with disabilities to participate in large scale data collection programs

Inclusion of Students with Disabilities in National and State Data Collection Programs

Assessment is a multimillion dollar enterprise in America today. Corporations assess individuals to identify who should be employed or promoted. The military assesses new personnel to determine the types of jobs to assign to them. Schools assess children and youth to ascertain the extent to which they are progressing in school and to identify who may have unique educational needs. Federal and state governments use information collected by means of assessment to describe the status of students and schools in the nation and in each state. This information is used for accountability purposes, and for setting educational policy.

Within the past decade, concern has been expressed in reports about education in America. Since the publication of A Nation at Risk (National Commission on Excellence in Education, 1983), many reports have lamented the failure of our schools to adequately educate students. While data were used to demonstrate education's lack of success (e.g., Kirsch & Jungeblut, 1986; NAEP, 1985; National Governors' Association, 1986; Office of Educational Research and Improvement, 1988), there was still dissatisfaction expressed about the adequacy of available educational data.

During the 1980s, a wave of state legislation was passed on school report cards, merit schools, and interstate achievement comparisons (Kirst, 1990). To some extent, it was believed that better educational data would pressure schools to improve, which in turn would lead to better student performance on tests, and even better preparation for students as they became members of the American workforce and competitors in the international marketplace. The validity of this logic has been questioned by many (e.g., Darling-Hammond, 1991; Oakes, 1991).

Still, the need for data to describe the status of students is generally accepted. There are many who argue that the data collection methods need to be improved to reflect assessment of higher order thinking or to incorporate authentic assessment procedures, but most still would agree on the need for the data. Even a cursory review of the current educational reform literature leads to the conclusion that there is a hunger for policy-relevant information on the performance of students in our educational system.

In response to a legislative mandate (Hawkins-Stafford Education Amendments of 1988: PL 100-297), the National Forum on Educational Statistics was formed to examine the extent to which overlap or gaps existed in data collected by federal or state entities. This effort was initiated specifically to influence the production of a national cooperative statistics system of useful and comparable data. In a report entitled A Guide to Improving the National Education Data System (NESAC, 1990), one of the Forum's committees presented 36 recommendations for improving the national education data system. These recommendations covered the data domains of (a) background/demographics, (b) education resources, (c) school processes, and (d) student outcomes.

Another related effort occurring about the same time, and also sponsored by the Hawkins-Stafford Education Amendments, was the formation of a Special Study Panel on Educational Indicators. The purpose of this group was to identify educational indicators and organize them into logical groupings. In their report (Special Study Panel on Education Indicators, 1991), entitled Education Counts, it was argued that the success of reform efforts depends upon the development of a "comprehensive education indicators information system capable of monitoring the health of the enterprise, identifying problems, and illuminating the road ahead" (p. 6). Further, in arguing that educational indicators should define the educational agenda rather that reflect an educational agenda (i.e., "An indicator system organized around today's goals cannot respond to tomorrow's"; p. 10), the Panel proposed six issue areas in which educational indicators are needed: (1) learner outcomes, (2) quality of educational institutions, (3) readiness for school, (4) societal support for learning, (5) education and economic productivity, and (6) equity.

A third major effort was the National Education Goals Panel, formed to create a mechanism to monitor the nation's and states' progress toward meeting the Bush administration's six national education goals (e.g., school readiness, 90% high school graduation rate, etc.). Six resource groups

were established to identify the kinds of data needed to track progress on each of the goals, what kinds of data were available, and to produce annual reports containing data relevant to each goal. In their first annual report (National Education Goals Panel, 1991b), the Panel indicated that it had brought together "the most recent quality data available on how well the nation and individual states are doing in achieving the National Education Goals" (p. 191). However, it also noted that "while the report is relatively comprehensive in some areas, there are also many gaps" (p. 191). The gaps identified in the report consist primarily of lists of indicators needed for each goal, such as data on how prepared children are to learn when they enter school (physically, emotionally, socially, and intellectually), comparable state-level data on high school completion and dropout rates, and so on.

The NESAC report, Education Counts, and the Goals Report all highlighted the need for better data on students, particularly on specific subsets of students. Students considered to be disadvantaged and students from specific cultures (e.g., Hispanic, African American, Native American) were specifically targeted for special attention. Only the NESAC report, however, highlighted the need to disaggregate data for students with disabilities who are served by the nation's special education system. Students with a wide array of disabilities fit within this population, including those with learning disabilities, emotional disabilities, and speech and language impairments, those with sensory disabilities such as hearing impairments and visual impairments, and those with multiple and more severe disabilities, typically involving significant mental impairments. Given that over 4.5 million school-age youngsters receive some form of special education services, services that are provided at significant expense to our educational system, it is imperative that we examine how these students are performing.

There are several sources of information that can be used to evaluate the performance of students with disabilities. First, our nation has a wide array of data bases with information that is relevant to individuals with disabilities. These data bases cover such areas as education, housing, employment, and health. Individual states also maintain data bases, usually within their Departments of Education. When the Council of Chief State School Officers (CCSSO, 1990) examined state education indicators, it again stressed the importance of identifying and collecting data that are technically, financially, and educationally feasible, and of improving the comparability of education data collected across states. In discussing the future, the report stated:

> The years ahead will be difficult for those trying to develop better data, but the results of their efforts will be critical if we are to do a better job of monitoring our educational system. (p. 6)

The CCSSO report does not address the need for data on specific subsets of students.

The purpose of this report is to examine national and state data bases to determine the extent to which students with disabilities are included in these collection systems. We first examine national data bases, then shift to state educational data bases. These data collection systems are examined in terms of inclusion rates, criteria for exclusion of a student from a data base, and special accommodations that are made to ensure inclusion of students. This is followed by a summary of the critical concerns that are generated as a result of the analysis of the inclusion of students with disabilities in national and state data collection systems. Several recommendations for dealing with the exclusion of students with disabilities in national and state data collection programs conclude the report. More comprehensive solutions and guidelines will be addressed in a subsequent report.

National Education Data

The United States has long recognized the value of large-scale federally funded studies to assess student progress. Data collection programs such as the National Assessment of Educational Progress (NAEP - the "Nation's Report Card"), the National Longitudinal Study (NLS), High School and Beyond (HSB), and the recent National Education Longitudinal Study (NELS) are some of the more recognizable efforts.

Unfortunately, it appears from our analysis that the sampling plans for these studies have not allowed for adequate analyses of subgroups of students with disabilities, a situation that has resulted in students with disabilities being considered "outliers" in our national education data collection system (Allen, 1989). Until the recent funding by the Office of Special Education Programs of the National Longitudinal Transition Study of Special Education Students (NLTS), there has been relatively little national level policy-relevant knowledge produced about the educational status and performance of students with disabilities. Given the magnitude of federally-mandated educational programs for students with disabilities, "the lack of adequately designed national studies of handicapped youth is particularly alarming" (Allen, 1989, p. 469).

The first step in our analysis of national data sets was to identify a preliminary list of "target" data sets for critical review. Two current national initiatives were considered in selecting data bases for further study: (1) the national education goals, and (2) the educational indicator movement.

National educational goals. Probably no single force has served to spearhead the current wave of education reform as have the activities that surround the National Education Goals Panel (NEGP) and the new national education strategy of the Bush administration, America 2000. If students with disabilities are to be included in current and future discussions of national education goals, it is important to assess the degree to which this population is included in the measurement strategies that drive these reform initiatives. This is important because it may be, as suggested by the Special Study Panel on Education Indicators (1991), that "we begin to value only what we can measure" (p. 5). There is the related concern that we may begin to value only who we can measure. "To categorically exclude students with disabilities perpetuates the myth of inherent differences. It makes students with handicaps non-students and perhaps non-people" (NASDSE, 1988, p. 10).

NEGP formed panels to develop recommendations for identifying potential indicators and strategies for assessing progress toward the six national goals. In its report entitled Measuring Progress Toward the National Education Goals: Potential Indicators and Measurement Strategies (National Education Goals Panel, 1991a), the Panel frequently recommended the use of indicators from such data sets as NELS and NAEP. Although many of the recommended data sets and indicators focus on context, input, and process indicators, some do include outcome indicators. Listings of the data sets are presented by goal in Table 1.

Educational indicator movement. It is clear that the current educational reform initiatives have produced a flurry of activity focused on the identification of educational indicators to measure progress toward goals. The development of educational indicator systems has become a big "business" in the United States (Odden, 1990), with nearly all national or state level groups and agencies becoming involved in these activities (Smith, 1988). As a result, many data sets included in the national education data system have been examined for possible indicators with which to monitor the effectiveness of reform activities.

Among the major organizations that are examining existing national data bases to identify indicators to measure the progress of children and youth are the Council of Chief State School Officers, the National Governors' Association (NGA), and Joining Forces (a coalition of the American Public Welfare Association and the Council of Chief State School Officers), to name but a few. These groups have all turned to national data bases for indicators to monitor progress during the current wave of reform. Most have produced reports on their recommendations. These reports were reviewed to identify those national data collection programs that are receiving significant attention (e.g., NAEP, NELS) in the educational indicator and reform literature.

Summary. Based on a review of the national goals activities and education indicators movement, 30 national data collection programs were identified. This list of national data sets is presented in Table 2.

Nine of the data sets in Table 2 have been reviewed for this report. We sampled these

Table 1: Data Sets Identified by the National Education Goals Panel As Possible Sources of Information for Six National Goals

Data Source	Sponsor	Indicator
Goal 1: School Readiness		
Nationwide Food Consumption Survey, Continuing Survey of Food Intakes	Dept of Agriculture	Nutritional status
Division of Vital Statistics	Dept of Health and Human Services	Birthweight; prenatal care timing; prenatal status
National Health Interview Survey	Dept of Health & Human Services, Nat'l Center for Health Statistics	Children's access to health care
National Household Education Survey	Dept of Education, National Center for Education Statistics	Home activities and attitudes
Current Population Survey, School Enrollment Supplement;	Dept of Commerce, Census Bureau	Preschool program participation
Head Start Program Information Report Questionnaire	Dept of Education	Preschool program quality
Goal 2: Dropouts		
Current Population Survey	Dept of Commerce, Census Bureau	Proportion of 19, 20, 24, 25 yr olds with high school credential
High School & Beyond; National Education Longitudinal Study	Dept of Education, National Center for Education Statistics	Drop out precursors; dropout rates; dropouts who complete high school
General Educational Development Testing Service	American Council on Education	Dropouts who complete high school
Common Core of Data	Dept of Education, National Center for Education Statistics	State-by-state school completion data
Goal 3: Achievement/Citizenship		
National Assessment of Educational Progress (NAEP)	Dept of Education, National Center for Education Statistics	Student achievement
Advanced Placement Tests	The College Board	Number of Advanced Placewment Tests; scores earned
High School Transcript Studies	Dept of Education, National Center for Education Statistics	High school course enrollments
International Evaluation of Educational Achievement	Dept of Education, National Center for Education Statistics	Internat'l achievement comparisons
State Provided Data		State reports on student achievement
Goal 4: Science and Math		
National Assessment of Educational Progress (NAEP)	Dept of Education, National Center for Education Statistics	Instructional practices; student achievement, attitudes; teacher characteristics
International Evaluation of Educational Achievement	Dept of Education, National Center for Education Statistics	Achievement; instructional practices; teacher characteristics
International Assessment of Educational Progress	Dept of Education, National Center for Education Statistics	Achievement
National Survey of Science and Mathematics Education	Dept of Education, National Center for Education Statistics	Teacher and principal attitudes

Table 1 (continued)

Data Source	Sponsor	Indicator
Schools and Staffing Survey	Dept of Education, National Center for Education Statistics	Instructional practices and teacher characteristics
National Education Longitudinal Survey	Dept of Education, National Center for Education Statistics	Instructional practices and teacher characteristics
Longitudinal Study of American Youth	National Science Foundation	Student and parent attitudes
Integrated Postsecondary Education Data System	Dept of Education, National Center for Education Statistics	Number of postsecondary graduates
Survey of Graduate Students and Post Doctorates	National Science Foundation	Number of graduate student enrollees
Goal 5: Literacy and Lifelong Learning		
National Assessment of Educational Progress (NAEP)	Dept of Education, National Center for Education Statistics	Literacy
Special Populations' Literacy Skills	Dept of Labor	Literacy
Armed Service Vocational Aptitude Battery	Dept of Defense	Vocational aptitudes
Workforce Participation Survey	Dept of Labor	Skills preparation
National Household Education Survey (NHES)	Dept of Education, National Center for Education Statistics	Participation in adult learning programs
Integrated Postsecondary Education Data System	Dept of Education, National Center for Education Statistics	Enrollment and graduation from postsecondary training programs
National Longitudinal Study; High School & Beyond; National Education Longitudinal Study	Dept of Education, National Center for Education Statistics	Number of associate and bachelors degree recipients
National Adult Literacy Survey	Dept of Education, National Center for Education Statistics	Literacy
Goal 6: Safe, Disciplined, Drug-Free Schools		
Monitoring the Future	Dept of Health and Human Services	Drug use; crime victimization; attitude toward drugs
National Adolescent School Health Survey	Dept of Health and Human Services	Drug use
National Crime Survey, School Crime Supplement	Dept of Justice	Perceptions of safety in school; weapons in school
National Education Longitudinal Study	Dept of Education, National Center for Education Statistics	Student & teacher views of noise, disruption on learning
Fast Response Survey System	Dept of Education, National Center for Education Statistics	Crime victimization; student & teacher views of noise, disruption on learning
Youth Risk Behavior Surveillance System	Dept of Health & Human Services, Center for Disease Control	Drug and tobacco use, weapon ownership and use

Table 2: Preliminary List of NCEO Targeted National Data Collection Programs

DEPARTMENT OF EDUCATION

- High School and Beyond
- Transcript Study
- Integrated Postsecondary Education Data System
- International Assessment of Educational Progress
- National Adult Literacy Survey
- National Assessment of Educational Progress: 1988, 1990
- National Assessment of Educational Progress: Trial State Assessment
- National Education Longitudinal Study
- National Longitudinal Transition Study of Special Education Students
- National Household Education Survey
- International Evaluation of Educational Achievement
- Young Adult Literacy Survey
- Beginning Postsecondary Students
- Baccalaureute and Beyond

DEPARTMENT OF COMMERCE

- Current Population Survey
- Survey of Income and Program Participation

DEPARTMENT OF LABOR

- Workforce Participation Survey
- Workplace Literacy Assessment

DEPARTMENT OF JUSTICE

- National Crime Survey

DEPARTMENT OF HEALTH AND HUMAN SERVICES

- National Health Interview
- National Health and Nutrition Examination Survey
- National Survey of Personal Health Practices and Consequences
- National Survey of Family Growth
- National Adolescent School Health Survey
- Youth Risk Behavior Surveillance System
- National Household Survey of Drug Abuse
- Monitoring the Future

NATIONAL SCIENCE FOUNDATION

- Survey of Graduate Students and Post Doctorates
- Longitudinal Study of American Youth

AMERICAN COUNCIL OF EDUCATION

- General Education Development Testing

THE COLLEGE BOARD

- Advanced Placement Tests

data sets from several sponsoring agencies so that we could examine variability among sponsoring agencies as well as among data sets. The specific data sets (and their sponsoring agencies) that were selected for review in this report are:

- National Adult Literacy Survey (Department of Education) — NALS
- National Assessment of Educational Progress: 1988 (Department of Education) — NAEP:88
- National Assessment of Educational Progress: 1990 (Department of Education) — NAEP:90
- National Assessment of Educational Progress: Trial State Assessment Program (Department of Education) — NAEP:Trial State
- National Education Longitudinal Study of 1988 (Department of Education) — NELS:88
- Current Population Survey, March Supplement (Department of Commerce) — CPS
- National Health Interview Survey (Department of Health and Human Services) — NHIS
- National Health and Nutrition Examination Survey, National Health Epidemiological Follow-up Study (Department of Health and Human Services) — NHANES, NHEFS
- Longitudinal Study of American Youth (National Science Foundation) — LSAY

These data sets are described in Table 3.

For each of the nine national data collection programs, all relevant reports and manuals (methodological and technical) were obtained from the sponsoring agency and reviewed as part of a detailed "disability sensitivity review" (McGrew, Spiegel, Thurlow, Ysseldyke, Bruininks, Deno, & Shriner, 1991a). The relevant documentation for each targeted data set was reviewed to extract the following information:

1. Descriptive information - Title, collection cycle, research design, sponsor, contact, and general purpose of the data collection program.
2. Source and method of data collection - Information on who (student, parent, teacher, administrator) provides the data, and how the data are collected (questionnaire, test, records review, etc.)
3. Sample analysis - Description of sampling design, with particular attention to the use of disability related exclusionary procedures, operational definition of disability categories, and the disability characteristics of the final sample.
4. Description of indicators - Description of the general domains of variables contained in the data set.

For the current report, the information of interest was obtained as part of the "sample analysis" component. This component provided information on: (a) the disability-related inclusion and exclusion guidelines used in the data collection program, (b) who makes the inclusion and exclusion decisions during data collection, and (c) the rates of inclusion or exclusion of students with disabilities in the final sample.

Criteria for Exclusion

The information presented in Table 4 is a summary of the different disability-related exclusion guidelines and procedures used by the nine selected national data collection programs. A number of conclusions can be drawn from the information presented in Table 4.

First, there is considerable variability in the extent to which exclusion guidelines are operationalized and reported in the different data collection programs. However, some of these differences are understandable based on the different assessment methods that are used. In data collection programs that only require the completion of an interview protocol (viz., NHIS, NHEFS, CPS), exclusion of

Table 3: Descriptions of Nine Data Sets

National Adult Literacy Survey (Department of Education) -- NALS

A nationally representative cross-sectional study designed to collect information on the types and levels of literacy skills adults living in the United States possess and how these skills are distributed across major subgroups. This study is assessing the prose, document, and quantitative literacy of approximately 15,000 adults (16 to 64 years of age) in 1992.

National Assessment of Educational Progress: 1988 (Department of Education) -- NAEP:88

National Assessment of Educational Progress: 1990 (Department of Education) -- NAEP:90

National Assessment of Educational Progress: Trial State Assessment Program (Department of Education) -- NAEP:Trial State

NAEP is a nationally representative cross-sectional study designed to monitor the knowledge, skills, understanding, and attitudes of the nation's children and youth. This data collection program began in 1969 and currently assesses different curriculum areas (e.g., reading, writing, mathematics, science, citizenship, U.S. history, geography, social studies, art, music, literature, career and occupational development) in grades 4, 8, and 12 every two years. Two years (1988 and 1990) as well as the voluntary state program started in 1990 (the State Trial) were reviewed for this report. The State Trial provided state-level mathematics data for eighth graders for 40 participating jurisdictions.

National Education Longitudinal Study of 1988 (Department of Education) -- NELS:88

A nationally representative longitudinal study designed to assess the baseline experiences of eighth grade students and to relate these experiences to current academic achievement and to later achievement in school and life. The 1988 base year data collection program gathered data in a variety of areas such as work status, values, school characteristics, school atmosphere, school work, school performance, guidance, special programs, after-school supervision, involvement with community, after school activities, educational and occupational life goals, and financial assistance. Follow-up assessments are being completed every two years from 1990 to 1996.

Current Population Survey, March Supplement (Department of Commerce) -- CPS

A nationally representative cross-sectional study designed to collect information on the employment situation and demographic status of the complete U.S. population (birth through adulthood). The March Supplement is specifically designed to gather data on work experience, income, noncash benefits, and population migration. Data collection in this program has been conducted annually since the 1940s.

National Health Interview Survey (Department of Health and Human Services) -- NHIS

A nationally representative cross-sectional study designed to provide information on the health of the civilian noninstitutionalized U.S. population (birth through adulthood). This survey has been completed annually since 1957. While the same basic demographic and health-related information is collected each year, additional information on special health topics (e.g., AIDS, aging, etc.) may be covered in any one survey.

National Health and Nutrition Examination Survey, National Health Epidemiological Follow-up Study (Department of Health and Human Services) -- NHANES, NHEFS

A nationally representative longitudinal study designed to (a) provide information on the prevalence of health conditions and risk factors, (b) monitor changes over time in health, functional status, and utilization of hospitals, and (c) track the incidence of various medical conditions in the U.S. population (birth through adulthood). The base year data are drawn from the National Health and Nutrition Examination Survey 1 (NHANES 1), with the follow-ups in 1982-84, 1986, 1987, and 1991.

Longitudinal Study of American Youth (National Science Foundation) -- LSAY

A nationally representative longitudinal study of seventh and tenth graders designed to assess student attitudes toward science and mathematics as areas of study and possible career choices. Base year data collection started in 1987, with annual follow-ups.

Table 4: Decision Makers and Guidelines Used in Inclusion Decisions for Students with Disabilities in Select National Data Collection Programs

Agency/Program	Who	Disability-Related Exclusion Guidelines
DEPARTMENT OF EDUCATION (NCES)		
National Assessment of Educational Progress, 1988 (NAEP:88)	School staff	Students on sampling roster who were deemed to be untestable and unable to participate meaningfully in the assessment. Disability related ineligibility categories used were: • Mild retardation (educable) • Functional disability
National Assessment of Educational Progress, 1990 (NAEP:90)	School staff	Students on sampling roster who were deemed to be unassessable and unable to participate meaningfully in the assessment. Disability-related ineligibility categories used were: • Student is in special education with an IEP and is mainstreamed less than 50% of the time in academic subjects and is judged incapable of participating meaningfully in the assessment. • Student is in special education with an IEP and the IEP team or equivalent group has determined that the student is incapable of participating meaningfully in the assessment.
NAEP Trial State Assessment Program, 1990	School staff	(Same as NAEP:90 above)
National Education Longitudinal Study of 1988 (NELS:88), Base Year	School staff	Students on sampling roster designated eligible if determination was made that student was capable of completing the survey instruments, and designated as ineligible if judged that student would be incapable of doing so. Disability related ineligibility categories used were: • Severe mental disability • Physical disability
National Adult Literacy Survey, 1992 (NALS:92) Base Year	Interviewer	Individuals on sampling roster who are unable to complete the background questionnaire or literacy exercises due to: • Language problem • Physical or mental disability • Reading difficulty
DEPARTMENT OF HEALTH AND HUMAN SERVICES (NCHS)		
National Health Interview Survey, 1989 (NHIS:89)	NA	No disability-related exclusion guidelines; individuals with disabilities are not systematically excluded. Information regarding individuals with disabilities who are included on the sampling roster is collected through adult proxies.
Nat'l Health & Nutrition Examination Survey I (NHANES I) Epidemiologic Followup Study, 1986 (NHEFS:86)	NA	(Same as NHIS:89 described above)
DEPARTMENT OF COMMERCE (Census Bureau)		
Current Population Survey (CPS)	NA	(Same as NHIS:89 described above)
NATIONAL SCIENCE FOUNDATION		
Longitudinal Study of American Youth, 1987 (LSAY:87), Base Year	School staff	No formal disability-related exclusion guidelines reported. Only mention of exclusion of students on sampling rosters was in regards to students who declined or refused to participate.

individuals with disabilities is not a significant issue, since third party informants or proxies can provide the necessary information for individuals who have disabilities that preclude their own communication with interviewers. In the case of these data collection programs, no specific exclusion guidelines are specified, nor are they needed.

This contrasts with those data collection programs that require the selected individuals to respond themselves to actual test items or survey instruments (viz., NAEP, NAEP Trial State Assessment, NELS, NALS, LSAY). In all of these data collection programs there is an expressed concern for individuals who are unable to participate meaningfully in the assessment due to some form of disability. The rationales advanced for exclusion typically revolve around a concern for not submitting individuals to a very stressful and often futile attempt at data collection, a situation that may also produce results of questionable quality.

Although there is generally a common ground in the reasons stated for exclusion and for who makes the exclusion decisions (usually local school staff), little in the way of common ground is found across the sampled data collection programs in operational guidelines. At one extreme is the inability to find any formal documentation of disability-related exclusion guidelines for LSAY:87. Inspection of the relevant LSAY:87 documents, including a search of the files by the current staff managing the program, failed to find any specific mention of disability-related exclusion guidelines for LSAY:87. In contrast, NAEP:88 and NELS:88 were similar in that they each suggested two possible categories (although the categories were different) for exclusion based on some form of disability; NAEP:88 used "mild retardation (educable)" and "functional disability" while NELS:88 used "severe mental disabilities" and "significant physical disabilities." Reflecting changes in methodology, the more recent NAEP:90 and NAEP Trial State Assessment Program of 1990 use exclusion categories tied to whether students are on an active Individualized Educational Plan (IEP), and the extent to which they participate in mainstream education.

NALS:92 plans to use yet a different set of exclusion guidelines, probably because it does not focus on a captive school population with readily available special education information on students. Individuals unable to complete the background questionnaire or literacy exercises due to language, visual impairment, physical or mental disability, or reading difficulties will be excluded. These determinations apparently will be made by the trained interviewers who have contact with the sampled individuals. Ironically, individuals with reading difficulties, who by definition are of greatest concern in the area of literacy, will be excluded because of their difficulty in reading.

In summary, national data collection programs appear to vary markedly in the extent and type of disability-related exclusion guidelines used during data collection. Some of this variability can be attributed to differences in data collection methods, with those not requiring respondents to complete survey instruments independently specifying no exclusion rules. However, even among those data collection programs that use similar data collection methods (e.g., tests), significant variability is noted in exclusion guidelines for individuals with disabilities.

Rates of Exclusion

Information on estimated exclusion rates for the nine data sets is presented in Table 5. Similar to the observed dichotomy in the nature of disability-related exclusion guidelines, the data collection programs appear either to exclude relatively few individuals with disabilities (e.g., NHIS, NHEFS, CPS) or to exclude approximately 1/3 to 1/2 of school age students with disabilities. Again, this dichotomy can be attributed to the difference in methods of data collection noted previously.

With the exception of NALS:92, for which figures are not yet available, and LSAY:87, for which no figures were reported, all data collection programs listed in Table 5 that require direct testing of students (NAEP, NAEP Trial State Assessment Program, NELS) exclude approximately 1/3 to 1/2 of all school age students with disabilities. As noted in Table 5, these estimates are based only on that

Table 5: Inclusion Rates for Individuals with Disabilities in Select National Data Collection Programs

Agency/Program	Estimated Percent with Disabilities Excluded[a]
DEPARTMENT OF EDUCATION (NCES)	
National Assessment of Educational Progress, 1988 (NAEP:88)	Average exclusion rate for total sample: approx. 5.7%, which includes students with limited English proficiency (LEP). If assume 1/3 of excluded students were LEP[b], 3.8% of the originally sampled students were excluded due to disability-related guidelines. Approximately 40% of students with disabilities are excluded from total sample.[c]
NAEP Trial State Assessment Program, 1990 (NAEP: Trial State)	Average exclusion rate across samples: approximately 4.4% for students with IEPs. Average of 52.7% of students with IEPs selected for the samples excluded.
National Assessment of Educational Progress, 1990 (NAEP:90)	(Assume same as NAEP Trial State Assessment Program, 1990)[d]
National Education Longitudinal Study of 1988 (NELS:88)	Exclusion rate for total sample: approximately 5.4%, which includes 1.9% students with limited English proficiency (LEP). 3.4% of the originally sampled students excluded due to disability-related guidelines. Approximately 36% of students with disabilities excluded from total sample.[c]
National Adult Literacy Survey, 1992 (NALS:92) Base Year	Information not yet available since survey will not be implemented until 1992.
DEPARTMENT OF HEALTH AND HUMAN SERVICES (NCHS)	
National Health Interview Survey, 1989 (NHIS:89)	No figures reported since no disability-related guidelines are used. Exclusion is probably negligible in amount.
Nat'l Health & Nutrition Examination Survey I (NHANES I) Epidemiologic Followup Study, 1986 (NHEFS:86)	(Same as NHIS:89 described above)
DEPARTMENT OF COMMERCE (Census Bureau)	
Current Population Survey (CPS)	(Same as NHIS:89 described above)
NATIONAL SCIENCE FOUNDATION	
Longitudinal Study of American Youth, 1987 (LSAY:87), Base Year	No figures reported since no disability-related guidelines are documented. Any informal/formal exclusion guidelines most likely resulted in exclusion rates similar to those of programs with similar data collection instruments (e.g., NAEP, NELS).

[a] Estimates only reflect exclusion of student population attending regular schools (does not reflect students with disabilities in separate facilities).

[b] 1988 NAEP technical report only provides exclusion figures as a total and does not report a breakdown by different exclusion categories. The use of the 1/3 figure for LEP students is drawn from the rate of LEP exclusion reported for the NELS:88 study which occurred at the same time.

[c] Estimate of total percent of individuals with disabilities excluded calculated by comparing reported exclusion percent (for disability-related reasons) in total sample with average percent of student population with disabilities (not including separate facilities). Since average values reported during recent years indicate approximately 10% of the student population can be classified as having a disability, and since approximately 7% of this population receives services through separate facilities, a value of 9.3% was used in these calculations. This represents the percentage of students with disabilities who were included in the sampling frame.

[d] Final technical report for NAEP:90 not available at time of this report. The procedures used were like those employed in the 1990 NAEP Trial State Assessment, suggesting that exclusion rates are probably comparable. However, analysis by Spencer (1991) suggests that NAEP:90 may have excluded more than the 1990 NAEP Trial State Assessment, based on a comparison of percent of tested students in the samples with IEPs.

portion of the school age population with disabilities that is receiving special education services through regular education, resource room, or separate special education classes. These figures do not reflect the additional exclusion of students with disabilities that occurs as the result of data collection programs starting with a sampling universe that excludes separate special education facilities (e.g., residential, homebound, hospital, separate school settings). It is estimated that approximately 7% of the school age population receives special education services in such separate environments. In numbers, this is approximately 315,000 students. The exclusion of these separate environments from the sampling universe is based on the rationale that essentially all students served in these settings have serious or multiple disabilities that preclude their meaningful participation in data collection, and that all such students would probably qualify under a data collection program's exclusion guidelines. However, such a blanket assumption is flawed. Many students in residential settings are capable of participating in standardized testing programs, and often do so on a regular basis (NASDSE, 1988).

Regardless of the exact proportion of students in separate facilities who could participate in data collection programs (even if this number were quite small relative to the entire sample size), the important point is that a significant portion of the student population with disabilities (i.e., all students in separate facilities) is ignored in much of the national education data collection system as currently designed. **The undercoverage of the total school age population of students with disabilities is estimated to be between 40% and 50%.**

Variability in Implementation of Exclusion Guidelines

Although certain of the data collection programs listed in Table 4 specify exclusion guidelines, a review of the relevant manuals from these programs suggests that little, if any, information is typically collected on the degree to which the guidelines are implemented consistently. Recently, information from two of the data collection programs (viz., NAEP Trial State Assessment; NELS) suggests that there may be significant variability in application of exclusion guidelines within data collection programs.

As reported in Table 4, approximately 52.7% of all selected students with IEPs were excluded from the 1990 NAEP Trial State Assessment data collection. State-by-state exclusion figures (Houser, personal communication, February, 1991) found exclusion rates for students with IEPs to be as low as approximately 33% (Minnesota) and as high as approximately 71% (Arkansas), with an even higher rate of 87% for the District of Columbia. Given that the same design procedures and exclusion guidelines were used in both the national and state 1990 NAEP assessments, one could conclude that considerable variability in the implementation of disability-related exclusion guidelines appears to occur in both the national and state NAEP assessments.

The follow-up study being completed on a sample of students determined to be ineligible for the NELS base year data collection (NELS:88) is providing important insights into the exclusion process (Ingels, 1991). In the case of NELS, as well as the other data collection programs listed in Table 1 that have exclusion guidelines, the application of the exclusion guidelines and all related decisions are made by local school staff. In the directions to local staff during NELS base year data collection, schools were asked to apply the exclusion guidelines on an individual basis, and not to exclude students categorically. In the case of uncertainty, school personnel were asked to include the student.

Preliminary results from the NELS Ineligible Study indicated that despite these directions, schools often resorted to the categorical exclusion of students (Ingels, 1991). Evidence for this categorical exclusion was observed in the exclusion of all students within a specific category on some of the school sampling rosters. In addition, if schools had been making individual exclusion decisions about students with disabilities and including students when in doubt, one would expect to find a number of students who would experience difficulty during the actual testing. In contrast, exceptionally high

completion rates (i.e., exceeding 99%) were reported for all tests. This leads to the conclusion that "the screening out of students was too effective in that one would expect more borderline cases had schools taken with full seriousness the injunction 'when in doubt, include'" (Ingels, 1991, p. 11).

Further evidence reflecting arbitrary exclusion during some of the data collection is found in preliminary results from the NELS Ineligible Study during the first follow-up survey (S. J. Ingels, personal communication, June 25, 1991). These results indicated that of the approximately 94% of the students in the NELS base year (1988) ineligible sample that were studied during the first follow-up (1990), over half (58.5%) were reclassified as eligible. More importantly, approximately 94% of those reclassified as eligible were able to successfully complete the data collection instruments. Unfortunately, the preliminary information available at the time this report was written did not provide for a disaggregation of the ineligible follow-up results for just those excluded due to disability-related guidelines (e.g., the data also include students excluded due to limited English proficiency). However, one probably can assume from all of the preliminary evidence reported from the NELS Ineligible Study that significant numbers of students who were deemed "ineligible" during base year data collection due to a specific disability should not have been excluded and apparently can successfully participate in these types of data collection activities. Most likely these results also generalize to other data collection programs (e.g., NAEP) since many are directed by the same agencies and many rely on local implementation of exclusion rules.

How Does Exclusion Occur?

The exclusion of certain segments of the population from large scale data collection programs occurs in a variety of ways. For example, the increased use of computer-assisted telephone interviewing in place of face-to-face interviewing has resulted in the omission of households without telephones. The recent National Household Education Survey (NHES; National Center for Education Statistics, 1991), which uses computer-assisted telephone interviews, estimates that approximately 8% of households were excluded because they did not have telephones (West, Hausken, Chandler, & Collins, 1991). Households without telephones tend to differ in substantive ways from those with telephones, usually in terms of including a higher percentage of low-income households. Other persons excluded when telephone surveys are used include persons who are deaf or who use telecommunication devices (Kiecolt & Nathan, 1985).

As mentioned previously, a substantial proportion of students with disabilities are regularly excluded from national education surveys. Special schools for individuals with disabilities, which are attended by as many as 7% of all students with disabilities, are not included in most sampling frames. In addition, students with disabilities attending regular schools are frequently excluded. For example, in HSB three sampling constraints have led to the non-inclusion of a segment of students in the data base (Owings & Stocking, 1985). First, only students who were involved in high school programs leading to graduation and a diploma were included. Consequently, those students in nondegree programs (such as students in special education who were working toward a certificate of completion) were not included in the sample. Second, those students who could not read and complete the questionnaire on their own were excluded, eliminating, among others, students who were blind or who had difficulty using pencils. Third, those students judged by teachers to be "at risk" for experiencing discomfort during the testing situation also were excluded. Consequently, a substantial proportion of students with mental, emotional, and/or physical disabilities were excluded from the HSB sample.

Exclusion also probably occurs for a number of reasons that have not been verified in the literature. As noted above, students may be excluded if it is perceived that they would experience discomfort during the testing situation. Some have suggested that an extension of this may be that students who do not want to take a test (translated to the test situation would cause discomfort) may similarly be excluded. Beyond this, it also has been suggested that administrators may exclude students if it is perceived that the

students' test scores would lower the performance levels reported for a school or a district.

Points of Exclusion

A detailed review of the sample selection and exclusion procedures used for NELS:88 illustrates those study design points where students with disabilities typically are excluded from national data collection programs. NELS:88 is only used for illustrative purposes, and is not being singled out as "the" example to illustrate exclusionary practices. In fact, the NELS:88 survey is probably the most prominent example of recent attempts to address the issue of exclusion of students with disabilities from survey samples. Even though exclusion rules were used during the collection of the base year data for NELS:88, extensive follow-up studies are being completed to reassess the continued status of excluded students, and where appropriate, to add such students back into the study at subsequent follow-ups (Ingels, 1991). The sample exclusion process of NELS:88 shown in Figure 1 is used only as an example to identify those points where students with disabilities may be systematically excluded from large scale sampling plans.

The five shaded boxes in Figure 1 are decision-making points that frequently result in the exclusion of students with disabilities from large scale assessment programs. First, a significant proportion of the student population with disabilities is excluded when separate schools for individuals with disabilities are excluded from the sampling frame. Second, the use of school rosters based on grade placement also results in the exclusion of any students in ungraded special education classes. Additional exclusion occurs through the application of formal exclusion criteria at two other decision-making points (pre-roster ineligibility decision; sample updating). These result in students being excluded if physical or mental disabilities are present that, in the judgment of school staff, preclude independent completion of the survey instruments. The final exclusion point may occur at the time of the administration of the survey if students are observed to experience difficulty independently completing the instruments.

Of the possible exclusion points highlighted in Figure 1, the last three are those for which data typically are gathered for reporting exclusion rates. These reported exclusion rates only represent the percent of students with disabilities who are excluded from the selected sample, and are most likely underrepresentations of the total number of students with disabilities who are excluded. This number is already lowered when entire schools are excluded and when students are in ungraded programs. The information conveyed by Figure 1 suggests that a sizable portion of the total student population with disabilities is excluded from most national data collection programs.

State Education Data

Just as federal initiatives have engendered interest in assessing student progress, so have state-level initiatives. Increasing numbers of states are involved in collecting outcomes data from schools in order to determine whether state goals are being met.

Data relevant to state exclusion policies are available from a survey of states conducted by the National Center on Educational Outcomes (NCEO, 1992; Shriner, Bruininks, Deno, McGrew, Thurlow & Ysseldyke, 1991). In the Spring of 1991, state directors of special education or their designees responded to a telephone survey on a variety of issues related to the assessment of outcomes. Seven survey questions addressed the issue of exclusion of students with disabilities from state-level outcomes assessments. These questions were:

- Do any students with disabilities take part in tests given to general education students that are collected and/or reported at the state level?
- What decision rules or guidelines are used to determine which students participate and which are excluded?
- Who makes the decision to include or exclude a student with disabilities regarding the outcomes assessment in general education?

Figure 1: NELS:88 Base Year Sample Selection/Exclusion Decision Making

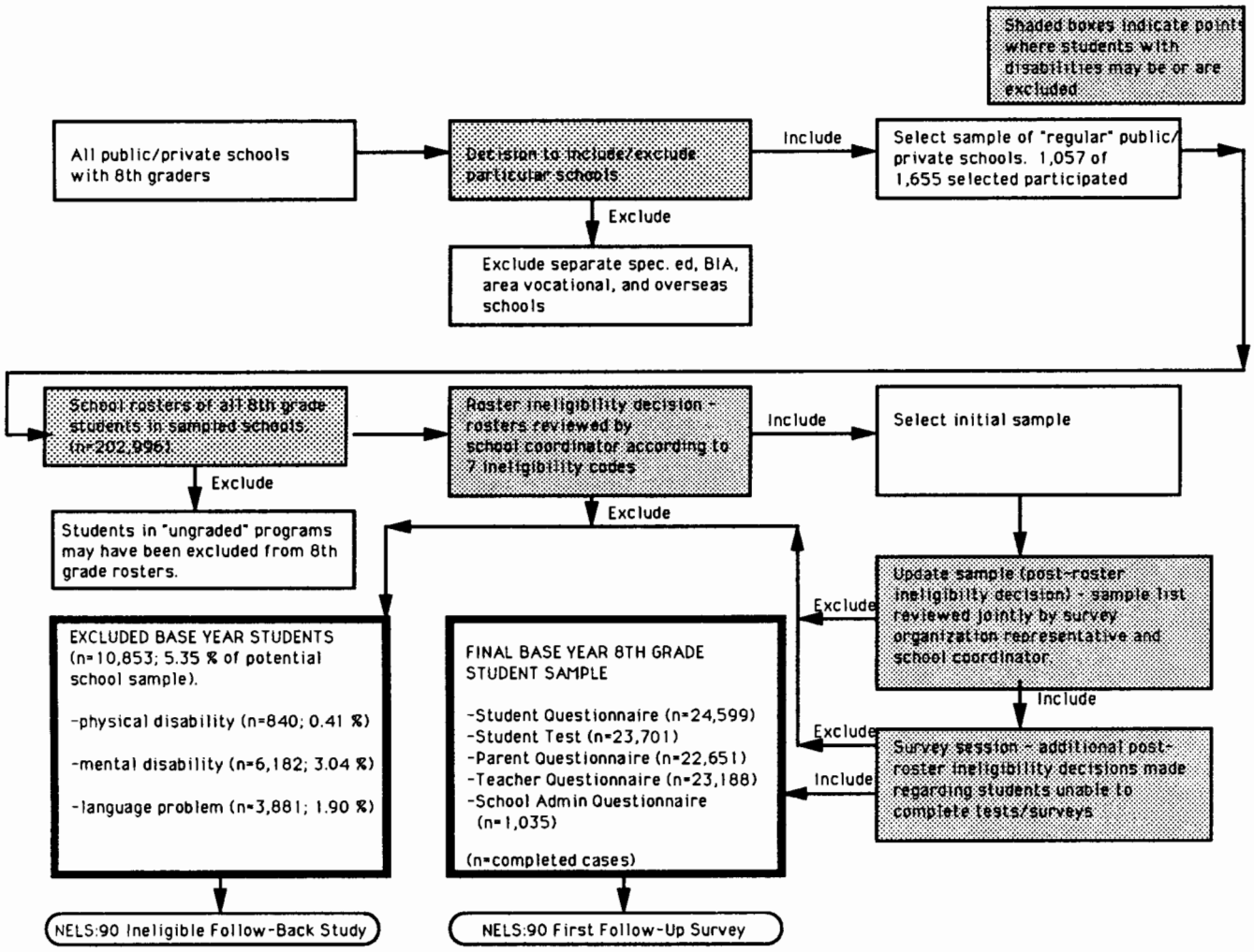
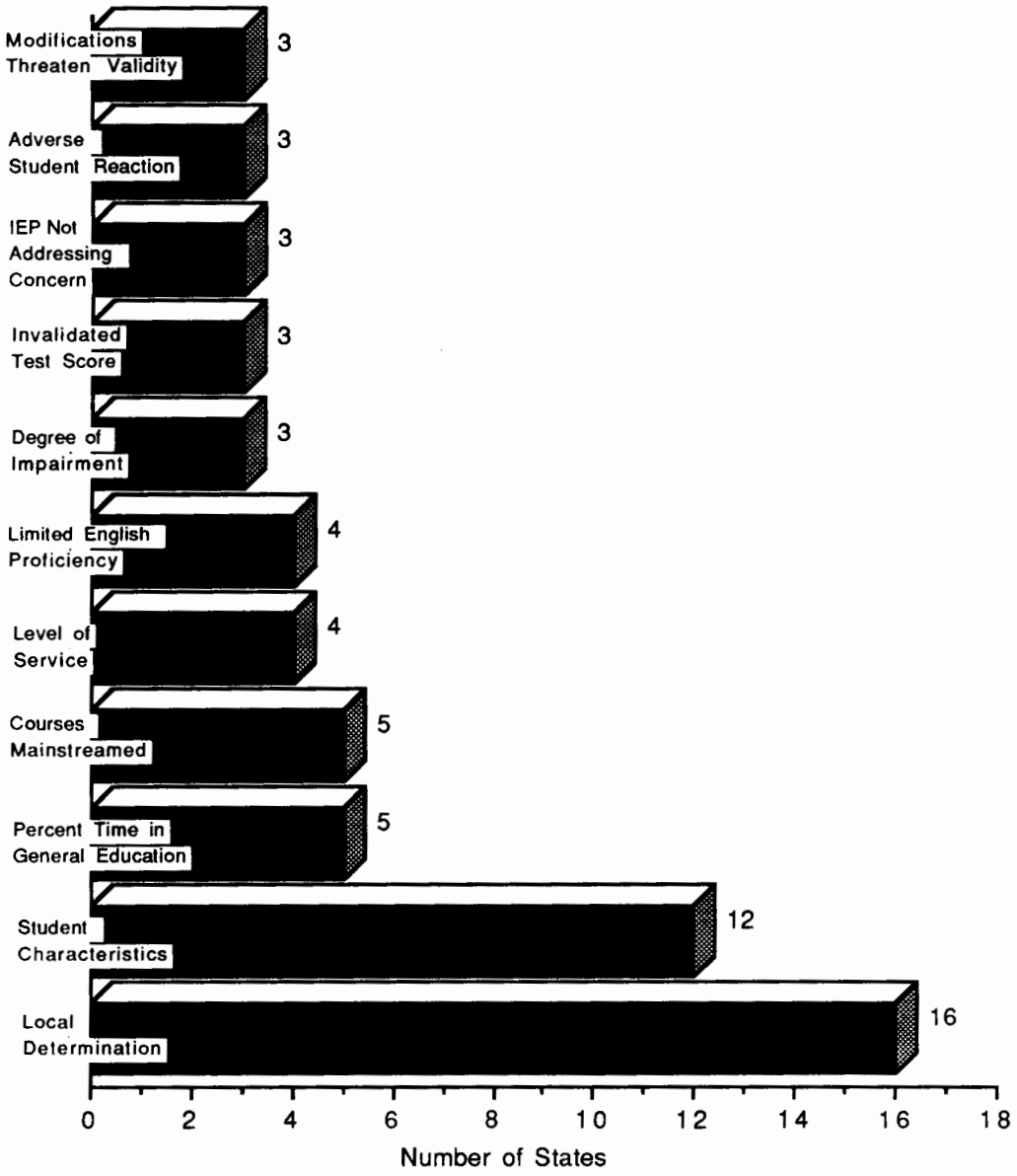


Figure 2: Number of States with Formal or Written Rules (n=34) Reporting Guidelines for Participation in General Education Assessments*



* Source: Shriner, Bruininks, Deno, McGrew, Thurlow, & Ysseldyke, 1991.

- Are these decision rules formal or informal?
- Would you share a copy of these guidelines with the Center?
- Can students with disabilities who participate in the assessments be identified in the data set?
- About how many students with disabilities do you believe participate each year in the assessments conducted for students without disabilities?

Criteria for Exclusion

The NCEO survey revealed that 49 of the 50 states reported that students with disabilities participate in the assessment of student achievement. Within these states, 34 states indicated that they have formal or written decision rules for the exclusion of students with disabilities. The decision rules generally are guidelines for districts and local schools. Eight types of guidelines encompass those used by most states to determine whether a student will participate in the general education assessment. These guidelines, which do overlap, include:

- Locally-determined rules for participation
- Student specific characteristics
- Percent of time in general education classes
- Courses for which student is mainstreamed
- Level of service received
- Documented circumstances that invalidate test scores
- IEP does not address test content
- Degree of student impairment

The number of states using each of these guidelines is displayed in Figure 2.

When states were asked to report on who makes the decision about inclusion of students with disabilities in state testing, 37 of the 49 states with students participating in some type of outcomes assessment identified local IEP teams. Other states indicated a principal (n = 8), parents (n = 9), or some other unique decision maker (n = 15); five states indicated that inclusion decisions were made by the state education agency. Of the 34 states with formal or written guidelines that addressed the issue of inclusion of students with disabilities in general education achievement testing, 27 (79%) identified the IEP team as the decision maker.

The consistency with which exclusion guidelines are implemented is another issue. While considerable effort has been expended in states to develop guidelines to assure the proper inclusion of students with disabilities, some State Directors of Special Education had significant doubts about how guidelines were operationalized and implemented. The special education units of most state agencies do not directly monitor inclusion decisions since it is typically the general education unit that conducts state assessments. As such, the inclusion/exclusion decision is influenced by factors not directly related to the student. For example, in states where scores determine various types of funding allocations, students with disabilities may be categorically excluded in order to increase the probability of higher or lower school scores, depending on which would better serve the purposes of a particular school.

Even when students with disabilities participate in state assessments, there is no guarantee that their data are included in state-level analyses and/or reports. Inconsistent handling of data after it leaves the local school site is quite common. Although not formally asked in the survey of states, respondents frequently expressed concern that many students are being excluded post hoc. In other words, their test protocols are simply discounted or discarded. Students are exposed to assessments that produce possibly valuable information, yet these data from the assessments are not used to enhance the students' education in any way.

Testing Accommodations for Students with Disabilities

The NCEO state survey results also revealed that one of the most important determinants of whether a student with a disability actually participated in state-level testing was whether the state or the local

school allowed special testing accommodations for the student. Forty-two states reported that they provide one or more forms of testing accommodations for students with disabilities. However, only 28 states had written guidelines about these accommodations. The nature of the accommodations identified by the 42 states is shown in Table 6. The most frequent accommodations were alternative presentation modes (e.g., Braille test, oral reading of test to student, etc.). Flexible testing times and flexible settings also were identified by a large proportion of the states.

Exclusion Rates

The rates at which students with disabilities are excluded from state assessment programs seems to be largely unknown at this time. When State Directors of Special Education or their designees were asked to specify the number or percent of students with disabilities who participated in their state-level general education testing programs, many states (n = 34; 68%) did not have an idea of the extent of participation. Some respondents were uncertain whether students with disabilities could be identified in their data sets. The 19 states that did report participation rates are shown in Table 7. Of the 13 states giving percentages, 5 states indicated that more than 90% of their students with disabilities participate in state testing, 5 indicated that between 50% and 90% participate, and 3 states indicated that less than 50% of their students with disabilities participate in testing. Of the three states with the lowest percentages, two had less than 5% of students with disabilities participating in state-level testing.

Conclusions

Calls for reform in American education during the past decade have resulted in raised expectations, attempts to develop uniform and "world class" standards, and increased emphasis on school accountability and the measurement of educational outcomes. Reform initiatives, both at the state and national level, are increasingly focusing on outcomes and quantifiable data. The measurement of educational indicators is playing a central role in the current wave of educational reform as various groups seek to produce policy-relevant information on the educational performance and status of children in our nation's schools. With increasing frequency, the data needed to monitor and evaluate these reform initiatives are being drawn from national and state data bases.

The Nature and Context of Exclusion

As currently designed, most of the existing national and state data collection programs exclude large portions of the student population with disabilities. At the national level, it is estimated that approximately 40% to 50% of all school-age students with disabilities are excluded from the most prominent national data collection programs (e.g., NAEP, NELS) that are playing a critical role in the evaluation of the current reform initiatives. Indicative of the low status that the school-age disability population has in data collection programs is the finding that state-level data documenting the extent of exclusion is, with few exceptions, largely unavailable at this time.

It is clear that the ability to extract useful national and state policy-relevant information on the outcomes of students with disabilities is seriously hampered by the extensive exclusion of portions of this population. Given the current assessment technology, some exclusion of students with unique needs is understandable and cannot be avoided. Unique testing accommodations cannot address all disability-related problems in large-scale national and state assessment programs.

Our review of national and state data collection programs suggests, however, that a sizable portion of excluded students should not have been excluded, and could readily participate (some with testing accommodations; others without) in such data collection programs. Contributing to this problem is the significant variability in the types of exclusion criteria used in national and state data collection systems. As indicated by the NELS Ineligible Study and the state NAEP results, even when exclusion criteria are in place, implementation of the criteria are variable, and may be affected by local decisions, which sometimes are driven by a concern for producing high scores in assessment results. In general, little effort has been made to

Table 6: Testing Accommodations Used by States

Accommodation	Number of States
Flexible Time	22
Flexible Setting	22
Alternate Presentation Mode	
IEP Determined	22
Braille	21
Oral Reading	15
Sign Language	9
Large Print	4
Alternate Response Mode	
IEP Determined	10
Sign Language	9
Oral Response	8
Computer	6

Table 7: Estimated Participation Numbers and/or Rates by States

State	Number	Percentage
California	--	90%+
Connecticut	--	65%
Delaware*	--	104%
Georgia	3,000-4,000	--
Indiana	20,000	--
Louisiana	5,000	--
Maine	--	95%
Michigan	--	2%
Montana	--	50%
New Jersey	--	50-60%
New York	28,000	--
North Carolina	--	98%
Pennsylvania	7,794	6%
South Carolina	9,500	--
Tennessee	--	90-95%
Texas	--	70%
Utah	30,000-35,000	71-83%
Virginia	6,000	--
Wisconsin	5,145	45%

* Delaware percentage verified in Special Education Effectiveness Development System (SEEDS). Delaware Department of Public Instruction, 1991. Ratio for percentage formed by dividng the number of students tested (May) by the number of students eligible (September).

systematically study ways to increase the rate of participation of students with disabilities in many, if not most, large-scale assessment programs.

When and How Exclusion Occurs

The exclusion of students with disabilities from national and state data collection programs occurs at a number of different stages: (a) during the development of assessment instruments, (b) when the data are collected, and (c) during the analysis, interpretation, and reporting of the results.

Our review of the methodology reports of most national assessment programs revealed little description of how students with disabilities are included in the item development phase of assessment instruments. Furthermore, most sampling plans for national data collection programs systematically exclude certain segments of the school-age population with disabilities (e.g., students in residential, homebound, hospital, or separate school settings). The logic in assuming that all of these students are incapable of participating in standardized testing programs is flawed, since many of these students actually are involved in standardized testing on a regular basis (NASDSE, 1988). This situation is analogous to the exclusion of students with disabilities in the development and standardization of many norm-referenced tests (Fuchs, Fuchs, Benowitz, & Barringer, 1987).

Commercial tests commonly used in state assessments (e.g., Stanford Achievement Test) typically do not include students with disabilities in the development stages of the test, or if they do, they do not report this information. Although the inclusion of representative samples of students with disabilities is not necessarily required for the calculation of accurate norms in well standardized instruments, the inclusion of such individuals in test development is critical from the perspective of learning what items or test administration procedures may need to be dropped or modified in order to accommodate the unique testing needs of this population (Daniels, 1989).

The exclusion of students with disabilities during the actual collection of data is probably the most obvious type of exclusion. The stated and unstated reasons for not giving many students with disabilities the opportunity to participate in the assessment are varied. Reasons range from a concern over the inability to provide proper accommodations (e.g., in test administration mode, in response mode, in flexible time or setting) to a concern about the potential aversiveness of the assessment situation for the student. The most common state-level criteria used for these exclusion decisions are based on (a) local rules about participation, (b) information regarding student specific characteristics, or (c) amount of time the student spends in general education classes. Although the exclusion criteria for national data collection programs may differ somewhat from those used by state programs, in both cases, exclusion criteria are typically implemented by local school personnel. As noted by Ingels (1991), the specification of formal exclusion criteria by no means guarantees accurate implementation. Local implementation variability appears to be commonplace. It is possible that many local school staff who make these decisions, who understandably are most concerned about their immediate setting, do not appreciate the "bigger picture" of national and state testing, and therefore do not incorporate into their decisions the usefulness of such information for school improvement and the development of educational policy.

Exclusion in the analysis and interpretation phases of assessments occurs when students who were included in the assessments are excluded from the data analysis. This type of exclusion is common in state-level assessments, possibly due to a desire to maintain higher school, district, or state level results in the current era of high stakes testing. From the results of the NCEO state survey, it seems that students with disabilities present many challenges to state agencies that try to include them in state assessments. Often, an easy solution is to remove their data from those of the "regular students." States are aware that data on students with disabilities might be retrieved and analyzed, but have not directed consistent efforts toward this task. In addition, few special education sections in state departments of education use the data

separately from general education for special education related policy decisions. Even in those states that collect data that can be disaggregated separately for students with disabilities, state directors of special education sometimes are unaware that this capability exists.

Impact of Exclusion

The exclusion of large numbers of students with disabilities from national and state assessment programs raises questions about the inferences that can be drawn from the results of these data collection programs. This exclusion may result in significant problems in estimation of national and state level statistics, the ability to obtain representative samples, and the ability to complete accurate policy studies (Ingels, 1991).

Not only does the treatment of students with disabilities as outliers in national and state data collection programs make it difficult to produce accurate national and state statistical estimates for this population, it also raises questions about bias being present in most national and state education statistical estimates that are reported. For example, dropout rates for certain subgroups (e.g., students with disabilities, students with limited English proficiency) of the population may be significantly higher than the general population. If large portions of these subgroups are excluded from the calculation of national and state dropout statistical estimates, the estimates may be biased. When the exclusion occurs for subgroups that typically have the most difficulty in learning and functioning within the schools (e.g., students with disabilities), the resulting estimates most likely will paint a more optimistic picture of educational progress, especially for those subgroups that are not adequately represented. Ironically, in such situations it is those who are excluded who are typically targeted for intervention programs and are those groups most often of greatest policy interest (Ingels, 1991).

Other problems can also result from the exclusion of students with disabilities from national and state data sets, particularly when data collection programs use different exclusion procedures and criteria. Samples that result from the implementation of different exclusion guidelines can cause problems with sample comparability (McGrew, Spiegel, Thurlow, Ysseldyke, Bruininks, Deno, & Shriner, 1991b). That is, differences in sample composition can make it difficult to make comparisons across data bases. In addition, exclusion procedures may make it difficult to study small groups (e.g., subgroups of individuals with disabilities) because of the resulting small sample sizes, and more critically, because of the significant problems this introduces into the estimation of appropriate standard errors of the estimate for statistics for these small subgroups (McGrew, Spiegel, Thurlow, Ysseldyke, Bruininks, Deno, & Shriner, 1991b).

Finally, the treatment of most students with disabilities as "outliers" in our national and state data collection programs is a concern from an equity and philosophical perspective. The categorical exclusion of students with disabilities "perpetuates the myth of inherent differences. It makes students with handicaps non-students and perhaps non-people" (NASDSE, 1988, p. 10). Given the magnitude of federal and state support for educational programs for students with disabilities, support that reflects the valuing of this population in our society, it is time that this implied value is matched by the commitment of resources to address the numerous political and technical hurdles that must be overcome in order for these students to more fully participate in our national and state data collection programs. Current and future activities of the NCEO are being focused in this direction.

Implications and Recommendations

Identifying and describing a problem is much easier to accomplish than is developing plans to effectively address the problem. In this report, we have demonstrated that large numbers of students with disabilities are excluded from national and state assessment programs. Multiple political and technical issues underlie this exclusion. Addressing those issues is more difficult than simply describing them. While we believe that further study and empirical research are needed before it is possible to produce a

comprehensive set of guidelines for including all students in assessment programs, it is possible to make some initial recommendations.

Seven recommendations are presented here. This list is not exhaustive and is continually being modified and expanded as part of ongoing NCEO activities. More detailed reports addressing these and additional recommendations will be produced in the future.

Develop broader and more uniform definitions of sample eligibility.

The need for a broad and inclusive definition of sample eligibility that would be uniform across data sets has been pointed out by others (e.g., Ingels, 1991). We found that significant variability exists between and within national and state data collection programs in their definitions and criteria for sample eligibility. This makes comparisons of results across data sets difficult. With a common system of disability definitions and categories, it would be possible not only to make comparisons across data sets but also to develop larger data sets constructed by integrating information from different data sets.

Increase adherence to inclusion guidelines.

Significant exclusion occurs in data collection programs as a result of inappropriate implementation of inclusion and exclusion guidelines. This is evident in the observed variability in the inclusion of students with disabilities among states in the NAEP Trial State Assessment Program, for example. It is also evident in the findings from the NELS Ineligible Follow-up Study, which revealed frequent categorical exclusion of students with disabilities (often involving students who were capable of participating in the data collection). It is clear that the "if in doubt, include" component of most inclusion guidelines is not consistently followed. The use of procedures that involve contacting school staff who have the most direct knowledge about students with disabilities (e.g., special education teachers instead of school office personnel) may help local staff be more consistent in their inclusion decisions (Ingels, 1991). Mechanisms need to be developed to insure greater adherence to uniform implementation of the "if in doubt, include" component of inclusion guidelines in national and state data collection programs. Possible mechanisms include monitoring systems and incentive programs.

Develop sampling frames that are more inclusive.

Many students with disabilities are never considered for inclusion in national and state data collection programs due to sampling plans that routinely exclude them. For example, separate schools for students with disabilities are almost always excluded from large scale assessment sampling plans. Yet, many of these schools serve students who are able to complete standardized assessment instruments. The use of grade-based school rosters can also result in the exclusion of students in ungraded special education programs. Efforts need to be directed toward insuring the inclusion of these excluded portions of the student population in data collection sampling plans.

Include follow-up studies of ineligible students as a standard component of data collection programs.

Greater efforts are needed to study the characteristics of students with disabilities who are excluded from data collection programs in order to accurately estimate the effect of this exclusion on important statistical estimates and to allow re-entry into the sample when the data collection effort is longitudinal. The NELS Ineligible Follow-up Study is an excellent example of attempts to address more systematically the issue of exclusion of individuals from data collection programs. This study has demonstrated the value of following samples of excluded students to (a) determine whether eligibility status changes with time so that initially excluded students can be folded back into the sample, and (b) gather information on the characteristics of excluded students to determine whether adjustments or qualifications need to be made in statistical estimates drawn from the complete sample.

Increase partial participation in data collection programs.

Given the current assessment technology, it is unreasonable to expect that all students with all forms of disabilities will be able to participate in the same way in all components

of national and state data collection programs. Yet, there are opportunities for the collection of partial information that are overlooked in many data collection programs. Typically, data collection programs that require the independent completion of tests or surveys (e.g., NAEP, NELS) exclude a student from all components of the data collection program if the student is unable to participate in those components requiring independent completion. Yet, these data collection programs often collect additional information from third party proxies (e.g., teachers or parents) on important outcome variables. Students with disabilities should be included in data collection components that rely on proxies or other indirect methods (e.g., administrative record reviews). Attention needs to be directed to insuring that these data collection opportunities are not lost for students with disabilities.

Include students with disabilities in instrument development.

The deliberate inclusion of students with disabilities, particularly those for whom an assessment may be relevant and appropriate, during the initial stages of instrument development, is encouraged. Through such involvement, those responsible for the development of assessment instruments can discover items, questions, tasks, or procedures that may need to be eliminated or modified in order to allow these students to participate in the data collection activities.

Develop assessment modifications, accommodations,and alternatives.

Probably the most critical barrier to the inclusion of more students with disabilities in national and state data collection programs is our limited knowledge of what modifications or accommodations can be made to current assessment tools. Research and development activities must focus on investigating the extent to which modifications and/or accommodations can be made to instruments used in large-scale data collection programs, without destroying the basic psychometric properties of the instruments. A variety of modifications and accommodations should be studied, including flexible scheduling, flexible settings, revised test formats, revised test directions, revised response formats, and/or the use of aides (NASDSE, 1988). In addition, research and development efforts need to be focused on the use of, or development of, alternative assessment methodologies or techniques (e.g., computer assisted testing; use of item response technology) that would allow greater numbers of students with disabilities to participate more fully in large-scale assessment programs.

References

Allen, T. E. (Ed.). (1989). Handicapped youth and the outcomes of their schooling — Resent research advances in the United States [Special Issue]. International Journal of Educational Research, 13.

CCSSO. (1990). State education indicators: 1989. Washington, DC: Author.

Daniels, M. H. (1989, April). Issues and advances in norming ability tests. Paper presented the annual convention of the National Association of School Psychologists, Boston.

Darling-Hammond, L. (1991). Measuring schools is not the same as improving them. In Voices from the field: 30 expert opinions on America 2000, the Bush administration strategy to "reinvent" America's schools. Washington, DC: Youth and America's Future, William T. Grant Foundation.

Fuchs, D., Fuchs, L. S., Benowitz, S., & Barringer, K. (1987). Norm-referenced tests: Are they valid for use with handicapped students? Exceptional Children, 54, 263-272.

Ingels, S. J. (1991, April). The problem of excluded baseline students in a school-based longitudinal study: Correcting national dropout estimates and accommodating eligibility change over time. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Kiecolt, K. J., & Nathan, L. E. (1985). Secondary analysis of survey data. Beverly Hills, CA: Sage.

Kirsch, I. S., & Jungeblut, A. (1986). Literacy: Profiles of America's young adults. Princeton, NJ: Educational Testing Service.

Kirst, M. W. (1990). Accountability: Implications for state and local policy makers. Policy perspectives. Washington, DC: Information Services office of Educational Research and Improvement, U. S. Department of Education.

McGrew, K. S., Spiegel, A., Thurlow, M. L., Ysseldyke, J. E., Bruininks, R. H., Deno, S. L., & Shriner, J. G. (1991a). Plan for identifying and analyzing national and state data bases by the National Center on Educational Outcomes. Minneapolis, MN: National Center on Educational Outcomes, University of Minnesota.

McGrew, K.S., Spiegel, A., Thurlow, M.L., Ysseldyke, J.E., Bruininks, R. H., Deno, S.L., & Shriner, J.G. (1991b). Secondary data analysis: A review of major conceptional, measurement, and technical issues. Minneapolis, MN: National Center on Educational Outcomes, University of Minnesota.

NASDSE. (1988). NAEP testing for state comparisons: Issues related to inclusion of handicapped students. Washington, DC: National Association of State Directors of Special Education.

National Center for Education Statistics. (1991). National Household Education Survey (NHES). Washington, DC: U.S. Department of Education.

NAEP (National Assessment of Educational Progress). (1985). The reading report card. Princeton, NJ: Educational Testing Service.

National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform. Washington, DC: U. S. Government Printing Office.

National Education Goals Panel. (1991a). Measuring progress toward the national education goals: Potential indicators and measurement strategies (Discussion Document). Washington, DC: Author.

National Education Goals Panel. (1991b). The national education goals report 1991: Building a nation of learners. Washington, DC: U.S. Government Printing Office.

National Governors' Association. (1986). Time for results: The governors' 1991 report. Washington, DC: Author.

NCEO. (1992). State special education outcomes 1991: A report on state activities in the assessment of educational outcomes for students with disabilities. Minneapolis, MN: National Center on Educational Outcomes.

References

NESAC (National Education Statistics Agenda Committee of the National Forum on Education Statistics). (1990). Improving our national education data systems: An agenda for action. Washington, DC: Author.

Oakes, J. (1991). The many-sided dilemmas of testing. In Voices from the field: 30 expert opinions on America 2000, the Bush administration strategy to "reinvent" America's schools. Washington, DC: Youth and America's Future, William T. Grant Foundation.

Odden, A. (1990). Educational indicators in the United States: The need for analysis. Educational Researcher, 19 (5), 24-28.

Office of Educational Research and Improvement. (1988). Youth indicators 1988: Trends in the well-being of American youth. Washington, DC: U. S. Government Printing Office.

Owings, J.A., & Stocking, C. (1985). High School and Beyond: Characteristics of high school students who identify themselves as handicapped. Washington, DC: U.S. Department of Education.

Shriner, J. G., Bruininks, R. H., Deno, S. L., McGrew, K. S., Thurlow, M. L., & Ysseldyke, J. E. (1991). State practices in the assessment of outcomes for students with disabilities (Technical Report). Minneapolis: National Center on Educational Outcomes.

Smith, M. S. (1988). Educational indicators. Phi Delta Kappan, 69 (7), 487-491.

Special Study Panel on Education Indicators. (1991). Education counts: An indicator system to monitor the nation's educational health. Washington, DC: National Center for Education Statistics.

Spencer, B.D. (1991). Eligibility/exclusion issues in the 1990 Trial State Assessment. Evanston, IL: Methodology Research Center, NORC, Northwestern University.

West, J., Hausken, E. G, Chandler, K. A., & Collins, M. (1991). Experiences in child care and early childhood programs of first and second graders prior to entering first grade: Findings from the 1991 National Household Education Survey. In Statistics in Brief. Washington, DC: Office of Education Research and Improvement, U.S. Department of Education.