



## ETS Research Spotlight

A Summary of Applications That Use Student Academic Growth  
for K–12 Accountability and Implications for Students With Disabilities

ISSUE 4, MARCH 2011

# Foreword

## ETS Research Spotlight ISSUE 4 MARCH 2011

Editor: Jeff Johnson

Contributors: Heather Buzick,  
Cara Laitusis

Reviewers: Dan Eignor, Wendy Yen,  
Dianne Henderson-Montero, Ed Shea

Copy Editor: Eileen Kerrigan

Layout: Claude M. DeProspo

Cover Design: George Barrett

The views expressed in this report are those of the authors and do not necessarily reflect the views of the officers and trustees of Educational Testing Service.

Copies can be downloaded from:  
**[www.ets.org/research](http://www.ets.org/research)**  
**March 2011**  
**Research & Development Division**  
**Educational Testing Service**

To comment on any of the articles in this issue, write to ETS Research Communications via e-mail at:  
**[RDWeb@ets.org](mailto:RDWeb@ets.org)**



At ETS, technical quality is the primary consideration in all of the tests and products that we create. That means asking questions such as these in our Research & Development efforts:

- Are the test's results comparable no matter when or where it was administered?
- Does the test measure what we think it measures?
- Is it valid to use the results for their intended purposes?
- Did the test's design allow all test takers a fair opportunity to demonstrate what they know or can do in the area being measured?

These questions — which, by the way, are examples of only the most general questions one could ask related to technical quality in testing — are especially important when people with disabilities take the test. This issue of ETS Research Spotlight provides a sample of the research that ETS is conducting in this area.



This issue's Featured Research Synopsis summarizes two recent reports by ETS research scientists Heather Buzick and Cara Laitusis. In their work, Buzick and Laitusis examined the challenges associated with measuring year-to-year growth for students with disabilities — an important topic, given the increasing interest that educational policymakers are showing in the use of growth in measures of educational accountability and the large proportion of public school students who receive special-education services.

Pages 6–7 of this issue contain a list of recent publications that ETS researchers have authored related to the subject of assessments for people with disabilities, as well as a list of selected current research projects aimed at:

- developing alternative testing formats, such as computer-voiced test versions, for test takers who are blind or visually impaired
- using evidence-centered design to make tests more accessible to students with disabilities
- examining the comparability and predictive validity of admissions tests
- determining the impact of different testing accommodations, such as extended testing time and read-aloud delivery, on the validity of test results

If you'd like to learn more about the research we conduct in these and other areas in order to support and improve assessments, visit us on the web at **[www.ets.org/research](http://www.ets.org/research)**.

A handwritten signature in dark ink, appearing to read "Ida Lawrence".

**Ida Lawrence**  
Senior Vice President  
Research & Development

## FEATURED RESEARCH SYNOPSIS

# An Overview of Applications That Use Student Academic Growth for K–12 Accountability and Implications for Students with Disabilities\*

**\*Editor's note:** This is a synopsis of two longer reports that appeared in other forums. The reference lists and technical details regarding this research appear in the original works, which are:

- 1) Buzick, H. M., & Laitusis, C. C. (2010). *Using growth for accountability: Measurement challenges for students with disabilities and recommendations for research*. *Educational Researcher*, 39, 537–544.
- 2) Buzick, H. M., & Laitusis, C. C. (2010). *A summary of models and standards-based applications for grade-to-grade growth on statewide assessments and implications for students with disabilities* (ETS Research Report No. RR-10-14). Retrieved from: [http://www.ets.org/research/policy\\_research\\_reports/rr-10-14](http://www.ets.org/research/policy_research_reports/rr-10-14)

## Introduction

Local, state, and federal policymakers have shown increased interest in using student academic growth on statewide annual K–12 assessments to support decisions and actions that hold schools and teachers accountable for providing students with a high-quality education. Organizations that advocate for students with disabilities also have become more interested in the measurement of growth and its use for teacher and school accountability. The use of growth in accountability is particularly important for students with disabilities, who are receiving standards-based instruction and participating in K–12 accountability assessments in greater numbers, but are still performing, on average, well below their nondisabled peers. While the use of growth holds promise for students with disabilities, obtaining longitudinal measures of their academic performance on annual accountability assessments — that is, measures of academic performance across two or more grades — is complex. To help states use the academic growth

of students with disabilities appropriately, we summarize uses of growth in accountability and models that states currently use, identify practical challenges in measuring and modeling growth for students with disabilities, and provide recommendations for research to find solutions to the challenges of using longitudinal measurements from this subpopulation.

## Purpose

In the United States, the Elementary and Secondary Education Act — a key component of national education policy — is due for reauthorization by lawmakers at the federal level. Meanwhile, the states, which have jurisdiction over most education policy decisions, are in the early stages of working together to develop common assessment systems. It is in the context of these and other discussions of U.S. education policy that policymakers, measurement researchers, and advocates for individuals with disabilities are discussing the use of student academic growth. This synopsis summarizes two articles by the authors that:

- provide information on how student academic growth, as measured by annual statewide K–12 assessments, is modeled and used;
- describe challenges with using longitudinal data from students with disabilities; and
- discuss recommendations for research to help ensure that decisions based on growth measures are meaningful and appropriate for students with disabilities.

## A Summary of Models for Academic Growth

Models for student academic growth can be used for federal and state school accountability (i.e., growth models) and for evaluations of teachers and schools for educational effectiveness (i.e., value-added models). The federal Growth Model Pilot Program, which started in 2005, has permitted states to consider students who are on track to proficiency as proficient for the purpose of making school accountability decisions.

States use different approaches to growth modeling, but all take into account individual student test scores from the previous one or more years when comparing current or projected scores to a policy-driven growth target. States aggregate individual student growth across schools, districts, and the entire state, and they use it to hold all students to the same standard of being proficient by a fixed point in time.

We can categorize the growth models that states currently use for school accountability in several different ways. Models differ on whether the scores that are used to make decisions are observed or predicted; whether the models are descriptive or statistical; how growth targets are calculated; the chosen time horizon; and the final grade in the model. Examples of growth models include value tables, gain scores, linear regression, projection models, and student growth percentiles. Some states also use or are considering using value-added models as a measure of teacher effectiveness. Value-added models differ from state growth models in that they take into account student demographic information (e.g., race and socioeconomic status) and estimate the contribution of teachers (or schools) to student academic growth.

*There are several challenges to measuring growth for this heterogeneous subpopulation that may affect decisions and actions that are based on student academic growth.*

## Measurement Challenges for Students with Disabilities

Students with disabilities are a heterogeneous subpopulation (e.g., different disability subtypes impact learning and test taking in different ways, testing conditions may differ due to the use of testing accommodations and modifications, and some students take modified or alternate assessments). There are several challenges to measuring growth for this heterogeneous subpopulation that may affect decisions and actions that are based on student academic growth. We identified the following five challenges:

- a) use of and changes in testing accommodations and modifications;
- b) a large percentage of students performing significantly below grade level;
- c) tracking and linking of student scores across testing programs;
- d) low-incidence disability subgroups and changing disability classifications; and
- e) the psychometric properties of alternate assessments.

*States use different approaches to growth modeling, but all take into account individual student test scores from the previous one or more years when comparing current or projected scores to a policy-driven growth target.*

The identification of these five challenges highlights the need for psychometric research to help ensure that high-stakes decisions based on growth measures for students with disabilities are appropriate and that score interpretations have their intended meaning. If states are going to continue to use growth models and introduce value-added models for school and teacher accountability, it is imperative to empirically evaluate the impact of testing accommodations on growth measures, the impact of test difficulty on growth measures, and the longitudinal characteristics of the population of students with disabilities.

## Potential Solutions and Areas for Future Research

While some of the measurement challenges described above represent logistical challenges (e.g., linking data), others have potential solutions that will require additional research. Changes in accommodation use across years may affect a student's measured academic growth across two or more years. Possible solutions include accounting for accommodation use in the model or providing evidence that some accommodations do not affect interpretations from some growth-based applications. It is important to conduct research on the impact of low reliability and low precision of proficiency estimates on results from such models, since growth-based applications use multiple test scores that are likely to be measured with higher error for students with disabilities relative to the general population. In addition, research on the impact that low initial performance has on academic growth can potentially help inform state accountability growth targets.

The following are some specific questions about the longitudinal characteristics of the subpopulation of students with disabilities that can be answered empirically:

- a) What are the characteristics of students excluded from growth models and value-added models (e.g., do early exit disability subtypes change the average proficiency level of the subgroup of students with disabilities in later grades)?
- b) How much and in what ways does disability classification change across years in longitudinal databases?
- c) To what extent do students with disabilities move between general and alternate assessments across years?

Answers to such questions would inform follow-up research on, for example:

- statistical approaches for dealing with exclusion;
- substantive and empirical research on the shape of growth trajectories by disability subtype;
- the accuracy of growth-based decisions by disability subgroup; and
- best practices for linking scores between assessments (e.g., alternate and general).

Furthermore, delineating the content being measured on general and alternate assessments would help inform the feasibility of measuring the growth of students with disabilities who switch from one assessment to another over time. Research also is needed on the feasibility of using value-added models in the presence of coteaching, which occurs primarily for students with disabilities who are in a general education classroom.

*Assessment developers and policymakers should consider the complexities of using growth measures in accountability as they develop new policies and assessments and as states create future longitudinal data systems.*

Assessment developers and policymakers should consider the complexities of using growth measures in accountability as they develop new policies and assessments and as states create future longitudinal data systems. While students with disabilities should continue to be included in growth models for accountability purposes, research is needed to address the challenges

associated with obtaining longitudinal measurements from this subpopulation. Growth-based interpretations can add to the body of evidence about student learning and performance and about the contribution of teachers and schools to student academic progress. Conducting empirical research to understand growth measures for students with disabilities and the consequences of their use will contribute to the reasonableness of relying on such measures to make meaningful and appropriate decisions.

## References

See the original reports, cited in the editor's note on page 3, for a list of references that the authors used when conducting and reporting on the research described in this synopsis.

### About the Authors

**Heather Buzick** is an associate research scientist at ETS.

**Cara Laitusis** is a senior research scientist at ETS.

Both authors work in the Foundational & Validity Research area of the Research & Development Division.

## You May Also Be Interested In ...

*Below is a selected list of recent ETS-authored publications related to the assessment of people with disabilities, as well as some ongoing projects that our staff members are leading or contributing to.*

### PUBLICATIONS

---

- Cook, L. L., Eignor, D., Sawaki, Y., Steinberg, J., & Cline, F. (2010). Using factor analysis to investigate accommodations used by students with disabilities on an English-language arts assessment. *Applied Measurement in Education*, 23(2), 187–208.
- Cook, L. L., Eignor, D., Steinberg, J., Sawaki, Y., & Cline, F. (2009). Using factor analysis to investigate the impact of accommodations on the scores of students with disabilities on a reading comprehension assessment. *Journal of Applied Testing Technology*, 10(2). Retrieved from <http://www.testpublishers.org>
- Hansen, E. G., Mislevy, R. J., & Steinberg, L. S. (2008). *Evidence-centered assessment design for reasoning about accommodations for individuals with disabilities in NAEP reading and math* (ETS Research Report No. RR-08-38). Princeton, NJ: ETS.
- Hansen, E. G., Shute, V. J., & Landau, S. (2010). An assessment-for-learning system in mathematics for individuals with visual impairment. *Journal of Visual Impairment and Blindness*, 104(5), 275–286.
- Laitusis, C. C. (2010). Examining the impact of audio presentation on tests of reading comprehension. *Applied Measurement in Education*, 23(2), 153–167.
- Laitusis, C. C., Cook, L. L., Buzick, H. M., & Stone, E. A. (2011). Adaptive testing options for accountability assessments. In Russell, M. (Ed.), *Assessing students in the margins: Challenges, strategies, and techniques*. Charlotte, NC: Information Age Publishing.
- Laitusis, C. C., Maneckshana, B., Monfils, L., & Ahlgrim-Dezell, L. (2009). Differential item functioning comparisons on a performance-based alternate assessment for students with severe cognitive impairments, autism and orthopedic impairments. *Journal of Applied Testing Technology*, 10(2). Retrieved from <http://www.testpublishers.org>
- Steinberg, J., Cline, F., Ling, G., Cook, L. L., & Tognatta, N. (2009). Examining the validity and fairness of a state standards-based assessment of English-language arts for deaf or hard of hearing students. *Journal of Applied Testing Technology*, 10(2). Retrieved from <http://www.testpublishers.org>
- Stone, E., Cook, L. L., Laitusis, C. C., & Cline, F. (2010). Using differential item functioning to investigate the impact of testing accommodations on an English-language arts assessment for students who are blind or visually impaired. *Applied Measurement in Education*, 23(2), 132–152.

## Initiatives and Projects

---

### **Technology Assisted Reading Assessment (TARA)**

*Funding Source: U.S. Department of Education*

TARA is one of three projects funded as part of the National Accessible Reading Assessment Projects (NARAP) and includes a program of research and development to improve reading assessments for students with visual impairments or blindness. TARA is examining the properties of existing assessments for students with visual impairments and is developing an assessment of reading with a particular focus on independent technology assisted reading.

### **Feedback-and-Revision on Alternate Assessment Based on Modified Achievement Standards in Mathematics**

*Funding Source: U.S. Department of Education*

The aim of this project is to develop a new test administration format for alternative assessment based on modified achievement standards in mathematics that allow students to receive immediate feedback on their answers to test items and revise their answers for partial credit. Researchers are examining whether providing immediate feedback and opportunity to revise their responses to test items to students with disabilities will improve the psychometric quality of the assessment.

### **Projects Within the ETS Strategic Research Initiatives**

*Funding Source: ETS Research Investment*

ETS is undertaking foundational research that aims to foster better measurement of growth and well-designed computer-based assessments for students with disabilities. This work includes projects that focus on:

- measuring and modeling growth for students with disabilities;
- identifying factors critical to carrying out fair and valid item-level adaptive testing for students with disabilities; and
- identifying the characteristics of accessible computer-based testing platforms.

## About ETS

---

At nonprofit ETS, we advance quality and equity in education for people worldwide by creating assessments based on rigorous research. ETS serves individuals, educational institutions and government agencies by providing customized solutions for teacher certification, English language learning, and elementary, secondary and post-secondary education, as well as conducting education research, analysis and policy studies. Founded in 1947, ETS develops, administers and scores more than 50 million tests annually — including the TOEFL® and TOEIC® tests, the GRE® tests and *The Praxis Series*™ assessments — in more than 180 countries, at over 9,000 locations worldwide.

---

756860



*Listening. Learning. Leading.*®

[www.ets.org](http://www.ets.org)