ies NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

Measuring the Status and Change of NAEP State Inclusion Rates for Students with Disabilities

Research and Development Report

NCES 2009-453

U.S. DEPARTMENT OF EDUCATION

U.S. Department of Education
NCES 2009-453

Measuring the Status and Change of NAEP State Inclusion Rates for Students with Disabilities

Research and Development Report

November 2008

Sami Kitmitto
Victor Bandeira de Mello
American Institutes for Research

Taslima Rahman
Project Officer
National Center for Education Statistics

U.S. Department of Education
Margaret Spellings
Secretary

Institute of Education Sciences
Grover J. Whitehurst
Director

National Center for Education Statistics
Mark Schneider
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. Unless specifically noted, all information contained herein is in the public domain.

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

November 2008

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
Kitmitto, S., and Bandeira de Mello, V. (2008). *Measuring the Status and Change of NAEP State Inclusion Rates for Students with Disabilities* (NCES 2009-453). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

For ordering information on this report, write to
U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794-1398

or call toll free 1-877-4ED-Pubs or order online at http://www.edpubs.org.

Content Contact

Taslima Rahman
(202) 502-7316
taslima.rahman@ed.gov

FOREWORD

The Research and Development (R&D) series of reports at the National Center for Education Statistics has been initiated to

- Share studies and research that are developmental in nature. The results of such studies may be revised as the work continues and additional data become available;
- Share the results of studies that are, to some extent, on the cutting edge of methodological developments. Emerging analytical approaches and new computer software development often permit new and sometimes controversial analyses to be done. By participating in frontier research, we hope to contribute to the resolution of issues and improved analysis; and
- Participate in discussions of emerging issues of interest to educational researchers, statisticians, and the Federal statistical community in general.

The common theme in all three goals is that these reports present results or discussions that do not reach definitive conclusions at this point in time, either because the data are tentative, the methodology is new and developing, or the topic is one on which there are divergent views. Therefore, the techniques and inferences made from the data are tentative and subject to revision. To facilitate the process of closure on the issues, we invite comment, criticism, and alternatives to what we have done.

Such responses should be directed to

Marilyn Seastrom
Chief Statistician
Statistical Standards Program
National Center for Education Statistics
1990 K Street NW
Washington, DC 20006-5651

EXECUTIVE SUMMARY

Since the late 1990s, participation rates of students with disabilities (SDs) in the National Assessment of Educational Progress (NAEP) from different states have fluctuated. To address concerns that these changes may affect the validity of reports on achievement trends, NAEP has

- instituted policies for providing test accommodations for students with disabilities;
- developed a methodology to correct for the bias resulting from changing inclusion rates, and
- implemented procedures to increase the number of students with disabilities who are included as test-takers, such as better training of field staff, better procedures to assign proper accommodations for students, and improved communications with schools.

States' procedures for including and accommodating students with disabilities are also evolving.

To measure whether these strategies and changes are associated with higher state-by-state inclusion rates, we have developed two distinct approaches for comparing state inclusion rates with one another and gauging progress in their improvement over time. Both approaches rely on regression analysis to estimate the relationship between a student's characteristics and the probability that the student is included on the NAEP assessment. One approach, the nation-based one, estimates one regression using data pooled from all states. The other, the state-specific approach, estimates the regression separately for each state. The relationships are estimated using individual-level data and are then used to establish expectations (or predicted probabilities) for the inclusion of students with disabilities with different characteristics. Individual-level predicted probabilities are aggregated to the state level to form state-level expected inclusion rates. The two approaches examined changes in inclusion rates from 2003 to 2005 and from 2005 to 2007 for grades 4 and 8 mathematics and reading assessments.

For the comparison between 2005 and 2007 described in this report, the two approaches produced similar results when comparing the indices of baseline status of inclusion and change over time:

- The majority of states did not make a statistically significant change in the rate of inclusion.
- Among states that did show a significant change, most were less inclusive in 2007 than in 2005.
  - For the nation-based approach: 8 out of 15 states for mathematics grade 4 were less inclusive in 2007 than in 2005; 17 out of 19 states for mathematics grade 8; 18 out of 26 states for reading grade 4; 21 out of 25 states for reading grade 8.
  - For the state-specific approach: 17 out of 19 states for mathematics grade 8 were less inclusive in 2007 than in 2005; 12 out of 22 states for reading grade 4; 14 out of 18 states for reading grade 8.
  - The exception was for the state-specific approach for mathematics grade 4, where more of the states with significant changes had increases: 8 out of 15.
- Most of the states whose inclusion rate significantly increased in 2007 had a relatively low inclusion rate in 2005.
  - All states with significant increases in inclusion rates in 2007 had relative inclusion rates in the bottom 50 percent in 2005 with the exception of one state for the nation-based method for grade 8 mathematics.

- States whose inclusion rate significantly decreased in 2007 had varied relative inclusion rates in 2005.
- The expected (predicted) inclusion rates varied from state to state by grade and subject.

ACKNOWLEDGMENTS

This report was designed and developed with the help and feedback of many people. We thank members of a panel that was convened to provide a review of the work. We also thank the Education Information Management Advisory Consortium (EIMAC) participants for their feedback and comments on numerous occasions. Finally, thanks go to the many NAEP State Coordinators who also provided a good deal of comments and feedback to preliminary drafts of this report. In particular, we appreciate the contributions of Wendy Geiger (Virginia Department of Education), Marcie Hickman (North Carolina State Department of Education), and Angie Mangiantini (Washington State Department of Education).

CONTENTS

LIST OF TABLES

Table **Page**

Table **Page**

Table **Page**

LIST OF FIGURES

Figure **Page**

INTRODUCTION

The purpose of the National Assessment of Educational Progress (NAEP) is to provide a reliable measure of achievement and trends in achievement at the national and state levels in several grades and subjects. Additionally, NAEP is supposed to report on the achievement of students with disabilities and students identified as limited English proficient (National Assessment of Educational Progress Authorization Act of 2002). Since the late 1990s, the rates at which sampled students with disabilities (SDs) participate (i.e., are included) in NAEP have fluctuated. Reporting of trends requires consistency in practices across years, and the lack of consistency in the inclusion of students with disabilities has called the validity of NAEP trends into question (Forgione 1999; McLaughlin 2000, 2001, 2003). To address these concerns, the National Center for Education Statistics (NCES), the administrator of NAEP, instituted policies for providing test accommodations and has supported the development of a methodology to correct for any bias resulting from changing inclusion rates.[1]

In July 2005, the U.S. Government Accountability Office (GAO) released the report *No Child Left Behind Act: Most Students with Disabilities Participated in Statewide Assessments, but Inclusion Options Could Be Improved*. In the report, the GAO recommended that NAEP "work with the states, particularly those with high exclusion rates, to explore strategies to reduce the number of students with disabilities who are excluded from the NAEP assessment." NCES responded with four actions:

- researched the local decision-making process for participation and accommodation decisions of students with disabilities on NAEP;
- implemented a decision tree that asks whether students could participate in NAEP without their normal state accommodations;
- improved training of NAEP administrators and field staff for 2007 assessments; and
- commissioned this study to develop a methodology for comparing state inclusion rates to one another and gauging progress in improving inclusion rates over time.

This report describes the methodological approach which calculates for each state an expected inclusion rate based on (a) its previous inclusion rates, (b) changes in the distribution of types of students with disabilities in the state, and (c) the set of accommodations offered by the state on its own tests. The method developed is applied to measuring changes from 2005 to 2007 for grades 4 and 8 mathematics and reading assessments.

This report is the first in a series of reports that explore methodologies to measure state-level changes in inclusion rates of students with disabilities as well as English language learners (ELLs). This report focuses on the inclusion of students with disabilities who are not English language learners. In the 2005 and 2007 mathematics and reading NAEP assessments, students with disabilities who were also English language learners made up 13.5 to 15.3 percent of all grade 4 students with disabilities and 16.1 to 19.2 percent of all grade 8 students with disabilities. However, because the factors influencing the inclusion of SDs and ELLs are distinct, we investigate their inclusion processes separately prior to modeling them jointly. We expect SDs who are also ELLs to be included on NAEP under a different process; hence expect that the model and, possibly, results will change by including them. Therefore, findings in this report may not be applicable to SDs who are ELLs or may be different when SDs who are ELL are included. The inclusion of ELLs and of SDs who are also ELLs will be addressed in subsequent reports.

1 For a methodological approach to correct for bias see McLaughlin (2003).

DISTRIBUTION OF TYPES OF STUDENTS WITH DISABILITIES

Provisions for the participation of an SD on NAEP differ by each student's characteristics. Table 1 shows weighted inclusion rates of students with disabilities with different characteristics: different types of disabilities, different severity levels of those disabilities, different grade levels of instruction on the subject being assessed (relative to their grade of enrollment), and whether the student received an accommodation on his or her state assessment that was not allowed on NAEP. For example, among all students with disabilities who were sampled for the 2005 mathematics grade 4 NAEP assessment and were identified with *a specific learning disability,* 85.1 percent participated on NAEP. From this table, it is clear that inclusion rates on NAEP vary by

- different types of disabilities;
- different severity levels of disabilities;
- different grade levels of instruction in the subject being assessed; and
- whether a student receives an accommodation on his or her state assessment that is not allowed on NAEP.

Table 1. Percentages of grades 4 and 8 public school students with disabilities who are not English language learners included in NAEP reading and mathematics assessments, by type of disability, severity level of disability, grade level of instruction, and use of non-NAEP accommodation on state assessment: 2005 and 2007

	2005				2007			
	Mathematics		Reading		Mathematics		Reading	
Characteristics	Grade 4	Grade 8	Grade 4	Grade 8	Grade 4	Grade 8	Grade 4	Grade 8
Disability type (not mutually exclusive)								
Learning disability	85.1	81.7	64.0	73.5	83.5	75.5	64.4	71.3
Speech impairment	85.8	71.1	75.0	63.4	85.6	66.7	77.6	58.7
Mental retardation	38.2	36.1	20.4	25.5	30.3	24.4	19.4	20.4
Emotion disturbance	76.6	76.8	62.7	72.1	72.2	67.2	61.4	70.5
Other disabilities	75.0	72.3	61.6	64.6	75.3	67.5	62.8	65.6
Disability severity level								
Severe	52.9	42.2	36.0	32.8	41.7	29.7	33.7	30.1
Moderate	79.1	72.8	61.4	63.5	76.8	63.2	59.9	59.6
Mild	92.3	86.6	77.1	80.1	91.2	80.8	79.3	77.5
Not reported	74.4	67.5	61.9	58.2	78.9	71.0	66.8	65.3
Grade level of instruction								
Same grade level or above	94.8	89.8	85.9	84.0	93.6	82.7	85.5	81.4
One year below grade	83.0	83.7	68.7	81.2	84.4	75.1	69.5	75.4
Two years or more below grade	51.9	58.1	39.9	51.2	52.2	50.8	41.7	48.4
Not reported	76.4	68.8	63.7	63.0	77.3	70.2	65.7	63.6
Received accommodation on state assessment that is not allowed on NAEP								
No	86.2	83.3	76.8	78.4	86.6	80.3	76.2	75.5
Yes	58.4	56.4	42.1	47.4	51.5	39.3	31.9	34.9

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2005 and 2007 Mathematics and Reading Assessments.

Variation in the distribution of characteristics of students with disabilities across states is demonstrated in table 2, which shows the average, standard deviation, and range of state-level percentage of students with disabilities with each characteristic.[2] For example, the average state-level percentage of students with disabilities with a *specific learning disability* in the 2007 NAEP grade 4 mathematics assessment was 45.8 percent, but this ranged from 14.7 percent in Kentucky to 63.3 percent in the District of Columbia.

Table 2. Average, standard deviation, and range of state-level percentage of students with disabilities by each characteristic, NAEP grade 4 mathematics: 2007

Characteristics	Average (percent)	Standard deviation	Range of percents	
			Min	Max
Disability type (not mutually exclusive)				
Learning disability	45.8	8.7	14.7	63.3
Speech impairment	28.3	8.2	7.2	48.0
Mental retardation	6.0	2.8	1.6	13.2
Emotion disturbance	5.0	2.7	1.2	13.3
Other disabilities	30.2	6.4	16.4	42.7
Disability severity level				
Severe	7.7	3.3	2.3	15.0
Moderate	35.1	9.3	12.3	57.6
Mild	47.5	10.8	25.9	74.5
Not reported	9.8	4.8	2.5	26.2
Grade level of instruction				
Same grade	46.5	9.0	18.4	63.9
One year below grade	19.8	4.2	10.0	27.0
Two years or more below grade	22.6	6.4	7.8	44.3
Not reported	11.1	3.9	4.2	26.2
Received accommodation on state assessment that is not allowed on NAEP				
Yes	15.2	9.7	2.0	47.5

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2007 Mathematics Assessments.

PARTITIONING CHANGES IN STATE-LEVEL INCLUSION RATES

For NCES to track changes in state-level inclusion rates of students with disabilities on NAEP, the goal is to decompose those changes into a portion explained by changes in the distribution of SD characteristics (type of disability, severity of disability, and grade level of instruction) and another, an *unexplained* portion, capturing changes in NCES policy and practices, state efforts, and other factors. As such, in this report, we develop a *partitioning* technique that is based on and akin to Oaxaca-Blinder decomposition techniques (Oaxaca 1973; Blinder 1973).[3] Oaxaca-Blinder decomposition techniques are typically used in studies of discrimination in which differences in an outcome variable, such as different wages for Blacks and Whites, are broken

2 The average given is an average of state-level figures. It is not weighted by the number of students with disabilities in each state and hence does not represent the average prevalence of a characteristic across the country.

3 Fairlie (2003) extended Oaxaca-Blinder decomposition techniques to nonlinear models.

down into two portions, one that can be explained by differences in underlying characteristics thought to affect the outcome, such as years of education and experience, and another that is explained by differences in how those characteristics are treated/rewarded, which is interpreted as discrimination. Our *partitioning* employs Oaxaca-Blinder decomposition techniques but does not delve as deeply in to explanation or interpretation of the portion not explained by differences in underlying characteristics.

The Oaxaca-Blinder decomposition technique measures the portion of mean group differences attributed to differences in underlying characteristics by fixing the individual-level relationship between observed characteristics and outcome. Similarly, our *partitioning* technique fixes the individual relationship between observed characteristics and outcome. This relationship provides a predicted outcome for each individual that is based on his or her characteristics. The difference between two groups' predicted outcomes (Predicted Outcome for Group 2 minus Predicted Outcome for Group 1) determines the portion of the actual difference in outcome (Actual Outcome for Group 2 minus Actual Outcome for Group 1) that is attributed to differences in observed characteristics. The Oaxaca-Blinder analysis simultaneously goes on to analyze and explain the remaining portion as differences in treatment and, hence, discrimination. Here, in our *partitioning,* the remaining portion of the actual difference that is not explained by differences in observed characteristics is attributed to *other* factors.

In the application of our partitioning methodology, we are comparing two groups: a state's SD sample in the initial period and that same state's SD sample in the second period. Here, the initial period is the 2005 NAEP administration and the second period is the 2007 NAEP administration. We use student-level logistic regression models to estimate the relationship between the probability of inclusion on NAEP (dependent variable or outcome) and the observable SD characteristics (the control variables). The estimated coefficients from this regression are used to calculate predicted probabilities of inclusion for all students in each year. These predicted probabilities are aggregated to the state level to get a state-level predicted inclusion rate. The difference between the state's 2007 and 2005 predicted inclusion rates determines how much of the overall difference in inclusion rates is due to differences in the distribution of SD characteristics. The remaining portion of the overall difference is called our *change measure.* The change measure is the primary focus of this report because it is the portion of change that is not due to factors that we expect to cause natural variation in inclusion rates. The methodology is illustrated in figure 1 for the linear case. Our application is to a nonlinear case, which is more complex, but the principles illustrated are the same.

Variation in the application of the Oaxaca-Blinder technique is found in how the individual-level relationship between observed characteristics and outcome is estimated. Different approaches to fixing this relationship can lead to different results. One variation that has been used is to pick one of the two groups as a reference group and estimate the relationship using only individuals in that group. Another variation is to pool the two groups and estimate the relationship using all individuals. Other variations on the estimation of the individual-level relationship exist. Though different variations may lead to different results, each result is still interpreted as a decomposition of the differences between groups. We have the same potential for variation in the application of our methodology for partitioning.

In this study, we develop two approaches to fixing the relationship between observed characteristics and outcome. One is the nation-based approach, in which all the students with disabilities in the initial period (2005, in this report) NAEP sample are used as the reference group for fixing the individual-level relationship between the characteristics of a student and his or her probability of inclusion on NAEP. The second, the state-specific approach, fixes the relationship between the characteristics of a student and his or her probability of inclusion separately for each state, using only that state's initial period (2005, in this report) NAEP SD sample. The benefit of the nation-based approach is that owing to pooling the data across states for estimation, it is possible to use more interactions between the control variables when establishing the relationship between student characteristics and probability of inclusion. The benefit of the state-specific approach is that a separate relationship is estimated for each

state, thus circumventing potential bias resulting from differences between states that might be systematic, such as different definitions of disability used in each state. Neither approach pools data across time periods. Both approaches can, hence, be interpreted as using the initial period as the reference period and as the basis for forming expectations for the second period.

Figure 1. Algebra of the partitioning technique for a linear case

For a given state, let $\overline{Y}_1$ be the inclusion rate in 2005 and $\overline{Y}_2$ the inclusion in 2007. Let $\overline{X}_1$ and $\overline{X}_2$ be the mean vector of control characteristics in 2005 and 2007, respectively. If $\hat{\beta}^*$ is the reference vector of coefficients relating control characteristics to inclusion probability, then $\overline{X}_1\hat{\beta}^*$ is the predicted inclusion rate for state in time period 1 and $\overline{X}_2\hat{\beta}^*$ is the predicted inclusion rate for state in time period 2.

The actual inclusion rate equals the predicted inclusion rate plus a difference, D_1 and D_2, for time periods 1 and 2 respectively.

$$\overline{Y}_1 = \overline{X}_1\hat{\beta}^* + D_1$$
$$\overline{Y}_2 = \overline{X}_2\hat{\beta}^* + D_2$$

The actual difference in inclusion rates, $\overline{Y}_2 - \overline{Y}_1$, can be partitioned in a portion explained changes in controls, $\left(\overline{X}_2\hat{\beta}^* - \overline{X}_1\hat{\beta}^*\right)$, and a portion not explained, $\left(D_2 - D_1\right)$:

$$\underbrace{\overline{Y}_2 - \overline{Y}_1}_{\text{Difference in inclusion rates}} = \underbrace{\left(\overline{X}_2\hat{\beta}^* - \overline{X}_1\hat{\beta}^*\right)}_{\text{Portion explained by differences in controls}} + \underbrace{\left(D_2 - D_1\right)}_{\text{Portion not explained by differences in controls}}$$

The change measure factors out the portion explained by differences in control variables and, hence, is

$$Change = D_2 - D_1 = \left(\overline{Y}_2 - \overline{X}_2\hat{\beta}^*\right) - \left(\overline{Y}_1 - \overline{X}_1\hat{\beta}^*\right).$$

In addition to providing measures of change in inclusion rates over time, we provide a context for this change by comparing states' inclusion rates on NAEP in the initial period. We refer to this as the measure of the *starting point* for each state. Even when we hold constant the different types and severities of disabilities and the different accommodations offered by the states for their own state assessments, not all states start with the same inclusion rate of SDs on NAEP. We expect to observe less change in NAEP inclusion rates in states that initially include SDs at relatively higher rates than other states. Hence, the starting point measure is intended to be a context for understanding the change measure. The use of the starting point measure vis-à-vis the change measure is discussed in detail below.

SUMMARY

Students with disabilities with different characteristics are included at different rates and the distribution of such characteristics differs across states and across time. Hence, the rate of inclusion of SDs on NAEP is expected to vary across states and across time. At the same time, because of NCES's efforts, states are improving their procedures for including and accommodating SDs on NAEP. We estimate an expected rate of inclusion for each state on the basis of its distribution of SD characteristics. Using this predicted inclusion rate, we then partition the total change in inclusion rates over time into a portion explained by observed changes in the distribution of SD characteristics (i.e., the expected change) and into a remaining portion attributable to other factors (e.g., NCES's efforts). If a state's change in the inclusion rate for SDs on NAEP is greater than the change expected (i.e., due to changes in the distribution of SD characteristics), it is considered a positive change or *progress.* The measure of change developed here provides NCES with an indicator for how its efforts and other factors relate to state-level NAEP inclusion rates holding SD characteristics and state assessment accommodations constant.

Two approaches were developed for applying this methodology. Both approaches were tested using 2003 and 2005 NAEP data. In this report, they are applied to 2005 and 2007 NAEP data to look at changes in state-level inclusion rates from 2005 to 2007. Additionally, we develop a method for comparing inclusion rates across states at a given point to provide context for the measure of change.

The concepts and measurement methods applied in these analyses are limited by existing data. NAEP data about students' disabilities do not have the level of detail necessary to create absolute rules for inclusion of SDs (i.e., a normative determination of whether any given student should be included). The concepts for measuring improvement in inclusion rates are relative to a set benchmark, the predicted inclusion rate, which is intended to be a *point of reference*, not a *goal*. The measures are relative in content (e.g., "NAEP's inclusion rate in State *X* is *higher/lower* than the benchmark"), not normative (e.g., "NAEP's inclusion rate in State *X* is a *better/worse* inclusion rate than it *should be*").

This report is limited to the discussion and application of methods for measuring change in state-level inclusion rates. Not included here are discussions of the explanations, other than methodological, behind reported results or the implications of these results for policy.