

Opinion: Teacher Evaluation and Retention of Teachers Based On Student Achievement in Rural Alaska

Gary H. Jacobsen

Abstract: Many states, including Alaska, have recently developed a requirement to include student achievement in the performance evaluation of teachers (State Board of Education Policy 4 AAC 04.200). The view of classroom teachers, administrators, and teacher educators is that there need to be multiple measures of student achievement and, in some cases, multiple years of data to determine the achievement level of students. There are many variables affecting student achievement over which the teacher has no control, and these may negatively skew the achievement levels used in the teacher's performance evaluation. These variables will be identified and examined in this article. This application of student achievement measures through standardized achievement affect the recruitment and retention of teachers in Rural Alaska, particularly in schools where there are very low performing students as measured by standardized tests.

Introduction

As a former teacher, principal, assistant superintendent, and superintendent, I have been involved in education in Rural Alaska over a span of twenty years, and I have seen a number of significant changes. Recruiting and retaining qualified staff members is perhaps the most important task for school administrators and school districts. In the early to mid-1980s, replacing departing teachers with equally qualified teachers was challenging, but could be accomplished because there was a large enough pool of applicants. Over the last decade this important endeavour by school administrators has become increasingly difficult. There are several reasons for this change including stagnation of salary schedules, which is a result of State funding levels. Changes to the retirement system have also played a part in discouraging teachers from applying for positions in Alaska. When

these are added to concerns related to living conditions, cultural differences, and isolation and remoteness from family and friends in the Lower 48, many prospective teachers have sought employment elsewhere leading to a limited pool of qualified applicants in Alaska. This situation has been exacerbated with the requirement for teachers in very small schools to be highly qualified in multiple fields of instruction. Adding student achievement information to the teacher evaluation process is going to make a difficult situation even more challenging since, historically, student achievement levels in Rural Alaskan schools have been lower when compared with "road system" and urban school districts (DEED).

There has been little published research looking at the effects of incorporating student achievement data in teacher evaluation. Most of the studies involve models developed by individuals, such as Robert Marzano, or models developed by various states, such as North Carolina and Tennessee.

However, Wright, Horn & Sanders (1997) looked at determining the magnitude of teacher effects on student academic growth and at the same time looked at the variables of heterogeneity among students, classroom size, and student achievement level. They found that heterogeneity, classroom size, and student achievement level had very little effect on student academic growth, while the teacher effects were dominant in factors affecting student academic gains. Thus, teachers make a difference in student academic gains (Wright, Horn, & Sanders, 1997).

Another study looked at the effect of individual teachers on student achievement and found a great variety among the teachers studied (Rockoff, 2003). This research showed the variability among teachers and how they affect student achievement. There are many variables that teachers have no control over, but they try to overcome those that present obstacles to student achievement. Some teachers are more successful than others, but their efforts may not show up until later as the student progresses through the grade levels.

Many commentators have written opinions or descriptions of various models, but most agree there are variables other than just student achievement to measure teacher effectiveness. Most of these articles are not written in scholarly journals, but more often published on the Internet under various websites that encourage opinions such as a blog by Scott McLeod and an Internet posting by Lindsay Humphery. Resolving the controversy over the inclusion of student achievement in teacher evaluation will have to wait until there is a body of research that either supports or repudiates the need for the inclusion of student achievement data in teacher evaluations.

Policy Changes

As test scores and student graduation rates in Alaska have fallen over the last decade, accountability for student learning has become an ever-increasing concern in public education as the public and policy-makers have initiated programs and policies to hold educators accountable for student achievement. Accountability of student achievement efforts has resulted in increased types and numbers of assessments administered to students every year, where some students might have as many as twelve days of testing during a single school year. This increase in accountability through the assessment of student achievement has been driven by federal grants and, in particular, the Leave No Child Behind Act (2001).¹

To policy-makers, it seems the drive for increasing student achievement as measured by standardized test scores has stagnated the last couple of years. This has caused policy-makers, educators, and administrators to examine other areas of the instructional process to determine the implications of each on student achievement. Jonah E. Rockoff (2003) states that “school administrators, parents, and students themselves widely support the notion that teacher quality is vital to student achievement, despite the lack of evidence linking achievement to observable teacher characteristics” (1). His observation supports the notion that the classroom teacher has a major impact on student learning and has resulted in discussions focused on measuring the effectiveness of teachers as it relates to student achievement (Goe & Stickler, 2008). Policy-makers in the US have brought pressure on state departments of education and local school districts to adopt measures requiring that student achievement be a part of the teacher evaluation process (NCTQ, 2011).

For decades, formal evaluation of teachers as conducted by the assigned school administrator has focused on the teacher’s knowledge of content, delivery methodology, classroom management, and communication with educators, parents, and community. Although student achievement has not been explicitly included as a component of a teacher’s evaluation, many school administrators, myself included, looked at student achievement as a part of the overall effectiveness of the school and the educational process. Student achievement has recently been included as a part of the teacher evaluation system when Race to the Top (RTTT) made it a priority as a part of the American Recovery and Reinvestment Act (ARRA).² Part of the criteria for RTTT funding required that states adopt measures to eliminate any barriers that prevent linking student achievement to teacher and principal performance evaluations (USDE, 2009). This formalization of the role of student achievement in a teacher’s evaluation has raised questions about

how and when to measure student achievement, what instruments should be used, and what proportion of the overall evaluation should be based on measured student achievement. The unanswered questions about student achievement and teacher evaluation have resulted in several different models being available for consideration by state agencies and local school districts, while trying to determine how best to incorporate student achievement data into the teacher evaluation process.

At its June 2012 meeting, the Alaska State Board of Education & Early Development adopted the Educator Accountability Policy (4 AAC 04.200) that requires Alaska school districts to revise their evaluation systems to include student learning data. The implementation timeline has the new evaluation system in place during the 2015-16 academic year with 20 percent of a teacher's evaluation based on achievement levels of his/her students and the same for the 2016-17 academic year. By 2017-18 the percentage increases to 35 percent, and increases again in 2018-19 to 50 percent. Individual school districts must develop a model approved by the Department of Education & Early Development (DEED) prior to the 2015-16 school year.

Models

One type of model that has gained attention within the discussion of student achievement and teacher evaluation is the *value-added model* developed through the National Comprehensive Center for Teacher Quality (McCaffrey, et al., 2003). An alternative model is one developed by Robert Marzano, prodigious author of educational reform and CEO of Marzano Research Laboratory at Englewood, CO (Marzano, 2012). The value-added models have similar components and are designed to gauge the effectiveness of a particular program, teacher, or administrator on a student's academic progress. This is done through a comparison of a student's projected test scores (on class, district, state, or national assessments) against actual test scores earned at the end of the school year or an alternative time period. The difference between the projected and actual scores reveals a student's growth. A student's growth data are a significant part of the evaluation of a teacher's performance, but it is used in conjunction with other evaluative tools (SREB, 2011). Other recommended components of teacher evaluation are planning, instructional strategies, professional behaviours, and delivery of content, all of which should be considered along with student academic growth for a comprehensive evaluation of teacher's performance.

The Marzano Teacher Evaluation Model is based on an aggregation of elements that have traditionally been shown to correlate in a positive way with student academic achievement. The model consists of four

domains with sixty-one elements within the domains (Marzano, 2012). The domains are Classroom Strategies and Behaviours, Preparing and Planning, Reflecting on Teaching, and Collegiality and Professionalism. Embedded in the various elements are components of planning, instructional strategies, and assessments, which would require teachers and administrators to use student achievement data as a means of determining the effectiveness of the teacher. Both models are designed to use multiple measures of student growth and achievement as a part of the evaluation process. As well, both models encourage flexibility and adaptability to allow school districts and states to adjust the model to meet their identified needs (Marzano, 2012).

Districts and states face a number of challenges as they develop a process for teacher evaluation that includes student achievement data as an integral part of evaluation. First, will be determining what measures of student achievement will be utilized as a part of this process. Alaska's policy requires at least two measures of student achievement but no more than four. Second, will be the issues of reliability and validity of the assessment instruments to be implemented within the evaluation process. Third, will be determining the role of subject areas not included in the state testing battery, which includes math, reading, writing, and science for selected grade levels. Whatever assessment instruments are utilized, they should be examined as an ongoing measure of a student's academic growth rather than as a single snapshot on a particular day.

The final challenge is the need to utilize multiple years of data to provide a more comprehensive view of a student's academic growth. This area will be of particular importance to small, rural schools in Alaska, which frequently experience very high rates of teacher turnover.

Alaska Schools

Schools in Rural Alaska are remote, isolated, and for the most part very small—ranging in size between 10 and 250 students in grades K through 12, with many of the schools having fewer than ten teachers. Many classrooms are multi-graded and have teachers presenting instruction in multiple content areas. Teachers in these remote schools must be identified as highly qualified in two content areas even though they might be teaching in three or four different areas. Many of the students and their families live a subsistence lifestyle where students may be absent from school for periods of time in order to participate in these activities. Although school schedules and calendars attempt to be sensitive to the cultural mores of the community, there are still times when the school schedule and cultural activities are at cross-purposes.

For a variety of reasons, the turnover of teachers can frequently be up to 100 percent from year to year.

The number of students in these schools is so small that it is very difficult to determine statistically the growth of student achievement as it relates to the teacher evaluation system. In addition, teachers within rural sites frequently move between districts, which makes it more difficult to have a consistent process for incorporating student achievement data into teacher evaluation processes.

Although students are spending considerable time being tested during the school year, not all tests are being utilized to determine a student's academic growth or achievement level. There is an assumption that all school districts in Alaska have developed curricula that are based on the Alaska State Standards, and therefore the Standards Based Assessments (SBA) that are aligned with the state standards are aligned with all districts' curricula. However, the standards are written quite broadly, which has resulted in a wide variation in what each district requires students to know, raising the issue of how well the district's curricula actually do align with the SBAs.

Testing

In Alaska, student achievement for students in grades three through ten is determined through SBA results. This one test, on one day, is used in a wide variety of ways to say what a student has learned and how that student has progressed, without consideration for the number of variables that can and do affect the student's performance on that one day, on that one test. This raises questions regarding student absences during the test, the impact of absences on student achievement, and the possible reflection on the effectiveness of the teacher. Perhaps the most notable of these variables is that the SBAs and other standardized tests are paper-and-pencil tests, which are limited in the type and scope of knowledge they can measure. Critical thinking, creativity, and application of problem-solving strategies are examples of student learning that cannot be measured with a paper-and-pencil standardized test. In addition, paper-and-pencil tests fail to measure many social skills such as co-operation, respect, and social-emotional learning as well as traditional knowledge and cultural ways of learning. Not all students perform at their best on a paper-and-pencil test, yet this is the one measure used for all students.

SBA tests are limited to Reading, Writing, and Mathematics tests in grades 3 through 10, and Science in grades 4, 8, and 10. Other content areas are not assessed, which sends a message that these content areas are not

important. This raises the issue of how would student achievement levels be measured in these content areas and where would data on student achievement come from for those teachers in these untested content areas such as social studies, art, music, physical education, and career and technical education. Without standardized tests in these content areas, these teachers face greater uncertainty as to what measures of student achievement might be used as a part of their evaluations.

Alaska has established four categories to identify the effectiveness of teachers—exemplary, proficient, basic, and unsatisfactory. Each district must report the number and percentage of educators at each of the overall performance levels to DEED. When sample sizes are very small, as will be the case for teachers in small village schools, confidentiality will be of critical concern. Policy-makers want information presented in terms of numbers that are easily understood and allow for comparisons between different groups to determine who is better and who is lacking. When presented in this manner, the data lends itself to comparisons that are simplistic in nature and do not present a picture of student achievement that is both clear and complete. Alaska makes every effort to assure standardized tests are culturally sensitive; however, students' experiences in the small, rural communities are very different from those students living in urban areas. These different life experiences can and do show in the results of standardized tests. Will the results of these standardized tests reflect an opportunity for educational equality or will the achievement levels of students in rural areas be a reflection of the different lifestyles rather than the effectiveness of the teacher?

There are a number of variables that must be considered when analyzing the results of student achievement tests. These variables include test-taking skills, reading skills, cultural background, and the student's environment. When creating tests, student test-taking skills are not a consideration in the development of those tests. Many students do not have the standardized test-taking skills needed to effectively demonstrate their achievement level. Teachers can help by improving the test-taking skills of students, but that comes at a price in the amount of time available for the delivery of content and other skills, which are a part of student learning. Although efforts to control for cultural bias and student environment on standardized tests are considerable, a number of unintended biases remain.

Another aspect that impacts student achievement is the amount of time a student spends in that teacher's classroom. Irregular attendance and mobility by a student disrupts the instructional process in the classroom; and therefore, the standardized test results do not reflect nor predict the

effectiveness of the teacher. Using multiple means of evaluating a teacher's performance would help alleviate the problem of poor test-taking skills, below-grade-level reading, and inconsistent school attendance of some of the students.

As the State and local school districts move forward with this initiative, it is crucial the variation between urban and rural educational settings be given full consideration as models and processes are developed. With schools and districts that consist of small numbers of students, it is important to have multiple years' worth of data to make decisions about teacher evaluations. Unfortunately, in rural districts where teacher turnover is consistently high, this will present a significant challenge. One approach to meeting this challenge could be multiple measures of student achievement within each year.

In all but a few rural schools where there are small enrolments, teachers are assigned multiple subject areas in order to prepare all of the students for the SBAs as well as to meet the graduation requirements. In many cases, teachers are teaching in subject areas with minimal preparation and training. This may result in lower student achievement in those areas of under preparation. At the same time, students may be achieving at high levels in subject areas where the teacher has more than adequate preparation and training. Therefore, the question is, what data do you use for evaluating the effectiveness of the teacher.

For many school districts, professional development takes two different forms. One form of professional development is the one-size-fits-all, where all certified staff members participate in the same professional development. This is appropriate when a new initiative is being implemented in the school district; everyone needs to have the same information and training at the same time. The other form of professional development is for the individual, and typically is selected by the individual teacher to meet individual needs. When student achievement data are included for consideration in the teacher's evaluation, it will be important to provide professional development that will address specific areas identified to improve the quality of instruction provided by each teacher. It is important to distinguish between improving teacher quality and implementing a plan of improvement. Improving teacher quality looks at improving student achievement through improved instruction, while the purpose of a plan of improvement is to address deficiencies in a teacher's performance of job responsibilities.

Implications for Rural Alaska

The State of Alaska is moving forward with requirements to include student achievement data in the teacher evaluation process. The challenge for school districts and educators is to provide insight into the process in order to create a model that will be flexible enough to address the needs of the larger, urban districts as well as the needs of the small, rural districts. Each setting has its own unique situations, which will require a teacher evaluation process that will accommodate the individual needs of all schools and districts within the state. Specific challenges will be to recognize the impact of student absences and student mobility, to recognize cross-cultural differences, and to include longitudinal data and multiple measures rather than just a brief snapshot. Without the focus on longitudinal data and multiple measures, rural districts could face even greater challenges in recruiting and retaining qualified teachers. A final phase in the development of the process must focus on the types of professional development that are available for teachers to address specific needs for each teacher.

Author

Gary H. Jacobsen is assistant professor of secondary education in the School of Education at the University of Alaska Fairbanks.

Notes

1. The Elementary and Secondary Education Act (ESEA) 2001, also known as the Leave No Child Behind Act, had bipartisan support with the aim of reducing the achievement gap between students in successful schools and failing schools.
2. American Recovery and Reinvestment Act (ARRA) 2009 was an economic stimulus package to respond to the Great Recession of 2008. Its primary object was to save and create jobs, and its secondary objective was to provide temporary relief programs for those most impacted by the recession, which included education. Race to the Top (RTTT) provides funds for competitive grants to encourage and reward states that are creating conditions for educational innovations.

References

- Goe, Laura, and Leslie M. Stickler. (2008). "Teacher Quality and Student Achievement: Making The Most of Recent Research." Washington DC: National Comprehensive Center for Teacher Quality. Retrieved September 14, 2012 from www.tqsource.org/publications/March2008Brief.pdf.

- Marzano, Robert. (2012). "Bridging the Gap Between Teacher Evaluation, Leadership Evaluation, and Student Achievement." Retrieved September 19, 2012 from <http://www.marzanoevaluation.com>.
- McCaffrey, Daniel F., Daniel Koretz, J. R. Lockwood, and Laura S. Hamilton. (2004). "Evaluating Value-Added Models for Teacher Accountability." Santa Monica, CA: RAND Corporation. Retrieved from <http://www.rand.org/pubs/monographs/MG158>.
- National Council on Teacher Quality (NCTQ). (2011). "Executive Summary. State of the States: Trends and Early Lesson on Teacher Evaluation and Effectiveness Policies." Retrieved from www.edweek.org/media/nctq_stateofthestates_embargoed.pdf.
- Rockoff, Jonah E. (2003). "The Impact of Individual Teachers on Student Achievement: Evidence From Panel Data." Retrieved from <http://scholar.google.com/scholar>.
- Southern Regional Education Board (SREB). (2011). "Focus on Teacher Reform Legislation in SREB States: Evaluative Policies." Retrieved September 14, 2012 from publications.sreb.org/2011/11S07_Focus_Teach_Eval.pdf.
- U.S. Department of Education. (2012). "Race to the Top, Executive Summary." Retrieved from <http://www2.ed.gov/programs/racetothetop-district/index.html>.
- Wright, S. Paul, Sandra P. Horn, and William L. Sanders. (1997). "Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation." *Journal of Personnel Evaluation in Education* 11:57-67.