Balasubramanian, N., & Muth, R. (2010, April). Instructional & program effectiveness: From data to action. Paper to be presented at the AERA Annual Meeting, Denver, CO.

Instructional and Program Effectiveness: From Data to Action

ABSTRACT

This paper introduces a new method of utilizing assessment data to identify instructional and program effectiveness, starting with a four-quadrant approach. Participants will learn ways to make maximal use of data to help teachers improve their craft by applying a graphic, growth-by-achievement-performance summaries. The feedback tool discussed includes an innovative and transparent growth model that employs a test-independent, normalized gain and performance index. Use of the tool provides a practical way to help educators get a better and rapid understanding of what students know and must be able to do from year-to-year or pretest-to-post-test in school, district, and state assessments. With constructive and timely feedback, teachers can be empowered to frame issues around 4+2 essential questions in professional learning communities.

OBJECTIVES

Are we teaching as many students as we can? Are students learning what we intended to teach them? Do we have ways of using and monitoring ongoing formative assessments? These questions, among several others (Balasubramanian, 2006), continue to guide teachers in the trenches as they strive to improve the performance of students with diverse needs and abilities. Providing timely feedback to all students to maximize learning, while simultaneously informing and adjusting one's classroom instruction based on school and district benchmark assessments, is a challenge.

This paper elaborates the use of longitudinal student-assessment data to improve the quality of teaching and learning in schools by using a simple, graphic growth-by-achievement performance summary. This feedback tool builds on the work of the physics-education-research (PER) community that uses normalized pretest-to-post-test gains "g" (Hake, 2007) on the force concept inventory (Hestenes, Wells, & Swackhamer, 1992) to quantify student improvement (Paden & Moyer, 1969).

THEORETICAL FRAMEWORK

Normalized gains can be used to quantify student growth to monitor their progress over time (Balasubramanian, Frieler, & Asp, 2008). The normalized gain "g" of a student is the actual gain (post-test %—pretest %) divided by the maximum possible gain (100%—pretest %). The gain can vary from 0 to 1 (or 0% to 100%). It is an estimate of how much students have learned regardless of their initial level of competence. The post-test % is a measure of student achievement. A scatterplot with these two dimensions——growth on the x-axis and achievement along the y-axis——produces four quadrants:

- High-achieving students doing better in the post-test than the pretest [North-East (NE) Quadrant];
- High-achieving students doing poorly in the post-test than the pretest [North-West (NW) Quadrant];

Submitted 7/8/2009
Accepted 11/6/2009
Correspondence to: Nathan Balasubramanian, Ph.D.; E-mail: nbala@cboces.org
Centennial BOCES, 830 South Lincoln Street, Longmont, CO 80501

- Low-achieving students doing poorly in the post-test than the pretest [South-West (SW) Quadrant];
- Low-achieving students doing better in the post-test than the pretest [South-East (SE) Quadrant].

The averages on growth and achievement across the school/district represent the two axes. The NE quadrant is clearly the most desirable—but the SE quadrant also shows student improvement, suggesting a positive growth trajectory over time. Teachers and administrators can view scatterplots like this and see a broad profile of how different kinds of students perform over time.

Like the normalized gain "g," the performance index (PI) estimates the extent of proficiency within a performance level. It is a measure of student achievement in a standardized test (Perie, Weiss, Kurtz, & Dunn, 2009). The PI is calculated by performing linear transformations of students' scale scores on large-scale assessments.

Knowing students' PI is helpful because, if we know their PI in two consecutive years, their growth from one year to the next (or change-in-PI) will be their PI in the current year minus their PI in the prior year. As before, a scatterplot with these two dimensions—growth on the x-axis and achievement along the y-axis—produces four quadrants:

- High-achieving students doing better this year than last year [NE Quadrant];
- High-achieving students doing more poorly this year than last year [NW Quadrant];
- Low-achieving students performing more poorly this year than last year [SW Quadrant];
- Low-achieving students doing better this year than last year [SE Quadrant].

The horizontal axis is set at 3.00 (proficient) and the vertical axis is set at 0.00. A student with a PI ≥ 3.01 is considered high achieving. A student with a PI ≤ 2.99 is considered low achieving. The standard error of measurement on growth with change-in-PI averages to zero with large sample sizes. Therefore, the criterion for growth is zero. A student with change-in-PI ≥ 0.01 is doing better this year. A student with change-in-PI ≤ minus 0.01 is doing poorly this year.

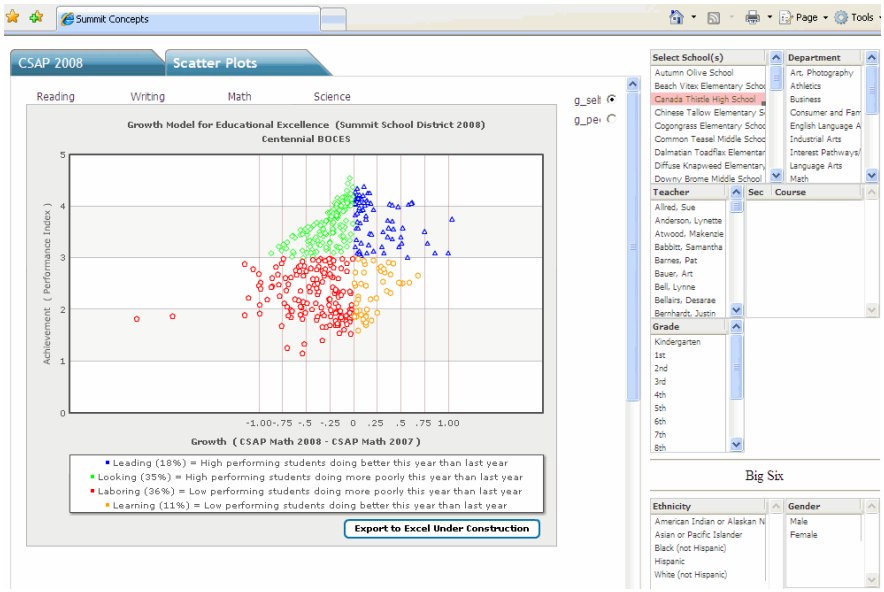

Fig. 1. Scatterplots to observe high level patterns across each performance level (Advanced 4.00-4.99; Proficient 3.00-3.99; Partially Proficient 2.00-2.99; Unsatisfactory 1.00-1.99)

Submitted 7/8/2009
Accepted 11/6/2009
Correspondence to: Nathan Balasubramanian, Ph.D.; E-mail: nbala@cboces.org
Centennial BOCES, 830 South Lincoln Street, Longmont, CO 80501

In the tool that we have developed (all the school, teacher, and student names are fictitious in Figs. 1 and 2), building and district administrators can also see these scatterplots further disaggregated by the six subgroups—gender, ethnicity, free-and-reduced lunch, special education, talented-and-gifted, and English language learners—to observe growth and achievement gaps and/or excellence across teachers, departments, schools and programs.

POSSIBLE MODES OF INQUIRY

Educators across the country meet periodically in Professional Learning Communities (PLCs) at their respective institutions to purposefully discuss how students learn (DuFour & Eaker, 1998). The data inquiry usually focuses on four questions—Q1: What do we want our students to know and be able to do? Q2: What evidence do we have that they have learned it? Q3: What do we do when students have not learned it (SW quadrant)? Q4: What do we do with students who have already learned it (NE quadrant)?

The growth-by-achievement summary (Fig. 1) contributes two necessary and essential questions that were absent until now in PLC conversations. These are—Q5: How do we engage and promote learning among high-achieving students who are performing more poorly this year than last year (to move them from the NW quadrant to the NE quadrant), and Q6: How do we develop proficiency among low achieving students who are doing better this year than last year (to move them from the SE quadrant to the NE quadrant). With these two questions, teachers can identify the "lowest hanging fruit" for effective data inquiry and achieve improved results.

Further, using this growth-by-achievement performance summary, teachers (as researchers) might investigate the following questions: What instructional strategies worked and did not work for me? Which groups of students (in the six subgroups) improved (or declined) in each performance level? What do we notice about the growth of these students (individually and collectively) in each performance level? Which groups of students pose additional questions? How does my performance relate to my colleagues who teach the same course within the school and across the district? What does student work look like? Do we have exemplars to share with our colleagues? Informed by specific strategies that worked, what is my plan of action for the next/current school year in terms of instructional interventions for students (individually and collectively) in the four quadrants?

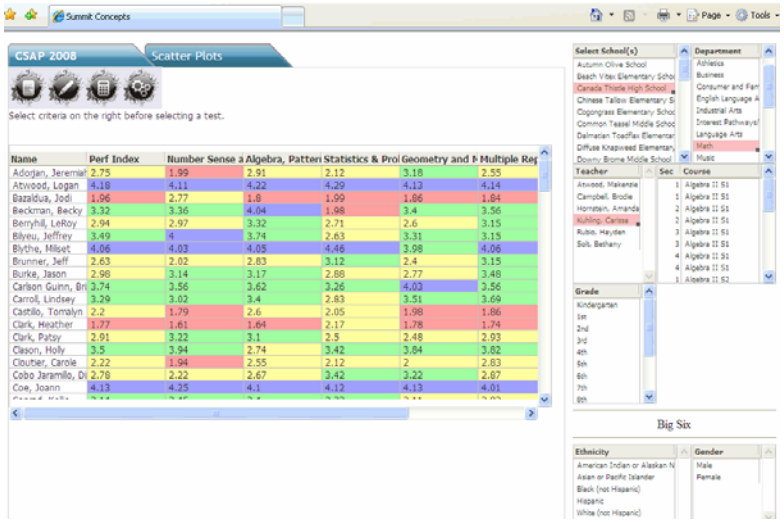

Fig. 2. Identifying students' strengths and needs

Submitted 7/8/2009
Accepted 11/6/2009
Correspondence to: Nathan Balasubramanian, Ph.D.; E-mail: nbala@cboces.org
Centennial BOCES, 830 South Lincoln Street, Longmont, CO 80501

After seeing the high-level patterns, teachers might examine in depth the specific content and sub-content standards shown in Figure 2. They can identify the strengths and needs of individual students and craft relevant instructional interventions.

This feedback tool, with its visually compelling data representation, can be used productively by schools, departments, and individual teachers to evaluate programs, instructional strategies, and classroom interventions using classroom-based and formative assessments year-round. With the ability to drill-down to teacher and department level, good/best practices can be identified and shared to improve instruction school-wide. Teachers can design interventions as a department; and school administrators likewise can craft evidence-centered decisions. They can set (what educators refer to as) SMART (Specific, Measurable, Attainable, Research-based, and Time-phased) goals to determine what might be reasonable year's growth across the different performance levels.

Administrators (as evaluators) might ask: What evidence do we have to determine which programs and interventions are working, which ones are not, which ones need adjusting, and who needs support? How can we leverage resources to maximize the learning opportunities for all students? What (intrinsically meaningful) performance indicators could we use to evaluate the effectiveness of processes and implementation across the district? Looking back at last year's district improvement plan and visually accessible growth-by-achievement data (on formative and summative school, district, and state assessments), can we identify the root cause(s) and reflect on them—for celebrations, roadblocks, opportunities—and work ahead to address problem areas? What resources, including professional development and training, might be necessary to bring about sustainable learning district-wide?

DATA SOURCES

Based on the real-life models in Figures 1 and 2, the approach encapsulated here offers numerous benefits to teachers and administrators using student assessment data. The assessment data in a typical school or district could come from year-to-year large scale state assessments. It could be pretest-to-post-test measures on school-wide/district-wide formative, benchmark, or summative assessments. The feedback tool (Figures 1 & 2) presented results from archival student assessment data on state assessments.

While promising, stakeholders should interpret the results of all assessment data with some care. Still more work needs to be done as we co-develop training materials with educators and integrate this approach into school-improvement efforts during our implementation studies. The interpretations might be skewed by factors such as regression toward the mean associated with testing errors; test-ceiling effects for some high achieving students; and improvements in test-taking skills associated with repeated, within-year assessments. Popham (2008), for example, makes a compelling case about the inherent limitations of large-scale assessments. These include the grain size of curricular aims, which make all the difference to the effectiveness of an assessment.

> There's not time enough to teach everything educators would like their students to learn. Similarly, there's not time enough to *test* everything educators would like to know if their students have learned (p. 128).

Osborne (2008) makes the same point more broadly: "Evidence-based decisions are only as good as the evidence they are based on" (p. x).

Submitted 7/8/2009
Accepted 11/6/2009
Correspondence to: Nathan Balasubramanian, Ph.D.; E-mail: nbala@cboces.org
Centennial BOCES, 830 South Lincoln Street, Longmont, CO 80501

CONCLUSIONS

Improving the performance of students with diverse needs and abilities amidst the data deluge in K-12 education can be a challenge. This paper offers a graphic tool to represent test performance of groups of students based on two factors: achievement *and* growth. The tool can be used to provide constructive and timely feedback to students and teachers. It can be used to identify gaps in learning and inform classroom instruction. In Figures 1 and 2 we showed how it can be used to track progress toward learning goals and targets on achievement and growth.

Additionally, the growth-by-achievement performance summary contributes two necessary and essential questions for effective data inquiry using PLCs: How do we engage and promote learning among high-achieving students who are performing more poorly this year than last year? And, how do we develop proficiency among low achieving students who are doing better this year than last year? The graphic tool in the figures presents a pragmatic approach to access a wealth of useful information within the full spectrum of student assessment data. Further, the possibilities and pitfalls have been identified. Based on experience, it is probable that the feedback tool discussed here will lead to better discussions, better strategies, better interventions, better instruction, and consequently better learning that can sustain continuous improvement practices.

SIGNIFICANCE OF THE STUDY

This proposal points to an innovative approach to promote meaningful conversations among teachers (in PLCs) and administrators (in professional leadership and learning teams) using growth-by-achievement performance summaries. Teachers and administrators, with these performance summaries, can obtain “at-a-glance” snapshots of what students know and are able to do (KAAATD) over time with effective instruction. For instance, by examining the scatterplots disaggregated by the six subgroups, growth and achievement excellence and/or gaps in excellence across teachers, departments, schools and programs can be observed. Using the two essential questions, teachers and administrator can identify the “lowest hanging fruit” for effective data inquiry and achieve improved results quickly, while engaging in the entire process over time can lead to significant learning gains for all students. With such results, schools and districts can build and then sustain a culture of continuous improvement systemwide.

WORD COUNT: 1870

Submitted 7/8/2009
Accepted 11/6/2009
Correspondence to: Nathan Balasubramanian, Ph.D.; E-mail: nbala@cboces.org
Centennial BOCES, 830 South Lincoln Street, Longmont, CO 80501

REFERENCES

Balasubramanian, N. (2006). Increasing student achievement through meaningful, authentic assessment. In C. Crawford et al. (Eds.), *Proceedings of Society for Information Technology and Teacher Education International Conference 2006* (pp. 3-8). Chesapeake, VA: Association for the Advancement of Computing in Education. Retrieved from www.editlib.org/d/21994/proceeding_21994.pdf

Balasubramanian, N, Frieler, J., & Asp. E. (2008) Designing learning. *Principal Leadership*, *9*(2), Reston, VA: National Association of Secondary School Principals. 34-39.

Colorado Department of Education (2009). *Performance Level Scale Ranges for CSAP Assessments*. Retrieved from http://www.cde.state.co.us/cdeassess/documents/csap/manuals/2008/20080718_scalescoreranges.pdf

DuFour, R., & Eaker, R. (1998). *Professional learning communities at work: Best practices for enhancing student achievement*. Bloomington, IN: National Education Service.

Hake, R. (2007). Six lessons from the physics education reform effort. *American Journal of Physics*, *1*(1), 24-31.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, *30*(3), 141-158.

Osborne, J. W. (Ed.). (2008). *Best practices in quantitative methods*. Thousand Oaks, CA: Sage Publications.

Paden, D. W., & Moyer, M. E. (1969). The relative effectiveness of three methods of teaching principles of economics. *The Journal of Economic Education*, *1*(1), 33-45.

Perie, M., Weiss., J., Kurtz, T., & Dunn, J. (2009, June). *Comparing the use of a growth model and an index in a state accountability system*. Session 041 presented at the meeting of the National Conference of Student Assessment, Los Angeles, CA.

Popham, W. J. (2008). *Transformative assessment*. Alexandria, VA: Association for Supervision & Curriculum Development.

Submitted 7/8/2009
Accepted 11/6/2009
Correspondence to: Nathan Balasubramanian, Ph.D.; E-mail: nbala@cboces.org
Centennial BOCES, 830 South Lincoln Street, Longmont, CO 80501