Academic Growth and Gaps for Specific Student Subgroups on State Accountability Tests

> Presenters: Gerald Tindal Joseph J. Stevens Ann C. Schulte Discussant: Martha Thurlow

Presented at the National Conference on Student Assessment conference, San Diego, CA, June, 2018.

Friday, June 29, 2018: 11:00 AM - 12:00 PM, Hilton San Diego Bayfront, Aqua Salon C (Level 3)





Session Organization:

Paper 1: Performance of Students with Specific Learning Disabilities (SLD) on a State Accountability Test: Classification and Performance Changes. Gerald Tindal, PhD and Daniel Anderson, PhD, University of Oregon.

Paper 2: Academic Growth and Gaps for Specific Student Subgroups on a State Accountability Reading Test. Joe Stevens, PhD, University of Oregon.

Paper 3: College and Career Ready Standards and Students with Disabilities: Using Longitudinal Data to Inform Systemic Improvement Efforts. Ann Schulte, PhD, North Carolina State University.

Discussion: Martha Thurlow, National Center on Educational Outcomes, College of Education and Human Development, University of Minnesota

Questions and Open Discussion

Presentation available on NCAASE web site: <u>http://www.ncaase.com/</u>

This research was funded in part by a Cooperative Service Agreement from the Institute of Education Sciences (IES) establishing the National Center on Assessment and Accountability for Special Education - NCAASE (PR/Award Number R324C110004); the findings and conclusions expressed do not necessarily represent the views or opinions of the U.S. Department of Education.

Visit our website for other publications, research briefs, and onepage research summaries called "Did You Know"





Did You Know? Ο

READING

NCAASE

ACCOUNTABLITY FOR SPECIAL FOLICATION

ontact

ation

Reading comprehension growth across elementary and middle school grades is similar between students with disabilities and students without disabilities.¹

The good news: Students with disabilities who took a state general assessment in reading began with lower average Grade 3 reading comprehension scores, but across Grades 4-7, made average annual gains similar to students without disabilities.

Research Note No.1 - Dec. 2017

The challenge ahead: Students with disabilities need more intensive instruction in the earliest grades to reduce reading gaps already evident at Grade 3.



 Reading achievement differences (achievement gaps) seen in Grade 3 remained stable through Grade 7. . The use of longitudinal data and a vertically-scaled test means that both students and the test were constant across grades. This constancy is critical for a clear picture of achievement growth.

' For more information, see:

Hom

90.00 80.09

70.00

60.00

Schulte, A. C., Stevens, J. J., Eliott, S. N., Tindal, G., & Nese, J. F. T. (2016). Achievement gaps for students with disabilities: Stable, widening, or narrowing on a statewide reading comprehension test? Journal of Educational Psychology, 108, 925-9.42. doi: 10.1037/edu0000107 or visit our website: www.ncaase.com.

Acknowledgement: This research was funded through the Institute of Education Sciences (IES) (http://les.ed.gov) through a Cooperative Service Agreement establishing the National Center on Assessment and Accountability for Special Education - NCAASE (PR/Award Number R324C110004). The findings and conclusions expressed do not necessarily represent the views or opinions of the U.S. Department of Education.





5

Did You Know?

READING

ion

act

Research Note No. 4 - Feb. 2018

Despite great interest in academic achievement gaps, there is little consistency in how gaps are measured and reported. The size and even the presence of gaps may be misunderstood in many instances because different methods are used for measuring gaps.'

The good news: There are well established methods called effect size (ES) measures that express group differences using a common yardstick (standard deviation units). Use of ES measures can reduce subjectivity and foster better understanding of group differences.

The challenge ahead: Many educators, analysts, and policy-makers will need additional professional development to learn about ES and better ways to represent achievement gaps.



- For example, in the figure above, differences in percent proficient (PP) on the Arizona state reading/language test for fifthgrade students with disabilities (SWD) and English learner (EL) students (N = 61,713 total) seem about the same.
- ES is calculated as the mean difference on the reading/language test scale score divided by the standard deviation (SD; for additional detail see: ES_Details.pdf)
- Converted to ES, the gap on the left is 1.08, and the gap on the right is 1.27, almost .20 SD larger for EL students, revealing
 a noteworthy difference (equal to almost half [45%] of an academic year of growth) in the size of the achievement gap for
 SWD vs. EL students.²
- General rules of thumb for interpreting ES are: zero is equivalent to no difference; ES of about 0.20 is considered "small," about 0.50 is "medium," and 0.80 or more is "large."
- To see examples using ES measures to report achievement gaps, see DYK No. 1 and DYK No. 2.

¹ For more information, see:

Ο

Hom

00.00

Stevens, Anderson, Nese, & Tindal (2016). Using Effect Size Measures to Estimate and Report Achievement Gaps, paper presented at NCME; available at our website: <u>www.ncaase.com</u> ¹Note that AZEL testing policies differ from many other states and so this result may not generalize. Acknowledgement: This research was funded through the Institute of Education Sciences (IES) (http://esed.gov) through a Cooperative Service Agreement establishing the National Center on Assessment and Accountability for Special Education – NCAASE (PR/Award Number R324Cm0004). The findings and conclusions expressed do not necessarily represent the views or opinions of the U.S. Department of Education.



one-Did

Performance of Students with Specific Learning Disabilities (SLD) on a State Accountability Test: Classification and Performance Changes

Gerald Tindal, PhD and Daniel Anderson, PhD

National Center on Assessment and Accountability for Special Education

Behavioral Research and Teaching – College of Education – UO

Never • Sometimes • Always SLD and Test Performance

| Cohort | SLD | Grade 6 | | Grade 7 | | Grade 8 | |
|--------|-----------|---------|-------|---------|-------|---------|-------|
| Conort | | Μ | SD | Μ | SD | Μ | SD |
| 8 | Always | 215.73 | 10.83 | 224.36 | 10.61 | 226.99 | 11.25 |
| 8 | Sometimes | 216.99 | 10.31 | 225.38 | 9.91 | 228.11 | 10.68 |
| 8 | Never | 227.56 | 10.67 | 234.87 | 10.26 | 237.77 | 10.93 |
| 9 | Always | 216.35 | 11.19 | 224.71 | 9.71 | 228.34 | 10.75 |
| 9 | Sometimes | 218.04 | 10.4 | 225.89 | 9.45 | 229.38 | 10.21 |
| 9 | Never | 228.52 | 10.81 | 234.82 | 9.5 | 239.06 | 10.29 |
| 10 | Always | 217.31 | 9.91 | 226.7 | 9.01 | 227.87 | 10.8 |
| 10 | Sometimes | 218.8 | 10.69 | 228.22 | 9.56 | 229.74 | 10.98 |
| 10 | Never | 228.38 | 10.04 | 235.99 | 8.98 | 239.06 | 10.42 |



Never • Sometimes • Always SLD and Test Proficiency

| Grade 6 | Grade 7 | Grade 8 | Always | Sometimes | Never | Total |
|---------|---------|---------|-----------|-------------|-------------|-------------|
| Non-SLD | Non-SLD | SLD | 42 (1%) | 125 (2%) | 156 (3%) | 323 (6%) |
| Non-SLD | SLD | Non-SLD | 9 (0%) | 14 (0%) | 15 (0%) | 38 (1%) |
| Non-SLD | SLD | SLD | 58 (1%) | 170 (3%) | 191 (3%) | 419 (7%) |
| SLD | Non-SLD | Non-SLD | 143 (3%) | 133 (2%) | 74 (1%) | 350 (6%) |
| SLD | Non-SLD | SLD | 10 (0%) | 20 (0%) | 36 (1%) | 66 (1%) |
| SLD | SLD | Non-SLD | 212 (4%) | 151 (3%) | 95 (2%) | 458 (8%) |
| SLD | SLD | SLD | 860 (15%) | 1,557 (28%) | 1,542 (27%) | 3,959 (71%) |



Student Characteristic for Changing SLD Status

| Characteristic | Cohort | Always | Sometimes | Never |
|----------------|--------|---------------|--------------|---------------|
| | 08 | 1,343 (0.05) | 573 (0.02) | 23,570 (0.92) |
| SLD | 09 | 1,306 (0.05) | 563 (0.02) | 23,299 (0.93) |
| | 10 | 1,310 (0.05) | 518 (0.02) | 22,836 (0.93) |
| | 08 | 9,669 (0.38) | 3,643 (0.14) | 12,174 (0.48) |
| FRPL | 09 | 10,276 (0.41) | 3,207 (0.13) | 11,685 (0.46) |
| | 10 | 10,718 (0.43) | 2,842 (0.12) | 11,104 (0.45) |
| | 08 | 1,650 (0.06) | 952 (0.04) | 22,884 (0.90) |
| LEP | 09 | 1,373 (0.05) | 1,205 (0.05) | 22,590 (0.90) |
| | 10 | 929 (0.04) | 1,441 (0.06) | 22,294 (0.90) |

10

Differences in Intercept for Changing SLD Status

| Parameter | Estimate | Std. Error | 95% LCI | 95% UCI |
|----------------------|----------|------------|---------|---------|
| Initial Achievement | 231.16 | 0.07 | 231.02 | 231.31 |
| Cohort 09 | 0.81 | 0.08 | 0.66 | 0.97 |
| Cohort 10 | 1.01 | 0.08 | 0.86 | 1.17 |
| American Indian | -2.32 | 0.25 | -2.82 | -1.82 |
| Asian | 4.73 | 0.15 | 4.43 | 5.03 |
| Black | -3.71 | 0.20 | -4.11 | -3.31 |
| Hispanic | -1.19 | 0.10 | -1.39 | -1.00 |
| Multiethnic | 0.20 | 0.19 | -0.16 | 0.56 |
| Decline | -1.31 | 0.41 | -2.12 | -0.50 |
| Male | -1.22 | 0.07 | -1.35 | -1.09 |
| FRL | -4.94 | 0.07 | -5.08 | -4.80 |
| LEP | -7.13 | 0.14 | -7.40 | -6.85 |
| SLD-Sometimes | -8.36 | 0.23 | -8.80 | -7.92 |
| SLD-Always | -9.40 | 0.15 | -9.69 | -9.11 |



Differences in Slope for Changing SLD Status

| Parameter | Estimate | Std. Error | 95% LCI | 95% UCI |
|----------------------|----------|------------|---------|---------|
| Growth | 9.08 | 0.05 | 8.99 | 9.17 |
| Cohort 09 | 0.12 | 0.05 | 0.02 | 0.22 |
| Cohort 10 | 0.39 | 0.05 | 0.29 | 0.49 |
| American Indian | -0.31 | 0.16 | -0.63 | 0.01 |
| Asian | 1.07 | 0.10 | 0.88 | 1.26 |
| Black | -0.24 | 0.13 | -0.50 | 0.01 |
| Hispanic | 0.29 | 0.06 | 0.16 | 0.41 |
| Multiethnic | 0.13 | 0.12 | -0.10 | 0.36 |
| Decline | 0.34 | 0.26 | -0.17 | 0.85 |
| Male | 0.30 | 0.04 | 0.22 | 0.38 |
| FRL | 0.02 | 0.05 | -0.07 | 0.11 |
| LEP | 1.39 | 0.09 | 1.21 | 1.56 |
| SLD-Sometimes | 0.70 | 0.14 | 0.42 | 0.98 |
| SLD-Always | 0.80 | 0.09 | 0.62 | 0.99 |



Proficiency Change in Performance Level Classifications from Grade 7 to 8



Differences in Proficiency by SLD Status Change AcrossGradesGrade 6Grade 7Grade 8





Model-based Growth Trajectories Compared with Raw Means by Cohort



Academic Growth and Gaps for Specific Student Subgroups on a State Accountability Reading Test

> Joseph J. Stevens University of Oregon

© Stevens, 2018





Presentation Purpose

- Discuss issues in estimating and understanding achievement gaps.
- Discuss importance of directly testing interaction effects (see Stevens & Schulte, 2016), i.e., precise disaggregation of groups.
- Describe results of an ongoing study of reading achievement growth for students with learning disabilities (LD) and English Learners (EL) on Arizona state reading test.
- Because of our short time, we only present some highlights:
 - Reading achievement over time for LD vs. not-LD and for EL vs. not-EL (what you get with usual regression models).
 - Reading achievement over time for the LD-EL interaction effect.
 - Differences in LD-EL subgroup performance (i.e., achievement gaps) expressed as effect sizes (ES).

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

Interactions of Student Characteristics

- Many studies do not directly test the interaction of SWD status and other student characteristics of interest (e.g., LD status and sex of student).
- Usually these variables are examined as one of several predictors in a regression model (i.e., a partial regression effect).
- Even though two predictors (LD-sex) are in the same regression model, they do not estimate an actual interaction effect (e.g., combined characteristics like LD-male versus LD-female).
- This can be very misleading and may result in incorrect interpretations as well as incomplete understanding of group differences. See:
 - Stevens, J. J., & Schulte, A. C. (2017). The interaction of learning disability status and student demographic characteristics on mathematics growth. *Journal of Learning Disabilities*. DOI: 10.1177/0022219415618496



Interactions of Disability Status and Student Characteristics

- Our purpose in the current study, therefore, was to test true interactions of SLD status with several other student characteristics.
- We were also interested in the size of the achievement gap in these comparisons and in whether the achievement gap was increasing or decreasing over grades.
- In this presentation we present selected results showing SLD interactions with EL status, although we have also analyzed interactions of SLD status with economic disadvantage, and with Hispanic vs. White race/ethnicity.

National Center on Assessment and

Accountability for Special Education



www.ncaase.com/

Analytic Methods

- We used hierarchical linear models (HLM) to test the interaction effects over Grades 3 to 6.
- Briefly the form of the two-level (grades and students) random intercepts and random slopes HLM model was:

Level-1 Model:
$$Y_{ti} = \pi_{0i} + \pi_{1i}^* (\text{Time}_{ti}) + \pi_{2i}^* (\text{Time}_{ti}^2) + e_{ti}$$
 (1)
Level-2 Model: $\pi_{0i} = \beta_{00} + \beta_{01}^* (\text{Predictor}_i) + r_{0i}$ (2)

$$\pi_{0i} = \beta_{00} + \beta_{01}^* (\operatorname{Predictor}_i) + r_{0i}$$
(2)

$$\pi_{1i} = \beta_{10} + \beta_{11}^{*} (\text{Predictor}_{i}) + r_{1i}$$
(3)

$$\pi_{2i} = \beta_{20} + \beta_{21}^{*} (\text{Predictor}_{i}) + r_{2i}$$
(4)



Method

- Student scores on the reading subtest of the Arizona Instrument to Measure Standards (AIMS) used for analyses.
- Sample details:
 - Sample size, N = 82,675 in Grade 3
 - Race/ethnicity composition in percent was 2.8 Asian, 5.6 Black, 43.3 Hispanic, 5.2 American Indian, 43.1 White.
 - 48.8% of the students were female; 12.9% were SWD; 5.6% LD; 19% were EL; 51.2% were economically disadvantaged.
- We examined attrition of the sample over grades; compared to Grade 3, 94% were present in Grade 4, 91% in Grade 5, and 87% in Grade 6.

Further details on sample, methods and procedures available on request from the author.



Results

- In all the analyses we conducted, the interaction effects were statistically significant (i.e., LD x EL, LD x ECD, EL x ECD, LD x Hispanic, EL x Hispanic).
- For brevity, we only present graphical displays of the key results here for illustration.
- We then provide summaries of the size of achievement gaps expressed as effect sizes (ES).





Single Predictor Comparison of LD Status



Single Predictor Comparison of EL status



Interaction of LD Status with EL Status



Results

- Note there is generally parallel growth over grades.
- Some closing of the gap for NotLD-EL students.
- <u>Good news</u>: all students groups are progressing in a similar way over grades.
- <u>Bad news</u>: on the whole the gap is not closing.







www.ncaase.com/

http:/

Interaction Subgroups (solid lines) vs. Partial Regression Effects **LD** Subgroups (dashed lines)



Interaction Subgroups (solid lines) vs. Partial Regression Effects **EL** Subgroups (dashed lines)



Results

- There are many follow-up analyses of interest, but here we only present some examples of achievement gaps expressed as Effect Sizes (ES).
- We have produced a series of brief, one-page summaries of our NCAASE research results called "Did You Know"; several describe the use of ES information to report achievement gaps (see DYK's 1, 2, 4, 13).
- There is also a research brief on ES:
 <u>http://www.ncaase.com/publications/view?id=138</u>

National Center on Assessment and

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

Accountability for Special Education



www.ncaase.com/

Results: Achievement Gap ES

- Using the findings illustrated graphically above, we now examine the size of achievement gaps using estimates of ES.
- Cohen's rules of thumb for interpreting ES are:
 - □ zero is equivalent to no difference between groups,
 - □ about 0.20 is considered a "small" effect,
 - □ about 0.50 is a "medium" effect,
 - 0.80 or higher is a "large" effect.



Results: Achievement Gap ES

- The ES for the partial regression comparisons were:
 - For students who are LD versus not LD (slide 9), -1.30 in Grade 4 and
 -1.29 in Grade 6; no appreciable change in the ES achievement gap.
 - For students who are EL versus not EL (slide 10), -0.97 in Grade 4 and -0.83 in Grade 6, a narrowing of the achievement gap.
- For the <u>interaction effects</u> of LD and EL, students who were not LD and not EL were the comparison group and achievement gaps for the remaining subgroups were:

| | Grade 3 | Grade 6 |
|----------|---------|---------|
| LD-NotEL | -1.32 | -1.30 |
| NotLD-EL | -1.01 | -0.81 |
| LD-EL | -1.75 | -1.73 |

National Center on Assessment and

Accountability for Special Education

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



www.ncaase.com

http:/

Conclusion

- Importance of investigating achievement gaps more carefully than usual methods (i.e., description of differences in percent proficient):
 - Longitudinal not cross-sectional, so <u>change</u> in gaps can be evaluated.
 - Use objective measures of achievement gap size (e.g., ES) rather than "eyeballing" differences in percent proficient.
 - Test true interaction effects to correctly evaluate combinations of student characteristics and to further disaggregate results.
- Our results demonstrate that using these methods, previously unexamined student subgroups may emerge with substantially larger achievement gaps.





Conclusion

- These results exactly parallel Stevens & Schulte (2017).
 - Replication with a different state sample of students and a different state testing system (AZ vs. NC).
 - Extension of previous analyses to an important student subgroup (EL students).
- Through the use of interaction effects, important student subgroups are more clearly identified and evaluated.
- This can increase attention to the need for intervention for certain student subgroups who may be at greater risk academically.





College- and Career-Ready Standards and Students with Disabilities: Using Longitudinal Data to Inform Systemic Improvement Efforts

> Ann Schulte Research Professor Arizona State University Professor Emerita North Carolina State University





The Changing Context of Accountability for Students with Disabilities (SWD)

- **ESSA-More flexibility and broader focus than proficiency**
- IDEA-Results Driven Accountability (RDA)
 - State Identified Measurable Result (SIMRs)
 - □ SIMR multi-year plans include baseline and targets
- Improved state student data systems and annual data collection as a result of NCLB
- Introduction of College- and Career-Ready Standards raises bar for all students, but very challenging for SWD



Leveraging Data and Flexibility for Improved SWD Outcomes

- Changing context and policies offer opportunities for long-term goal setting and multi-year change strategies
- Longitudinal data can play an important role in developing progress indicators, monitoring the impact of change efforts





Cross Sectional vs. Longitudinal Samples

- Cross sectional reporting incorporated into original NCLB achievement reporting requirements
- Problematic for monitoring groups where membership may change annually
- When membership is related to outcome variable (i.e., achievement), potential bias introduced
- Multiple studies have documented differences in outcomes with longitudinal vs. cross sectional samples (e.g., Schulte & Stevens, 2014; Tindal & Anderson, this session; Ysseldyke & Bielinski, 2002)

Issues Using Cross Sectional Samples for Examining SWD Outcomes

- Students enter special education at different times and remain for different time spans, as illustrated by point (11.8%, Grade 3) vs. period prevalence (16.3%) rates for SWD (Grade 3-7; Schulte & Stevens, 2015)
- Students exiting special education have higher achievement than students who enter or remain (Thurlow et al., 2016)
- Interplay of interventions across time is lost
 - □ 50% of SWD retained before identified (Beebe et al., 2004).
 - Special education lowers probability of retention after placement (Moser et al., 2012)
 - Costly interventions: \$8,000 sped, \$12,000 retention (McCoy et al., 2017)



Longitudinal Studies

- Although ESSA requires SWD subgroup to include only students in special education for reporting annual achievement outcomes, longitudinal approaches are an option for systemic improvement progress indicators
- Given concerns about instability in annual SWD subgroup, the National Center for Systemic Improvement (NCSI, 2016) issued a brief on using longitudinal designs to capture students targeted for intervention in SIMRs created as part of RDA efforts
- NCSI suggested defining the target population broadly to include students likely to be served in special education during baseline and monitoring period
- Remainder of presentation illustrates issues with cross sectional and longitudinal approaches, and an example of selecting a longitudinal sample likely to be a target of systemic improvement efforts



The Long & Winding Road...

- Example takes a descriptive "pathways" approach--following an entire state cohort for 6 years, starting at Grade 3 and looking at special education participation and outcomes at endpoint (Grade 8 for most students)
- Variables of interest
 - Stability of specific exceptionality and special education status
 - **Retention**
 - Tracked actual grade retentions Grade 3 and beyond
 - Due to lack of early elementary data, used used "Above Modal Age for Grade" (age 9 before entering Grade 3) as a proxy for late school entry or retention
 - Reading and mathematics performance relative to College- and Career-Ready Standards



The Long & Winding Road...

Four touchpoints

- □ 2007/08 (all Grade 3)
- □ 2009/10 (modal Grade 5)
- □ 2011/12 (modal Grade 7)
- □ 2012/13 (modal Grade 8)
- Endpoint coincided with the first year of implementation College- and Career-Ready Standards in NC and aligned state assessment.
- Are there patterns to special entrances and exits that can inform school interventions or guide selection of target group or progress indicators?





Sample Description

- Initial sample was 112,955 students—all students who completed 3rd grade for first time in 2007/08, as indicated in the state student demographic file for that school year
- Students followed forward to 2013. Analyses were completed only with children who were in NC at all of the four touchpoint years. Final sample size was 90,259 (80.15% of initial sample)
- General, alternate (1% & 2%) assessments were in use at all touchpoints



Exceptionality by Touchpoint



Special Education Pathways

| YR 1 | Above | 12,097 (GR = 0%) Above Modal Age for Grade = 27.0% | | | | 78,432 (GR=0%) Above Modal Age for Grade = 6.8% | | | |
|---------|-----------|--|------------|------------|------------|---|------------|------------|--|
| | | | | | | | | | |
| 3 | 9,: | 9,219 | | 2,867 | | 75,565 | | | |
| | (GR= | (GR=4.9%) (GR=3.4%) | | (GR=11.9%) | | (GR=2.9%) | | | |
| | | $\overline{}$ | | | \frown | | \land | | |
| 5 | 8,036 | 1,183 | 131 | 2,747 | 2,378 | 489 | 1,210 | 74,355 | |
| | (GR=6.7%) | (GR=6.9%) | (GR=19.1%) | (GR=4.62%) | (GR=13.6%) | (GR=11.9%) | (GR=13.3%) | (GR=3.8%) | |
| | Ļ | Ļ | Ļ | Ļ | Ļ | Ļ | Ļ | Ļ | |
| 6 | 7,637 | 1106 | >121 | 2719 | 2212 | 468 | 1,144 | 74,046 | |
| | (GR=7.4%) | (GR=7.8%) | (GR=19.2%) | (GR=5.1%) | (GR=14.2%) | (GR=12.0%) | (GR=19.0%) | (GR=4.1%) | |
| | 399 | 77 | <10 | 28 | 166 | 21 | 66 | 309 | |
| | (GR=5.3%) | (GR=10.4%) | (GR= 0.0%) | (GR=17.9%) | (GR=10.8%) | (GR=4.8%) | (GR=10.7%) | (GR=22.0%) | |

Blue = Special Ed, Gold = General Ed

GR = Grade Retention (Cumulative %)

Special Education at Year 6



Blue = Special Ed, Gold = General Ed

GR = Grade Retention (Cumulative %)

Percent of Students Meeting Career- and College-Ready Proficiency Standards for Grade 8



Note: Percent proficient figures will underestimate figures reported by state because denominators include all students in the cohort and specified category, including retained students and those participating in alternate assessments (1% or 2%).

Stability of Initial Exceptionality



Year 1 Exceptionality by Year 6 Status



Percent of Students Meeting Career- and College-Ready Proficiency Standards for Grade 8



Note: Percent proficient figures will underestimate figures reported by state because denominators include all students in the cohort and specified category, including retained students and those participating in alternate assessments (1% or 2%).

Percent of Students Meeting Career- and College-Ready Proficiency Standards for Grade 8



Note: Percent proficient figures will underestimate figures reported by state because denominators include all students in the cohort and specified category, including retained students and those participating in alternate assessments (1% or 2%).

Creating a High Risk Longitudinal Sample for Monitoring to Grades 3-8

- Goal: Select group of students with high likelihood of participating in special education services across time span of interest to address issue of annual entrances and exits
 - □ Based on information available at baseline (in this case, Grade 3)
 - Include substantial portion of students actually SWD at endpoint (sensitivity)
 - Exclude most students who were not served in special education (specificity)
- Proposed Group: SWD at Year 1 (excluding Speech-Language Impaired), plus lowest 15% of students in reading at Year 1



High Risk Group Compared to SWD at Yr 6

| | Not SWD at Yr 6 | SWD at Yr 6 | Total |
|-------------------|-----------------------|----------------|--------|
| Not at Risk | 69,651 | 2,052 | 71,703 |
| At Risk | 9,325 | 9,501 | 18,826 |
| Total | 78,976 | 11,553 | 90,529 |

- Includes 82.2% of SWD at Year 6
- Relative risk for SWD at YR 6 is 17 times higher for At Risk than Not at Risk group
- Dovetails well with lowest 15% often targeted for Tier 2 in Multitiered Systems of Support
- Percent proficient rates similar to those for SWD at Yr 6 (5-6% proficient)



A Word About Retentions...



Blue = Special Ed, Gold = General Ed

GR = Grade Retention (Cumulative %)

Take-Aways From the Pathways Approach

- Cross sectional and longitudinal approaches produce different results because SWD group membership changes over time
- The group most likely to exit special education, Speech-Language Impairment, is also the group most likely to meet grade level proficiency at Grade 8, which partially accounted for differences in cross sectional and longitudinal outcomes at endpoint
- Longitudinal results are helpful in discerning patterns that may be missed by examining SWD group based on annual membership
- Defining target population more broadly and using a longitudinal evaluation design may better capture students exiting special education or who benefited from preventive interventions
- Suggested high risk group here is an illustration of how longitudinal results can be used to select a stable target group. Selected group would depend on targeted outcome, intervention focus, and age group

Thank You!

Contact information: Gerald Tindal, <u>geraldt@uoregon.edu</u> Joe Stevens, (541) 346-2445, <u>stevensj@uoregon.edu</u> Ann Schulte, <u>schulte@ncsu.edu</u> Martha Thurlow, <u>THURL001@umn.edu</u>

Presentation available on NCAASE web site: <u>http://www.ncaase.com/</u>



