

# Toward a Growth-Centric Assessment Model

A white paper from Pearson Educational Measurement

Thomas H. Fisher, Ed.D.

Jon S. Twing, Ph.D.

April 2006



PEMwp\_GCA\_y06n02

*Using testing and  
assessment to  
promote learning*

## **Table of Contents**

<b>Introduction</b>	<b>3</b>
<b>Developing Growth-Centric Assessment Models</b>	<b>4</b>
<b>Preconditions for Growth-Centric Assessment</b>	<b>5</b>
<b>Growth-Centric Analysis Strategies</b>	<b>5</b>
<b>Criteria for Evaluating Growth-Centric Assessment Systems</b>	<b>6</b>
<b>Conclusion</b>	<b>8</b>
<b>About the Authors</b>	<b>9</b>
<b>About Pearson Educational Measurement</b>	<b>9</b>
<b>Endnotes</b>	<b>10</b>

## **Toward a Growth-Centric Assessment Model**

**April 2006**

### **Introduction**

There is no single definition of “value-added” in education, and there is no common method of performing the analyses. The concept is not synonymous with the Tennessee Value-Added Assessment System (TVAAS), although this is one method.

In other arenas, the definitions of “value-added” are as varied as the industries. “Economic value-added” is a company’s financial performance based on residual wealth calculated by deducting cost of capital from its operating profit (adjusted for taxes). Economic value-added attempts to capture the true economic profit of a company.<sup>1</sup>

“Value-added taxation” is used in the European Common Market as a type of sales tax borne by the consumer. It is a levy imposed on business at all levels of the manufacture and production of a good or service and is based on the increase in price or value provided by each level.<sup>2</sup>

“Value-added” agriculture is a way to add value to a farmer’s products. The farmer implements a new process or service that makes an agricultural product more valuable. One writer refers to it as, “Fry it, pie it, or put it in cider” to increase the basic value of the commodity.<sup>3</sup>

“Value-added business activities” are those that enhance the value of the company while eliminating those that do not. A retail Web site may attract viewers, but if there is no increase in sales, the activity may need to be replaced.<sup>4</sup>

“Value-added” also refers to the process by which a company systematically reviews all of its manufacturing processes and services and eliminates unnecessary steps. This is often called “lean manufacturing.” The company adds value by eliminating things that do not directly contribute to increased profit.<sup>5</sup>

“Value-added” is often used as an advertising phrase describing additional and unexpected “value” a customer may receive. For example, a hospital may assert additional value to its services by the amenities within its patients’ rooms.<sup>6</sup>

Then, in education, measuring “value-added” means monitoring individual student growth in academic skills from year to year. The gains are aggregated for individual schools and districts (and, sometimes, by teacher) with the intent of determining which school or district (or teacher) is or is not producing gains in student achievement.<sup>7</sup>

With the myriad definitions of “value-added” both in education and beyond, it is better to change the terminology to focus on the ultimate goal of testing: improved student achievement. That is the “value” that assessments need to add in education. When the different definitions of “value-added” are aggregated, it quickly becomes apparent that today’s educational environment demands a “growth-centric assessment model.”

A growth-centric assessment model must include a standard measure of performance. The standard could be the median score on a nationally normed achievement test, the overall average score on a statewide assessment test, or the average score in a district assessment test. Each of these approaches is based on current educational levels at a time when there is criticism that education in the nation is lacking. The alternative is to develop a system based on a performance standards adopted by decision-makers through a “standard-setting” (similar to socially moderated achievement standards) procedure. The No Child Left Behind (NCLB) Act requires adoption of such standards.

### **Developing Growth-Centric Assessment Models**

The design of a growth-centric model in education is directly related to the underlying assumptions and the purposes to be achieved. Entities differ in their conceptualizations of the task of measuring growth.

Student achievement results at the state, district, and school levels are routinely compared in the news media and other official reports and analyses. Typically, these are status reports comparing, for example, the fourth graders in 2003-2004 to the fourth graders in 2004-2005. Educators often complain that such comparisons are unfair because the students are in different cohorts, and the schools should not be held accountable for things beyond the control of educators.

Dissatisfaction with status reports gives motivation for analyses that follow students over time to monitor growth. Accountability programs may be designed to make adjustments for the characteristics of the schools and districts. These adjustments attempt to “level the playing field” for factors such as prior achievement, race/ethnicity, socio-economic level, mobility, gender, language proficiency, etc. Critics of NCLB object to this approach because it comes close to suggesting that some children cannot learn and provides schools a ready excuse for not trying to educate them.

NCLB requires comparisons based on status and asserts that schools are responsible for educating all children. This orientation can be honored while monitoring student progress over time but appears to provide some disincentive for “leveling the playing field.” In other words, NCLB does not encourage schools to lower expectations or goals because of the population they serve.

Growth-centric analyses may also be conducted to identify the best and worst performing schools or teachers. Researchers can then conduct in-depth observations of these schools or teachers to gain insights into successful and unsuccessful instructional strategies. This knowledge can form the basis for improved pre-service and in-service education programs.

### **Preconditions for Growth-Centric Assessment**

To implement growth-centric assessments, it is necessary for the assessment system to include:

- well-constructed student achievement tests in designated content areas.
- achievement tests in adjacent grade levels.
- student achievement scores reported on a vertical scale.
- vertically articulated content standards.
- unchanging identification numbers for students, schools, and districts.
- a data warehouse with large complex databases to track students for multiple years, even if they move from school to school.
- identification numbers assigned to teachers to permit them to be matched to individual students.
- a set of rules and procedures to address what will be done with missing records, students who move into and out of a school or district, and classrooms that are team-taught.
- procedures for entering corrections to records in the data warehouse.
- rules to address unusual situations such as a “school within a school” or a school that may straddle boundary lines between two districts.
- stable testing programs that permit the same student assessment test series to be used for multiple years even though the test support contractor may change due to competitive bidding.

### **Growth-Centric Analysis Strategies**

There are various ways to prepare growth-centric analysis models. The results of the different approaches will differ, and, in addition, the results from a growth model will differ from a status model.

Growth-centric analyses can be performed by simply monitoring student achievement growth over time. Such analyses can be summarized at the district, school, or teacher level. The average gains for each school can be compared with the average gains of students in other schools, the district, or the state. If desired, the student gain values can be aggregated by classroom teacher and compared with other teachers.

Alternatively, a school’s “expected score,” against which the current average score is to be compared, can be calculated through use of a simple regression approach. A prediction equation can be created using the current achievement level as the dependent variable (output) and various demographic variables as independent variables (inputs). Another version would include the previous year’s achievement average as another input variable. Or, the gain score between two years could be the dependent variable, and various demographic variables could be the independent variables.

The next level of complexity appears to be the use of hierarchical linear modeling (HLM). This method allows for analyses of gain scores as a function of the characteristics of the school as well as the characteristics of the classrooms and students. This method is used in Dallas, Texas.<sup>8</sup>

A more complex approach is one used with the Tennessee Value-Added System (TVAAS), which follows students over time for multiple years, monitors the teachers who taught each student, and uses a methodology for analysis known as the “linear, mixed-effects with layering” approach. The “layering” is an attempt to calculate the effects of the current teacher as well as the previous teachers. No other external covariates are used since the TVAAS assumes each child is under his/her own control.<sup>9</sup>

The more sophisticated “effects” models are mathematically complicated, require special computer programs, and sometimes result in situations in which the mathematical algorithms do not generate a

clean solution. Each output score has a standard error or area of uncertainty about it that makes it difficult, if not impossible, to uniquely separate all schools into non-overlapping values. This effectively means that the results are not likely to distinguish statistically between every school or teacher.

### **Criteria for Evaluating Growth-Centric Assessment Systems**

Legislators and business leaders are urging the educational community to embrace growth or value-added accountability systems. The intentions are good, in the sense that schools exist for the purpose of educating children, and that monitoring cohorts of students as they progress through the grade levels is useful.

Unfortunately, there is no single way to design and implement a growth-centric system. There are conflicting purposes, competing models, and confusing solutions. It is, therefore, useful to consider some criteria with which to evaluate a proposed growth-centric assessment system:

#### **1. Growth-centric analyses are believable only when the underlying testing program has technical quality, produces valid scores, and maintains a longitudinal database.**

The quality of the statewide and district assessments vary. Item quality varies and reliability may be weak. The alignment of the tests to the specified content standards may be undocumented or inadequate. Systems may have no unique student identification numbers, student demographic information is unstable, and data warehouses do not exist. Tests may not be in contiguous grade levels, and vertical scales may not be available. Cheating can occur, and schools can learn to “game the system.” The most sophisticated value-added system cannot compensate for such deficiencies.

#### **2. Adjustments to compensate for the characteristics of the student population in a school should be undertaken only with extreme care and due consideration.**

Educators often argue that schools should not be compared because their student bodies differ in ways that the school cannot control. However, NCLB mandates improved educational services for all students with the expectation that all students can become “proficient” with proper instruction. Adjusting scores can lead to lower academic expectations for some students.

#### **3. Growth-centric systems must be designed to anticipate fundamental changes in the assessment program over time.**

Statewide assessment and accountability systems are seldom stable over time, and significant changes of direction are the norm. Elected officials create statewide assessment programs to serve a public need for information, but the programs change as governors, legislators, and education officials leave office. In addition, the programs are constantly buffeted by those who are opposed to them.

Tests become outdated, norm-referenced tests may be discontinued, national norms may be recalculated, score scales may be adjusted, content standards can be changed, and cut-scores often are redefined. Contracts for testing programs are rebid, and a new contractor is hired. Any growth-centric system must be capable of withstanding these changes.

#### **4. Growth-centric systems must be transparent.**

An educational accountability system must be understandable. Depending on its design, it may be simple enough for every parent and educator to understand it, or it may be so complex that only a few researchers can understand the calculations. The former is preferable, but the latter could be acceptable as long as the system is clearly understood by researchers and statisticians so they evaluate it. To meet this criterion, the system must not be “secret,” and all of its inner workings must be visible. This is especially true if individual teachers are evaluated by the system.

#### **5. Growth-centric analyses must be verifiable.**

Regardless of the model being used, the results must be verifiable to guarantee 100 percent accuracy. There is no room for mistakes when pronouncements are made about school accountability findings or the performance of an individual teacher. The system must be amenable to independent audit for accuracy.

**6. Growth-centric analyses should not be overly complex.**

Statewide assessment and accountability programs are created to inform students, parents, educators, citizens, and decision-makers about what students know and can do. To be effective in using data to make personal decisions about success in school or policy decisions about the educational system, its programs, and its effectiveness, end-users must have a basic understanding of how the accountability system works. The primary function of a value-added system is to inform the public and not to conduct large-scale research projects. Using overly complex systems may not be as effective as using systems that allow an intuitive understanding of the results.

**7. Growth-centric analyses must fit into a framework that includes rewards and sanctions.**

It appears highly likely that various rewards and sanctions will be based on growth-centric results, directly affecting students, schools, teachers, principals, and superintendents. NCLB requires that schools be measured on their “annual yearly progress,” and there will be public disclosure of summary test results and designation of schools that are not being successful. The implication is that any growth-centric system must deliver timely data that are absolutely correct.

**8. Growth-centric systems at the teacher level should be approached cautiously.**

Schools are social organizations, and the data to describe their effectiveness can be quite elusive. It is appropriate for a school principal to monitor how effective the teachers are in helping students grow. But, in the end, if a student gains skill, it may be the result of the classroom teacher, a friendly teacher in another class, a parent helping at home, a private tutor, or the student’s own hard work. Adding complexity to a value-added system cannot distinguish who was effective with the child. For these reasons, growth-centric systems at the teacher level must be considered developmental, yielding only tentative conclusions.

## **Conclusion**

With its wide variety of definitions and approaches, “value-added” assessment is a confusing model. A clearer concept is “growth-centric” assessment, where all of the available data are used to drive student growth. A growth-centric model allows educators to use the standard units of measurement available from today’s assessment tools and other available data to manage student progress toward achievement goals. However, because of the complexities and various influences on student achievement, it is important to carefully develop and evaluate a growth-centric assessment model so that it produces results that are 100 percent accurate. This is particularly critical when communicating the results to local, state, and federal legislators who are placing increasing demands for accountability on schools. When a well-developed, growth-centric assessment model is in place, educators will have the information they need to monitor student progress, forecast future test results, and develop instructional interventions to ensure that all students achieve at the highest levels.

## **About the Authors**

### **Thomas H. Fisher, Ed.D.**

Thomas H. Fisher earned a bachelor of science in mathematics from Middle Tennessee State University, a master's of education in secondary administration from the University of Toledo, and a doctorate of education in curriculum development from Wayne State University in Detroit, Mich. He started his career as a mathematics teacher and worked for 35 years in large-scale academic assessment, teacher licensure examinations, and school accountability in Michigan and Florida. He served six years as a member of the National Assessment Governing Board.

Fisher retired from the Florida Department of Education and formed Fisher Education Consulting, LLC, where he provides student assessment and school accountability consulting services across the nation.

### **Jon S. Twing, Ph.D.**

Jon S. Twing is senior vice president, test and measurement services at Pearson Educational Measurement (PEM). With more than 20 years of assessment experience, Twing leads the company's psychometric and content development groups.

While at PEM, Twing and his staff have been responsible for all or parts of more than 15 state assessment programs, including the Texas Assessment of Knowledge and Skills (TAKS); the Texas Assessment of Academic Skills (TAAS); the Michigan Educational Assessment Program (MEAP); and the Florida Comprehensive Assessment Tests (FCAT). He earned a doctorate of philosophy and a master's of arts in educational measurement and statistics (psychometrics) from the University of Iowa, and a bachelor of science in psychology and business from Central Michigan University.

## **About Pearson Educational Measurement**

Pearson Educational Measurement, the largest comprehensive provider of educational assessment products, services and solutions. As a pioneer in educational measurement, PEM has been a trusted partner in district, state and national assessments for more than 50 years. PEM helps educators and parents use testing and assessment to promote learning and academic achievement.

PEM's full-service offerings for K-12 and other assessment organizations include PEMSolutions™ (Pearson Educational Measurement Solutions) for custom assessments, both online and on paper; PASeries™ (Progress Assessment Series™) for formative assessments; Perspective™ for performance reporting; EDWARD™ for assessment-based education data management and reporting, as well as other essential educational assessment products and services. Pearson Educational Measurement operates as a business of Pearson Education, the world's largest education company. More information is available at [www.pearsonedmeasurement.com](http://www.pearsonedmeasurement.com).

## **Endnotes**

- <sup>1</sup> Answers.com (n.d.) *Economic Value Added*. Retrieved August 25, 2005 from <http://www.answers.com/topic/economic-value-added>
- <sup>2</sup> Encyclopedia.com (2005) *Value-Added Tax*. Retrieved August 26, 2005, from <http://www.encyclopedia.com/html/v1/valueadd.asp>
- <sup>3</sup> Holland, R. (2005). *Considerations for a Value-Added Farm Business*. Retrieved August 25, 2005, from University of Tennessee, Center for Profitable Agriculture: <http://cpa.utk.edu/pptpresentations/CPA103considerfarmbus.pdf>
- <sup>4</sup> Martin, D.F. (2001). *Demystifying Value Added*. Retrieved August 25, 2005, from THE CEO Refresher: <http://www.refresher.com!/demystifying.html>
- <sup>5</sup> Industrial Solutions, Inc. (ISI) (n.d.). *Lean Manufacturing*. Retrieved August 25, 2005, from <http://www.isiworld.net/Lean.html>
- <sup>6</sup> Arkansas BlueCross BlueShield. (2001-2005). *Health Plans and Services*. Retrieved from August 25, 2005, from [http://www.arkbluecross.com/health\\_plans/](http://www.arkbluecross.com/health_plans/)
- <sup>7</sup> Hoover Institution (2002). *Education Next: Accountability Gains*. Retrieved August 25, 2005, from <http://www.educationnext.org/20022/9.html>
- <sup>8</sup> Dallas Independent School District. (2005). *DISD's Balanced Scorecard for Improvement*. Retrieved August 26, 2005, from [http://www.dallasisd.org/inside\\_disd/documents/dip/](http://www.dallasisd.org/inside_disd/documents/dip/)
- <sup>9</sup> Tennessee Comptroller of the Treasury. (2005). *Audits and Reports: K-12 Education*. Retrieved August 25, 2005, from <http://www.comptroller.state.tn.us/K12rpt.htm>

PA Series, Progress Assessment Series, PEMSolutions, Perspective and EDWARD are trademarks or U.S. registered trademarks of Pearson Education, Inc. and/or one or more of its direct or indirect affiliates. The names of other companies and products mentioned herein may be the trademarks of their respective owners.  
© 2006 Pearson Education, Inc. or its affiliate(s).