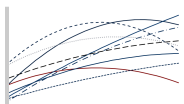


---

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

---

**George Washington University  
Washington, DC  
Wednesday, October 22, 2014**



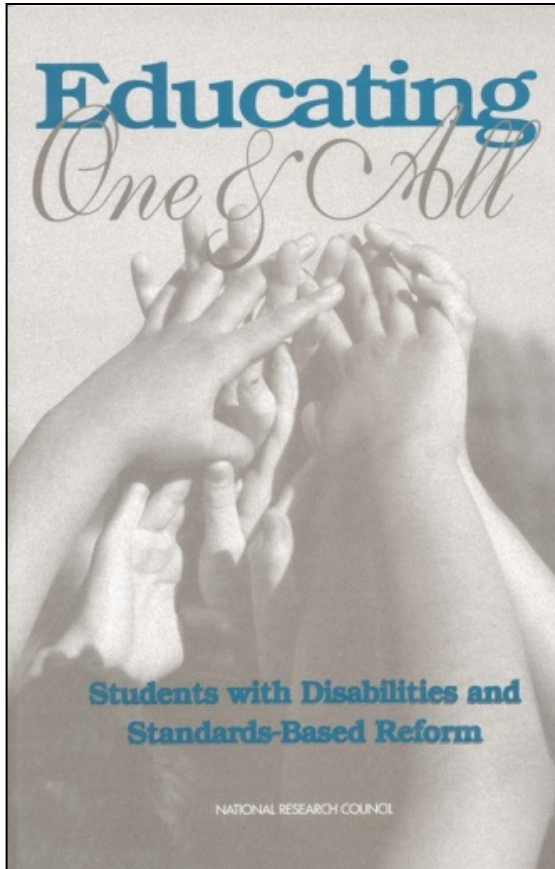
---

# **Educating One & All 1994-2014: Context for NCAASE Research and Its Implications**

---

**Stephen N. Elliott  
Arizona State University**

# Educating One & All Report 1994-1997



(McDonnell, McLaughlin, & Morison, 1997)

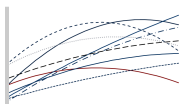
- The National Academy of Sciences' Committee on Education Goals 2000 and Services to Student with Disabilities was established "to conduct a comprehensive study of the inclusion of children with disabilities in school reforms assisted under Goals 2000: Educate America Act" (PL103-227, sec. 1015).
- EOA report, *"There is a scarcity of research evidence directly bearing on the effects of standards-based reforms, much less their impact on students with disabilities. In addition, the research base on instructional practices and achievement contains few studies that include populations of students with and without disabilities, making systematic comparisons difficult."* (Executive Summary, 1997, p. 2)

# EOA Recommendations

- The EOA Committee offered 12 recommendations, of which several are highly relevant to the work of NCAASE. Recommendation #12 called for
  - “a long-term research agenda to address the substantial gaps in knowledge about the schooling of students with disabilities and the impact of standards-based reforms. Areas needing particular attention include research on the school experiences of students with disabilities, the potential of computer-based technologies, how local decisions are made about students’ curricular opportunities, alternative student credentials, and the relationship between testing accommodations and validity.” (p. 209)
- There was virtually no discussion in the EOA Report about measuring students’ achievement growth.

# Since the EOA Report

- Over the past 12 years, there has been a shift in accountability models in special education from one that emphasized compliance with policies and procedures set forth in the IDEA, to one that focuses on measuring student outcomes to gauge the quality of academic programming provided to students with disabilities.
- Yet, students with disabilities (SWDs) continue to lag significantly behind their peers without disabilities. For example, many states report that over 70% of SWDs perform below proficiency on annual statewide reading and mathematics tests (Center on Educational Policy, 2009). Similarly, the NAEP results for 2013 indicated only 8% and 7% of 8<sup>th</sup> graders with disabilities performed at or above the proficient level on the NAEP reading and mathematics assessments, respectively.



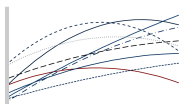
# Shift from Achievement Status to Growth

- A number of educational scientists have argued for shifting our accountability metric away from achievement status to students' achievement growth (e.g., Betebenner, 2008; Hanushek & Raymond, 2005; Schulte & Villwock, 2004).
- The argument for achievement growth is based on the dual premise that (a) schools should be held accountable for achievement outcomes they can control, rather than their prior achievement, and (b) status models incentivize schools to focus on students near the threshold of proficiency rather than focusing on the achievement growth of *all* students, including those functioning well below these thresholds.
- With high levels of student participation and the systematic collection of individual student annual test data by states, it is possible to conduct longitudinal growth analyses of the academic achievement of all students. Thus, since 2007 the U.S. Department of Education has allowed states to develop growth models that provide longitudinal characterizations of student achievement.

# IES's Call for Research to Answer Questions

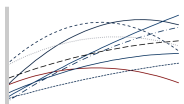
- In 2010, IES put out an RFA for a National Special Education Research and Development Center on Assessment and Accountability. It stated:

*“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?

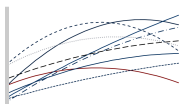
# Capacity

## ■ 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
<b>TOTAL</b>				<b>185,100</b>
<b>NWEA-PA</b>	2010-2012	3	11,577	<b>34,731</b>
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
<b>TOTAL</b>				<b>3,628,954</b>
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
<b>TOTAL</b>				<b>10,989,500</b>
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
<b>TOTAL</b>				<b>15,705,000</b>
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
<b>TOTAL</b>				<b>11,256,398</b>
<b>Grand Total</b>				<b>41,799,683</b>

# Resolution for NCAASE

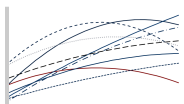
Findings are at the end of a long logic chain

- 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

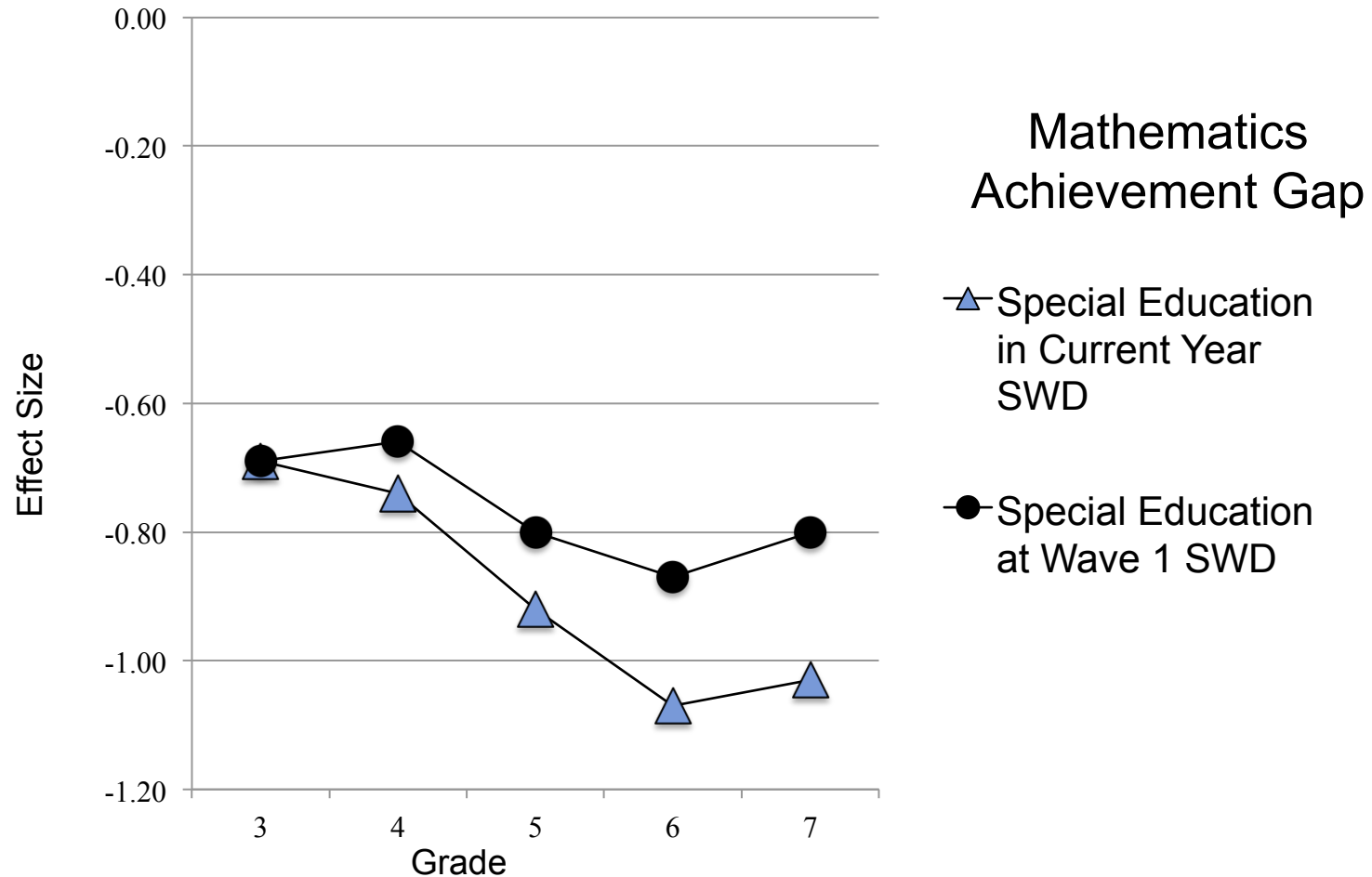
# Study 1: Once, Sometimes, or Always in Special Education

- 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**

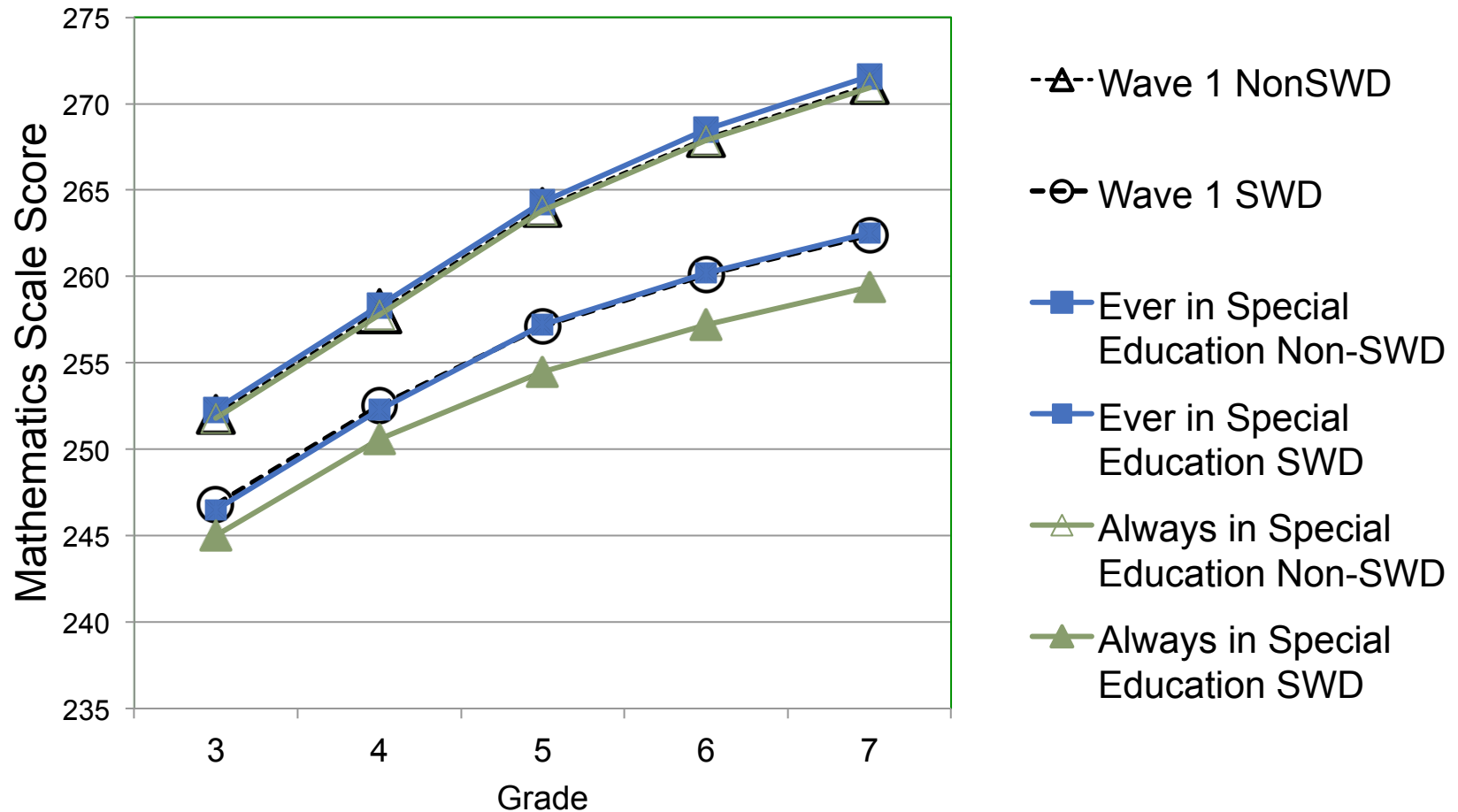
# Special Education Membership Grades 3-7

<b>SWD Subgroup Identification Method</b>	<b>Percent</b>
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



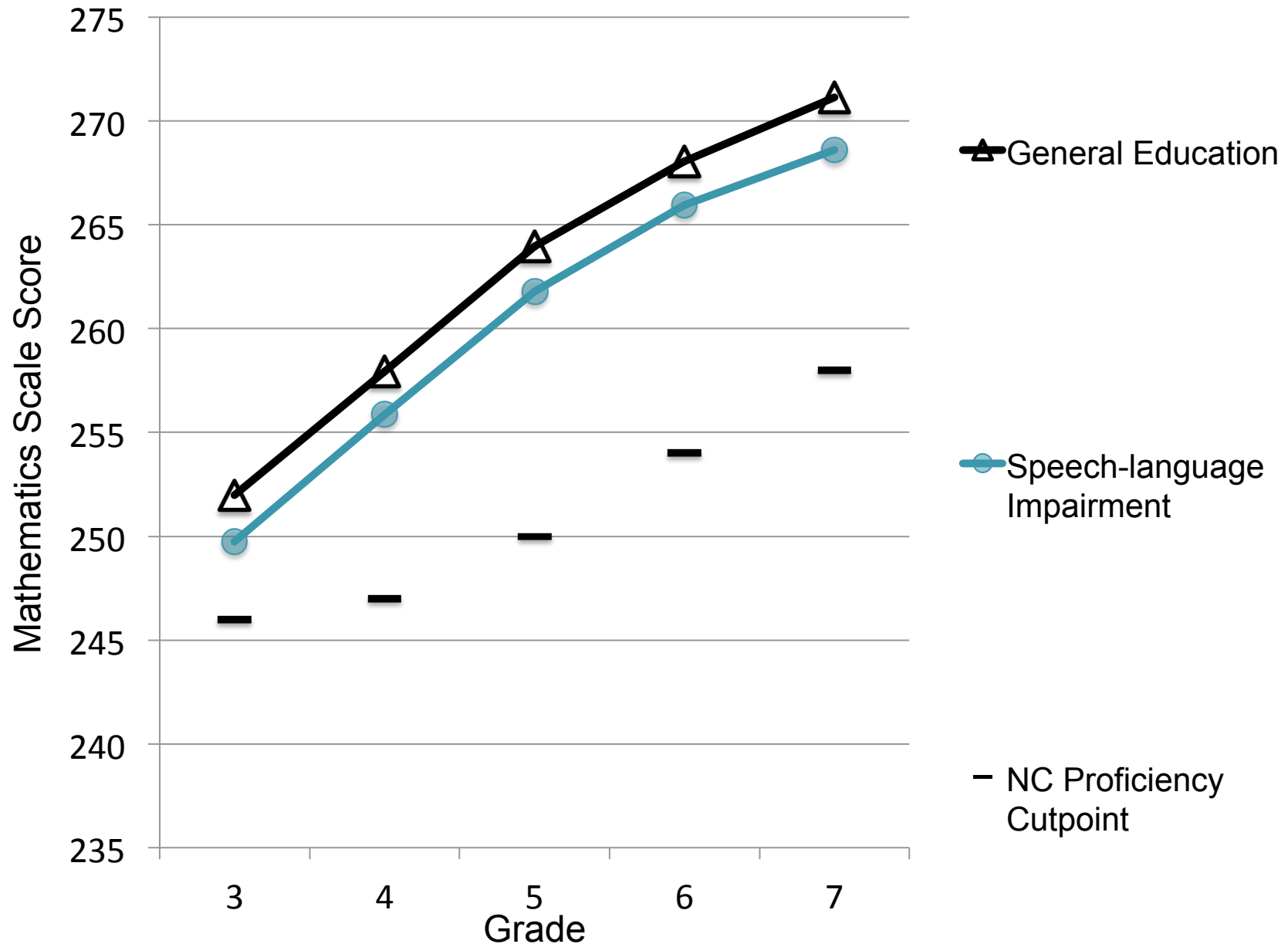
# Observed Means by SWD Identification Method



# Studies 2 & 3: Growth Across Grades

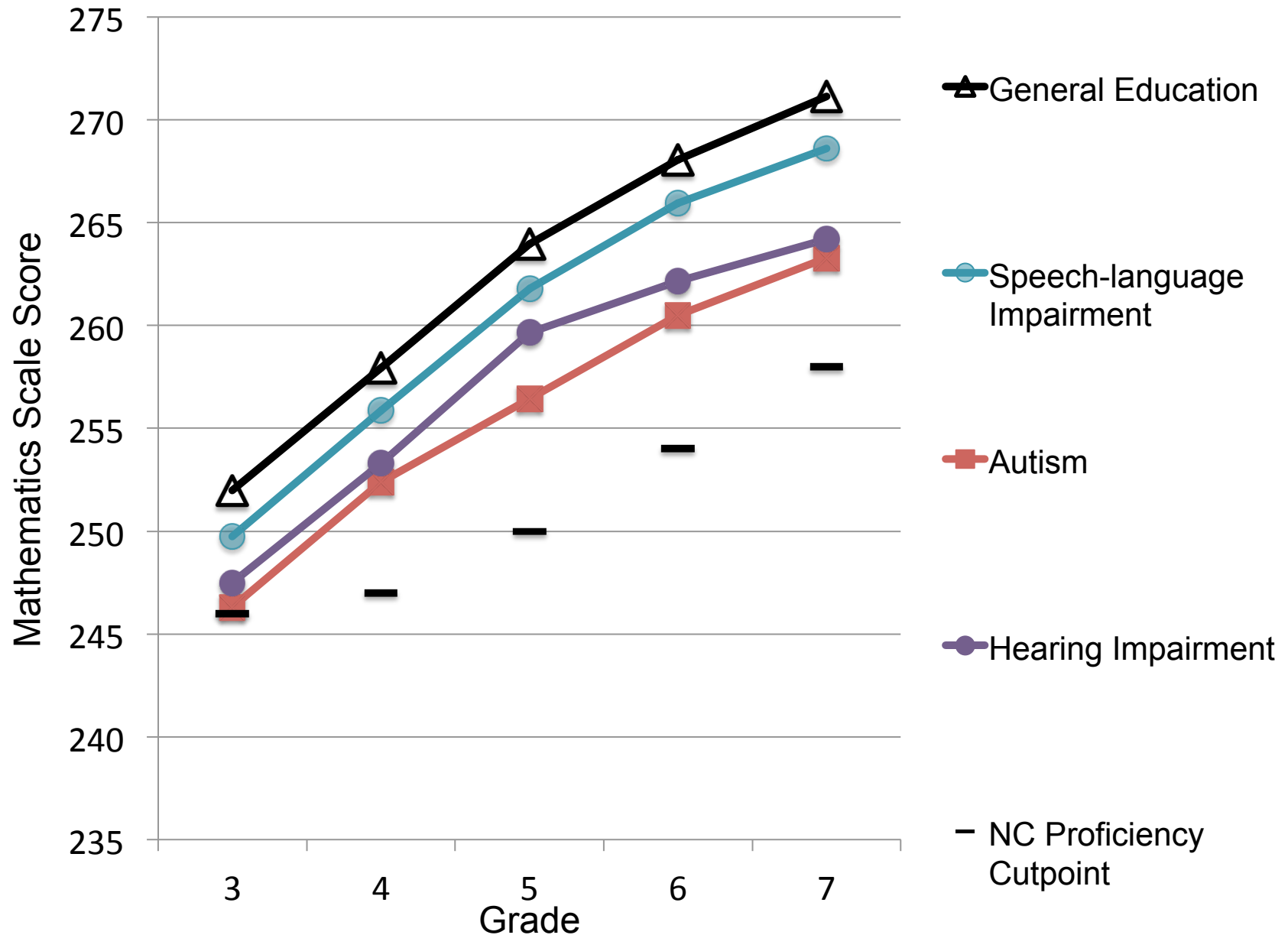
- Examine the developmental progress in mathematics and reading comprehension for general education students (GE) and students in specific exceptionality groups on a statewide achievement test
- Two longitudinal cohorts followed across grades three to seven
- Entire state cohorts,  $N > 100,000$ ;  $N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

# Mathematics Growth by Exceptionality

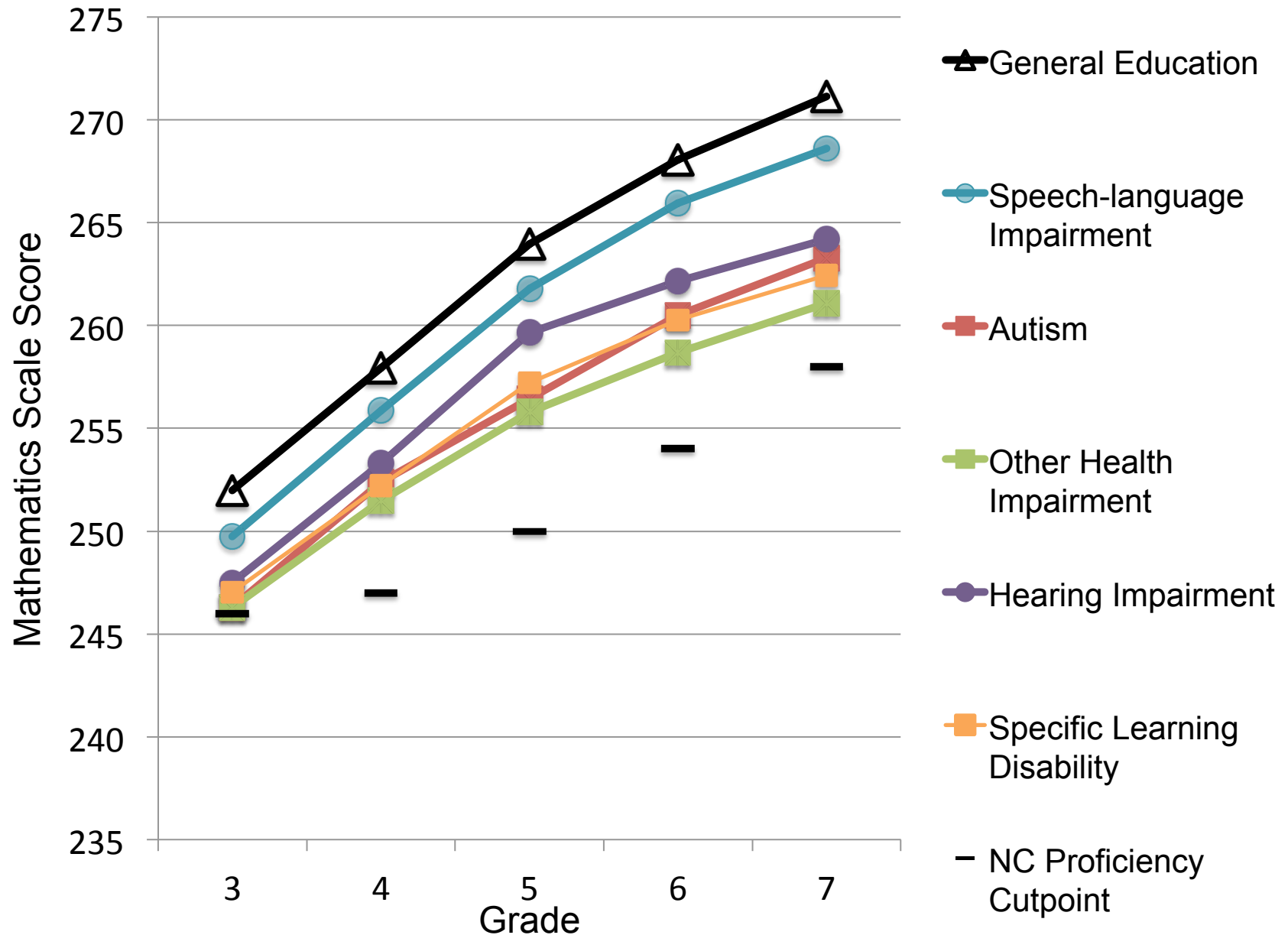




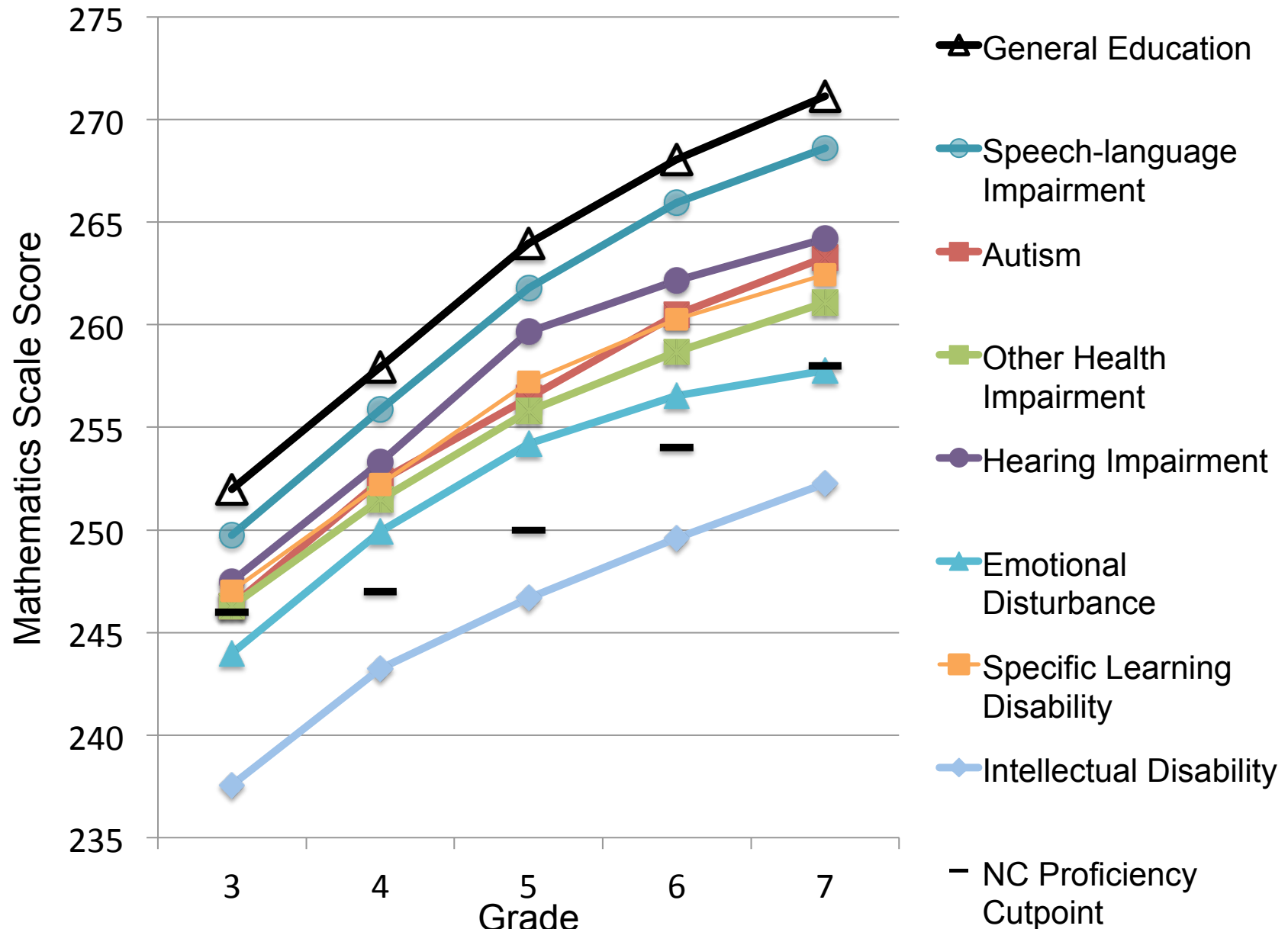
# Mathematics Growth by Exceptionality



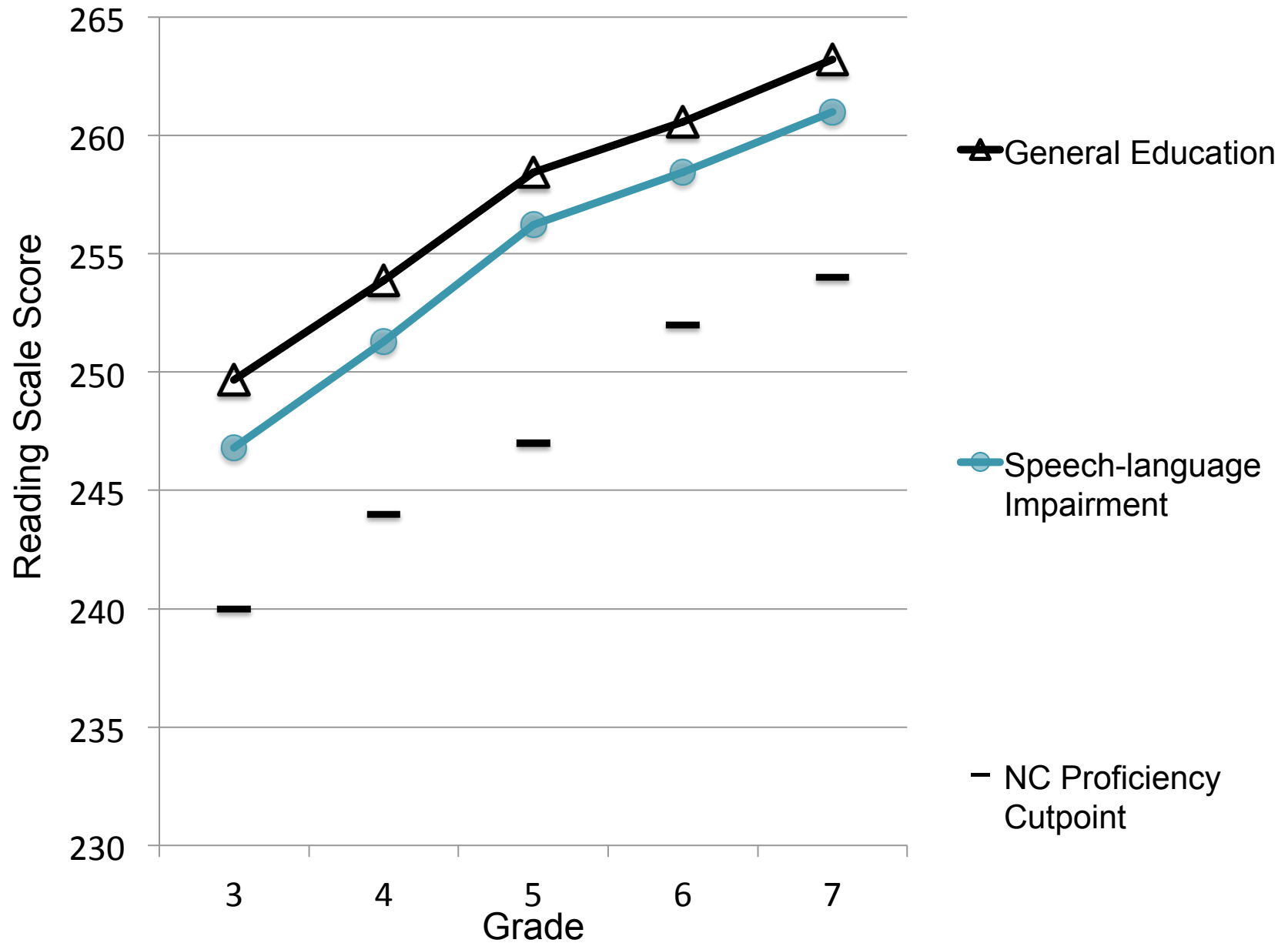
# Mathematics Growth by Exceptionality



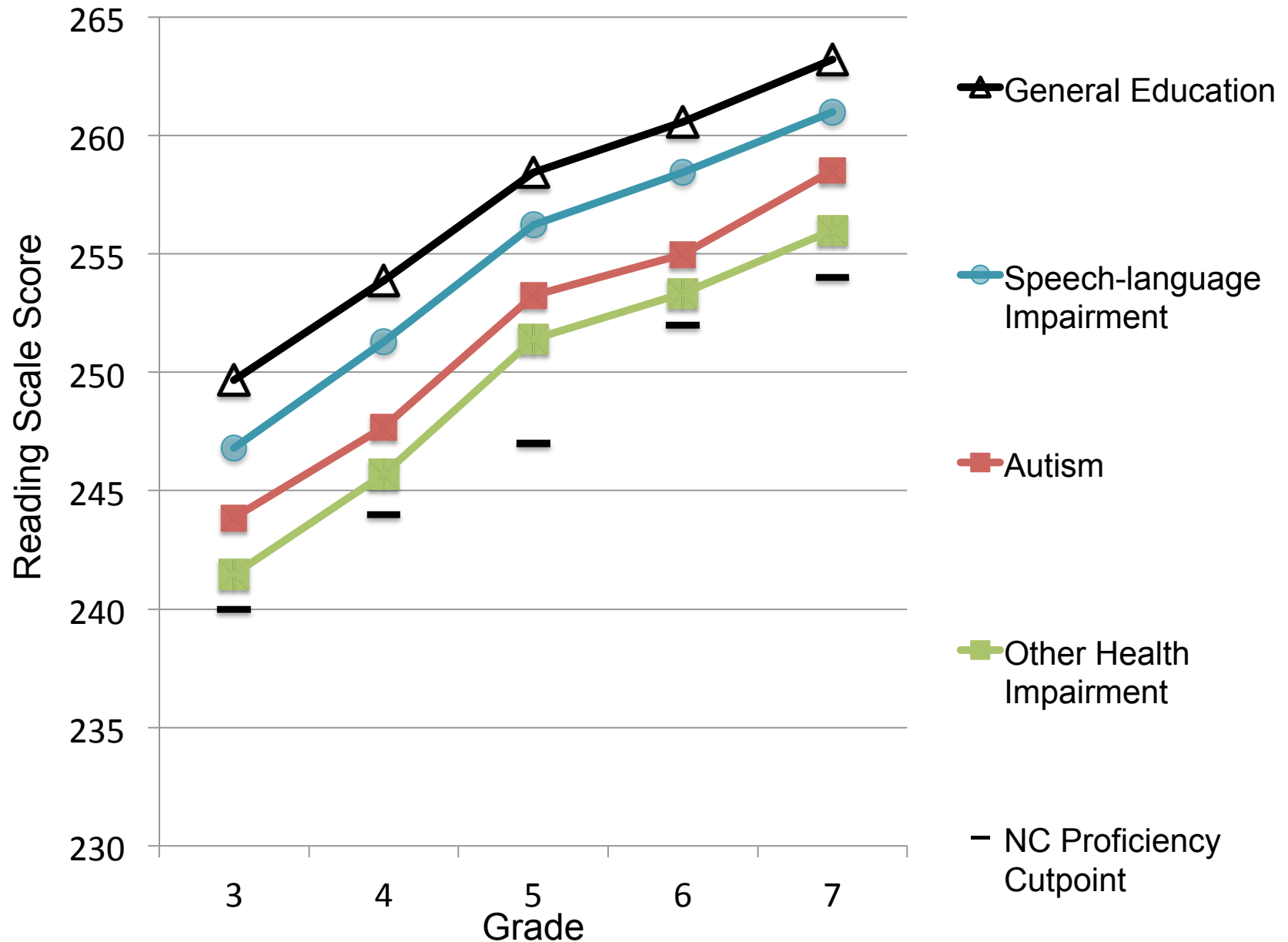
# Mathematics Growth by Exceptionality



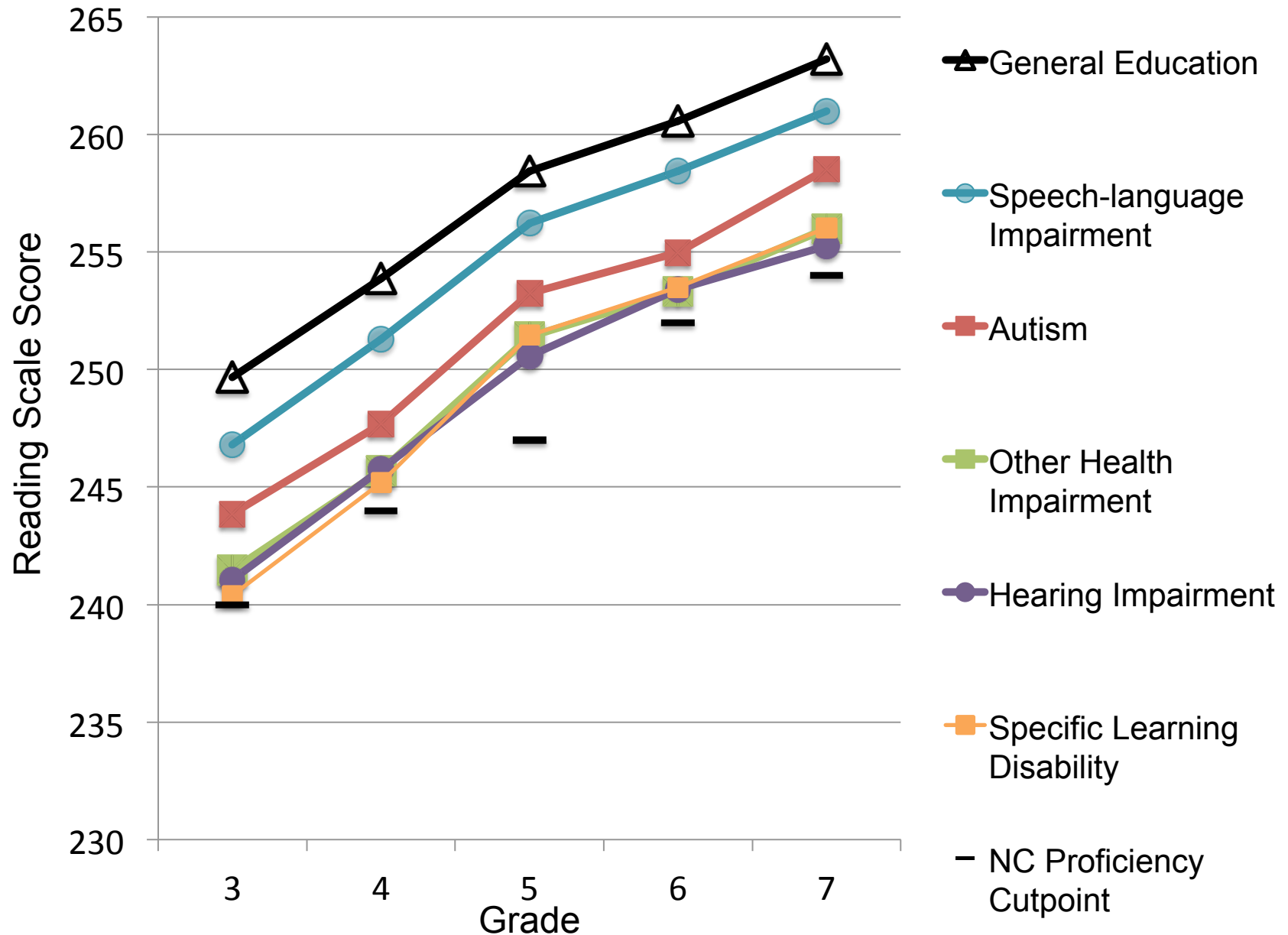
# Reading Growth by Exceptionality



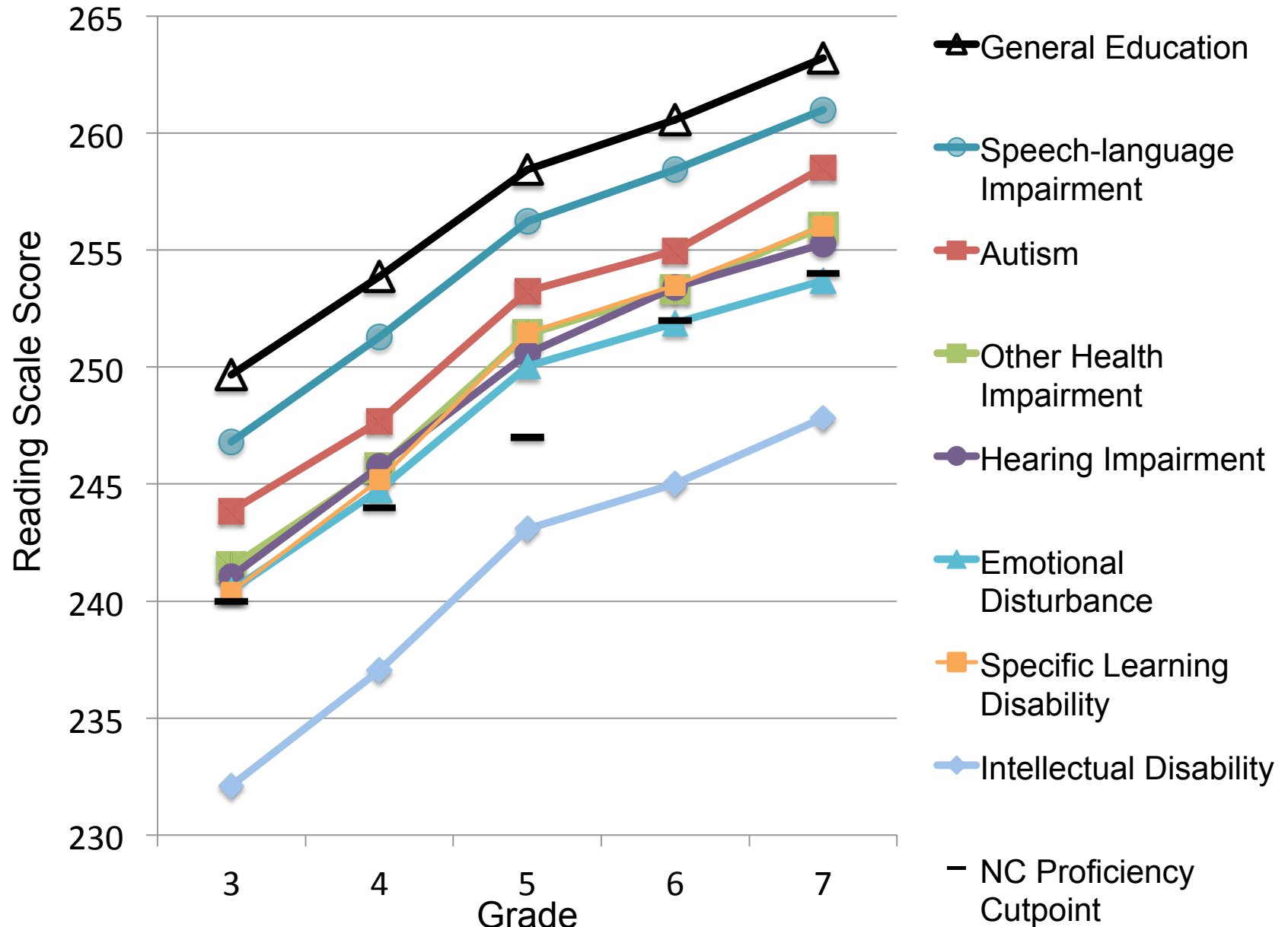
# Reading Growth by Exceptionality



# Reading Growth by Exceptionality



# Reading Growth by Exceptionality



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

Joe Stevens  
University of Oregon

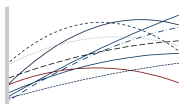


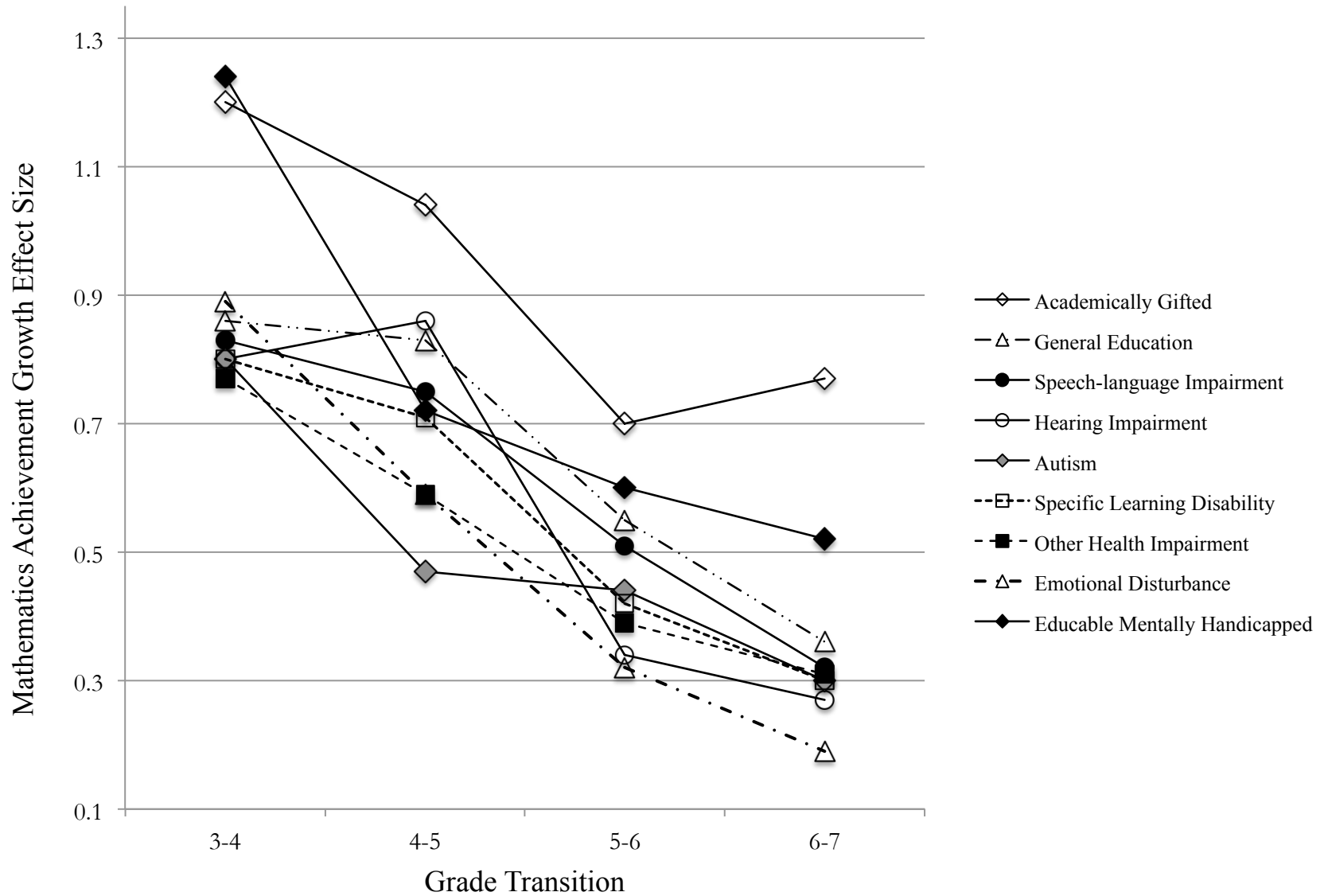
# Purpose

- 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

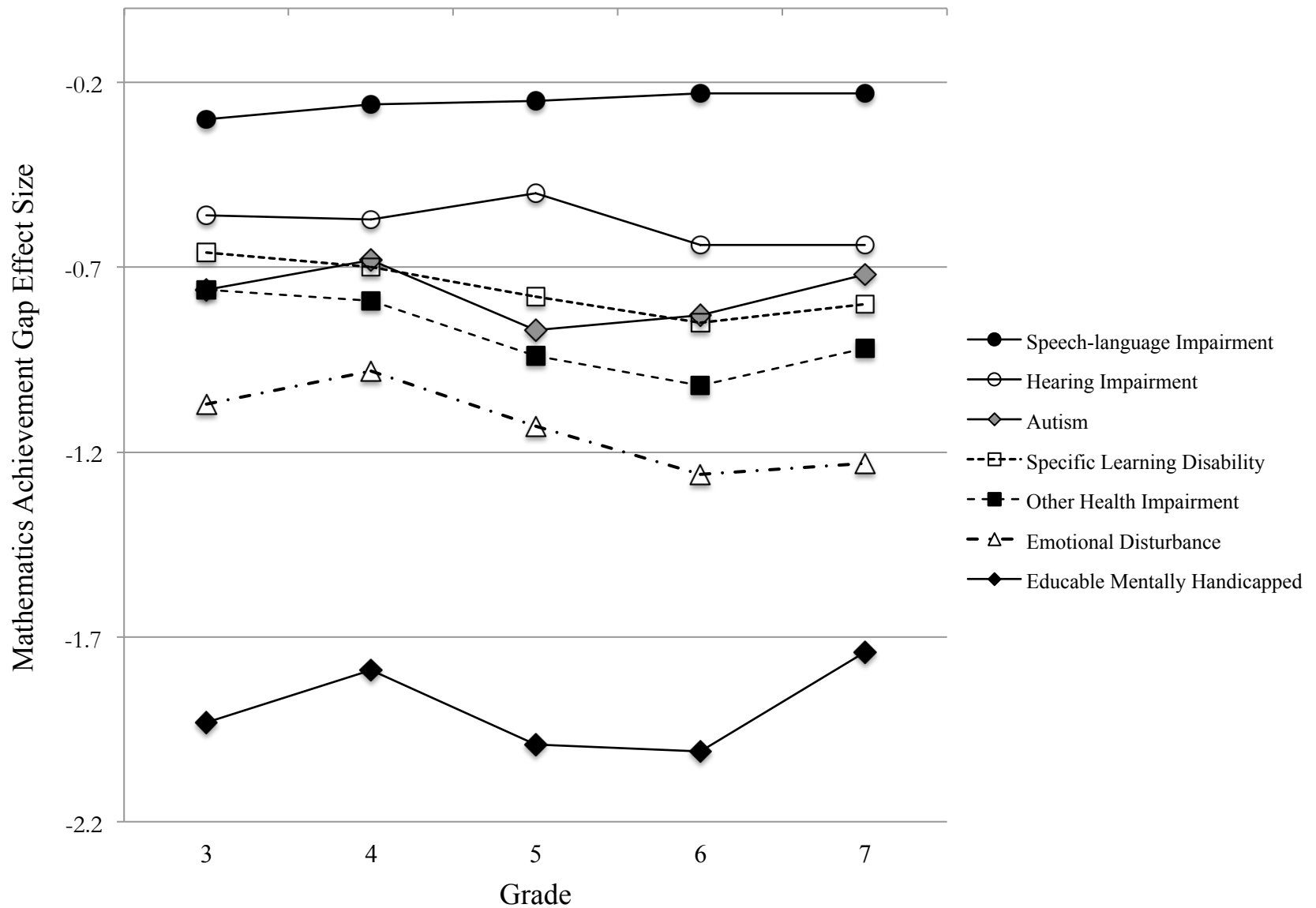
# Mathematics Achievement Gaps

- 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)





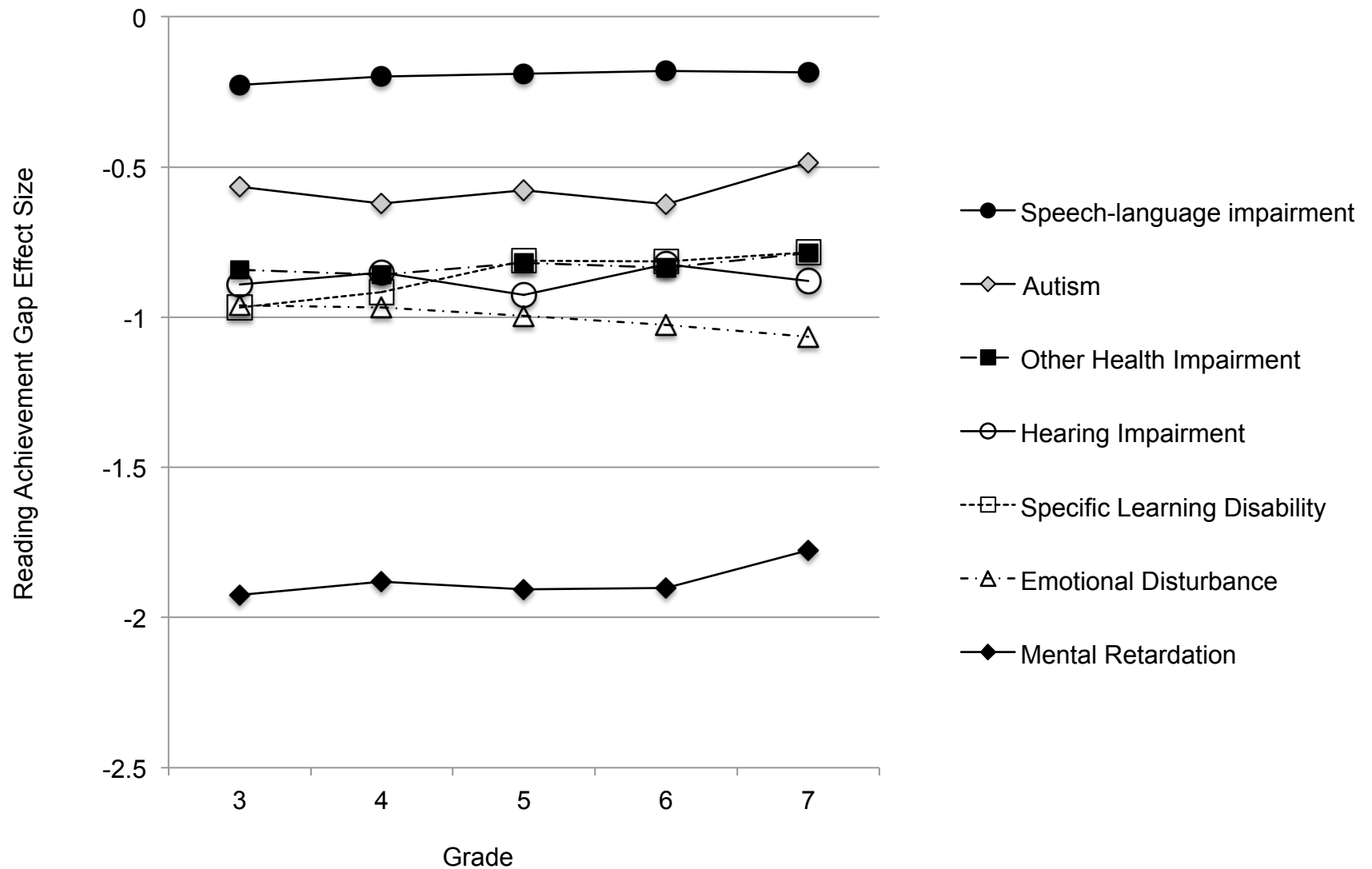
*Figure.* Mathematics achievement growth effect size at each grade transition by student group (from Stevens et al., in press).



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

# Reading Achievement Gaps

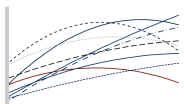
- What is the size of the achievement gap in reading for students in specific exceptionality categories?
- Does the gap increase, decrease or stay the same over time?

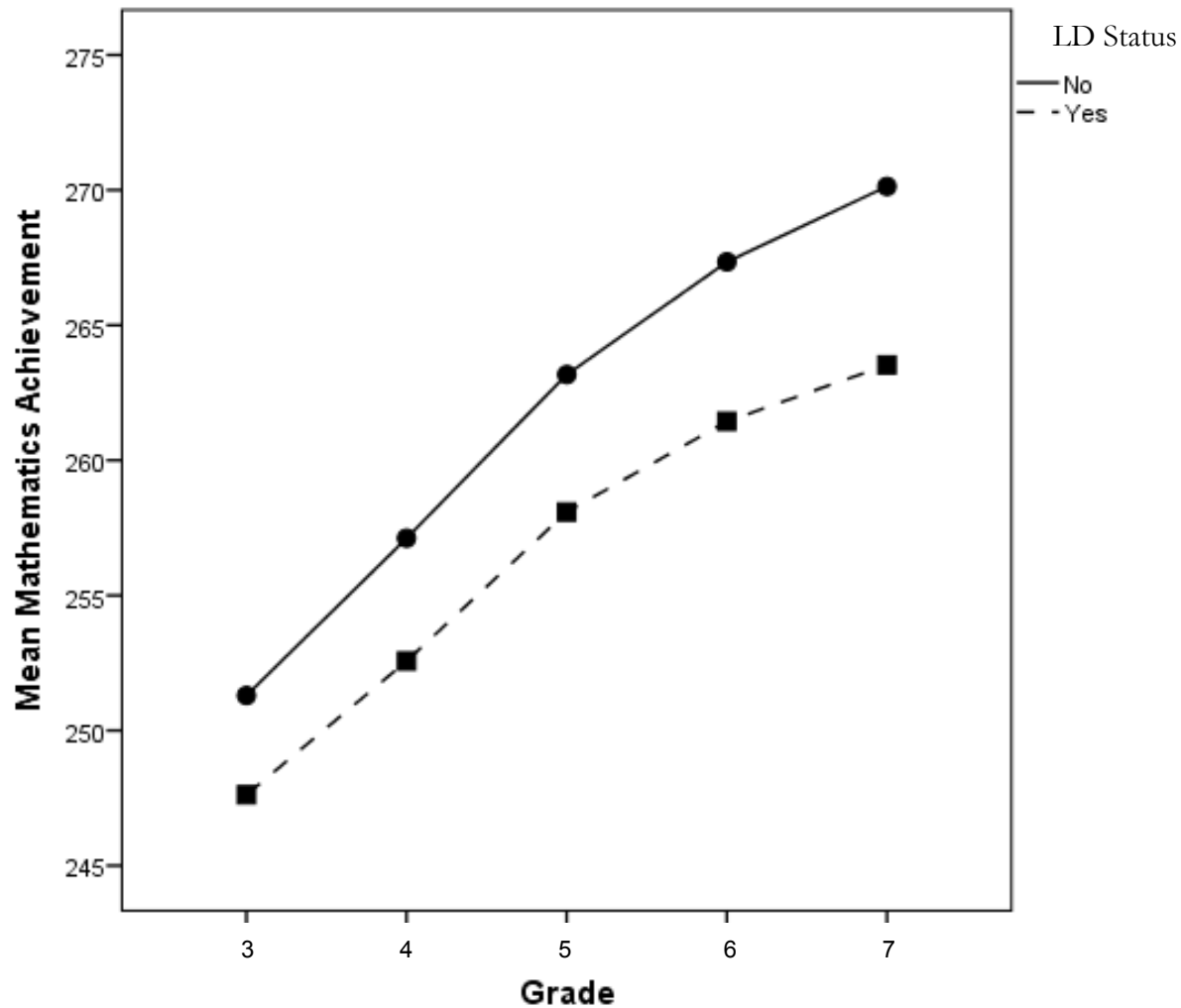


*Figure.* Reading achievement gap effect sizes between all SWoD and each exceptionality group by grade (from Schulte & Stevens, manuscript submitted for publication).

# Interactions of SWD status and Other Student Characteristics

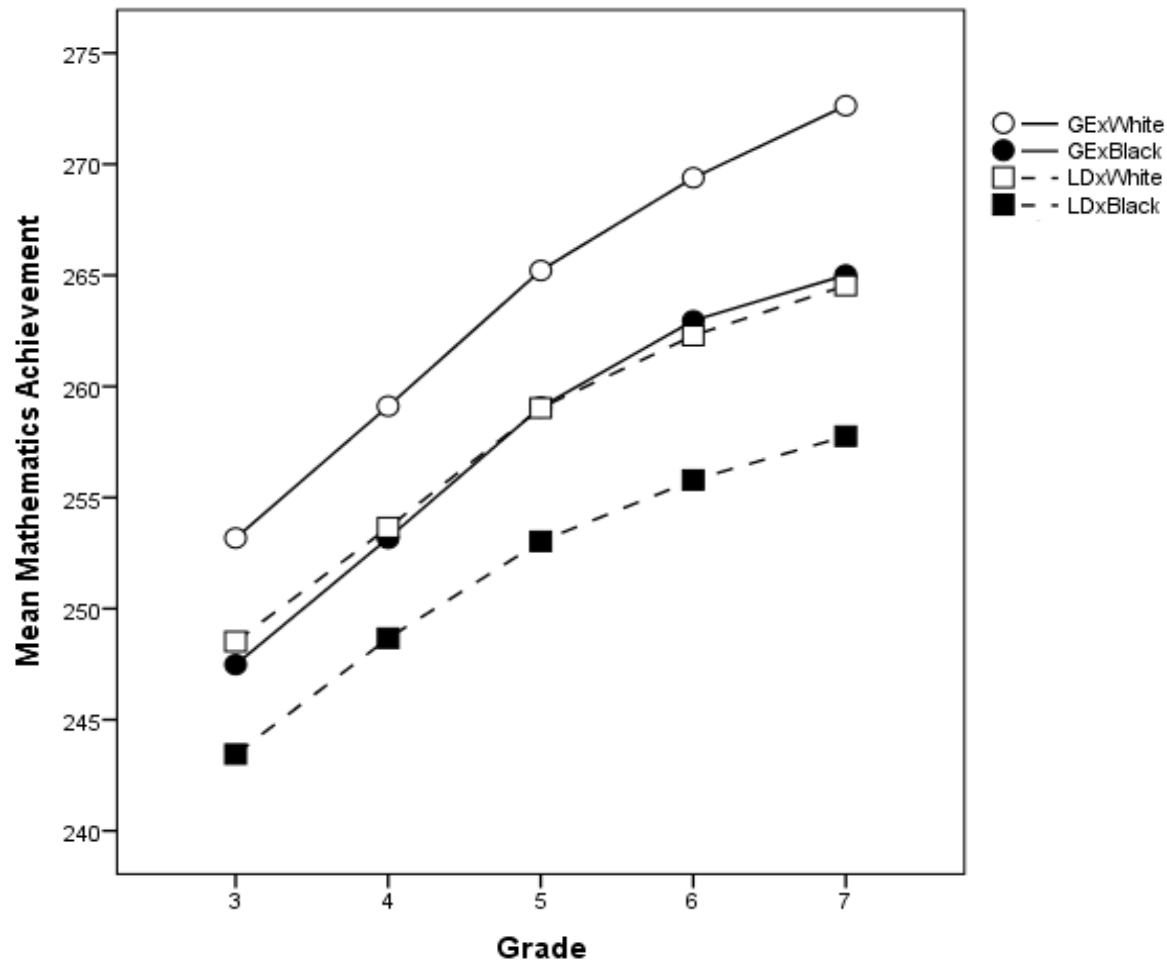
- Many studies do not directly test the interaction of SWD status and factors thought to be related to student performance (e.g., LD status and sex of student)
- When these factors are included in statistical models (especially regression and HLM models), only partial regression effects not the actual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



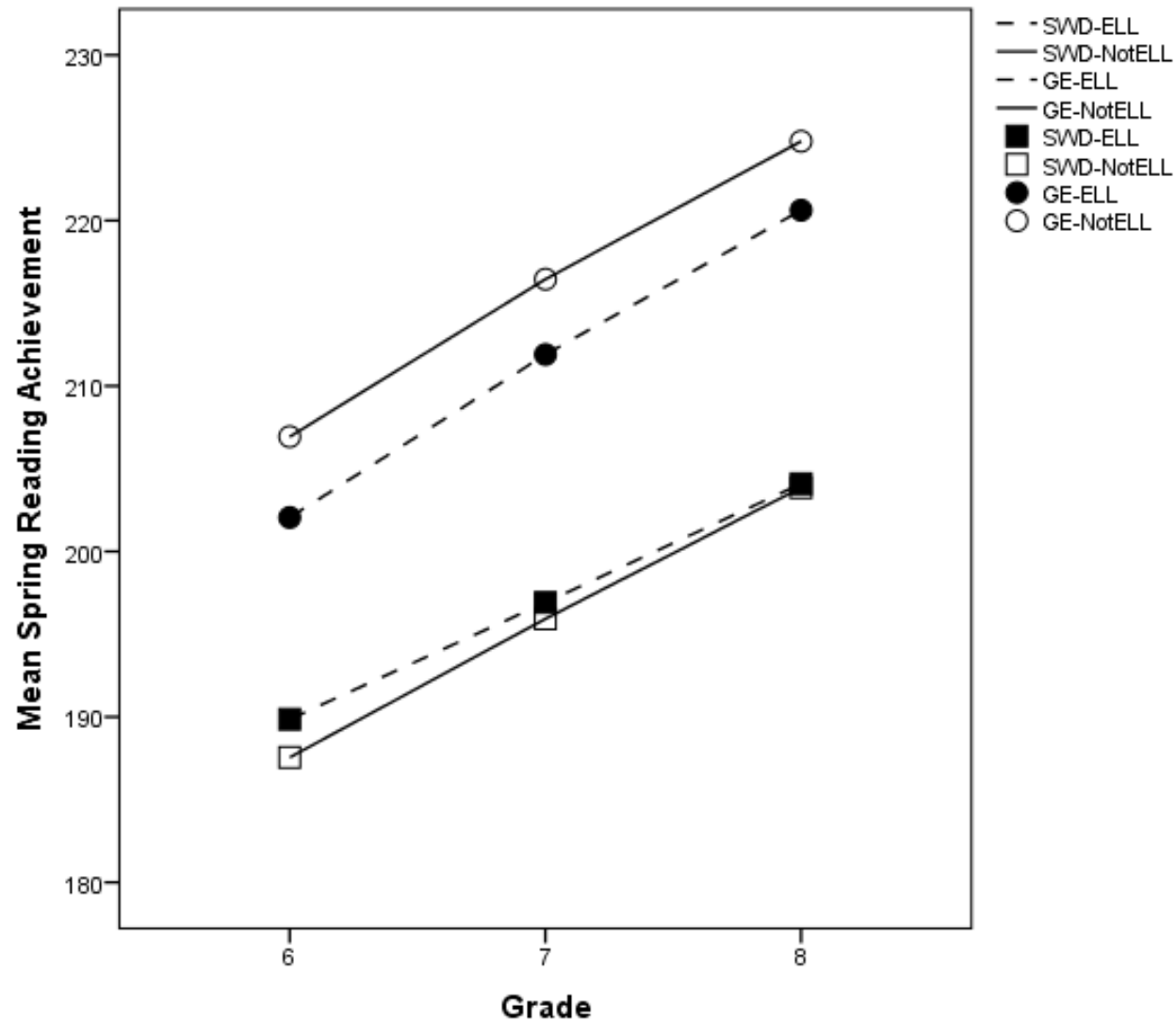


*Figure.* Mean mathematics achievement by grade and LD status (note increasing achievement gap; Stevens & Schulte, manuscript submitted for publication).





*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).



*Figure.* Interaction of SWD and ELL status on interim assessment performance for Tempe, AZ middle school students (from Stevens & Schulte, manuscript in preparation).

---

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

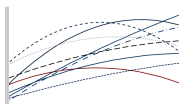
---

Stephen N. Elliott  
Arizona State University

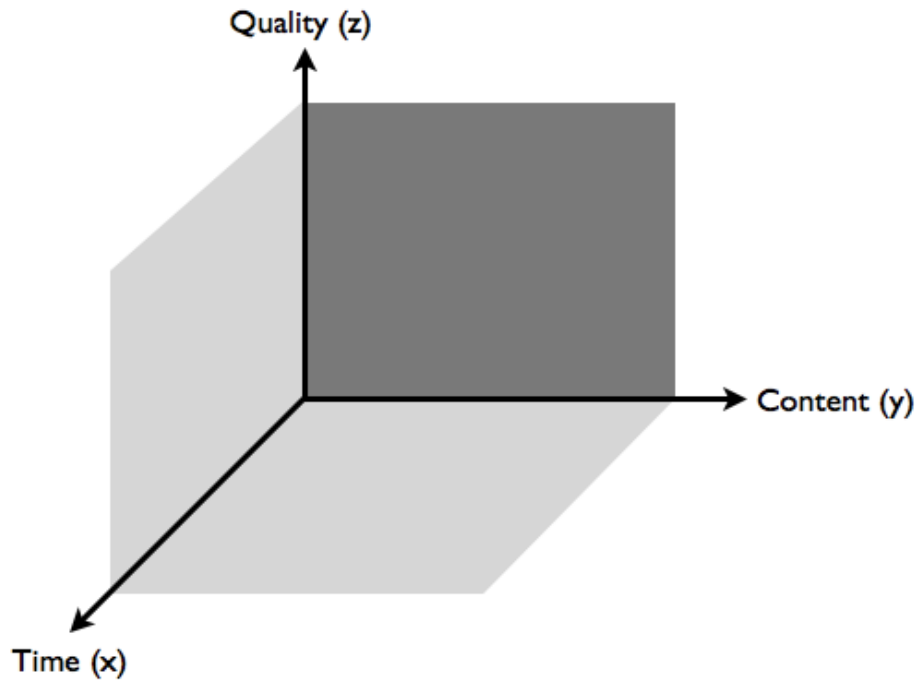
# Research Questions

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

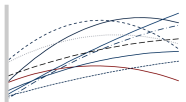


## Definition: Opportunity to Learn

*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.**



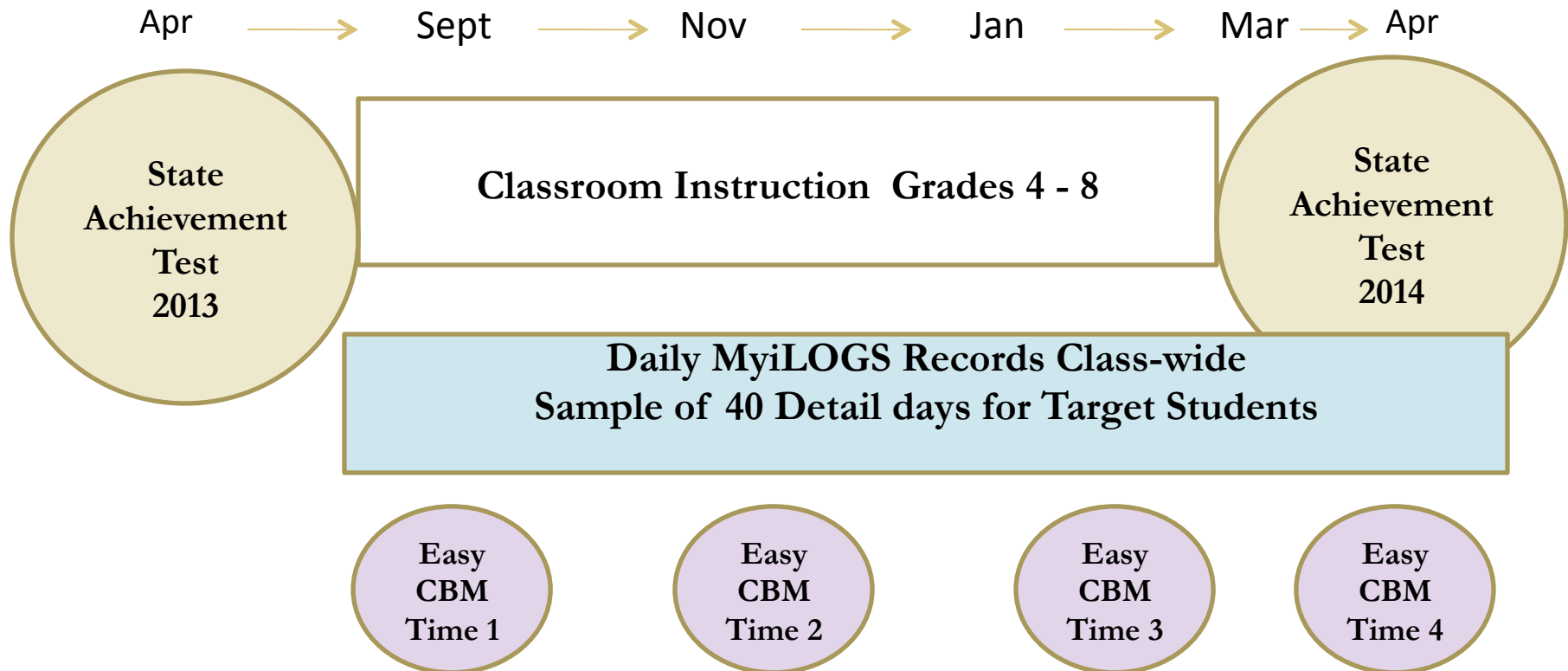
**NCAASE**

National Center on Assessment and  
Accountability for Special Education

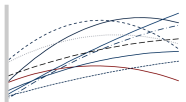
Advancing research on growth measures, models, and policies for improved practice

# Multiple Measures Study Design\*

Teachers (N = 67; AZ 35, OR 32) and students (N = 261; 136 SWD + 125 SWoD)  
from AZ & OR schools grades 4<sup>th</sup>-8<sup>th</sup>



\*A 3-year study with longitudinal student cohorts



**NCAASE**

National Center on Assessment and  
Accountability for Special Education

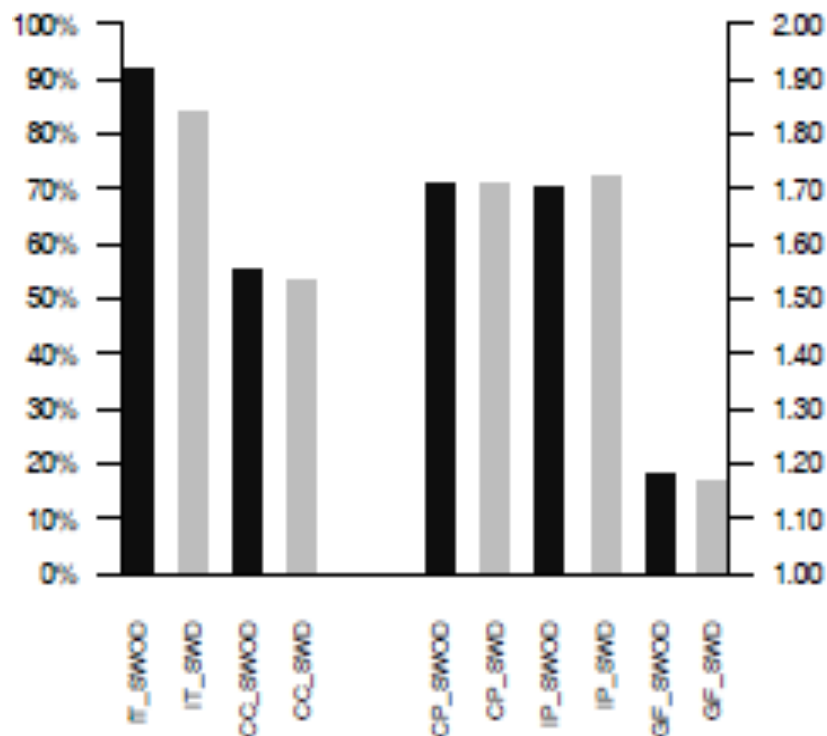
*Advancing research on growth measures, models, and policies for improved practice*

# Year 1 Findings

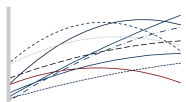
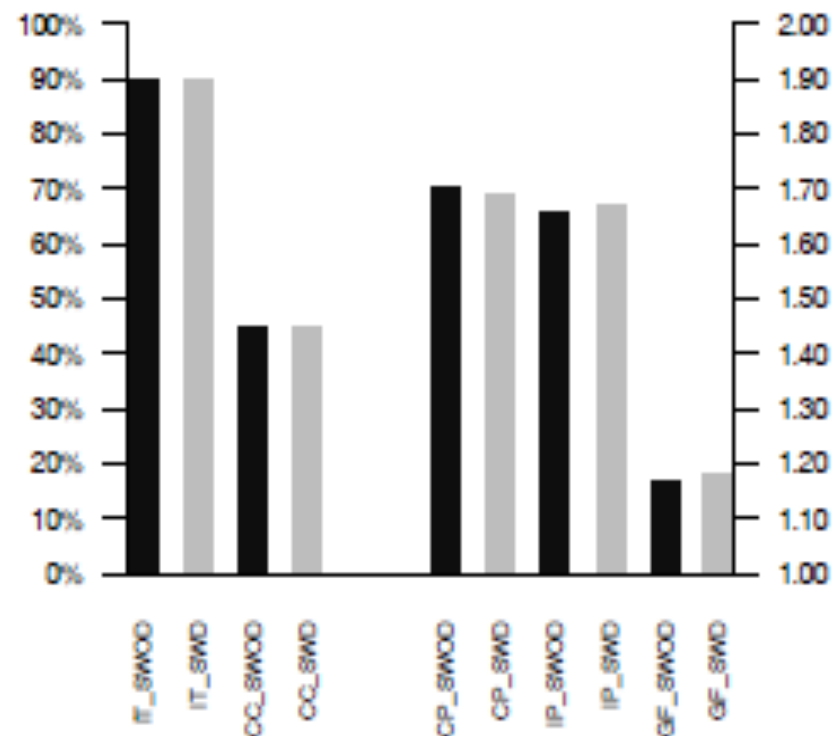
- We observed very similar instructional processes for students with and without disabilities learning mathematics in the same elementary or secondary classrooms in AZ and OR schools. Significant achievement gaps between these groups of students, however, existed on the four interim CBM assessments and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

# Comparison of OTL Indices

Oregon Elementary SWOD vs. SWD  
Comparison of MyLOGS Indices



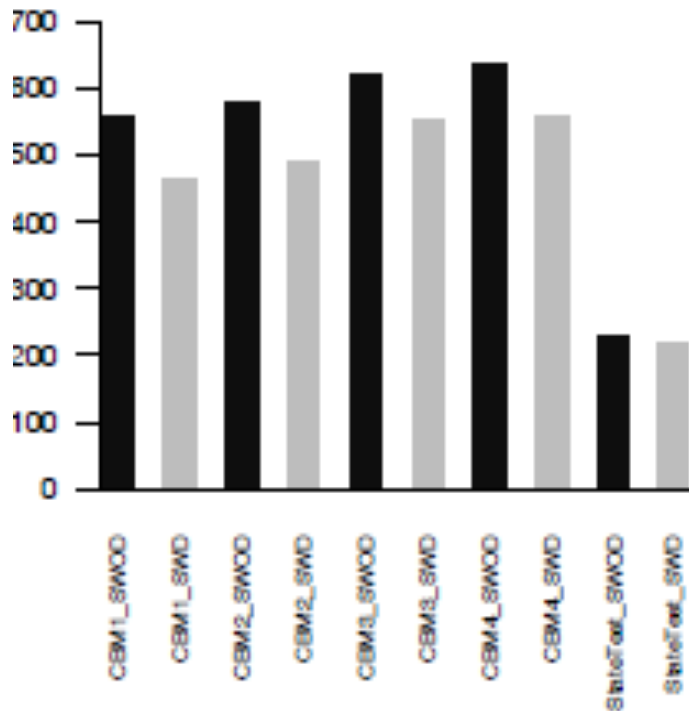
Oregon Secondary SWOD vs. SWD  
Comparison of MyLOGS Indices



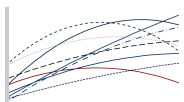
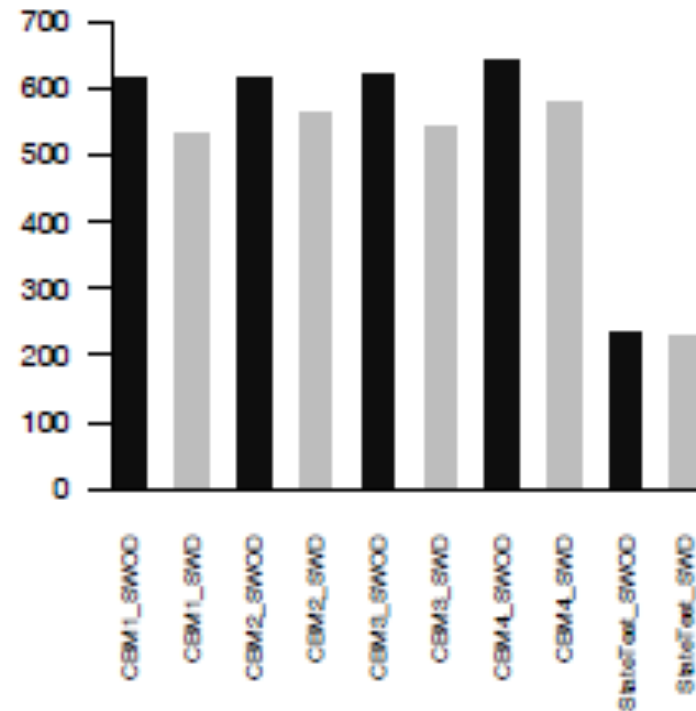


# Comparison of Interim & End-of-Year Test Results

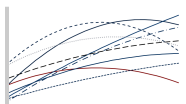
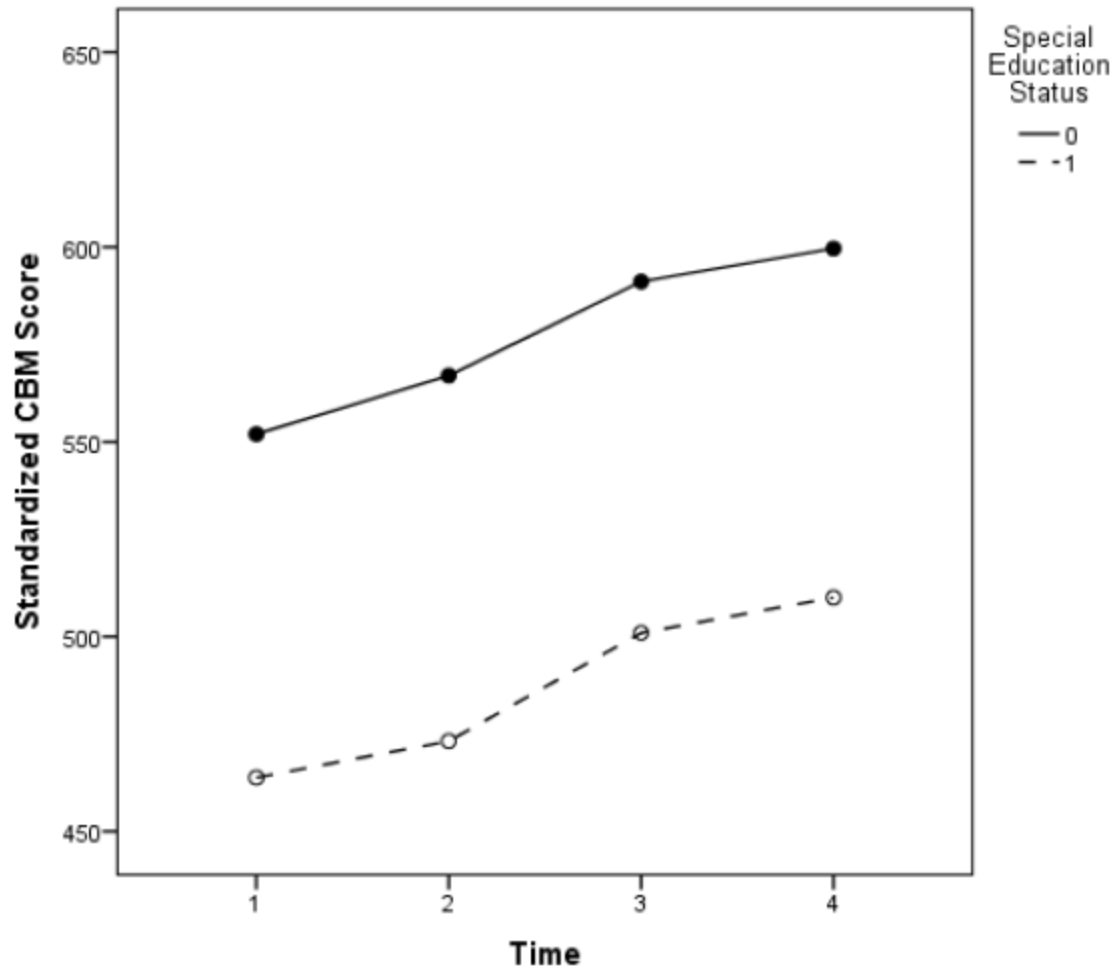
Oregon Elementary SWOD vs. SWD  
Comparison of EasyCBM & State Test



Oregon Secondary SWOD vs. SWD  
Comparison of EasyCBM & State Test



# Within Year Standardized Mathematics CBM Growth



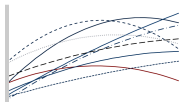
# Conclusion

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.

If the findings in Year 1 of this study are replicated in subsequent years and other studies, 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.

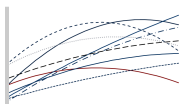
# Influence of Testing Procedures on Documenting Growth

Joseph F. T. Nese  
University of Oregon



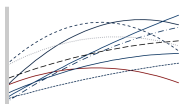
# Influence of Testing Procedures on Documenting Growth

- 1) Testing Students with Significant Cognitive Disabilities: Patterns of Participation in a General or Alternative State Assessment
- 2) Is Once Enough? The Effects of Multiple Administrations of a State Achievement Test



# Influence of Testing Procedures on Documenting Growth

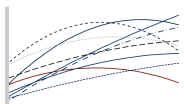
- Testing Students with Significant Cognitive Disabilities: Patterns of Participation in a General or Alternative State Assessment
  - Jessica L. Saven, Daniel Anderson, Joseph F. T. Nese, Dan Farley, Gerald Tindal
- Purpose
  - Explore how students with significant cognitive disabilities switch between the General and Alternate state assessments over time.



# Patterns of Participation:

## General (GA) or Alternative Assessment (AA)

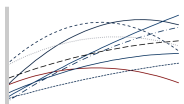
- States may develop AA based on alternate achievement standards.
- 9% of SWD, or 1% of all students.
- AA must meet technical adequacy requirements, and link with state academic content standards.
- Eligibility criteria and implementation vary; so some students “switch” test types between years.
- Accountability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



# Patterns of Participation:

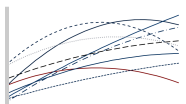
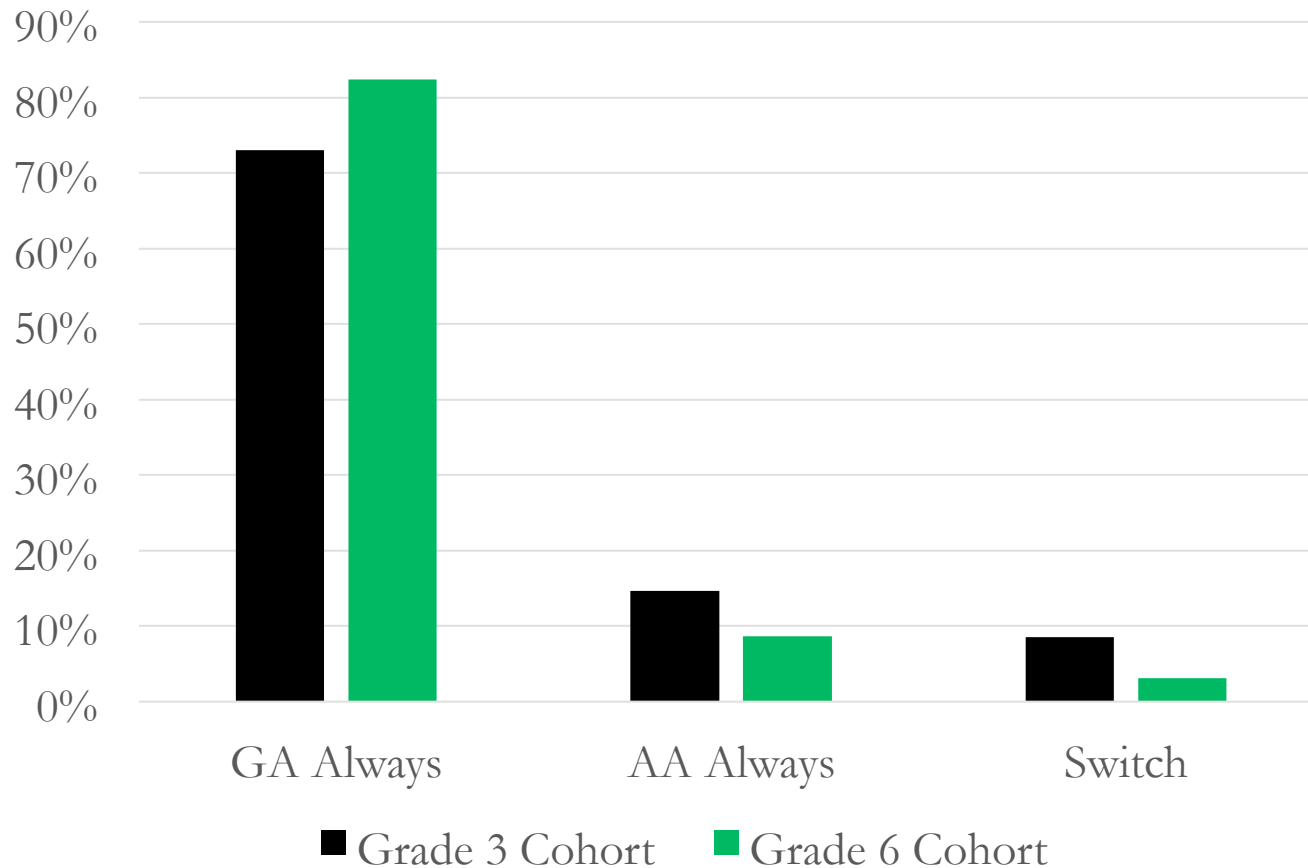
## General (GA) or Alternative Assessment (AA)

- 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?

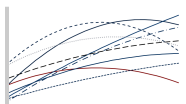
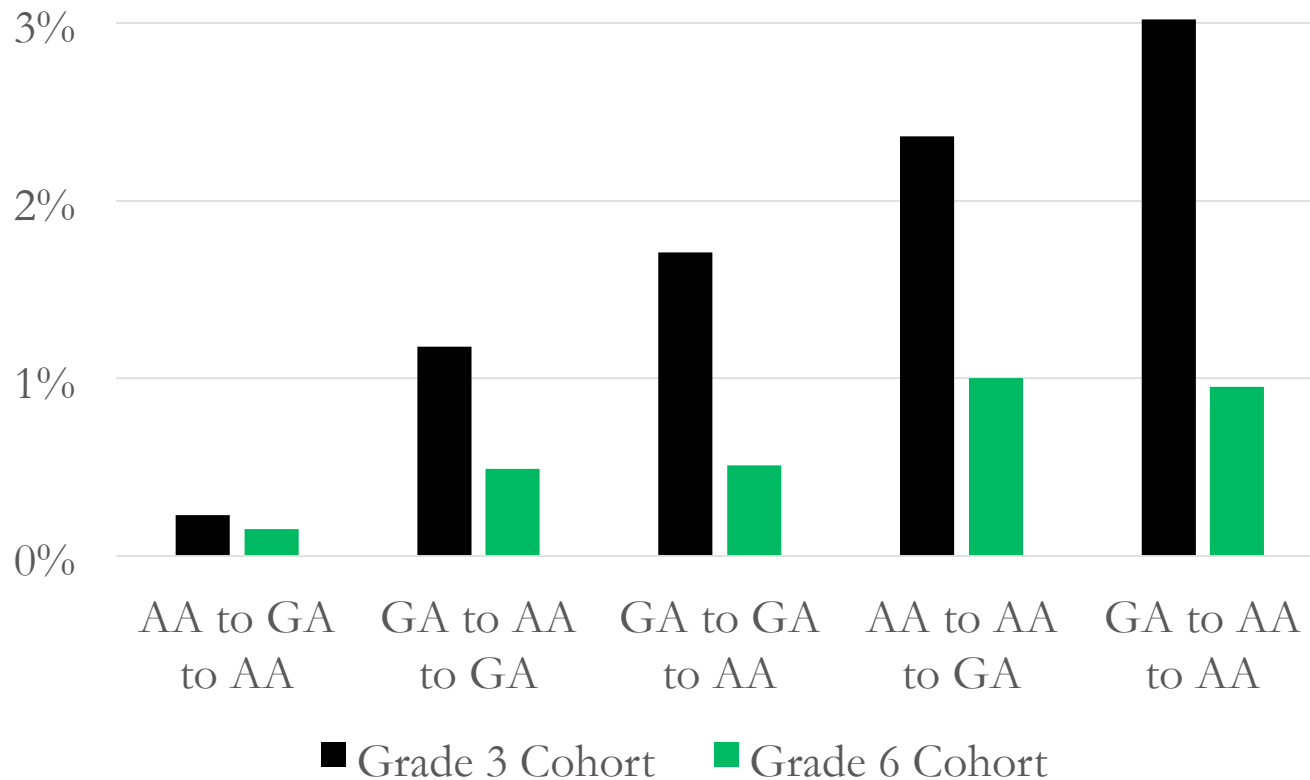




# Grade 3 ( $n = 3,048$ ) & Grade 6 ( $n = 3,911$ ) Cohort Test Patterns 2009/10 – 2011/12



# 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.

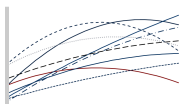
---

# Influence of Testing Procedures on Documenting Growth

- Is Once Enough? The Effects of Multiple Administrations of a State Achievement Test
  - Joseph F. T. Nese, Gerald Tindal, Joseph Stevens, Stephen N. Elliott

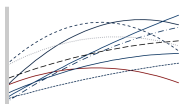
# The Effects of Multiple Administrations of a State Achievement Test

- No reference to multiple tests in ESEA.
- Multiple tests on NCLB summative tests (e.g., Delaware, Oregon).
- Multiple administrations may increase validity, decrease the false negative results, and increase false positive results.



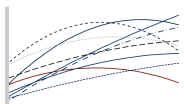
# The Effects of Multiple Administrations of a State Achievement Test

- **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).



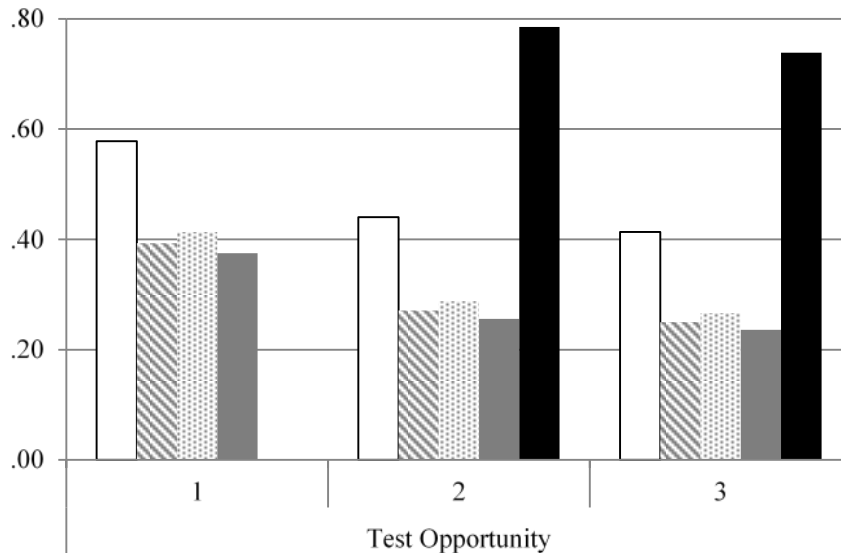
# The Effects of Multiple Administrations of a State Achievement Test

- 1) Are Multiple Administrations More Likely For Certain Students?
  - ❑ **non-LEP** vs. LEP students;
  - ❑ **GenEd** vs. SpEd students; and
  - ❑ **Bubble** vs. Below Bubble students.
    - These results were consistent across grades and subjects in direction, magnitude.
- 2) More Likely to Pass On Successive Attempts
  - ❑ **Females (reading); Males (math)**
  - ❑ **White** vs. Hispanic students.
    - No difference in passing rates between Whites vs. Other ethnic minority students
  - ❑ **Non-FRL** vs. than FRL students
  - ❑ **Non-LEP** vs. LEP students
  - ❑ **GenEd** vs. SpEd students
  - ❑ **Bubble** vs. Below-Bubble students

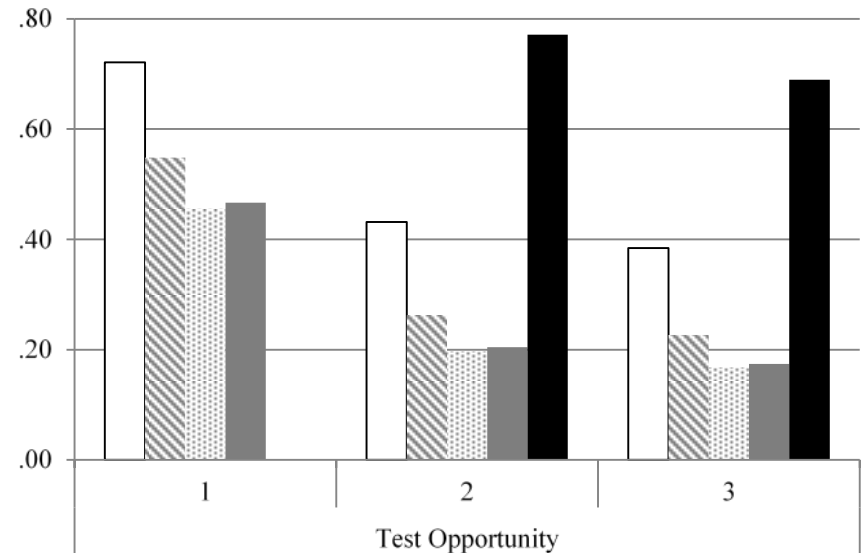


# Estimated probabilities of passing the Grade 3 math or reading test for specific student subgroups

a) Mathematics



b) Reading



□ Reference    ▨ FRL    ▩ LEP    ■ SpEd    ■ Bubble

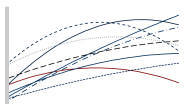
**Reference** = White, male, non-FRL (free/reduced priced lunch recipient), non-LEP (limited English proficiency status), GenEd (general education), BelowBubble (lower than one standard error of measurement below the proficiency cut score on the previous test).

**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).





# Implications

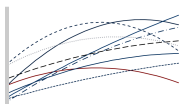
A multiple test policy can provide improved prospects for additional instruction, opportunity to learn, student development, and success for students and schools, but it remains uncertain whether large-scale state achievement are meaningful learning events. Multiple test opportunities may enable more fair and equitable proficiency reports and decrease the likelihood of false negatives.

# 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)

