

Growth and Value-added Issues in Minnesota

Mark L. Davison

University of Minnesota

Leslie J. Davison

St. Cloud State University

Paper present to the Council of Chief State School Officers (CCSSO) meeting entitled the *Use of Growth Models Based on Student-level Data in School Accountability*, November 15 – 16, Washington, DC.

Correspondence concerning this paper should be addressed to Mark L. Davison,
Department of Educational Psychology and Office of Educational Accountability,
University of Minnesota, 178 Pillsbury Dr. SE, Minneapolis, MN 55455

Abstract

While Minnesota has not adopted a value-added approach in its statewide accountability system, both the legislature and the administration are committed to do so in the future. Meanwhile, some districts have adopted student growth as a component of their accountability or reporting system. Several such systems are built around the expectation that students will make one year's worth of growth. In discussions about growth and value-added, a number of issues have arisen: linking growth expectations to the achievement of standards, including all students in the computation of school indicators, and combining growth indicators with status indicators. These issues are described and possible solutions are discussed.

Growth and Value-added Issues in Minnesota

Mark L. Davison

University of Minnesota

Leslie J. Davison

St. Cloud State University

While there is a great deal of interest in value-added, it is not yet a part of Minnesota's statewide accountability system. In what follows, we will briefly describe the discussion at the state and district levels concerning value-added systems. Then we will discuss some possible policy objections to value-added: a lack of connection to standards and a loss of students in the computation of indicators. To address these concerns, we describe an "on-track" growth indicator tied to standards, a statewide tracking system to minimize the loss of students, and an indicator that combines status and growth so as to include as many students as does the current adequate yearly progress system.

Tekwe, Carter, Ma, Algina, Lucas, Roth, Ariet, Fisher, and Resnick (2004, p. 12) have defined the term value-added as "methods of assessment of school/teacher performance that measure the knowledge gained by individual students from one year to the next and then use that measure as the basis for a performance assessment system." As we use the term here, value-added refers to a longitudinal approach that uses a student's prior year achievement and the current year's achievement to evaluate gains over time. Demographic factors, such as eligibility for free or reduced price lunch, can be added into the evaluation in various ways to adjust for background factors that place students at risk. In what follows we will distinguish value-added, which takes students'

prior achievement into account, from risk adjustment, which takes demographic background factors into account. The term “risk adjustment” is taken from the medical field where differences in patient risk are taken into account when evaluating hospitals (Pearson & Stecher, 2004).

Minnesota and Its Districts

Somewhat different approaches are being pursued at the state and district levels. At the state level, much of the interest centers on the value-added approach of William Sanders developed at the University of Tennessee (Ballou, Sanders, & Wright, 2004; Sanders, Saxton, & Horn, 1997).

State Level

Value-added has not been incorporated into the state’s accountability system. However, in the legislature, the governor’s office, and the Minnesota Department of Education, there is a commitment to incorporate value-added into the state’s accountability system. Governor Pawlenty recently stated “As a state, it’s time to move to the next level of accountability in our schools. Based on new technologies, we now have the ability to measure individual students compared to where they were the year before. I have instructed the Department of Education to move forward in developing and implementing this “value-added” system.”

http://www.governor.state.mn.us/Tpaw_View_Article.asp?artid=1144

Minnesota statute (MS 120b.30) calls for the incorporation of value-added indicators into the state’s accountability system by 2006. At the legislature, the value-added system of Dr. William Sanders has attracted the most attention, and Dr. Sanders has been invited by the legislature several times (Ballou, Sanders, & Wright, 2004).

Some legislators have been particularly attracted to the Tennessee model because of its potential application to the evaluation of both schools and teachers.

While the state is moving forward with the development of a longitudinal value-added approach based on growth scores, it is unclear how such scores would be incorporated into the evaluation of schools. While we cannot speak for all educators in Minnesota, it is our impression that many would like to see growth indicators incorporated into the adequate yearly progress component of the state accountability system. If it could not be incorporated as the primary indicator, then maybe it could be incorporated into a safe harbor process or the school appeals process; that is, if the school did not meet its annual measurable objective, then it could still be considered as having made adequate yearly progress based on a growth indicator. Even if a growth indicator is not incorporated into the adequate yearly progress component of the state's accountability system, it could be used to identify high performing schools.

District Level

At the state level, the primary focus has been on the Tennessee value-added approach. Districts, however, have been developing other approaches based on year-to-year gain scores and growth norms for those gain scores. In Minnesota, much of this effort has been led by Dr. James Angermeyr, Director of Research and Evaluation in the Bloomington, Minnesota schools. However, similar efforts exist in the state of Idaho, and in districts elsewhere that have adopted the *Northwest Achievement Levels Test (NALT)*.

In this section we will describe the growth norm approach assuming the *NALT* as the achievement test, because growth norms already exist for this test. Once a state has

had annual testing in grade 3 – 8 for two consecutive years, the state can readily develop its own growth norms. Once growth norms are in place, a state can implement the approach described below. For growth norms on the *NALT*, see http://www.meridianschools.org/pdf/isat_growth.pdf, the website of the Meridian, Idaho public schools.

Almost half of the school districts in Minnesota have adopted the *Northwest Achievement Levels Test (NALT)* as their district norm referenced test¹. All forms of the test have been vertically equated so that a pre-test score can be subtracted from a post-test score to obtain a measure of growth or gain. If students in the district have been tested twice, either fall to spring or spring to spring, the score report for each student includes a growth score for each tested subject, current test score minus prior test score. Districts that use the *NALT* employ these growth scores as a measure of the value-added to a student's learning from fall to spring or from spring to spring.

If one adopts a growth score, then one must define a performance standard for that growth score, a statement of how much growth is expected. Typically, the expected amount of growth is one year's worth of growth defined as the mean or median amount of gain in some well defined norm group. Because of regression to the mean due to measurement error (Butterbaugh, 2003; Campbell & Kenny, 1999) and other factors, students with low pre-test scores will typically make greater gains than students with higher pre-test test scores. Consequently, the *NALT* has separate growth norms for each grade and within a grade, there are separate growth norms for students falling within various score intervals on the pre-test.

¹ Some might object to our referring to the *NALT* as a norm-referenced test because its norm group is a user group, rather than a nationally representative sample.

For each pre-test score interval and grade combination, the norm table shows the mean amount of growth in the norm group among students in that grade and pre-test score interval. As an example, the table shows that 5th grade students in the norm group scoring between 160 and 169 on the pre-test had gained, on average, 11.6 points; therefore 5th graders with pre-test scores between 160 and 169 must make at least 11.6 points of gain to be considered as making one year's growth. A 5th grader with a pre-test score between 170 and 179 would only have to make 10.7 points of growth to be considered as having made one year's worth of growth according to the table. Generally, the expected amount of growth goes down with each increasing grade and pre-test score.

To form school indicators from individual student growth scores, schools can use the growth scores and growth indicators in several ways, three of which are described here. First, the school can compute the percentage of students making at least one year's worth of growth. That is, the school can compute the percentage of students whose gain score exceeded the mean gain shown in the growth norm table for the students' respective grade and pre-test score interval.

Second, the school can compute the mean gain for students in a school and use the mean as the school indicator. In this system, each student has an actual gain score and an expected gain score. The expected gain score is taken from the growth norm table for that student's grade and pre-test score interval. An expected gain for the school can be computed by averaging the expected gain for each student in the school. If the school's average actual gain exceeds its average expected gain, the school can be said to have met its target. That is, the average gain in the school is said to meet or exceed one year's worth of growth.

Third a school can use the actual gain and expected gain (from the growth norm table) from each child to compute a percent of expected gain (PEG): the actual gain divided by the expected gain multiplied by 100; that is, $PEG = (\text{actual gain} / \text{expected gain}) \times 100$. If a child has an actual gain of 10 points, and the growth norms indicate that his expected gain is 8 points, then the child's percent expected gain is $(10/8) \times 100 = 125$. If the child just makes his/her expected gain, then PEG will equal 100. PEG will be less than 100 if the child makes less than the expected gain shown in the norm table; PEG will be more than 100 if the child achieves more than the expected gain. A school indicator can be created by averaging the PEG over all children in the school. Schools strive for an average PEG over 100.

In short, using growth norms and growth scores, schools can compute one of three indicators for schools: (a) the percentage of students meeting or exceeding one year's worth of growth, or (b) the mean growth in the school, which can be compared to the mean expected growth, or (c) the average percent of expected gain with a goal of achieving an average PEG over 100². These indicators can be disaggregated by subgroup if desired. Many districts using the *NALT* and its growth norms have incorporated growth indicators into their district accountability systems and/or into their district reporting system (e.g., see the average growth results for the Computerized Achievement Levels Test (CALT) scores in Bloomington, Minnesota

<http://re.bloomington.k12.mn.us/results>.)

Addressing Limitations of a Longitudinal Growth Score Approach

² Of these three indicators, two require vertically equated tests: the mean growth indicator and the percent expected gain (PEG). However, computing the percentage of students making one year's worth of growth does not require vertically equated tests. If scores have not been vertically equated, however, the process for computing the percentage of students making more than one year's worth of growth is different than the process described above because the necessary norm tables are different.

While there is a great deal of interest in value-added systems, both at the state and district levels, a number of concerns have been expressed, two of which will be considered here: the link to standards and full inclusion of all students.

Link to Standards: It can be argued that a value-added approach is not clearly linked to performance standards in that students may make one year's worth of growth and yet not achieve a performance standard. In our opinion, this limitation applies, not to growth or value-added approaches per se, but rather to the "one year's worth of growth" criterion for evaluating growth. In any longitudinal growth based system, there must be a benchmark against which to judge the growth. The one year's-worth-of-growth benchmark is not clearly attached to the attainment of standards.

Concern about the "one-year's-worth-of-growth" concept has led to discussion of an alternative benchmark for growth, the "on track" concept. The on-track concept requires that a student's progress be sufficient to reach a designated academic standard in a future grade should growth continue at the present rate.

The on track system requires a testing system in which tests contain content aligned with grade level standards. The tests must have a clear grade-to-grade progression of content. Furthermore, the tests must be developmentally scaled (vertically equated) over grades, so that scores on any two adjacent grades can be subtracted in order to compute an estimate of the amount that a student has grown from one year to the next. Second, for each subject area, it requires the identification of a critical grade (or grades) and a proficiency level in that grade. The critical proficiency level is an achievement target that all students are expected to reach in that critical grade. In Minnesota, students first take the high school graduation test in 8th grade, and we talk

about students being “on-track” to pass the high school graduation test on their first attempt. If an “on-track” system were built around passing the high school graduation test in 8th grade, then the critical grade is 8th grade and the proficiency level is the passing score on that test, a score of 600. Third, the indicator requires two years of data from each student. That is, each student must have been tested last year and the current year somewhere in the state.

In this system, students are evaluated, not on whether they have achieved a fixed proficiency standard in the current year, but rather on whether they can be considered “on track” to reach the proficiency level in the critical grade. A student is considered “on track” if the student’s achievement gain from last year to this year is sufficient to reach the critical proficiency standard in the critical grade should growth continue at the present rate.

As an illustration, consider a reading test system in which the tests have been vertically equated to form a developmental scale over 3rd through 8th grades. Over those grades, scores can range from a low of 100 to a high of 1,000 scale score points. If a student achieves a score of 600 in 8th grade, the student has met the high school graduation requirement.

To see how this “on track” evaluation would work, consider a 5th grade student, Rhonda, who earned a score of 450. As a fifth grader, Rhonda has three more years to reach the high school graduation benchmark of 600 in 8th grade; to reach that benchmark in three years, Rhonda’s reading score must increase by 150 points, 50 points per year over the next three years. When Rhonda takes the reading test in 6th grade, she will be considered on track if she scores at least 50 points higher than she did in 5th grade, i.e., a

6th grade score of at least 500. If she improves by less than 50 points, she will be considered “not on track.” In this system, the value-added over the school year is a simple gain score, this year’s score minus last year’s; and the student is considered on track if the amount gained exceeds the score increase needed per year to reach the proficiency level in 8th grade.

This “on-track” system sets unequal growth and proficiency targets for students based on their achievement scores last year. That is, each student has a unique growth target and achievement score target based on his/her grade level and prior year achievement score. Appendix A contains a formula for computing a student’s target achievement level in any given year.

If the “on-track” criterion is used to evaluate growth, then the value-added system is clearly tied to a performance standard, i.e., the proficiency level in the critical grade (8th grade in the illustration above). However, it does not require that young students immediately catch up but rather gives them several years--until 8th grade in the illustrations--to catch up. The on-track value-added system recognizes that some students may start out well behind and may require more than one year to catch up (Davison, Seo, Davenport, Butterbaugh, and Davison, 2004; Hart & Risley, 1995; Lee & Burkham, 2002; Snow, Burns, & Griffin, 1998; <http://www.myschools.com/News/2000/csab.htm>). At the same time, its goal is to make sure that no child falls behind in the sense of failing to meet the critical grade’s target proficiency standard.

One goal of a value-added system is to level the playing field for schools with large numbers of disadvantaged and initially low achieving students. The on-track

concept does so by taking the student's prior achievement into account. Except in the critical year, students with lower achievement scores in the prior grade have somewhat lower achievement targets in the current grade. This has the effect of partially leveling the playing field for schools with large numbers of disadvantaged and initially low achieving students. However, the on-track concept does not completely level the playing field because it requires greater growth from students with lower achievement scores the prior year. The on-track system does not seek to completely level the playing field because doing so means abandoning the goal of having all children reach standards. While this system does set different performance targets based on prior achievement, it does not set different performance targets based on ethnicity, economic status, English proficiency, or disability. While the example above was constructed around the goal of having all children passing a high school graduation test on their first attempt, a critical grade and proficiency level can be chosen based on other goals: having all students ready for higher education, having all students ready to read a high school textbook before entering high school, having all students ready to take algebra by a designated grade, or having all children reach the state's 8th grade proficiency level (when annual testing at every grade ceases in many states).

Inclusion of All Students. Because a longitudinal system requires at least two years of data from each student, fewer students can be fully included in the calculation of school indicators. However, this loss of students can be mitigated by implementation of a tracking system at the state level, rather than at the school or district level.

Furthermore, the longitudinal approach can be combined with a status based approach to yield a system that is just as inclusive as the current adequate yearly progress system.

Student Tracking Systems. In states with a system of uniform identification numbers that follow the student who moves from school to school within the state, the loss of data can be minimal. That is, one can base each school's indicator on all students enrolled in the school on October 1 and who took the test somewhere in the state last year. The only students lost for lack of test data last year will be ones who were not in the state last year or for whom their last year's test score cannot be located. Even though the indicator requires student test scores for two consecutive years, students need not be enrolled in the same school or district for two consecutive years to be included in the calculation of the indicator for that school or district.

Combining On-Track and Status Indicators. All states have (or soon will have) proficiency standards at grades three through eight and at some high school grades. It is possible to use these grade-by-grade proficiency standards along with a critical grade proficiency standard to produce school indicators that combine elements of status and growth. For instance, one can consider a student to have succeeded if they meet one of two conditions: (a) the student scores above his/her grade level proficiency standard or (b) the student is on track to meet the proficiency level in the critical grade.

Such a combined status and longitudinal growth indicator is highly inclusive. All students who have been tested in the current year can be used to compute the indicator. Students who have a test score for this year, but not last year, have only one way to succeed; they must score above their grade level proficiency standard. Students who have a test score for both years have two ways to succeed: (a) achieving their grade level standard or (b) achieving on-track status. The number of students included in the computation of this hybrid status/on-track indicator will be the number of students tested

in the current year. Therefore, the number included will be virtually the same as for the status indicators in the current adequate yearly progress system.

Summary and Conclusions

Longitudinal growth indicators of school and district performance have attracted a great deal of interest in Minnesota. At the state level, the value-added system of William Sanders (Sanders, Saxton, & Horn, 1997) has received the most attention. Districts, however, have developed a variety of approaches around the concept of growth norms: the percentage of students making one year's worth of growth, the average amount of student growth as compared to the expected average, and the average percentage of expected growth (PEG).

In the discussions around longitudinal approaches, various concerns have arisen including a concern that the growth approach may not be clearly linked to standards and may not be as inclusive as the current adequate yearly progress approach. Whether or not a longitudinal approach is clearly linked to standards would seem to depend on the criterion used to evaluate growth. If an "on-track-to-meet-standards" criterion is used as a benchmark against which to evaluate student progress, then the longitudinal growth approach can be as clearly linked to standards as is the current adequate yearly progress approach. Because longitudinal approaches require two years of data, they may not be as inclusive as the current adequate yearly progress approach. However, a statewide record keeping system for tracking students over time can minimize loss of students as compared to district wide or school wide student record systems. Further, longitudinal

indicators can be combined with a status indicator to create a system that is just as inclusive as the current adequate yearly progress system.

In evaluating longitudinal, growth based indicators, we must be careful to distinguish between properties of those indicators and properties of the criteria used to evaluate those indicators. For instance, it has been argued that, by taking student prior achievement into account, longitudinal growth based measures level the playing field for schools with large numbers of at-risk students (Kupermintz, 2003; Sanders & Horn, 1998). We are beginning to suspect, however, that the degree to which a growth based system levels the playing field will heavily depend on the benchmark used to evaluate growth. For instance, in an as yet unfinished master's project, Johnson (in progress) is finding that the poverty concentration of a school is more highly correlated with the percentage of students "on track" than with the percentage of students "making one year's worth of growth." That is, the on-track indicator does not level the playing field for high poverty schools to the same extent as does a "one-year's-worth-of"-growth" indicator, yet both of these indicators use the same longitudinal growth data; they differ only in the criterion used to evaluate longitudinal growth. To some degree, any leveling of the playing field arises from the criterion used to evaluate growth and is not an inherent property of growth itself.

Likewise, it can be argued that growth indicators are inconsistent with standards. We suspect that this statement says more about the criterion used to evaluate growth than it does about the growth approach itself. The on-track concept illustrates that a growth-based approach can be clearly linked to a standard.

While the on-track approach links growth to a future standard in a critical grade rather than the current year grade, using a future grade standard may not be unreasonable. There is an increasing body of evidence to suggest that some students from disadvantaged backgrounds enter kindergarten and first grade less well prepared than their peers (Davison, Seo, Davenport, Butterbaugh, and Davison, 2004; Hart & Risley, 1995; Lee & Burkham, 2002; Snow, Burns, & Griffin, 1998; <http://www.myschools.com/News/2000/csab.htm>). In the words of Lee and Burkham (2002), such students are behind at the starting gate; that is, they are behind even before entering school. Inequalities of cognitive ability, i.e. gaps in achievement or school readiness, exist before school begins. Students who enter Kindergarten with limited English proficiency are the most obvious example, but such children are not the only example.

For children who enter school with limited readiness, the task faced by schools is not simply one of preventing the students from falling behind; they are already behind. Rather the task is to help those students catch up. To catch up, students with low readiness must make *more* progress than their better prepared peers. Davison et al. (2004) present data suggesting it is difficult, or at least rare, for groups of children from disadvantaged backgrounds to make so much more progress than their more advantaged peers that they do indeed catch up in a short amount of time (e.g., two years). Individual children may do so, but they are the exception.

Lee and Burkhan (2002, p. 1) conclude that “We should expect schools to increase achievement for all students, regardless of race, income, class, and prior achievement. But it is unreasonable to expect schools to completely eliminate any large

pre-existing inequalities soon after children first enter the education system, especially if those schools are under-funded and over-challenged.” The on-track value-added indicator implicitly recognizes that some students do enter school with lower readiness or lower achievement. We should expect those children to catch up, but catching up may take more than a single year. An on-track indicator requires that such students make more progress than their initially higher performing peers as they must if they are to catch up. It sets a critical grade at which we expect such students to reach standards; and until that grade, the indicator longitudinally monitors student progress toward that standard by evaluating whether each student is on track to meet the standard at the critical grade.

Appendix

In the on-track approach, each person has an individualized achievement target that is based on the student's score last year, the number of years remaining until the student reaches the critical grade, and the critical score:

$$\text{Current Year Target Score} = \text{Last Year Score} + \frac{\text{Critical Score} - \text{Last Year Score}}{\text{Critical Grade} - \text{Last Year Grade}} \quad (1)$$

The ratio in the last term on the right side of Equation 1 gives the current year target gain for the student.

Using Equation 1, a proficiency target is generated for each student. For instance, for a 5th grade student whose last year score equaled 200, whose critical score is 600, whose critical grade is 8th, and whose last year grade was 4th, the current year target score according to Equation 1 is

$$\text{Current Year Target Score} = 200 + \frac{600 - 200}{8 - 4} = 300 \quad (2)$$

If, in the current year, the student scores above the target, the student is said to be “on track.” If the student falls below the target, the student is said not to be “on track.” An indicator would be calculated for each school as the percentage of students on track.

References

Balou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics*, 29(1), pp. 37 – 65.

Butterbaugh, D. J. *Differential reading and mathematics growth rates among demographic groups: Fact or artifact?* Unpublished doctoral dissertation, University of Minnesota, Minneapolis, MN. (UMI Number 3113167).

Campbell, D. T. & Kenny, D. (1999). *A primer on regression artifacts*. New York: Guilford Press.

Davison, M. L., Seo, Y. S., Davenport, E. C., Jr., Butterbaugh, & Davison, L. J. (June, 2004). When do children fall behind? What can be done? *Phi Delta Kappan*, 85(10), 752 – 761.

Elmore, R. F. (2004). Conclusion: The problem of stakes in performance-based accountability systems. In S. H. Fuhrman & R. F. Elmore (Eds.), *Redesigning accountability systems for education*. New York: Teachers College Press, pp. 73 – 95.

Hart, B. & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H. Brookes.

Johnson, K. E. (in progress). *Do value-added models level the playing field for high poverty schools? A comparison of four school indicators*. Unpublished paper, University of Minnesota, Minneapolis, MN.

Kupermintz, H. (2003). Teacher effects and teacher effectiveness: A validity investigation of the Tennessee value-added assessment system. *Educational Evaluation and Policy Analysis*, 25(3), pp. 287 – 298.

Lee, V. E. & Burkam, D. T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. Washington, DC: Economic Policy Institute.

Linn, R. L. (2004). Accountability models. In S. H. Fuhrman & R. F. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 73–95). New York: Teachers College Press.

Pearson, M. & Stecher, B. (2004). Risk adjustment methods in health care accountability. In B. Stecher & S. N. Kirby (Eds.), *Organizational improvement and accountability: Lessons for education from other sectors* (pp. 95 – 105). Santa Monica, CA: Rand. .

Sanders, W. L. & Horn, S. P. (1998). Research findings from the Tennessee value added assessment system (TVAAS): Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education*, 12(3), 247 – 256.

Sanders, W. L., Saxton, A. M., Horn, S. P. (1997). The Tennessee value-added assessment system: A quantitative outcomes-based approach to student assessment. In J. Millman (Ed.), *Grading teachers, grading schools. Is student achievement a valid evaluation measure?* (pp. 137 – 162). Thousand Oaks, CA: Corwin.

Snow, C. E., Burns, S., & Griffin, P. (Eds., 1998). *Preventing reading difficulties in young children*. Washington, D. C.: National Research Council.

Tekwe, C. D., Carter, L. R., Ma, C., Algina, J., Lucas, M. E., Roth, J., Ariet, M., Fisher, T., & Resnick, M. B. (2004). An empirical comparison of statistical models for value-added assessment of school performance. *Journal of Educational and Behavioral Statistics*, 29(1), 11 – 35.