No Child Left Behind and the Testing of Deaf Students

Michael A. Karchmer
Gallaudet Research Institute
and
Department of Educational Foundations and Research

First Wednesday Research Seminar
Gallaudet University
March 3, 2004
Goals of the Presentation

- To briefly give a national picture of how well deaf and hard of hearing students read in comparison to their hearing peers
- To describe the accountability provisions found in the No Child Left Behind Act (NCLB)
- To list some oddities and state-to-state inconsistencies concerning testing
- To share some personal perspectives
Before considering the testing provisions of the NCLB, let’s review some GRI research concerning the academic achievement of deaf students.

Here are data from the 1996 national norming for deaf and hard of hearing students of the 9th edition of the Stanford Achievement Test (SAT-9).
The Stanford Achievement Test

- **Norm-referenced**
  - Scores are relative to other students, not absolute measurements
  - Published by Harcourt Educational Measurement; now in 10th edition

- Tests in many areas; 8 levels (from 1st – 9th grades)

- Content reflects generalized national standards
  - Because of this, it can suffer misalignment with individual state standards

- Permits comparison among groups of students as to relative ability levels
  - Including with deaf and hard of hearing students
Characteristics that have an impact on the achievement of hearing students likewise with deaf and hard of hearing students, though with some twists:

- Race/ethnic status; languages other than English used at home; socioeconomic status
- Similarly, presence of additional conditions negatively predicts achievement
Some Basic Findings

- Degree of hearing loss is inversely related to reading achievement.
- Early use of ASL or contact with the sign language community has a positive impact on achievement.
  - Deaf children with deaf parents are known to read better than children with hearing parents, all else equal.
Comparisons of Deaf and Hearing Students by Age

- One persistent finding across many studies is that by age 15, typical deaf students read about 6 grade equivalents lower than their same age hearing peers.

- Let’s look at the data in more depth as we compare scores of hearing and deaf groups:
  - We need to consider both the central tendency (median scores), but how scores in each age group are dispersed.
Reading Comprehension for Deaf & Hard-of-Hearing/ Hearing Students Age 8/ 2nd Grade through Age 15/ 9th Grade: SAT-9

Note: Bottom of each line is first decile; top is 9th decile; medians also shown.
Comparisons of Deaf and Hearing Students by Age: Three Findings

- First, median score increases for each cohort year for both the hearing and the deaf and hard of hearing groups.
Comparisons of Deaf and Hearing Students by Age: Three Findings

- Second, the median score is consistently higher for hearing students than for deaf and hard of hearing students.

- The difference is fairly constant across cohorts.
Third, the dispersion (distance between the 1st and 9th deciles) for hearing students decreases for successive cohorts, up to age 12/grade 6, while the dispersion for deaf and hard of hearing students increases.
Comparisons of Deaf and Hearing Students by Age

- This implies that the higher performing deaf and hard of hearing students are likely to be making the same amount of annual achievement growth as hearing students (though the level of performance of only a few deaf students exceeds typical hearing students) while the lower performing deaf and hard-of-hearing students are further and further behind.
What about the relationship between instructional setting and achievement?

- There are differences in academic achievement by instructional setting, but it’s very difficult to attribute them to the programs themselves.
- It is more likely that achievement differences reflect sorting and selecting decisions done at placement more strongly than instructional efficacy.
The metaphors of education have changed and so has the role of testing

- A shift has occurred over the past 15 years from talking about the process of education to a focus on the outcomes of schooling
  - We no longer want to graduate students with minimum competencies; we want students “to meet high standards reflective of skills needed to compete in a global economy”*

- What’s propelled this in part has been a widespread belief that our schools are failing us, that standards for graduation were low, and “social promotion” was out of control

*Helwig et al. (1999)
The metaphors of education have changed and so has the role of testing

- These calls for school improvement were bipartisan
  - The Clinton administration called for greater school accountability
  - He called for appropriate use of tests and other indicators of academic performance in determining whether students should be promoted (see Hauser, 2001)
Participation by Disabled Students in Testing Programs

- The 1997 revision of IDEA is clear in its expectation that all students with disabilities participate in state testing programs.
So Now We Have the No Child Left Behind Act*

- This law calls for all students to meet increasingly high academic standards as gauged by statewide tests.
  - Until that time in 2014 when **100% of all** students will be judged “proficient” in key academic areas.

- The law includes all students, measured together and according to membership (1) in major racial/ethnic groups; (2) economically disadvantaged students; (3) students with limited English; and (4) students with disabilities.

*Actually the 2001 reauthorization of the Elementary and Secondary Assistance Act
So Now We Have the No Child Left Behind Act

- When any of these groups’ test performance falls short, the school must improve by providing more qualified personnel and better programs and services
  - Or face an escalating series of consequences
- Consequences of failure ripple from student to school to LEA to State

Let’s turn now to the elements of a statewide testing program, but before we do I want you to keep in mind three important things about the Law:
Three Important Things to Keep in Mind When Discussing NCLB

- Deaf and hard of hearing students are not mentioned in the Law; they are simply lumped together with all other students with disabilities

- Each state has a different implementation of NCLB
  - Different state standards, tests, procedures, test accommodations, criteria for showing compliance
Three Important Things to Keep in Mind When Discussing NCLB

- NCLB involves high stakes testing, but not necessarily of students
  - Students take tests to measure the progress of subgroups
  - In order to make decisions about the quality of schools and school districts (LEAs)
Elements of a Statewide Accountability System

- NCLB requires each state to have challenging academic content and achievement standards grade by grade
- In at least the following areas:
  - Reading/ language arts
  - Mathematics
  - Science (beginning 2005 – 06)
Elements of a Statewide Accountability System

- A state must design and implement a statewide testing system for measuring how well students have mastered the curriculum content.

- A State must design or choose a test of at least:
  - Reading/Language Arts standards and Math standards (now) and Science (2007-08)

- States use a variety of tests developed by themselves or under contract with companies:
  - Most are “criterion-referenced”, though at least 10 use “norm-referenced” tests like the Stanford and adapt it.
  - Typically involve both “selected” and “constructed” response.
Elements of a Statewide Accountability System

- Alternative assessment instruments have to be available
  - Meant to be used with very small number of students for whom the regular tests aren’t appropriate
  - Some confusion: are these meant to be a measure of same standards with a different instrument or measure of alternative standards for cognitively impaired students???

- Every two years, samples of 4th and 8th grade students must take the federal National Assessment of Educational Progress
  - The only means of comparing state to state performance
Elements of a Statewide Accountability System

How often are students to be tested?

- For reading/language arts:
  - Yearly testing in grades 3 – 8, plus at least once in 10 – 12 (beginning 2005 – 06)

- For science:
  - Beginning 2007 – 08, at least once in each grade span 3 – 5, 6 – 9, 10 – 12

Test results must be stated according to at least 3 performance levels: “Basic,” “Proficient,” and “Advanced” that reference the standards

- Levels are defined differently in each state
- Most states have more than 3 and are called different names
Elements of a Statewide Accountability System

- Who is tested? **All students in virtually every school setting must participate in the testing**
  - If they have been in school all year
  - Only a few are excused
    - Example: Parental religious objections
  - Until last week, even students with limited English proficiency must take the tests
Elements of a Statewide Accountability System

- Testing accommodations must be available for those who need them
  - Adaptations of the testing procedures or test meant to “level the playing field” [i.e., interpreting the instructions]
- Generally, no “out of level” testing permitted; students must take age/grade appropriate test
Elements of a Statewide Accountability System

- **Reporting of Results:**
  - When? Beginning of school year to permit needed changes in schools and LEAs; also to inform education
  - To whom:
    - Individual results in understandable language to parents
    - Individual school, LEA, and State publicly available “report cards” – **showing percentages of students scoring at proficient or above**
      - Minimum contents of report cards are specified in the Law
      - Single out schools in need of improvement, etc.
Elements of a Statewide Accountability System

- Reporting of Results (continued):
  - Results by school, LEA, and state have to be “disaggregated” according to the recognized subgroups in a state
    - Each state has different subgroups and the operational definitions of each subgroup may differ
    - Disaggregation is only done for subgroups in a school or LEA that are a certain size (for statistical reliability)
      - Varies widely by state, ranging from 5 to 50
  - Some states use confidence intervals in reporting school and LEA results; others don’t
Key Provision in the Law: AYP

- Remember that the overall goal is that all (100%) students in each state should be able to demonstrate proficiency by 2013 – 14 in all key curriculum areas.

- NCLB recognizes that this is not likely to happen overnight, so has established intermediate goals.
  - Successfully achieving these according to a negotiated timetable is called making “Adequate Yearly Progress” (AYP).
Key Provision in the Law: AYP

- **Who must demonstrate AYP:**
  - All sub-groups within schools
  - Schools
  - LEAs/ School District
  - States must demonstrate AYP

- **Continued failure to demonstrate AYP results in sanctions**
To demonstrate AYP all of the following must happen:

- Meet or exceed some specified target for proficient performance in ALL key areas
  - At any level, each subgroup must meet these in order to demonstrate AYP
    - Or “Safe Harbor”: Reduction of Non-Proficient Students by 10%
- Demonstrate 95% participation, strictly calculated
- Satisfy other indicators:
  - Increase high school graduation rate each year
  - One other at elementary middle school
    - Example: increased attendance
Here are typical AYP Targets by Year for Math (Left) and Reading (Right)

This is for Pennsylvania, but every state is different!
Sanctions and Rewards

- Failure to meet AYP can result in sanctions
  - After two years, a school is identified for improvement, requiring
    - School choice for parents
    - Staff training
    - Supplemental services must be offered
  - Continued failure results in various corrective actions, including replacement of staff, restructuring, or a host of other “remedies”

- Rewards for meeting AYP at least two years in a row
  - Varies by state, but not well spelled out in the Law
Some Oddities in AYP

Suppose Hispanic students fail AYP for math in Year 1, but pass in Year 2. Disabled students in the same school pass Reading in Year 1, but fail in Year 2. Does the school achieve AYP?

ANSWER: NO! All sub-groups must pass all tests both years.

A Hispanic deaf student comes from an economically disadvantaged family. In which subgroup is student counted for AYP purposes?

Answer: THREE subgroups, meaning this person’s test scores carry disproportionate weight.
Huge State Variation by:

- AYP Targets and Timetables
- Performance standards (and labels) acceptable for demonstrating AYP
- The size of reportable subgroups for AYP purposes
  - 5 – 50
- Whether confidence intervals can be used in reporting of scores
- What is allowable as a test accommodation
Personal Perspectives: Some Positives about NCLB

- There’s a lot to admire:
  - It asks each of the states to establish meaningful and high standards of education
  - It aims to measure, then reduce, social disparities
  - It helps provide schools with funding and program support
  - It promotes the idea of exposing students to programs “that have been proven to be effective”
    - “Scientifically-based practices”
  - It requires openness of information and allows for public scrutiny
Personal Perspectives: Some Positives about NCLB

- Of most relevance to us:
  - The law seeks to raise the ante for students with disabilities and the programs that serve them by not treating them as separate and unequal
A Personal Perspective

- That there is so much fantasy and wishful thinking, starting with the name of the Act itself:
  - You saw the test scores I showed for deaf students: Is AYP probable, much less the notion of a 100% pass-rate?
  - It’s mathematically challenged: Take a subgroup at a school struggling to make AYP. It can do so for 12 years straight and still wind up with 32% unsatisfactory. How?
A Personal Perspective

- It’s really negative and heavy on educator blame!
  
  - Example: Check the websites of the various states to see how they comply. NCLB talks about rewards and sanctions. I see the latter; the former is not so clear...
    
    - How will we bring new and qualified teachers into our field when they see what may happen to them?
In regard to the testing of our students (particularly disabled students), I worry that:

- NCLB asks state tests to carry weight that no single measure should carry:
  - Even when the tests aren’t high stakes for a student, they are very high for the school, the district, and the state
  - There is little recognition that tests are not perfect measures of achievement

- It turns the traditional role of testing upside down:
  - Instead of tests designed to support instruction, instruction now is expected to support testing
In regard to the testing of our students (particularly disabled students), I worry that:

- It can encourage narrowing the curriculum by focusing instruction to the test and thereby promoting a tense, joyless type of education
  - Those most in need of joy in education will be most deprived of it

- There will be a lot of unintended consequences of how test scores will be used for the organization of the education of our disabled students, some of them may not be what we want
In regard to the testing of our students (particularly disabled students), I worry that:

- The provisions and goals of IDEA in some ways seems at odds with NCLB
  - IDEA requires the use of appropriate assessment, matched to a student’s individual needs and individually-defined educational goals
  - NCLB places emphasis on testing of performance of grade standards based on a general education curriculum
    - All we may learn from such an exercise is that many of our students aren’t in the ball park!
Where does that leave us?

- Hard to say

- Despite using the rhetoric of science in the law, the insights of research or, for that matter, the standards testing experts have agreed upon about fair testing practices may or may not provide guidance for current assessment practice under NCLB.

- The need to be “accountable,” to end “social promotion” and establish “high standards for all students” overwhelms the discussion.
Where does that leave us?

- Hoping that when wishful thinking collides with reality, we will chart a saner course to accomplish the same commendable purposes.

- In the meantime, let’s do our best as professionals to use tests that can guide practice: the aim should be instructionally supportive tests for all students, not just for those who are on par with grade standards.
Resources


- Check out the extensive website on test accommodations and how different states are implementing NCLB: http://education.umn.edu/NCEO/. A great website that includes a position paper on NCLB and links to other sites is run by the National Conference of State Legislatures: http://ncsl.org. Also, check out the fine website of the Education Commission of the States: http://www.ecs.org.

- To see how tests can be constructed and used to support instruction, see the report developed by the Commission on Instructionally Supportive Assessment, convened by 5 educational associations. Go to http://www.principals.org/pdf/bldg_test.pdf

- An excellent resource explaining NCLB to Principals is the K – 12 Principals Guide to No Child Left Behind, published in 2003 by the Educational Research Service in Arlington, VA.