Do Not Confuse Assessment with Accountability!

By Ronald W. Costello, Ed.D. and Barbara Shuey

As schools become more standards based, we struggle with the vast amount of student performance data that is being gathered. In Indiana, as in other states in the United States, students take a state test to measure whether they have mastered the 2000 Indiana Academic Standards in English/language arts, mathematics and science. The results of this test, the Indiana Statewide Testing of Educational Progress (ISTEP), are made public by the media, state government, local districts, and schools. Rarely do we get beyond simply reporting and rank ordering the data. Schools find it difficult to use the data to set priorities and determine if they are making a difference in student achievement. Furthermore, we want to be accountable for more than test scores or school rankings; we want to be accountable for improved student achievement.

In his keynote address, at the 1999 Fall ASCD Teaching and Learning Conference, Michael Fullan said educators must become "assessment literate." To accomplish this, he stated we must be able to do three things: (1) examine student data and make sense of it; (2) make changes in teaching and schools derived from those data; and, (3) commit to engage in external assessment discussions. Fullan stated that educators really do not have a choice about whether to engage in the national conversation about assessment because it will take place with or without them. In order for schools to be successful in this environment, educators must become "assessment literate."

A major problem with becoming assessment literate is that educators seem to confuse assessment with what they should be accountable for accomplishing. We look for other ways to measure student skills for which we already have assessment data instead of trying to use the

data we already have in ways to improve student achievement. In order to be more accountable, we suggest the following steps: 1) Agree upon how to use the results of assessments already in place; 2) Align programs based upon the results of those assessments; 3) Relate assessment results to interventions to improve student achievement; and, 4) Be accountable for improved student achievement. Each step will be addressed separately.

How to use the results

In order to use the assessment results, we need to create an information system which supports the efficient and effective management of the data. In the Archdiocese of Indianapolis, with 71 schools and approximately 24,000 students, we were able to implement an information system through a school improvement grant from the Lilly Endowment, Inc. One of the goals of our grant was to improve student achievement on the ISTEP. Because most of our schools (40) are accredited by the North Central Association (NCA), we have used their data analysis method, outlined below, as the basis for assessing gains in student performance. This has allowed us to meet the basic assessment requirements needed by our schools to reorganize the individual student data and to be able to generate reports at the student, classroom, school, or district level.

The need to create the different reports became evident when we created a template for a standards based report. The first student we looked at, at the first school for more than 5,000 students tested, had met the state's performance expectation in English/language arts but not in mathematics. Looking further at his mathematics results at the academic standard level, we found he passed all of the academic standards except problem solving. If the teacher had this

information readily available in August when school started, she would have been able to work with the student on problem solving before he took the ISTEP examination in September.

Based upon experiences like this, we decided we need to provide our staff with a method to use the performance data at the classroom level so they had the tools to generate reports on the students they served. We used a basic statistical method of converting the performance data to a standard score then comparing the standard scores to determine an "*effect size*" or difference. This is a critical aspect of deciding how to use the assessment data because it provides both a method to determine areas where interventions are needed and whether those interventions were effective.

Below is an example of a grade level report for a sample school in the archdiocese used in the first step of the process to identify areas where improvement is needed.

ISTEP+ Grade Level Effect Size Report
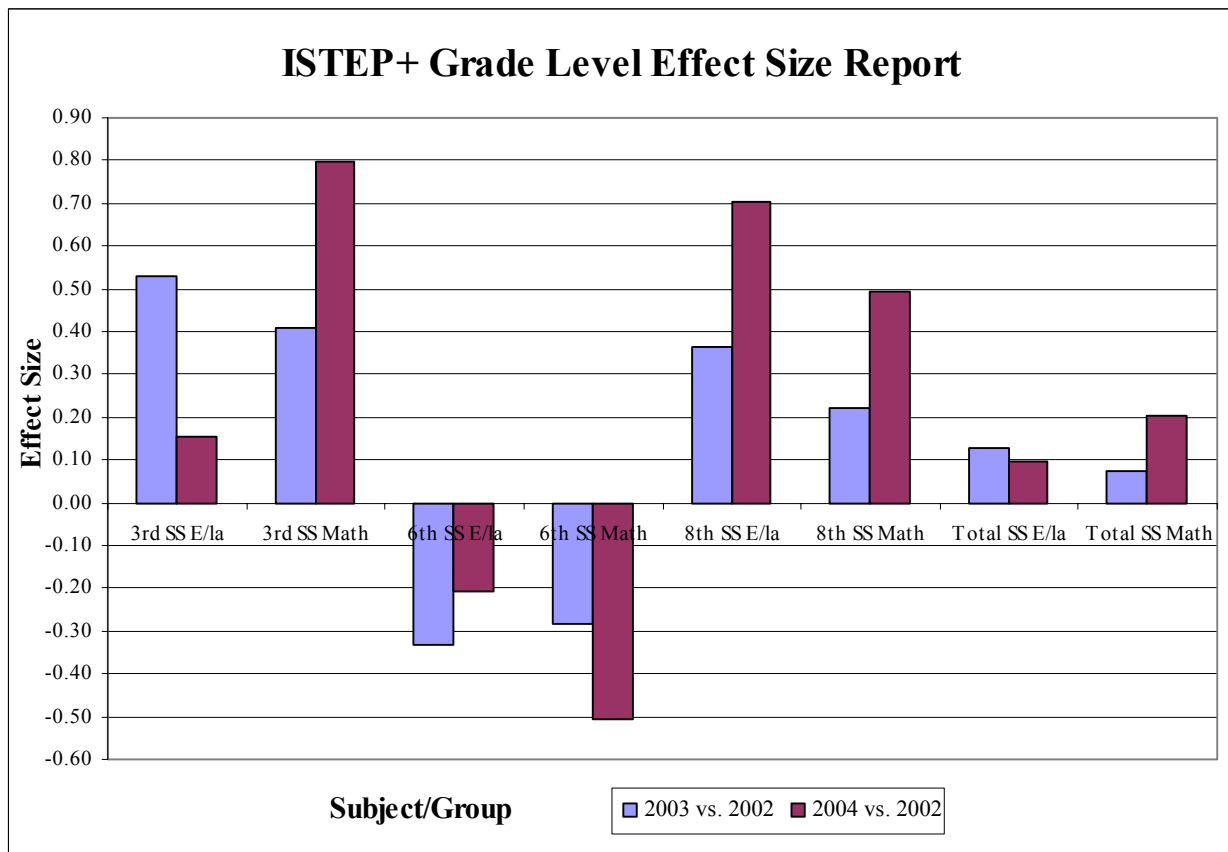
The school *Effect Size* report was created at the subject level using the scale score. According to Yeow Meng Thum (2001) in *Measuring Progress towards a Goal*, the scale score is a good measure to show growth in student achievement because the scale score provides a larger range of variation in scores which is more likely to show growth or lack of it.

In the graph, there is great deal of variation in *effect size* scores at the different grade levels. Based upon their experience evaluating thousands of school improvement plans over many years, NCA COS has made the following suggestion for interpreting *effect size*:

- ✓ An effect size of .3 or greater shows growth that is substantial.
- ✓ An effect size of .2 or greater but less than .3 shows growth that is quite good.
- ✓ An effect size of .1 or greater but less than .2 shows growth that is enough to mention.
- ✓ An effect size less than .1 but greater than -.1 shows growth that is not enough to mention.
- ✓ An effect size between -.1 and -.2 shows a decline that is enough to mention.
- ✓ An effect size of between -.2 and -.3 show a decline that is quite bad.
- ✓ An effect size of -.3 or greater shows a decline that is substantial.

Allowing the above interpretations to *effect size* scores for the 6th grade at the standard level for our sample school, we have the following findings.
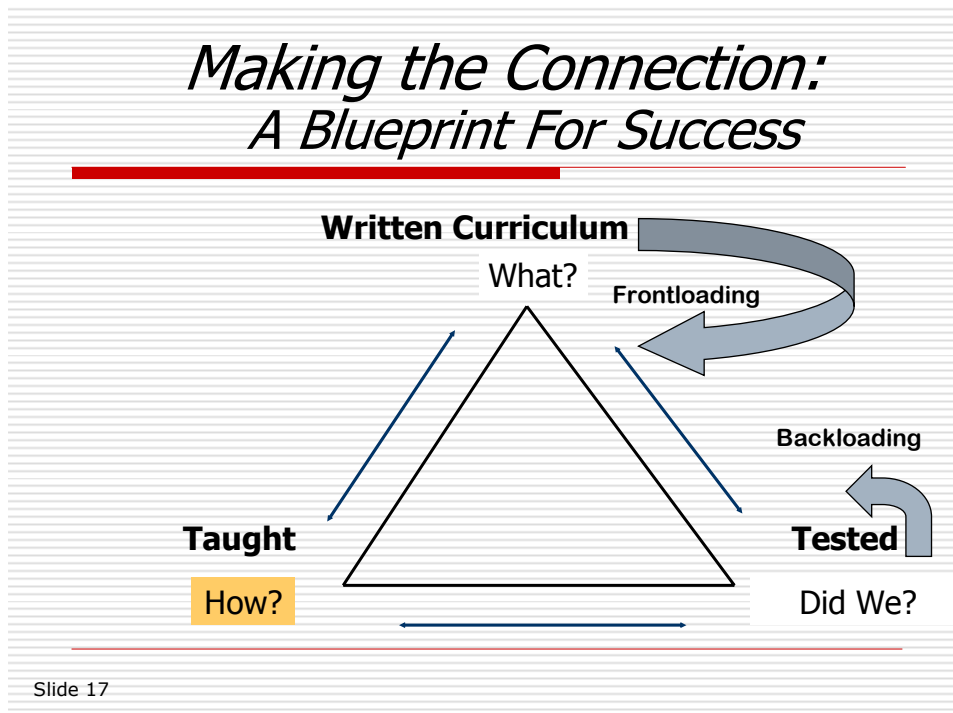
	St. Sample School	Year/IPI Score		Standard Score		Effect Size 2004 vs. 2002	Comment 2004 vs. 2002:
Grade	Subject	2002	2004	2002	2004		
6th	Rd. Voc.	79.8	93.2	-0.11	0.83	0.94	Growth Substantial
6th	Rd. Comp.	87.3	84.7	0.40	0.22	-0.19	Decline Enough to Mention
6th	Lit. Resp.	87.3	88.4	0.23	0.32	0.08	Change Not Enough to Mention
6th	Writ. Proc.	87.2	88.2	0.48	0.55	0.06	Change Not Enough to Mention
6th	Writ. App.	62.3	68.0	-0.35	0.21	0.55	Growth Substantial
6th	Lang. Conven.	84.3	91.6	-0.06	0.62	0.68	Growth Substantial
6th	Num. Sen.	85.7	86.9	0.46	0.54	0.07	Change Not Enough to Mention
6th	Compu.	78.8	90.8	0.13	0.71	0.58	Growth Substantial
6th	Alg. & Fun.	78.1	87.9	0.28	0.82	0.54	Growth Substantial
6th	Geom.	61.1	64.1	-0.45	-0.30	0.15	Growth Enough to Mention
6th	Measure.	75.9	87.3	0.18	0.81	0.63	Growth Substantial
6th	Dat. An. & Prob.	63.5	59.7	-0.27	-0.48	-0.21	Decline Quite Bad
6th	Prob. Solv.	61.8	68.4	0.22	0.50	0.28	Growth Quite Good
6th	Total Tested	29	28				

This table was created by drilling down in each subject area using the Indiana Performance Indicator (IPI), the electronic score provided on each student at the Indiana academic standard level by the state. In English/language arts, reading comprehension is the only standard area

where student achievement has declined. In mathematics, data analysis and probability is the only standard area where student achievement has declined. With this information in hand, teachers and administrators can quickly align the instructional program with student achievement needs.

Aligning program with achievement results

Educators have always tried to align testing with instruction. According to Fenwick English, in order to align instruction today, we must backload from test results because the test is now designed by the state not the instructor. The process is shown in the diagram below.

Making the Connection:
A Blueprint For Success

Written Curriculum
What?
Frontloading

Backloading

Taught
How?

Tested
Did We?

Slide 17

In the past when educators designed the test after a lesson, we had lesson-test alignment because we decided what part of the curriculum we wanted to teach and test. Now that the test is being

designed for us by the state, we need to make sure that what is taught to students from the written curriculum is aligned with the state test. This additional step is needed to make sure our curriculum decisions are based upon student achievement data. From this data, we can decide the sequence of what is taught and how we need to plan our instructional interventions. Looking at the data and aligning it with instruction is a critical part of our school improvement efforts.

Being accountable for improving achievement

Our assessments and interventions must be aligned because this will allow us to track and report student progress over time. By doing this, we will be more accountable to our stakeholders. Using the same assessments and method of analysis to determine improvement, we can track student progress and make adjustments as we implement our improvement efforts. This data analysis will give us a longitudinal view of student achievement that we have not previously had.

Using our sample school, we can compare our 2004-2005 testing results back to testing results over the last two years where the same state assessments have been used. Because not all grades were tested in the two previous years (only grades 3, 6 and 8), we only have four grade levels where we can make this comparison.

ISTEP+ Year to Previous Testing Effect Size Report
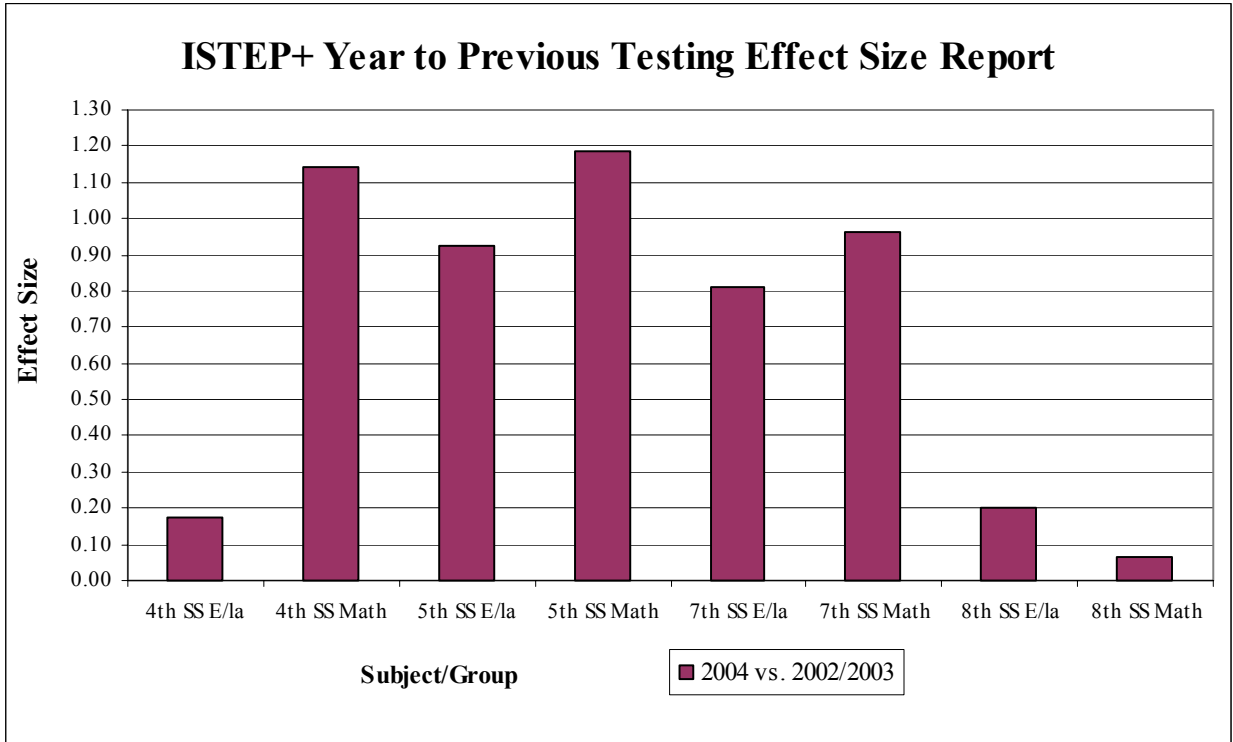
As shown in the graph, student achievement has improved at all grade levels. In five of the eight observations, improvements in student performance would be at a level which is significant given our original criteria. When we receive results back for the 2005-2006 school year, we will be able to compare growth in performance for all students except those who will be tested for the first time at grade 3. We believe this comparison will be our most important because we expect all students to improve their achievement each year.

Summary

We have focused on the use of the ISTEP to measure student performance on the Indiana Academic Standards not because it is the only measure, but because it is the primary indicator of success in our state and how the Indiana Department of Education reports statewide achievement under the No Child Left Behind Act. Schools have for years been committed to assessing the

performance of their students. We have used a number of assessments to record and report student progress – standardized tests, grades, and other measures. What we have not done effectively is to decide how we are going to analyze these results to set priorities, develop interventions, and determine if we have made a difference in student achievement. By drilling (disaggregating) down into test scores, and looking at scores over time, we have created an effective method to manage assessment data and achieve gains in student performance. Our instructional content is more closely aligned with student needs, and we can report and be held accountable for gains in improved student performance. We strongly believe that only through an assessment system with accountable will we be able to achieve the desired improvement in achievement for all students.

References

English, Fenwick W. Deciding What to Teach and Test. Corwin Press, Inc. 2000.

Fullan, Michael. "Managing the Process of Change" Presentation ASCD Teaching & Learning Conference Reno, Nevada. October 1999.

Tools for School Improvement Software Suite, North Central Association Commission on Accreditation and School Improvement, Temple, Arizona, 2002.