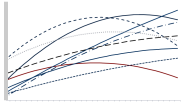
# Research on the Academic Growth of Students with Disabilities and its Implications for Educational Policies and Practices

National Conference on Student Assessment
San Diego, CA
June 23, 2015



NCAASE National Center on Assessment and Accountability for Special Education

*Advancing research on growth measures, models, and policies for improved practice*

# IES's Call for Research to Answer Questions

- In 2010, IES put out an RFA for a National Special Education Research and Development Center on Assessment and Accountability. It stated:

*"At minimum, however, to raise academic achievement for students with disabilities, schools need data that will accurately measure individual student progress from year to year and some means to gauge whether or not individual students are making reasonable progress. Understanding the yearly progress made by students will provide teachers and schools with information necessary to make important instructional and programmatic decisions for students with disabilities. Measuring progress for students with disabilities raises many questions about what is expected for their achievement over the course of a year. The recent emphasis on accountability has raised expectations for the performance of students with disabilities to achieve the same academic standards as their peers without disabilities. It could also be argued however, that by nature of having a disability, students with disabilities cannot be expected to learn at the same rate as their peers without disabilities. Questions remain as to what progress can be expected within a specified time frame."* (IES, 2010, p. 10)

# NCAASE's Role

- As specified by the IES call, NCAASE focuses *"on conducting a program of research that identifies the academic growth trajectories of students with disabilities, and develops and tests practical and relevant methods of accurately measuring academic growth for students with disabilities to be used in accountability systems. The ultimate objective of such work would be to develop assessment methods that schools can use to (1) accurately assess the academic progress of students with disabilities and (2) improve the quality of education provided to students with disabilities to lead to improved student outcomes. …In addition … the … Center will conduct supplementary studies and engage in national leadership activities relevant to assessment of students with disabilities"* (IES, 2010, p.11)

- Goals for this meeting are to
  - Share NCAASE research
  - React to this research by addressing questions such as: What are the implications of this work for educational policies? Educational practices? Future research?

# Research on Growth

Six Research Questions

Two Critical Concepts

# Six Areas for Research

- 1. What is the natural developmental progress in achievement for students with disabilities?
- 2. What models best characterize achievement growth for students with disabilities who are participating in general achievement tests?
- 3. How do various growth models represent school effects for students with and without disabilities, and how do results compare to those derived from status models now in use?
- 4. What are the reliability and validity of estimates of school effectiveness for students with disabilities produced by alternative growth models and how are these estimates influenced by contextual differences among schools and students?
- 5. How do results from different types of interim assessments of students' achievement meaningfully contribute to a model of academic growth for students with disabilities?
- 6. How can information about opportunity to learn and achievement growth be used to enhance academic outcomes for students with disabilities?

# Capacity

- Partnerships at Various Levels
  - States: OR • NC • AZ • PA
  - IHEs: UO (5 faculty • 5 Doc Students) and ASU (3 faculty • 1 Doc Student)
  - Consultants (3) and Advisors (7)
  - IES (NCSER)
- Three Individuals Critical for Success
  - Jackie Buckley (IES) for guidance and support
  - Raina Megert (UO) for contracts and finances
  - Aaron Glasgow (UO) for technology backbone

# Resolution

## Running the numbers – Chris Jordan

http://www.chrisjordan.com/gallery/rtn/#prison-uniforms-set

# Resolution

## Running the Numbers – NCAASE

| Dataset | Years | # Sep. Files/ Years | Students/ File | Total Records |
|---|---|---|---|---|
| Tempe NWEA Dates | 2006-2013 | 8 | 10,000 | 80,000 |
| Tempe Demog & Assessment | 2006-2014 | 9 | 10,000 | 90,000 |
| Students Retained | 2006-2013 | 8 | 100 | 800 |
| Tempe Teacher ID | 2009-2010 | 2 | 7,150 | 14,300 |
| **TOTAL** | | | | **185,100** |
| | | | | |
| **NWEA-PA** | 2010-2012 | 3 | 11,577 | **34,731** |
| | | | | |
| OR Disability codes | 2006-2013 | 9 | 71,835 | 646,515 |
| OR Extended Assess | 2002-2014 | 13 | 10,000 | 130,000 |
| OR RL/M files | 2005-2011 | 1 | 1,500,000 | 1,500,000 |
| OR RL/M files | 2012 | 1 | 450,813 | 450,813 |
| OR RL/M files | 2013 | 1 | 450,813 | 901,626 |
| **TOTAL** | | | | **3,628,954** |
| | | | | |
| PA General Assessment RL | 2006-2012 | 7 | 780,000 | 5,460,000 |
| PA General Assessment Math | 2006-2012 | 7 | 780,000 | 5,460,000 |
| PA Alternate Assessment RL/M | 2008-2012 | 5 | 13,900 | 69,500 |
| **TOTAL** | | | | **10,989,500** |
| | | | | |
| NC RL/M files | 2001-2009 | 9 | 650,000 | 5,850,000 |
| NC Demographic | 2001-2012 | 12 | 650,000 | 7,800,000 |
| NC AYP files | 2010-2012 | 3 | 685,000 | 2,055,000 |
| **TOTAL** | | | | **15,705,000** |
| | | | | |
| AZ RL/M files | 2007-2011 | 1 | | 4,933,142 |
| AZ LEP test scores | 2007-2011 | 1 | | 454,681 |
| AZ Primary Disability | 2007-2011 | 1 | | 349,094 |
| AZ accommodations files | 2007-2011 | 1 | | 1,951,539 |
| AZ Student demo file | 2007-2011 | 1 | | 2,680,084 |
| AZ growth percentiles | 2010-2011 | 1 | | 887,858 |
| **TOTAL** | | | | **11,256,398** |
| | | | | |
| **Grand Total** | | | | **41,799,683** |

# **Resolution** for NCAASE

Findings are at the end of a long logic chain

- From State Accountability Systems
  - Standards based test development process
  - Integrity in data collection systems
  - Assemblage of data with directories
- From NCAASE Researchers
  - Data rendering with… missing data • varying participation rates • time varying changes in categories( e.g., disabilities and/or English language status) • cohort configurations • test and policy changes….

# Mathematics and Reading Growth Across Grades

Ann C. Schulte
Arizona State University

# Cornerstone Study Strand

- Basic questions about the population of students with disabilities and their achievement growth have yet to be answered
- This information needed to form the basis for an accountability system that includes SWDs, but actually captures schools' performance with the students, not construct irrelevant variance
- Select results from three longitudinal studies presented—tracking mathematics and reading growth across 5 years

# Illustrative Complexities

- Identification as a SWD is not stable from year-to-year (Ysseldyke & Bielinski, 2002)
  - *Export* the success stories to general ed
  - *Import* struggling students from general ed
- Specific exceptionality can change across years
- Mobility and grade retentions affect interpretation of outcomes, and SWDs are likely to have elevated levels of both

# THE GRADE 3 LD DIASPORA

| Yr | Gen ED | | | | | OHI | | | | | LD | | | | | EMH | | | | | BEH | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | | | 5,272 | | | | | | | | | | | | | | |
| 2 | 403 | | | | | 154 | | | | | 4,544 | | | | | 85 | | | | | 36 | | | | |
| 3 | 299 | 14 | 85 | 14 | 2 | 4 | 111 | 36 | 2 | 1 | 410 | 85 | 3,927 | 48 | 28 | 3 | 0 | 22 | 59 | 1 | 2 | 2 | 3 | 0 | 29 |

| Yr | Gen ED | | | | | OHI | | | | | LD | | | | | EMH | | | | | BEH | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | | | | | | | | | | | 3,927 | | | | | | | | | | | | | | |
| 4 | 423 | | | | | 60 | | | | | 3,347 | | | | | 43 | | | | | 20 | | | | |
| 5 | 355 | 5 | 57 | 3 | 0 | 4 | 38 | 16 | 0 | 1 | 283 | 45 | 2,667 | 20 | 18 | 1 | 0 | 19 | 23 | 0 | 1 | 1 | 3 | 0 | 14 |

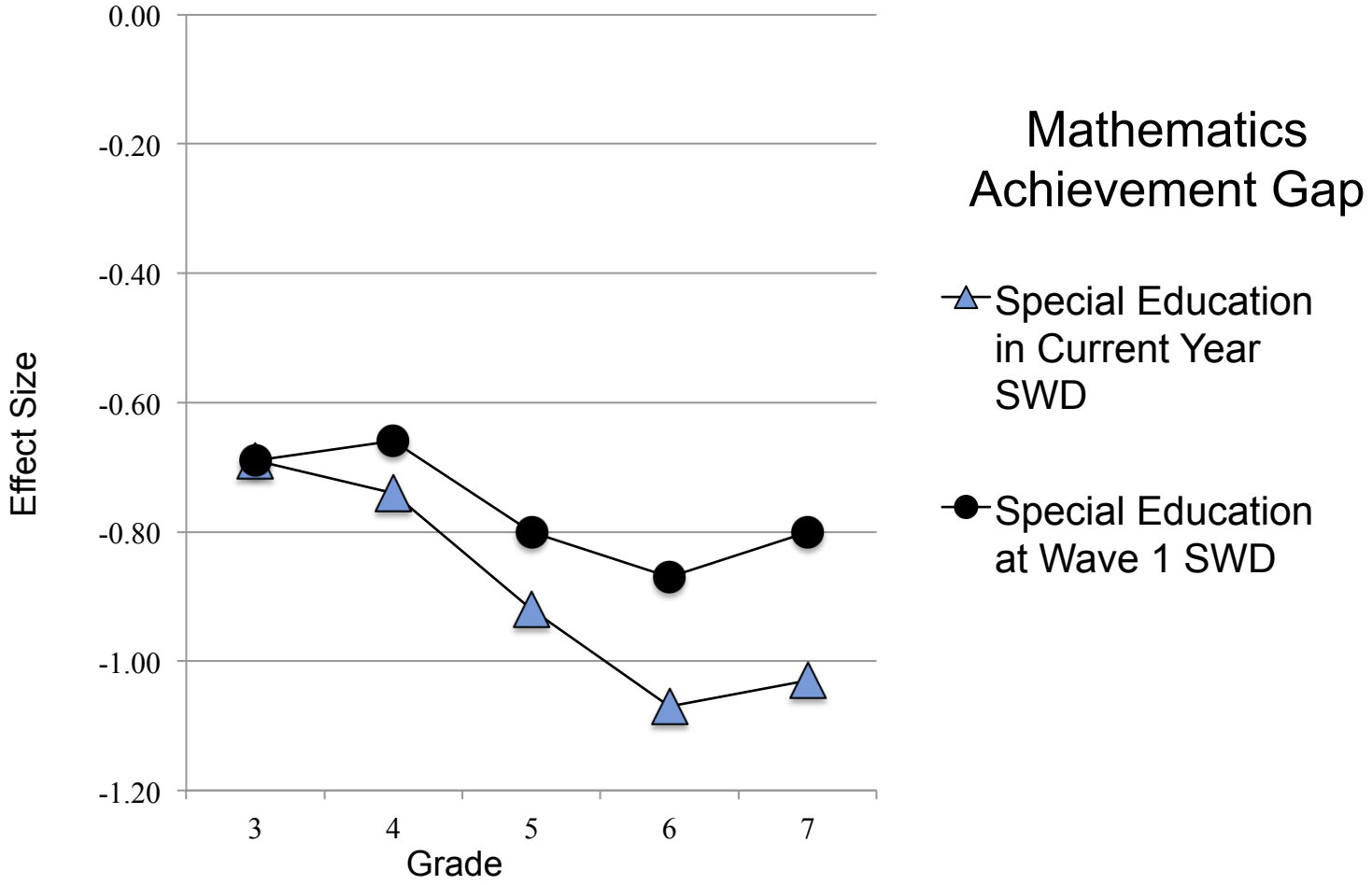
# Study 1: Once, Sometimes, or Always in Special Education

- What is the impact of entrances and exits from special education on portrayal of mathematics achievement gap and growth?
- Cross sectional
  - **Current Year**: As in NCLB, annual determination
- Longitudinal
  - **Wave 1**: SWD or non-SWD at initial data collection time point
  - **Ever in Special Education**: Student presence in special education at any time during grades 3-7
  - **Always in Special Education**

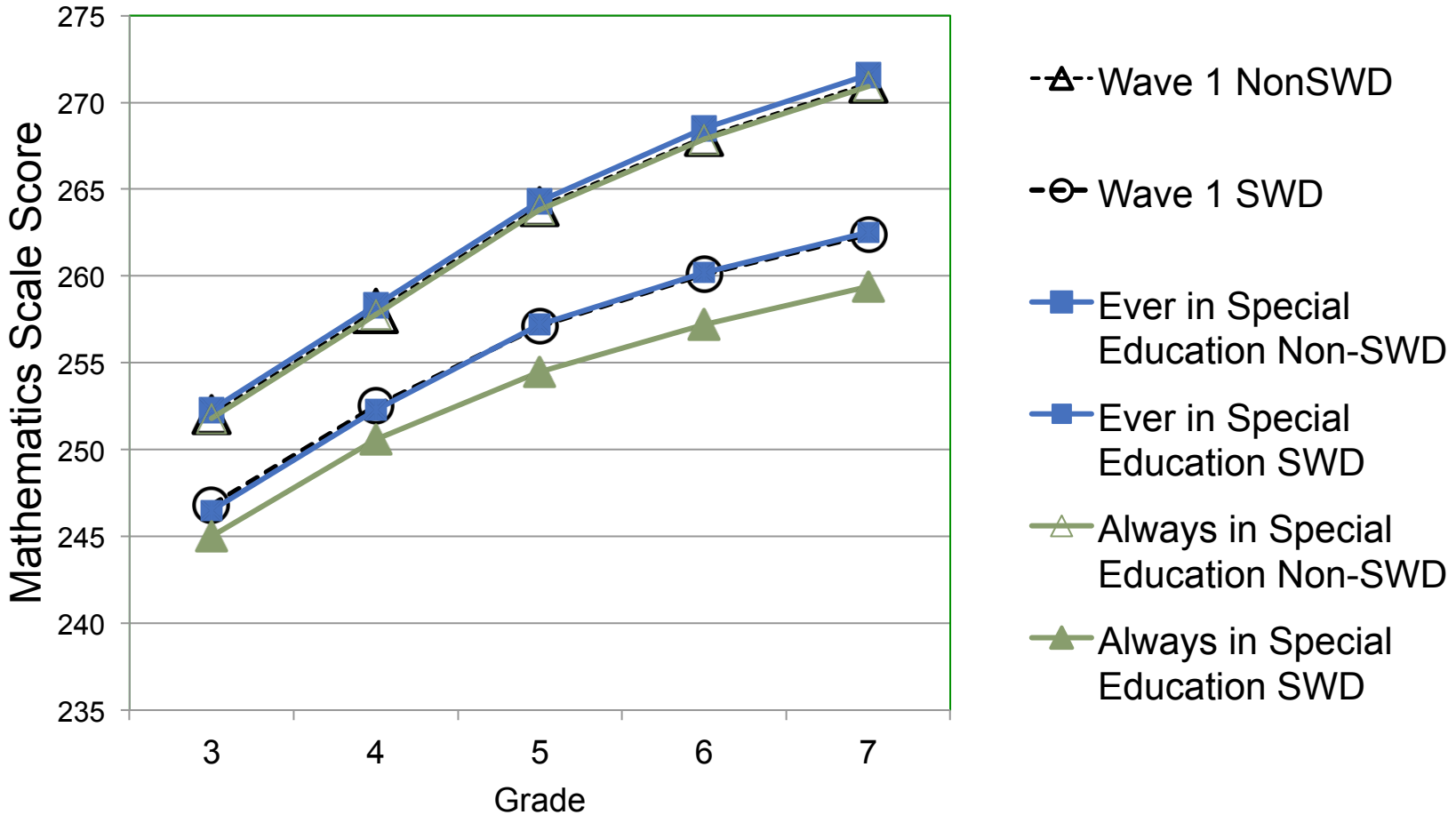
# Special Education Membership Grades 3-7

| SWD Subgroup Identification Method | Percent |
|---|---|
| Current Year | 11.1 to 12.4 |
| Wave 1 | 11.8 |
| Ever in Special Education | 16.1 |
| Always in Special Education | 6.0 |

# Stable Subgroup Membership Matters



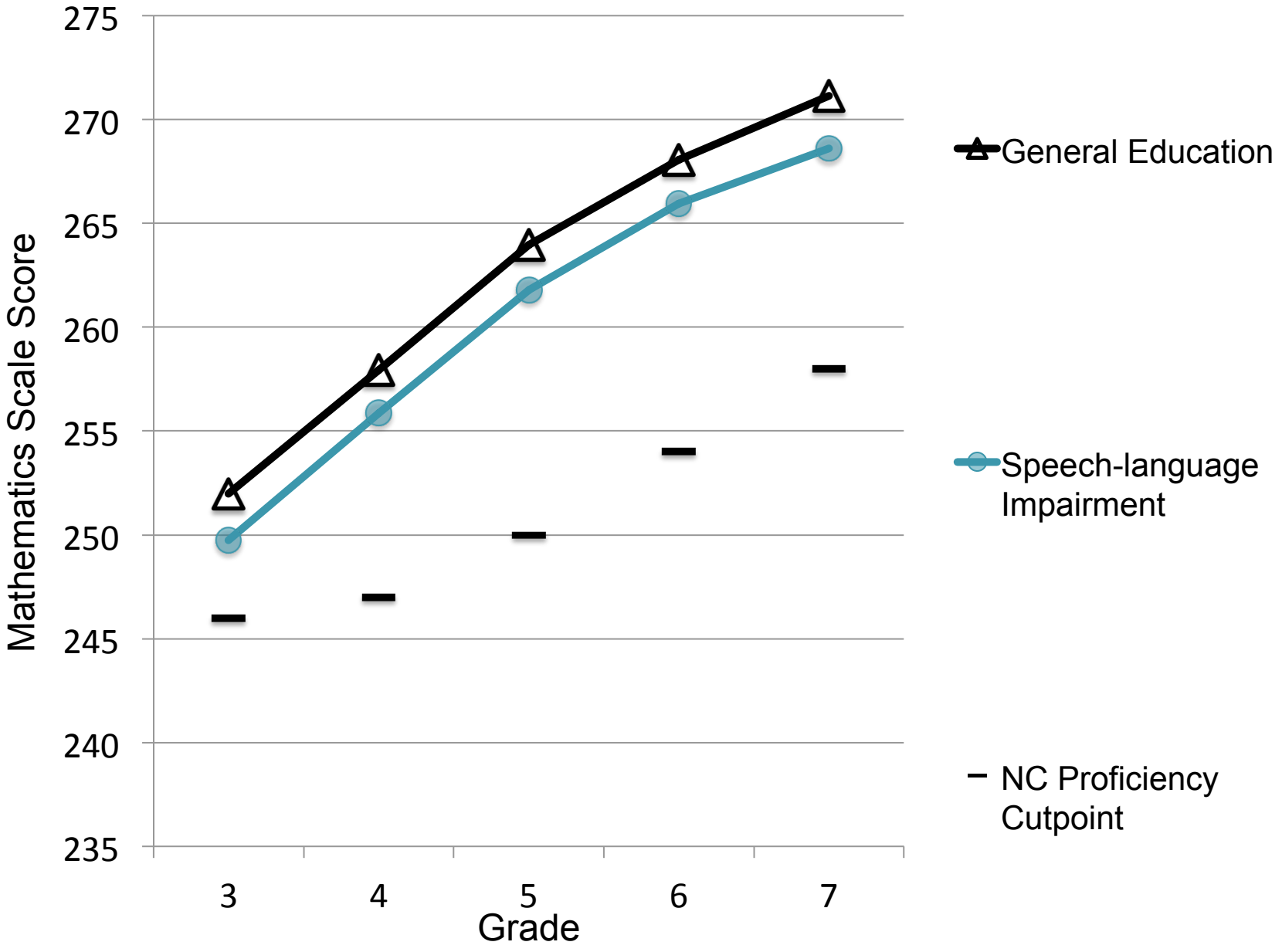
# Observed Means by SWD Identification Method



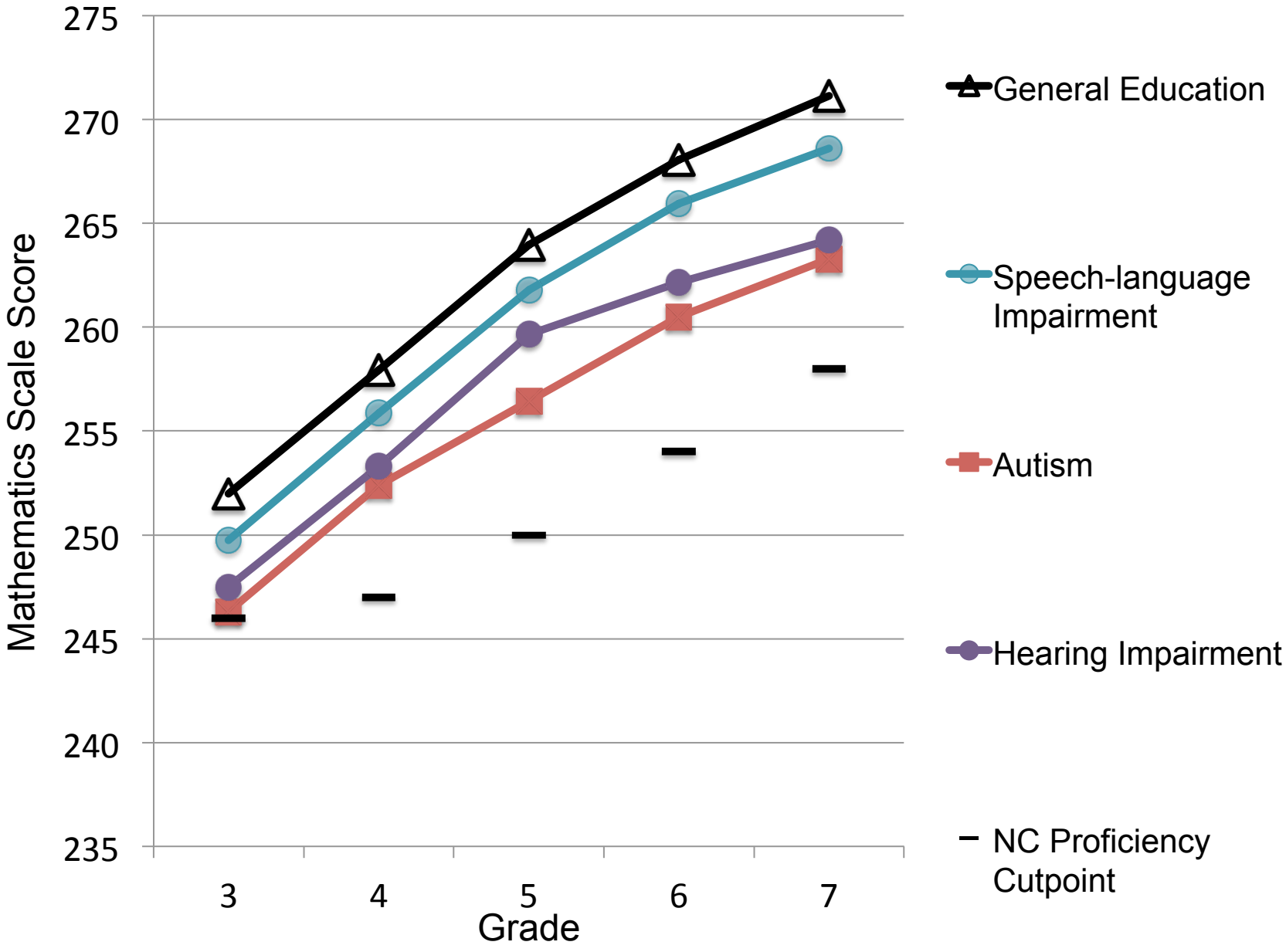
# Studies 2 & 3: Growth Across Grades

- Examine the developmental progress in mathematics and reading comprehension for general education students (GE) and students in specific exceptionality groups on a statewide achievement test
- Two longitudinal cohorts followed across grades three to seven
- Entire state cohorts, N > 100,000; N > 90,000 for analytic samples
- Students never taking general assessment, retained, or where exceptionality category N <100 were excluded

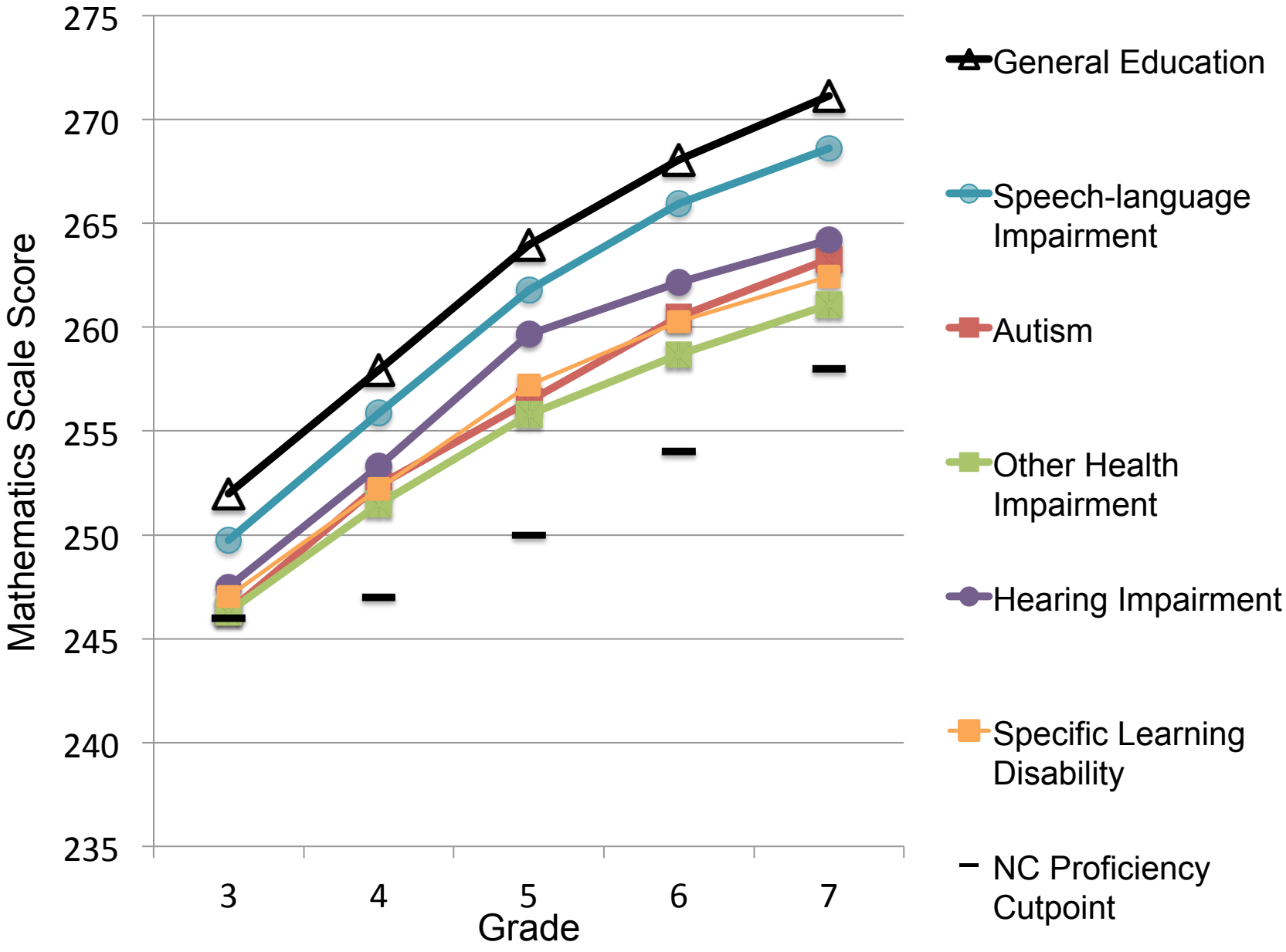
# Mathematics Growth by Exceptionality



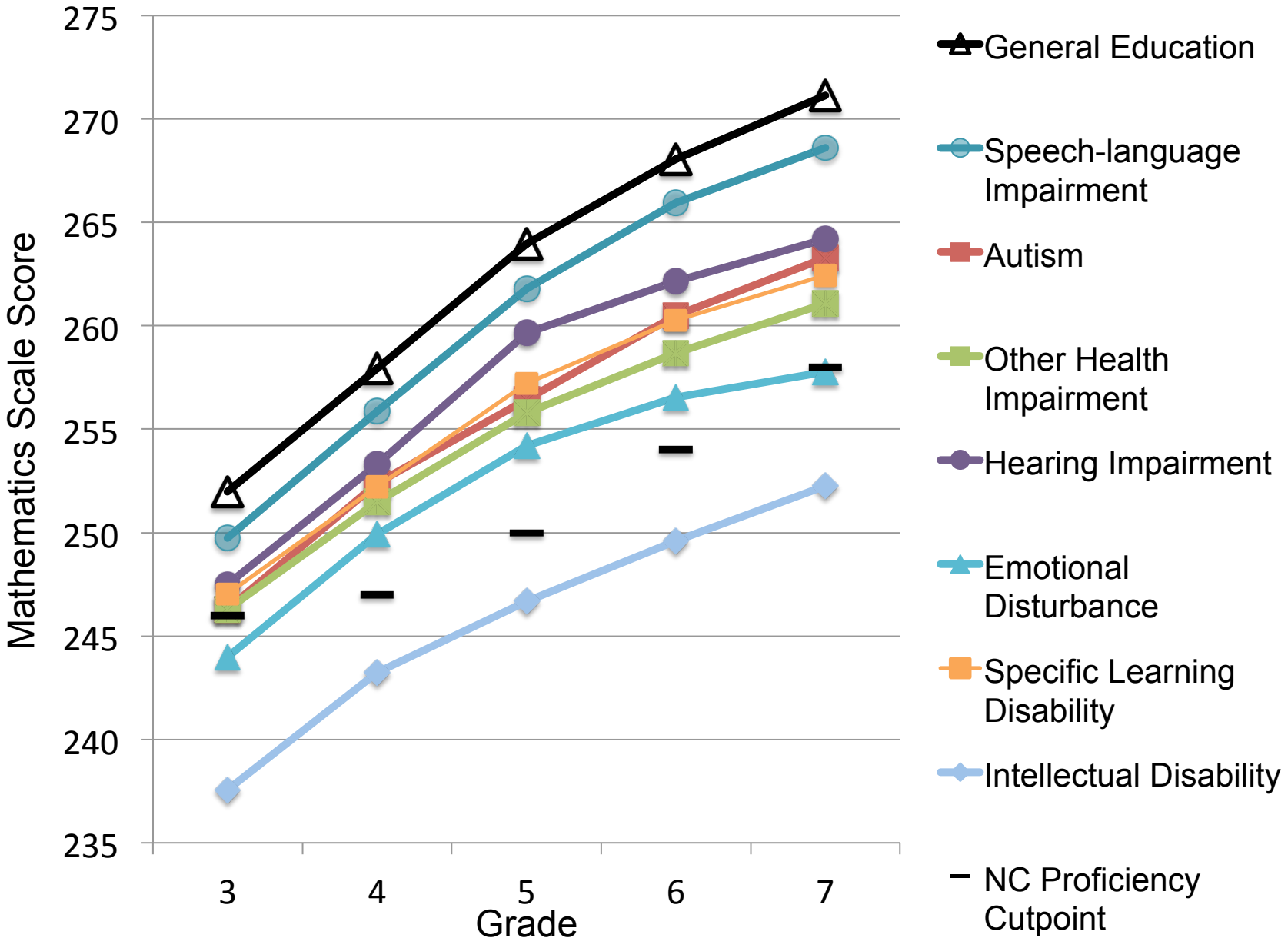
# Mathematics Growth by Exceptionality



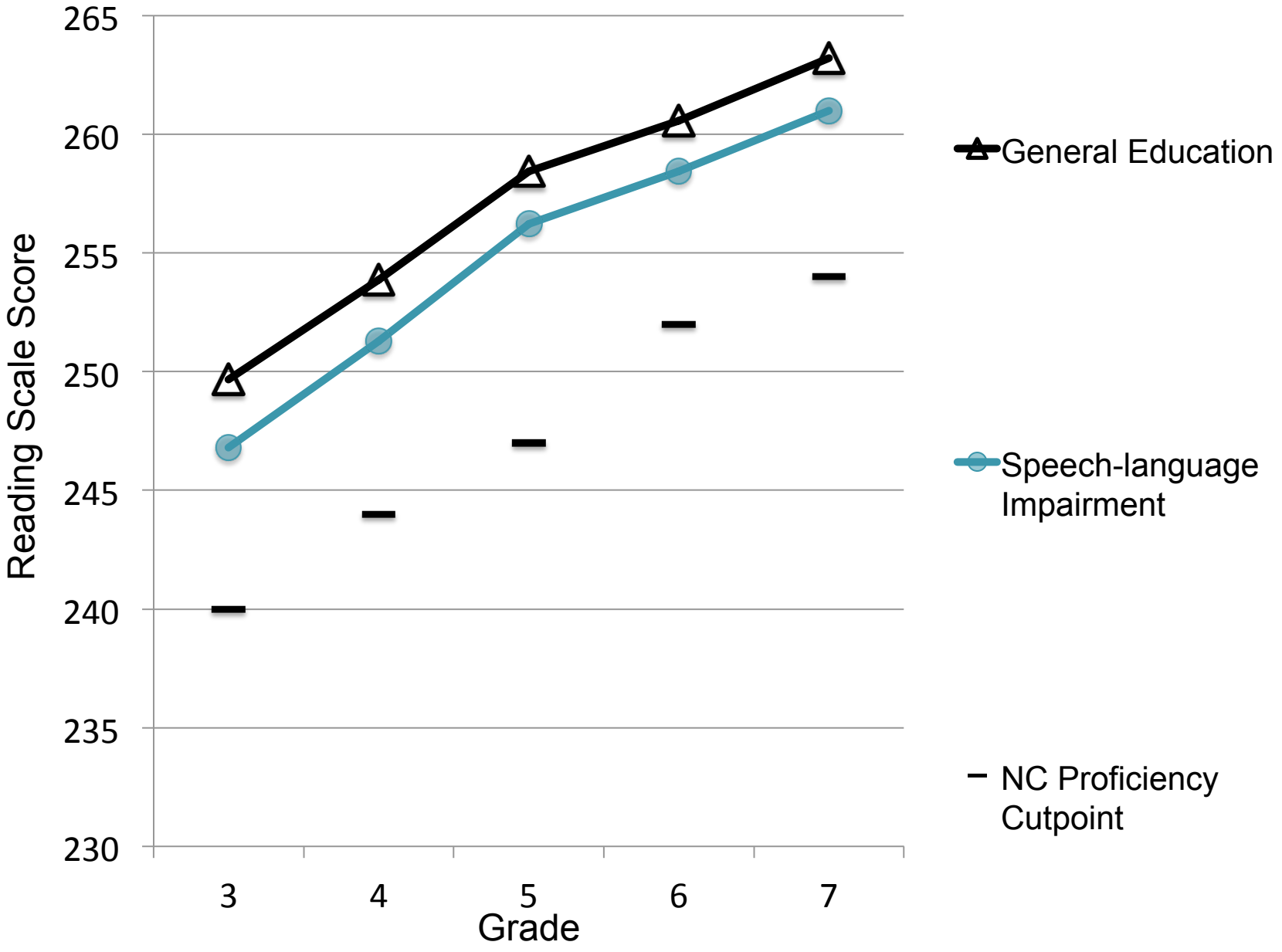
# Mathematics Growth by Exceptionality



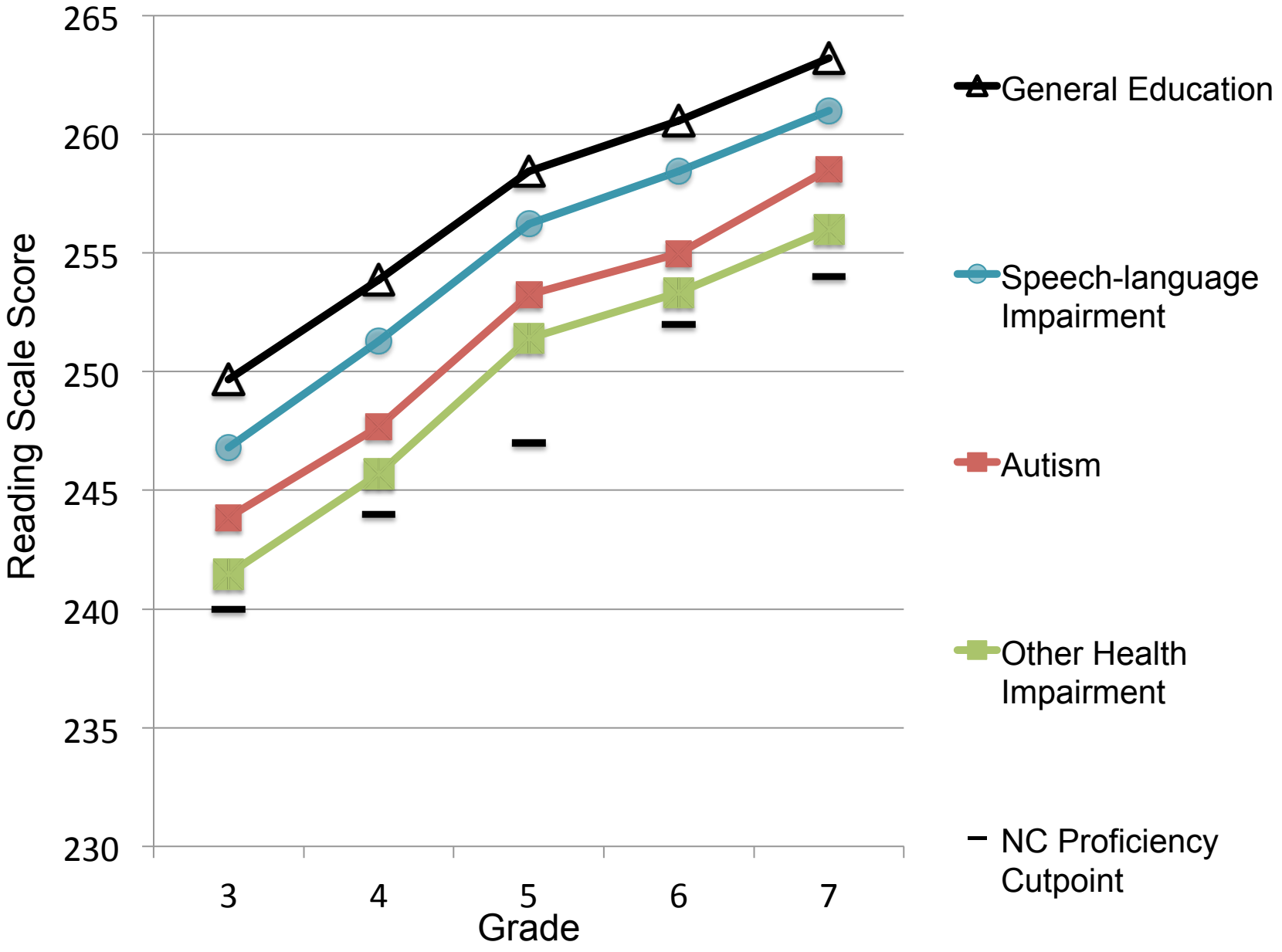
# Mathematics Growth by Exceptionality



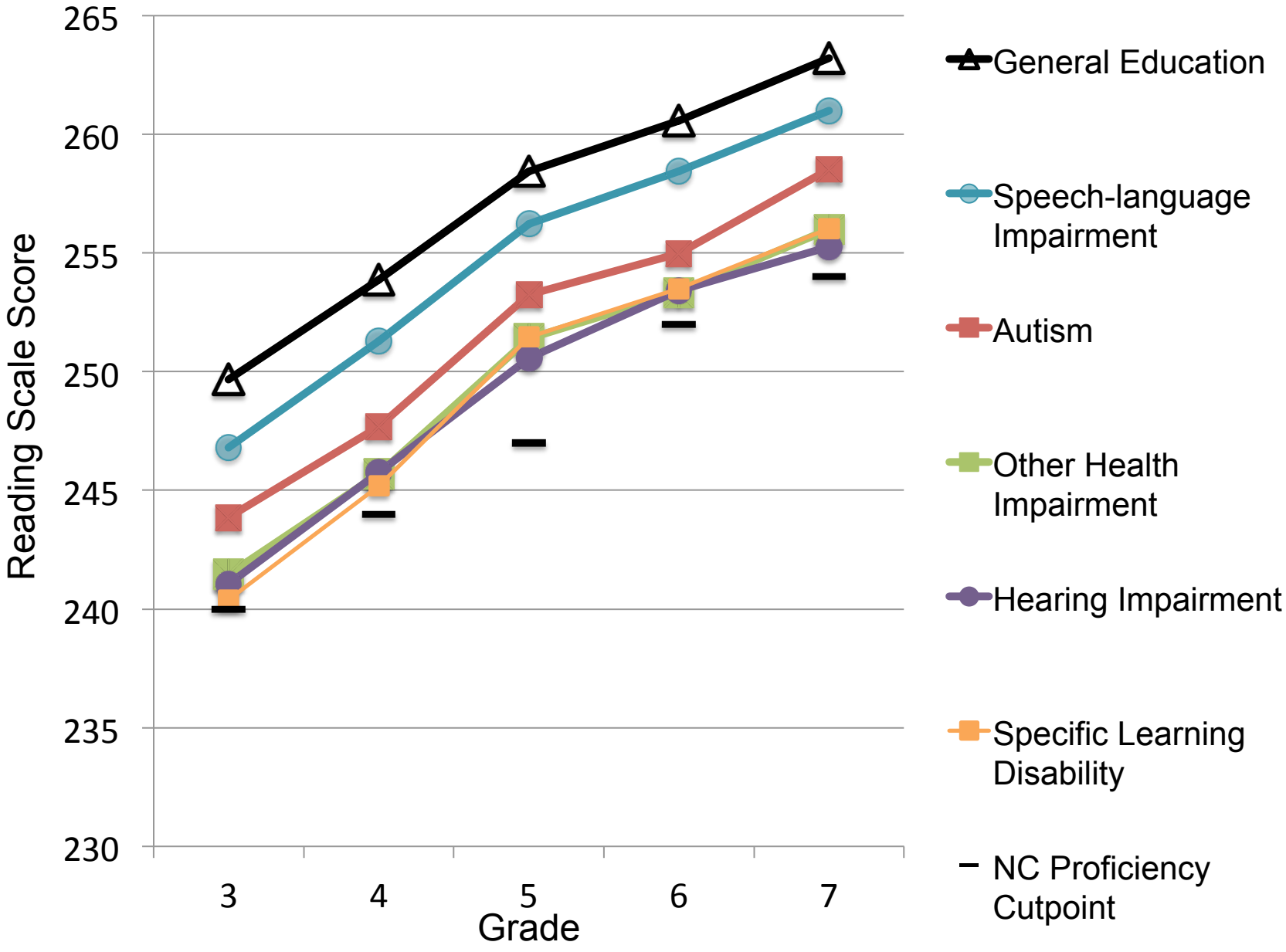
# Reading Growth by Exceptionality



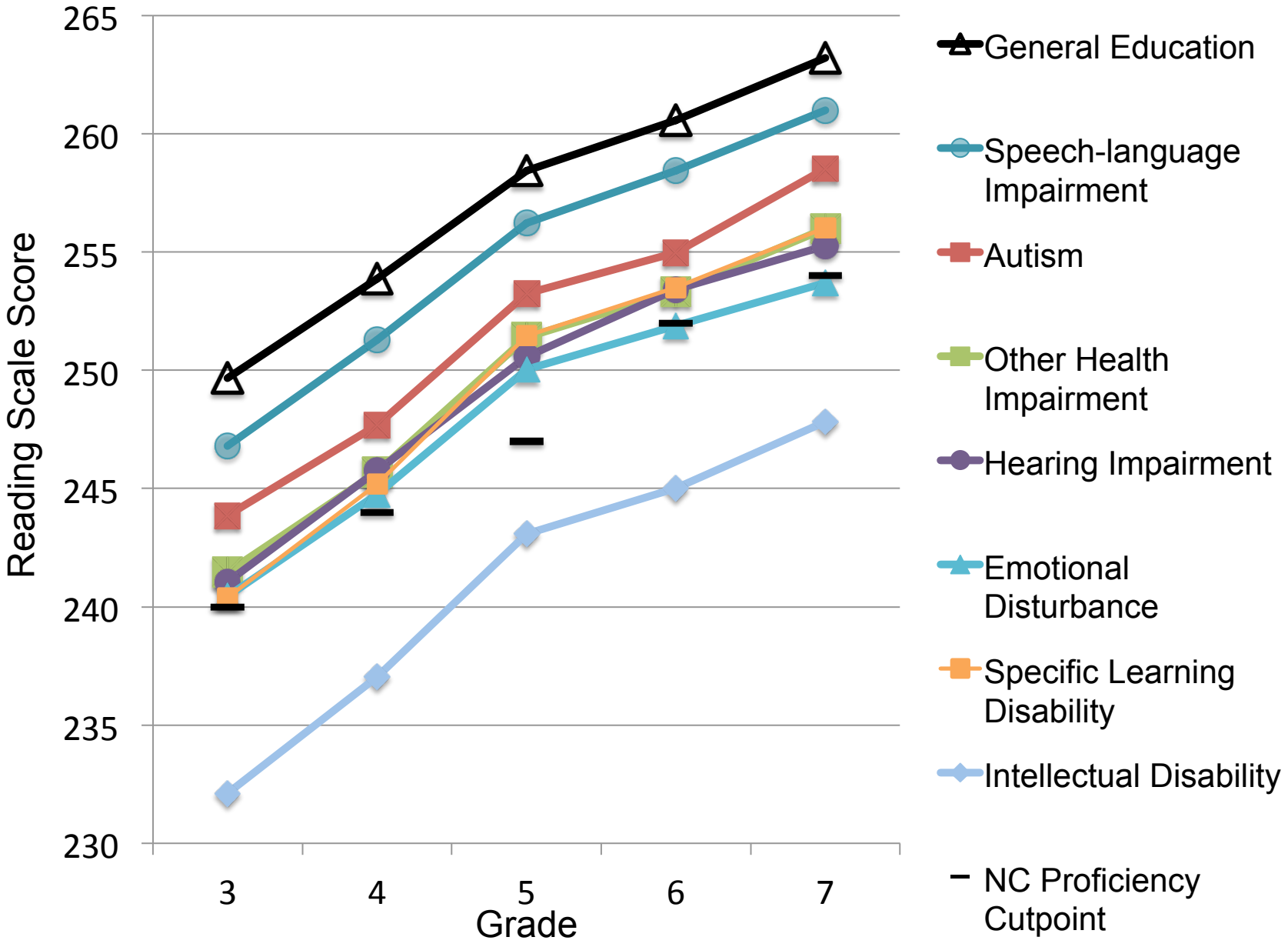
# Reading Growth by Exceptionality



# Reading Growth by Exceptionality



# Reading Growth by Exceptionality



# Individual Differences and Achievement Gaps in Math and Reading for SWD

Joe Stevens

University of Oregon

# Purpose

- Purpose of this presentation to summarize a number of our study results that focus on individual differences in academic performance
- Draw attention to and quantify achievement gaps in mathematics and reading especially for students with and without disabilities (SWD and SWoD)
- A central goal of NCLB and RTTT is universal proficiency and the reduction of achievement gaps between SWoD students and protected subgroups including SWD

# Mathematics Achievement Gaps

- What is the size of the achievement gap in mathematics for students in specific exceptionality categories?
- Does the gap increase, decrease or stay the same over time?
- Previous research on achievement gaps has limitations:
  - Often gaps are not evaluated empirically, visual inspection rather than statistical testing; no common, empirical metric (effect size) to describe differences
  - Interactions not tested (more on this below)

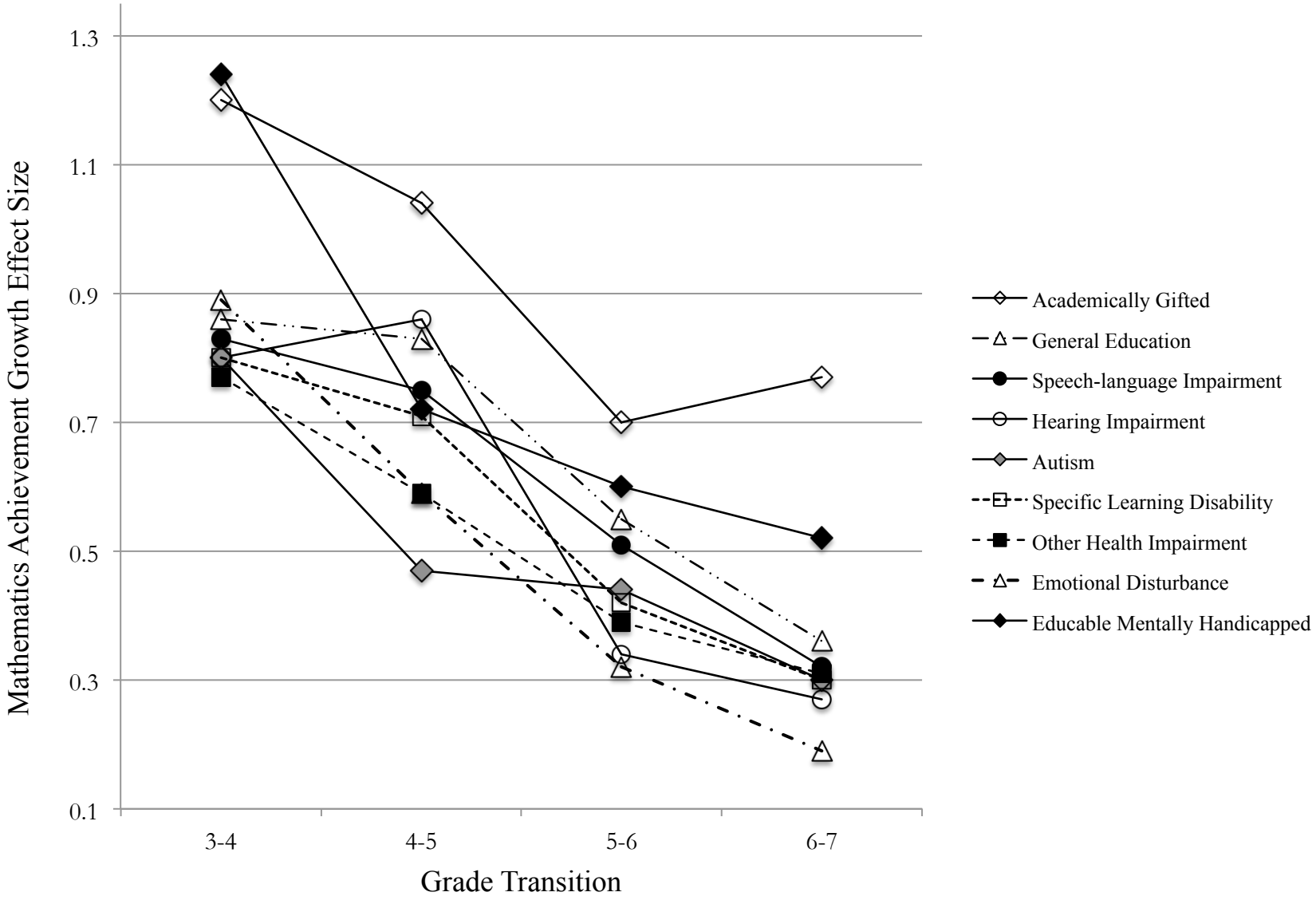



*Figure*. Mathematics achievement growth effect size at each grade transition by student group (from Stevens et al., in press).

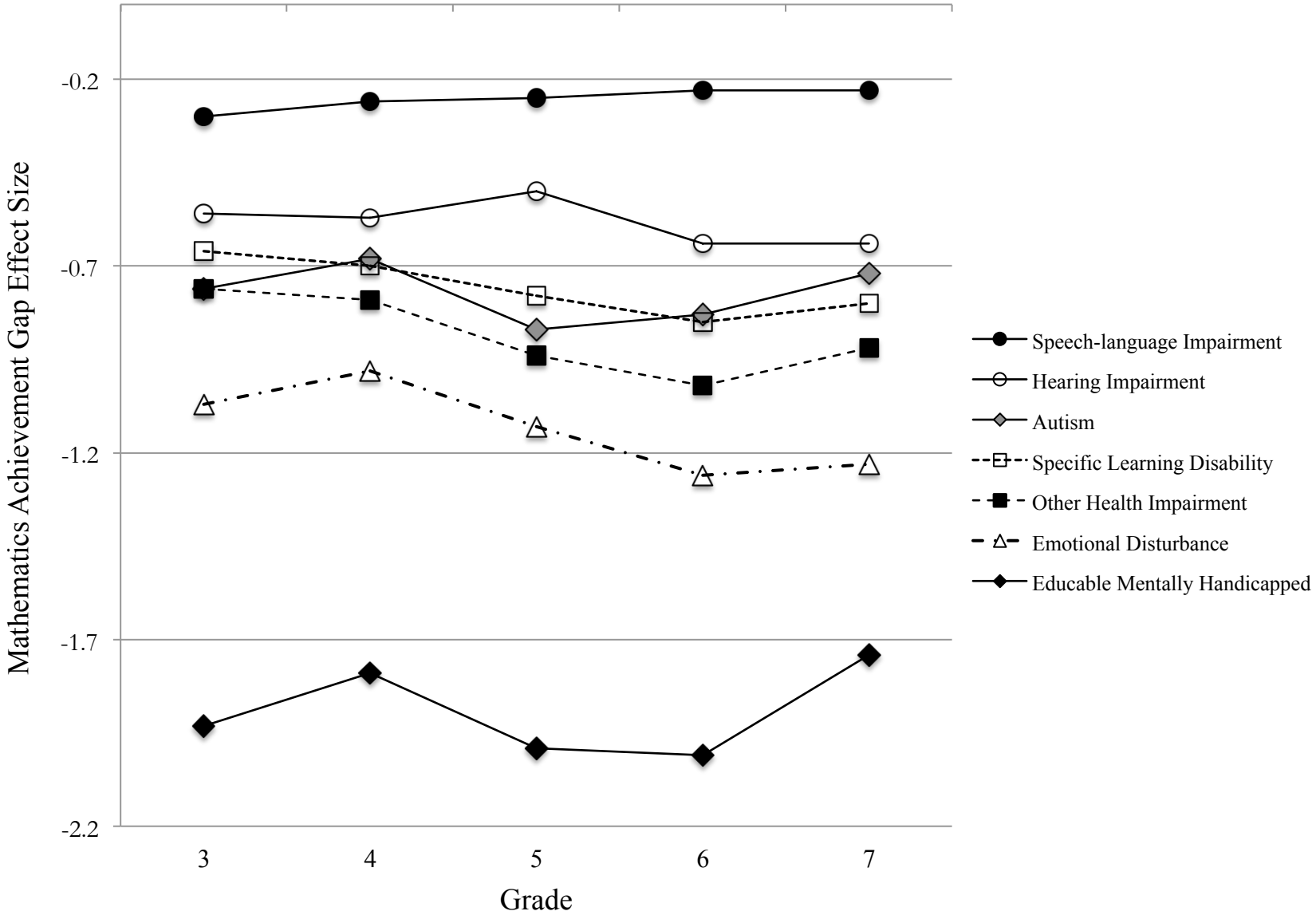



*Figure*. Mathematics achievement gap effect sizes between all SWoD and each exceptionality group by grade (from Stevens et al., in press).

# Reading Achievement Gaps

- What is the size of the achievement gap in reading for students in specific exceptionality categories?
- Does the gap increase, decrease or stay the same over time?

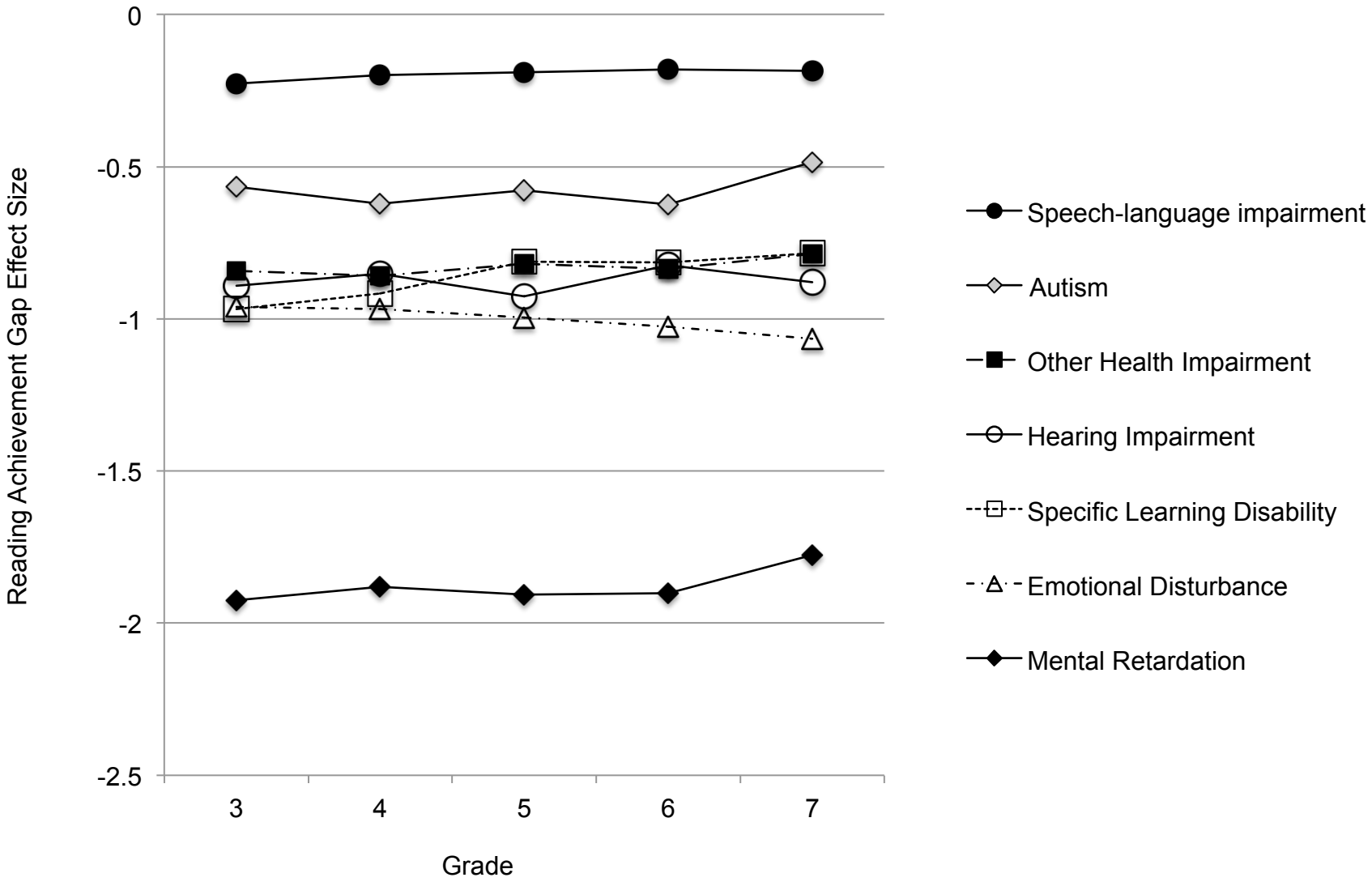



*Figure*. Reading achievement gap effect sizes between all SWoD and each exceptionality group by grade (from Schulte & Stevens, manuscript submitted for publication).

# Achievement Gaps as Differences in Proficiency Rates

- In order to examine the achievement gap between students in specific exceptionality groups and SWoD, we also computed effect size (ES) in other ways
- Difference in percent proficient (*P-P*)
  - Most common method in public dissemination (e.g., report cards); district, state, and federal reports
  - Easy to interpret
- Problems with this metric, however:
  - Size of gap depends on test used and location of cutscore
  - Size of gap depends on shape of score distributions for the two groups
  - Proportions are ordinal, units may be different at different locations on the scale (i.e., not an interval scale)

| North Carolina | Mathematics | | | | | | Reading | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Grade | 3 | 4 | 5 | 6 | 7 | 8 | 3 | 4 | 5 | 6 | 7 | 8 |
| Student Group | | | | | | | | | | | | |
| General Education | 78.3 | 79.3 | 77.2 | 74.4 | 75.4 | 76.8 | 61.7 | 66.1 | 64.1 | 67.7 | 60.0 | 62.3 |
| *N* | (97680) | (94162) | (92973) | (91406) | (90642) | (91668) | (97625) | (94111) | (92935) | (91370) | (90607) | (91627) |
| All SWD | 53.2 | 49.9 | 45.3 | 39.7 | 38.5 | 39.9 | 30.8 | 32.1 | 27.8 | 28.5 | 23.0 | 24.1 |
| *N* | (11208) | (11046) | (9934) | (9310) | (8728) | (8613) | (10759) | (10475) | (9466) | (9059) | (8510) | (8459) |
| *h* | .54 | .63 | .67 | .72 | .76 | .77 | .63 | .69 | .75 | .81 | .77 | .79 |
| Autism | 62.4 | 64.9 | 59.7 | 62.1 | 64.8 | 57.4 | 41.9 | 48.3 | 50.6 | 47.2 | 51.6 | 47.4 |
| *N* | (351) | (365) | (365) | (330) | (244) | (284) | (346) | (360) | (360) | (335) | (252) | (289) |
| *h* | .35 | .32 | .38 | .27 | .23 | .42 | .40 | .36 | .27 | .42 | .17 | .30 |
| Communication | 67.3 | 66.3 | 59.7 | 51.8 | 51.2 | 48.4 | 46.6 | 50.8 | 43.2 | 42.8 | 32.2 | 29.1 |
| *N* | (3842) | (2501) | (1354) | (651) | (369) | (223) | (3838) | (2496) | (1353) | (649) | (370) | (223) |
| *h* | .25 | .29 | .38 | .47 | .51 | .60 | .30 | .31 | .42 | .51 | .57 | .68 |
| Emotional | 44.4 | 43.9 | 35.1 | 29.2 | 26.5 | 22.6 | 28.8 | 44.6 | 26.3 | 22.4 | 22.7 | 17.9 |
| *N* | (331) | (394) | (453) | (510) | (475) | (562) | (323) | (166) | (453) | (510) | (476) | (570) |
| *h* | .71 | .75 | .88 | .94 | 1.02 | 1.15 | .67 | .44 | .78 | .95 | .78 | .95 |
| Hearing | 49.0 | 50.0 | 56.0 | 46.7 | 41.0 | 44.1 | 20.6 | 30.9 | 32.3 | 34.3 | 21.9 | 24.1 |
| *N* | (143) | (130) | (134) | (105) | (100) | (111) | (136) | (123) | (133) | (102) | (96) | (112) |
| *h* | .62 | .63 | .45 | .58 | .71 | .68 | .86 | .72 | .65 | .68 | .80 | .79 |

| | Mathematics | | | | | | Reading | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Grade | 3 | 4 | 5 | 6 | 7 | 8 | 3 | 4 | 5 | 6 | 7 | 8 |
| General Education | 78.3 (97680) | 79.3 (94162) | 77.2 (92973) | 74.4 (91406) | 75.4 (90642) | 76.8 (91668) | 61.7 (97625) | 66.1 (94111) | 64.1 (92935) | 67.7 (91370) | 60.0 (90607) | 62.3 (91627) |
| Intellectual | 7.9 | 4.8 | 3.7 | 4.4 | 4.2 | 6.1 | 1.7 | 3.2 | 2.8 | 2.2 | 0.0 | 1.7 |
| *N* | (252) | (229) | (215) | (225) | (240) | (296) | (239) | (218) | (216) | (230) | (250) | (295) |
| *h* | 1.60 | 1.76 | 1.76 | 1.66 | 1.69 | 1.64 | 1.55 | 1.54 | 1.52 | 1.63 | 1.59 | 1.56 |
| Orthopedic | 55.6 | 40.5 | 62.5 | 47.7 | 58.8 | 54.7 | 42.1 | 45.9 | 50.0 | 42.6 | 58.8 | 40.0 |
| *N* | (36) | (37) | (48) | (44) | (34) | (53) | (38) | (37) | (48) | (47) | (34) | (55) |
| *h* | .49 | .82 | .32 | .56 | .36 | .47 | .39 | .41 | .29 | .51 | .02 | .45 |
| Other | 44.1 | 41.7 | 38.5 | 36.0 | 33.8 | 36.7 | 24.7 | 27.7 | 26.7 | 29.9 | 22.2 | 25.6 |
| *N* | (1663) | (2085) | (2218) | (2347) | (2299) | (2172) | (1625) | (2043) | (2204) | (2371) | (2313) | (2189) |
| *h* | .72 | .79 | .81 | .79 | .86 | .83 | .77 | .79 | .77 | .78 | .79 | .76 |
| Language Disability | 47.1 | 46.4 | 45.5 | 40.8 | 40.9 | 43.6 | 19.6 | 24.2 | 22.7 | 26.1 | 21.9 | 23.7 |
| *N* | (4524) | (5246) | (5085) | (5058) | (4909) | (4865) | (4148) | (4761) | (4638) | (4776) | (4662) | (4680) |
| *h* | .66 | .70 | .67 | .69 | .72 | .69 | .89 | .87 | .86 | .86 | .80 | .80 |
| TBI | 30.0 | – | 35.3 | 38.5 | 36.8 | 20.0 | – | – | 43.8 | 25.0 | 15.8 | 6.7 |
| *N* | (10) | (8) | (17) | (13) | (19) | (15) | (9) | (7) | (16) | (12) | (19) | (15) |
| *h* | 1.01 | – | .87 | .74 | .80 | 1.21 | – | – | .41 | .89 | .95 | 1.30 |
| Visual Impairment | 62.8 | 75.0 | 57.1 | 60.9 | 62.9 | 74.1 | 50.0 | 53.3 | 57.1 | 39.1 | 54.3 | 46.2 |
| *N* | (43) | (48) | (42) | (23) | (35) | (27) | (44) | (45) | (42) | (23) | (35) | (26) |
| *h* | .34 | .10 | .43 | .29 | .27 | .06 | .24 | .26 | .14 | .58 | .12 | .32 |

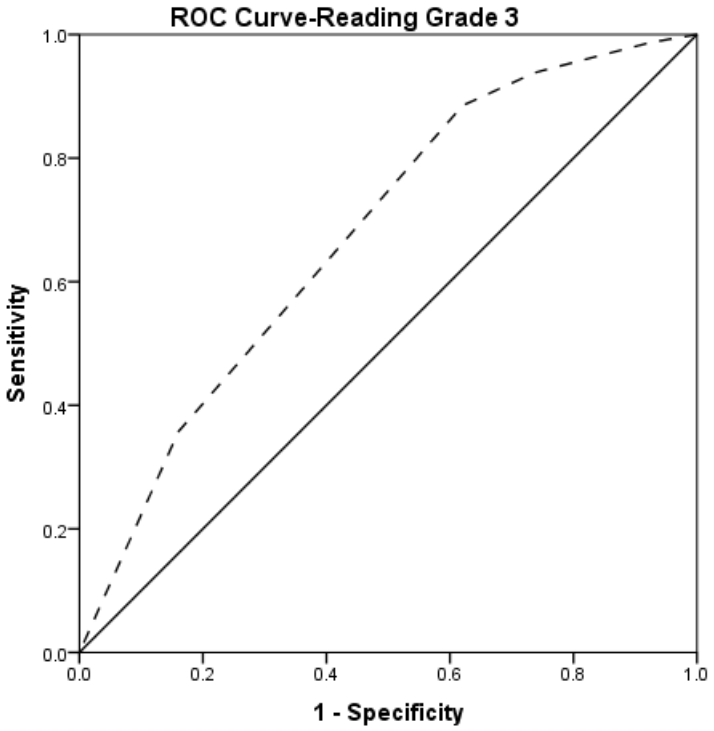
# Achievement Gaps as Areas Between Score Distributions

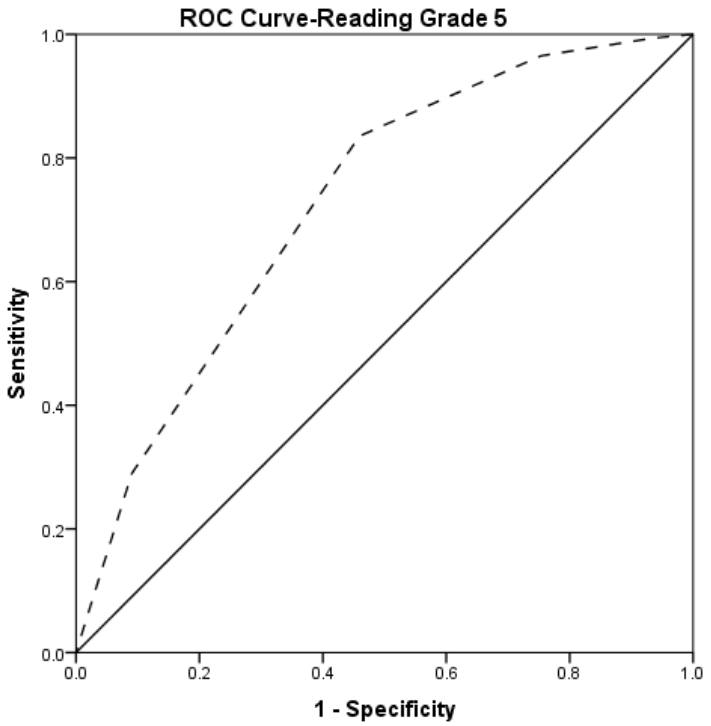
- A limitation of traditional ES measures is that they only compare groups at the mean or at the proficiency cutpoint, possibly overlooking important group differences lower or higher on the score scale
- Alternatives are effect size measures based on nonparametric methods that examine group differences for all proficiency categories (see Ho & Reardon, 2012):
  - Area under the curve in Receiver Operating Curve (ROC) analysis
  - $V'$ statistic
- Because of time constraints, we report proportions and Cohen's $h$ here and examples of ROC analysis (our $V'$ statistic tables are in preparation)

# Whole distribution Achievement Gaps

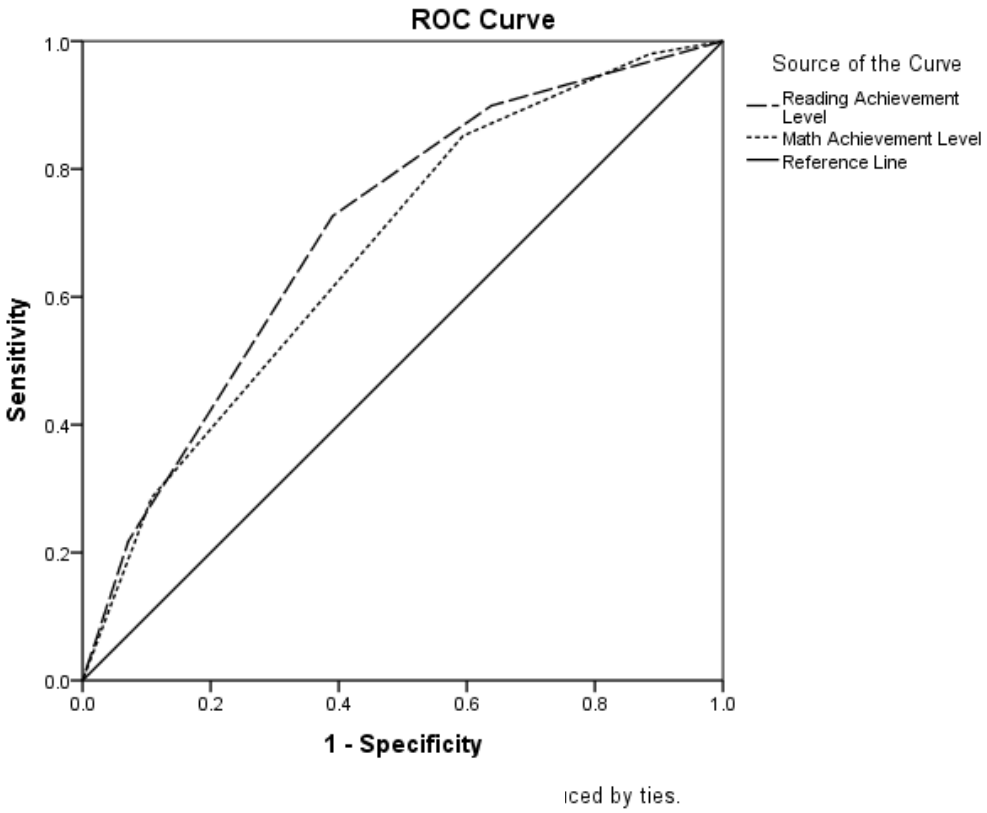
- ROC analysis (and $V'$ ) use nonparametric methods to relieve problems associated with characteristics of score distributions
- Advantage is estimation of gap across all proficiency levels
- ROC curve diagonal line represents no difference between reference group (SWoD) and focal group (SWD)
- Size of area between SWD group curve and diagonal is the area under the curve or the size of the difference between the two groups
- In following examples, note differences:
  - at different proficiency levels
  - for math vs. reading
  - by exceptionality subgroup

# Achievement Gap for SWD vs. SWoD in Oregon Reading in Grade 3 (on left) and Grades 5 (on right)

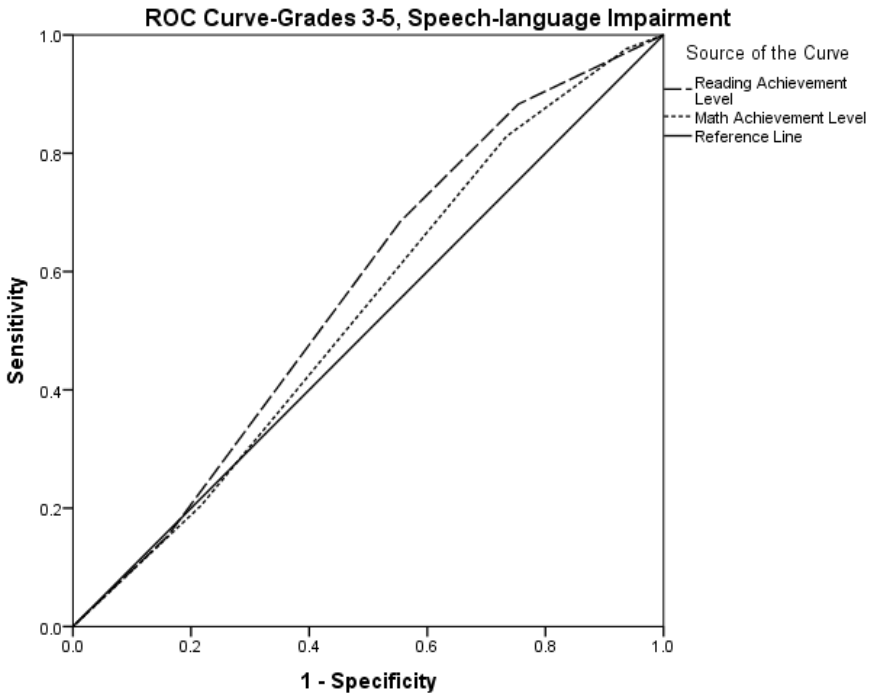


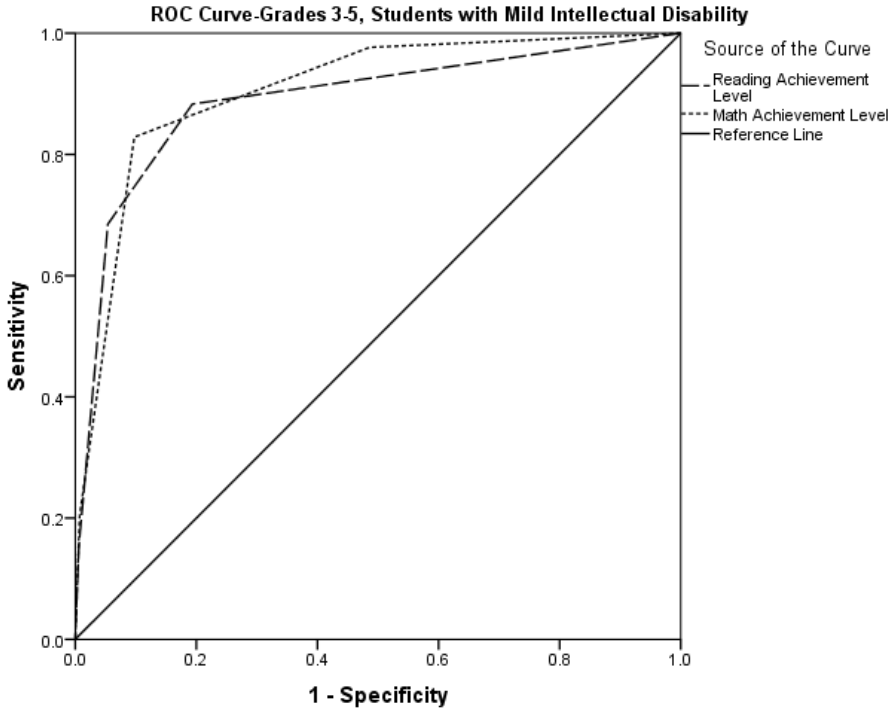



# Achievement Gap for SWD vs. SWoD in NC Math and Reading Grades 3-5



# Achievement Gap for SWoD vs. Speech-language Impairment (on left) or Mild Intellectual Disability (on right) on NC Math and Reading Grades 3-5



# Interactions of SWD status and Other Student Characteristics

- Many studies do not directly test the interaction of SWD status and factors thought to be related to student performance (e.g., LD status and sex of student)
- When these factors are included in statistical models (especially regression and HLM models), only partial regression effects not the actual interactions are analyzed
- This can be very misleading and result in incorrect interpretations
- We explicitly test interactions of SWD with student characteristics

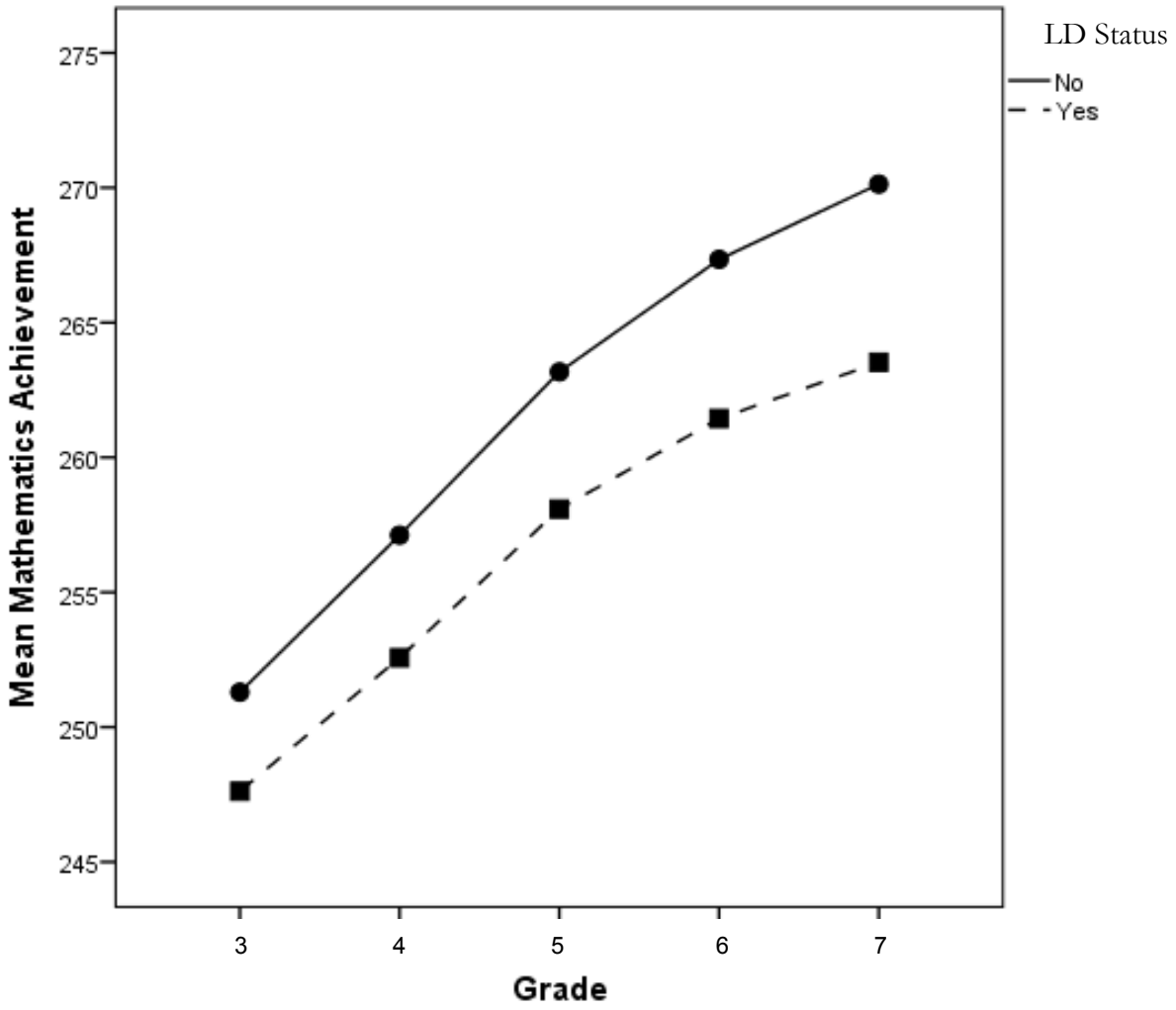



*Figure*. Mean mathematics achievement by grade and LD status (note increasing achievement gap; Stevens & Schulte, manuscript submitted for publication).

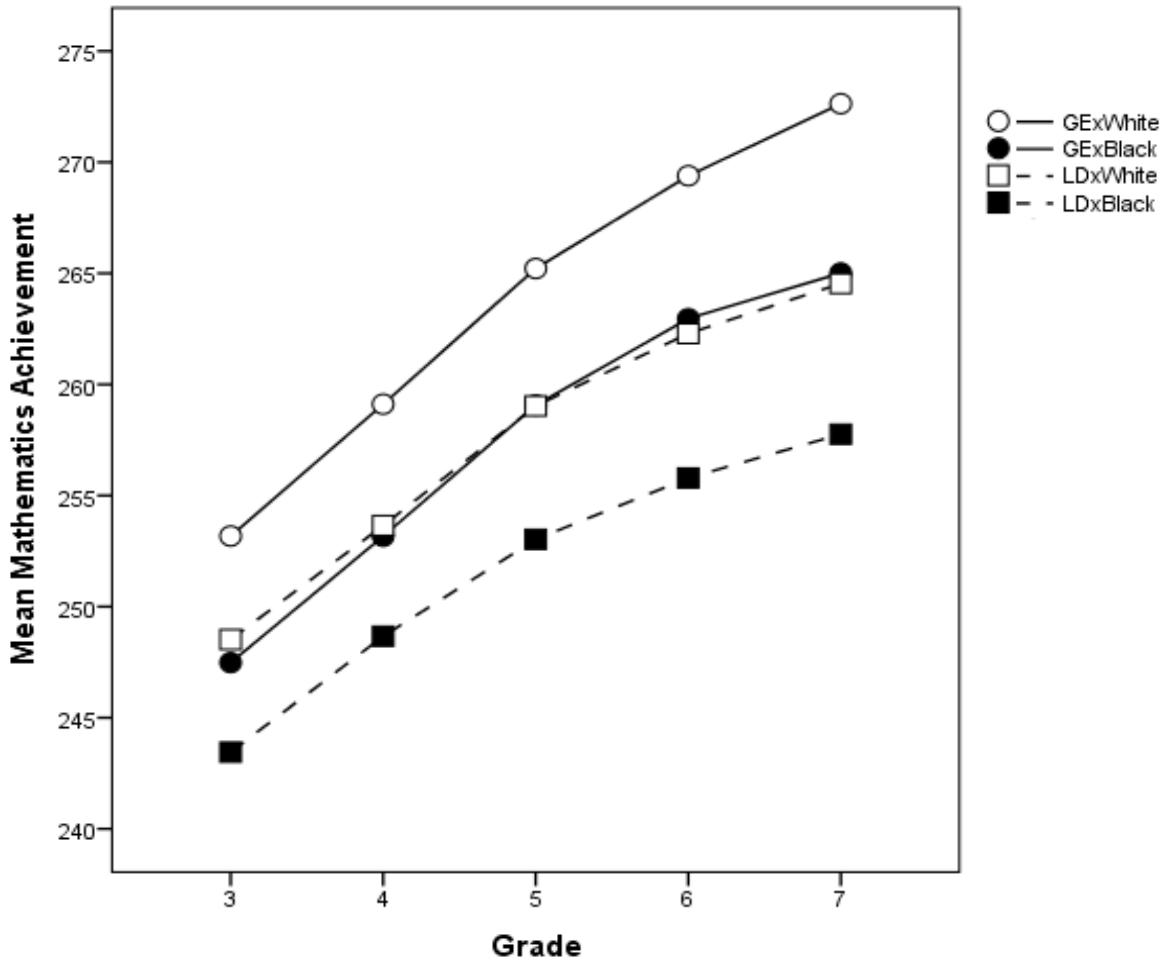



*Figure*. Interaction of LD Status With Black Race/ethnicity on Mathematics Achievement Growth (almost identical results for interaction of LD and FRL status; from Stevens & Schulte, manuscript submitted for publication).

# Mathematics Achievement Gaps for Elementary and Secondary Students: The Influence of Opportunity to Learn and Special Education Status

Stephen N. Elliott

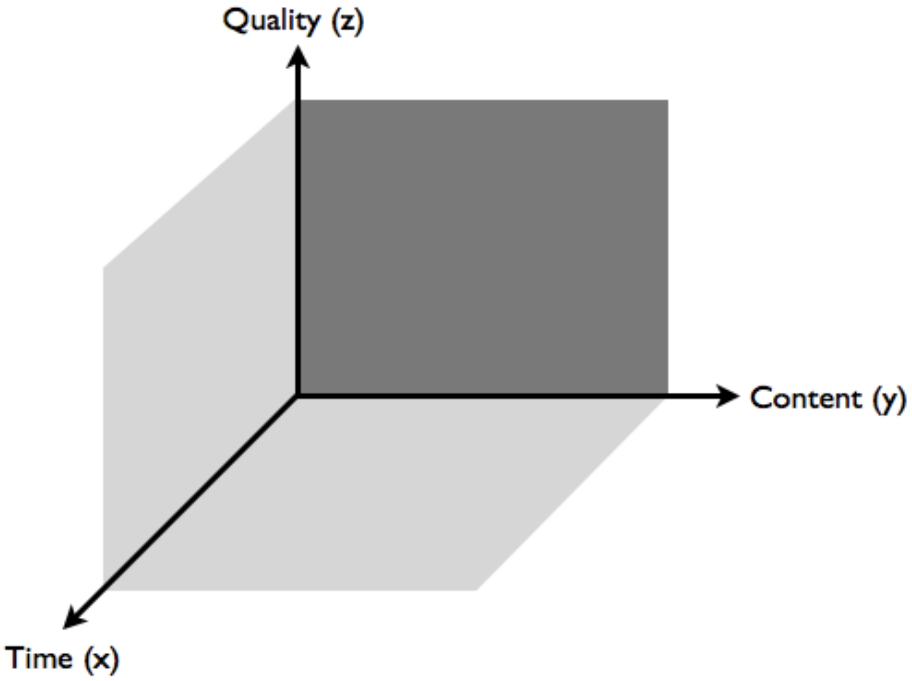
Arizona State University

# Research Questions

Specific research questions motivating the study were:

1. Do students with and without disabilities who received instruction in the same general education classrooms have an equal opportunity to learn mathematics?
2. What is the relationship among five instructional variables (characterized as OTL) and within year academic growth on an interim assessments?
3. What is the predictive relationship among five instructional OTL variables and students' end-of-year mathematics achievement?

# Opportunity to Learn (OTL) the Intended Curriculum




**Definition: Opportunity to Learn**

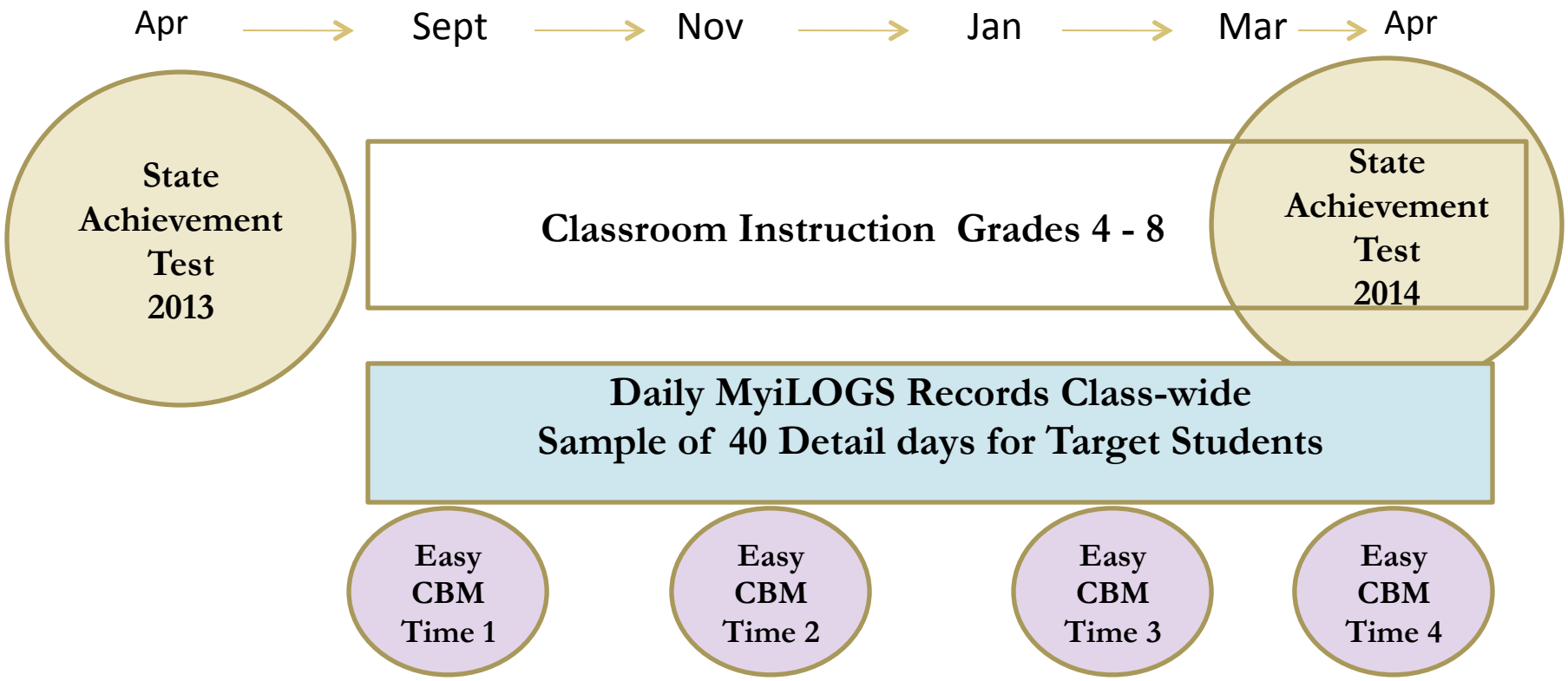
*The degree to which a teacher dedicates instructional time and content coverage to the intended curriculum objectives emphasizing higher-order cognitive processes, evidence-based instructional practices, and alternative grouping formats.*

(Kurz, 2011)

**A unified conceptualization of OTL based on 50+ years of empirical research.**

# Multiple Measures Study Design*

Teachers (N = 78; AZ 49, OR 29) and students (N = 327; 162 SWD + 165 SWoD) from AZ & OR schools grades 4th-8th




*A 3-year study with longitudinal student cohorts

# Summary of Year 1: Key Findings

- We observed very similar instructional processes for students with and without disabilities learning mathematics in the same elementary or secondary classrooms in AZ and OR schools. Significant achievement gaps between these groups of students, however, existed on the four interim CBM assessments and the end-of-year achievement state test.

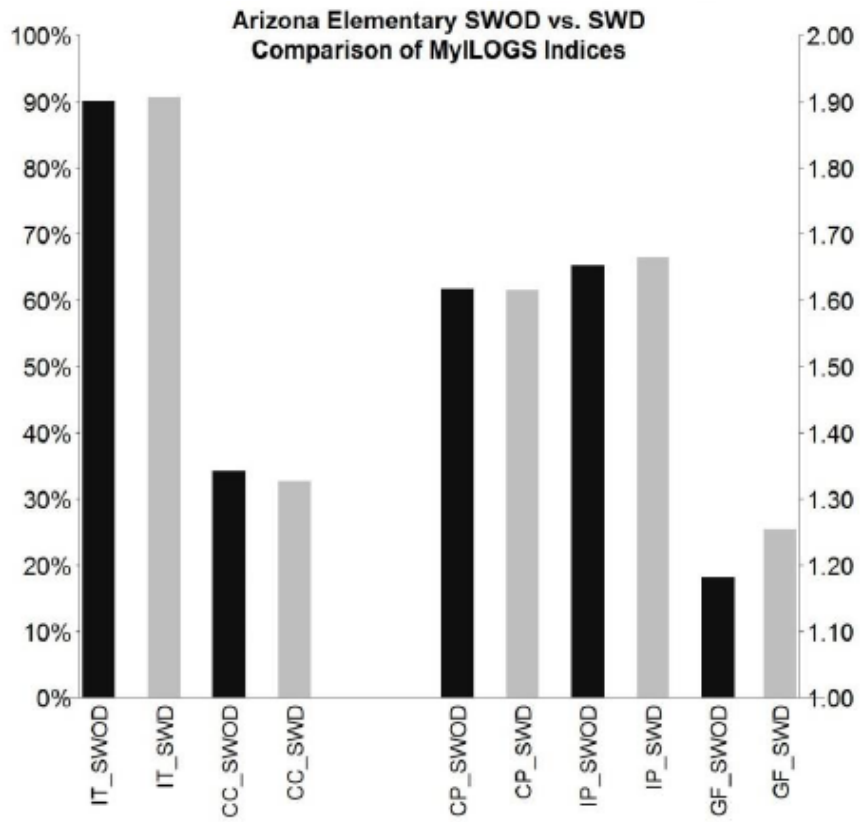
- We found that the collection of five MyiLOGS scores, along with grade level and special education status, accounted for a substantial amount (i.e., 43% to 44%) of the variance in student's end-of-year mathematics scores. A subset of OTL indices explained a statistically significant, although relatively small portion of unique variance in the end-of-year mathematics scores. The particular OTL scores found to be significant contributors varied across AZ and OR.

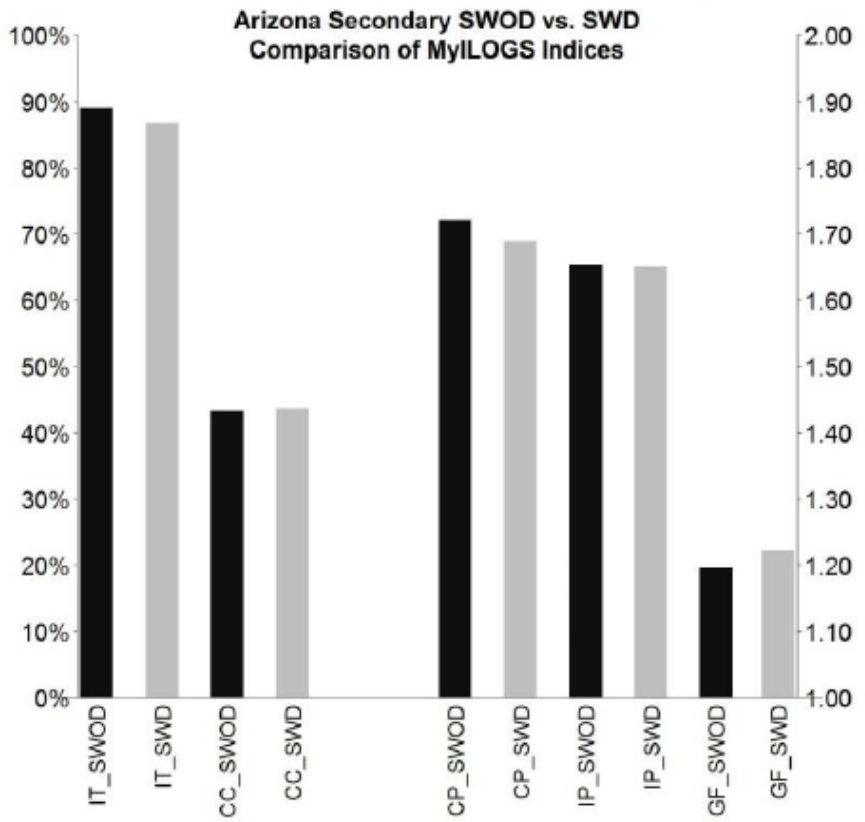
# Year 2 Findings

- AZ teachers reported an average of 164 days and OR teachers reported 158 days of instruction; 25% of these days were Detail Days where instructional information on cognitive processes, practices, and grouping for SWD and SWOD was documented. Based on these Detail Days, we observed **very similar mathematics instructional processes for students with and without disabilities in the same elementary or secondary classrooms** in AZ and OR schools. Yet, there were **significant achievement gaps between these groups of students** on the four interim CBM assessments and the end-of-year achievement state test.

- We also found that Grade Level and Special Education Status, along with the collection of five MyiLOGS scores, accounted for a substantial amount (i.e., 30% OR, 39% AZ) of the variance in student's end-of-year mathematics scores. OTL indices explained a relatively small portion of unique variance in the end-of-year mathematics scores.

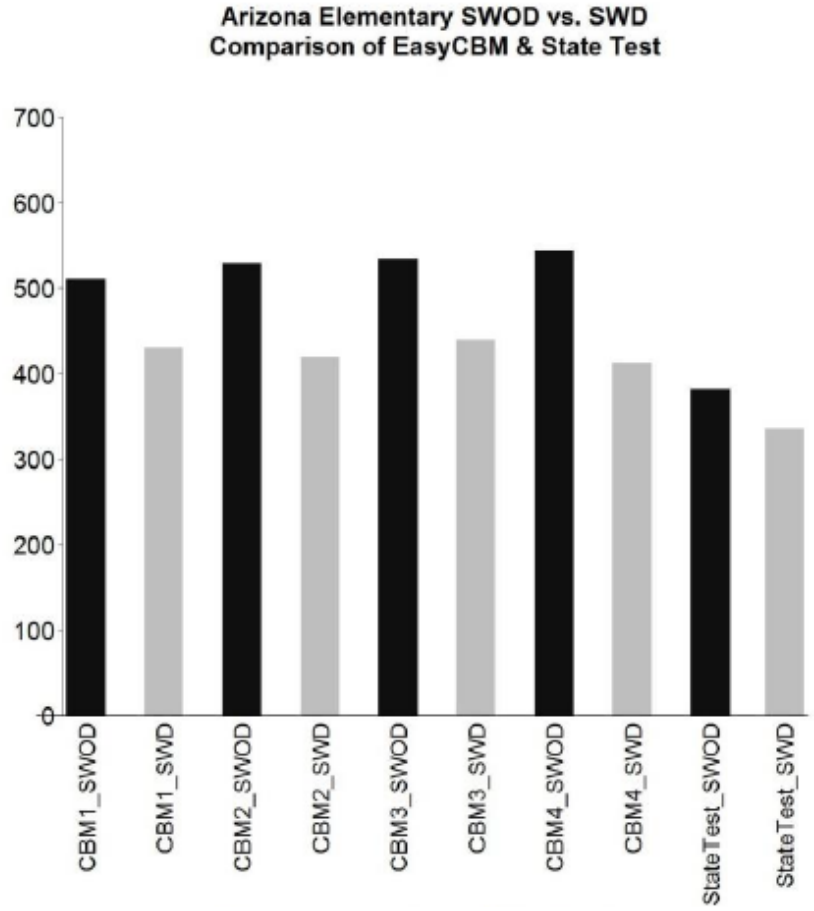
# Comparison of OTL Indices for AZ Students

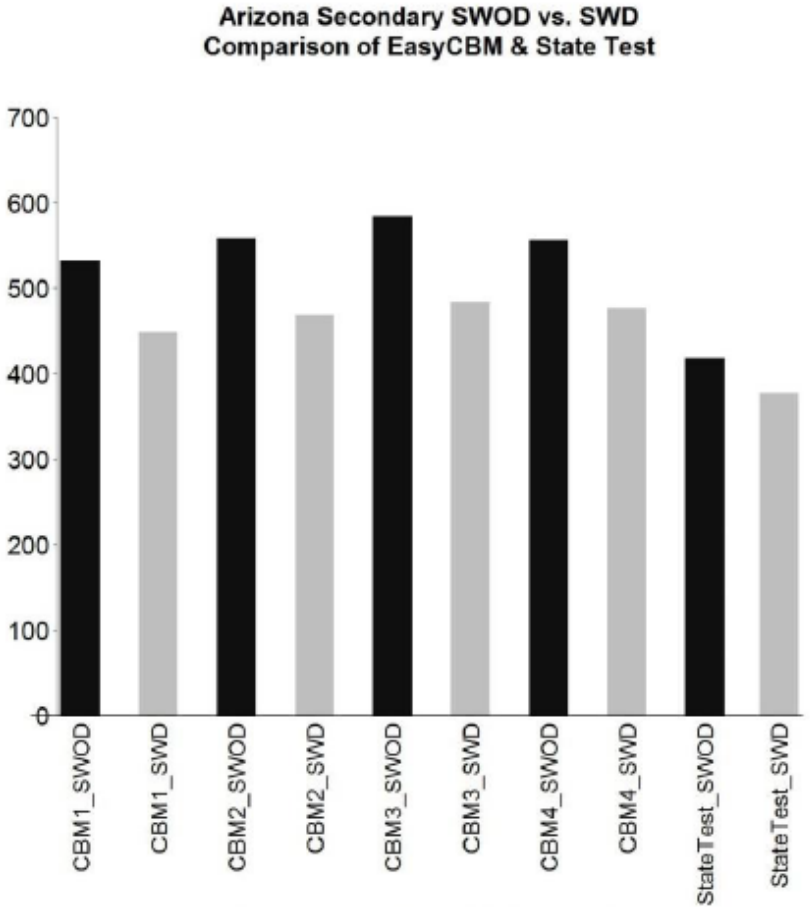



**SWOD Black** **SWD Gray**

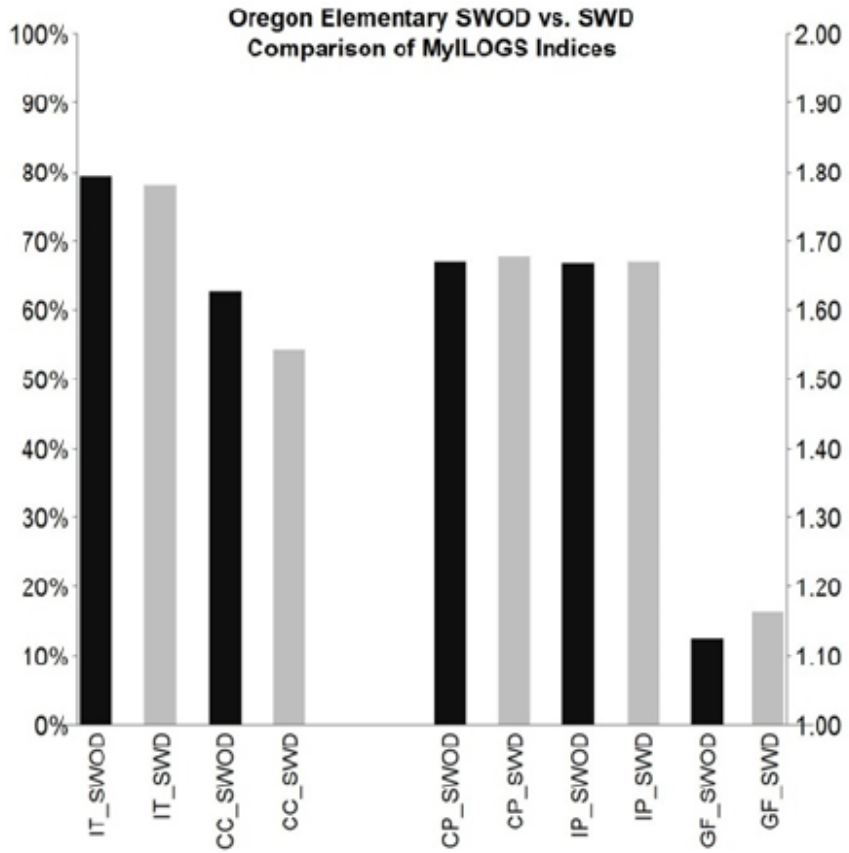
# Comparison of Interim & End-of-Year Test Results for AZ Students

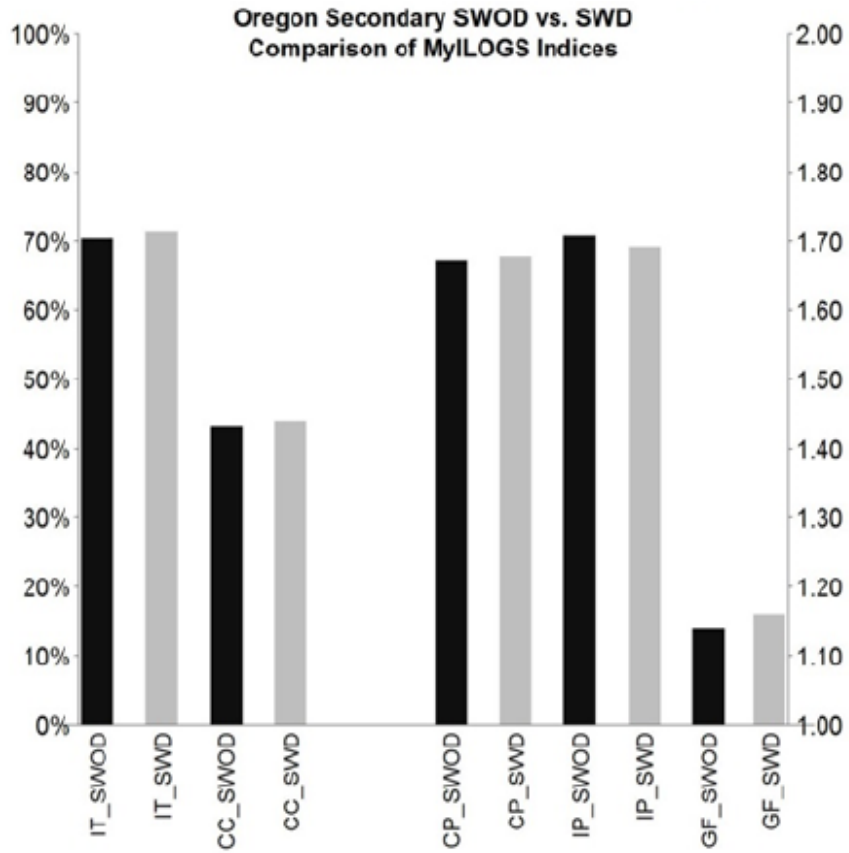



**SWOD Black** **SWD Gray**

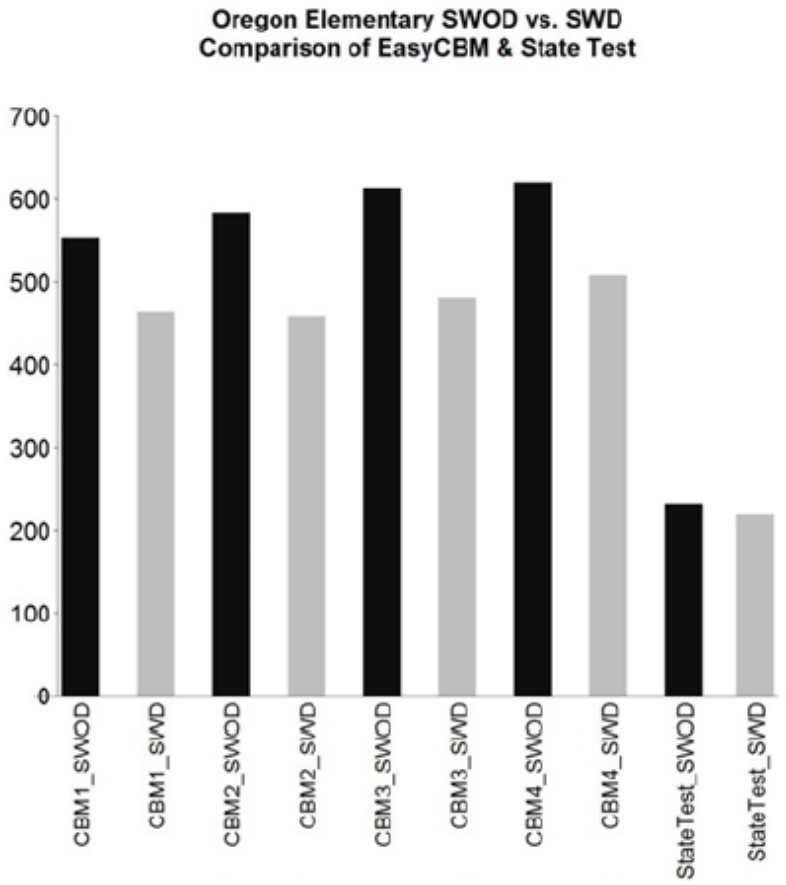
# Comparison of OTL Indices for OR Students

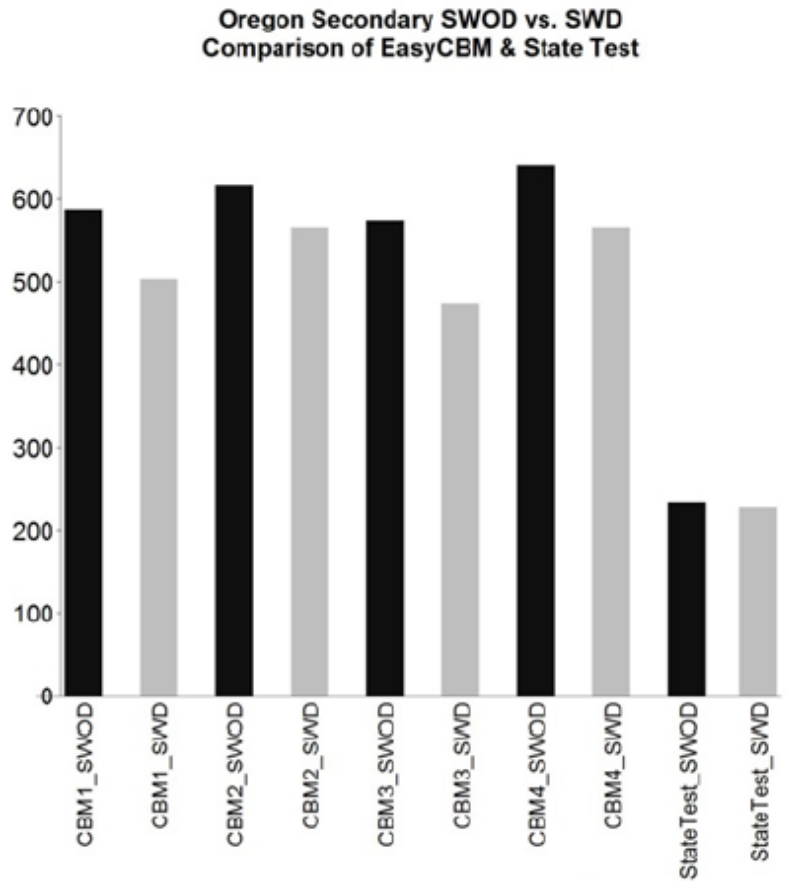



**SWOD Black** **SWD Gray**

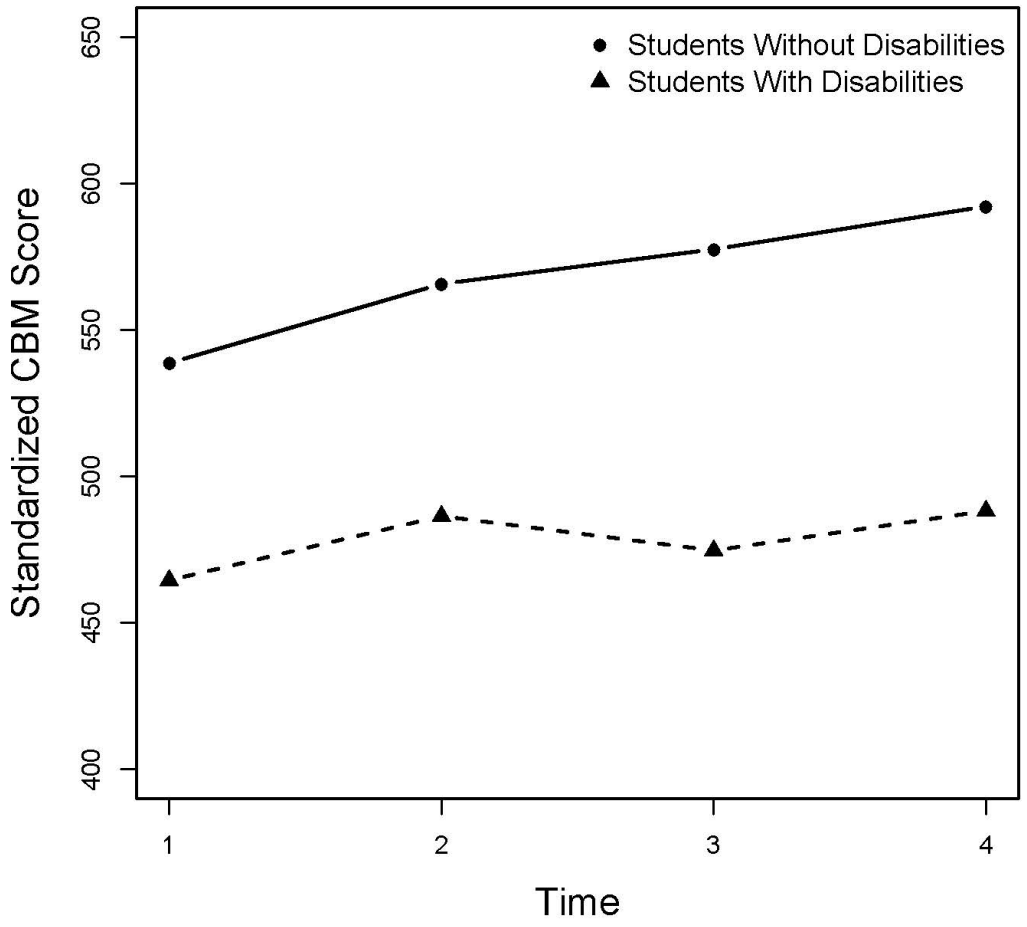
# Comparison of Interim & End-of-Year Test Results for OR Students




**SWOD Black** **SWD Gray**

# Within Year Standardized Mathematics CBM Growth



# Conclusion

Offering students with disabilities the same amount of instruction on the same content standards in the same general education classrooms was found to offer **the same historic results—large and persistent gaps in achievement -- in comparison to students without disabilities.**

The findings in Year 2 replicated those from Year 1. Thus, it indicates that students with disabilities will need more instructional time on the intended curriculum, and perhaps more differentiated instruction to increase their rate of achievement enough to close gaps that currently exist between them and students without disabilities.

# Influence of Testing Procedures on Documenting Growth

**Multiple Testing**
**Within Year Growth**
**Test Participation**
**Growth on Alternate Assessment**

Gerald Tindal
Joseph F. T. Nese
University of Oregon

NCAASE National Center on Assessment and Accountability for Special Education
*Advancing research on growth measures, models, and policies for improved practice*

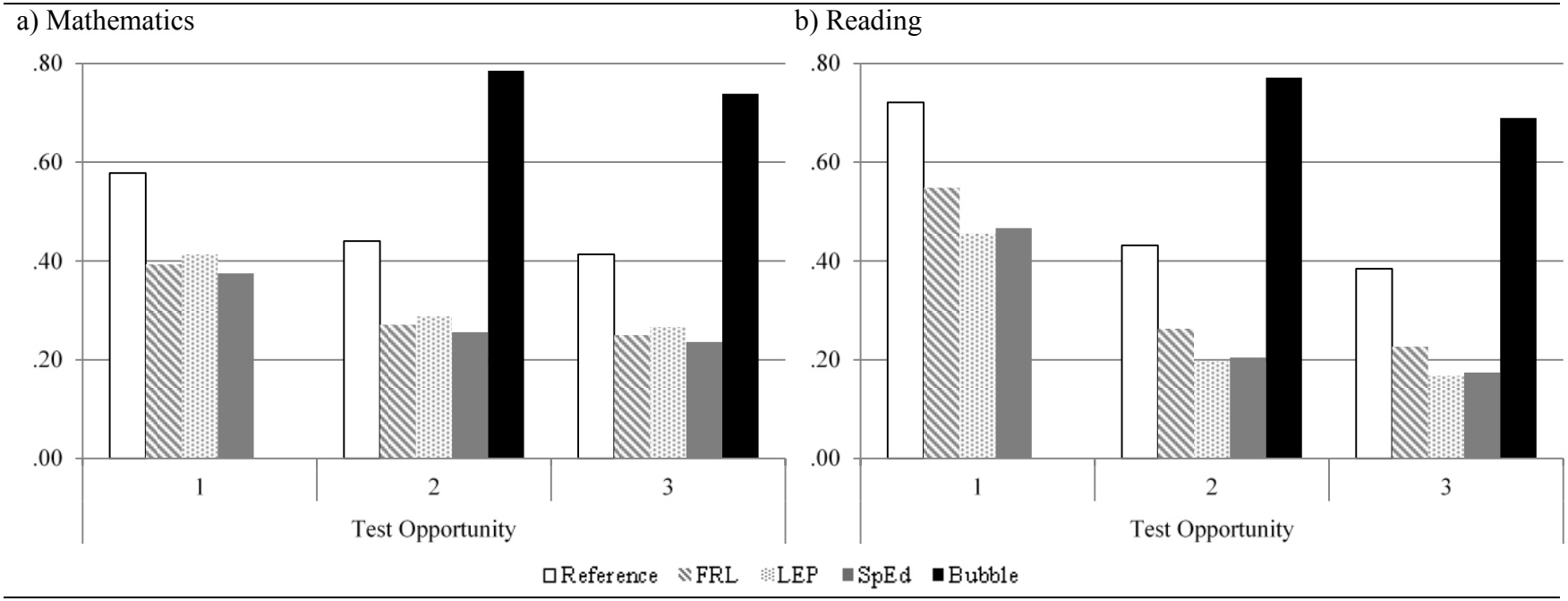
- **The Influence of Multiple Administrations of a State Achievement Test on Passing Rates for Student Groups**

  - Joseph F. T. Nese, Gerald Tindal, Joseph J. Stevens, Stephen N. Elliott

# The Influence of Multiple Administrations of a State Achievement Test

- **Purpose:** Explore outcomes from the use of multiple test administrations in reaching proficiency.
  1. Are student characteristics associated with how many times a student takes the state test?
  2. For various student subgroups, what is the likelihood of passing the test given previous failure(s)?
- Performance of students on the "bubble" of proficiency (potential false-negatives).

# Estimated probabilities of passing the Grade 3 math or reading test for specific student subgroups




**Reference** = White, male, non-FRL (free/reduced priced lunch recipient), non-LEP (limited English proficiency status), GenEd (general education), BelowBubble (lower than one standard error of measurement below the proficiency cut score on the previous test).
**FRL** = White, male, FRL, non-LEP, GenEd, BelowBubble.
**LEP** = White, male, non-FRL, LEP, GenEd, BelowBubble.
**SpEd** = White, male, non-FRL, non-LEP, Special Education, BelowBubble.
**Bubble** = White, male, non-FRL, non-LEP, GenEd, Bubble (one standard error of measurement below the proficiency cut score on the previous test).

Article

HAMMILL INSTITUTE ON DISABILITIES

Remedial and Special Education
1–13

DOI: 10.1177/0741932515590234
rase.sagepub.com
SAGE

# Growth on Oral Reading Fluency Measures as a Function of Special Education and Measurement Sufficiency

**Gerald Tindal, PhD[1], Joseph F. T. Nese, PhD[1], Joseph J. Stevens, PhD[1], and Julie Alonzo, PhD[1]**

## Abstract

For 30 years, researchers have investigated oral reading fluency as a measure of growth in reading proficiency. Yet, little research has been done with these measures in the context of progress monitoring in Tier 2 systems. First, we document teachers' progress-monitoring decisions on type of passage (on-grade or off-grade) and how often to administer them. Then, we use a two-level hierarchical linear model to document the effects on both intercept and slope as a function of student special education status and measurement sufficiency. Across Grades 3 to 5, teachers diagnostically document student performance with different grade-level measures and also target a group of Tier 2 students to monitor early and systematically throughout the year. This latter group starts out much lower but has a significantly different slope than those for whom progress monitoring is more diagnostic and infrequent.

## Keywords

academic achievement, curriculum based, assessment, reading, elementary, special education, oral reading fluency

**Table 3.** Unconditional Growth Model Results for Linear Growth Across Three Grades.

| Fixed effect | Coefficient | *SE* | *t* ratio | *df* | *p* value |
|---|---|---|---|---|---|
| Grade 3 | | | | | |
| Intercept, $\beta_{00}$ | 80.72 | 1.113 | 72.54 | 1280 | <.001 |
| Linear slope, $\beta_{10}$ | 0.73 | 0.03 | 23.78 | 1280 | <.001 |
| Grade 4 | | | | | |
| Intercept, $\beta_{00}$ | 100.59 | 0.99 | 101.78 | 1237 | <.001 |
| Linear slope, $\beta_{10}$ | 0.65 | 0.02 | 32.62 | 1237 | <.001 |
| Grade 5 | | | | | |
| Intercept, $\beta_{00}$ | 125.00 | 1.169 | 107.00 | 1096 | <.001 |
| Linear slope, $\beta_{10}$ | 0.68 | 0.024 | 28.45 | 1096 | <.001 |

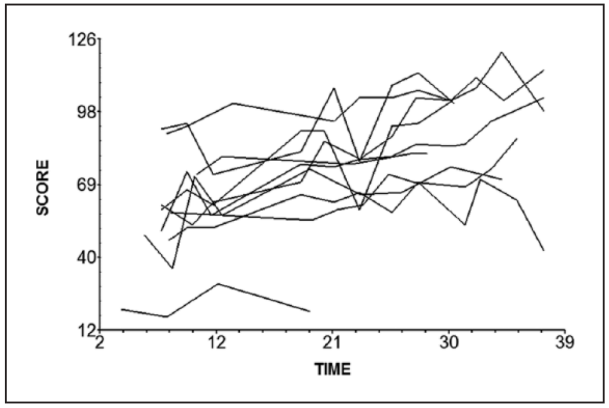



**Figure 1.** Plot of a random sample (.05%) of students' progress monitoring unconditional growth.

**Table 4.** Final Conditional Model With Special Education Status and Measurement Condition.

| Fixed effect | Coefficient | *SE* | *t* ratio | *df* | *p* value |
|---|---|---|---|---|---|
| Grade 3 | | | | | |
| Intercept, $\beta_{00}$ | 88.23 | 1.286 | 68.61 | 1278 | <.001 |
| Special education, $\beta_{01}$ | −19.70 | 3.041 | −6.48 | 1278 | <.001 |
| Sufficient measurement, $\beta_{02}$ | −21.72 | 2.23 | −9.74 | 1278 | <.001 |
| Slope, $\beta_{10}$ | 0.67 | 0.03 | 20.87 | 1278 | <.001 |
| Special education, $\beta_{11}$ | 0.08 | 0.12 | 0.67 | 1278 | .505 |
| Sufficient measurement, $\beta_{12}$ | 0.16 | 0.05 | 2.99 | 1278 | .003 |
| Grade 4 | | | | | |
| Intercept, $\beta_{00}$ | 107.56 | 1.09 | 98.69 | 1235 | <.001 |
| Special education, $\beta_{01}$ | −23.71 | 2.70 | −8.79 | 1235 | <.001 |
| Sufficient measurement, $\beta_{02}$ | −19.89 | 2.14 | −9.31 | 1235 | <.001 |
| Slope, $\beta_{10}$ | 0.62 | 0.03 | 24.31 | 1235 | <.001 |
| Special education, $\beta_{11}$ | 0.01 | 0.05 | 0.16 | 1235 | .872 |
| Sufficient measurement, $\beta_{12}$ | 0.12 | 0.04 | 2.71 | 1235 | .007 |
| Grade 5 | | | | | |
| Intercept, $\beta_{00}$ | 133.24 | 1.30 | 102.37 | 1094 | <.001 |
| Special education, $\beta_{01}$ | −28.42 | 2.70 | −10.53 | 1094 | <.001 |
| Sufficient measurement, $\beta_{02}$ | −18.27 | 2.53 | −7.22 | 1094 | <.001 |
| Slope, $\beta_{10}$ | 0.70 | 0.03 | 21.93 | 1094 | <.001 |
| Special education, $\beta_{11}$ | −0.08 | 0.05 | −1.53 | 1094 | .126 |
| Sufficient measurement, $\beta_{12}$ | −0.10 | 0.05 | −2.03 | 1094 | .043 |

- **Patterns of Statewide Test Participation for Students with Significant Cognitive Disabilities**

  - *Journal of Special Education* **(in press)**
  - Jessica L. Saven, Daniel Anderson, Joseph F. T. Nese, Dan Farley, Gerald Tindal

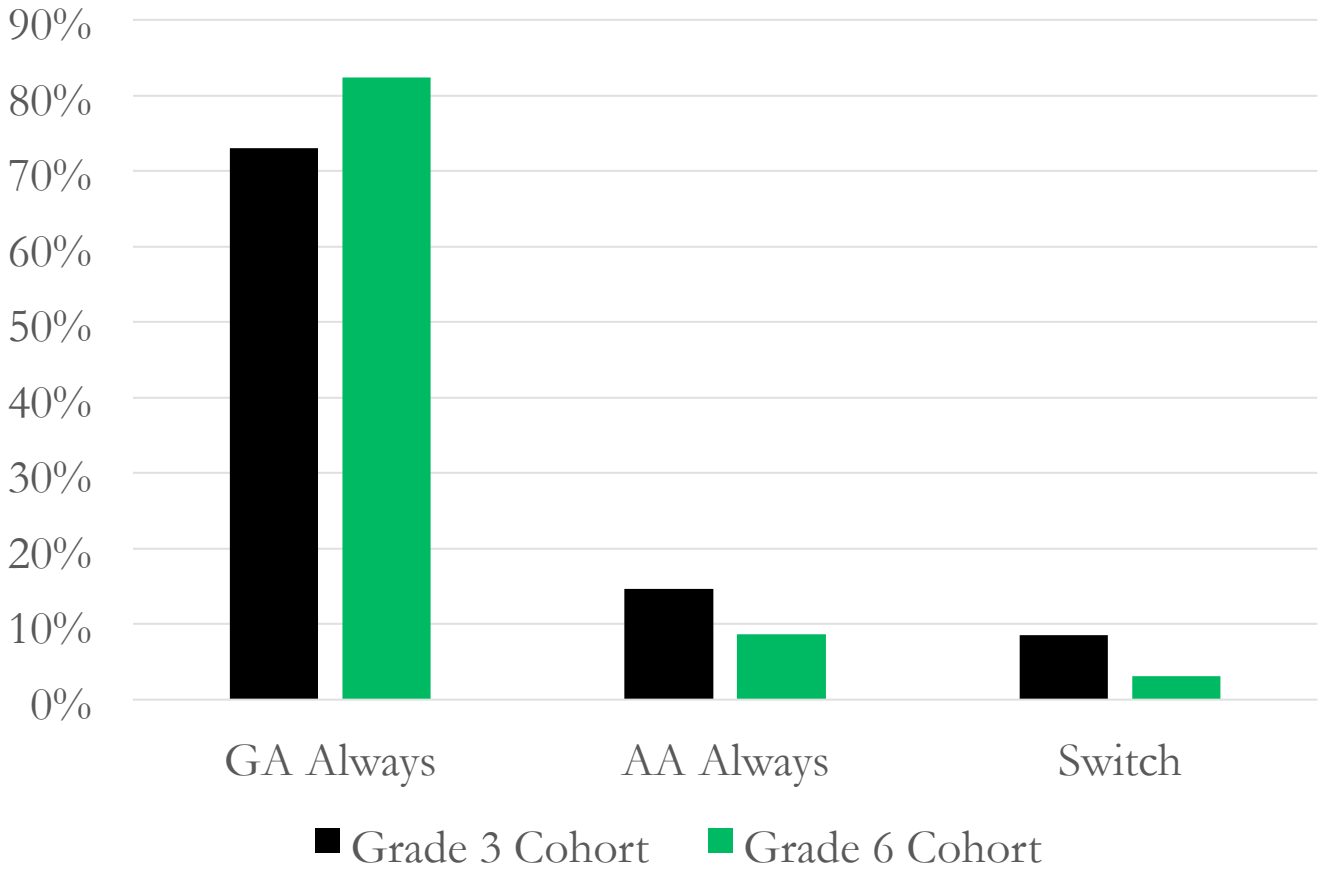
# Patterns of Participation: General (GA) or Alternative Assessment (AA)

- States may develop AA based on alternate achievement standards.
- 9% of SWD, or 1% of all students.
- AA must meet technical adequacy requirements, and link with state academic content standards.
- Eligibility criteria and implementation vary; so some students "switch" test types between years.
- Accountability implications.

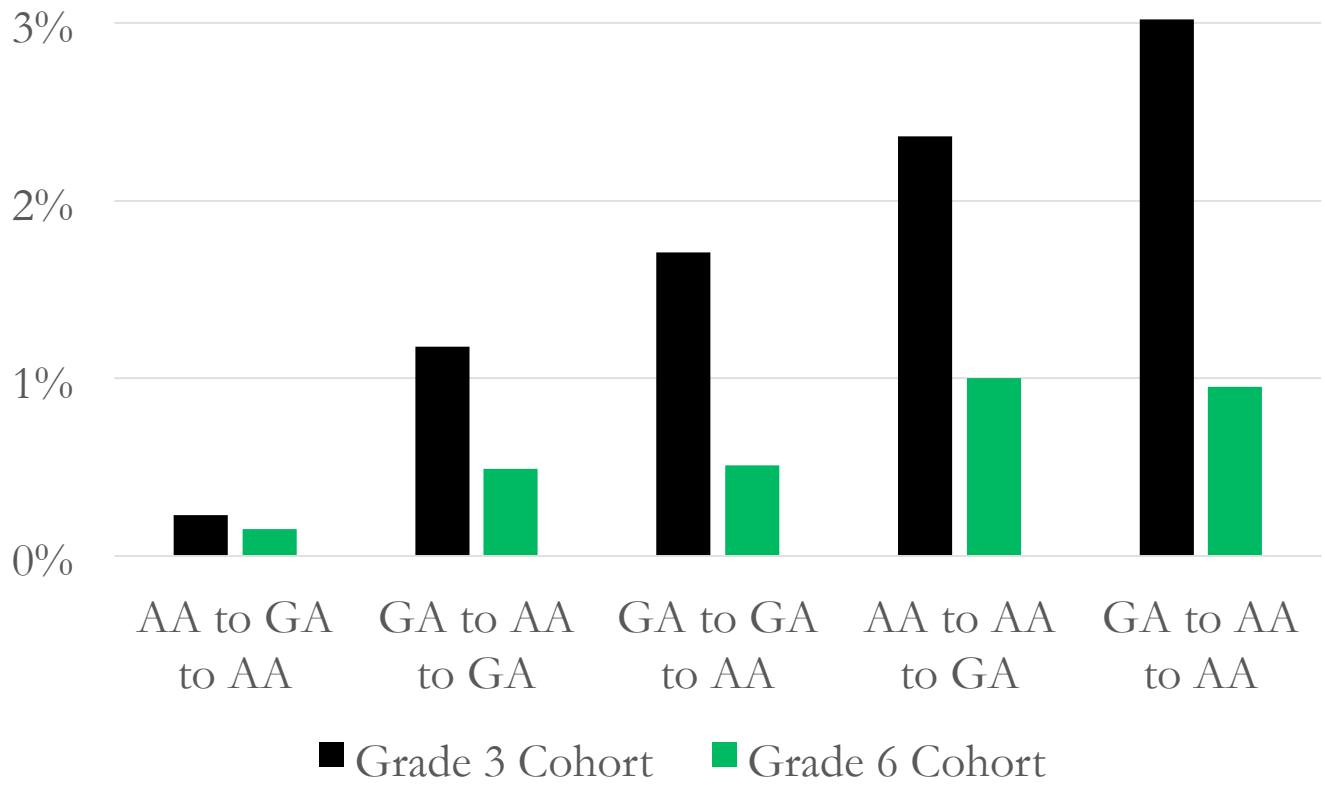
# Patterns of Participation: General (GA) or Alternative Assessment (AA)

- We tracked test participation for two cohorts (elementary and middle school) of students with a documented disability over three years.
  - Students with intellectual disabilities (ID), autism (ASD), or learning disabilities (LD).
- Research Questions
  1. What is the likelihood of test switching on the reading portion of the AA and the GA across consecutive years, over a three-year span for students with intellectual disabilities, autism, or learning disabilities?
  2. Do students performing highly on the AA or poorly on the GA (i.e., students on the "bubble") have an increased likelihood of switching test type as compared to other students with the same disability?
  3. Is the observed pattern the same across cohorts of students in middle school as compared to elementary school?

# Grade 3 ($n$ = 3,048) & Grade 6 ($n$ =3,911) Cohort Test Patterns 2009/10 – 2011/12



# Grade 3 ($n$ = 3,048) and Grade 6 ($n$ =3,911) Cohort Test Patterns 2009/10 – 2011/12



# Implications

Mechanisms must be found to include SWSCDs and ensure appropriate participation in the testing program over time. Otherwise, high percentages of students switching test types necessarily limit the accuracy of estimates of growth for these students and complicates interpretations of students' levels of proficiency and growth.

Exceptional Children
1–16

DOI: 10.1177/0014402915585492
ec.sagepub.com
SAGE

# Documenting Reading Achievement and Growth for Students Taking Alternate Assessments

**Gerald Tindal[1], Joseph F. T. Nese[1], Dan Farley[1], Jessica Saven[2], and Stephen N. Elliott[2]**

## Abstract

Students with disabilities have been included in state accountability systems for more than a decade; however, only in the past few years have forms of alternate assessments of alternate achievement standards (AA-AAS) become stable enough to allow for an examination of these students' achievement growth. Using data from Oregon's AA-AAS for Reading during the period 2008–2009 to 2010–2011, we examined the achievement growth for a sample of 1,061 elementary students using two growth models: a transition matrix and a multilevel linear growth model. We found with the transition matrix model that a majority of students remained at the same performance level from one year to the next, whereas with the multilevel linear growth model, students' scores revealed small, but statistically meaningful, growth year to year. We conclude by noting advantages and disadvantages of these models to characterize growth and their implications for policy and practice.

**Table 6.** Unconditional and Final Conditional Model Parameters With Robust Standard Errors.

| | Unconditional model | | | Final model | | |
|---|---|---|---|---|---|---|
| Fixed effects | Coefficient | *SE* | *p* value | Coefficient | *SE* | *p* value |
| Intercept, $\beta_{00}$ | 102.15 | 0.61 | <.001 | 86.30 | 1.46 | <.001 |
| Sex, $\beta_{01}$ | | | | −1.72 | 0.90 | .056 |
| Ethnicity, $\beta_{02}$ | | | | −1.08 | 0.86 | .208 |
| EconDis, $\beta_{03}$ | | | | 3.04 | 1.04 | .004 |
| Intellectual disability, $\beta_{04}$ | | | | 0.13 | 1.30 | .918 |
| GenEd40, $\beta_{05}$ | | | | −5.34 | 1.03 | <.001 |
| PerfLevel, $\beta_{06}$ | | | | 26.35 | 1.12 | <.001 |
| Slope, $\beta_{10}$ | 4.64 | 0.24 | <.001 | 6.30 | 0.65 | <.001 |
| Sex, $\beta_{11}$ | | | | −0.50 | 0.48 | .299 |
| Ethnicity, $\beta_{12}$ | | | | −0.09 | 0.50 | .863 |
| EconDis, $\beta_{13}$ | | | | 0.93 | 0.53 | .078 |
| Intellectual disability, $\beta_{14}$ | | | | −0.49 | 0.58 | .397 |
| GenEd40, $\beta_{15}$ | | | | −1.86 | 0.52 | <.001 |
| PerfLevel, $\beta_{16}$ | | | | −1.89 | 0.52 | <.001 |

*Note*. EconDis = economic disadvantage; GenEd40 = program placement; PerfLevel = performance level.

**Table 4.** Transition Matrix From Grade 4 to Grade 5.

| | Grade 5 (2010–2011) | | | |
|---|---|---|---|---|
| Grade 4 (2009–2010) | Low | Nearly meets | Meets | Exceeds |
| Low | 168 (26.0) | 14 (2.2) | 3 (0.5) | 0 (0) |
| Nearly meets | 44 (6.8) | 49 (7.6) | 29 (4.5) | 4 (0.6) |
| Meets | 15 (2.3) | 28 (4.3) | 101 (15.6) | 49 (7.6) |
| Exceeds | 3 (0.5) | 2 (0.3) | 42 (6.5) | 96 (14.8) |

*Note*. $N = 647$. Frequencies shown with percentages in parentheses. There were no students in the lowest level (*very low*).

**Table 5.** Level Change From Grade 4 to Grade 5.

| −3 | −2 | −1 | 0 | +1 | +2 |
|---|---|---|---|---|---|
| 3 (0.5) | 17 (2.6) | 114 (17.6) | 414 (64.0) | 92 (14.2) | 7 (1.1) |

*Note*. $N = 647$. Frequencies shown with percentages in parentheses.

# Upcoming NCAASE Studies

Joe Stevens
University of Oregon

# School Performance

- One of our central goals is to compare different models of estimating school performance
- We will compare commonly used models of school performance to determine how model choice and model characteristics impact characterizations of school performance
- We begin this work this fall using Oregon data
- We will then replicate using AZ, NC, and PA data

# Models of School Performance

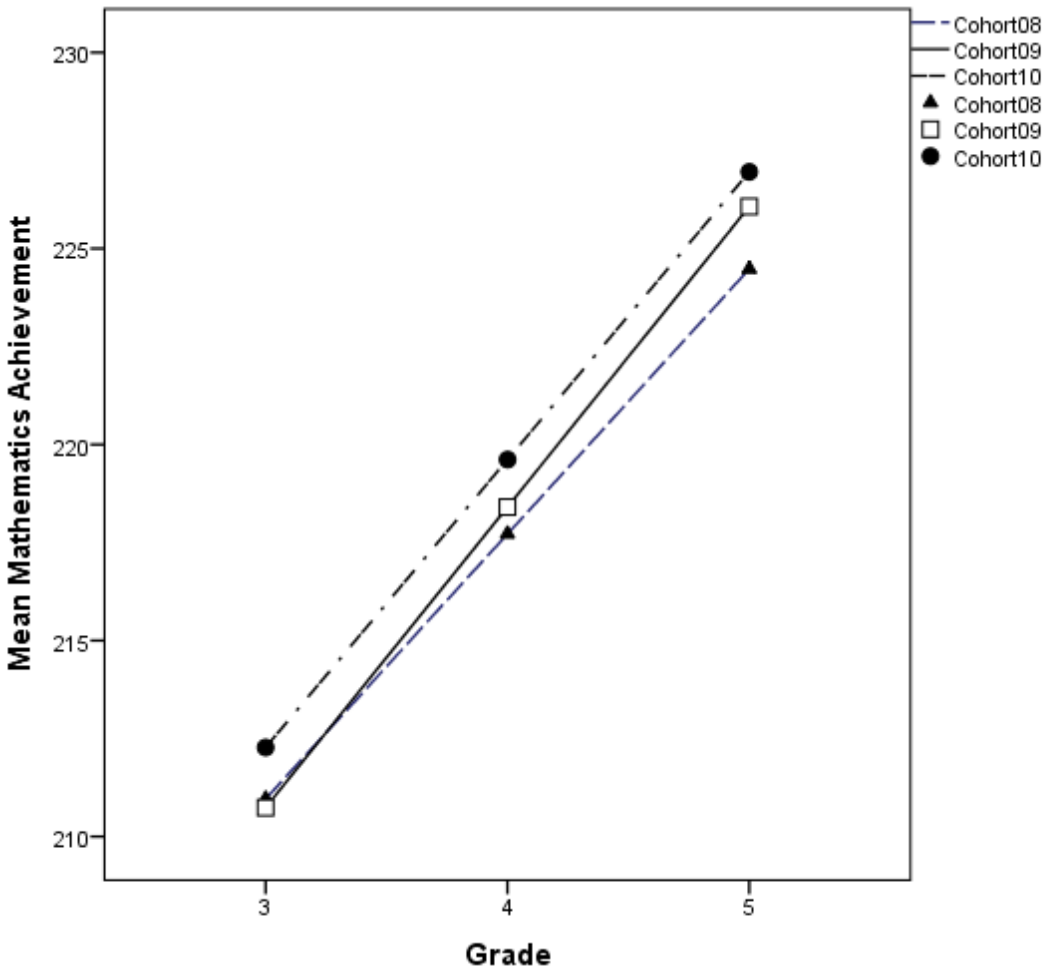
- Status; gain and residual scores; projection models
- Transition matrix
- Value-added models
- Student Growth Percentiles
- Hierarchical linear growth models
- Latent Growth curve models

# Model Variations

- Two grade levels studied: elementary schools, middle schools
- Focused study of impact of models on schools serving SWD
- Three cohorts studied for each analysis to determine cohort stability
- Unconditional vs. conditional models (school size, student composition of school)
- For some models different estimation methods examined (OLS, EB, Fully Bayesian)

# Some Preliminary Results

- There are substantial differences in estimates of school effects depending on the model chosen
  - SGPs almost perfectly correlated with residual regression
  - Low to moderate correlation of SGPs with HLM growth
  - Low correlation of status models with HLM Growth
- There are significant differences in model results by cohort, that is, results vary from year to year undermining any estimates based on a single year or single cohort


Mean Mathematics Achievement
230
225
220
215
210
3
4
5
Grade
Cohort08
Cohort09
Cohort10
Cohort08
Cohort09
Cohort10

# Research on the Academic Growth of Students with Disabilities and its Implications for Educational Policies and Practices

National Conference on Student Assessment

San Diego, CA

June 23, 2015

Thank You for Your Participation

NCAASE National Center on Assessment and Accountability for Special Education

*Advancing research on growth measures, models, and policies for improved practice*