Fostering Productive Conversations with Growth Data to Improve Educator Effectiveness

Damian W. Betebenner

National Center for the Improvement of Educational Assessment
Dover, NH

The Role of Short Cycle Assessment in Educator Evaluation,
Chicago, IL May 5th & 6th, 2011
Data Data Everywhere (The Economist, 2010)

- More and better organized data: *mega, giga, tera, peta, exa, zetta, yotta*!
- Historical records of student achievement.
- Linkage of student achievement to student demographics, teachers, schools, educational programs, . . .
- Stakeholder interest in examining student achievement over time (student growth) derives from data availability..
## Lessons Learned

<table>
<thead>
<tr>
<th>On Questions</th>
<th>John Tukey</th>
</tr>
</thead>
<tbody>
<tr>
<td>It is better to have an approximate answer to the right question than a precise answer to the wrong question.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>On Statistical Models</th>
<th>George E. P. Box</th>
</tr>
</thead>
<tbody>
<tr>
<td>All models are wrong but some are useful.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>On Understanding</th>
<th>Aristotle</th>
</tr>
</thead>
<tbody>
<tr>
<td>We understand best those things we see emerge from their very beginnings.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>On Accountability</th>
<th>Sherlock Holmes</th>
</tr>
</thead>
<tbody>
<tr>
<td>It is a capital mistake to theorize before you have all the evidence. It biases the judgment.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>On Theories of Action</th>
<th>Confucius</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tell me and I will forget, show me and I will remember, involve me and I will understand.</td>
<td></td>
</tr>
</tbody>
</table>
On Questions & Models

**Growth models address specific questions**

- Different growth analysis techniques often answer different, but related, questions.
- It is critical to understand these different questions.
- Different questions lead to different conversations which lead to different uses and outcomes.

**Changing conversations about education**

- Starting with the right questions simplifies development – motivating the proper use of the growth model results.
- The questions set the table for those conversations.
- Yen (2007) provides an excellent list of questions derived from a survey of parent, teachers and administrators
- Different stakeholders have different questions.
What are the relevant questions for parents?

Yen (2007), from a state survey of parents, teachers and administrators, compiled a list of frequently voiced questions/concerns by stakeholder group.

### 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?
What are the relevant questions for teachers?

Yen (2007), from a state survey of parents, teachers and administrators, compiled a list of frequently voiced questions/concerns by stakeholder group.

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?
What are the relevant questions for administrators?

Yen (2007), from a state survey of parents, teachers and administrators, compiled a list of frequently voiced questions/concerns by stakeholder group.

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?
What are the relevant questions for policy makers?

**Administrator Questions**

- Can/how should growth be used in judgments of educator quality?
- Can/how should growth be used in judgments of education (e.g., school) quality?
- Are students on track for career and college readiness?
- What adjustments for student demographics need to be considered?
Growth & Accountability

Growth versus Status

- Enthusiasm for growth in accountability stems from the belief that growth and teacher/school quality are more closely related than status.

- Enthusiasm for growth also stems from its potential diagnostic uses.
Growth & Accountability

**Fundamental Premise**

“Good” schools/teachers bring about growth in student achievement in excess of that found at “bad” schools/teachers.

- The inherent believability of this premise is at the heart of current enthusiasm to incorporate growth models into accountability systems.
- Note that current NCLB accountability invokes the same premise (with achievement substituted for growth).
- “Good schools/teachers” are commonly referred to as highly effective schools/teachers.
- Race to the Top has pushed teacher effectiveness.
- It is critical to understand the distinction between growth and effectiveness.
Who/What is responsible for student growth?

- Often, analyses of student growth attempt to determine the amount of student progress/growth that can be attributed to the school or teacher.
- Called *Value-added* analyses, these techniques attempt to estimate the teacher/school contribution to student achievement.
- Value-added models quantify *effectiveness*—purporting to measure the impact on achievement that this school or teacher would have, on average, upon similar students assigned to them.
- These models and their effectiveness claims are controversial for a number of reasons (EPI (2010), NRC/NAE (2010), Braun (2005), Rubin and Stuart (2004)).
Value-added Analyses/Inferences

Who/What is responsible for student growth?

- Value-added models are better described as **effectiveness** models than as **growth** models.
- Value-added is an **inference** – a causal conclusion drawn from the data.
- Value-added isn’t a type of statistical model or type of assessment.
- All growth models can be used for value-added purposes but not vice-versa.
Descriptive Questions

- Note that the questions from Yen’s (2007) survey are primarily descriptive.
- The questions are only peripherally associated with causality.
- High stakes accountability has transformed questions about student growth into questions about responsibility/cause: Teacher and School Effectiveness.
- Beginning with description of growth doesn’t prevent subsequent discussions about responsibility/cause.
- Beginning with responsibility/cause (i.e., rendering a verdict) often prevents discussions about the data.
Accountability system results can have value without making causal inferences about school quality, solely from the results of student achievement measures and demographic characteristics. Treating the results as descriptive information and for identification of schools that require more intensive investigation of organizational and instructional process characteristics are potentially of considerable value. Rather than using the results of the accountability system as the sole determiner of sanctions for schools, they could be used to flag schools that need more intensive investigation to reach sound conclusions about needed improvements or judgments about quality [Linn, 2008, p. 21].
Describing Student Growth

- Measuring student growth, even with a vertical scale, is not a simple task.
- Some believe a vertical scale simplifies the task of measuring student growth.
- Even with an interval (or ratio) scale, growth is not easy to interpret. Consider, for example, height.
  - A child might grow 4 inches between ages 3 and 4.
  - 4 inches is a well understood quantity.
  - The 4 inch increase becomes really meaningful only when understood alongside the growth of other 3 to 4 year olds.
- **Student growth percentiles** were developed to provide a normative context for describing student growth.
Student Growth Percentiles

Should we be surprised with a child’s current achievement given their prior achievement?

- Given a student’s prior scale scores and the associated conditional density, their current scale score corresponds to a percentile of that conditional distribution.
- This percentile is the student’s growth percentile.
- Growth percentiles are closely related to estimating the probability of observing a student’s current achievement taking account of their past achievement:

\[ \Pr(\text{Current Achievement} | \text{Past Achievement}) \]

- As such, growth percentiles describe the rarity of a student’s current achievement conditional upon their prior achievement.
Student Growth Percentiles

Should we be surprised with a child’s current achievement given their prior achievement?

- Student growth percentiles answer this question.
- Consider a low achieving student with 90th percentile growth and a high achieving student with 10th percentile growth.
  - The low achieving student grew at a rate exceeding 90 percent of similar students.
  - The high achieving student grew at a rate exceeding just 10 percent of similar students.
  - The low achiever’s growth is more *exemplary* (probabilistically) than the high achiever’s.

- Judgments about the *adequacy* of student growth require external criteria.
Establishing growth-standards based upon growth-norms

- The most common adequacy criterion are judging growth toward an achievement goal (i.e., growth-to-standard).
- Results from student growth percentile analyses can be used to calculate percentile growth trajectories for each student.
- These trajectories indicate what future rates of growth will lead to and are used to make adequacy judgments.
- This growth-to-standard approach was approved as part of Colorado’s successful application to the Growth Model Pilot Program.
It’s of interest to examine schools where students demonstrate, on average, extraordinarily high and low student growth.

To summarize the student growth percentiles associated with a school (or other grouping) calculate the median of the student growth percentiles.

If students were randomly assigned to schools, expect to see a median of 50.

Values greatly above or below 50 are of interest in identifying best practices or providing extra support.

Examining growth with achievement sheds new light on school performance.
District C: 2008 CSAP Math School Results

Student Growth versus Student Achievement by Percent Free/Reduced Lunch

- Median of Student Growth Percentiles in School
- Percent at/above Proficient in School

School Percent Free/Reduced Lunch:
- Less than 20 percent
- 20 to 40 percent
- 40 to 60 percent
- 60 to 80 percent
- More than 80 percent

School Size:
- 50 Students
- 100 Students
- 200 Students
- 500 Students
- 1,000 Students
“Good teachers/schools” are often called highly effective teachers/schools.

What’s the relationship between growth and effectiveness?

Effectiveness indicates who/what is responsible for the growth (value-added models).

RttT has placed teacher responsibility for student progress as a primary issue.

In teacher performance evaluations, even when large scale assessment results are available, depending upon the state they constitute, on average, about a third of the overall teacher rating.
“This is the difference between a retrospective question of identifying fault as opposed to a prospective strategy to engineer some corrective measure, almost independent of considering whether there was blame-worthiness. And to move away from the blame-worthiness paradigm toward something that is more regulatory in nature where one might seize upon disparities or circumstances that are for some reason deemed unacceptable and engineer the interventions needed to bring about the necessary change. . . . It’s the no-fault gap closing strategy in which the effort is to build a consensus about a vision of an improved society rather than figure out where’s the person we want to pillory.”

C. Edley (2006)
The Colorado Department of Education and the Center for Assessment have been working for the last two years on developing next generation data visualization to accompany growth model data.

The goal: Transform conversations about education through active engagement with data (i.e., evidence).

Our efforts have received tremendous interest and recognition:

- Recognized by Adobe for innovative uses of their technology as an Adobe Max Award finalist in October, 2009.
- 2010 NCME Award for Outstanding Dissemination of Educational Measurement Concepts to the Public.
- 14 states signing MOUs to co-develop a cloud-based reporting platform in a non-proprietary fashion.
- Colorado recently devoted $2.5 million of stimulus funds to the development efforts
The Decade of Data Visualization


Hans Rosling/Gapminder
http://www.youtube.com/watch?v=jbkSRLYSoho

Journalism in the Age of Data
http://datajournalism.stanford.edu/
Web 2.0: Data Visualization and Social Networking

With a collaborative spirit, with a collaborative platform where people can upload data, explore data, compare solutions, discuss the results, build consensus, we can engage passionate people, local communities, media and this will raise—incredibly—the amount of people who can understand what is going on.

And this would have fantastic outcomes: the engagement of people, especially new generations; it would increase knowledge, unlock statistics, improve transparency and accountability of public policies, change culture, increase numeracy, and in the end, improve democracy and welfare.

E. Giovannini, Chief Statistician, OECD. June 2007
The Colorado Growth Model

- Admitted Model
- Near Model Adoption
- Interested in Model
References


