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

MEETING MINUTES

Meeting Date: December 13, 2013
Meeting Time: 10:00 A.M.
Meeting Place: State House, 200 W. Washington St., 233
Meeting City: Indianapolis, Indiana
Meeting Number: 8

Members Present: Superintendent Glenda Ritz, Co-Chairperson; Dr. Steve Yager, Co-Chairperson; Steve Baker; Derek Redelman; Dr. Jim Snapp; Robert Lugo; Claire Fiddian-Green; Dr. Shane Robbins; Jessica Dunn Feeser; Scott Bess; Dr. E. Ric Frataccia; Michele Walker.

Members Absent: Cheryl Ramsey; Keith Gambill; Sheila Seedhouse; Melanie Park; Casandra McLeod.

Co-Chairs Ritz and Yager called the meeting to order at 10:08 a.m. They shared the accountability language adopted by the State Board of Education (attached as Exhibit 1). They indicated that an Addendum to the Memorandum of Understanding creating the Accountability System Review Panel had been prepared. When signed by all parties, the Addendum will extend the Panel through December 31, 2014. The Panel is requested to continue its work in accordance with the original agreement and to review and enhance its recommendations. They provided the Panel with a summary of the Panel’s projected

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.
Debbie Daley, Assistant Director of Information Services for the Department of Education, provided a status report to the Panel and summarized the results of the A-F performance grading of schools for the 2012-2013 school year. Her presentation was summarized in a PowerPoint presentation. (Attached as Exhibit 3).

Damien Betebenner, National Center for the Improvement of Educational Assessment, presented an explanation of his work concerning student growth models, student growth percentiles, student growth projections and targets, and the applicability of these measures to assessing the performance of schools. He summarized his presentation in a PowerPoint presentation (Attached as Exhibit 4). He focused on issues related to a trajectory model of measuring student performance. A trajectory model is a growth model based on a student's previous test scores compared to proficiency at a later point in time. To create a growth trajectory, a state must determine the gap between a student's current achievement level and proficiency. From there, a linear path is created that closes that achievement gap over time. School performance is determined by aggregating the statistics for each individual student.

The Department of Education provided the members with handouts explaining growth model elements, describing growth model value questions, suggesting a hybrid approach toward measuring student growth that includes trajectory growth model and value table growth model components, and providing the resume of Derek C. Briggs (a national expert that the Department has consulted in addition to Damien Betebenner). (Attached as Exhibits 5, 6, and 7).

Molly Chamberlin, Chief Accountability and Assessment Officer for the Center for Education and Career Innovation, provided the Panel with a Sample Status Improvement Value Table. (Attached as Exhibit 8). These tables are used in the value table growth model for assessing school performance. Individual student growth metrics are established and points are awarded to individual students based on their growth between performance categories (e.g., as used in Exhibit 8: did not pass-1; did not pass-2; did not pass-3; pass-1; pass-2; pass plus-1; pass plus-2; pass plus-3). Student performance from year-to-year is studied, and each student is assigned a growth metric based on the relationship between last year’s performance level and the current year’s performance level. School performance is determined by aggregating the statistics for each individual student. The Panel members broke into small groups to discuss relative values that should be assigned to various levels of improvement by students.

The co-chairs adjourned the meeting at 2:30 p.m. and indicated that the Panel might next meet in February after the results of some statistical analysis of various growth models is completed.
“The State Board of Education adopts the Accountability System Review Panel's recommendations for new accountability categories, and affirms the Panel’s vision for the framework.

The State Board of Education further recognizes that a validation and statistical analysis process informed by technical experts regarding the indicators, formulas and metrics that support the categories as they are built and validated through the beta testing period of the new model design must occur. This work will lead to follow up recommendations by the Panel to the Board.

Board staff and Department staff will collaborate with technical experts to inform the work of the Panel and ultimately the Board.
Projected scope of work for new A-F implementation, using panel recommendations:

- A-F panel meets to define scope of future work and role of the panel in rolling out A-F recommendations; A-F panel reviews Performance data; define guidelines for Categorical Improvement Value table.
  - SBOE to review.

- IDOE (with assistance from outside experts, if necessary) preliminarily sets cut scores for additional categories and creates preliminary value table; IDOE engages US DOE for NCLB waiver renewal.
  - A-F panel to review/adjust value table (for purpose of running simulations); review growth to target formulas.
  - SBOE to review.

- IDOE works with A-F panel to finalize definitions of college and career readiness and graduation rate improvement.
  - SBOE to review.

- IDOE runs simulations based on performance and growth as preliminarily created with support from the panel.
  - SBOE to review.

- IDOE shares high-level results with A-F panel and selected experts.
  - SBOE to review.

- IDOE approves conceptual A-F model; begins rulemaking process.
  - IDOE works with practitioners and outside experts to create, pilot, and implement official additional categories for ISTEP+.
  - IDOE works with statistical experts and practitioners to analyze results of simulations, including reliability, stability, and validity of model.
  - IDOE begins work on implementation of business logic and programming for running final model.

- A-F panel to review results.
  - SBOE to review.

- IDOE shares preliminary ("practice") A-F grades with schools (using 2013-2014 and prior years' data).
  - IDOE makes adjustments as necessary as a result of school feedback, etc.

- A-F panel to review results.
  - SBOE to review (Final rule adoption; December 2014)

- Preliminary new "official" A-F results are released to schools.
  - A-F appeals are received and reviewed.

- Final new A-F grades are released (October 2015).
Overview of Progress Since Last Meeting:

- The SBOE has affirmed the framework recommended by the Accountability Panel. The Board vote carried to adopt the framework concept defined by the Accountability Panel, including:
  - **Scale**: 100 points
  - **Category Placement**: 90, 80, 70, 60
  - **Measure Domains**: Performance and Growth (May be adjusted as outcome of statistical analysis)

- The DOE and CECI have engaged in conversations with 2 experts in the areas of Accountability and Growth:
  - **Damian Betebenner**: Damian Betebenner is an associate at The National Center for the Improvement of Educational Assessment in Dover, New Hampshire. His areas of specialization include applied statistics, particularly with regard to large scale data analysis. His current research interests center around longitudinal data analysis, specifically with regard to state and federal performance mandates.
  - **Derek Briggs**: Derek Briggs is Associate Professor of quantitative methods and policy analysis at the University of Colorado at Boulder. His research agenda focuses upon building sound methodological approaches for the valid measurement and evaluation of growth in student achievement.
Data Overview:

- The Accountability Panel will continue to engage in recommendations throughout the A-F Rule making process.

- As data becomes available, it will be shared with Accountability Panel for review and formula refinement; then it will be shared with the State Board of Education.

- An estimated scope of work has been created to outline the process and engagement of each group.

Data Availability:

- Data for the two Domains will be shared individually, then combined for overall impact analysis.

- Performance
  - Data is currently available for 2012 and 2013 school years.
  - Summary information is provided in this presentation; however, de-identified school data is also available.
  - Please note that 2013 data used for this review is preliminary information and may adjust as data is finalized.

- Growth
  - Data for Students Meeting Targeted Growth will be available Late Winter
  - Data for Categorical Status Improvement will be available Early Spring
Performance Data Calculations:

Total Points: \( (\text{ELA Points} \times 0.2) + (\text{Math Points} \times 0.2) + (\text{Grad Points} \times 0.3) + (\text{CCR Points} \times 0.3) \)

Sum of Weights

Example Elem/Middle: \( \frac{(100 \times 0.2) + (75 \times 0.2)}{0.4} = 87.5 \)

Example HS: \( \frac{(75 \times 0.2) + (75 \times 0.2) + (100 \times 0.3) + (100 \times 0.3)}{1.0} = 90.0 \)

Performance Summary:

* Performance data breakdown by point ranges.

<table>
<thead>
<tr>
<th>POINT RANGE</th>
<th>TOTAL CNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>90.0-100.0</td>
<td>660</td>
</tr>
<tr>
<td>80.0-89.9</td>
<td>880</td>
</tr>
<tr>
<td>70.0-79.9</td>
<td>802</td>
</tr>
<tr>
<td>60.0-69.9</td>
<td>355</td>
</tr>
<tr>
<td>0.0-59.9</td>
<td>123</td>
</tr>
</tbody>
</table>

\* Does not include Growth data

2013 with New Model; Performance Only

<table>
<thead>
<tr>
<th>POINT RANGE</th>
<th>TOTAL CNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>90.0-100.0</td>
<td>588</td>
</tr>
<tr>
<td>80.0-89.9</td>
<td>468</td>
</tr>
<tr>
<td>70.0-79.9</td>
<td>470</td>
</tr>
<tr>
<td>60.0-69.9</td>
<td>167</td>
</tr>
<tr>
<td>0.0-59.9</td>
<td>125</td>
</tr>
</tbody>
</table>

\* Does not include Growth data

2012 with 2012 A-F Rules

<table>
<thead>
<tr>
<th>POINT RANGE</th>
<th>TOTAL CNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.51-4.00</td>
<td>856</td>
</tr>
<tr>
<td>3.50-3.50</td>
<td>420</td>
</tr>
<tr>
<td>2.00-2.99</td>
<td>425</td>
</tr>
<tr>
<td>1.00-1.99</td>
<td>242</td>
</tr>
<tr>
<td>0.00-0.99</td>
<td>144</td>
</tr>
</tbody>
</table>

\** Includes both Growth and Performance data
Performance Summary:

* Performance data by previous model designation

2012 with New Model; Performance Only

<table>
<thead>
<tr>
<th>ACCOUNTABILITY MODEL CODE</th>
<th>POINT RANGE</th>
<th>TOTAL CNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMS</td>
<td>90.0-100.0</td>
<td>372</td>
</tr>
<tr>
<td>EMS</td>
<td>80.0-89.9</td>
<td>651</td>
</tr>
<tr>
<td>EMS</td>
<td>70.0-79.9</td>
<td>385</td>
</tr>
<tr>
<td>EMS</td>
<td>60.0-69.9</td>
<td>119</td>
</tr>
<tr>
<td>EMS</td>
<td>0.0-59.9</td>
<td>66</td>
</tr>
</tbody>
</table>

**Does not include Growth data**

2012 with 2012 A-F Rules

<table>
<thead>
<tr>
<th>ACCOUNTABILITY MODEL CODE</th>
<th>POINT RANGE</th>
<th>TOTAL CNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMS</td>
<td>3.51-4.00</td>
<td>670</td>
</tr>
<tr>
<td>EMS</td>
<td>3.00-3.50</td>
<td>294</td>
</tr>
<tr>
<td>EMS</td>
<td>2.00-2.99</td>
<td>202</td>
</tr>
<tr>
<td>EMS</td>
<td>1.00-1.99</td>
<td>201</td>
</tr>
<tr>
<td>EMS</td>
<td>0.00-0.99</td>
<td>126</td>
</tr>
</tbody>
</table>

**Includes both Growth and Performance data**
Growth Models, Student Growth Percentiles, Student Growth Projections/Targets and Accountability

Damian W.Betebenner

National Center for the Improvement of Educational Assessment
Dover, NH

Indianapolis, Indiana December 13th, 2013
Our work in the last few years investigating growth/value-added models has led us to believe that there are two general types of growth based accountability.

- Norm-referenced growth based accountability.
- Criterion-referenced growth based accountability.

These two approaches to growth impart different levels of stringency.

Norm-referenced Is the group of students growing better/worse (on average) than statistical expectation?

Criterion-referenced Is each student in the group growing at a sufficient rate (i.e., targeted growth) to reach/maintain a desired level of achievement (e.g. Pass or Pass +)?
Criterion-referenced student growth accountabilities

- Former Secretary of Education Margaret Spellings initiated the Growth Model Pilot Program allowing states to include growth as a means of satisfying AYP.
- The "Bright Line Principle" of all students being proficient by 2014 remained.
- 15 states were accepted as part of the pilot. None saw a significant improvement in their AYP numbers because of the growth-to-proficiency goals.
- Indiana's proposed growth-to-standard approach put forward in the Report of the Accountability System Review Panel is a trajectory gain score model utilized by many of the 15 states as part of the growth model pilot program.
- To my knowledge, none of the states adopting this approach outlined by Indiana continue to use that model.
The reality in all state education systems I have worked with is that students who start behind tend to stay behind (Did Not Pass) and students who start ahead tend to stay ahead (Pass or Pass +).

Targeted growth reflects this unfortunate reality: Schools serving the lower achieving students will likely have low percentages of students reaching targeted growth and schools serving high achieving students will likely have high percentages of students reaching targeted growth.

Along, growth to proficient provides an impoverished view of student growth.

This is not to say that targeted growth is unimportant. I must be balanced with a norm-referenced view of student progress.

The SGP methodology's norm and criterion-referenced components attempt to strike that balance allowing stakeholders to understand: What is?, What should be? and What is Reasonable?
<table>
<thead>
<tr>
<th>Grade 3</th>
<th>Grade 4</th>
<th>Grade 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>2013</td>
<td>2014</td>
</tr>
<tr>
<td>Scale Score</td>
<td>286</td>
<td>357</td>
</tr>
<tr>
<td>Achievement Level</td>
<td>Did Not Pass</td>
<td>Did Not Pass</td>
</tr>
<tr>
<td>Growth Percentile</td>
<td>32</td>
<td>Low</td>
</tr>
</tbody>
</table>
Math

Achievement

ISTEP+ Math
Scale Score

Growth Percentiles

Typical 35th - 64th
High 65th - 99th
Low 1st - 34th

Achievement Level Did Not Pass Pass

Achievement Target Pass (485)/Pass + (532)

Grade 3 2012
398

Grade 4 2013
460

Grade 5 2014

Growth Level

High 65th - 99th
Typical 35th - 64th
Low 1st - 34th

Growth Percentile

Typical 64

Growth Target

Keep Up (43)/Move Up (91)
Recommendations

Consistent with the duties charged to the Accountability System Review Panel the state should perform thorough impact analyses on the outcomes of the results including:

- How does growth relate to achievement at both the individual and group level?
- Do schools serving lower achieving students demonstrate markedly different growth than schools serving high achieving students?
- Do growth results differ by school level (elementary, middle, high)?
- Do growth results differ by important demographic subgroups in the state? Are any differences observed consistent with expectations and acceptable?
- To what extent do the results of the growth model change the outcomes of the final A-F ratings based just upon status?
- Does growth answer questions of relevance to the various stakeholder groups and are those questions consistent with larger policy goals?
- Is the impact of growth on A-F outcomes coherent with policy?
Why do we want to look at Growth?

- Movement to proficiency
- Movement beyond proficiency
- Movement relative to proficiency
- Growth for ALL students
- Growth matters MOST
- Accuracy
- Growth honors successful effort
- Growth which allows us to differentiate from starting point

What do we want to do with it?

- Reward
- Recognize
- Acknowledge
- Use for instruction (Growth reporting for individual students)
- Replicate
- Communicate (to parents, teacher, administrators and school)
- Use to inform curriculum
- Comprehensive FOCUS on student/schools

Growth Model Elements

- Align growth with grades
- Categories based on fairness (expand DO NOT pass and Pass Plus categories)
- Expected growth should be differentiated
- Considerations include student background
- Hybrid Model (use portions of different models)
- Backwards progress counted adversely
- Maintain or Move up
- Encourages excellence
- Reward performance over Zip Code
- Do Not LOWER expectations
Key Policy Questions Answered by Proposed Accountability Framework

**Performance:**
- How are students performing on state assessments?
- How are students performing on measures of college readiness, including high school graduation and other key metrics?

**Growth:**
- How much progress are students making toward a target ("trajectory growth")
- Are students maintaining or changing their levels of proficiency ("categorical growth")
Growth: Value Questions to Answer and Work to be Done

- How much progress are students making toward a target ("trajectory growth")
  
  VALUE Question: How much targeted growth is sufficient?
  
  ANSWER: Define sufficient growth targets

- Are students maintaining or changing their levels of proficiency ("categorical growth")
  
  VALUE Question: What categorical changes (or maintenance) do we want to see most? What status changes (or maintenance) concern us?
  
  ANSWER: Set priority levels for value table

TODAY'S FOCUS:

- Are students maintaining or changing their levels of proficiency?
  
  VALUE Question: What category changes (or maintenance) do we want to see most? What category changes (or maintenance) concern us?
  
  ANSWER: Set priority levels for value table
**VALUE TABLE:**

- Groups of four

- For each category change, identify whether it is evaluated at:
  
  - +2 (exceeds expectations)
  - +1 (meets expectations)
  - 0 (neutral)
  - -1 (below expectations)
  - -2 (far below expectations)

- NOTE: After cut score setting, there may not be 8 categories, but value statements can be modified to meet set categories (do not need to wait until cut scores are set to identify what we value)
CURRICULUM VITAE
Derek C. Briggs

University of Colorado, School of Education, 249 UCB, Boulder, CO, 80309
Phone: (303) 492-6320 • E-mail: derek.briggs@colorado.edu
http://www.colorado.edu/education/faculty/derekbriggs/

RESEARCH EXPERTISE

Learning Progressions, Vertical Scaling, Growth Models, Test Validity, Large-Scale Assessment, Diagnostic Assessment, Item Response Theory, Value-Added Models, Causal Inference

PROFESSIONAL EXPERIENCE

2013-present	Professor, Research and Evaluation Methodology, School of Education, University of Colorado at Boulder

2009-2013	Associate Professor, Research and Evaluation Methodology, School of Education, University of Colorado at Boulder

2003-2009	Assistant Professor, Research and Evaluation Methodology, School of Education, University of Colorado at Boulder

2002-2003	Visiting Postdoctoral Scholar, Berkeley Evaluation and Assessment Research Center, Graduate School of Education, University of California at Berkeley

1998-2002	Graduate Student Researcher, National Center for Research on Vocational Education (98-00); Berkeley Evaluation and Assessment Research Center (00-02), Graduate School of Education, University of California at Berkeley

1996-1997	Research Assistant, National Association for State Community Services Programs, Washington, D.C.

1993-1996	Assistant Analyst, Macroeconomic Analysis Division, Congressional Budget Office, Washington, D.C.

EDUCATION

Ph.D. University of California, Berkeley 2002
Education, Quantitative Methods and Evaluation

Dissertation: SAT Coaching, Bias and Causal Inference
Chair: Mark Wilson

B.A. Carleton College 1993
Economics
DEREK C BRIGGS

HONORS/AWARDS

- Outstanding Reviewer Award, Journal of Educational and Behavioral Statistics, 2013
- University of Colorado Provost’s Award for Faculty Achievement, 2012
- Outstanding Reviewer Award, Educational Researcher, 2012
- AERA Division D Mary Catherine Ellwein Outstanding Dissertation Award, 2004
- UC Berkeley Graduate School of Education Commencement Address, 2003
- Educational Testing Service Summer Associate, 2002
- RAND Summer Associate, 2000
- Graduate School of Education Research Centers Coordinator, 1999-2001
- National Center for Educational Statistics, NELS-88 Training Fellowship, Summer 1999
- Graduate School of Education Regents Fellowship, 1999-2002
- Berkeley Evaluation and Assessment Research Center Coordinator, 1998-99
- Departmental Distinction in Economics, Carleton College, 1993

TEACHING EXPERIENCE

- EDUC 8230: Quantitative Methods in Educational Research I
- EDUC 7316: Intermediate Statistical Methods
- EDUC 7326: Experimental Design
- EDUC 7386: Educational Evaluation
- EDUC 8710: Measurement in Survey Research
- EDUC 8720: Advanced Topics in Measurement

GRANTS


Pearson. Application of a Diagnostic Classification Model to Learning Progressions in Science. $43,513 (PI) (2/1/13 – 1/31/14)

Denver Public School District. Student Outcomes Analysis in Support of Educator Effectiveness Evaluation. $65,028 (PI) (6/1/12-5/31/13)

University of Colorado Department of Continuing Education. An Evaluation of CU’s Online Summer Session Courses. $23,000 (PI) (6/1/12-1/31/13)
DEREK C BRIGGS

Institute of Education Sciences. An Exploration of Novice Teachers' Core Competencies: Impacts on Student Achievement, and Effectiveness of Preparation. $987,152 (co-PI) (03/01/2012- 02/28/2015)

Colorado Department of Higher Education. Analysis of Educator Preparation and K-12 Placement in Colorado. $15,000 (PI) (January 2012 to June 2012)


National Science Foundation. Undergraduate Science Course Innovations and their Impact on Student Learning. $121,000 (PI) (January 2007 to January 2008).

University of Colorado. Junior Faculty Development Grant. Vertical Scaling in Value-Added Models for Student Learning. $4,000 (July 2006 to August 2006).

National Science Foundation. Learning Assistant model for Teacher Education in Science and Technology. $2,500,000. (Co-PI) (September 2006 to September 2011).


PUBLICATIONS

Refereed Journal Articles


Refereed Book Chapters


Reports


DEREK C BRIGGS


Commentary and Reviews

## SAMPLE Status Improvement Value Table

<table>
<thead>
<tr>
<th>Previous Year Level</th>
<th>Did Not Pass-1</th>
<th>Did Not Pass-2</th>
<th>Did Not Pass-3</th>
<th>Pass-1</th>
<th>Pass-2</th>
<th>Pass Plus-1</th>
<th>Pass Plus-2</th>
<th>Pass Plus-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pass Plus-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pass Plus-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pass Plus-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pass-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pass-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did Not Pass-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did Not Pass-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did Not Pass-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>