Delivering on the Promise:
An Overview of Value-Added

May 2011
Who is Battelle for Kids?

- Not-for-profit educational improvement organization
- We believe in:
  - **The Right People**
    - Having highly effective educators throughout the system to maximize student opportunities.
  - **The Right Measures**
    - Measuring educator, employee and program effectiveness
  - **The Right Practices**
    - Researching and supporting effective educational practices
  - **The Right Messages**
    - Engaging stakeholders for strategic improvement and managing change
Harnessing the Power of Data
National Context

“With increased accountability, American schools and the people who work in them are being asked to do something new – to engage in systemic, continuous improvement in the quality of the educational experience of students and to subject themselves to the discipline of measuring their success by the metric of students’ academic performance”

– Richard Elmore, Bridging the Gap Between Standards and Achievement
Student Academic Outcome Measures

- **Achievement** – Measuring status at a single point in time
  - **Passage Rates** – the percentage of students reaching a standard
  - **Performance Index** – assigning a value for each student reaching various levels of achievement

- **Growth Models** – Measures of changes in student performance
  - **Improvement Models** - measures the change in the percent of students meeting a certain benchmark
  - **Simple Growth Models** - compares students’ scaled scores this year to their scores last year; or change in performance categories
  - **Growth to Proficiency Models** - are designed to show whether students are on a linear track to meet standards some time in the future
  - **Value-Added Models** – estimates the influence of teachers, schools, and programs on the academic growth of students by using rigorous statistical models.
The Power of Two: Achievement & Progress

Achievement
- Compares students’ performance to a standard
- Does not factor in students’ background characteristics
- Measures students’ performance at a single point in time
- Critical to students’ post-secondary opportunities

Growth
- Compares students’ performance to their own prior performance
- May factor in students’ background characteristics
- Measures students’ academic growth between two points in time
- Critical to ensuring students’ future academic success

A more complete picture of student learning
The Power of Two: Achievement & Progress

- School A
- School B
- School C
- School D
- School E
- School F
- School G
- School H
- School I
- School J
- School K
- School L
- School M
- School N
- School O
- School P

75% Pass

High Achievement
- Low Progress

Average Progress

High Progress

Low Achievement

Low Progress

Average Progress

High Progress

Low Progress

© 2010, Battelle for Kids
### Reeves Leadership Model Adaptation

(ANTECEDENTS OF EXCELLENCE AND ORGANIZATIONAL RESULTS)

<table>
<thead>
<tr>
<th>LUCKY</th>
<th>LEADING</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Achievement</td>
<td>High Achievement</td>
</tr>
<tr>
<td>Low Progress</td>
<td>High Progress</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LOSING</th>
<th>LEARNING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Achievement</td>
<td>Low Achievement</td>
</tr>
<tr>
<td>Low Progress</td>
<td>High Progress</td>
</tr>
</tbody>
</table>
Understanding Value-Added

- Technical Basis
  - Oak Tree Analogy
  - Basic Example
  - Data Considerations and Quality
- Harnessing the Power of Value-Added
  - Reporting – What is happening?
  - Analysis – Why is this happening?
  - Responding – How do we improve?
An Analogy: Explaining the concept of Value-Added by evaluating the performance of two gardeners

• For the past year, these gardeners have been tending to their oak trees trying to maximize the height of the trees.

• Each gardener used a variety of strategies to help their own tree grow... which of these two gardeners was more successful with their strategies?
To measure the performance of the gardeners, we will measure the height of the trees at the end of summer.

- Using this method, who might we infer to be the more effective gardener?

This method is analogous to using an Attainment Model.
Who was more effective at Gardening?

• By finding the difference between these heights, we can determine how many inches the trees grew during the year of gardener’s care.

• Based on this measure, **who might we infer is the more effective Gardener?**

This is analogous to a **Simple Growth Model**, also called **Gain**.
... but this **simple growth** result does not tell the whole story either.

- We do not yet know how much of this growth was due to the strategies used by the gardeners themselves.
- This is an “apples to oranges” comparison.
- For our oak tree example, three environmental factors we will examine are: **Rainfall**, **Soil Richness**, and **Temperature**.
Predicting Growth

The predicted height for trees in Oak A’s conditions is 59 inches.

The predicted height for trees in Oak B’s conditions is 74 inches.
Comparing the actual height of the trees to our predictions.

Oak A’s actual height of 61 inches is 2 inches more than we predicted. We attribute this above-average result to the effect of Gardener A.

Oak B’s actual height of 72 inches is 2 inches less than we predicted. We attribute this below-average result to the effect of Gardener B.
Now do this for all trees under each gardener’s care.

Apply statistical techniques to look for significance.

This is analogous to a Value-Added Model.
## Growth – Simple Subtraction and Average

<table>
<thead>
<tr>
<th>Student</th>
<th>Prior Year</th>
<th>Current Year</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>100</td>
<td>130</td>
<td>+30</td>
</tr>
<tr>
<td>Tom</td>
<td>75</td>
<td>95</td>
<td>+20</td>
</tr>
<tr>
<td>Emily</td>
<td>110</td>
<td>141</td>
<td>+31</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>115</strong></td>
<td><strong>132</strong></td>
<td><strong>+17</strong></td>
</tr>
</tbody>
</table>

Imagine 3 students and following their scores year over year.

In this example, all students are demonstrating growth by subtracting their scores from year-to-year and averaging for overall growth of +27 points.

Each student in this example shows growth!

Yeah! Celebrate! Or wait…
Basic Predictions and Observations

<table>
<thead>
<tr>
<th>Student</th>
<th>Prior Year</th>
<th>Predicted</th>
<th>Observed</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>100</td>
<td>125</td>
<td>130</td>
<td>+5</td>
</tr>
<tr>
<td>Tom</td>
<td>75</td>
<td>101</td>
<td>95</td>
<td>-6</td>
</tr>
<tr>
<td>Emily</td>
<td>110</td>
<td>137</td>
<td>141</td>
<td>+4</td>
</tr>
</tbody>
</table>

Value-Added +1

Imagine all students with similar performance (testing) histories and using those students’ prior performance results to determine a predicted score.

• E.g., the population of students like **John** who scored **100** last year, on average, score **125** this year.

Imagine using tests from other subjects as predictors for John’s performance.

• Do you think Reading can be predictive of Math?
• Could Math and Reading have predictive power in Science?

The differences between actual and predicted scores are used to compute “value-added.”
Commonly Examined Student Factors

1. Prior Test Scores?
2. School/Grade Level?
3. Gender?
4. Race/Ethnicity?
5. Low-Income Status?
6. ELL Status?
7. Special Education Status?
8. Homelessness?
9. Student Attendance?
10. Credit status?
11. Co-teaching?

Value-Added models can determine if there is a relationship between factors we may think describe variability in student performance and those that do describe variability in student performance.
Data Quality & Decisions

- Data quality of the inputs into value-added models extremely important
  - Student-Teacher Data Link
    - Snapshots and intra-year mobility
    - Multiple teachers
  - Full Academic vs. Non-Full Academic Year
  - ELL Status
    - Policy
    - Proficiency
  - Free & Reduced Lunch Status
    - Stigma
Why is linkage important?

Discrepancies and Variation in Data

Working with hundreds of districts and thousands of schools across the country, Battelle for Kids has learned that virtually every school has required significant revisions or modifications to teaching assignments, number of students taught, the mobility of students and the instructional responsibility attributed to a single teacher.

In spring 2010, more than 125,000 rosters were verified by educators in South Carolina, Texas, Ohio and Oklahoma. Recent analyses of linkage results from schools across the country yield alarming results, including:

<table>
<thead>
<tr>
<th>Large Urban Schools</th>
<th>Small Rural Schools</th>
<th>Nearly</th>
<th>Nearly</th>
</tr>
</thead>
<tbody>
<tr>
<td>33%</td>
<td>11%</td>
<td>1 in 10</td>
<td>40%</td>
</tr>
<tr>
<td>of students linked to more than one teacher in a subject</td>
<td>of students linked to more than one teacher in a subject</td>
<td>of the teachers in the source data had incomplete or incorrect content area association</td>
<td>of rosters were changed from district-reported data through the teacher linkage process</td>
</tr>
</tbody>
</table>
Summary of Value-Added

- Develops predictions of students’ achievement based on
  - Multiple prior test scores in the same subject or
  - Subjects shown to have predictive power in another subject
- Often includes other school or student factors to more reliably estimate the influence of educators on student learning.
  - E.g., demographics, student attendance
- Applies statistical techniques to handle such things as missing test data, measurement error and the sharing of instructional responsibilities by multiple educators
- Produces comparative results of schools or classrooms that perform above, at, or below predictions
Harnessing the Power of Value-Added

- Value-added reports tell you **what is happening**, but not **why**.
- With knowledge of **what is happening**, various techniques can be employed to determine **why** we are producing the results.
  - Reflective examination & analysis (e.g., Focus)
  - Diagnostic observation (e.g., Instructional Rounds)
- Once we know why, we can **plan for improvement**.
Interpretation of Data

- More than expected growth
- Expected growth
- Less than expected growth
What results are we producing?

What theories do you have about why this might be happening?

Let’s take a look at 6th and 10th grade?

### District Math Value-Added Report: The What

<table>
<thead>
<tr>
<th></th>
<th>Estimated Region Mean NCE Gain</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade:</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>Mean NCE Gain over Grades Relative to Growth Standard</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth Standard:</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006 Mean NCE Gain:</td>
<td>1.6 G</td>
<td>2.0 G</td>
<td>3.4 G</td>
<td>-2.5 R*</td>
<td>4.4 G</td>
<td>7.2 G</td>
<td>4.4 G</td>
<td>1.7 G</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std Error:</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.4</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007 Mean NCE Gain:</td>
<td>3.9 G</td>
<td>2.6 G</td>
<td>4.9 G</td>
<td>-0.7 R*</td>
<td>7.7 G</td>
<td>5.8 G</td>
<td>3.0 G</td>
<td>-0.5 R*</td>
<td>3.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std Error:</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.4</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008 Mean NCE Gain:</td>
<td>6.3 G</td>
<td>1.6 G</td>
<td>3.9 G</td>
<td>-2.5 R*</td>
<td>5.1 G</td>
<td>3.0 G</td>
<td>1.6 G</td>
<td>-1.8 R*</td>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std Error:</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.4</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.4</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-Yr-Avg NCE Gain:</td>
<td>3.9 G</td>
<td>2.1 G</td>
<td>4.1 G</td>
<td>-1.9 R*</td>
<td>5.7 G</td>
<td>5.3 G</td>
<td>3.0 G</td>
<td>-1.6 R*</td>
<td>2.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std Error:</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6th Grade Math: Macro to Micro: Why?

- As you continue to drill down, what new information are you finding?
- Are there pockets of success? If so, what might be the reasons for this?
- What other information do you need?
10th Grade Math: Diagnostic: Why?

- Are we producing different results with students of different prior achievement groups?
- Are the curriculum, processes, assessment practices and instruction working for all students?
10th Grade Math: Diagnostic: Subgroups

by African American, Male

Gain

2008 Gain Previous Cohort(s) Reference Line Standard Error

1 (Lowest) 2 3 (Middle) 4 5 (Highest)

Gain
From Micro to Macro

- Do AYP groups have the same patterns?
- Why might this be?
- What do we do when students have already mastered the content?
District, School and Teacher Influence on Student Progress

- 65% To the Teacher quality
- 30% To the School quality
- 5% To the District quality
What do teacher estimates tell us?

Normal Curve

Distribution of Scores

Typical

-2 SE
-1 SE
1 SE
2 SE
Cautions and Conversations

- Value-added is only **one measure** of teacher effectiveness and should be used in combination with other information.
- The analysis becomes more stable and reliable **over time**.
- There is power **diagnostic information** included and given the information teachers can make significant gains in a short period of time.
- Focus the conversation on the things teachers can control and begin with the **teachers strengths**!
# Teacher-Level Value-Added Report

**Subject:** TAKS/Stanford Mathematics, Grade 7

<table>
<thead>
<tr>
<th>Year</th>
<th>Teacher NCE Gain</th>
<th>Tch Std Error</th>
<th>Reference Gain</th>
<th>Teacher Comparison to Ref Gain</th>
<th>Teacher Gain Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>11.6</td>
<td>1.5</td>
<td>5.9</td>
<td>Above</td>
<td>3.8</td>
</tr>
</tbody>
</table>

**Average Simple Gains by Prior-Achievement Subgroup**

<table>
<thead>
<tr>
<th>Year</th>
<th>Ref Gain</th>
<th>1 (Lowest)</th>
<th>2 (Middle)</th>
<th>3 (Highest)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>5.9</td>
<td>5.9</td>
<td>5.9</td>
<td></td>
</tr>
</tbody>
</table>

- **Ref Gain**
- **Avg Gain**
- **Std Error**
- **Nrof Students**
# Teacher Value-Added Summary Report

## Overall Distribution

<table>
<thead>
<tr>
<th>Grade</th>
<th>Teacher</th>
<th>Reading</th>
<th>Math</th>
<th>Science</th>
<th>Social Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>03</td>
<td>Johnson, Michael</td>
<td>NDD</td>
<td>NDD</td>
<td>NDD</td>
<td>NDD</td>
</tr>
<tr>
<td></td>
<td>Veranda, Maria J.</td>
<td>Below</td>
<td>NDD</td>
<td>Below</td>
<td>NDD</td>
</tr>
<tr>
<td></td>
<td>Jonas, Catalina</td>
<td>NDD</td>
<td>NDD</td>
<td>Below</td>
<td>NDD</td>
</tr>
<tr>
<td></td>
<td>Ortega, Mary Ann K.</td>
<td>NDD</td>
<td>NDD</td>
<td>NDD</td>
<td>NDD</td>
</tr>
<tr>
<td>04</td>
<td>Schroff, Abigail</td>
<td>Below</td>
<td>NDD</td>
<td>NDD</td>
<td>Below</td>
</tr>
<tr>
<td></td>
<td>Molina, Heather L.</td>
<td>Below</td>
<td>NDD</td>
<td>NDD</td>
<td>NDD</td>
</tr>
<tr>
<td></td>
<td>Cheney, Donald J.</td>
<td>Below</td>
<td>NDD</td>
<td>NDD</td>
<td>NDD</td>
</tr>
<tr>
<td></td>
<td>Flanagan, Tennessee</td>
<td>NDD</td>
<td>NDD</td>
<td>NDD</td>
<td>NDD</td>
</tr>
<tr>
<td></td>
<td>Weil, Julie</td>
<td>Above</td>
<td>NDD</td>
<td>Above</td>
<td>NDD</td>
</tr>
<tr>
<td>05</td>
<td>Kerry, Joanne</td>
<td>NDD</td>
<td>NDD</td>
<td>Above</td>
<td>Above</td>
</tr>
<tr>
<td></td>
<td>Ingalls, Catherine</td>
<td>Above</td>
<td>NDD</td>
<td>Above</td>
<td>Above</td>
</tr>
<tr>
<td></td>
<td>Radcliff, Jose A.</td>
<td>Above</td>
<td>NDD</td>
<td>Above</td>
<td>Above</td>
</tr>
</tbody>
</table>
Student Projections — What can they tell us?
Using Value-Added - What have we learned?

- **Alignment** – value-added needs to align with all the district’s school improvement efforts.

- **Data Quality** - Get the data right including teacher linkages.

- **Communication** – Develop a communications strategy and do not assume understanding by all stakeholders.

- **Response** - Have a planned response to the data. How will you use the measures and information to improve your system?

- **Capacity** – Provide multiple opportunities and incentives to build educator capacity to use the information.
Table Group Questions

The Promise - What are the potential uses of or benefits of having this kind of information?

The Risk - What are cautions associated with having this kind of information?
The Promise of Value-Added

- Better identify effective teachers, leaders, and programs
- Make more informed decisions and create a lever for real change in the system
- Rewarding excellence and create more opportunities for teachers
- Target professional development and support where it is needed
- Better placement of students and teachers
The Risk of Value-Added

- The data going into the models are not reliable.
- That educators will not use the information.
- The information will be used improperly.
- The measures are unstable or unreliable from year to year.
- The measure won’t align with your understanding of staff through observation and evaluation.
- Unwanted public release of the information.
Questions?
Contact Information

John Hussey
jhussey@battelleforkids.org
(614) 668-9488

Todd Hellman
thellman@battelleforkids.org
(614) 596-4510