Members

Rep. Robert Behning, Co-Chairperson
Rep. Rhonda Rhoads
Rep. Timothy Brown
Rep. Edward Clere
Rep. David Frizzell
Rep. Kathleen Heuer
Rep. Cindy Noe
Rep. Jeffrey Thompson
Rep. Greg Porter
Rep. David Cheatham
Rep. Clyde Kersey
Rep. Vernon Smith
Rep. Shelli Vandenburgh
Sen. Dennis Kruse, Co-Chairperson
Sen. Carlin Yoder
Sen. James Banks
Sen. James Buck
Sen. Luke Kenley
Sen. Jean Leising
Sen. Scott Schneider
Sen. Earline Rogers
Sen. Frank Mrvan
Sen. Timothy Skinner

LSA Staff:

Allen Morford, Attorney for the Select Commission
Erma Reinumagi, Attorney for the Select Commission
Chuck Mayfield, Fiscal Analyst for the Select Commission
David Lusan, Fiscal Analyst for the Select Commission

Authority: P.L. 160-2012

SELECT COMMISSION ON EDUCATION

Legislative Services Agency
200 West Washington Street, Suite 301
Indianapolis, Indiana 46204-2789
Tel: (317) 233-0696 Fax: (317) 232-2554

MEETING MINUTES[1]

Meeting Date:	April 24, 2012
Meeting Time:	2:00 P.M.
Meeting Place:	State House, 200 W. Washington St., House Chambers
Meeting City:	Indianapolis, Indiana
Meeting Number:	1

Members Present: Rep. Robert Behning, Co-Chairperson; Rep. Rhonda Rhoads; Rep. Edward Clere; Rep. Kathleen Heuer; Rep. Cindy Noe; Rep. Jeffrey Thompson; Rep. Greg Porter; Rep. David Cheatham; Rep. Vernon Smith; Rep. Shelli Vandenburgh; Sen. Dennis Kruse, Co-Chairperson; Sen. Carlin Yoder; Sen. James Banks; Sen. James Buck; Sen. Luke Kenley; Sen. Jean Leising; Sen. Earline Rogers; Sen. Timothy Skinner.

Members Absent: Rep. David Frizzell; Rep. Timothy Brown; Rep. Clyde Kersey; Sen. Scott Schneider; Sen. Frank Mrvan.

Co-chairperson Kruse called the meeting to order at 2:05 p.m., and stated that the

---

[1] These minutes, exhibits, and other materials referenced in the minutes can be viewed electronically at http://www.in.gov/legislative Hard copies can be obtained in the Legislative Information Center in Room 230 of the State House in Indianapolis, Indiana. Requests for hard copies may be mailed to the Legislative Information Center, Legislative Services Agency, West Washington Street, Indianapolis, IN 46204-2789. A fee of $0.15 per page and mailing costs will be charged for hard copies.

Commission will have four additional meetings, with the next meeting to be held on May 21, 2012, starting at 1:00 p.m. Co-chairperson Behning explained that today's meeting would be an opportunity for the Department of Education (DoE) to explain the matrices used to establish the A-F school designations and discuss the growth model of measuring student achievement He called upon the members of the Commission to introduce themselves.

State Superintendent of Public Instruction Dr. Tony Bennett thanked the Commission for the opportunity to explain DoE's processes in implementing recent statutory changes. He stated that DoE's work has been carried out with transparency, engaging in dialogue with all interested parties in the rule-making process. He feels that Indiana's growth model for student improvement demonstrates the current best practices in education. DoE distributed material that contained a policy overview, myth versus fact and FAQ documents, copies of DoE's presentations, sample A - F report cards, and supporting research on student growth and school accountability (Exhibit A).

Dr. Wes Bruce, Chief Assessment Officer, DoE, discussed the history and implementation of the growth model for student measurement (a copy of Dr. Bruce's presentation is contained within Exhibit A). Under a growth model, a student's academic growth is measured on an annual basis to ascertain whether the student has achieved a year's worth of academic growth during a school year. Dr. Bruce discussed norm and criterion referenced testing and models and explained the timeline during which Indiana's growth model developed, beginning in 2005. He explained the key features of Indiana's growth model, such as adding a dimension by which to understand student achievement and allowing easier determination of the progress a school makes to address achievement gaps within the school.

Dr. Damian Betebenner, Senior Associate at the National Center for the Improvement of Educational Assessment, Dover, NH, developed the student growth percentiles and percentile growth trajectories methodology, which is both a norm and criterion referenced model. He works with states, including Indiana, to develop student growth models for assessment (a copy of Dr. Betebenner's presentation is contained within Exhibit A). Dr. Betebenner began working with Indiana in 2008. He discussed the development and use of growth models, including its analysis and reporting elements.

Dale Chu, Assistant Superintendent, DoE, presented information concerning the new accountability metrics adopted as a rule by the State Board of Education (a copy of Mr. Chu and Mr. Gubera's presentation is contained within Exhibit A). Mr. Chu explained that the goals of the accountability system were to:

- ensure that every student counts;
- close the achievement gap;
- provide transparent results;
- merge state and federal accountability;
- use new tools to provide a stronger system; and
- ensure post-secondary success for all students.

He explained the timeline under which the rule was developed from the fall of 2010 to the approval of the final rule by the State Board in February, 2012. Under the rule, there are different standards by which schools are measured:

- for elementary and middle schools, performance and growth; and
- for high schools, performance and improvement.

The rule adds designations for schools of A - F, incorporates the Indiana growth model, adds a new subgroup of the bottom 25% of students, and includes college and career readiness as a criterion.

Jon Gubera, Chief Accountability Officer, DoE, discussed the importance of high growth in academic progress, particularly for the lowest performing 25% of students in elementary and middle school. He also presented information concerning the need for postsecondary education for employment, which is reflected in the new measures of college and career readiness included in the accountability provisions. Mr. Gubera presented more detailed information on how the A - F accountability system will work, including:

- the system of base scores, bonuses, and deductions for elementary and middle schools; and
- the weighted system of passage rates on English 10 and Algebra I, the graduation rate, and college and career readiness for high schools.

Mr. Chu, Mr. Gubera, Dr. Betebenner, and Dr. Bruce responded to questions from Commission members concerning:

- the availability of testing statistics;
- types and scoring of tests given;
- the use of rigid cut scores;
- the use of peer groups in setting improvement criteria;
- how long a student must be enrolled in a school for the student's score to be counted for a school;
- how best practices can be identified and shared; and
- whether the accountability models will be expanded beyond math and English/language arts.

In addition, Commission members commented upon:

- the use of, time spent on, and cost of assessments;
- whether the assessments should be used to evaluate schools and teachers; and
- whether the testing helps children learn.

Co-chairperson Behning adjourned the meeting at 5:30 p.m.

Indiana Department of Education

Indiana Department of Education
2012 Select Committee on Education Issues
A through F School Accountability Metrics

SELE

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

April 24, 2012

Distinguished Members of the Select Commission on Education Issues:

Thank you for this opportunity to share the Department's work to update and improve the state's approach to raising student achievement. The State Board of Education has guided a deliberative process that has resulted in a more comprehensive and student-centered A through F school accountability system. I am incredibly proud of the hard work and careful consideration that has gone into this important task over the last few years.

We have gathered some key information to be sure you have the research and details necessary to accurately assess Indiana's progress toward implementing a school accountability system that is fair, thorough, and transparent. This system is a direct outgrowth of ongoing discussions the Department has had with a broad based group of engaged stakeholders; it reflects my firm commitment to ensuring Indiana leads the nation in preparing our students to compete on the global stage.

Within this folder, you will find the following information:

- Policy overview
- Myth v. Fact and FAQ document
- PowerPoint presentations
    - Indiana Growth Model
    - Indiana's A-F School Accountability Metrics
- Sample A-F Report Cards
- Supporting research on growth and school accountability

We realize you may have additional questions that are not covered by this material and may not be addressed in today's meeting. Please feel free to reach out to my Director of Legislative and Intergovernmental Affairs, Ashley Gibson, at agibson@doe.in.gov or 317-232-6618.

Finally, I hope you are as proud as I am of the tremendous work that has been done to implement the trailblazing legislation enacted by you and your colleagues. By every measure, Hoosier students are improving. As a result, other states are now beginning to emulate Indiana and the steps we've taken together to dramatically improve the educational opportunities for all students.

Sincerely,

Tony Bennett
Superintendent of Public Instruction

Indiana Department of Education

SUPPORTING STUDENT SUCCESS

Indiana's Growth Model History and Implementation

April 24, 2012

Wesley Bruce

Chief Assessment Officer

Norm vs. Criterion

- Norm referenced – interpret performance relative to other students
  - Percentile scores
    - *Examples: SAT, Child Development*
- Criterion referenced - interpret performance relative to a standard
  - Category placement
    - *Examples: ISTEP+, Medical Boards*

Indiana Growth Model Timeline

- Nov 2005 – USED announces "growth model" pilot of AYP
- Winter 05-06 – Indiana applies for pilot
- Aug 2008 – Present Growth and Value Added models to State Board
- April 2009 – State Board's action to implement Student Growth Percentiles
- Sep 2009 – Presented to the Roundtable
- Nov 2009 – Growth data available to corps

Indiana Growth Model Development

- Extensive literature search of existing growth model
  - NCLB approved
  - Models without NCLB caps
- Constructed working models
- Consulted variety of experts
  - Damian Betebenner – NCIEA
  - Derek Briggs – U of Colorado
  - Andrew Ho – Boston College
  - Dan Lewis & Karen Barton – CTB
  - Glenn Ligon – ESP Solutions Group
- Committee participation
  - Accountability Systems and Reporting (CCSSO)

Key Features of the Growth Model

- Growth adds a second dimension to understanding student achievement
- Growth allows easy determination of the progress being made to address achievement gaps
  - We have added "the bottom 25%" disaggregation
  - This serves as an clear benchmark of sound school growth
  - Allows monitoring of the progress of the students in greatest academic need in every school
  - Metric – is growth of the "bottom 25%" greater than "top 75%"?

Student Growth Percentile Adoption

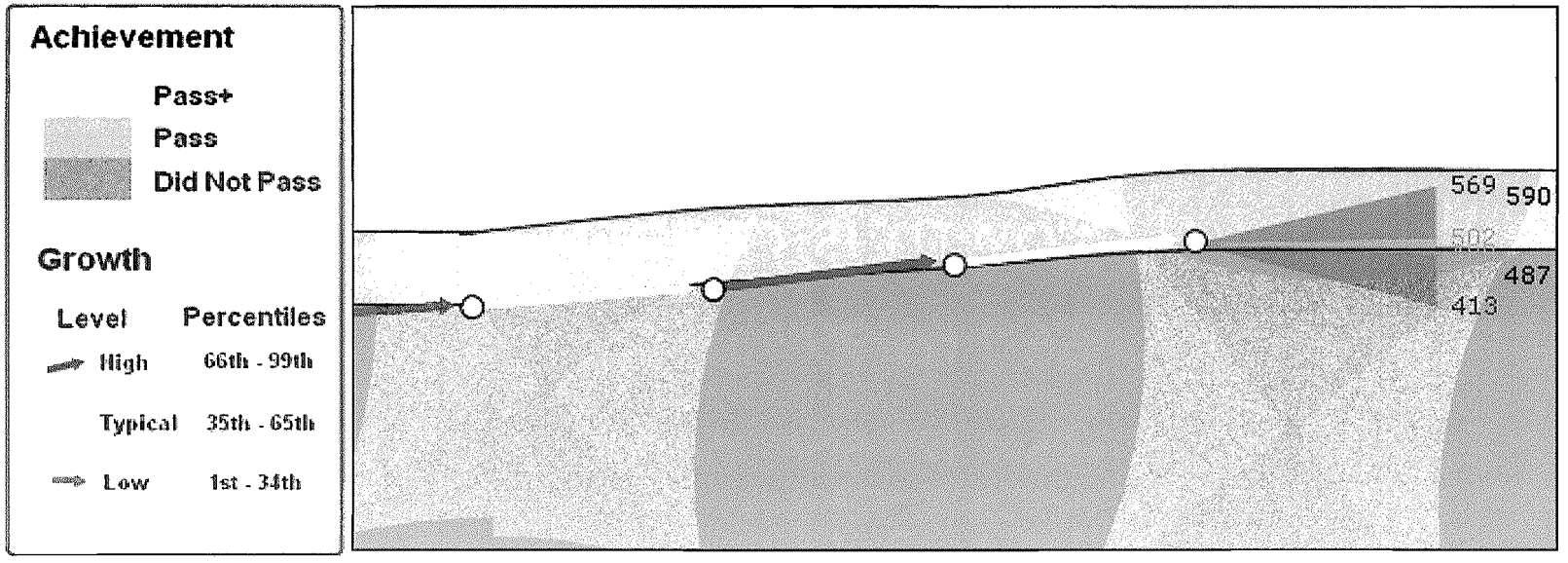
Indiana Growth Examples

- Growth begins at the student level
- Growth is expected for every student
- Student level growth is reported in 3 categories
  - High 66th to 99th percentile
  - Typical 35th to 65th percentile
  - Low 1st to 34th percentile
- Students are grouped into classes and schools, results are reported in median

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

Student Growth Report

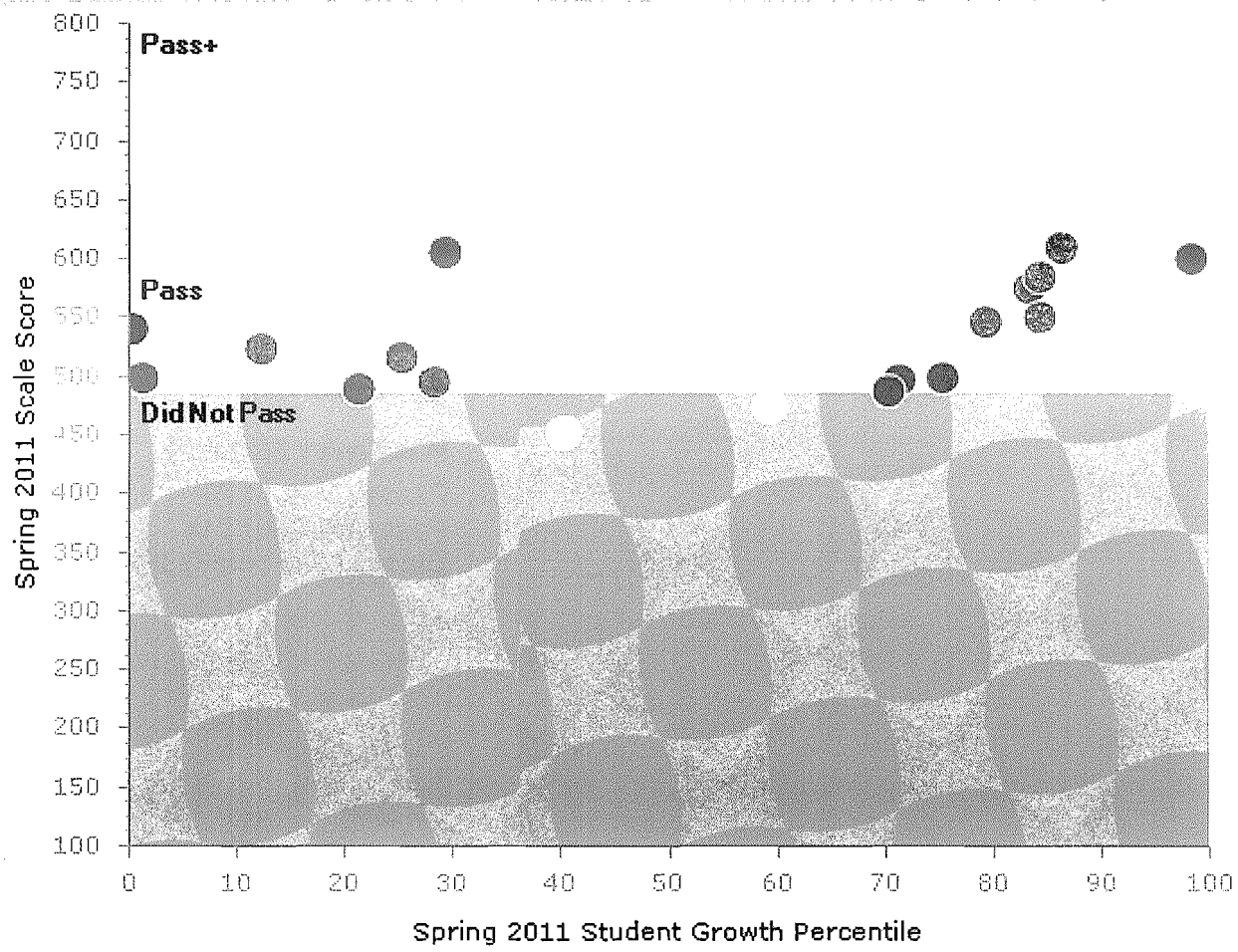
Mathematics



	Grade 4 Fall 2008		Grade 4 Spring 2009		Grade 5 Spring 2010		Grade 6 Spring 2011		Next Year
Scale Score	410		433		466		496		
Growth Percentile		39.0		33.0		57.0			

Class Growth Report (Student)

STUDENT BY ACCOUNTABILITY BY CLASS



Subject Area

- (●) Math
- ( ) English Language Arts

Growth Summary

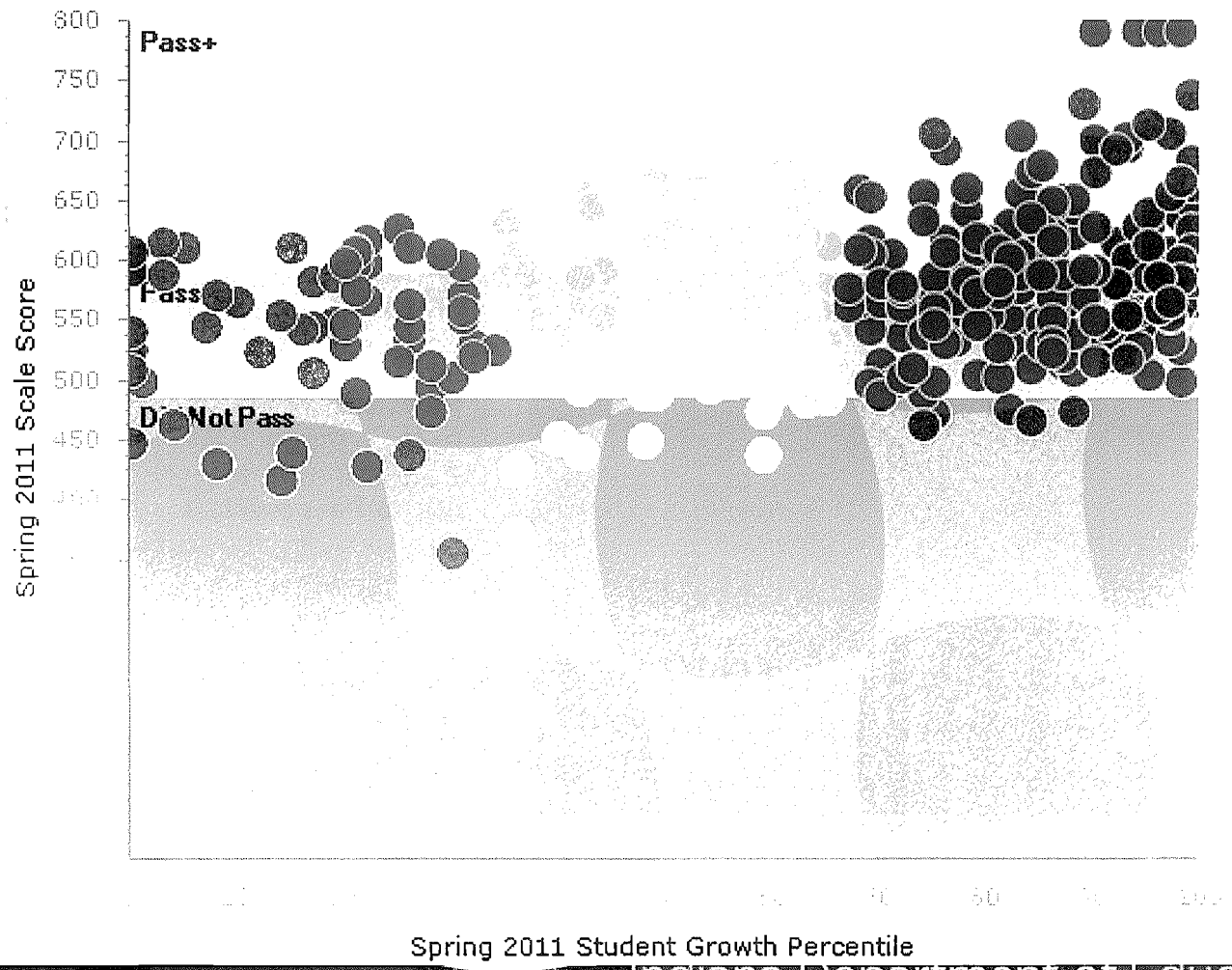
- [x] High Growth (39.1%): 9
- [x] Typical Growth (34.8%):
- [x] Low Growth (26.1%): 6
- [x] Growth N/A: 1

Proficiency Summary

Pass+ (>590): 3
Pass (487-590): 18
Did Not Pass (<487): 3

School Growth Report (Student)

GRADE ACHIEVEMENT AND GROWTH



Subject Area

- (•) Math
- ( ) English Language Arts

Growth Summary

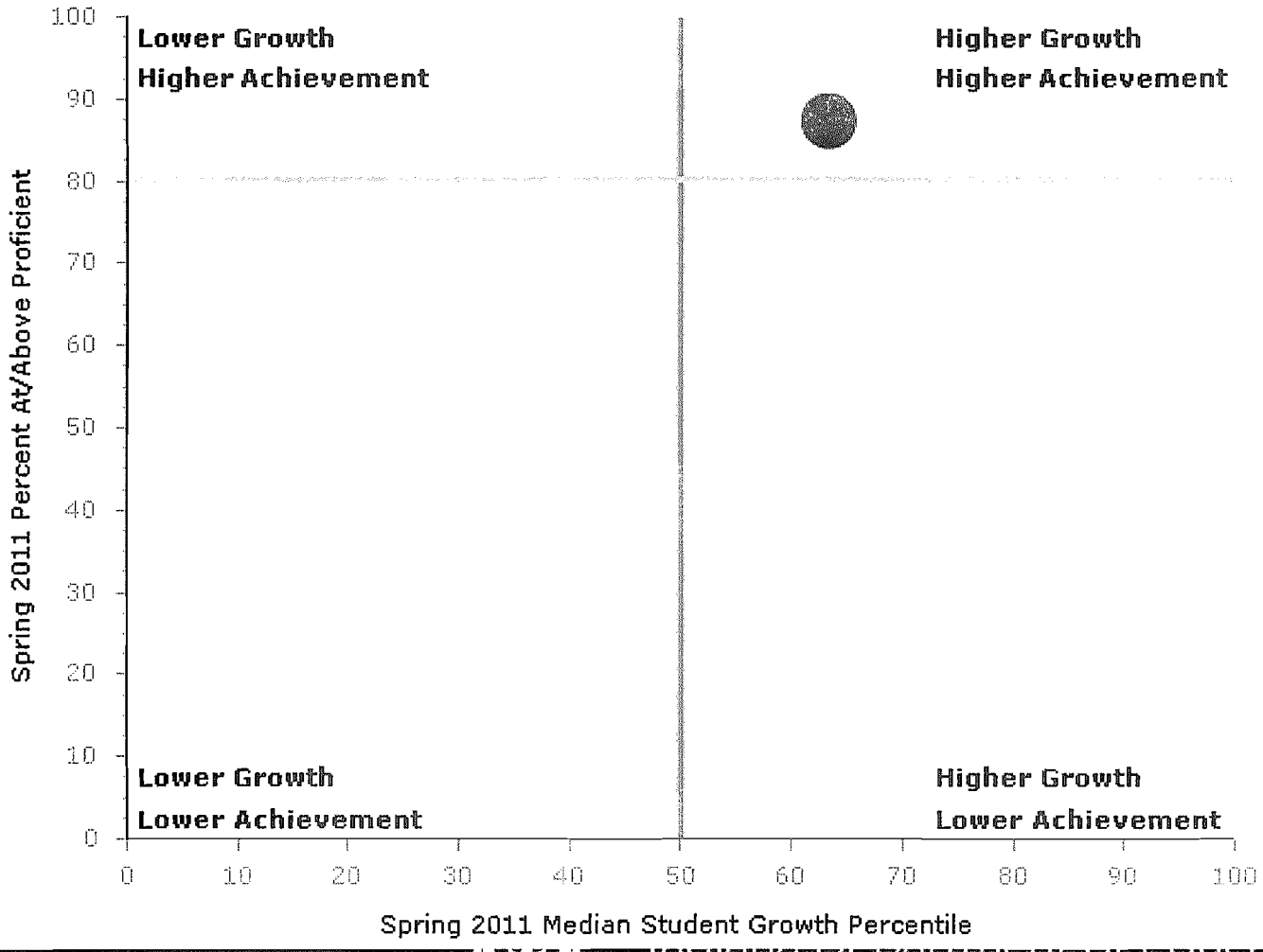
- [x] High Growth (58.3%): 201
- [x] Typical Growth (26.7%): 92
- [x] Low Growth (15.1%): 52
- [x] Growth N/A: 7

Proficiency Summary

Pass+ (>590): 113
Pass (487-590): 214
Did Not Pass (<487): 25

Corporation View

FRANKLIN TOWNSHIP COM SCH CORP (5310): SCHOOL ACHIEVEMENT & GROWTH



Year

Spring 2011

Subject Area

- (●) Mathematics
- ( ) English Language Arts

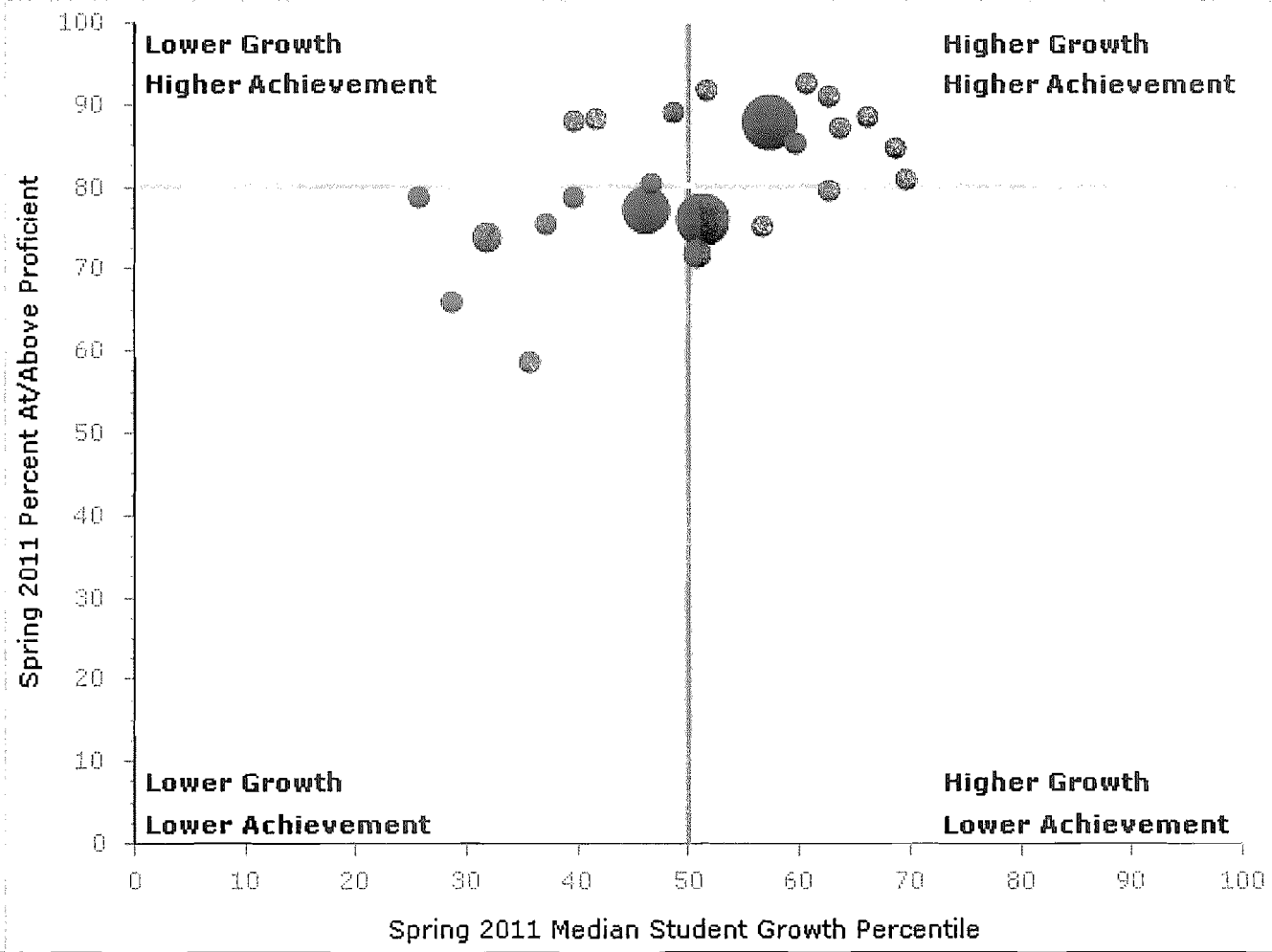
Corporation Overview

Students Enrolled: 0
Students Tested: 3968
Pass Percent: 87.6
Median Growth: 63.0

Return to Search

School Level

VIGO COUNTY SCHOOL CORP (8030): SCHOOL ACHIEVEMENT & GROWTH



Year

Spring 2011

School Level

- [x] Elementary
- [x] Middle

Subject Area

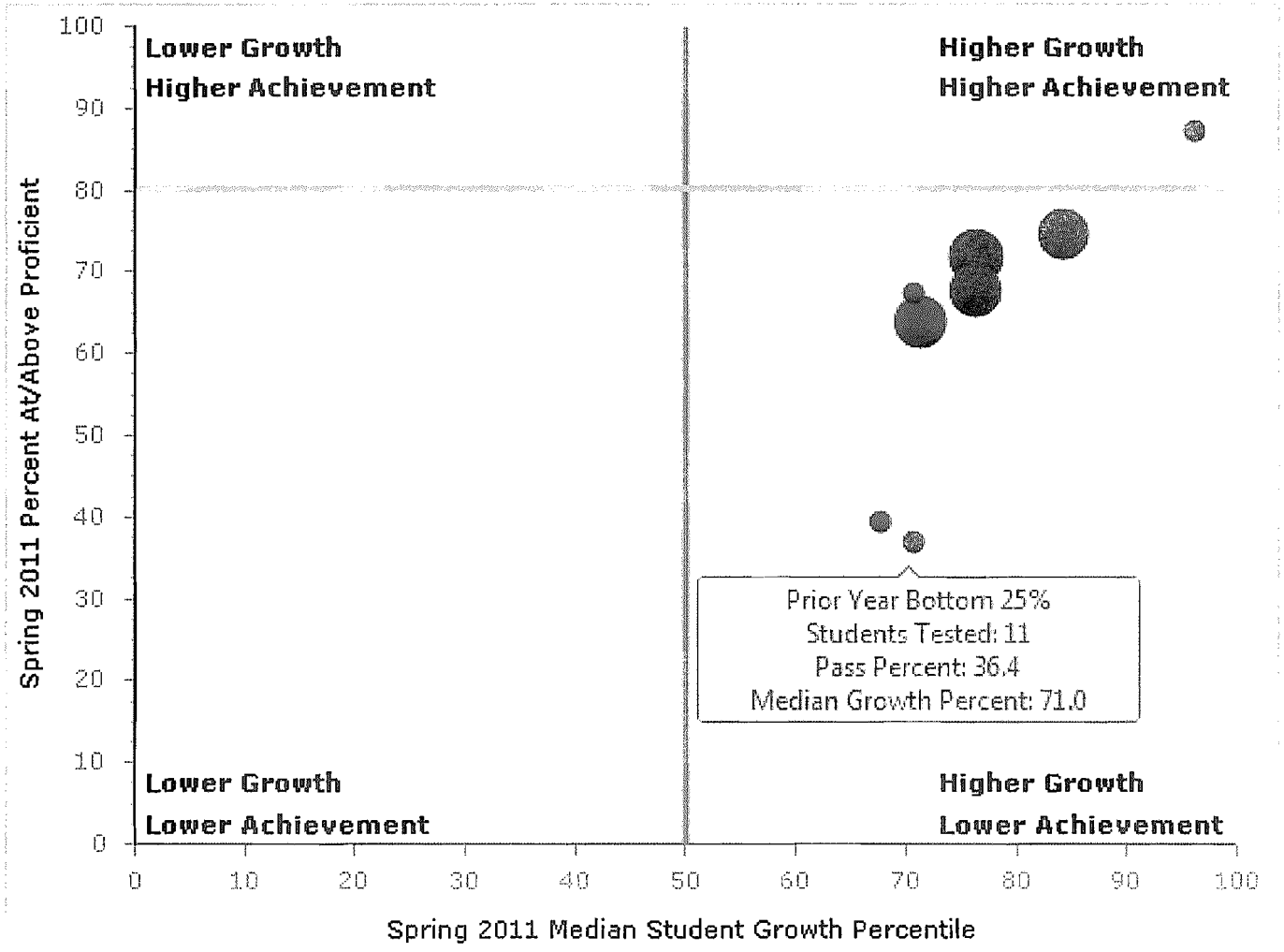
- (•) Mathematics
- ( ) English Language Arts

Corporation Overview

Students Enrolled: 0
Students Tested: 6373
Pass Percent: 79.5
Median Growth: 48.0

Up One Level
Return to Search

School Level (Drilldown)



Year

Spring 2011

Subject Area

- Mathematics
- English Language Arts

Corporation Overview

Students Enrolled: 0
Students Tested: 8114
Pass Percent: 64.9
Median Growth: 46.0

School Overview

Students Enrolled: 0
Students Tested: 93
Pass Percent: 67.7
Median Growth: 73.0

Up One Level
Return to Search

Turn over to Damian

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

Indiana Department of Education

SUPPORTING STUDENT SUCCESS

Student Growth Percentiles: A Developmental Overview

Damian W. Betebenner

National Center for the Improvement of Educational Assessment
Dover, NH

Center for Assessment

Indianapolis, Indiana April 24, 2012

About Me

Dr. Damian Betebenner, PhD

- Senior Associate at the Center for Assessment (NCIEA).
- Developed student growth percentiles and percentile growth trajectories to help states and educational associations employ student growth in decision making [Betebenner, 2008, Betebenner et al., Betebenner, 2012].
- Currently refining and sharing these techniques with other states including Colorado, Massachusetts, Indiana, Nevada, Hawaii and 25 other states in various stages of investigation/adoption.

Student Growth Percentiles

2007	Began develop of the Colorado Growth Model (student growth percentiles).
2008	Began work with Indiana on development of growth model.
2008	Successfully application of Colorado Growth Model (criterion-referenced component) as part of growth model pilot program.
2009-2010	Began/continued work with approximately two-dozen other states on developing state growth models.
2010	Worked with states on incorporating growth into RttT applications and with successful states on their use (Massachusetts, Georgia, New York, Rhode Island).
2011	Worked with a number of states on incorporating student growth percentiles into the ESEA waiver applications and teacher evaluation systems.

About the growth model

- The Student Growth Percentiles Methodology is a norm- and criterion-referenced growth analysis method.
- The norm-referenced component of the model is the student growth percentile which indicates the relative growth a student makes.
- The criterion-referenced component of the model is the student growth percentile target which indicates the amount of growth required by the student to reach/maintain desirable achievement goals in a specified time frame. For example, catching-up to proficiency within 3 years.
- The norm- and criterion-referenced components together allow stakeholders to understand what is, what should be, and what is reasonable simultaneously.
- The goal of the metric was to assist states to move beyond debates over what thermometer is best to discussions of how best to use the thermometer.

Built to maximize sharing

A number of analytic tools (for doing calculations) and visualization tools have been developed with the goal to promote maximum sharing. Data isn't valuable! It's what one does with data that is potentially valuable!

- The analytic tools to calculate all the quantities associated with student growth percentiles are open source and freely available to all states.
- The visualizations associated with student growth percentiles are open source for non-commercial use.
- Conduct training sessions at national conferences to help other calculate these quantities (like Indiana) on their own.
- Build enhanced visualization and reporting systems to make data come alive to stakeholders and lead to better stakeholder engagement with data.

References

- Betebenner, D. W. (2008).
  Toward a normative understanding of student growth.
  In Ryan, K. E. and Shepard, L. A., editors, *The Future of Test-Based Educational Accountability*, pages 155–170. Taylor & Francis, New York.
- Betebenner, D. W. (2009).
  Norm- and criterion-referenced student growth.
  Educational Measurement: Issues and Practice, 28(4):42–51.
- Betebenner, D. W. (2012).
  Growth, standards, and accountability.
  In Cizek, G., editor, *Setting Performance Standards: Foundations, Methods, and Innovations*, pages 439–450. Routledge, New York.

New Accountability Metrics

A – F

Transparency
&
Student Success

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

Goals of Updating Accountability

- Ensure **EVERY** student counts
- Generate dramatic closure of the achievement gap
- Provide transparent results
- Merge state and federal accountability
- Utilize new tools to provide a more robust system
- Engender post-secondary success for **ALL** students
- *These goals address the limitations of the old model*

Project Participants

- Indiana Association of School Principals
- Indiana Association of Public School Superintendents
- Indiana Parent/Teacher Association
- Indiana Charter Schools Association
- Indiana Chamber of Commerce
- Indiana University-Bloomington
- Indianapolis Mayor's Office
- Central Office School Officials
- Elementary and High School Principals
- Teachers
- Community Members
- IDOE staff
- Indiana State Board of Education
- And others…

Timeline on passage of new rule

- Fall, 2010 *Workgroup formed to develop new school accountability frameworks for elementary/middle and high schools. Group comprised of school leaders from around the state, along with the Director of IASP, the Director of CEEP (Indiana University), a member of the Indianapolis Mayor's education division, and IDOE

- Jan 5, 2011 *Different models and ideas from the workgroup were presented to the State Board of Education (SBOE). SBOE encourages further exploration

- Jan thru September, 2011 *Further workgroup deliberations on proposed models, and discussions by IDOE with various groups around the state

- October 5, 2011 *A single new school accountability framework proposed to SBOE. Amendments requested after discussion

Timeline on passage of new rule (Cont.)

- October 6-28, 2011 *Workgroup meetings to address concerns. Recommended changes formulated. IDOE gives presentations throughout the state on the proposed elementary/middle & high school models
- November 7, 2011 *Recommended changes offered for proposed rule. Approved by SBOE to initiate rulemaking
- November 2011 thru January 2012 *IDOE gives more presentations and hosts statewide WebEx on the proposed models
- December 2011 thru January 2012 *Public testimony
- February 8, 2012 *Additional changes offered for proposed rule. Final rule approved by SBOE

Two New Models

- Elementary & Middle Schools
  - ***Performance & Growth***
- High Schools
  - ***Performance & Improvement***

Key Improvements from Old to New

- Simplifies designations awarded to schools (A-F)
- Incorporates **Indiana Growth Model**
  - ***Focuses on ALL students, not just the "bubble kids"***
- Adds a **Bottom 25%** subgroup
  - ***Focuses on closing achievement gaps***
- Includes **College & Career Readiness**

Transparent School & District Performance

Many People Can't Put These in Order

Academic Watch

Commendable Progress

Exemplary School

Commendable School

Academic Probation

Academic Progress

Exemplary Progress

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

Transparent School & District Performance

Almost Everyone Can Put These in Order

C

D

A

F

B

Indiana Department of Education

SUPPORTING STUDENT SUCCESS

Transparent School & District Performance

Exemplary Progress	=	**A**
Commendable Progress	=	**B**
Academic Progress	=	**C**
Academic Watch	=	**D**
Academic Probation	=	**F**

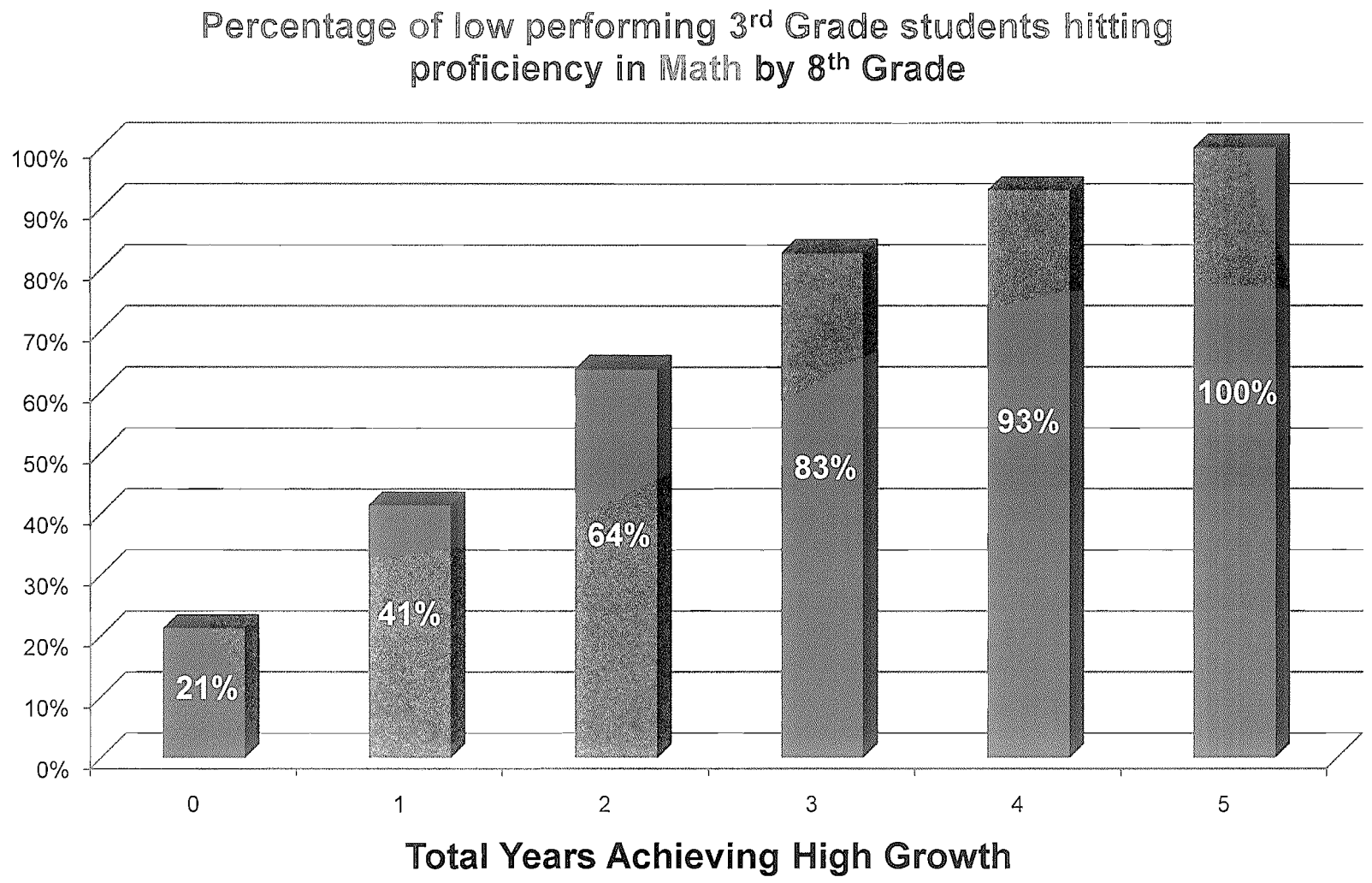
Indiana Department of Education
SUPPORTING STUDENT SUCCESS

Power of the Growth Model

High Growth Matters

Every year our lowest scoring elementary & middle school students demonstrate high growth the more likely they will hit proficiency before high school

Bottom 25% Subgroup



Indiana Department of Education
SUPPORTING STUDENT SUCCESS

Outgrowing AYP

The "N" Size Problem

Under AYP, schools needed at least 30 students in an ESEA subgroup (*e.g. Special Education*) before they were held accountable for them. That threshold meant that in Indiana thousands of minority, LEP, Special Education and Free/Reduced lunch students were lumped into the overall group and effectively lost in accountability.

Power of the Bottom 25% Subgroup

Percentage of Indiana schools held accountable for student performance by traditional ESEA subgroup: Old AYP vs. New PL 221 Model

Traditional ESEA Subgroup	Old	New
American Indian	0%	**16%**
Asian	3%	**31%**
Limited English Proficient	19%	**59%**
Black	23%	**62%**
Hispanic	22%	**71%**
White	91%	**97%**
Free/Reduced Priced Lunch	90%	**99%**
Special Education	57%	**99%**

Closing the Achievement Gap

Progress on closing the achievement gap under the old PL221 and AYP accountability models was incremental.

Sticking with the olds models is simply unacceptable for Hoosier students.

Under the old models…

The achievement gap closure rates over the past decade on ISTEP E/LA and math results are small:

Achievement Gaps Between:	Annual Rate of Closure
General Education and Special Education	0.9%
Paid Lunch and Free/Reduced Lunch	0.3%
Whites and Minorities	0.5%
English Speaking and ELL Students	0.7%

Under the old models…

Maintaining those rates of closure would leave the gaps open for far too long for too many of our kids:

Achievement Gaps Between:	Time Still Needed
General Education and Special Education	23 years
Paid Lunch and Free/Reduced Lunch	43 years
Whites and Minorities	47 years
English Speaking and ELL Students	245 years

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

Closing the Achievement Gap

Progress under the new A-F model raises expectations based on closing the achievement gap.

Coupling the Bottom 25% Subgroup along with the Indiana Growth Model dramatically improves the approach to closing the achievement gap…

Under the NEW model's expectations...

The projected achievement gap closure rates over the next decade on ISTEP E/LA and math are increased:

Achievement Gaps Between:	Annual Rate of Closure
General Education and Special Education	1.8%
Paid Lunch and Free/Reduced Lunch	1.3%
Whites and Minorities	1.6%
English Speaking and ELL Students	1.5%

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

Old versus New…

The new model more than doubles the annual rate of achievement gap closure for all groups:

Achievement Gaps:	Annual Rate of Closure: New	Annual Rate of Closure: Old
General Education and Special Education	1.8%	0.9%
Paid Lunch and Free/Reduced Lunch	1.3%	0.3%
Whites and Minorities	1.6%	0.5%
English Speaking and ELL Students	1.5%	0.7%

Success Beyond High School

According to a recent study from the Georgetown University *Center on Education and the Workforce:*

"55 percent of Indiana's jobs will require postsecondary education by 2018."

-Georgetown University Center on Education and the Workforce, *Help Wanted: Projections of Jobs and Education Requirements Through 2018,* http://www9.georgetown.edu/grad/gppi/hpi/cew/pdfs/State-LevelAnalysis-web.pdf

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

Success Beyond High School

And according to the Lumina Foundation for Education:

"In 2010, only **33 percent** of Hoosiers held at least a two-year degree."

-Lumina Foundation for Education, *A Stronger Nation Through Higher Education*, http://www.luminafoundation.org/state/indiana/

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

Success Beyond High School

College and career readiness measures are our strongest indicators of student post-secondary success:

- Pass an Advanced Placement (AP) exam
- Pass an International Baccalaureate (IB) exam
- Earn college credit (dual credit)
- Earn an industry certification

Setting Goals Matters

Since the establishment of the 25% goal by Dr. Bennett in 2009, Indiana schools have achieved:

(1) the highest increase in the nation in AP exam *participation* for graduates

(2) the second highest increase in the nation in AP exam *passage* for graduates

-The College Board, *AP Report(s) to the Nation*, http://professionals.collegeboard.com/data-reports-research/ap/nation

Strengths of Indiana's Accountability System

Two independent, non-profit and nationally renowned educational organizations have recently given very high praises to Indiana's new accountability system:

- Achieve
- Thomas B. Fordham Institute

Achieve's Review

"Equity Aims: One of the most critical purposes your state's differentiation and classification system is to illuminate and incentivize districts and schools to address gaps in achievement, attainment, and course completion and success among groups of students based on race/ethnicity, income, disability status and English language learner status. (See example on next page)."

Example on next page is Indiana's Statewide AMO (Annual Measurable Objectives) for the Hispanic Subgroup

- *Common Core Implementation Workbook: Take Action: Implementation Action IV, Transition Accountability and Data Reporting System*, Ch. 8 pgs 27-28, updated March, 2012.
http://www.achieve.org/files/Chapter-8-Mar1.pdf

Fordham Institute's Review

"Robust accountability metric. By combining measures of proficiency, growth, and college and career readiness, Indiana captures a detailed picture of student, school and district performance. It also disaggregates the lowest-performing quartile of students to ensure that schools and districts are held to account for those student's performance. In addition, the A-F grading scale for schools and districts is straightforward and easily understood by policymakers, parents, and community members."

-*Defining Strong State Accountability Systems: How Can Better Standards Gain Greater Traction*, April 4, 2012

http://www.edexcellencemedia.net/publications/2012/20120404StateAccountability/040412INAccountabilityFINAL.pdf

Indiana Department of Education

SUPPORTING STUDENT SUCCESS

How do the new models work?

BASES and BONUSES*

- For each model the *preliminary* grade is BASED on student PERFORMACE

*For each model the *preliminary* grade may go up or down due to student GROWTH or IMPROVEMENT and PARTICIPATION

A-F

The BASES are aligned with the state goal:

90 – 25 – 90

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

Indiana's Three Goals

- 90 percent of students pass math and E/LA ISTEP+
- 25 percent of grads pass an AP or IB exam or earn college credits during high school
- 90 percent of students graduate with a diploma

Indiana's Education Scoreboard

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

Elementary & Middle School Model

- **Performance** on ISTEP+, IMAST, & ISTAR (90)

90.0 – 100.0%	4.00 points
85.0 – 89.9%	3.50 points
80.0 – 84.9%	3.00 points
75.0 – 79.9%	2.50 points
70.0 – 74.9%	2.00 points
65.0 – 69.9%	1.50 points
60.0 – 64.9%	1.00 points
0.00 – 59.9%	0.00 points

Elementary & Middle School Model

- **Improvement** on ISTEP+, IMAST, & ISTAR (90)

Absolute improvement is annually recognized because the proficiency scale has five point increments for scores.

When a school's total proficiency rises between 0.1 – 5.0% then their overall score increases.

Elementary & Middle School Model

- **Improvement** on ISTEP+, IMAST, & ISTAR (90)

For example:

In 2011, Smith Middle School had an overall proficiency rate of 78.4 which scores 2.50 points

In 2012, Smith Middle School earns an overall proficiency rate of 80.0 (**1.6% improvement**) which scores 3.00 points

80.0 – 84.9%	*3.00 points*
75.0 – 79.9%	*2.50 points*

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

Elementary & Middle School Model

- **Growth** on ISTEP+
  - Potential bonuses for **High** Growth
  - Potential penalty for **Low** Growth

Elementary & Middle School Model

There are **four** opportunities for the **High** Growth bonus across the two subgroups:

		Criteria
1)	Bottom 25% on E/LA	(42.5%)
2)	Bottom 25% on math	(44.9%)
3)	Top 75% on E/LA	(36.2%)
4)	Top 75% on math	(39.2%)

Elementary & Middle School Model

There are **two** circumstances for the **Low** Growth deduction:

	Criteria
1) Overall group on E/LA	(39.8%)
2) Overall group on math	(42.4%)

Elementary & Middle School Model

Summary Impact of Growth

The preliminary grade based on proficiency and absolute improvement may be **raised** based on **HIGH** academic **growth on E/LA and/or math** for students who performed in the bottom 25% on ISTEP+.

The preliminary grade based on proficiency and absolute improvement may be **raised** based on **HIGH** academic **growth on E/LA and/or math** for students who performed in the top 75% on ISTEP+.

The preliminary grade based on proficiency and absolute improvement may be **lowered** due to **LOW** academic **growth on E/LA and/or math** of 100% of students for performance on ISTEP+.

Elementary & Middle School Model

Impact of Participation

The preliminary grade may also be **lowered** if **less than 95%** of students in any of the subgroups **participate** in the ISTEP+, IMAST, or ISTAR exams. The possible deduction may occur in both E/LA and math assessments.

Elementary & Middle School Model

Calculating Grades

- Separate grades (points) are determined for English/Language Arts and Math

- A **final grade** is determined by averaging the English/Language Arts and Math grades (points)

Elementary & Middle School Model

Convert the averaged E/LA and math grades (points) to a **FINAL** letter grade for a school, using the following scale:

3.51 - 4.00 points	=	A
3.00 - 3.50 points	=	B
2.00 - 2.99 points	=	C
1.00 - 1.99 points	=	D
0.00 - 0.99 points	=	F

High School Model

- Schools assessed in four weighted areas:
  - English 10 ECA (30%)
  - Algebra 1 ECA (30%)
  - Graduation Rate (30%)
  - College & Career Readiness (10%)

High School Model

- English 10 and Algebra I ECAs (90)
  - **Performance** (*base*)

90.0 – 100.0%	4.00 points
85.0 – 89.9%	3.50 points
80.0 – 84.9%	3.00 points
75.0 – 79.9%	2.50 points
70.0 – 74.9%	2.00 points
65.0 – 69.9%	1.50 points
60.0 – 64.9%	1.00 points
0.00 – 59.9%	0.00 points

High School Model

- **Improvement** on ISTEP+ & ISTAR (90)

For example:

In 2011, Smith High School had an overall proficiency rate of 68.4 which scores 1.50 points

In 2012, Smith High School earns an overall proficiency rate of 70.0 (**1.6% improvement**) which scores 2.00 points

70.0 – 74.9%	*2.00 points*
65.0 – 69.9%	*1.50 points*

High School Model

- English 10 and Algebra 1 ECAs

Improvement *(bonus)*

- 8th grade to 10th grade
- 10th grade to graduation

High School Model

Improvement (*bonus*)

- 8th grade to 10th grade — Criteria
  - E/LA possible bonus (10.3%)
  - Math possible bonus (17.1%)
- 10th grade to graduation
  - E/LA possible bonus (59.3%)
  - Math possible bonus (62.8%)

High School Model

Improvement

	Criteria
• 8th grade to 10th grade	
– E/LA possible deduction	(-0.1%)
– Math possible deduction	(-0.1%)

High School Model

- College & Career Readiness (25)
  - Pass an AP exam (*base*)
  - Pass an IB exam (*base*)
  - Earn College Credits (3) (*base*)
  - Earn an Industry Certification (*base*)

High School Model

- College & Career Readiness

			(base)
25.0 -	100%	=	4.00 points
18.4 -	24.9%	=	3.00 points
11.7 -	18.3%	=	2.00 points
05.0 -	11.6%	=	1.00 points
00.0 -	04.9%	=	0.00 points

High School Model

- Graduation Rate (90)

4-year rate (*base*)

90.0 – 100.0%	4.00 points
85.0 – 89.9%	3.50 points
80.0 – 84.9%	3.00 points
75.0 – 79.9%	2.50 points
70.0 – 74.9%	2.00 points
65.0 – 69.9%	1.50 points
60.0 – 64.9%	1.00 points
0.00 – 59.9%	0.00 points

High School Model

- Graduation Rate

  Improvement (*bonus*)*

  - Diploma type*
  - 5-year rate*

*These components of the model begin in 2014-15

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

Remediation by Diploma Type

The recently published "College Readiness Report" by the *Indiana Commission for Higher Education* shows that diploma type is a strong indicator of remediation needs in college:

The 2010 report stated that 31% of Indiana high school graduates required remediation in college. The remediation rates were higher for particular diplomas types:

69% for *General &Waiver* diploma recipients

39% for *Core 40* diploma recipients

7% for *Honors* diploma recipients

- Indiana Commission for Higher Education, *College Readiness Reports*, http://www.in.gov/che/2489.htm

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

High School Model

Calculating Grades

- *Weighted* grades (points) are determined for:
  - English 10 ECA — **Score x 30%**
  - Algebra I ECA — **Score x 30%**
  - Graduation Rate — **Score x 30%**
  - College and Career Readiness — **Score x 10%**
- A **final grade** is determined by adding the four grades (points) together

High School Model

Convert the combined grades (points) of the Eng. 10 ECA, math ECA, Grad Rate, & CCR to a **FINAL** letter grade for a school, using the following scale:

3.51 - 4.00 points	=	A
3.00 - 3.50 points	=	B
2.00 - 2.99 points	=	C
1.00 - 1.99 points	=	D
0.00 - 0.99 points	=	F

Combined Schools & Corporation Grades

- Schools that serve grades that encompass both models will receive one **FINAL** grade by adding the two scores from both models together after multiplying those scores by the enrollment percentages weighted accordingly.

- Corporations are treated as a single large school. They will receive one **FINAL** grade using the same methodology outlined above.

What's My Grade?

IDOE developed a workbook with all of the proficiency measures and bonuses already incorporated which allows schools to simply input their achievement data and have their grade automatically calculated.

Found here:

http://www.doe.in.gov/improvement/accountability/f-accountability

What does the Report Card look like?

IDOE worked with stakeholders to build sample report cards for both the Elementary/Middle and High School models.

They are found here:

http://www.doe.in.gov/improvement/accountability/f-accountability

Goals of Updating Accountability

- Ensure **EVERY** student counts
- Generate dramatic closure of the achievement gap
- Provide transparent results
- Merge state and federal accountability
- Utilize new tools to provide a more robust system
- Engender post-secondary success for **ALL** students
- *These goals address the limitations of the old model*

Additional Questions?

schoolaccountability@doe.in.gov

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

A-F ACCOUNTABILITY

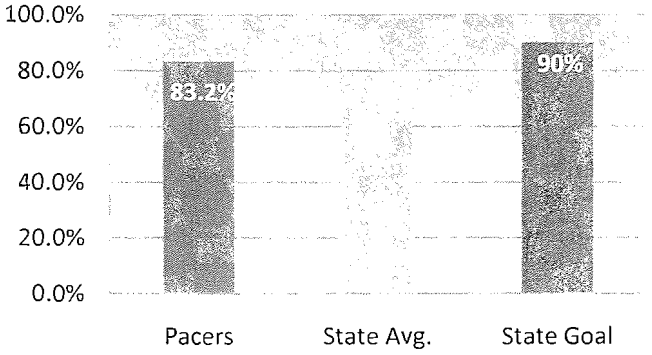
Pacers High School Report Card, 2011-2012

Grade Points by Category		Overall Grade
English10 ECA:	*3.50 pts (B)*	3.05 Points
Algebra 1 ECA:	*2.00 pts (C)*	**B**
Graduation Rate:	*4.00 pts (A)*	
College & Career Readiness:	*2.00 pts (C)*	

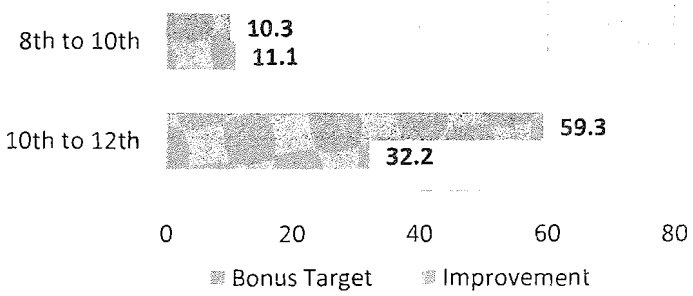
The final grade reflects student performance and improvement on Indiana's basic skills graduation exams English 10 and Algebra I End of Course Assessments (ECA) along with performance on College and Career Readiness indicators and on-time (4-year) Graduation Rate. Student results on Indiana's alternate assessment, ISTAR, are included in the performance calculations. *Click on the subject areas to learn more about the scoring and calculations used to determine the final grade.*

English 10

Performance on ECA and ISTAR

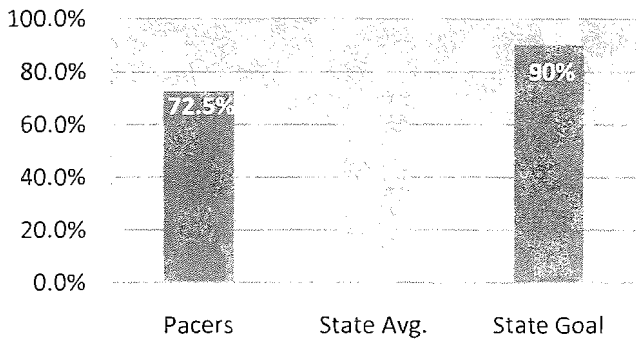


Student Improvement

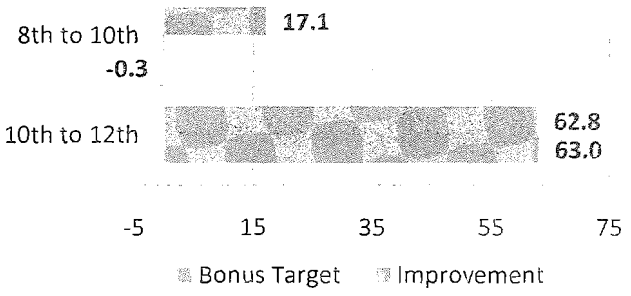


Algebra I

Performance on ECA and ISTAR

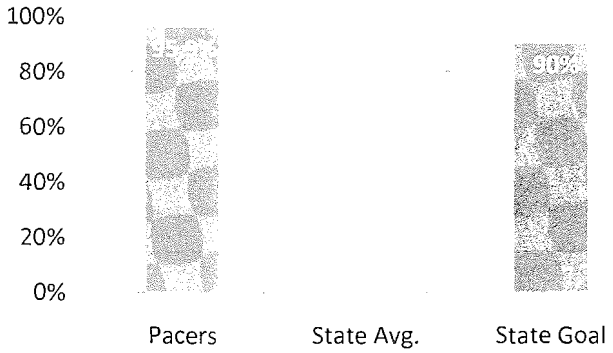


Student Improvement



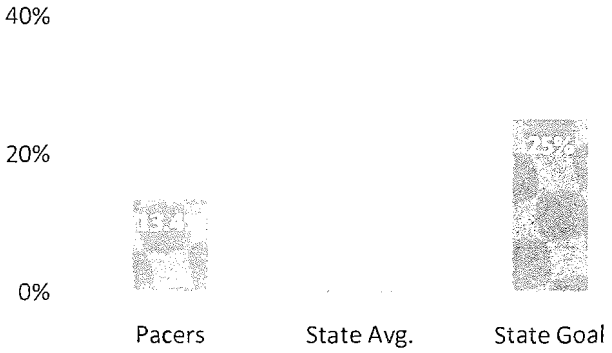
Graduation Rate

4-Year Grad Rate



College & Career Readiness

% of Graduates Successful



Indiana Department of Education
SUPPORTING STUDENT SUCCESS

A-F ACCOUNTABILITY

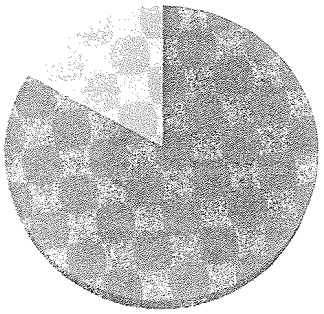
Pacers High School Report Card, 2011-2012

English 10 ECA (30% of Final Grade):	Overall Grade:
Performance: 3.00 Points	3.05 Points
Bonus: 0.50 Points	**B**
Total: 3.50 Points	

	Preliminary Score
Performance	
% of Students Passing the English 10 ECA and ISTAR:...83.2%	3.00 points

Points Scale

90.0 – 100%	=	4.00 points
85.0 – 89.9%	=	3.50 points
80.0 – 84.9%	**=**	**3.00 points**
75.0 – 79.9%	=	2.50 points
70.0 – 74.9%	=	2.00 points
65.0 – 69.9%	=	1.50 points
60.0 – 64.9%	=	1.00 points
00.0 – 59.9%	=	0.00 points



	Bonus Points
Improvement	
8th grade to 10th grade Improvement: 11.1%	+0.50 points
Target = > 10.3	
10th grade to 12th grade Improvement: 32.2%	+0.00 points
Target = ≥ 59.3	
Final Score	**3.50 points**

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

A-F ACCOUNTABILITY

Pacers High School Report Card, 2011-2012

Algebra 1 ECA (30% of Final Grade):
Performance: 2.00 Points
Bonus: 0.00 Points
Total: 2.00 Points

Overall Grade:
3.05 Points
B

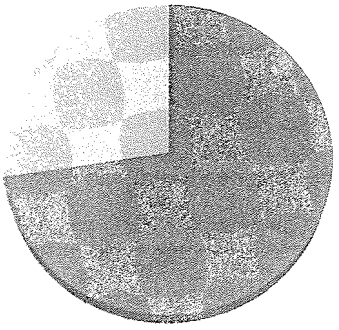
Preliminary Score

Performance

% of Students Passing the Algebra I ECA and ISTAR: .....72.5% — 2.00 points

Points Scale

90.0 – 100%	=	4.00 points
85.0 – 89.9%	=	3.50 points
80.0 – 84.9%	=	3.00 points
75.0 – 79.9%	=	2.50 points
70.0 – 74.9%	**=**	**2.00 points**
65.0 – 69.9%	=	1.50 points
60.0 – 64.9%	=	1.00 points
00.0 – 59.9%	=	0.00 points



Bonus Points

Improvement

8th grade to 10th grade Improvement: ..... -0.3% — -0.50 points

Target = > 17.1

10th grade to 12th grade Improvement:.....63.0% — +0.50 points

Target = ≥ 62.8

Final Score ..... **2.00 points**

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

A-F ACCOUNTABILITY

Pacers High School Report Card, 2011-2012

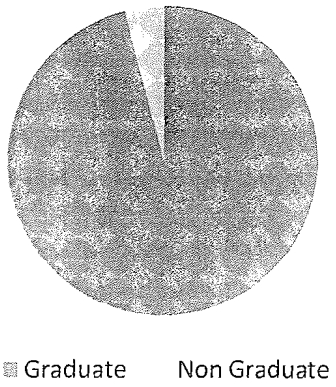
Graduation Rate (30% of Final Grade): *Performance: 4.00 Points* *Bonus: NA* ***Total: 4.00 Points***	Overall Grade: 3.05 Points **B**

Performance

	Score
4-Year Graduation Rate: 95.9%	4.00 points

Points Scale

90.0 – 100%	=	**4.00 points**
85.0 – 89.9%	=	3.50 points
80.0 – 84.9%	=	3.00 points
75.0 – 79.9%	=	2.50 points
70.0 – 74.9%	=	2.00 points
65.0 – 69.9%	=	1.50 points
60.0 – 64.9%	=	1.00 points
00.0 – 59.9%	=	0.00 points



Final Score ..... **4.00 points**

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

A-F ACCOUNTABILITY

Pacers High School Report Card, 2011-2012

College & Career Readiness (10% of Final Grade):	Overall Grade:
Performance: 2.00 Points *Bonus: NA* ***Total: 2.00 Points***	3.05 Points **B**

Total number of Graduates 268

Number of Graduates who:

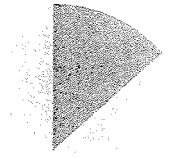
Passed an Advanced Placement Exam 21
Passed an International Baccalaureate Exam.....6
Received at least 3 Hours of College Credit 7
Received an Industry Certification 2

Performance

	Score
% of graduates who passed an AP or IB Exam, or received at least 3 hours of college credit, or received an Industry Certification: 13.4%	2.00 points

Points Scale

25.0 – 100% = 4.00 points
18.4 – 24.9% = 3.00 points
11.7 – 18.3% = 2.00 points
5.0 – 11.6% = 1.00 points
0.0 – 4.9% = 0.00 points



CCR without CCR

Final Score ..... 2.00 points

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

A-F ACCOUNTABILITY

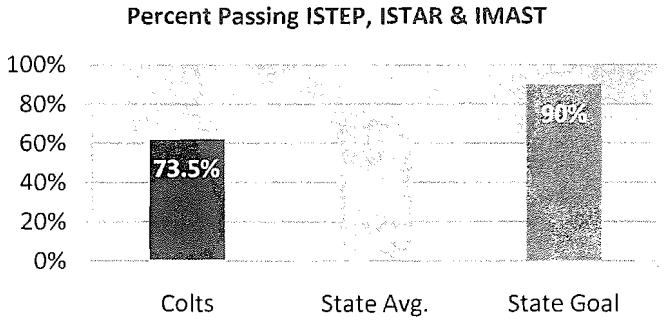
Colts Middle School Report Card, 2011-2012

Grade Points by Category	Overall Grade:
English/Language Arts: *2.00 pts (C)*	3.00 Points **B**
Math: *4.00 pts (A)*	

The final grade reflects student performance and growth on Indiana's English/Language Arts and Math basic skills tests called ISTEP+. Student performance on Indiana's alternate assessment, ISTAR, and modified assessment, IMAST, are included in the performance calculations. *Click on the subject areas to learn more about the scoring and calculations used to determine the final grade.*

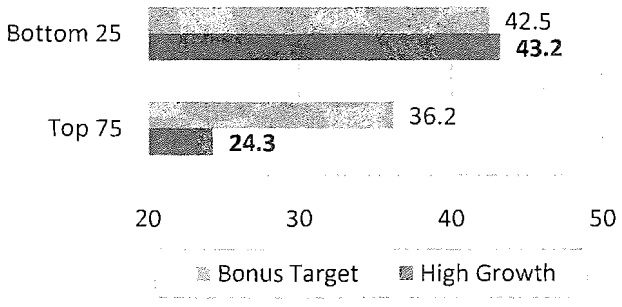
English/Language Arts

Student Performance

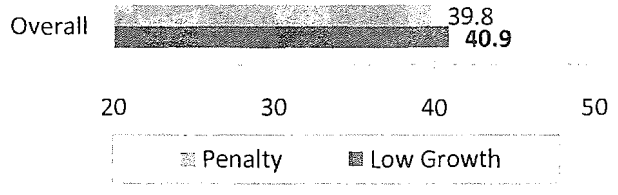


Student Growth

Percent of Students Achieving **High** Growth

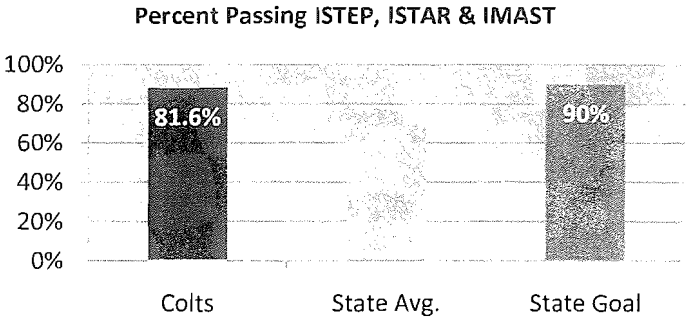


Percent of Students Showing **Low** Growth



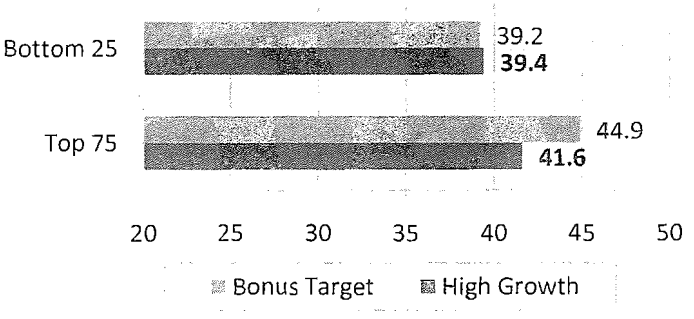
Mathematics

Student Performance

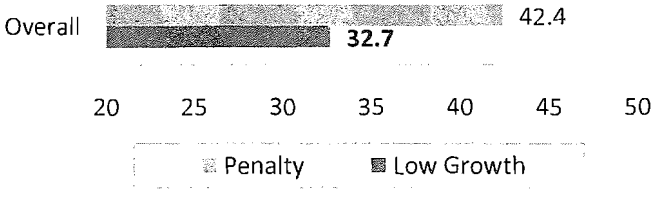


Student Growth

Percent of Students Achieving **High** Growth



Percent of Students Showing **Low** Growth



Indiana Department of Education
SUPPORTING STUDENT SUCCESS

A-F ACCOUNTABILITY

Colts Middle School Report Card, 2011-2012

Math (50% of Final Grade):	Overall Grade:
Performance: 3.00 Points	3.00 Points
Bonus: 1.00 Points	**B**
Total: 4.00 Points	

Performance

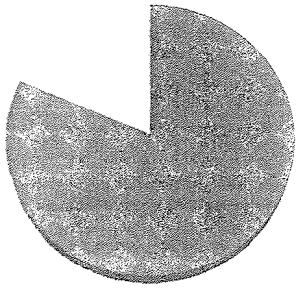
% of Students Passing Math ISTEP, ISTAR & IMAST:.....81.6%

Preliminary Score

3.00 points

Points Scale

90.0 – 100%	=	4.00 points
85.0 – 89.9%	=	3.50 points
80.0 – 84.9%	**=**	**3.00 points**
75.0 – 79.9%	=	2.50 points
70.0 – 74.9%	=	2.00 points
65.0 – 69.9%	=	1.50 points
60.0 – 64.9%	=	1.00 points
00.0 – 59.9%	=	0.00 points



Bonus Points

Growth

% of lowest performing students with *high growth*:.....41.6% — 0.00 points
Target is ≥ 44.9%

% of highest students with *high growth*: .....39.4% — +1.00 points
Target is ≥ 39.2%

% of ALL students with *low growth*: .....32.7% — 0.00 points
Target is < 42.4%

Participation

% of lowest performing students participating in Math:..96.8% — -----
Target is ≥ 95.0%

% of remaining students participating in Math:.....95.2% — -----
Target is ≥ 95.0%

Final Score ..... **4.00 points**

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

A-F ACCOUNTABILITY

Colts Middle School Report Card, 2011-2012

English / Language Arts (50% of Final Grade):	Overall Grade:
Performance: 2.00 Points *Bonus: 0.00 Points* ***Total: 2.00 Points***	3.00 Points **B**

Performance

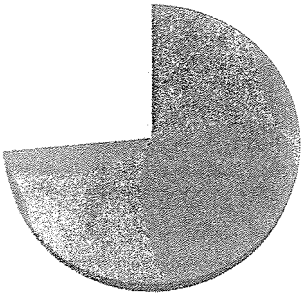
% of Students Passing E/LA ISTEP, ISTAR & IMAST:.....73.5%

Preliminary Score

2.00 points

Points Scale

90.0 – 100%	=	4.00 points
85.0 – 89.9%	=	3.50 points
80.0 – 84.9%	=	3.00 points
75.0 – 79.9%	=	2.50 points
70.0 – 74.9%	**=**	**2.00 points**
65.0 – 69.9%	=	1.50 points
60.0 – 64.9%	=	1.00 points
00.0 – 59.9%	=	0.00 points



Pass Did Not Pass

Bonus Points

Growth

% of lowest performing students with *high growth*:.....43.2% — +1.00 point
Target is ≥ 42.5%

% of highest students with *high growth*: .....24.3% — 0.00 points
Target is ≥ 36.2%

% of ALL students with *low growth*: .....40.9% — -1.00 point
Target is < 39.8%

Participation

% of lowest performing students participating in E/LA...96.4% — -----
Target is ≥ 95.0%

% of remaining students participating in E/LA:.....95.1% — -----
Target is ≥ 95.0%

Final Score ..... **2.00 points**

A policy brief from Lumina Foundation for Education

A stronger nation through higher education — and Indiana's role in that effort

In Indiana, 33 percent of the state's nearly 3.4 million working-age adults (25-64 years old) hold at least a two-year degree, according to 2008 Census data. This compares to a national average of around 38 percent.

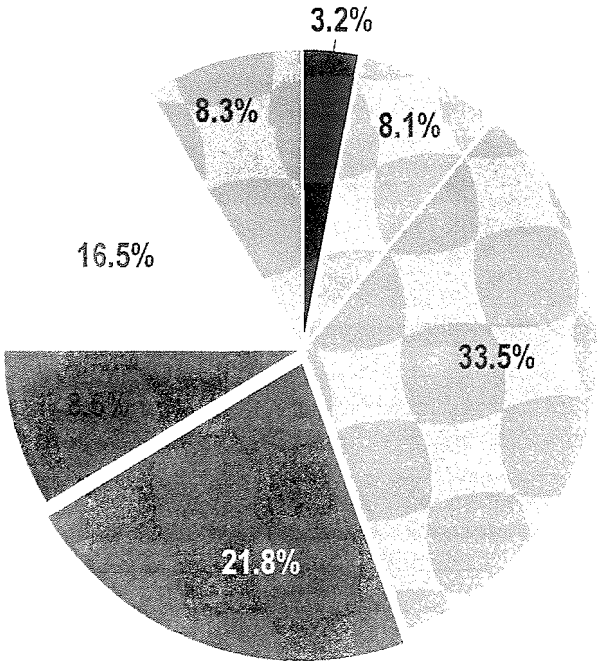
Attainment rates in Indiana are increasing modestly, even though the proportion of degree-holding young adults — those 25-34 years old — mirrors that of the overall adult population. If Indiana continues to increase attainment at the rate it did over the last decade (2000-2008), the state will have a college-attainment rate of 44 percent in 2025 — far short of the Big Goal of 60 percent.

However, this gap *can* be closed. The key is to begin increasing degree production, and to continue to increase it each year until 2025. By increasing production by 6,454 associate or bachelor's degrees each year between now and 2025 — an annual increase of 6.3 percent — Indiana will reach the Big Goal.

One excellent place to begin looking for these additional graduates is in the ranks of Indiana residents who have completed some college without earning a degree. In 2008, 730,000 Indiana residents fit into this category — representing nearly 22 percent of the state's adult population. If only a small portion of this group could be enticed to return to college to complete either a two- or four-year degree, it would go a long way to helping Indiana reach the goal of 60 percent higher education attainment.

Also, by looking at the geographic distribution of college graduates within the state (see chart on reverse), policymakers and other stakeholders can begin to work strategically and systematically to close achievement gaps. They can target counties and regions that show the greatest need and focus their efforts on those specific areas.

How do we know that Indiana's economy will demand more college graduates? A recent analysis by the Georgetown University Center on Education and the Workforce provides the answer. According to the center's analysis of occupation data and workforce trends, 55 percent of Indiana's jobs will require postsecondary education by 2018. Between now and 2018, Indiana will need to fill about 930,000 vacancies resulting from job creation, worker retirements and other factors. Of these job vacancies,



Levels of education for Indiana residents, ages 25-64

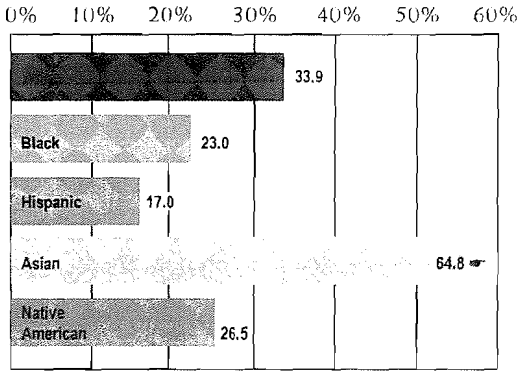
Less than ninth grade	108,812	3.2%
Ninth to 12th grade, no diploma	273,086	8.1%
High school graduate (including equivalency)	1,125,166	33.5%
Some college, no degree	734,541	21.8%
Associate degree	290,493	8.6%
Bachelor's degree	554,593	16.5%
Graduate or professional degree	277,639	8.3%
TOTAL	3,364,330	100%

Source: U.S. Census Bureau, 2008 American Community Survey

506,000 will require postsecondary credentials, while only 424,000 are expected to be filled by high school graduates or dropouts.

It will be impossible to reach the Big Goal without significantly increasing college success among the groups that can accurately be called 21st century students, including working adults, low-income and first-generation students and students of color. Closing attainment gaps among racial and ethnic groups is a particular challenge in Indiana, as in other states. The current higher education attainment rates in Indiana are displayed in the accompanying bar graph.

Degree-attainment rates among Indiana adults (ages 25-64), by population group



Source: U.S. Census Bureau, 2008 American Community Survey

These gaps have persisted for decades, and closing them is obviously a big challenge. Still, it can and *must* be done. It will require concerted and strategic efforts over many years — efforts based on solid evidence about what works to increase attainment. Information about successful strategies to increase the number of students who complete higher education is available on Lumina Foundation's Web site. The site also provides specific information about Indiana's degree-attainment rates at **www.luminafoundation.org/state_data/**. From there, you can find links to data from all 50 states.

Still more information is available at a Web-based resource created by the National Center for Higher Education Management Systems (NCHEMS). The NCHEMS Information Center provides detailed comparative data for all states and counties, as well as other contextual information that can help higher education policymakers and analysts make sound policy decisions. We urge you to visit the site (**www.higheredinfo.org**).

Percentage of Indiana adults (25-64) with a two- or four-year degree, by county:

Adams	22.7	Fulton	21.9	Kosciusko	27.2	Ripley	25.4
Allen	37.1	Gibson	28.4	LaGrange	14.2	St. Joseph	36.2
Bartholomew	39.2	Grant	27.2	Lake	29.1	Scott	21.0
Boone	48.5	Greene	24.0	LaPorte	26.1	Shelby	27.4
Cass	20.1	Hamilton	63.6	Lawrence	20.4	Spencer	31.1
Clark	28.4	Hancock	38.3	Madison	25.4	Starke	18.6
Clay	28.4	Harrison	22.9	Marion	36.2	Steuben	26.9
Clinton	20.9	Hendricks	43.8	Marshall	25.5	Sullivan	28.5
Daviess	22.6	Henry	24.3	Miami	20.2	Tippecanoe	45.1
Dearborn	28.0	Howard	27.8	Monroe	50.2	Vanderburgh	33.0
Decatur	22.6	Huntington	26.6	Montgomery	24.4	Vigo	32.2
DeKalb	28.3	Jackson	23.6	Morgan	24.8	Wabash	27.1
Delaware	32.0	Jasper	23.9	Noble	23.2	Warrick	40.1
Dubois	30.7	Jay	20.0	Owen	16.2	Washington	16.7
Elkhart	24.4	Jefferson	25.6	Porter	35.6	Wayne	25.4
Fayette	18.9	Jennings	16.7	Posey	29.4	Wells	27.6
Floyd	33.0	Johnson	36.7	Putnam	25.6	White	21.9
Franklin	28.9	Knox	34.2	Randolph	20.6	Whitley	30.1
						Other counties	20.9*

*This percentage is an average for the 20 Indiana counties with fewer than 20,000 residents.

Source: U.S. Census Bureau, 2008 American Community Survey

Lumina Foundation for Education is committed to enrolling and graduating more students from college — especially low-income students, students of color, first-generation students and adult learners. Our goal is to increase the percentage of Americans who hold high-quality degrees and credentials to 60 percent by 2025. Lumina pursues this goal in three ways: by identifying and supporting effective practice, through public policy advocacy, and by using our communications and convening power to build public will for change. For more details on the foundation, visit our Web site at **www.luminafoundation.org**.

August 2010 © Lumina Foundation for Education, Inc.

LUMINA

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

OVERVIEW OF IMPLEMENTATION

SUBJECT: A-F Metrics
PL 221-1999
http://www.in.gov/legislative/ic/code/title20/ar31/ch8.html

SUBMITTED BY: Jon Gubera, Chief Accountability Officer

The Indiana General Assembly passed Public Law 221 in 1999. The first rules to implement the statute were finalized by the State Board of Education (SBOE) in 2001. The first category placements were made for the 2004-2005 school year and were based on the rules established by the SBOE.

The updated (2012) metrics were designed to aggressively attack the achievement gap, take advantage of the latest advancements in assessment technology, and better meet the needs of the 21st century Hoosier student. Reflecting nearly two years of deliberations, the metrics were thoroughly reviewed by a workgroup comprised of a broad range of stakeholders. The new metrics have already garnered national accolades. Two independent, non-profit, nationally renowned educational organizations, Achieve and the Fordham Institute, have lauded the Hoosier state's new model, citing its "equity aims" and "robust accountability metric."

Requirements

IC 20-31-10-1 Adoption of rules

Sec. 1. The state board may adopt rules under IC 4-22-2 to implement this article.

IC 20-31-8-3 Categories or designations of school improvement established

Sec. 3. The state board shall establish a number of categories or designations of school improvement based on the improvement that a school makes in performance of the measures determined by the board with the advice of the education roundtable.

IC 20-31-8-4 Placement of school in category or designation of school improvement

Sec. 4. The state board shall place each school in a category or designation of school improvement.

Implementation

The nomenclature for categorical placement was changed by the SBOE in 2010, shifting from "Exemplary Progress, Commendable Progress, Academic Progress, Academic Watch and Academic Probation" to a more transparent system of "A, B, C, D and F." The shift from the previous difficult-to-understand categories to a letter grading system has been well received by the public.

In the summer of 2010, the Indiana Department of Education (IDOE) began conversations about updating the metrics for determining categorical placements based in large part on two significant developments: (1) the successful launch and implementation of the powerful Indiana Growth Model, which provides valuable information to educators and parents on the progress and achievement of **ALL** students in English/Language Arts and Math on the ISTEP+ exams, and (2) the need to incorporate college and career measures, knowing that most jobs and careers now require the completion of post-secondary education.

In early fall of 2010, the IDOE formed a workgroup comprised of school leaders from around the state to guide the development of updating the metrics to incorporate the latest technologies and address the needs for student success in the 21st century.

The goals for updating the school accountability system were to do the following: (1) ensure **EVERY** student counts, (2) generate dramatic closure of the achievement gap, (3) provide transparent results, (4) merge state and federal accountability into a single system, (5) utilize new tools to provide a more robust system, and (6) engender post-secondary success for **ALL** students.

The workgroup developed various models and shared them throughout the state to generate discourse and solicit feedback. The group presented its early findings to the SBOE on January 5, 2011. The SBOE encouraged further exploration.

Eventually two models – one for elementary/middle schools and one for high schools – were formulated and introduced to the SBOE on October 5, 2011. Although these models resonated with SBOE members, they asked for additional fine-tuning.

The workgroup gathered again to address concerns articulated by the SBOE and other stakeholders. Recommended changes were presented to and formally approved by the SBOE on November 7, 2011, thus initiating the rulemaking process.

Public comment was open from December 2011 thru late January 2012. On February 8, 2012, additional changes were recommended and subsequently approved by the SBOE, finalizing the rule.

There are three new additions to the state's A to F metrics. Specifically, the incorporation of Indiana's Growth Model, a laser-like focus on a new subgroup called the "Bottom 25%," and a commitment to college and career readiness indicators represent a substantive departure from the old accountability model. These changes focus on generating success for **ALL** students beyond high school.

Support

Throughout the nearly two years of deliberations on the updating of the school accountability metrics, the IDOE has provided numerous communications, including supporting documents, resources, and tools. The lengthy process involved exemplifies IDOE's commitment to an open dialogue regarding the development of the metrics.

IDOE's "A-F Accountability" website, http://www.doe.in.gov/improvement/accountability/f-accountability, provides valuable information including the following: (1) a basic summary, (2) PowerPoint, (3) WebEx, (4) FAQs, (5) workbook (to calculate approximate grades), and (6) sample report cards.

The response from school officials, teachers, parents and other stakeholders concerning the new metrics and accompanying tools for understanding and incorporating the models has been extremely positive.

*Indiana's updated accountability system provides a better way of measuring and reporting student achievement and will engender greater long-term educational and career outcomes for **ALL** Hoosier students.*

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

Myth v. Fact and FAQ: The Truth about Indiana's A to F School Accountability Metrics

After gathering input for nearly two years from school leaders statewide, the Indiana Department of Education (IDOE) developed new metrics to help communities understand how their schools are performing. The enhanced metrics reflect the latest advances in the way we measure student achievement and they set high expectations for student learning. These metrics were approved by the State Board of Education at its February 2012 meeting and formed a key piece of Indiana's No Child Left Behind waiver application recently approved by the U.S. Department of Education.

The metrics take advantage of new tools for measuring and interpreting student performance that were not available when Public Law 221 took effect more than 12 years ago. They strike an effective balance by looking at both proficiency and growth, the combination of which provides a better picture of school effectiveness. Schools receive credit for the percentage of students passing state assessments as well as for the progress students make during the course of the school year. Indiana's new accountability metrics give credit to schools and educators who are driving growth for all students and not just those resting at or near the proficiency bar, known as the "bubble kids." In short, the new metrics ensure that every student counts, solving a limitation of the old model.

Unfortunately, some misconceptions about the effect of the new metrics have spread throughout the state. This document attempts to provide clarity so the new metrics can be judged based on their true merit rather than on rumors and innuendo.

Myth #1: The Indiana Growth Model causes a fixed percentage of schools to "fail" each year.

This statement couldn't be further from the truth. In fact, under the new model it is possible for zero percent of schools to "fail."

Specifically, the Indiana Growth Model is a tool with many uses; one such use is as a factor in the accountability metrics. The Indiana Growth Model is not the name of the A to F metrics, but it informs the A-F metrics. The Indiana Growth Model does not cause schools to fail (or succeed). More accurately, the Growth Model is just one of several components for measuring the success of Indiana's elementary and middle schools.

Moreover, there are more opportunities for a school to improve (rather than lower) its grade with student growth. A school's final grade is raised for driving high growth in a significant number of students in the bottom 25 percent and/or the top 75 percent of students. A school's grade will only be lowered if a significant number of students in the overall student population show low growth.

Myth #2: A pre-determined number of schools will receive an A, B, C, D or F in the new model.

Wrong again. A quick study of the new metrics shows this claim to be mathematically impossible. Since several calculations factor in to the final grade at the elementary, middle and high school levels, there is no way to assign a fixed number of schools to each grade. IDOE believes every school succeeding in driving academic excellence should be rewarded. A system that sets artificial ceilings on the number of schools receiving high grades runs contrary to this belief.

Myth #3: Indiana is the only state attempting to incorporate student growth into its accountability metrics.

This claim is also false. In 2006, the U.S. Department of Education began encouraging the use of growth models for school accountability. Since then, the federal government has approved the use of growth models in accountability

systems in 15 states. Indiana is joining the national movement to utilize growth for school accountability, and the rest of the country is following suit.

Myth #4: The new metrics ensure hundreds of schools will be eligible for state takeover in the next few years.

Once again, the facts demonstrate otherwise. The accountability metrics and Indiana's school accountability timeline are two separate items. Indiana's accountability metrics consist of all the measurements used to assign grades; the accountability timeline deals with the consequences for repeatedly failing to raise student achievement. The fact is only one school may be eligible for state board intervention at the end of the 2011-12 school year and only two schools face the possibility of intervention at the end of the 2012-2013 school year.

Frequently Asked Questions

Q: When it comes to growth models, what is the difference between norm referenced (as in Colorado's growth model) and criterion referenced (as in the growth models approved by USED for AYP)?

A: "Norm referenced" and "criterion referenced" are terms that are used in describing growth models, but they also apply in other statistical models. Generally, a "norm referenced" model compares one individual to other individuals. A "criterion referenced" one compares something (here a test score) to an established standard.

The SAT exam and IQ tests are examples of norm referenced tests. They consider student performance in relation to the performance of other students who take the same test. Examples: Wes scored at the 25th percentile in Math (in other words, he performed the same or better than 25 percent of students who took the exam). Norm referencing is also widely used in medicine. "Your daughter's height is at the 80th percentile for 24-month-olds" (In other words, she is as tall as or taller than 80 percent of 24-month-olds).

Indiana's Growth Model is norm referenced like Colorado's model. Once a student has taken two consecutive ISTEP+ exams, the model can measure how much the student's performance improved compared to other students who achieved the same initial score. Example: Mike scored a 365 on his third grade Math ISTEP+ and a 475 on his fourth grade Math ISTEP+. Mike's growth is measured by looking at all other third grade students who achieved a 365 in Math. Most third graders who achieved a 365 in Math only grew by 50 points; Mike grew by 110. He is considered "high growth" under Indiana's Growth Model.

The ISTEP+ exams, the National Assessment of Educational Progress (NAEP) and the Medical Board exam are all criterion referenced tests. They set a standard to which all who take the test are held and measured. ISTEP+ exams measure students' ability to meet or surpass state established cut scores; students either pass or fail these exams.

The growth models approved for use in AYP accountability (federal accountability) are criterion referenced. Instead of measuring students' improvement year over year in comparison with other students at similar levels of academic achievement, these criterion referenced models measure students' growth according to state established standard (criterion) for growth. Example: The state sets a criterion that every student must show average growth. Sue scored a 368 in math on her state assessment in the third grade. If the average score change in the state between third and fourth grade is 30 points, Sue must score 398 in grade four to meet the criterion. A criterion referenced assessment does not consider whether gaining those 30 points is a reasonable or unreasonable target for Sue.

Indiana Department of Education
SUPPORTING STUDENT SUCCESS

Q: Why did Indiana choose a norm referenced growth model?

A: The Indiana Department of Education selected a norm referenced growth model like Colorado's after years of careful consideration. It is the fairest, most inclusive, and simplest model for Indiana's students. Indiana's model allows us to measure the change in any student's achievement in comparison to students with similar test results. A high achieving fifth grade student's growth will be measured against other high achieving fifth grade students and not against that of the average fifth grade student.

Our growth model provides a way to assess the rate at which all students are growing, and it expects and measures the growth of every student who has taken at least two consecutive tests. This powerful tool starts with individual student growth, showing parents and teachers how a student's achievement is changing over time. It adds another dimension to the picture of student achievement; we know not only how students are doing on their mastery of content standards but also how their change in achievement compares to students with a similar history of test scores.

Q: When it comes to measuring students, what is the difference between "improvement" and "growth"?

A: Before Indiana moved to the new metrics for PL 221, the state only measured simple "improvement," which meant how many more students in a school passed state assessments over a three-year period. This measure was school focused rather than student focused and fostered greater attention on the "bubble kids," often leaving others behind. The Growth Model allows us to measure the growth of every elementary and middle school student who takes ISTEP+ at least two years in a row. Incorporating this growth measurement as one component of PL 221 metrics creates an accountability system focused on individual students and schools' ability to help each of them grow. In short, now every student counts.

Q: What is "growth to standard"?

A: Growth to standard is a special case of growth that focuses on the change needed to get students to a particular **criterion** (Pass or Pass Plus) in a set amount of time (say three years). Example: In fifth grade Bill gets a scale score of 402 in Math; passing in eighth grade is a 537. Bill needs to increase his scale score by 135 points (45 points each year from sixth grade to eighth grade) to reach the standard. While this method makes sense in this simple example, no state has yet figured out a statistically valid way to measure growth to standard.

Section IV

IMPLEMENTING Common Core State Standards and Assessments

A Guide for State and District Leaders

To download the full workbook, go to www.achieve.org/ImplementationWorkbook

achieve

IEDI
U.S. EDUCATION DELIVERY INSTITUTE

IN THIS SECTION

Getting Started ____ 8.3

Vision for Accountability
That Drives College and Career Readiness ____ 8.4

Overview of Core Systems ____ 8.6

Statewide Student Performance Goals ____ 8.12

Differentiation and Classification System ____ 8.24

System of Supports and Interventions ____ 8.29

Data Reporting Systems ____ 8.29

Resources ____ 8.30

Conclusion ____ 8.30

8. Implementation Action IV: Transition Accountability and Data Reporting System

Diagnostic questions to guide your team's reading of this chapter:

- What is your state's aspiration for improving student achievement and educational attainment **outcomes** as you implement the Common Core State Standards (CCSS) and transition to the Partnership for Assessment of Readiness for College and Careers (PARCC) assessment system? How is this aspiration shared among state leaders, including the governor, policymakers, educators, parents, students and the public? How is it communicated throughout the state?
- How will your state's implementation of the CCSS and transition to PARCC change your aspiration for the state's **accountability system?** What new purposes will it need to serve to drive college and career readiness? To what extent has your state advanced planning for such a system through the Elementary and Secondary Education Act waiver process?
- What are the most critical **indicators** of student performance within such a system? How are they defined? Do state leaders, policymakers, educators, parents, students and others understand what these indicators are and how their actions have an impact on them? What other indicators are needed to fully understand students' readiness for college and career? How will the transition to the CCSS and PARCC change how these indicators are measured?
- How well is your state doing now as measured by these indicators? What is your state's target level of performance, and is it **on track** to reaching it? If not, what is it doing to **adjust course?** What kind of evidence are you using to assess progress?
- How does your state **differentiate and classify** districts and schools based on students' level of college and career readiness? How will CCSS and PARCC implementation change the measures included in this system, including measures of growth and status?
- How do these classifications connect to the intensity and kind of **support and intervention** that the state provides? How will the statewide system of support and intervention connect to your state's overall strategy to implement the CCSS and PARCC?
- How do you, state leaders, policymakers, educators, parents, students and the public find **information** about student preparation for and success in postsecondary education and the workforce, and is it easy to find, interpret and use this information to take action? How will the shift to the CCSS and PARCC change the way your state presents data to each of these groups?

Getting Started

The purpose of this chapter is to help you think through the purpose, design and implementation of your state's accountability systems as your state transitions to the Common Core State Standards (CCSS) and the Partnership for Assessment of Readiness for College and Careers (PARCC) assessment system. The new accountability systems should look significantly different from those most states developed in response to the No Child Left Behind Act. While most states centered those systems around the need to ensure that all students meet a minimal level of proficiency, these new systems will center around the need to make **ambitious but achievable progress in student performance at a much higher level linked to college and career readiness.**

The chapter is organized around the following sections:

- **Vision for Accountability that Drives College and Career Readiness;**
- **Overview of Core Systems;**
- **Statewide Student Performance Goals;**
- **Differentiation and Classification System;**
- **System of Supports and Interventions;**
- **Data Reporting Systems; and**
- **Resources.**

Feel free to move about the chapter in the way that makes the most sense given your state's position, particularly in light of Elementary and Secondary Education Act (ESEA) waivers. A critical assumption underlying the chapter, however, is that even with progress through waivers, most states will find that their accountability systems are farther behind their standards, assessments and data systems in the capacity to drive college and career readiness. For instance, if your state has already adopted statewide student performance goals but has not made significant changes to its accountability formula, you may want to go directly to the section on differentiation and classification systems. If the statewide student performance goals do not incorporate college- and career-ready indicators, or if they do not roll down to the district and school levels, that section may still be applicable to your state.

Throughout the chapter you will see illustrative examples from states as well as key questions and exercises for you and your colleagues to consider. There are also references to Deliverology 101 and links to external resources to provide further information and guidance. In particular, you should consult the Council of Chief State School Officers' ***Roadmap for Next-Generation State Accountability Systems*** for a comprehensive guide to accountability systems. The intent is to draw on what is known and what states are learning about driving performance.

Vision for Accountability That Drives College and Career Readiness

Implementation of the CCSS and PARCC assessments will require states to advance toward an accountability approach that truly drives toward college and career readiness for all students. Three major changes should result from the implementation:

- A change in accountability **systems** from those that focus on improving rates of minimal proficiency toward those that are geared toward improvement in college and career readiness;
- A change in accountability **measures** — including new measures that harness longitudinal P–20 data as well as status and growth measures from new CCSS-aligned assessments such as PARCC; and
- A change in actual **outcomes** for students resulting from heightened instructional capacity brought about through effective professional learning, instructional tools, data systems and other implementation efforts. The anticipated improvement in student outcomes should be a driving force behind states' work to set performance goals and benchmarks within accountability systems based on college and career readiness measures.

Fortunately, several factors are creating a landscape shift in which such change can occur:

- An **evolving federal-state relationship** through ESEA flexibility;
- A **growing state commitment to** advance the state-district relationship from one based on compliance and sanctions to one focused on **delivery of results for students**; and
- **Enhanced state data systems** with the ability to link data from early childhood through postsecondary education and the workforce.

These shifts have given states the opening to move toward accountability systems that serve as critical levers to improve student preparation for postsecondary success — to firmly sync accountability with the right goals for kids.

In doing so, states can move toward a far broader approach to accountability. In the past, these systems have often been defined only by an accountability formula in which assessment and other student data go into one side of an equation and accountability determinations appear on the other side. A broader definition encompasses statewide performance goals; differentiation and classification systems for districts and schools; systems of supports and interventions; and compelling systems to report data to educators, policymakers, parents and the public. It also encompasses the routines and conversations among leaders, educators and stakeholders. For example, the simple act of asking questions — what are you trying to accomplish, how are you going to accomplish it, what are your major challenges, what are you doing about them — and following up can support a more positive, goal-oriented and accountable culture across the education system. A broader definition also brings more people into the process, building engagement and understanding through transparency from development to data reporting to continuous improvement of the system. It may also encompass other forms of accountability that are out of the scope of this chapter but nevertheless may be critical parts of states' overall theory of action about improving student outcomes:

- **Student accountability,** such as requiring students to attain a certain assessment score for promotion or graduation decisions, for placement into credit-bearing courses, or for postsecondary scholarships.
- **Educator accountability,** such as tying results of educator evaluations to decisions about placement, retention, tenure and compensation.
- **Organizational accountability,** such as holding officials accountable for meeting an organization's goals or project deliverables.
- **Peer-to peer-accountability.** In contrast to more vertical or hierarchical forms of accountability, this horizontal form takes place in a culture in which colleagues hold each other accountable for performance — such as teachers within a professional learning community. (See McKinsey & Co., *How the World's Most Improved School Systems Keep Getting Better*.)
- **Shared community accountability.** In this form of accountability, actors across a community hold each other accountable for student outcomes — a central tenant of the Promise Neighborhood and STRIVE approaches. (See Education Sector, *Striving for Student Success: A Model of Shared Accountability*.)

This broader notion of accountability moves the system beyond the kind of top-down, compliance-driven, punitive approach that has had only limited returns to student performance and into the realm of a system that is continuously improving its effectiveness at building capacity to deliver results for kids.

Overview of Core Systems

The primary scope of this chapter is the purpose, design and implementation of the first two of the major core systems that are critical for a college- and career-ready accountability approach. The major core systems include:

- **Statewide student performance goals** for educational achievement and attainment that are tied to college and career readiness aspirations; developed, shared and communicated by state leaders, including the governor; translated into meaningful district and school goals; and linked to policies, practices and resources to meet the goals.
- A system to **differentiate and classify districts and schools** based on student performance outcomes tied to college and career readiness; to identify the districts and schools in greatest need of supports and interventions to build their capacity to prepare students to this level; and to suggest both the intensity and kinds of interventions that districts and schools across the spectrum should pursue to lift student performance toward and beyond readiness.
- A system of **supports and interventions** for all districts and schools, but particularly for the lowest-performing schools, that prescribes and delivers the appropriate level and type of assistance to build capacity following deeper diagnostic review and is closely linked with the state's overall system of instructional support for implementing the CCSS.
- Data reporting systems to share **actionable, meaningful data** on college and career readiness with educators, policymakers, parents and the public to predict future outcomes, identify needs to adjust course and suggest solutions.

All of these systems share similar characteristics — the measurement, analysis and reporting of data tied to college and career readiness expectations; transparency and clarity about the results and what they mean for student preparation for postsecondary pursuits; and joint aims to focus the education system on a set of priorities to achieve real progress. They also differ in important ways. For example, the intensity of stakes varies, from low to moderate (such as a tough conversation between a state and district leader) in the performance goal system to potentially very high (such as district and school turnaround and takeover) in the differentiation and classification system.

Finally, these systems interact at many points, and your state should aim to create a strong degree of coherence across them. The data reporting system, for example, will likely be focused on results from the performance goal and differentiation and classification systems. The differentiation system could be leveraged as a strategy to meet statewide performance goals (such as closing achievement gaps). Finally, the supports and interventions system can be a central delivery mechanism for meeting the statewide performance goals while drawing heavily on timely, actionable data to serve the lowest-performing schools and districts.

Purpose

States should be very clear about what they expect to achieve through each of the systems. This clarity of purpose is essential for strong design and implementation, including continuous improvement processes. For each system, you should think through the outcomes it should bring about overall and how it should be used by various actors.

EXERCISE: CLARIFY PURPOSE OF EACH CORE SYSTEM

Purpose: To identify the outcomes you expect to achieve through each accountability and data reporting system in your state, both overall and for specific actors.

Who should participate? State policy leaders; district and school leaders; teachers; students; parents; and representatives from business, community and student advocacy organizations should complete this exercise.

Directions:

1. For each system, consider the overall outcomes you expect to achieve — e.g., clarify state college and career readiness aspirations, communicate progress, meet college and career readiness achievement gap performance goals, build capacity.
2. Repeat the exercise, thinking through what key actors from each level will get out of the system — e.g., forum to communicate priorities, data to more clearly target assistance, information for resource allocation. Key actors to consider include:
    a. State — governor, chief/state education agency, state board, legislature
    b. District — superintendent, school board
    c. School — principal, teachers, students, parents
    d. Community — business and community leaders, nonprofit organizations, advocacy organizations

	Overall purpose	State purposes	District purposes	School purposes	Community purposes
Statewide performance goals					
Differentiation and classification system					
System of supports and interventions					
Data reporting systems					

Basic Design Components

The purpose of each system should directly inform decisions about design. Most of these decisions will involve three major design components: indicators, metrics and determinations.

Indicators

Indicators are the measures of a goal or aspiration. A college and career readiness system will incorporate indicators that measure course participation and success, achievement, and attainment outcomes. To incentivize and support continuous improvement, you should design the indicators in a manner that reflects a continuum of whether students are progressing toward, achieving or exceeding college and career readiness. (See Achieve and The Education Trust, *Measures that Matter*.)

This continuum of indicators allows states to accomplish the dual goals of ensuring that students identified as off track receive the supports they need to get back on track while simultaneously avoiding a situation in which the floor becomes the ceiling for students who meet the college and career readiness requirements earlier in high school. The exceeding college and career readiness indicators, such as earning college credit through Advanced Placement (AP), International Baccalaureate (IB) or dual enrollment courses while in high school, provide incentives for students and schools to strive for more. The table below suggests some indicators that states may incorporate into the accountability systems.

Possible Indicators (EXAMPLE)

	Progressing toward college and career readiness	Achieving college and career readiness	Exceeding college and career readiness
Course completion and success	• Timely credit accumulation • Credit recovery	• Successful completion of college- and career-ready course of study	• Participation in AP, IB or dual enrollment
Achievement	• Performance on CCSS-aligned assessments	• Meeting standards on the college- and career-ready statewide anchor assessment • Postsecondary remediation rates	• College-level performance on AP and/or IB exams
Attainment	• Graduation	• Earning the college- and career-ready diploma	• Earning credits in dual enrollment courses • Application to and enrollment in postsecondary

It is important to note that this table focuses on high school-oriented indicators because many states now rely only on graduation rates and assessments at the 9th or 10th grade content level, while there are many opportunities to expand the range of indicators. College and career readiness, however, is an imperative for elementary and middle school students as well. States should consider what kind of indicators beyond assessment scores they can incorporate in these grades to signal that students are on track toward college and career readiness, including early warning indicators such as attendance and course completion. (See Data Quality Campaign, ***Measuring the Education Pipeline***.)

Your state may not currently collect all of the indicators you would like to include within the system (for a profile of your state's data system, see the Data Quality Campaign's 2011 ***State Analysis***.) For example, your state may not currently have student assessment data aligned to the CCSS. While planning to enhance your data collections, you can think about valid proxies for the indicator that might substitute in the interim — do not let the perfect be the enemy of the good. For example, if your system does not collect student-level course-taking data in advanced mathematics, focus on the numbers of students participating in and earning college credit on AP/IB/ dual enrollment courses. This presents a good opportunity to make the case for why the data are critical to release; otherwise, the education community will not have the necessary data to make informed decisions to drive improved student achievement. State data systems are most valuable when they can provide timely data to key education stakeholders to make meaningful changes in the preparation of students.

EXERCISE: PLAN FOR NEW DATA INDICATORS

Purpose: To clarify new indicators that will be needed in your state's accountability systems and the process for incorporating them into your systems.

Who should participate? State education agency accountability, assessment and data staff along with stakeholders from districts, schools and community should complete this exercise.

Directions:

1. Based on your state's aspirations for students, determine the indicators needed to measure progress toward meeting this aspiration. The resource from the Data Quality Campaign (Measuring the Education Pipeline) may be helpful in this process.
2. Of these indicators, identify those that your state currently does not use for accountability and reporting purposes and list them in the left-hand column.
3. For each of these new indicators:
    a. Define the data that will be needed to construct the indicator;
    b. Determine the data owner, data source/collection methodology and verification process;
    c. Identify a proxy indicator that can be used in the interim; and
    d. Identify any data that the state no longer needs to collect because of use of this indicator.

New indicator	Data needed for indicator	Data management • Owner • Collection • Verification	Proxy indicator for interim use	Data that no longer need to be collected

Metrics

This chapter uses the term "metric" to describe a measure of an indicator that can be used to assess progress. For example, an indicator within your state's performance goal system might be the percentage of students who score proficient or advanced in 5th grade mathematics on an assessment aligned to the CCSS, where proficiency is tied to a level of performance that is on track to college and career readiness. The metric would be the comparison between the percentage of students who score proficient or advanced in 5th grade mathematics in 2012 and the expected percentage proficient or advanced in 2012 given your state's trajectory.

In differentiation and classification systems, your state will likely use a combination of status metrics and growth metrics. For example, a status metric for elementary schools might be the percentage of students proficient or advanced in grades 3–5 compared to a target level of performance for that year. A growth metric for elementary schools would use individual students' historical assessment data to estimate how much growth, on average, 4th and 5th grade students made in that year as well as the likelihood that this growth will lead to students scoring proficient or advanced within a few years. Assessing progress using status metrics to illuminate where students are currently performing is critical — an importance that increases as students get closer to high school graduation and opportunities for growth diminish. Assessing progress using growth metrics — particularly "growth to standard" metrics linked to college- and career-ready levels of performance — sends the clear signal that all students are expected to progress academically along the continuum of readiness.

Determinations

Determinations are the judgments made about performance that lead to an action. A state, for example, may determine that a school is in Priority status given its rating in the differentiation and classification system, which would trigger a specific set of interventions. Or a state may determine that a district did not meet its expected level of performance in 2012 on its performance goal trajectory, which would trigger a wider range of actions, such as a series of technical assistance sessions to reallocate resources toward strategies more likely to help the district meet the goal. As well, a state may determine that a school with one of the highest rates of improvement in the percentage of students scoring college and career ready on a statewide anchor assessment is eligible for reward status, which would trigger recognition and a financial award. States should think carefully about how they use metrics to differentiate and classify districts and schools. Decisions about how much growth is sufficient, how much weight is assigned to different metrics and the ways in which different metrics are combined send important signals to educators that have enormous implications for student instruction and support.

Implementation Considerations

For each system, your state should think through how it will approach the following implementation elements to ensure that the system meets its intended purpose.

- **Stakeholder engagement and communications:** Accountability systems cannot be designed and carried out by a small number of state education agency staff in isolation. Only by authentically engaging the full range of stakeholders in the process from development to implementation to continuous improvement can the system have the necessary transparency, clarity and understandability. Are there structures for stakeholder engagement and communications your state is using for CCSS implementation that can be used for an accountability conversation? (See *Chapter 4* of this workbook on stakeholder engagement and communication.)

- **Governance and management:** It needs to be clear what entity or person has the decisionmaking responsibility for overseeing the success of the system and what entity or person will manage it day to day. How does this fit in with the overall structure for governing CCSS implementation in your state? (See ***pages 3.5–3.8*** of this workbook on governance of CCSS.)

- **Data:** Each system will put pressure on data collections, management, quality improvement and analysis. It is necessary to plan carefully exactly how your state will access, verify and calculate data, particularly data needed for college and career readiness indicators that come from data sources outside your state's assessment system.

- **Continuous improvement process:** A clear process for state leaders to evaluate the impact of the system and make adjustments to improve its impact will ensure that the accountability system continues to get better at driving college and career readiness throughout your state. It is often tempting to try to keep accountability systems constant to preserve reliability or comparability from year to year. However, as new data sources become available, more refined growth models are developed and new PARCC assessments come on line, adapting the system quickly will be critical. Putting into place the right processes from the beginning will help ensure that this continuous improvement becomes a reality. (See ***Chapter 11*** of this workbook on routines.)

Statewide Student Performance Goals

A clear, easily understood set of college and career readiness student performance goals, along with a system to manage against them, can serve as the central driver of not only your state's accountability system but also your entire reform agenda. They can serve numerous critical purposes: They can clarify the state's aspirations and priorities; they can focus policy, practice and resources on the most effective strategies to improve college and career readiness; and they can signal the need to adjust course along the way to ensure effective implementation toward results. More broadly, they can be used by your state's leaders, including the governor, to rally support for reform; bring stakeholders together for a common purpose; and communicate that what matters the most is real, measurable improvement in student outcomes. (See National Governors Association, ***Setting Statewide College- and Career-Ready Goals***.) They can be a guiding force to ensure that your state's implementation strategy for the CCSS and PARCC assessments brings about real results for students.

In the last few years, more and more states have begun to set these goals and use them to guide their college- and career-ready reforms. The eight states that participated in the College- and Career-Ready Policy Institute (CCRPI), a multistate collaborative sponsored by the Bill & Melinda Gates Foundation to support states in developing college and career readiness policy reform plans, began by developing performance goals. (See box on next page.)

Still more states developed them in response to Section A of the Race to the Top (RTTT) grant application. Today, states that have been applying for waivers under the U.S. Department of Education's ESEA Flexibility program have crafted them at the state, district and school levels under the Annual Measureable Objectives (AMO) requirements in Principle 2. As states develop and adjust their goals over time, they can consider a range of design components — aspirations; indicators; baselines, targets and trajectories; district- and school-level targets and trajectories; and routines to monitor and drive programs.

Louisiana Goals Developed through the College- and Career-Ready Policy Institute (EXAMPLE)

	Goal	Measure	2005–06 baseline	2009–10 target	2013–14 target
1	Reduce dropouts and increase high school graduation rates	Four-year cohort graduation rate[1]	64.8	67.0	80.0
2	Increase readiness for postsecondary education	% of students graduating with LA Core-4[2]	58.5	62.5	72.5
		% of graduating class with ACT score of 18 or higher in English and 19 or higher in math[3]	46.1	51.1	58.1
3	Increase career readiness of students	# of National Career Readiness Certificates (WorkKeys Platinum, Gold, Silver or Bronze)	2,652	4,000	7,000
		# of industry-based certifications earned by high school students in high-skill, high-wage and high-demand occupations as approved by the Louisiana Workforce Investment Council and BESE	3,600	7,500	10,000
4	Increase participation and completion rate in postsecondary education	% of public school 11th graders enrolling in an LA public postsecondary institution within four years (includes dual enrollment)[4]	51.4	54.4	63.4
		# of high school graduates enrolling in a technical college or two-year LA public postsecondary institution within two years of graduation	*	*	*
		# of public postsecondary[5] degrees and certificates awarded (one-year certificate, associate, bachelor's or higher)	32,416 (2007–08)	35,500	41,000
		# of public postsecondary degrees and certificates awarded (one-year certificate, associate, bachelor's or higher) in high-skill, high-wage and high-demand occupations as defined by the Louisiana Occupational Forecasting Conference	*	*	*
		# of credit hours enrolled in public postsecondary institutions by LA public high school students	*	*	*

*Historical data are currently being researched by the Board of Regents and Department of Education to determine the baseline and set targets.

[1] The percentage of students who entered the 9th grade and graduated four years later. Students who transfer from the LA public education system are not counted in this rate

[2] Baseline for this measure is TOPS Core.

[3] Baseline and targets provided by LA Board of Regents

[4] Baseline provided by LA Board of Regents using LDE 2002-03 grade 11 data file.

[5] Baseline and target provided by LA Board of Regents

Source: http://www.louisianaschools.net/lde/uploads/15403.pdf

Aspirations

A state's student performance goals should strongly reflect the state's aspirations, both for its students and for the future of the state as a whole. Often, these aspirations reflect state goals for economic growth, improvements in quality of life for its citizens, or easing inequality based on race/ethnicity and income. These aspirations often inform a state's major education policy agenda and are often referenced in speeches and other communications by the governor, state chief, higher education officials, and business and community leaders. Articulating them can not only focus the system in a clear direction but also bring stakeholders together toward a common purpose. (See Chapter 1A of Deliverology 101 on defining the aspiration.)

Indicators

Indicators are selected to best **measure** progress toward the state's aspirations. If the state has a goal to improve the quality of its science, technology, engineering and math (STEM) workforce, for example, the state could select a variety of indicators that would show if progress toward the goal is being made. As mentioned earlier, states should select several indicators for course completion and success, achievement, and attainment along a continuum of progressing toward, meeting and exceeding college and career readiness standards. For example, a state could select the percentage of all students meeting an on track to college- and career-ready benchmark in 8th grade math (achievement/progressing toward), the percentage of high school students completing four years of math (course completion and success/meeting), and the percentage of students who earn postsecondary credit in STEM courses while in high school (attainment/exceeding). Some states may wish to expand the range of actors engaged in this goal by broadening the continuum of college and career readiness to encompass more cradle-to-career indicators, such as the percentage of entering kindergarten students scoring at a certain level on a math readiness assessment and the number of students completing STEM degrees in postsecondary institutions.

Your state's priority indicators will change over time, particularly your mathematics and English language arts achievement indicators, as your state transitions to the CCSS and PARCC assessments. While recognizing this and communicating that indicators will shift over time is critical, selecting the best indicators your state has available now is imperative to maximize student progress through the transition process.

Student characteristics and program participation: All indicators should be disaggregated by student race/ethnicity, economic status, disability status and English language learner status. This disaggregation is critical for public reporting and to ensure that educators have the information they need to make clear ties between performance and systemic issues with instruction and support that affect particular subgroups. Your state may also have additional ways of grouping students, such as states that have "lowest quartile" types of subgroups following the ESEA waiver process.

Indicators defined as percentages: Many indicators will be rates with numerators and denominators. For college- and career-ready indicators, particularly at the meeting and exceeding areas of the continuum, being very careful about selecting the denominator is imperative. Denominators that include a select group of students can send an inaccurate picture of performance in a school. For example, the rate of students in a graduating cohort with scores of 3 or higher on an AP exam will appear far higher if the denominator is only those students who took an AP exam. It is advisable in this circumstance to include a denominator that includes all students in the 9th grade cohort as the denominator.

Indicators defined as numbers of students: In some cases, states may wish to define some indicators not only as percentages but also in numbers of students. This is particularly important in communicating the urgency of meeting the state's aspiration and in sending a clear message about what it will take to meet the targets. It can also clarify the process of building the trajectory. (See next section.)

Leading indicators: Some states may select not only summative, or lagging, performance indicators but also a set of leading indicators to give early signals about the direction of progress. Ideally, leading indicators would be able to be measured more than once a year and would send clear early warning signals to those who can take action to improve performance. They are critical components of your state's routines to monitor performance but can also serve as accountability metrics in their own right. Such an approach has been seen mostly at the district level, such as the high school freshmen on track measure that is reported by Chicago Public Schools to help meet graduation rate goals. (See *High School Freshmen On Track Rates by School* and *High School Freshmen On Track Rate Fact Sheet*.)

Leading Indicators (EXAMPLE)

Performance Indicator: The percentage of high school students graduating high school in four years having completed a college- and career-ready curriculum. The state's diploma requirements include:

- Successful completion of an Algebra II course and four years of math;
- Successful completion of four years of grade-level English; and
- Successful completion of three years of science, including a biology, chemistry and physics.

If your indicator is the percentage of high school students graduating having completed a college- and career-ready curriculum, **leading indicators** might include:

- The percentage of students enrolling in Algebra II by 11th grade, Geometry by 10th and Algebra I by 9th grade;
- The percentage of students who are enrolled in Algebra I in 9th grade with a grade of C or higher at the first grading period and an attendance rate above 89 percent; and
- The percentage of students completing Algebra II by 11th grade, Geometry by 10th and Algebra I by 9th grade.

Baselines, Targets and Trajectories

Once your state has selected a set of indicators, the next step is to understand where performance is now (baseline) and where it has been (historical progress), decide where your state wants student performance to be and by when (target), and clarify the expected path between the baseline and the target (trajectory). The targets and trajectories are used to gauge progress along the way. If your state has already completed this work through its waiver application, RTTT application or otherwise, you may wish to skip ahead or read through this section to see how your goals compare to the principles outlined here. (See Chapter 3B of Deliverology 101 for more in-depth treatment of these steps.)

Baseline: The baseline level of performance can be thought of in two ways. A simple way is to think about it as the current level of performance, usually as indicated by the most recent year's results. A more complex way is to think about it as the expected amount of future progress given the amount of progress over the previous several years. The second approach is preferable because it gives more information to help set an ambitious but achievable target.

In pursuing the second approach, your state should consider what the data patterns have been for the indicator historically, whether growth/decline has been predictable or nonlinear over time, and where the level of performance would be over time absent any interventions. As well, examine the factors that contributed to the metric's historical growth or decline.

It may be necessary to disaggregate the data by districts and schools and roll up these data in different ways to better understand what it will take to achieve your target at the state level. The more detail that can be accounted for early on in the process, the more accurate the estimations will be. You may also find that examining the data through different subgroup lenses allows for patterns to emerge that would otherwise be masked at an aggregate level. This might include breaking out the data by:

- **Individual comparisons** (e.g., students, teachers, schools, districts);
- **Characteristic** (e.g., special education, low income, English language learners, rural and urban schools, etc.); or
- **Performance band** (e.g., how do the top and bottom school quintiles compare along a metric, how do the top and bottom district quintiles compare?). (See pg. 98–99 for more information.)

Notice that in the example below, the state has chosen to use college- and career-ready diploma trend data from a performance band — its highest and lowest performing quintiles of schools — to better inform state target-setting.

Using Historical Data To Establish the Baseline (EXAMPLE)

	2007	2008	2009	2010	Average yearly percentage point change
Percentage of high school students earning college- and career-ready diploma in MeasuredState	32%	36%	37%	40%	+3 annually
Schools in top quintile	45%	50%	55%	65%	+5 annually
Schools in bottom quintile	24%	25%	25%	26%	+1 annually

Equipped with this knowledge of historical data trends, your state is better positioned to set realistic and ambitious targets.

Target: The target level of performance is the number that defines your state's aspiration — it is your state communicating to all of the actors in the education system that student performance not only will improve but also will improve to this specific number by this specific time. It enhances the sense of urgency to reform, brings people together with a common cause, and shows that the state is serious about implementing the CCSS in a way that creates real change and real impact on students.

Targets should be set according to the SMART criteria: Specific, Measurable, Ambitious, Realistic and Time Bound. It is critical that the targets are ambitious enough to stretch and motivate actors throughout the education system but achievable enough to be legitimate and meaningful to all. The language used in RTTT and the ESEA Flexibility requirements has been "ambitious but achievable."

The target-setting process can be carried out in a variety of ways depending on your state's context. The following represent three possible ways to set targets, and your state could choose to use any combination of these three:

- Your state may **have previously adopted targets** on one or more indicators through a strategic planning process, a grant application, its K–12 ESEA accountability system or a higher education accountability system. If so, your state can adopt these targets and use them to inform the target level of performance for other related indicators that you have selected.

- Your state may wish to **set targets based on aspirations** of state leaders, educators, parents and other stakeholders. If so, a process can be established in which you share baseline data with the group and have participants walk through guiding questions to reach targets that they can agree are ambitious but achievable.

- A third approach is to **use a data-rich benchmarking process** to determine the appropriate targets. There are four benchmarking techniques that can be helpful in target-setting, each with a set of questions that can be used to guide the process:

  - Historical comparisons: What have the data trend lines been at the state level, by subgroups of students, type of district, quartiles of schools, etc.? What would the impact be on the indicator if a group of students or schools grew by a particular percentage over the historical average?
  - Internal peer comparisons: What have been the data trend lines been at the state level, by subgroups of students, type of district, quartiles of schools, etc.? What would the impact be on the indicator if a group of students or schools achieved at the level of a higher-achieving group?
  - External peer comparisons: How do our data trend lines at the state level compare to other states, by subgroups of students, type of district, quartiles of schools, etc.? What would the impact be on the indicator if a group of students or schools achieved at the level of a neighbor state? For example, a state looking to estimate the target and trajectory for the percentage of students earning a college- and career-ready diploma could look to the experience of a top-performing state that has seen growth in the percentages of its students earning a college- and career-ready diploma in the past.
  - International comparisons: How do our data trend lines at the state level compare to other countries, by subgroups of students, type of district, quartiles of schools, etc.? What would the impact be on the indicator if a group of students or schools achieved at the level of a country of interest?

It is critical for your state to think about how it will establish baselines and targets once new assessment indicators are available under the PARCC assessments. For example, the state may be more likely to use internal peer or external peer comparisons to set targets at this point rather than historical comparisons. It may even be able to adjust targets based on international comparisons if that kind of information is available.

Subgroup-level targets: Your state should consider how to leverage targets to clarify state goals and drive strategy to close achievement gaps. There are several options for how your state could do this. Some states have decided to have the same target for all student subgroups, others have decided to apply the rate of improvement called for by the overall target to each subgroup so that each group has to make the same rate of progress, and others have had different targets for different subgroups but called on greater rates of progress for groups of students that start out farther behind (see example below). As you consider the options, it would be strongly advisable to consult with stakeholders about setting targets for individual groups of students with varying baseline levels of performance and historical rates of progress. Having this engagement from the beginning of the conversation will ensure that your state makes this difficult and complex decision with the support and guidance of those with the most insight into the drivers of performance for these groups of students.

Massachusetts Graduation Rate and MassCore Completion Baselines and Targets by Subgroup — Greater Improvement Expected from Subgroups with Lower Baselines (EXAMPLE)

MassCore Completion Rate, sorted by 2011 percent

	2011		2014		Percentage point change	Percent change
	Number	Percent	Number	Percent		
Asian	2,140	67.9	2,801	82.7	14.8	30.9
Black/African American	2,298	46.9	3,811	73.2	26.3	65.8
Hispanic	3,462	51.5	5,323	75.9	24.4	53.8
White	34,840	74.9	42,264	84.3	9.4	21.3
Low income	9,076	53.1	13,185	76.6	23.5	45.3
Students with disabilities	4,178	60.0	5,767	78.6	18.6	38.0
Limited English proficient	409	30.1	913	66.5	36.4	123.2
Female	22,459	71.0	28,145	82.9	11.9	25.3
Male	21,082	68.3	27,097	82.0	13.7	28.5
Overall	**43,541**	**69.6**	**55,241**	**82.5**	**12.9**	**26.9**

Five-Year Graduation Rate, sorted by 2010 percent

	2010		2014		Percentage point change	Percent change
	Number	Percent	Number	Percent		
Asian	3,216	89.4	3,294	92.4	3.0	2.4
Black/African American	5,110	73.5	5,445	79.9	6.4	6.6
Hispanic	6,898	65.9	7,390	72.5	6.6	7.1
White	48,169	89.5	49,836	92.1	2.6	3.5
Low income	21,257	72.3	22,668	77.9	5.6	6.6
Students with disabilities	10,091	68.6	11,173	76.3	7.7	10.7
Limited English proficient	2,755	63.7	3,475	71.4	7.7	26.2
Female	32,554	87.0	33,542	90.3	3.3	3.0
Male	32,049	82.4	33,585	86.4	4.0	4.8
Overall	**64,603**	**84.7**	**67,127**	**88.3**	**3.6**	**3.9**

Source: http://www.doe.mass.edu/boe/docs/0311/spec_item1_deliveryplan.pdf

Trajectories: The trajectory is the path that you expect performance will take from the current level of performance to the target. It is often made up of annual benchmarks from the current year to the target year. The trajectory is essential for monitoring progress and deciding where mid-course corrections need to be made. (See Chapter 3B in Deliverology 101.) For example, if one indicator is showing enough improvement to meet or exceed the trajectory, and another indicator has not shown enough improvement to meet the trajectory, the state will know that it needs to adjust its strategies for improvement on the second indicator.

Like baselines and targets, your state can develop trajectories in several different ways:

- **Develop a linear trajectory,** which is a straight line between the current level of performance and the target. The linear trajectory calls on steady, incremental progress each year.
- **Set annual benchmarks** based on what you know about historical progress following interventions. For example, if your state has just implemented a new assessment system with much lower scores in the first year, you may know that in the second and third years you can expect to see more improvement in scores than you will in years four and beyond.
- **Estimate the impact of future strategies and interventions** (see pages 5.4–5.5 and 6.5–6.6 of this workbook) using impact data from similar past interventions to help estimate the trajectory toward the target. It is critical in this case that states avoid trajectories that push the bulk of expected performance improvement toward the later years.

Establishing the Baseline (EXAMPLE)

If your state would like to estimate the trajectory by tying it to the strategies and interventions it is planning to implement to meet the target, the first step is to establish the baseline.

In this example, the baseline is 40 percent. With an average growth of 3 percentage points annually, we can expect the percentage of students earning a college- and career-ready diploma in 2018 to be 64 percent without any additional interventions. The state has set a target of graduating 80 percent of students with a college- and career-ready diploma by 2018, so there is a 16 percentage point gap to close.

	Baseline	2011	2012	2013	2014	2015	2016	2017	2018
Percentage of high school students earning college- and career-ready diploma	40%	43%	46%	49%	52%	55%	58%	61%	64%

As a next step, states can ask themselves some key questions about their strategies or interventions:

- What has the impact been?
- On whom is/was the impact greatest?
- Are the outcomes in line with what was intended?
- What actions should be taken upon reviewing these impact data?
- What works and does not work?
- Where does it work, and where might a better intervention be employed?

States can also think through how their strategies will have an impact over time. Some will have a lagged effect, while others have a more immediate impact on your trajectory. For example, if your state was looking to develop a 12th year bridge course in mathematics, it would be reasonable to expect that there would be an immediate impact in the number of students scoring college- and career-ready on a state assessment. However, you will find that some interventions take additional time to phase in and produce results.

States can also consider the differentiated effects of a strategy by individual, characteristic or performance band. Once you have a comprehensive understanding of your state's historical data patterns and interventions and their potential to help reach your goals, you can better assess whether they are the best use of your limited resources or whether your state might be better served reallocating resources or focusing on a specific subgroup of students.

The state would close the gap by estimating the impact of each of its strategies on the level of performance each year. For example, the following table includes estimates of the impact (none/low/medium/high) of each intervention on the level of performance.

You can also assign values in numbers of students who will be affected by the interventions, as Kentucky has done with its trajectory toward its college and career readiness target (see example).

Impact of Each Intervention on the Level of Performance (EXAMPLE)

Intervention	Baseline	2011	2012	2013	2014	2015	2016	2017	2018
12th grade bridge courses	-	-	L	L	M	M	M	L	L
New math and science teachers	-	-	M	H	H	M	M	L	L
9th grade dropout prevention program	-	-	None	None	None	M	M	M	M
Total	-	-	L	M	M	M	M	L	L

Impact of Each Intervention on the Level of Performance (EXAMPLE)

Kentucky currently has a graduation rate of 76 percent; however, only 34 percent of students are college and/or career ready as measured by the ACT benchmarks. To support the efforts of districts in ensuring that more graduates obtain 21st-century skills and are prepared to attend college, the Kentucky Department of Education (KDE) is implementing several strategies focused on increasing the college and career readiness rate.

This example provides an estimated number of additional students needed each year for each priority strategy for Kentucky to reach its goal of 67 percent of students college and/or career ready by 2015. If Kentucky continues business as usual, only 39 percent of students will be college and/or career ready by 2015.

The trajectory was created with a focus on the current 8th grade cohort (class of 2015) and their exposure/participation in KDE's specific initiatives for increasing college and career readiness. The current 8th grade cohort has approximately 48,000 students, of which 16,320 (48,000 * 0.34) are college and/or career ready. To meet the 2015 goal, Kentucky must have an additional 15,840 students college and career ready. The trajectory also assumes student gains will be maintained year over year (e.g., student impact from 2011will be sustained in subsequent classes).

Finally, the major impact on students will be in years 2012–13 and 2013–14 due to the implementation of the new accountability model and the cumulative effect of all eight strategies.

College and Career Readiness Trajectory — Additional Number of Students

Strategy	2010–11	2011–12	2012–13	2013–14	2014–15	Total
Target intervention	950	1,488	2,880	3,840	2,376	11,534
Acceleration	1,358	1,463	1,628	1,888	1,974	8,309
Course and assessment alignment	0	298	480	768	1,267	2,813
New accountability model	0	0	2,400	384	317	3,101
Innovative pathways/student success	0	0	34	30	30	93
Academic/career advising	0	149	480	1,536	1,901	4,066
District 180/turnaround	66	150	380	488	573	1,658
Career readiness definition	0	0	32	51	37	119
# of additional college- and career-ready students	2,374	3,548	8,313	8,984	8,475	31,693
# of unique college- and career-ready students *	1,187	1,774	4,157	4,492	4,237	15,847
Total # of college- and career-ready students	17,507	19,281	23,437	27,929	32,167	
% of students college and career ready	36%	40%	49%	58%	67%	

* Due to the duplicated student participation in the above eight strategies, the total number of additional college- and career-ready students has been reduced by 50 percent.

Source: http://www.education.ky.gov/NR/rdonlyres/8FD9B030-078E-4874-AC12-C69A2B100762/0/CCRRateTrajectoryDetail.pdf

District- and school-level targets and trajectories: States will not meet their student performance goals through actions taken only at the state level; rather, actions taken in districts and schools will far more directly affect student readiness for college and career. Taking state-level targets and trajectories to the district and school levels will go far in establishing the performance-driven relationships needed for truly transformational change. States may have set these through their AMO process, but if this process was geared toward setting minimal threshold targets, you may wish to consider setting stretch targets — those that require districts and schools to make greater progress.

Generally, there are three different approaches to managing the work of establishing district- and school-level targets and trajectories:

- A state might employ a **top-down** approach by mandating targets and trajectories for districts and schools.
- A state might employ a **bottom-up** approach in which districts and/or schools develop their own targets and trajectories and submit them to the state for consideration/approval.
- A **hybrid** approach features discussions between the state and districts/schools in an effort to build buy-in and understanding around the trajectories and targets, or a state might suggest targets and trajectories to districts and/or schools, which can then accept them, raise them or lower them based on their own goals.

Your state may decide to draw on these approaches in varying ways, depending on the indicator or on the state-district context. Regardless of the approach, your state should be clear about the connection between the district and school targets and the state's overall targets. For example, you should aim that if all districts meet their targets, the state as a whole will meet or exceed its target and that a district will meet or exceed its target if all of its schools meet their targets. This is not only mathematically necessary, but it reinforces the delivery relationship — how each district and school plays a role in the state's overall progress and how critical it is to the state that it support districts and schools in meeting their targets.

Furthermore, states should be thoughtful about the kind of improvement the targets would expect of districts and schools at various performance levels. For example, do you want districts and schools in the bottom half of performance to make tremendous improvement, and that this will lead to meeting the state goal? Or do you want all districts and schools to make improvement, and if so, do you want those that have lower performance to begin with to make more rapid progress?

Routines To Monitor and Adjust

While state performance goals can be powerful levers in their own right by communicating the state's aspirations and clarifying priorities, perhaps their greatest power arises from regular reviews of progress among those responsible for effectively executing the strategies to meet the goals. (See Deliverology 101, Chapter 4B.) For example, a state can use its trajectory work to frame a conversation about the impact of various strategies and interventions. If the state is not meeting its trajectory, it can dig into the impact data on strategies to determine where it needs to adjust course to have the expected impact. These routines can serve a role similar to the determinations in the differentiation and classification system — making clear where the state, districts and schools are against the targets and trajectories to which they have committed.

Some states are beginning to use their performance goals to frame routines between state chiefs, senior members of the state education agency and project managers that illuminate progress on key indicators and leading indicators, discuss the status of implementation efforts, and determine next steps to improve strategy and implementation.

Some states are also using the goals to frame conversations between state leaders and district leaders to reinforce the importance of the goals, clarify the key strategies for reaching them, decide whether adjustments need to be made given the rate of progress against expectations, determine needs for assistance and identify effective practices that can be shared more widely. (See example below.)

Rhode Island's Gap Analysis Meetings (EXAMPLE)

For the first part of the local education agency (LEA) RTTT Gap Analysis process, the Rhode Island Department of Education will populate LEA data into a Student Performance Outcomes Gap Analysis Template. An example from this template is shown below. These data will be used during the Gap Analysis meetings with the LEA leadership teams to monitor progress toward the student performance goals set by each LEA and to reinforce the frame of the RTTT work in the context of improving student academic outcomes.

District Goals and performance measures	2009 actual	2010 actual	2011 goals	Difference 2009-10	Progress 2009-10	Gap between 2010 actual and 2011 goal
Students entering the **4th grade** will be proficient in **reading** on NECAP	75%	72%	78%	-3%	▼	-6%
Students entering the **4th grade** will be proficient in **mathematics** on NECAP	64%	63%	66%	-1%	▼	-3%
Students entering the **8th grade** will be proficient in **reading** on NECAP	78%	79%	80%	1%	▲	-1%
Students entering the **8th grade** will be proficient in **mathematics** on NECAP	59%	52%	60%	-7%	▼	-8%
85% of students who first entered 9th grade four years prior will graduate from high school	80%	81%	81%	1%	▲	0%
77% of students who graduate from high school will enroll in an institution of higher education within 16 months of receiving a diploma	75%	75%	76%	0%	◀▶	-1%
90% of students who enroll in an institution of higher education will complete at least one year's worth of credit within two years of enrollment	80%	88%	82%	8%	▲	6%

Source: http://www.ride.ri.gov/commissioner/RaceToTheTop/docs/RIRTTT_Performance%20Monitoring%20Plan_Revised6.24.11.pdf

Differentiation and Classification System

Although your state performance goals and system to manage progress against them can go a long way in differentiating districts and schools based on the variation in student performance outcomes and in sending some real signals about the level of progress that needs to occur as well as areas that are progressing better than others, they likely don't have the ability to more finely distinguish overall school and district performance. To credibly evaluate the performance of institutions such as schools and districts at moving students toward and beyond college and career readiness, your state needs a more sophisticated set of metrics and more precise rules for determinations. It also must comply with federal requirements for accountability systems set through ESEA or, for states granted waivers, the requirements set through ESEA Flexibility.

The system to differentiate and classify schools and districts serves a number of purposes:

- It **identifies the districts and schools** in greatest need of state intervention and support;
- It **can rally community support** to improve districts and schools that are not making sufficient progress;
- It **can serve as a strategy for reaching state performance goals** such as those to close achievement gaps; and
- It **can provide strong incentives** for districts and schools to improve college and career readiness and receive rewards and recognition.

As your state thinks through the primary aims for its differentiation and classification system — given federal law, state policy and implementation context — and state priority goals, recognizing that the new system will be focused on driving ambitious but achievable progress in student college and career readiness is critical. The system will not be focused on improving rates of students meeting a minimal level of proficiency. The implications of this shift should be considered thoughtfully in the design and implementation of the system.

Indicators

The indicators chosen for the differentiation and classification system, like those for state performance goals, should reflect student performance in terms of course completion and success, attainment, and achievement. They should also flow along a continuum of progressing toward, meeting and exceeding college and career readiness. The indicators will be different — and there likely will be more of them — than those selected for use through the statewide performance goal system. For instance, states will have to include reading and math proficiency results in the 3rd–8th grades and once in high school in the differentiation and classification system, while they might have chosen to focus on reading in 4th grade and math in 8th grade for the performance goal system.

The indicators for the system will also be different from those in prior differentiation and classification systems. Achievement indicators will transition from state assessment results indicating minimal proficiency to state-developed and PARCC assessment results tied to college and career readiness. For example, some states have already implemented more rigorous academic standards and cut scores that reflect higher expectations for proficiency. Even as states work toward implementation of the PARCC assessments, some are beginning to incorporate the CCSS into current assessment systems. All states' achievement indicators will reflect college- and career-ready achievement once the consortia assessments are implemented in 2014–15.

States have traditionally included only one attainment measure in these systems — the high school graduation rate. College- and career-ready systems, however, will include a broader mix of attainment measures, including those showing that students have met college and career readiness expectations, such as graduation with a college- and career-ready diploma, and those showing exceeding college and career readiness, such as earning credit through dual enrollment.

As well, systems rooted in college and career readiness will include indicators of course completion and success, including those showing progress toward readiness, such as timely credit accumulation along a college- and career-ready course sequence, and those showing that students have exceeded readiness standards, such as completion of AP and IB courses.

Given that this system has potentially very high stakes for districts and schools, it is essential that all indicators maintain the highest standards for data quality and transparency. The incorporation of new college- and career-ready indicators for attainment and course completion and success, in particular, will likely require strong implementation planning to ensure high levels of data quality and to ensure that educators understand how indicators are defined and calculated. Partnering with data managers from higher education for indicators of dual enrollment, postsecondary enrollment and success, as well as providers of assessments such as AP and IB, will also be critical to obtain and use the right data.

Metrics

The kind of metrics used in the differentiation and classification system will be defined very precisely, encouraging schools to help more students reach an absolute level of college and career readiness as well as encouraging them to help all students make individual progress toward and beyond readiness. The system will include both status metrics, such as requiring a school's percentage of students proficient or advanced in English language arts to meet a school-specific or statewide benchmark (e.g., AMOs) or assigning points based on the percentage in an index system, and growth metrics, such as requiring that on average, students make enough progress from previous years to be on track to a performance level tied to college and career readiness within three years.

States are taking many different approaches to growth metrics within their differentiation and classification systems. Within a college and career readiness accountability system, it is absolutely critical that growth be incorporated in a way that drives all students forward to a clear standard of performance tied to college and career readiness. It is also important that the growth metrics provide more detailed information about how much growth individual students are making and their individual likelihood of reaching important college and career readiness measures. Given that PARCC states will transition their assessment system in 2014–15 as PARCC becomes operational, it is important to start planning now for how growth measures will be calculated and incorporated over this time. PARCC is planning to commission several research papers in coming months that will assist states in this planning process.

Determinations

Your state will then need clear and transparent business rules on how to combine the metrics to arrive at an ultimate determination of the district's or school's classification. Your state can choose to do so using a **conjunctive** approach, in which the school or district must meet acceptable performance on most or all metrics. Adequate yearly progress (AYP) is an example of a conjunctive approach. It could also do so using a **compensatory** approach, such as an index, in which all metrics are assigned points based on the level of performance.

An index presents numerous benefits because it is very transparent and the values assigned to different levels of performance can easily reflect the priority goals of the state. It needs to be constructed carefully, however, because higher performance in one area — such as AP course completion — could compensate for lower performance in another area — such as high school graduation rates.

The conversation about weights to assign to different metrics has critical implications for the system's results and therefore the kinds of incentives it communicates to educators. States should determine, for example, how much weight to assign to status metrics versus growth metrics. States should provide enough weight on college and career readiness metrics to communicate the central importance of students reaching a level of performance tied to college and career readiness. (See example below.) This could become a more complex task once assessments transition to measuring college and career readiness against the CCSS in mathematics and English language arts, but not in science and social studies. It will be important for states to ensure that the aims of differentiation based on college- and career-ready level of performance are not dampened because of other assessments.

Florida's High School Accountability Weights (EXAMPLE)

Reading	Mathematics	Writing	Science	Acceleration	Graduation rate	College readiness
Performance (100) 6.25%	Performance (100) 6.25%	Performance (100) 6.25%	Performance (100) 6.25%	Participation (175) 10.94%	Overall (200) 12.5%	Reading (100) 6.25%
Learning Gains (100) 6.25%	Learning Gains (100) 6.25%			Performance (125) 7.18%	At-Risk (100) 6.25%	Math (100) 6.25%
Lowest-Performing 25% Gains (100) 6.25%	Lowest-Performing 25% Gains (100) 6.25%					
300 Points 18.75%	300 Points 18.75%	100 Points 6.25 %	100 Points 6.25 %	300 Points 18.75%	300 Points 18.75%	200 Points 12.5%

Source: http://www.fldoe.org/pdf/FloridaProposal1-31-12.pdf

Classifications

States send strong signals through the classifications they bestow on schools and districts. The strongest signals are typically reserved for the schools and districts that the metrics and determinations indicate are the lowest performing in the state. For example, though ESEA Flexibility, states commit to identifying the bottom 5 percent of Title I schools as Priority. Another 10 percent of Title I schools, such as those with the highest achievement gaps, are deemed Focus schools. The top 5 percent of schools, meanwhile, are classified as Reward schools.

To inform their system of supports and interventions and help clarify parent and public understanding of school and district performance, states should also develop classifications and determination rules for the other 80 percent of schools (as well as all districts). When the central driver is college and career readiness, all districts and schools will need support to continually improve student outcomes, particularly for students in greatest need. In addition, a continuum of classifications heightens the degree of incentives in the system — all but the very top schools and districts can aim to reach higher classifications and be rewarded and recognized for doing so.

Equity Aims

It is likely that one of the most critical purposes of your state's differentiation and classification system is to illuminate and incentivize districts and schools to address gaps in achievement, attainment, and course completion and success among groups of students based on race/ethnicity, income, disability status and English language learner status. Two broad principles are at play in meeting this intended purpose.

First, your system should ensure that student groups that start out farthest behind make the most progress toward meeting college and career readiness standards. For status metrics, you may decide to use different AMOs for different subgroups with a higher slope for lower-performing subgroups, and the ultimate target could be the same or result in substantial gap-closing. (See example on next page.) For growth metrics, you may decide to hold lower-performing subgroups to higher growth scores to close gaps. Again, as in the setting of targets and trajectories for your student performance goals (which may correspond with your AMOs), it is advisable to deeply involve stakeholders in making these decisions and monitoring their impact.

Indiana's Statewide AMO for the Hispanic Subgroup (EXAMPLE)

School year	Benchmark	Benchmark goal	Annual state assessment proficiency goal	Pass % English language arts (ELA)	Pass % math	Annual college and career readiness rate goal	College and career ready %	Annual graduation rate goal	Graduation rate %
2011–12	Baseline			68%	70%		11%		76%
2012–13			Increase by 4 percentage points in ELA and math	72%	74%	Increase by 3 percentage points	14%	Increase by 1 percentage point	77%
2013–14			Increase by 4 percentage points in ELA and math	76%	78%	Increase by 3 percentage points	17%	Increase by 2 percentage points	79%
2014–15	Three-Year Benchmark	Achieve and 'A' or improve by one letter grade from the 2012 baseline	Increase by 4 percentage points in ELA and math	80%	82%	Increase by 3 percentage points	20%	Increase by 2 percentage points	81%
2015–16			Increase by 2 percentage points in ELA and math	82%	84%	Increase by 1 percentage point	21%	Increase by 2 percentage points	82%
2016–17			Increase by 2 percentage points in ELA and math	84%	86%	Increase by 1 percentage point	22%	Increase by 2 percentage points	84%
2017–18			Increase by 2 percentage points in ELA and math	86%	88%	Increase by 2 percentage points	24%	Increase by 2 percentage points	86%
2018–19			Increase by 2 percentage points in ELA and math	88%	90%	Increase by 2 percentage points	26%	Increase by 2 percentage points	88%
2019–20			Increase by 2 percentage points in ELA and maintain 90% and continue to improve in math	90%	92%	Maintain 25% and continue to improve	28%	Increase by 1 percentage point	90%

Source: http://www2.ed.gov/policy/eseaflex/approved-requests/in.pdf

Second, your system should be clear about how equity goals balance with overall student improvement goals in determinations and classifications. For states with systems that closely follow AYP, this balance is very clear — any subgroup that does not make AYP means that the school or district as a whole does not make AYP. For states with other types of systems, it is critical to think carefully to ensure that determinations are made in a way that singles out the most disadvantaged groups for focus and signals the types and intensity of needed interventions.

System of Supports and Interventions

With a shift to college and career readiness accountability systems, the statewide system of support and intervention also needs to shift its direction. It can no longer focus only on helping the state's lowest-performing schools and districts bring students up to minimal proficiency standards. Now, it needs to focus on connecting the full range of districts and schools with the right intensity and kind of interventions to move all students toward and beyond college and career readiness.

The system must be tightly integrated with your state's overall plan to implement the CCSS and PARCC assessments. How will diagnostic reviews uncover school and district capacity to help teachers implement the CCSS in their classrooms? How will schools and districts at various classifications participate in state and regional professional learning networks? How will technical assistance providers help schools and districts at various classifications use instructional tools and supports? While this chapter does not address the system of supports and interventions in detail, it is critical that your state think through these questions in its planning for CCSS implementation.

Data Reporting Systems

Data reporting is a critical piece of a broader approach to accountability tied to college and career readiness. Rather than accountability as a punitive system, strong presentation of data can support a culture of data use that drives continuous improvement and engagement at all levels. (See Data Quality Campaign, ***The Next Step: Using Longitudinal Data to Improve Student Outcomes***.) One of the most powerful levers to improve student performance is to simply report clear and meaningful data in such a way that they are used and understood by those who influence student achievement and attainment outcomes. A clear presentation of statewide student performance goals at the state, district and school levels can go a long way toward driving improvement through focusing decisionmakers at all levels on the priority goals. Likewise, thoughtful reports from your differentiation and classification system can ensure that the incentives built into the system are effective at creating the necessary responses throughout the system.

Although this chapter does not address data reporting systems in detail, your state should be very clear about how to leverage data reporting systems to improve the implementation of CCSS and common assessments:

- How will your state collect and report new data indicators that predict success on the CCSS?
- How will your state collect and report data that can illuminate specific areas of need for student performance aligned to the CCSS?
- How will your state collect and report data that could suggest solutions for improved strategy?

Your state can go beyond reporting by connecting data reports to clear actions. For example, your state's governor, chief or other system leader can engage in regular routines with accountable officials to review CCSS-aligned data indicators and metrics and make course adjustments to implementation strategy based on the data. Your state can connect the CCSS to student learning targets and instructional resources through a Learning Management System (LMS) made available to educators. It can also collect more data that can be used as leading indicators of progress toward college and career readiness, such as interim assessments and course grades. (See National Governors Association, *Using Data to Guide State Education Policy and Practice*.)

Your state can think through how these shifts have implications for reporting to policymakers, such as state board of education and legislative committees; to educators through early warning data systems and other student and teacher performance management systems; to parents; to business and community leaders; and to the general public. Finally, your plan should think through how your state will engage each of these actors in the process of developing, releasing and publicizing data reports to further ensure clarity and use.

Resources

U.S. Education Delivery Institute: *Deliverology 101: A Field Guide for Education Leaders*

Council of Chief State School Officers: *Roadmap for Next-Generation State Accountability Systems*

Achieve: *On the Road to Implementation: Common Core State Standards and Accountability*

Achieve/The Education Trust: *Making College and Career Readiness the Mission for High Schools: A Guide for Policymakers* and the full range of guides at *Measures that Matter*

National Governors Association: *Creating a College and Career Readiness Accountability Model for High Schools*

National Governors Association: *Setting Statewide Performance Goals*

National Governors Association: *Using Data to Guide State Education Policy and Practice*

Data Quality Campaign: *Measuring the Education Pipeline*

Data Quality Campaign: *The Next Step: Using Longitudinal Data to Improve Student Success*

Conclusion

You should now have a road map for how your state can develop an approach to accountability that will advance your state's success at improving student outcomes through the implementation of CCSS and PARCC assessments. The approach encompasses a system of statewide performance goals tied to the CCSS and other college and career readiness indicators, a system to differentiate and classify schools and districts based on progress on these indicators, a system to support and intervene in districts and schools based on their classification, and a system to report actionable data to a wide array of stakeholders. It is now time to consider how this success will be reinforced through effective engagement and partnership with higher education institutions in the implementation of CCSS.

NOTES

Hide photos

Journal Gazette

Published: December 15, 2009 3:00 a.m.

Florida shows grading schools on A-F scale works

Patricia Levesque

Bennett

The Journal Gazette recently criticized state schools Superintendent Tony Bennett for pursuing a plan to add meaning and clarity to the quality of education in Indiana schools. Bennett has proposed a policy to grade schools on a letter scale – A, B, C, D or F – just as students are graded. The grades are based on objective data, such as annual test scores. It is a model that has been used in Florida for nearly a decade with great success.

The A-F grading system makes it clear to parents and the public how well our schools are educating our children. No one has to explain the difference between an A and an F.

When you ask a Florida parent about the quality of education in his or her child's school, it is likely that the parent knows the school's grade and, if it is anything less than an A, the parent has demanded improvement from the principals, teachers, district administrators, legislators and even the governor.

Perhaps more importantly, there is no ambiguity about what the letter grade means because it is based solely on objective data. The more students in a school who are performing on grade level or making progress from one year to the next, the better the grade. Simple, clear, transparent.

The A-F scale brought a command focus to learning, and the results are stunning. Before the A-F scale was adopted, Florida ranked at the bottom on the National Assessment of Educational Progress, more commonly known as the Nation's Report Card. Nearly half of Florida's fourth-graders were functionally illiterate. Today, Florida is above the national average with more than 70 percent of its fourth- and eighth-graders reading on grade level, and math scores showed a similar turnaround. In fact, the U.S. Department of Education recently recognized Florida as one of only three states in the country that is narrowing the achievement gap between rich and poor students and between white and minority students in both core subjects.

Now compare Florida to Indiana. During the same time frame, Indiana's student performance on the same national test has remained flat. In 1992, Florida lagged behind Indiana by about a year and a half. Today, Sunshine State students have surpassed students in the Hoosier state.

It is disheartening that The Journal Gazette chooses to ignore the impact of the A-F scale on the quality of education, especially when 10 years of experience has proven its effectiveness in a state as large and diverse as Florida.

Thankfully, Superintendent Bennett has demonstrated the courage to pursue a policy that serves children and parents.

Patricia Levesque is the executive director of former Florida Gov. Jeb Bush's Foundation for Excellence in Education. She wrote this for The Journal Gazette.

© Copyright 2012 The Journal Gazette. All rights reserved. This material may not be published, broadcast, rewritten or redistributed.

INDIANA ECONOMIC DIGEST
Daily News on Business and Economic Events throughout Indiana

Brought to you by the REALTORS® of Indiana
Indiana Business Research Center
indiana

5/9/2010 10:18:00 AM

EDITORIAL: Assigning letter grades to schools a good way for judging performance

Times of Northwest Indiana

The State Board of Education was right Wednesday in adopting the controversial plan to assign letter grades to Indiana schools.

"We have to be clear about how schools perform," Superintendent of Public Instruction Tony Bennett said. "That clarity is imperative."

So if Gov. Mitch Daniels approves the plan, as he is expected to do, schools will receive the same grades, A-F, given to their students. The current grades "exemplary progress," "commendable progress" and "academic progress" would become A, B and C.

Using the latest data available, about 47 percent of Indiana's schools would be graded D or F.

Teachers and school administrators are up in arms over this idea. More than 200 told the board they believe it's unfair to base a grade to an entire school based solely on one test given to the student body. More than 300 additional comments opposing the letter grades were sent via e-mail.

There are indeed some ways in which the evaluation system is unfair. Students with disabilities, for example, can affect a school's grade. That factor should be dealt with appropriately.

But the same school officials who note socioeconomic factors as affecting a school's performance don't take them into account when grading individual students, and with good reason. The central mission of a school is to educate its students. Some factors can make that goal more difficult to achieve, but the simple fact is that those factors must be overcome if students are to succeed in life.

Giving schools letter grades, just like the schools do for students, is a good way for communities to judge the schools' performance. It's also a good way for school officials to exhort the school community to provide the support necessary to turn these schools around.

School officials will need help to improve these schools. The status quo isn't good enough at schools facing difficulties. The public needs to know how well their schools are performing so they can know how dire the situation is.

We give the State Board of Education's plan an A.

Related Stories:

- State board approves school grades
- Should state grade schools
- OPINION: The death of local control

© Copyright 2012, nwitimes.com, Munster, IN

Editor, John C. DePrez Jr.; **Executive Editor, Carol Rogers; Publishers: IBRC and IAR**

Software © 1998-2012 1up! Software, All Rights Reserved

courierpress.com

Read more at courierpress.com

EDITORIAL: Grading schools

Staff Reports

Wednesday, May 12, 2010

If the state tells us that our child's school received a performance grade of "commendable progress," we nod our heads and guess that it is probably good news. On the other hand, if the state says our child's school received a performance grade of "B," we would know for certain that it is good news, and we could have some sense of satisfaction with the school.

The reverse is true, as well. If we read that our child's school received a grade of "academic watch," we are not so sure what to think. Change that term of educational jargon to a grade of "D" and we know that all may not be right with the school.

But let's give parents some credit. Regardless of how schools may be graded, parents know the first measurement of performance is how their child is doing in that school. Report cards which use an A-F grading system leave no doubt how the child is performing. It is transparency personified — no confusion, no doublespeak.

And now, the Indiana Board of Education has approved an understandable new school performance rating system based on the old reliable A-F scale in rating overall school performance on standardized tests. It will replace the warm and fuzzy system that graded schools as exemplary progress, commendable progress, academic progress, academic watch and academic probation.

Some Indiana educators don't like the new system — they say it unfairly punishes schools with students who are more difficult to educate.

Again, let's give parents credit. Most understand that schools which serve low-income neighborhoods face special challenges. But there is no good reason to try and hide it behind a wall of jargon, other than to try and fool the people.

Indiana Superintendent of Public Instruction Tony Bennett said the A-F grades will improve transparency. According to the Associated Press, Bennett said, "Indiana citizens desire a clear picture of the performance of schools. Letter grades will provide a clear picture." We agree.

Still to be decided for the new system that takes effect next year is how to measure performance in assigning grades. But to that end, the board made another decision that will serve fairness. Officials agreed to disconnect the new grading system from the federal government's "adequate yearly progress" system, another device which

attempts to inform parents about school performance. Unfortunately, the standard for the No Child Left Behind law has been flawed from the beginning.

This measurement actually exaggerates low school performance by rating both overall school performance and individual subgroups. Those subgroups include low-income children, children for whom English is a second language and children in special education.

You can have a school that is remarkably successful overall, but if one of the subgroups fails on ISTEP, the school is listed as not making the mark. But the truth is, some special education students will never be able to pass ISTEP.

Hence, Indiana school officials did the next best thing last week in removing the "adequate yearly progress" system from state's new rating system.

That, we can understand.

© 2012 Scripps Newspaper Group — Online

courierpress.com

Read more at courierpress.com

State proposes new grading scale for schools

By John Martin

Friday, August 26, 2011

EVANSVILLE — The Indiana Department of Education on Monday will release its most recent school accountability data, and it will have a different look.

Schools and school districts in the past have been lumped into one of five categories — Exemplary Progress being the best and Academic Probation being the worst. Now, though, the state will give a report card that the parent of any child in school can easily understand, with letter grades of A, B, C, D and F.

But even more changes are coming, Superintendent for Public Instruction Tony Bennett told the Evansville Courier & Press editorial board on Friday.

In future years, the state will revamp metrics used to decide which letter grade schools receive. The new system will be what Bennett calls a "growth model," taking into account students' year-to-year progress on test scores and not merely pass/fail data.

The change, said Bennett, will be central to Indiana's request for a waiver from the federal government's No Child Left Behind law, which has an entirely different grading standard for schools.

"We can bring a new accountability system using A-F grades that truly differentiates schools," Bennett said. "We should be able to merge those two measures (federal and state) into one. So we're going to be making a waiver request to the feds that says accept our state accountability system for the federal accountability system."

"We want to have a very clear definition of what schools do. There shouldn't be two announcements that sometimes don't connect to each other."

Under Indiana's grading system, schools face sanctions, including the possibility of state takeover, the longer they remain in the bottom category.

Bennett said the growth model, once it is in place, will give a truer picture of how schools are performing. He called it a "much more sophisticated analytic."

"Are we looking at test scores? Yes. But we're looking at test scores more analytically. We can drill into this and literally see what students are how far behind and what we need to do to close that gap. As opposed to it being an autopsy, we can look at what we can do proactively to put that student on a trajectory to meet standard."

U.S. Secretary of Education Arne Duncan said this month that he will consider requests for waivers from No Child Left Behind, the landmark education reform legislation signed by President George W. Bush a decade ago.

Duncan has said that this fall, he will provide details on what will be required of states to obtain a waiver. Kentucky is among a few states that have applied already.

Bennett said Indiana embraces school accountability, and the states request for a waiver will be aimed at having a universal but fair standard statewide.

"Accountability will drive education reform in this nation," Bennett said. "As human beings ... we're all going to stay the same until the pain of staying the same is greater than the pain of change. That's not criticism. It's human nature."

© 2012 Scripps Newspaper Group — Online

Truth editorial: Ed changes good for schools, students

Truth editorial for Sunday, Feb. 12, 2012.

On the same day President Obama announced waivers from provisions of the federal No Child Left Behind Act, the Indiana State Board of Education approved new rules that will give grades to the state's schools.

Both could be positive steps forward in giving Indiana what it needs to respond to the changing educational landscape.

No Child Left Behind, passed under the George W. Bush administration, was well-meaning. Its goal was to raise the educational levels of all students in the United States.

What NCLB didn't take into account was those factors that are beyond the control of teachers and administrators — especially special education students likely never to meet the standards and students whose primary language is not English. For Elkhart County schools, those two segments of students impact the test scores of the entire group when they are factored into the NCLB equation.

Make no mistake: There should be standards in education. And those standards should be high. But we also have to acknowledge that a 100 percent success rate, as essentially required by NCLB in the next couple of years, is unattainable.

So how does the state's new grading system fit into this?

The Indiana Department of Education was already moving to an A-F grading system for all schools. Thursday's changes add a new — and more complicated — evaluation system that recognizes when students' test scores increase faster than their peers. Schools will get more credit for that.

Judging students based on their own growth is a better way to ascertain whether schools are doing the job, as opposed to comparing random groups of students from year to year.

It's also important that Indiana is adopting "common core standards" shared by 45 other states, which should mean greater uniformity in what students learn. Indiana students are already moving up in national rankings. Shared standards give schools clear goals for advancement.

School letter grades and the No Child Left Behind waiver should help fuel improvement in Indiana schools.

Defining Strong State Accountability Systems

How Can Better Standards Gain Greater Traction?

A FIRST LOOK

By Eileen Reed, Janie Scull,
Gerilyn Slicker, and Amber M. Winkler

April 2012

THOMAS B. FORDHAM INSTITUTE

INDIANA

Overview

Passed in 1999, Public Law 221 (P.L. 221) established Indiana's accountability system for K-12 education. The Hoosier State has continued to tweak and improve the system for over a decade since, most recently through 2010 legislation that restructured the accountability framework to adhere to an A-to-F grading scale for school classification (similar to Florida's system), among other changes.

Indiana premises its accountability system on the notion that, by focusing on progress and providing educators and administrators with diagnostic and monitoring tools, it can induce schools to improve learning outcomes for students. To that end, the state education department provides school districts with a number of optional diagnostic and formative assessments to measure student progress.

At the same time, the state's focus on using data to drive improvement means that only the very worst schools face any consequences and interventions—and only after many years. Interventions exist only for schools that persistently rank in the lowest performance category; no improvement actions are required of any other schools, and no accountability measures are required of districts. This may be a reasonable course of action for Indiana—after all, state education departments are limited in their capacity to intervene and intervene *well*—but it is significant tradeoff nonetheless.

Below, we map Indiana's progress against six key components of strong state accountability systems.

1) Adoption of demanding, clear, and specific standards in all core content areas, and rigorous assessment of those standards

Indiana recently adopted the Common Core standards in reading and math and uses the Indiana Academic Standards in all other subjects. The standards are considered well developed and rigorous.[55] To measure student progress, the Hoosier state administers the Indiana Statewide Testing for Educational Progress Plus (ISTEP+) assessments in reading and math in grades three through eight, as well as science in grades four and six, and social studies in grades five and seven.

[55] In reviews by the Thomas B. Fordham Institute, Indiana received an A-minus for its U.S. History standards in 2011 and an A-minus for its science standards in 2012. The Common Core standards for reading and math, which Indiana adopted in 2010, earn grades of B-plus and A-minus, respectively. See *The State of State U.S. History Standards 2011*, *The State of State Science Standards 2012*, and *The State of State Standards—and the Common Core—in 2010*, at www.standards.educationgadfly.net/.

ISTEP+ also includes high school end-of-course assessments for students completing Algebra I, Biology I, or English 10.

Indiana would do well to improve the rigor of its assessments. While many states establish lower cut scores of proficiency, Indiana's cut scores for eighth-grade reading and fourth- and eighth-grade math all equate to the *basic* level on the National Assessment of Educational Progress (NAEP); worse, its fourth-grade reading cut score equates to NAEP's *below basic* level.[56]

2) Reporting of accessible and actionable data to all stakeholders, including summative outcome data and other formative data to drive continuous improvement

Indiana fares moderately well on this component. The state presents information on school and district performance through an easy-to-navigate database called the DOE Compass. The database includes information on proficiency rates, graduation rates, dropout rates, AP/IB passing rates, SAT/ACT scores, and AYP data. Unfortunately, the state does not disaggregate any data by subgroup (outside of the AYP data, which only include subgroup proficiency rates), and it only displays student performance by proficiency rate instead of also by performance level.

The state does, however, allow parents direct online access to their children's data, including longitudinal results. Further, Indiana reports a measure of growth known as the Indiana Growth Model. The growth model evaluates individual-student growth relative to peers with similar academic histories, and classifies student growth into three bands:

- **High Growth** – 66th to the 99th percentile
- **Typical Growth** – 35th to the 65th percentile
- **Low Growth** – 1st to the 34th percentile

For school and district reporting purposes, the median student growth percentile in a school or district represents that entity's growth.

[56] National Center for Education Statistics, *Mapping State Proficiency Standards Onto the NAEP Scales: Variation and Change in State Standards for Reading and Mathematics, 2005–2009* (Washington, D.C.: U.S. Department of Education, August 2011), http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2011458. For reading and math, the state plans to participate in new Common Core assessments beginning in the 2014-15 school year, as it is a governing state in the Partnership for Assessment of Readiness for College and Careers (PARCC).

3) Annual determinations and designations for each school and district that meaningfully differentiate their performance

Indiana's state board of education approved a new metric in November 2011 by which the state will now evaluate schools and districts. While Indiana's previous metric included both state-required and federal AYP measures, the new metric exorcizes all federal measures.

Elementary and middle schools evaluations are now based on proficiency, growth, and assessment participation rates. Schools receive preliminary scores based on their proficiency rates; those scores can be raised if the lowest performing quartile or the remaining 75 percent of students show high growth, as measured by the Indiana growth model. The preliminary scores can also be lowered if a certain percentage of all students show low growth, or if assessment participation rates are low.

High schools are evaluated based on student performance on end-of-course assessments in English 10 and Algebra I, graduation rates, and post-secondary readiness indicators—including AP/IB passing rates, college credit attainment, and industry certification. All three areas are added together to form a final score. Schools receive preliminary assessment scores based on the percentage of students passing the end-of-course assessments; those scores can be raised or lowered depending on student progress from eighth to tenth grade, and from tenth grade to graduation. Schools also receive preliminary graduation scores based on four-year graduation rates—which can be raised or lowered depending on the number of honors and waiver diplomas[57] issued by the school, and by five-year graduation rates.

Based on these measures, Indiana assigns school ratings on an A-to-F scale:

- A (Exemplary Progress)
- B (Commendable Progress)
- C (Academic Progress)
- D (Academic Watch – Priority)
- F (Academic Probation – High Priority)

Districts are also scored according to the same metric, and are categorized by the same five performance designations.

[57] A student who does not meet graduation exam requirements may graduate with a waiver diploma if he passes all required courses, demonstrates to the satisfaction of his school that he has met the performance bar set by the state graduation exam, or completes an internship and workforce readiness assessment.

Indiana's removal of AYP in its accountability metric means that traditional subgroups of students are no longer incorporated into the accountability system. As one state representative explained, by focusing on the lowest-performing quartile of students, the state aims to target low performers regardless of race or disability.

4) A system of rewards and consequences to drive improvement at the school and district levels

Rewards

District representatives indicate that the greatest incentive for school improvement is the state's designation system and the accountability requirements tied to different labels. Beyond the school ranking system, Indiana has two incentive programs of note, one of which entails financial rewards:

- **Four Star Schools Awards** – The Four Star Schools Award recognizes schools making AYP that also rank in the state's top 25 percent of schools based on ISTEP+ results in reading and mathematics in all grades tested. Winning schools are recognized through various press releases and are featured on the state department of education's website.

- **Graduation Rate Incentive Award** – This new award program buttresses the state's focus on increasing graduation rates. The program is competitive and open to all Indiana public high schools. Monetary rewards go to the twelve public high schools with the greatest increases in graduation rates between the previous two school years: ten mid-to-large schools (more than three hundred students) receive $20,000 awards, and two small schools (fewer than three hundred students) receive $10,000 awards. The building principal and district superintendent of each winning school receive the funds and may distribute the cash awards to staff members whose work was critical to achieving the graduation-rate increases. The principal may receive no more than $5,000 of the award amount.

Sanctions

Indiana's system for intervening in low-performing schools is relatively light on details and relies heavily on the resolve of the state and local boards of education. To begin, the state only holds schools accountable for performance; districts, no matter their performance designations, face no sanctions. The only schools that face state sanctions are those that repeatedly continue to earn F grades. For the first few years a school earns Fs, the system primarily requires local school boards to help create and implement improvement plans. The local school board can request that the state board of education appoint an outside team to manage the

school or assist in developing a new improvement plan, but it is not required to do so.

Not until years four, five, and six do the requirements become more prescriptive. At this stage, the state board appoints an expert team to rework the school's improvement plan and recommend changes. The expert team must include representatives from the community or region that the school serves. According to a representative of the state department of education, this requirement is intended to engage the community in holding schools accountable for meeting student learning needs.

If a school is still on academic probation in year six, the state board is required to conduct a hearing and consider testimony on options such as closing the school, merging the school with another school, or assigning a third-party Turnaround School Operator to run all or part of the school. Other options, including continued revision of the improvement plan, are also on the table.

As such, Indiana's accountability system only addresses its lowest-performing schools, and allows them to languish in low-performance for years before taking any significant actions—and even then, it provides avenues through which schools can avoid those drastic interventions. Thus far, however, the state board of education has proven steadfast in its resolve and not opted for the easy way out (improvement-plan stasis). The 2011-12 school year marks the first year that schools could have earned F grades for six years running, and seven schools did so. Of those seven, the board chose the turnaround option for five schools and a slightly less intensive option for the remaining two schools. Of course, it remains to be seen whether the turnarounds will work or if the board will opt for closure or school merges—presumably even more severe options—in the future.[58]

Supports

As its relatively sparse outline of sanctions demonstrates, the Hoosier State's accountability system aims to drive improvement through progress monitoring and support, not through mandated interventions. On that side of the coin, Indiana provides its schools with a full kit of diagnostic and assessment tools to support local improvement efforts. See *Diagnostic and Assessment Tools* for a partial list of these tools made available by the state (some are required and some are optional).

One superintendent described how such tools support local school improvement efforts: "With the data from [diagnostic assessments], we've seen our buildings

[58] A state representative noted that the state is working to expand the outline of sanctions for schools earning both Ds and Fs, instead of just Fs; to eliminate the weaker options allowed for schools in year six; and to establish an outline for district accountability.

Diagnostic and Assessment Tools

IREAD-3 - The purpose of the Indiana Reading Evaluation And Determination (IREAD-3) assessment is to measure foundational reading standards. Based on the Indiana Academic Standards, IREAD-3 is a summative assessment that evaluates reading skills of third-grade students to ensure that all students can read proficiently before moving on to fourth grade. IREAD-3 is administered in addition to ISTEP+ and provides timely information in order to plan for summer intervention if needed, as well as for decisions relative to fourth-grade placements. IREAD-3 is mandatory.

mCLASS - The purpose of the mCLASS assessments is to provide diagnostic measures for K-2 students in literacy and numeracy. These assessments help identify students' foundational skills and provide teachers with instructional suggestions based on student performance on benchmark assessments and regular progress monitoring. These assessments are optional.

Acuity (3-8 and Algebra) - The purpose of the Acuity assessments is to provide diagnostic measures for students in grades three through eight in reading, mathematics, science, and social studies, as well as for students in Algebra I. Assessment reports provide standards-aligned performance data, which support educators' ability to inform instruction. These assessments are optional.

make real gains...Because the state provides these resources it saves us expense at the local level... It's a big help."

5) A system of rewards and consequences to drive improvement at the individual student level

Indiana requires that high school students pass end-of-course assessments in English 10 and Algebra I in order to graduate. In addition, beginning in spring 2012, all third-grade students will be required to pass the IREAD-3 assessment in order to be promoted to fourth grade (a similar requirement exists in Florida).

Like all states, federal NCLB regulations require Indiana to allow students in low-performing schools to attend other, high-performing schools. According to a state representative, however, because Indiana's low-performing schools are primarily

clustered in Indianapolis and Gary, few high-performing options exist for students in failing schools; the state does not enforce or encourage this option in any meaningful way. On the bright side, Indiana recently introduced a voucher program for students statewide. While the program is an option for low-income students, and not specifically those in low-performing *schools*, the overlap of those groups means that the program will in practice allow many students in low-performing schools to attend other private schools.

6) A system of rewards and consequences to drive improvement at the individual teacher and administrator level

In its 2011 legislative session, Indiana passed a landmark law requiring local districts to develop annual teacher and principal evaluations. The evaluations must be "significantly" informed by student achievement and growth measures and must differentiate performance across four categories: highly effective, effective, improvement necessary, or ineffective.

Principals will continue to be employed under one-year contracts. For teachers, districts must now incorporate evaluation results into their compensation structures and employment decisions. New teachers will automatically fall into a "probationary" category, and can be dismissed after just one ineffective rating. They are also eligible for "established" status, akin to receiving tenure, after earning three effective or highly effective ratings in a five-year period. Established teachers cannot be dismissed unless they receive two consecutive ineffective ratings, or three ineffective or "improvement necessary" ratings over a five-year period.

Indiana also recently introduced a competitive Excellence in Performance Award to provide districts with grants for establishing bonus programs for effective and highly effective teachers. To apply, districts must outline their new evaluation metrics and describe the expected impact of the bonus program on teacher recruitment, development, and retention.

What are the strengths and limitations of Indiana's accountability system?

Strengths

> **Strong academic standards.** Indiana's content standards are rated among the best in the nation. The state routinely uses national-level content experts to develop and review its academic standards. The state also annually tests in all four key content areas—reading, math, science, and social studies.

Robust accountability metric. By combining measures of proficiency, growth, and college and career readiness, Indiana captures a detailed picture of student, school, and district performance. It also disaggregates the lowest-performing quartile of students to ensure that schools and districts are held to account for those students' performance. In addition, the A-to-F grading scale for schools and districts is straightforward and easily understood by policymakers, parents, and community members.

A full toolkit of instruments to monitor student progress. The state's investment in diagnostic and progress monitoring tools to support the work of school improvement is key. Educators can access data that pinpoint individual-student learning needs, and building administrators have tools and resources to monitor progress toward success on summative assessments. The state has also embraced the use of technology; for example, parents have direct online access to their students' data, including longitudinal results.

Teacher evaluation system that informs salary and employment decisions. Indiana's new teacher (and principal) evaluation legislation revamps the state's traditional salary structure and requires evaluations to include student performance and growth indicators. Further, the state actively encourages districts to adopt model systems by offering competitive grants for teacher bonuses.

Limitations

Limited system of consequences. While it measures district performance, Indiana's accountability structure ties no supports, sanctions, or interventions to district results. The system only meaningfully addresses those schools designated as the lowest performing for multiple years. While the state has clearly invested in progress-monitoring tools to support improvement efforts, its accountability system could benefit from a wider range of interventions and greater urgency in implementing those strategies with low-performing schools.

Final Word

Indiana's focus on growth and progress monitoring appears to be its greatest asset. Still, the system rests on the assumption that simply giving districts the data and diagnostic tools to track and monitor progress can spur academic improvement. Perhaps. But the interventions for low performers must be robust and the implementation swift. In Indiana, the choice of state-level interventions is limited, and the powers that be could choose to implement the weakest available option. If

interventions are to be targeted to the most egregious cases, state leaders must be vigilant and bold in their prescriptions.

Information on Indiana's education-accountability system was primarily drawn from interviews with state representatives, district representatives, and local stakeholders, as well as from the Indiana Department of Education website at www.doe.in.gov. Additional information was drawn from the National Council on Teacher Quality's *2011 State Teacher Policy Yearbook.*

A Technical Overview of the Student Growth Percentile Methodology: Student Growth Percentiles and Percentile Growth Projections/Trajectories

Damian W. Betebenner
The National Center for the Improvement of Educational Assessment
Dover, New Hampshire

August 12, 2011

Introduction: Why Student Growth?

Accountability systems constructed according to federal adequate yearly progress (AYP) requirements currently rely upon annual "snap-shots" of student achievement to make judgments about school quality. Since their adoption, such *status measures* have been the focus of persistent criticism (Linn, 2003; Linn, Baker, & Betebenner, 2002). Though appropriate for making judgments about the achievement level of students at a school for a given year, they are inappropriate for judgments about educational *effectiveness*. In this regard, status measures are blind to the possibility of low achieving students attending effective schools. It is this possibility that has led some critics of No Child Left Behind (NCLB) to label its accountability provisions as unfair and misguided and to demand the use of growth analyses as a better means of auditing school quality.

A fundamental premise associated with using student growth for school accountability is that "good" schools bring about student growth in excess of that found at "bad" schools. Students attending such schools—commonly referred to as highly effective/ineffective schools—tend to demonstrate extraordinary growth that is causally attributed to the school or teachers instructing the students. The inherent believability of this premise is at the heart of current enthusiasm to incorporate growth into accountability systems. It is not surprising that the November 2005 announcement by Secretary of Education Spellings for the Growth Model Pilot Program (GMPP) permitting states to use growth model results as a means for compliance with NCLB achievement mandates and the Race to the top competitive grants program were met with great enthusiasm by states. (Spellings, 2005).

Following these use cases, the primary thrust of growth analyses over the last decade has been to determine, using sophisticated statistical techniques, the amount of student progress/growth that can be justifiably attributed to the school or teacher—that is, to disentangle current *aggregate* level achievement from effectiveness (Braun, 2005; Rubin, Stuart, & Zanutto, 2004; Ballou, Sanders, & Wright, 2004; Raudenbush, 2004). Such analyses, often called *value-added* analyses, attempt to estimate the teacher or school contribution to student achievement. This contribution, called the *school* or *teacher effect*, purports to quantify the impact on achievement that this school or teacher would have, on average, upon similar students assigned to them for instruction. Clearly, such analyses lend themselves to accountability systems that hold schools or teachers responsible for student achievement.

Despite their utility in high stakes accountability decisions, the causal claims of teacher/school

effectiveness addressed by value-added models (VAM) often fail to address questions of primary interest to education stakeholders. For example, VAM analyses generally ignore a fundamental interest of stakeholders regarding student growth: How much growth did a student make? The disconnect reflects a mismatch between questions of interest and the statistical model employed to answer those questions. Along these lines, Harris (2007) distinguishes value-added for program evaluation (VAM-P) and value-added for accountability (VAM-A)—conceptualizing accountability as a difficult type of program evaluation. Indeed, the current climate of high-stakes, test-based accountability has blurred the lines between program evaluation and accountability. This, combined with the emphasis of value-added models toward causal claims regarding school and teacher effects has skewed discussions about growth models toward causal claims at the expense of description. Research (Yen, 2007) and personal experience suggest stakeholders are more interested in the reverse: description first that can be used secondarily as part of causal fact finding.

In a survey conducted by Yen (2007), supported by the author's own experience working with state departments of education to implement growth models, parents, teacher, and administrators were asked what "growth" questions were most of interest to them.

Parent Questions:

- Did my child make a year's worth of progress in a year?
- Is my child growing appropriately toward meeting state standards?
- Is my child growing as much in Math as Reading?
- Did my child grow as much this year as last year?

Teacher Questions:

- Did my students make a year's worth of progress in a year?
- Did my students grow appropriately toward meeting state standards?
- How close are my students to becoming Proficient?
- Are there students with unusually low growth who need special attention?

Administrator Questions:

- Did the students in our district/school make a year's worth of progress in all content areas?
- Are our students growing appropriately toward meeting state standards?
- Does this school/program show as much growth as that one?
- Can I measure student growth even for students who do not change proficiency categories?
- Can I pool together results from different grades to draw summary conclusions?

As Yen remarks, all these questions rest upon a desire to understand whether observed student progress is "reasonable or appropriate" (Yen, 2007, p. 281). More broadly, the questions seek a description rather than a parsing of responsibility for student growth. Ultimately, questions may turn to who/what is responsible. However, as indicated by this list of questions, they are not the starting point for most stakeholders.

In the following student growth percentiles and percentile growth projections/trajectories are introduced as a means of understanding student growth in both normative and a criterion referenced ways. With these values calculated we show how growth data can be utilized in both a norm- and in a criterion-referenced manner to inform discussion about education quality. We assert that

the establishment of a normative basis for student growth eliminates a number of the problems of incorporating growth into accountability systems providing needed insight to various stakeholders by addressing the basic question of how much a student has progressed (Betebenner, 2008, 2009).

Student Growth Percentiles

It is a common misconception that to quantify student progress in education, the subject matter and grades over which growth is examined must be on the same scale—referred to as a vertical scale. Not only is a vertical scale not necessary, but its existence obscures concepts necessary to fully understand student growth. Growth, fundamentally, requires change to be examined for a single construct like math achievement across time—*growth in what?*.

Consider the familiar situation from pediatrics where the interest is on measuring the height and weight of children over time. The scales on which height and weight are measured possess properties that educational assessment scales aspire towards but can never meet.[1]

> An infant male toddler is measured at 2 and 3 years of age and is shown to have grown 4 inches. The magnitude of increase—4 inches—is a well understood quantity that any parent can grasp and measure at home using a simple yardstick. However, parents leaving their pediatrician's office knowing only how much their child has grown would likely be wanting for more information. In this situation, parents are not interested in an absolute criterion of growth, but instead in a normative criterion locating that 4 inch increase alongside the height increases of similar children. Examining this height increase relative to the increases of similar children permits one to diagnose how (a)typical such an increase is.

Given this reality in the examination of change where scales of measurement are perfect, we argue that it is absurd to think that in education, where scales are at best quasi-interval, one can/should examine growth differently.[2]

Going further, suppose that scales did exist in education similar to height/weight scales that permitted the calculation of absolute measures of annual academic growth for students. The response to a parent's question such as, "How much did my child progress?", would be a number of scale score points—an answer that would leave most parents confused wondering whether the number of points is good or bad. As in pediatrics, the search for a description regarding changes in achievement over time (i.e., growth) is best served by considering a norm-referenced quantification of student growth—*a student growth percentile* (Betebenner, 2008, 2009).

A student's growth percentile describes how (a)typical a student's growth is by examining his/her current achievement relative to his/her *academic peers*—those students beginning at the same place. That is, a student growth percentile examines the current achievement of a student relative to other students who have, in the past, "walked the same achievement path". Heuristically, if the state assessment data set were extremely large (in fact, infinite) in size, one could open the infinite data set and select out those students with the exact same prior scores and compare how the selected student's current year score compares to the current year scores of those students with the same prior year's scores—his/her academic peers. If the student's current year score exceeded the scores of most of his/her academic peers, in a normative sense they have done as well. If the student's

[1] Height and weight scales are interval (actually, ratio scales) where a unit increase reflects an equivalent increase in the underlying quality being measured no matter where on the scale the increase occurs.

[2] The scales on which students are measured are often assumed to possess properties similar to height and weight but they don't. Specifically, scales are assumed to be interval where it is assumed that a difference of 100 points at the lower end of the scale refers to the same difference in ability/achievement as 100 points at the upper end of the scale. See Lord (1975) and Yen (1986) for more detail on the interval scaling in educational measurement.

current year score was less than the scores of his/her academic peers, in a normative sense they have not done as well.

The four panels of Figure 1 depict what a student growth percentile represents in a situation considering students having only two consecutive achievement test scores.[3]

Upper Left Panel Considering all pairs of prior year and current year scores for all students in the state yields a bivariate (two variable) distribution. The higher the distribution, the more frequent the pair of scores.

Upper Right Panel Taking account of prior achievement (i.e., conditioning upon prior achievement) fixes a the value of the prior year scale score (in this case at 600) and is represented by the red slice taken out of the bivariate distribution.

Lower Left Panel Conditioning upon prior achievement defines a *conditional distribution* which represents the distribution of outcomes on the current year test assuming a prior year score of 600. This distribution is indicated with the solid red curve.

Lower Right Panel The conditional distribution provides the context against which a student's current year achievement can be examined and understood in a norm-referenced fashion. Students with achievement in the upper tail of the conditional distribution have demonstrated high rates of growth relative to their academic peers whereas those students with achievement in the lower tail of the distribution have demonstrated low rates of growth. Students with current achievement in the middle of the distribution could be described as demonstrating "average" or "typical" growth.

In Figure 1, the student scores approximately 650 on the current year test. Within the conditional distribution, the value of 650 lies at the 75th percentile. Thus the student's growth from 600 in the prior year to 650 in the current year met or exceeded that of approximately 70 percent of students starting from the same place. Thus, relative to others with the same prior achievement score, this 50 point increase is above average. It is important to note that characterizing a student growth percentile as "adequate", "good", or "enough" requires a qualitative judgment to be rendered—growth standard setting. Later in this paper growth adequacy standards are investigated *vis-à-vis* state achievement levels.

Figure 1 also serves to illustrate the relationship between a vertical scale and student growth percentiles. Using the vertical scale implied by Figure 1, the student grew 50 points (from 600 to 650) between the prior and current year. This 50 points represents the absolute magnitude of change. Quantifying the magnitude of change is scale dependent. For example, different vertical achievement scales in the prior and current year would yield different annual scale score increases: A scale score increase of 50 could be changed to a scale score increase of 10 using a simple transformation of the vertical scale on which all the students are measured. However, relative to other students, his/her growth has not changed—the growth percentile is invariant to scale transformations common in educational assessment (Briggs & Betebenner, 2009). Student growth percentiles normatively situate achievement change bypassing questions associated with the magnitude of change, and directing attention toward relative standing.

To fully understand how many states intend to use growth percentiles to make determinations about whether a student's growth is sufficient, the next section details specifics of how student growth percentiles are calculated. These calculations are subsequently used to calculate percentile growth

[3]By default, the SGP package (Betebenner & Iwaarden, 2011) uses the entire achievement history of the student subject to some suitability conditions. Figure 1 is presented with just a single prior score to facilitate representing a conditional distribution.

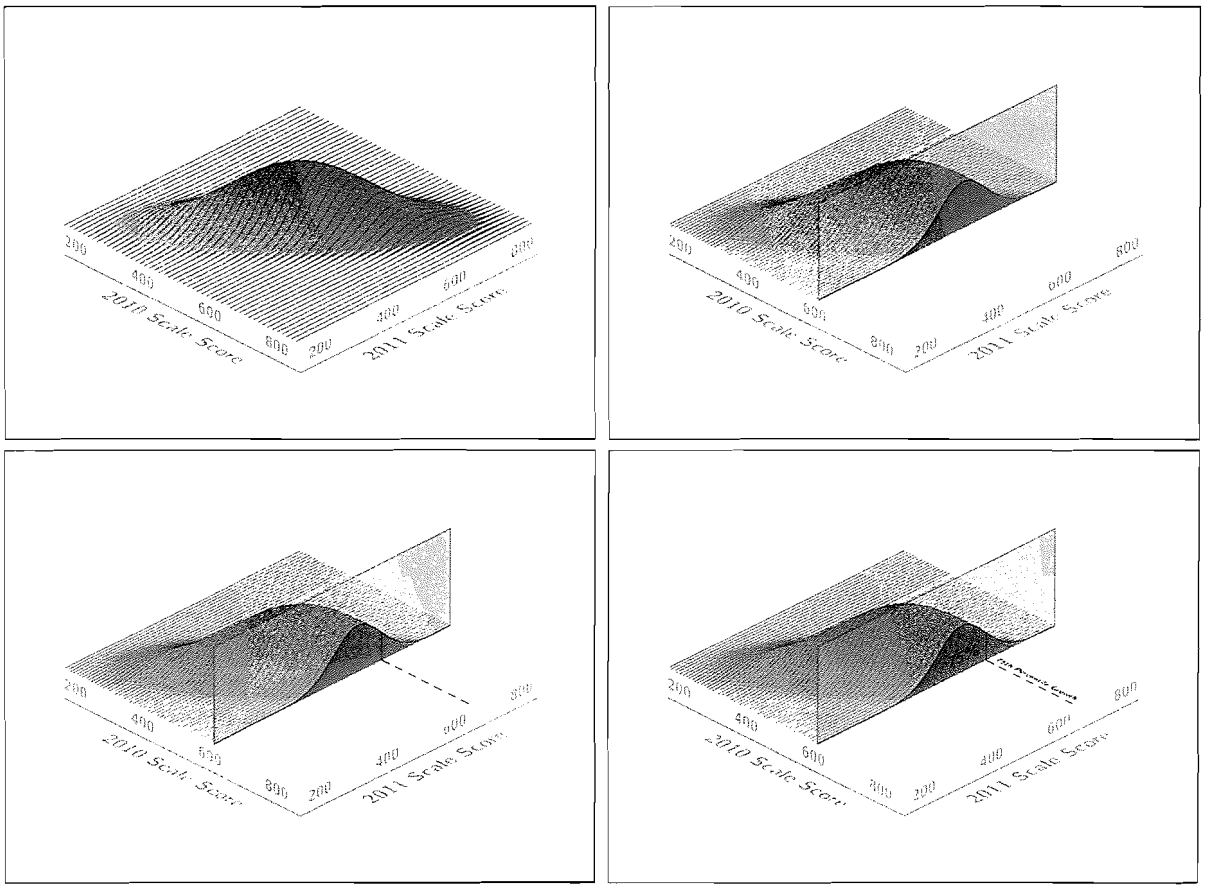


Figure 1: Figures depicting the distribution associated with 2010 (prior year) and 2011 (current year) student scale scores together with the conditional distribution and associated growth percentile

projections/trajectories that are used to establish how much growth it will take for each student to reach his/her achievement targets.

Student Growth Percentile Calculation

Quantile regression is used to establish curvi-linear functional relationships between the cohort's prior scores and the cohort's current scores. Specifically, for each grade by subject cohort, quantile regression is used to establish 100 (1 for each percentile) curvi-linear functional relationships between the students grade 3, grade 4, grade 5, and grade 6 prior scores and their grade 7 scores.[4] The result of these 100 separate analyses is a single coefficient matrix that can be employed as a look-up table relating prior student achievement to current achievement for each percentile. Using the coefficient matrix, one can plug in *any* grade 3, 4, 5, and 6 prior score combination to the functional relationship to get the percentile cutpoints for grade 7 conditional achievement distribution associated with that prior score combination. These cutpoints are the percentiles of the conditional distribution associated with the individual's prior achievement

Consider a student with the following reading scores:

[4]For the mathematical details underlying the use of quantile regression in calculating student growth percentiles, see the Appendix *Student Growth Percentile Estimation* on 16.

Grade 3/2004	Grade4/2005	Grade 5/2006	Grade 6/2007	Grade 7/2008
519	518	587	589	601

Table 1: Scale scores for a hypothetical student across 5 years in reading

Using the coefficient matrix derived from the quantile regression analyses based upon grade 3, 4, 5, and 6 scale scores as independent variables and the grade 7 scale score as the dependent variable together with this student's vector of grade 3, 4, 5, and 6 grade scale scores provides the scale score percentile cutpoints associated with the grade 7 conditional distribution for these prior scores.

1st	2nd	3rd	···	10th	···	25th	···	50th	51st	···	75th	···	90th	···	99th
514.8	534.9	543.9	···	566.9	···	584.8	···	600.5	601.3	···	616.9	···	630.1	···	653.8

Table 2: Percentile cutscores for grade 7 reading based upon the grade 3, 4, 5, and 6 reading scale scores given in Table 1

The percentile cutscores for 7th grade reading in Table 2 are used with the student's *actual* grade 7 reading scale score to establish his/her growth percentile. In this case, the student's grade 7 scale score of 601 lies above the 50th percentile cut and below the 51st percentile cut, yielding a growth percentile of 50. Thus, the progress demonstrated by this student between grade 6 and grade 7 exceeded that of 50 percent of his/her academic peers—those students with the same achievement history. States can qualify student growth by defining ranges of growth percentiles. For example, the Colorado Growth Model designates growth percentiles between 35 and 65 as being *typical*. Using Table 2, another student with the exact same grade 3, 4, 5, and 6 prior scores but with a grade 7 scale score of 530, would have a growth percentile of 1, which is designated as *low*.

This example provides the basis for beginning to understand how growth percentiles in the SGP Methodology are used to determine whether a student's growth is *(in)adequate*. Suppose that in grade 6 a one-year (i.e., 7th grade) achievement goal/target of proficiency was established for the student. Using the lowest proficient scale score for 7th grade reading, this target corresponds to a scale score of 619. Based upon the results of the growth percentile analysis, this one year target corresponds to 78th percentile growth. Their growth, obviously, is less than this and the student has not met this individualized growth standard.

Percentile Growth Projections/Trajectories

Building upon the example just presented involving only a one-year achievement target translated into a growth standard, this section extends this basic idea and shows how multi-year growth standards are established based upon pre-established achievement targets/goals. That is, by defining a future (e.g., a 3 year) achievement target for each student, we show how growth percentile analyses can be used to quantify what level of growth, expressed as a per/year growth percentile, is required by the student to reach his/her achievement target. Unique to the SGP Methodology is the ability to stipulate *both* what the growth standard is as well as how much the student actually grew in a metric that is informative to stakeholders.

Defining Adequate Growth

Establishing thresholds for growth for each student that can be used to make adequacy judgments requires pre-established achievement targets and a time-frame to reach the target for each student against which growth can be assessed (i.e., growth-to-standard). Adequacy in many contexts has been defined as *catching-up* and *keeping-up*:

Catch-Up Those students currently not proficient are expected to reach proficient within 4 years following the establishment of the achievement target or by grade 10, whichever comes sooner.[5]

Keep-Up Those students currently at or above proficient are expected to remain at or above proficient in all of the 4 years following the establishment of the achievement target or by grade 10, whichever comes sooner.

Move-Up Those students currently proficient are expected to reach advanced within 4 years following the establishment of the achievement target or by grade 10, whichever comes sooner.

The previous definitions specify "4 years following the establishment of the achievement target" as the time frame. For example, an non-proficient 3rd grader would be expected to be proficient by 7th grade, assuming a 1 grade/year progression. The first check of the student's progress occurs in 4th grade, when the student's growth over the last year is compared against targets calculated to assess their progress along a multi-year time-line. The question asked following the 4th grade for the student is: Did the student become proficient and if not are they on track to become proficient within 3 years?

Achievement targets are not unique. For example, in the currently un-implemented version the Colorado Growth Model approved by the federal government for AYP purposes, less rigorous achievement targets than those established by Colorado's Senate Bill 163 are proposed:

- Unsatisfactory students are expected to be NCLB proficient (Colorado partially proficient) within 4 years following the establishment of the achievement target or by grade 10, whichever comes sooner.
- NCLB proficient students are expected to remain NCLB proficient for at least the next 4 years following the establishment of the achievement target or by grade 10, whichever comes sooner.
- Colorado proficient students are expected to remain Colorado proficient for at least the next 4 years following the establishment of the achievement target or by grade 10, whichever comes sooner.
- Advanced students are expected to remain at or above Colorado proficient for at least the next 4 years following the establishment of the achievement target or by grade 10, whichever comes sooner.

It is important to note that each student's achievement targets and time-frame to reach these targets are fixed. However, depending upon the student's interim rates of growth, the growth percentiles required to reach his/her fixed achievement target are likely going to be adjusted. For example, a 3rd grade non-proficient student in reading (with an achievement target of proficient in reading by the 7th grade) might demonstrate sizable growth between 3rd and 4th grade and still

[5] The time frame of 4 years following the establishment of the achievement target is equivalent to within 3 years of when the first growth adequacy judgment is established for the student.

remain non-proficient. However, the question of relevance for state departments of education, given that they did not reach proficiency in 4th grade, is whether they are on track to become proficient within 3 years. Following the 5th grade, if the student isn't proficient a determination will be made as to whether they are on track to become proficient with 2 years. And following the 6th grade, a determination will be made as to whether they are on track to become proficient in 1 year if they are not already proficient. The destination and the time frame to reach it remain fixed.

Calculation of Growth Percentile Targets

As mentioned previously, the calculation of student growth percentiles across all grades and students results in the creation of numerous coefficient matrices that relate prior with current student achievement. These matrices constitute an annually updated statewide historical record of student progress. For the SGP Methodology, they are used to determine what level of percentile growth is necessary for each student to reach future achievement targets. For example, in the calculation of student growth percentiles in 2008 in Colorado, the following coefficient matrices are produced:[6]

Grade 4 Using grade 3 prior achievement.

Grade 5 Using grade 4 and grades 3 & 4 prior achievement.

Grade 6 Using grade 5, grades 4 & 5, and grades 3, 4, & 5 prior achievement.

Grade 7 Using grade 6, grades 5 & 6, grades 4, 5, & 6, and grades 3, 4, 5, & 6 prior achievement.

Grade 8 Using grade 7, grades 6 & 7, grades 5, 6, & 7, grades 4, 5, 6, & 7, and grades 3, 4, 5, 6, & 7 prior achievement

Grade 9 Using grade 8, grades 7 & 8, grades 6, 7, & 8, grades 5, 6, 7, & 8, and grades 4, 5, 6, 7, & 8 prior achievement.

Grade 10 Using grade 9, grade 8 & 9, grades 7, 8, & 9, grades 6, 7, 8, & 9, and grades 5, 6, 7, 8, & 9 prior achievement

To describe how these numerous coefficient matrices are used together to produce 1, 2, 3, and 4 year growth targets, consider, for example, a 2008 4th grade student in reading with 3rd and 4th grade state reading scores of 450 (Unsatisfactory) and 500 (Unsatisfactory), respectively. The following are the steps that transpire over 4 years to determine whether this student is on track to reach proficient.

August 2007 Accountability clock begins requiring students to reach state defined achievement targets within 4 years or by grade 10. In this example, the unsatisfactory 3rd grade (in 2007) student under consideration is expected to be proficient by grade 7 in 2011.

August 2008 Employing the coefficient matrices derived in the calculation of 2008 student growth percentiles:

- First, the coefficient matrix relating grade 4 with grade 3 prior achievement is used to establish the percentile cuts (i.e., one-year growth percentile projections/trajectories). If the student's actual 2008 growth percentile exceeds the percentile cut associated with

[6] Note that because testing began in 2003 in Colorado, at present there is a maximum number of 5 consecutive prior achievement scores.

proficient, then the student's one year growth is enough to reach proficient.[7] If the student reaches his/her achievement goal, then the accountability clock is reset for this student, with a new achievement goal of remaining proficient for the next 4 years.

- Next, the 2 year growth percentile projections/trajectories are calculated, from 2007 to 2009. The student's actual grade 3 scale score together with the 99 hypothetical one-year growth percentile projections/trajectories derived in the previous step are plugged into the most recently derived coefficient matrix relating grade 5 with grade 3 & 4 prior achievement. This yields the percentile cuts (i.e., 2 year growth percentile projections/trajectories) for the student indicating what consecutive two-year 1st through 99th percentile growth (based upon the most recent student growth histories in the state) will lead to. Using the August 2007 achievement targets (proficient by 7th grade, for this student), 2 year growth sufficient to reach the target is determined and the student's growth percentile is compared to this target. If the student's growth percentile exceeds this target, then the student is deemed on track to reach proficient.
- Next, the 3 year growth percentile projections/trajectories are established. The student's actual grade 3 scale score together with the 99 hypothetical 1 and 2 year growth percentile projections/trajectories derived in the previous step are plugged into the coefficient matrix relating grade 6 with grade 3, 4, & 5 prior achievement. This yields the percentile cuts (i.e., 3 year growth percentile projections/trajectories) for each student indicating what consecutive three-year 1st through 99th percentile growth (based upon the most recent student growth histories in the state) will lead to in terms of future achievement. Using the August 2007 achievement targets (proficient by 7th grade, for this student), 3 year growth sufficient to reach the target is determined and the student's growth percentile is compared to this target. If the student's growth percentile exceeds this target then the student is deemed on track to reach proficient.
- Last, the 4 year growth percentiles projections/trajectories are established. The student's actual grade 3 scale score together with the 99 hypothetical 1, 2, and 3 year growth percentile projections/trajectories derived in the previous step are plugged into the coefficient matrix relating grade 7 with grade 3, 4, 5, & 6 prior achievement. This yields the percentile cuts (i.e., 4 year growth percentile projections/trajectories) for each student indicating what consecutive four-year 1st through 99th percentile growth (based upon the most recent student growth histories in the state) will lead to in terms of future achievement. Using the August 2007 achievement targets (proficient by 7th grade for this student), 4 year growth sufficient to reach the target is determined and the student's growth percentile is compared to this target. If the student's growth percentile exceeds this target then the student is deemed on track to reach proficient.

August 2009 Employing the coefficient matrices derived in the calculation of 2009 student growth percentiles:

- First, with the student now presumably completing grade 5, the coefficient matrix relating grade 5 with grade 3 & 4 prior achievement is used to establish 99 percentile cuts (i.e., one-year growth percentile projections/trajectories). If the student's actual 2009 growth percentile exceeds the cut associated with proficient, then the student's one year growth

[7] Checking growth adequacy using one-year achievement targets is equivalent to confirming whether the student reached his/her one-year achievement target since the coefficient matrices used to produce the percentile cuts are based on current data.

was enough to reach proficient.[8] If the student reaches his/her achievement goal, then the accountability clock is reset for this student, with a new achievement goal of remaining proficient for the next 4 years.

- Next, the student's grade 3 & 4 actual scores together with the 99 hypothetical one-year growth percentile projections/trajectories derived in the previous step are plugged into the coefficient matrix relating grade 6 with grade 3, 4, & 5 prior achievement. This yields 99 percentile cuts (i.e., 2 year growth percentile projections/trajectories) for the student indicating what consecutive two-year 1st through 99th percentile growth (based upon the most recent student growth histories in the state) will lead to in terms of future achievement. Using the August 2007 accountability achievement targets (proficient by 7th grade for this student), 2 year growth sufficient to reach the target is determined and the student's growth percentile is compared to this target. If the student's growth percentile exceeds this target then the student is deemed on track to reach proficient.
- Finally, the student's grade 3 & 4 actual scores together with the 99 hypothetical 1 & 2 year growth percentile projections/trajectories derived in the previous steps are plugged into the coefficient matrix relating grade 7 with grade 3, 4, 5, & 6 prior achievement. This yields 99 percentile cuts (i.e., 3 year growth percentile projections/trajectories) for the student indicating what consecutive three-year 1st through 99th percentile growth (based upon the most recent student growth histories in the state) will lead to in terms of future achievement. Using the August 2007 accountability achievement targets (proficient by 7th grade for this student), 3 year growth sufficient to reach the target is determined and the student growth percentile is compared to this target. If the student's growth percentile exceeds this target then the student is deemed on track to reach proficient.
- No 4 year targets are utilized because they exceed the time-frame initially established for the student to reach proficient.

August 2010 Employing the coefficient matrices derived in the calculation of 2009 student growth percentiles:

- First, with the student now presumably completing grade 6, the coefficient matrix relating grade 6 with grade 3, 4, & 5 prior achievement is used to establish 99 percentile cuts (i.e., one-year growth percentile projections/trajectories). If the student's actual 2010 growth percentile exceeds the cut associated with proficient, then the student's one year growth was enough to reach proficient.[9] If the student reaches his/her achievement goal, then the accountability clock is reset for this student, with a new achievement goal of remaining proficient for the next 4 years.
- Next, the student's grade 3, 4, & 5 actual scores together with the 99 hypothetical one-year growth percentile projections/trajectories derived in the previous step are plugged into the coefficient matrix relating grade 7 with grade 3, 4, 5, & 6 prior achievement. This yields 99 percentile cuts (i.e., 2 year growth percentile projections/trajectories) for the student indicating what consecutive two-year 1st through 99th percentile growth (based upon the most recent student growth histories in the state) will lead to in terms of future achievement. Using the August 2007 accountability achievement targets (proficient by 7th grade for this student), 2 year growth sufficient to reach the target is determined and the

[8] Note, this is equivalent to just checking whether the student reached proficient in 2009 since the coefficient matrices used to produce the 99 percentile cuts are based on 2009 data.

[9] Note, this is equivalent to just checking whether the student reached proficient in 2010 since the coefficient matrices used to produce the 99 percentile cuts are based on 2010 data.

student's growth percentile is compared to this target. If the student's growth percentile exceeds this target then the student is deemed on track to reach proficient.

- No 3 or 4 year targets are utilized because they exceed the time-frame initially established for the student to reach proficient.

August 2011 Employing the coefficient matrices derived in the calculation of 2011 student growth percentiles:

- Because 2011 is the terminal year of the 4 year time frame established for the student to reach proficient the student is deemed to have grown sufficiently if they have reached proficient.
- No 2, 3, or 4 year targets are utilized because they exceed the accountability time-frame initially established for the student to reach proficient.

The complexity of the process just described is minimized by the use of the R software environment in conjunction with an open source software library SGP developed by the state Department of Education to calculate student growth percentiles and percentile growth projections/trajectories (R Development Core Team, 2010; Betebenner & Iwaarden, 2011). Every year, following the loading of the data into the state Department of Education data warehouse, student growth percentiles and percentile growth trajectories are calculated for each student. Once calculated, these values are easily used to make the yes/no determinations about the adequacy of each student's growth relative to his/her fixed achievement targets. These yes/no determinations are then used in aggregate to determine whether schools have met their AYP targets.

System-wide Growth and Achievement Charts

Operational work calculating student growth percentiles with state assessment data yields a large number of coefficient matrices derived from estimating Equation 2. These matrices, similar to a lookup table, "encode" the relationship between prior and current achievement scores for students in the norming group (usually an entire grade cohort of students for the state) across all percentiles and can be used both to qualify a student's current level growth as well as predict, based upon current levels of student progress, what different rates of growth (quantified in the percentile metric) will yield for students statewide.

When rates of growth necessary to reach performance standards are investigated, such calculations are often referred to as "growth-to-standard". These analyses serve a dual purpose in that they provide the growth rates necessary to reach these standards and also shed light on the standard setting procedure as it plays out across grades. To establish growth percentiles necessary to reach different performance/achievement levels, it is necessary to investigate what growth percentile is necessary to reach the desired performance level thresholds based upon the student's achievement history.

Establishing criterion referenced growth thresholds requires consideration of multiple future growth/achievement scenarios. Instead of inferring that prior student growth is indicative of future student growth (e.g., linearly projecting student achievement into the future based upon past rates of change), predictions of future student achievement are contingent upon initial student status (where the student starts) and subsequent rates of growth (the rate at which the student grows). This avoids fatalistic statements such as, "Student X is projected to be (not) proficient in three years" and instead promotes discussions about the different rates of growth necessary to reach future achievement targets: "In order that Student X reach/maintain proficiency within three years, she will have to demonstrate *nth* percentile growth consecutively for the next three years." The change

is phrascology is minor but significant. Stakeholder conversations turn from "where will (s)he be" to "what will it take?"

Parallel growth/achievement scenarios are more easily understood with a picture. Using the results of a statewide assessment growth percentile analyses, Figures 2 and 3 depict future growth scenarios in math and reading, respectively, for a student starting in third grade and tracking that student's achievement time-line based upon different rates of annual growth expressed in the growth percentile metric. The figures depict the four state achievement levels across grades 3 to 10 in shades of dark to light gray (e.g., unsatisfactory, partially proficient, proficient and advanced) together with the 2011 achievement percentiles (inner most vertical axis) superimposed in white. Beginning with the student's achievement starting point at grade 3 a grade 4 achievement projection is made based upon the most recent growth percentile analyses derived using prior 3rd to 4th grade student progress. More specifically, using the coefficient matrices derived in the quantile regression of grade 4 on grade 3 (see Equation 2), predictions of what 10th, 20th, 40th, 50th, 60th, 80th, and 90th percentile growth lead to are calculated. Next, using these seven projected 4th grade scores combined with the student actual 3rd grade score, 5th grade achievement projections are calculated using the most recent quantile regression of grade 5 on grades 3 and 4. Similarly, using these seven projected 5th grade scores, the 6 projected 4th grade scores with the students actual third grade score, achievement projections to the 6th grade are calculated using the most recent quantile regression of grade 6 on grades 3, 4, and 5. The analysis extends recursively for grades 6 to 10 yielding the *percentile growth trajectories* in Figures 2 and 3. The figures allow stakeholders to consider what consecutive rates of growth, expressed in growth percentiles, yield for students starting at different points.

Figure 2 depicts percentile growth trajectories in mathematics for a student beginning at the threshold between achievement level 1 and achievement level 2. i Based upon the *achievement* percentiles depicted (the white contour lines), approximately 7 percent of the population of 3rd graders rate as unsatisfactory. Moving toward grade 10, the percentage of unsatisfactory students increases dramatically to near 35 percent. The black lines in the figure represent seven different growth scenarios for the student based upon consecutive growth at a given growth percentile, denoted by the right axis. At the lower end, for example, consecutive 10th percentile growth leaves the student, unsurprisingly, mired in the unsatisfactory category. Consecutive 20th, 40th, 50th 60th, and 80th percentile growth also leave the student in the unsatisfactory category. This demonstrates how difficult probabilistically, based upon current rates of progress, it is for students to move up in performance level in math statewide. Considering a goal of reaching proficient or career and college readiness (next to top region) by 10th grade, a student would need to demonstrate growth percentiles consecutively in excess of 80 to reach this achievement target indicating how unlikely such a event currently is. In light of NCLB universal proficiency mandates, the growth necessary for non-proficient students to reach proficiency, absent radical changes to growth rates of students statewide, is likely unattainable for a large percentage of non-proficient students.

Figure 3 depicts percentile growth trajectories in reading for a student beginning at the partially proficient/proficient threshold in grade 3. In a normative sense, the performance standards in reading are less demanding than those in mathematics (particularly in the higher grades) with approximately 30 percent of students below proficient in grades 3 to 10. The black lines in the figure represent seven growth scenarios for the hypothetical student based upon consecutive growth at a the given growth percentile. Compared with the growth required in mathematics, more modest growth is required to maintain proficiency in reading. Typical growth (50th percentile growth) appears adequate for such a student to move up slightly into the proficiency category.

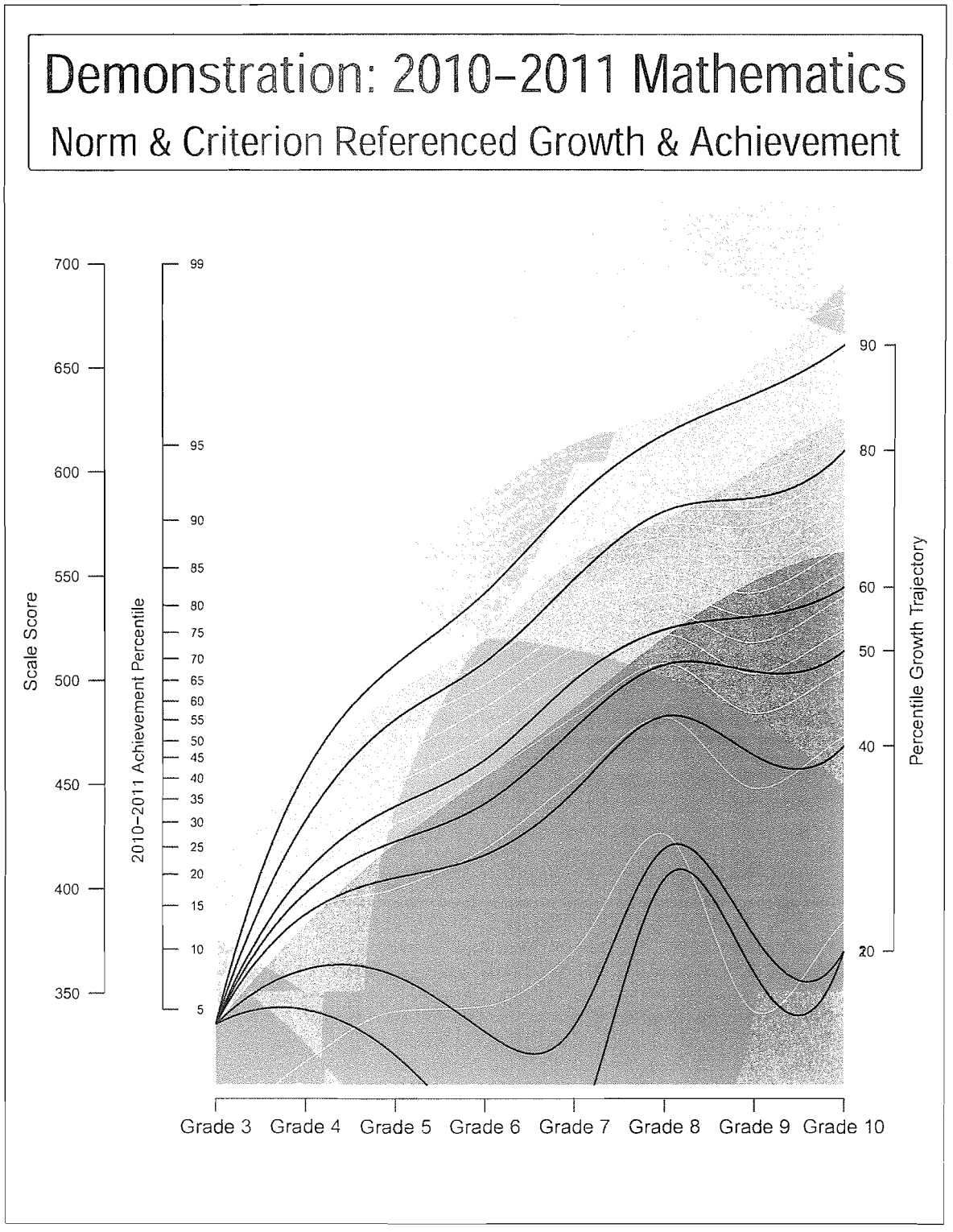


Figure 2: Growth chart depicting future mathematics achievement conditional upon consecutive 10th, 20th, 40th, 50th, 60th, 80th, and 90th percentile growth for a student beginning the third grade at the cutpoint between lowest and next to lowest achievement levels

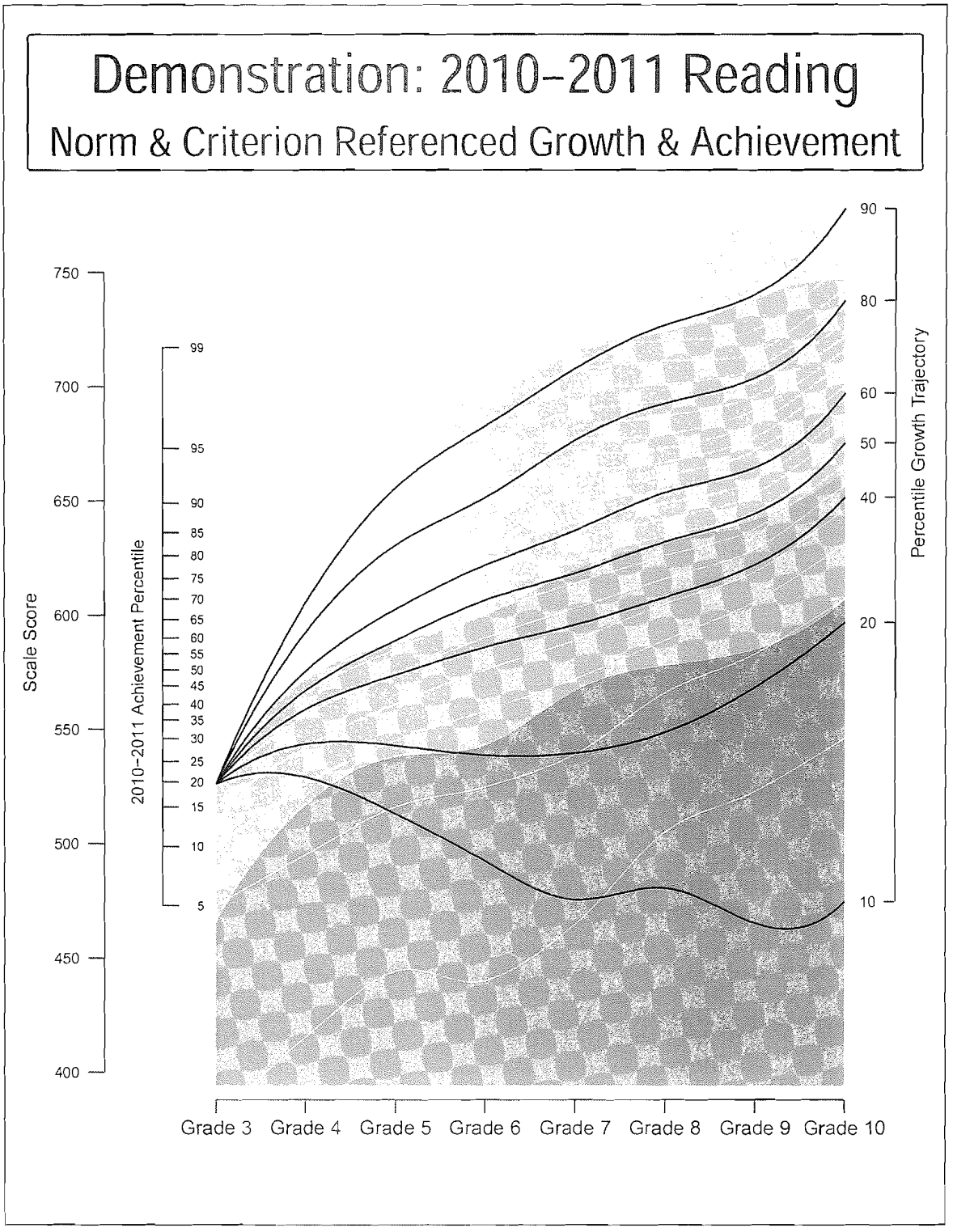


Figure 3: Growth chart depicting future reading achievement conditional upon consecutive 10th, 20th, 40th, 50th, 60th, 80th, and 90th percentile growth for a student beginning the third grade at the cutpoint between the achievement level 2 and achievement level 3

References

Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment for teachers. *Journal of Educational and Behavioral Statistics*, *29*(1), 37–65.

Betebenner, D. W. (2008). Toward a normative understanding of student growth. In K. E. Ryan & L. A. Shepard (Eds.), *The future of test-based educational accountability* (pp. 155–170). New York: Taylor & Francis.

Betebenner, D. W. (2009). Norm- and criterion-referenced student growth. *Educational Measurement: Issues and Practice*, *28*(4), 42–51.

Betebenner, D. W., & Iwaarden, A. V. (2011). SGP: An textupR package for the calculation and visualization of student growth percentiles [Computer software manual]. (R package version 0.4-0.0 available at `http://cran.r-project.org/web/packages/SGP/`)

Braun, H. I. (2005). *Using student progress to evaluate teachers: A primer on value-added models* (Tech. Rep.). Princeton, New Jersey: Educational Testing Service.

Briggs, D., & Betebenner, D. W. (2009, April). *The invariance of measurement of growth and effectiveness to scale transformation.* (Paper presented at the 2009 NCME Annual Conference, San Diego, CA.)

Harrell, F. E. (2001). *Regression modeling strategies.* New York: Springer.

Harris, D. N. (2007). *The policy uses and "policy validity" of value-added and other teacher quality measures* (Tech. Rep.). Princeton, NJ: Educational Testing Service.

Koenker, R. (2005). *Quantile regression.* Cambridge: Cambridge University Press.

Linn, R. L. (2003, July). *Accountability: Responsibility and reasonable expectations* (Tech. Rep.). Los Angeles, CA: Center for the Study of Evaluation, CRESST.

Linn, R. L., Baker, E. L., & Betebenner, D. W. (2002). Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001. *Educational Researcher*, *31*(6), 3–16.

Lord, F. M. (1975). The "ability" scale in item characteristic curve theory. *Psychometrika*, *20*, 299–326.

R Development Core Team. (2010). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Available from `http://www.R-project.org` (3-900051-07-0)

Raudenbush, S. W. (2004). What are value-added models estimating and what does this imply for statistical practice? *Journal of Educational and Behavioral Statistics*, *29*(1), 121–129.

Rubin, D. B., Stuart, E. A., & Zanutto, E. L. (2004). A potential outcomes view of value-added assessment in education. *Journal of Educational and Behavioral Statistics*, *29*(1), 103–116.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis.* New York: Oxford University Press.

Spellings, M. (2005, Nov). *Secretary Spellings announces growth model pilot* [Press Release]. U.S. Department of Education. (Retrieved August 7, 2006 from `http://www.ed.gov/news/pressreleases/2005/11/1182005.html`)

Wei, Y., & He, X. (2006). Conditional growth charts. *The Annals of Statistics*, *34*(5), 2069–2097.

Yen, W. M. (1986). The choice of scale for educational measurement: An IRT perspective. *Journal of Educational Measurement*, *23*, 299–325.

Yen, W. M. (2007). Vertical scaling and No Child Left Behind. In N. J. Dorans, M. Pommerich, & P. W. Holland (Eds.), *Linking and aligning scores and scales* (pp. 273–283). New York: Springer.

Student Growth Percentile Estimation

Calculation of a student's growth percentile is based upon the estimation of the conditional density associated with a student's score at time t using the student's prior scores at times $1, 2, \ldots, t-1$ as the conditioning variables. Given the conditional density for the student's score at time t, the student's growth percentile is defined as the percentile of the score within the time t conditional density. By examining a student's current achievement with regard to the conditional density, the student's growth percentile normatively situates the student's outcome at time t taking account of past student performance. The percentile result reflects the likelihood of such an outcome given the student's prior achievement. In the sense that the student growth percentile translates to the probability of such an outcome occurring (i.e., rarity), it is possible to compare the progress of individuals not beginning at the same starting point. However, occurrences being equally rare does not necessarily imply that they are equally "good." Qualifying student growth percentiles as "(in)adequate," "good," or as satisfying "a year's growth" is a standard setting procedure requiring external criteria (e.g., growth relative to state performance standards) combined with the wisdom and judgments of stakeholders.

Estimation of the conditional density is performed using quantile regression (Koenker, 2005). Whereas linear regression methods model the conditional mean of a response variable Y , quantile regression is more generally concerned with the estimation of the family of conditional quantiles of Y. Quantile regression provides a more complete picture of both the conditional distribution associated with the response variable(s). The techniques are ideally suited for estimation of the family of conditional quantile functions (i.e., reference percentile curves). Using quantile regression, the conditional density associated with each student's prior scores is derived and used to situate the student's most recent score. Position of the student's most recent score within this density can then be used to characterize the student's growth. Though many state assessments possess a vertical scale, such a scale is not necessary to produce student growth percentiles.

In analogous fashion to the least squares regression line representing the solution to a minimization problem involving squared deviations, quantile regression functions represent the solution to the optimization of a loss function (Koenker, 2005, p. 5). Formally, given a class of suitably smooth functions, $\mathcal{G}$, one wishes to solve

$$\arg\min_{g \in \mathcal{G}} \sum_{i=1}^{n} \rho_\tau(Y(t_i) - g(t_i)), \tag{1}$$

where t_i indexes time, Y are the time dependent measurements, and ρ_τ denotes the piecewise linear loss function defined by

$$\rho_\tau(u) = u \cdot (\tau - I(u < 0)) = \begin{cases} u \cdot \tau & u \geq 0 \\ u \cdot (\tau - 1) & u < 0. \end{cases}$$

The elegance of the quantile regression Expression 1 can be seen by considering the more familiar least squares estimators. For example, calculation of $\arg\min \sum_{i=1}^{n}(Y_i - \mu)^2$ over $\mu \in \mathbb{R}$ yields the sample mean. Similarly, if $\mu(x) = x'\beta$ is the conditional mean represented as a linear combination of the components of x, calculation of $\arg\min \sum_{i=1}^{n}(Y_i - x_i'\beta)^2$ over $\beta \in \mathbb{R}^p$ gives the familiar least squares regression line. Analogously, when the class of candidate functions $\mathcal{G}$ consists solely of constant functions, the estimation of Expression 1 gives the τth sample quantile associated with Y. By conditioning on a covariate x, the τth conditional quantile function, $Q_y(\tau|x)$, is given by

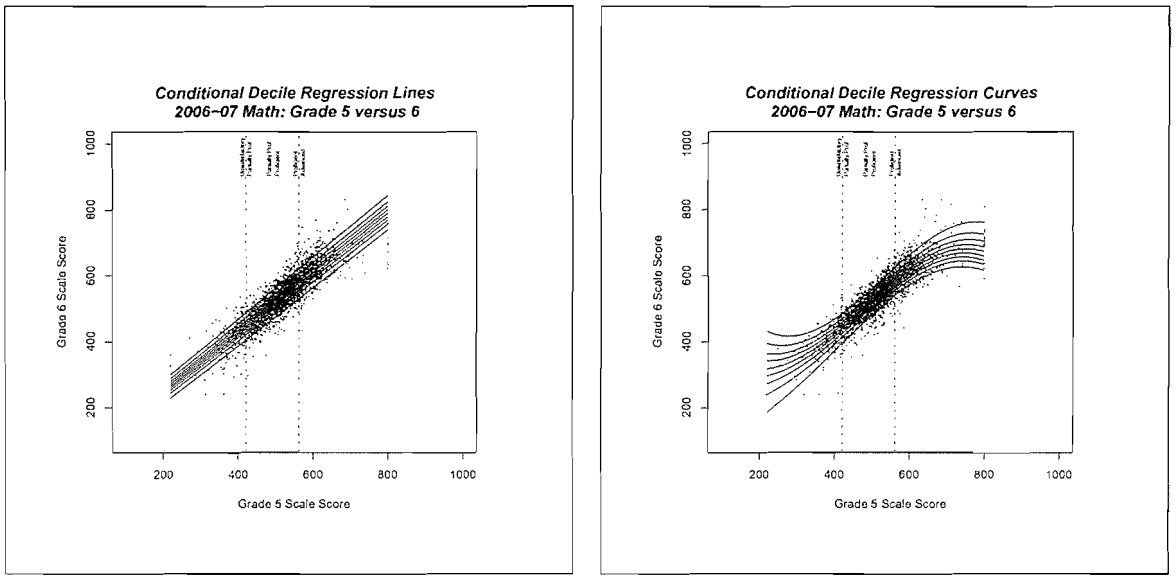


Figure 4: Linear and B-spline conditional deciles based upon bivariate math data, grades 5 and 6

$$Q_y(\tau|x) = \arg\min_{\beta\in\mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(y_i - x_i'\beta).$$

In particular, if $\tau = 0.5$, then the estimated conditional quantile line is the median regression line.[10]

Following Wei & He (2006), we parametrize the conditional quantile functions as a linear combination of B-spline cubic basis functions.. B-splines are employed to accommodate non-linearity, heteroscedasticity and skewness of the conditional densities associated with values of the independent variable(s). B-splines are attractive both theoretically and computationally in that they provide excellent data fit, seldom lead to estimation problems (Harrell, 2001, p. 20), and are simple to implement in available software.

Figure 4 gives a bivariate representation of linear and B-splines parametrization of decile growth curves. The assumption of linearity imposes conditions upon the heteroscedasticity of the conditional densities. Close examination of the linear deciles indicates slightly greater variability for higher grade 5 scale scores than for lower scores. By contrast, the B-spline based decile functions better capture the greater variability at both ends of the scale score range together with a slight, non-linear trend to the data.

Calculation of student growth percentiles is performed using R (R Development Core Team, 2010), a language and environment for statistical computing, with SGP package (Betebenner & Iwaarden, 2011). Other possible software (untested with regard to student growth percentiles) with quantile regression capability include SAS and Stata. Estimation of student growth percentiles is conducted using all available prior data, subject to certain suitability conditions. Given assessment scores for t occasions, $(t \geq 2)$, the τ-th conditional quantile for Y_t based upon $Y_{t-1}, Y_{t-2}, \ldots, Y_1$ is given by

$$Q_{Y_t}(\tau|Y_{t-1}, \ldots, Y_1) = \sum_{j=1}^{t-1} \sum_{i=1}^{3} \phi_{ij}(Y_j)\beta_{ij}(\tau), \tag{2}$$

[10]For a detailed treatment of the procedures involved in solving the optimization problem associated with Expression 1, see Koenker (2005), particularly Chapter 6.

where $\phi_{i,j}$, $i = 1, 2, 3$ and $j = 1, \ldots, t - 1$ denote the B-spline basis functions. Currently, bases consisting of 7 cubic polynomials are used to "smooth" irregularities found in the multivariate assessment data. A bivariate rendering of this is found is Figure 4 where linear and B-spline conditional deciles are presented. The cubic polynomial B-spline basis functions model the heteroscedasticity and non-linearity of the data to a greater extent than is possible using a linear parametrization.

Discussion of Model Properties

Student growth percentiles possess a number of attractive properties from both a theoretical as well as a practical perspective. Foremost among practical considerations is that the percentile descriptions are familiar and easily communicated to teachers and other non-technical stakeholders. Furthermore, implicit within the percentile quantification of student growth is a statement of probability. Questions of "how much growth is enough?" or "how much is a year's growth?" ask stakeholders to establish growth percentile thresholds deemed adequate. These thresholds establish growth standards that translate to probability statements. In this manner, percentile based growth forms a basis for discussion of rigorous yet attainable growth standards for all children supplying a normative context for Linn's (2003) existence proof with regard to student level growth.

In addition to practical utility, student growth percentiles possess a number of technical attributes well suited for use with assessment scores. The more important theoretical properties of growth percentiles include:

Robustness to outliers Estimation of student growth percentiles are more robust to outliers than is traditionally the case with conditional mean estimation. Analogous to the property of the median being less influenced by outliers than is the median, conditional quantiles are robust to extreme observations. This is due to the fact that influence of a point on the τ-th conditional quantile function is not proportional (as is the case with the mean) to the distance of the point from the quantile function but only to its position above or below the function (Koenker, 2005, p. 44).

Uncorrelated with prior achievement Analogous to least squares derived residuals being uncorrelated with independent variables, student growth percentiles are not correlated with prior achievement. This property runs counter to current multilevel approaches to measuring growth with testing occasion nested within students (Singer & Willett, 2003). These models, requiring a vertical scale, fit lines with distinct slopes and intercepts to each student. The slopes of these lines represent an average rate of increase, usually measured in scale score points per year, for the student. Whereas a steeper slope represents more learning, it is important to understand that using a normative quantification of growth, one cannot necessarily infer that a low achieving student with a growth percentile of 60 "learned as much" as a high achieving student with the same growth percentile. Growth percentiles bypass questions associated with magnitude of learning and focus on normatively quantifying changes in achievement.

Equivariance to monotone transformation of scale An important attribute of the quantile regression methodology used to calculate student growth percentiles is their invariance to monotone transformations of scale. This property, denoted by Koenker (2005) as *equivariance to monotone transformations* is particularly helpful in educational assessment where a variety of scales are present for analysis, most of which are related by some monotone transformation. For example, it is a common misconception that one needs a vertical scale in order to calculate growth. Because vertical and non-vertical scales are related via a monotone transformation, the student growth percentiles do not change given such alterations in the underlying scale.

This result obviates much of the discussion concerning the need for a vertical scale in measuring growth.[11]

Formally, given a monotone transformation h of a random variable Y,

$$Q_{h(Y)|X}(\tau|X) = h(Q_{Y|X}(\tau|X)).$$

This result follows from the fact that $\Pr(T < t|X) = \Pr(h(T) < h(t)|X)$ for monotone h. It is important to note that equivariance to monotone transformation does not, in general, hold with regard to least squares estimation of the conditional mean. That is, except for affine transformations h, $E(h(Y)|X) \neq h(E(Y|X))$. Thus, analyses built upon mean based regression methods are, to an extent, scale dependent.

[11] As already noted with regard to pediatrics, the existence of nice "vertical" scales for measuring height and weight still leads to observed changes being normed.