An overview of Value-Added Assessment

Ted Hershberg
Director, Operation Public Education
Professor, Public Policy and History
University of Pennsylvania
Without much thought, we answer:

“Good schools are those with high test scores.”
Where do we find these schools?

*In rich communities.*
Most people would reject the definition of a successful school as one with wealthy students.

They are correct.
**Family Income and SAT Scores**

- Under $10,000: 884
- $10,000 - $20,000: 906
- $20,000 - $30,000: 937
- $30,000 - $40,000: 967
- $40,000 - $50,000: 996
- $50,000 - $60,000: 1014
- $60,000 - $70,000: 1026
- $70,000 - $80,000: 1039
- $80,000 - $100,000: 1058
- Over $100,000: 1119

"2005 College-Bound Seniors: Total Group Profile Report" published by CollegeBoard SAT
Do good schools make good students or do good students make good schools?
Field Experiment
What would happen over time if we took ...

- Kids from the inner-city and educated them in affluent suburban schools
- Kids from the affluent suburbs and educated them in the inner-city schools

Since we can’t undertake this experiment, we need a statistical method that can do it for us.
The difference between *achievement* and *growth*
Achievement -- a score on a vertical scale at a single moment in time (absolute or raw score, status, proficiency) -- is best predicted by family income.
New technologies and data sets make new research possible

- James Coleman (1966) and Christopher Jencks (1972) concluded that family background was more important than schooling in explaining achievement.
- They did not have the technology to trace individual students over time (not cohorts).
- Nor did they have data sets that link the teacher for every subject and grade to an individual student’s record.
Growth (the progress students make over the course of the school year) is best predicted by the quality of instruction.
No Child Left Behind

● All students must reach proficiency in reading and math within 12 years

● Adequate yearly progress (AYP) measured for
  – All students
  – All major racial/ethnic subgroups
  – Low-income students
  – Limited English proficiency students
  – Students with disabilities
No Child Left Behind

- A powerful catalyst for change
- Admirable goal
- Serious design flaws
- Value-added assessment can improve AYP, strengthen instruction and increase student achievement
- Proposals from states to include growth in AYP will now be accepted by the Fed DOE
Identifying AYP’s shortcomings

- High Achievement
  - Low Growth
- Low Achievement
  - Low Growth
- High Achievement
  - High Growth
- Low Achievement
  - High Growth

Achievement

Growth
Defining a successful school

- Each year the performance of the students exceeds what is expected of them, given their academic background.
- Over time all students are able to achieve high standards (NCLB).
Value-Added: A New Lens
Value-Added Assessment

- First developed for Tennessee by William Sanders.
- Since 1992, tracks each of the state’s 885,000 students.
- 10 million records, grades 2-12 with test scores in every subject, every grade, every teacher.
- Largest data base ever assembled.
- Mandatory in Pennsylvania and Ohio as well as in over 300 districts and consortia across the U.S.
Philosophy Behind Value-Added Assessment

- Schools **can** and **should** add value for each student from September to June.
- This is true whether the student comes in above grade, at grade or below grade.
- Students are **entitled** to grow at least at a rate they have demonstrated in the past.
Value-Added: The Basics

- Value-added is not a test.
- It is a way of looking at the results that come from tests.
- Value-added lets us determine whether the students in a class, school or district are making enough academic growth each year.
Example - Student A

![Graph showing test scores vs. grade for Student A.](image-url)
Student A vs. District Average

Test Score vs. Grade

- **Student A**
- **District Average**
What can we conclude?

Not much. The student appears to be consistently near average, if a little below.
But we see that in fourth grade, the student advanced at a rate far above the district average...
…and in fifth grade, the student advanced at a rate below the district average.
We can draw no conclusions about the quality of classroom activity from these facts alone...
Now, if we instead think of the line as representing:
- the average of all students in the teacher’s classroom, and
- the average of three years of the teacher’s classrooms,
then we can draw meaningful conclusions about what’s happening in that class.
The Concept Behind Value-Added

- Value-added is statistically and computationally complex
- But the idea behind it is straightforward...
Projected and Actual Scores

- Value-added calculates a projected test score for a student in a given grade and subject.
- The projected score is based entirely on the student’s prior academic achievement.
- It is then compared to the actual score at the end of the year.
Value-added measures the difference between actual and projected.

- Actual exceeds projected
- Actual equals projected
- Actual falls below projected

Value-Added
Levels the Playing Field
Value-added measures the difference between actual and projected.
Don’t confuse value-added assessment with mere growth or gain

• Many people make this mistake.
• It sounds reasonable to think of the “growth” or “gain” a student makes from one year to the next as the “value” that’s been “added.”
• Value-added assessment is much more powerful than a simple growth or gain score.
Value-Added

Divides difference between projected and actual scores into two parts

● That which is contributed by the student

● That which is contributed by the teacher
Each child serves as his own statistical control

The Environmental Variables Remain the Same
- Family Income
- Ethnicity
- Gender
- Neighborhood

Grade 3

Grade 4
These records are then averaged for all students in the teacher’s classroom over a 3-year period.
What makes value-added fair?

- For children

  Value-added is fair to students because it bases their projected score only on their prior academic record. That ensures that all children are expected to make progress each year from wherever they start.
What makes value-added fair?

- For educators
  It is fair to administrators and teachers because prior academic achievement data already incorporate the student background characteristics that bias absolute test scores.
Value-Added in a Standards World

(Hypothetical Progression)

The Standard

Closing the Gap through Systemic Change

Student Achievement

Grade

Right Now

Not “losing ground”
Value-Added Findings

Patterns from the Data
Tennessee Schools and their Value-Added Scores

It is impossible to determine where a school falls just by knowing its location or the make-up of its student body.

Math: 1996-97

100 (on the horizontal axis) means a year’s worth of growth in a year.
Income has no effect on value-added

• Each dot represents one school.

• 100 (on the vertical axis) means a year's worth of growth in a year.

Cumulative Gain of a Large East Coast County’s Schools Compared with the Percentage of Students Receiving Free and Reduced-Priced Lunches

Each dot represents 1 school system
Horizontal line at 100% represents gain equal to national norm gain
3 Year Average Gain
July, 1997
Cumulative Gain of Tennessee Schools Compared with the Percent of Minority Students in the School

Minority status has no effect on value-added

- Each dot represents one school.
- 100 (on the vertical axis) means a year’s worth of growth in a year.
Teacher Effectiveness

First 10-12 years

Second 10-12 years

Typical Salary Schedule

After 20-24 years

Teacher Experience
Value-Added
Findings From Tennessee

The Teacher Effect
Importance of Teacher Sequence

Cumulative Effects of Teacher Sequence on Fifth Grade Math Scores for Two Metropolitan Systems
Probability that a bottom-quartile 4th grade student will pass the high-stakes graduation exam in 9th grade

- Poor teacher sequence: <15%
- Average teacher sequence: 38%
- Good teacher sequence: 60%
## Cumulative Effects of Value-Added

<table>
<thead>
<tr>
<th>Grade</th>
<th>TVAAS Scores</th>
<th>75%</th>
<th>100%</th>
<th>140%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>2.75</td>
<td>3</td>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3.5</td>
<td>4</td>
<td>4.8</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4.25</td>
<td>5</td>
<td>6.2</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>6</td>
<td>7.6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>5.75</td>
<td>7</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>6.5</td>
<td>8</td>
<td>10.4</td>
<td></td>
</tr>
</tbody>
</table>

**GRADE LEVEL IMPACT**

-1.5  0  +2.4

This means a difference of almost 4 grade levels by the end of middle school.
Predicting Student Learning Results

Achievement
Best predicted by family income

Growth
Best predicted by the quality of instruction

Tennessee research shows that teacher effectiveness is the single most powerful predictor of student progress – stronger than income, class size, race or family educational background.
Reducing class size is a poor policy choice to increase student learning

- It ranks 40\textsuperscript{th} among 46 options
- Feedback and direct intervention are the highest (effect sizes of 0.81)
- Where the average is 0.40, the effect size of reducing class size is 0.12

Source: John Hattie, Keynote, International Conference on Class Size, University of Hong Kong, May, 2005
Using Value-Added to Inform Instruction
Diagnostics 1
The Focus of Instruction
This pattern – of high early achievement followed by extremely low value-added – is quite common, especially in urban districts.
In this pattern – frequently found in suburban districts – the teacher is teaching to the high achievers at the expense of other students.
In this pattern, the teacher is teaching right down the middle.
Sustained Growth

Closing the gap

Still making gains

115-120%

105%

100%

Gain

Low Average High

Previous Achievement

Low Average High

Previous Achievement
Diagnostics 2

The Impact of Instruction
Value-Added: Three Results

- Three Results

Above

No Detectable Difference (NDD) – One year’s worth of growth

Below

(using 3-year running averages)
Diagnostics 3

Combining the Focus and Impact of Instruction
Example:
Four 5th Grade Classrooms

100%

No Detectable Difference

Reading  Language Arts  Math  Social Studies
Example:
High School English Dept.

No Detectable Difference
Tepee Pattern
Using Previous Academic Achievement Levels
Example 1

- No Detectable Difference

100%

Low Average High

{ No Detectable Difference }
Tepee Pattern
Using Previous Academic Achievement Levels
Example 2

100%

Low Average High

No Detectable Difference
Shed Pattern

Using Previous Academic Achievement Levels

Example

No Detectable Difference

100%
Reverse Shed Pattern
Using Previous Academic Achievement Levels
Example

No Detectable Difference
Value-added provides powerful diagnostic data

- Identify and improve the focus and impact of instruction
- End the isolation of teachers
- Build learning communities
- Improve data-driven decision making
- Differentiate instruction
- Create student growth trajectories to targets and develop intervention strategies
- Measure the success of schools through growth, not simply achievement
Tennessee NAEP Scores

Mathematics - Grade 4

- NAEP Scale Score
- Tennessee
- National Average
The limits of value-added in Tennessee

- May be used in individual teacher evaluations, but may not exceed 8%
- Lack of professional development to accompany statewide rollout
Value-added assessment is only a thermometer; if we don’t analyze the information and use it, nothing happens.
Why collect classroom level data?

- The variation in the quality of instruction is much greater within schools than between schools.
- Struggling students are not randomly distributed in classrooms – they are found disproportionately in classrooms where they receive poor instruction.
- It allows you to deal with underlying causes not symptoms.
For additional information on our package of reforms, please contact:

cgpinfo@pobox.upenn.edu

or (215) 746-6478