Analysis of the Virginia Teacher Evaluation System Through a Functionalist Lens

Brendan Bourdage, Kelly Erickson, Yi Hua

Abstract

Effective teacher evaluations serve the dual function of being both summative and formative (Danielson & McGreal, 2000). The current evaluation tool in Virginia uses a traditional approach of incorporating standardized test scores as forty percent of a measure of teacher effectiveness. By way of a literature review, this article applies a sociological perspective to the process of teacher evaluations in Virginia through a functionalist lens. Additionally, an examination of the pros and cons of a traditional versus authentic student assessment model is included. Evidence gathered suggests the current Virginia teacher evaluation tool fulfills the summative, but not the formative, function. This discrepancy could have lasting negative ramifications if not remedied, as the need to support teachers in achieving maximum professional performance is central to the educational system serving its inherent function of producing competent citizens capable of contributing to society.

Keywords: *authentic assessment, functionalism, teacher evaluation, traditional assessment*

Both the No Child Left Behind Act of 2001 (NCLB) and Race to the Top of 2009 (RTTT) emphasized the importance of improving teacher quality as a means of enhancing student performance. Moreover, RTTT proposed using students' academic growth over an academic year as an indicator to measure teacher effectiveness (U.S. Department of Education, 2009). In response to the law and solid research evidence (e.g., Muñoz & Chang, 2007; Rivkin, Hanushek, & Kain, 2005; Stronge, 2010), Virginia has connected teacher performance to student academic progress in teacher evaluation, basing 40 % of teachers' evaluations on student academic progress (Virginia Department of Education [VDOE], 2015). Virginia's evaluation model incorporates the characteristics of both traditional and authentic assessments serving the dual function of being both summative and formative (Danielson & McGreal, 2000). "Evaluation" and "assessment" are used interchangeably in this text. Simply speaking, traditional assessments focus on using quantitative indicators, such as

student academic achievement, to assess teachers, while authentic assessments emphasize assessing teachers' abilities through multiple aspects and data sources (Tanner, 2001). However, it is questionable whether the Virginia model appropriately balances "traditional" and "authentic," and how well it serves both the summative and formative functions. Starting with the sociological context in which functionalism is embedded, this article analyzes the Virginia evaluation model with respect to both traditional and authentic assessments. The authors selected sociology as the context through which to analyze this question, because perceptions of teacher evaluation are socially constructed by different social groups, such as teachers, students, parents, and education scholars, and each stakeholder has their own expectations for the teacher evaluation process. The outcome of teacher evaluation influences many aspects of social life and educational development within a society. Moreover, functionalism provides the authors a solid theoretical basis to explore the functions of teacher evaluation given different standards and educational contexts. In this article, the pros and cons of traditional and authentic assessments are discussed. Evidence gathered suggests that the current Virginia teacher evaluation tool fulfills the summative, but not the formative, function. This discrepancy could have lasting negative ramifications if not remedied, as the need to support teachers in achieving maximum professional performance is central to the educational system serving its inherent function of producing competent citizens capable of contributing to society. In the end, the authors summarize the concerns regarding the Virginia evaluation model.

Context of Sociology

Sociology is the study of social structures and institutions. The discipline is rooted in a "positivist tradition and aims to be led by evidence, studying individual lives to develop generalizations applicable to the greater society" (Bruce, 1999, p. 57). Modern sociology has evolved over time, making its first appearance in the philosophical writings of Greek and Arab philosophers in the 14th century. The discipline flourished in the 20th century after religion's stronghold as the primary source for legitimate social order was broken, and the societal ramifications of the Industrial Revolution took root across Europe.

Several notable figures have influenced the evolution of sociology, but three individuals in particular, each with differing perspectives on formative societal issues, stand out as having made substantial impacts in defining the theoretical underpinnings of the discipline. Karl Marx (1818-1883) believed that ongoing conflict between two social classes, the bourgeoisie and the proletariat, would lead to revolution, and that communism would eventually replace capitalism (Bruce, 1999). Max Weber (1864-1920) believed that the rise of rationality, the process of replacing traditional or emotionally driven thought with reason and practicality, was the driving force in society. He posited the concept of *Verstehen*, or understanding why people do what they do in their interactions, as a primary lens through which to view human interaction (Bruce, 1999). Emile Durkheim (1858-1917), a French sociologist considered the

"father of sociology," believed that "culture does for human beings what instinctual and environmental constraints do for other species" (Bruce, 1999, p. 35). His contention was that the breakdown of shared social norms, resulting from the evolution of a more complex society, was the issue bringing the most detriment to bear on society.

There are four guiding tenets of sociology. The first is the belief that "reality is socially constructed" (Bruce, 1999, p. 25). This means that there is not one single reality. Rather each person's reality is constructed based on individual experience. The second tenet is that human behavior has hidden social causes. This means that despite having the freedom to think and act independently of formal constraints, human behavior is driven by social norms and expectations. People have a biological need to belong to the larger social group, and tailor their behavior accordingly. Third is the notion that "much of social life is inadvertent or unintended" (Bruce, 199, p. 86). This tenet is important as sociologists seek to understand the causes of a given outcome, because an outcome may not in fact be an intended but rather "an unintended consequence of a large number" (Bruce, 1999, p. 87) of other factors. The fourth and final tenet of sociology is that the way "people see themselves is greatly affected by how others see them" (Bruce, 1999, p. 48).

The work of sociologists is focused on various broad facets of social life including class structure and institutions as well as the components within those institutions. The discipline is focused in large part on trying to understand "how humans come to live their realities" (Bruce, 1999, p. 48). Sociologists are interested in the "social causes of health, wealth, and happiness as well as poverty, illness, and depression" (Bruce, 1999, p. 56). They search for "patterns and use systematic comparisons to illuminate the cause of who we are and what we do" (Bruce, 1999, p. 54), attempting to not only identify the "what" of human behavior or interactions, but also to examine beliefs, values, motives, and intentions underlying the "why" of human actions.

Functionalism and Education

A discussion of educational issues from a sociological standpoint fits particularly well into the functionalist framework. Ontologically, functionalism has its roots in the positivistic organisms of the 19th century, as developed by Emile Durkheim (1858-1917). Functionalists look at the world as a system, and try to understand the contributions of its components.

As discussed by Martindale (1965), the industrialization of nations in the 19th century led to higher population densities, a movement from individual thought toward collectivist thought, and the development of "the institutions of the mass state" (p. 153). Sociological functionalism was primarily an invention of the American mass state in the second half of the 20th century. It could be seen as a response to the WWII-era success of functional military and economic systems, and an attempt to apply these models to human behavior. As sociological functionalism was further developed by Talcott Parsons (1902-1979) and Robert Merton (1910-2003), they established an epistemology focused more on teleology than causality. In other words, as Martindale (1965) explained, Parsons and Merton were

more interested in the *purposes* served by social phenomena, and less in what may have caused them. This is truly the root of a functionalist mindset—the *what* is more important than the *why*.

In its broadest sense, functionalism describes the social system (and its component institutions, like education) as "a set of social activities operating in equilibrium," and posits that "all recurrent social activities have the function of maintaining a social system" (Whitaker, 1965, p. 154). From an axiological standpoint, then, functionalists count any activity that helps maintain societal equilibrium as good, and anything that causes dysfunction as bad.

According to Whitaker (1965), there are many advantages to studying problems using this functionalist methodology. He discussed how functionalism excludes speculation on human motives and the origins of institutions, and brings us closer to a "natural science of society" (p. 143). Functionalism has its limitations as well. It can be seen as a very conservative mindset, only interested in maintaining the status quo, and consequently ignorant of how dysfunctional certain aspects of society can be. According to Whitaker (1965), functionalism also struggles to account for social change, and is seen as a "static closed system," only dealing with social phenomena "at one moment in [time]" (p. 140). These advantages and disadvantages are certainly valid in some situations, and with some thinkers and theorists. However, thinkers like Merton have found ways to modify functionalism to account for social change, and opened it up to discussions of dysfunction (Whitaker, 1965). In fact, that is just what is required as we turn to look at our education system, and more specifically the function of teacher evaluation within that system.

At its core, the function of the institution of education is to produce citizens able to contribute to the workforce and provide value to society, thereby maintaining societal equilibrium. Many factors impact learning at all levels, including school environment, curriculum quality, an individual student's background/culture, and most importantly for our purposes here, teacher quality. This paper examines the evaluation of teacher effectiveness in Virginia, and considers how the evaluation process could be better operationalized to affect student learning by way of increased teacher capacity. This increased student learning outcome would demonstrate the function of the evaluation process in service to society at large by providing the next generation of capable citizens.

Traditional vs. Authentic Assessment Models

Teacher assessment methods fall broadly into two categories, traditional and authentic. Traditional methods have roots in the positivist paradigm, which is based on the belief in an objective reality, attempting to rate teacher performance using quantitative measures, and judging teachers on their ability to perform key tasks, according to a normative standard that is applied across the board. Traditional methods often focus more on student performance on standardized tests, utilize single raters, and rate teachers using what Bambrick-Santoyo (2012) calls the "scoreboard" model (para. 8). This methodology is often driven by pressure to increase test scores,

stemming from legislation like the NCLB of 2001 (Valli et al., 2007). Authentic methods, on the other hand, are more in the functionalist paradigm, which regards society as a complex system whose parts work together to maintain social stability. It recognizes differences in students and schools, and attempts to account and control for all influences on student learning. In the words of Tanner (2001), authentic assessment "expects environmental conditions to vary," and seeks out "multiple ways for students to demonstrate their learning" (p. 27).

Strengths of Traditional Assessment Method

Using standardized test results, a traditional method to evaluate teachers, has both strengths and weaknesses. Proponents argue these assessments are norm referenced, valid, and reliable (Tanner, 2001). Also, standardized tests are posited to be an objective measure of student learning (Corcoran, 2010) and, assuming the integrity of the student growth percentile, afford a reliable indicator of student proficiency with grade-level material. A further argument supporting use of traditional assessments comes from a 2013 Associated Press-NORC Center for Public Affairs Research poll, which found that "75% of parents say standardized tests are a solid measure of their children's abilities. 69% say the tests are a good measure of a school's quality" (Elliott & Agiesta, 2013 as cited in ProCon.org, 2015, April 3, para. 9). Proponents seem to be arguing that if teachers can affect a baseline competency in reading and math, then they have performed their function within the education system.

Weaknesses of Traditional Assessment Method

If we stipulate for the moment that this is the proper function of a teacher, problems still exist. For instance, only about 50% of educators teach in grade levels requiring standardized assessments, leaving a large percentage of educators who do not have those assessments on which their evaluations can be based (Toch, 2008). Even though there are allowances in the evaluation tool for using other validated, quantitative measures, these assessments are often teacher or district made. As a result, there is neither uniformity in those assessments nor a reasonable way to ensure they are a valid measure of student learning, or if they will be implemented with fidelity. Additionally, opponents contend there is too much emphasis placed on the test score (Tanner, 2001), offering only a single snapshot of student learning (Toch, 2008), and not a comprehensive idea of teacher effectiveness.

A further argument against the use of standardized tests is that their use narrows the curriculum, as it is not possible to test all things a student should know (David, 2011). Teachers then attempt to raise test scores by teaching to the test. Teaching to the test not only invalidates the test itself, but takes time away from the teaching of those concepts/skills which are not tested. Also of particular concern is that these standardized tests, given once or twice a year, provide "few if any feedback opportunities for students or teachers to improve" (Mielke & Frontier, 2012, p.10).

Strengths of Authentic Assessment Method

Measuring the quality of teaching and quantifying the impact a teacher has on individual student learning can be challenging. This is especially true given that, as Valli, Croninger, and Walters (2007) state, "disentangling teacher effects from school effects is more complex than generally acknowledged" (p. 637). Factors like student mobility, team teaching, and other complex links with teachers have noticeable influence on teacher effect (Valli et al., 2007). Popham (1999) agrees that it is very "difficult to isolate the impact of a single teacher from the other factors that influence test scores. These other factors include what was taught in school in previous years or by other teachers, a child's innate intellectual ability, and the student's out of school learning" (p. 4). Finally, in contrast to data indicating parental support for the use of standardized tests cited above, other reports indicate the emergence of an "anti-testing sentiment as the number of tests given to students has risen, resulting in parents opting their students out of testing" (Sawchuk, 2015, p. 4).

Authentic assessment, conversely, may expand the teacher evaluation rubric to include non-academic measures, empowering teachers through self-diagnosis and self-evaluation, and providing multiple raters for a single teacher. Additionally, authentic assessment purports to coach for growth, not scores (Bambrick-Santoyo, 2012), and focuses on developing teachers and their expertise (Mielke & Frontier, 2012). Proponents of incorporating authentic assessments of student progress into evaluations of teachers cite the shift in learning to a more constructivist process wherein "teaching is done for understanding, not for rote learning" (Danielson & McGreal, 2000, p. 3). Supporters also laud authentic assessments because they inherently focus more on "learning than on passing a test, and on the transfer of learning as measured by application of skills on authentic tasks" (Tanner, 2001, p 24). If authentic assessments are more reflective of students' learning and ability to apply that learning, it stands to reason that these assessments may be better suited as formative evaluations, which is a more functional use of the evaluation tool than their traditional counterpart.

Reflecting on student performance of authentic tasks helps teachers focus on improving the quality of their instruction. According to Tanner (2001), any kind of assessment is defensible only to the extent that it "actively forwards and enhances a child's learning" (p. 25). Authentic assessments also tend to align more closely with diverse curricula, and afford students the opportunity to respond in a variety of ways. This flexibility is essential, given the diversity of learners in schools. Further, authentic assessments are generated locally by teachers or districts, not by third parties or testing companies. As a result, they are often a better reflection of what has been taught in a particular school. These assessments are generally scored manually, so even with a rubric in place, there is an opportunity for a teacher to exercise professional judgment in evaluating student responses. Results of authentic assessments can also provide valuable, timely feedback for teachers reflecting to improve the effectiveness of their practice, while

simultaneously allowing teachers to inform their instruction to meet students' needs.

Weaknesses of Authentic Assessment Method

Opponents of incorporating authentic assessments of student progress in teacher evaluations claim that they are not a reliable indicator of student achievement because they lack a normative basis, validity, and reliability. Additionally, grading of these assessments is subjective, and assessments themselves may not be given with fidelity, thereby skewing any results. Finally, opponents argue that using authentic assessment is both "time and cost intensive" (Tanner, 2001, p. 28). For each content area by grade level, divisions have to define essential learning competencies, and develop assessments and rubrics to evaluate learning. Teachers then require training on how to give and evaluate the assessment, and how to analyze the results to inform their instruction. In addition, this type of work requires consensus across districts/schools, which may be difficult to reach.

In the specific case of teacher evaluation in the Virginia K-12 public schools, the authors question whether the Virginia evaluation model strikes enough of a balance between "traditional" and "authentic" measures to accurately assess whether or not teachers are performing the sociologically required function of producing educated, competent citizens. In Tanner's (2001) words, can we "connect the classroom to life beyond the school, and advance the quality of teaching in the process" (p. 24)?

Overview of Virginia K-12 Teacher Evaluation Model

Previous traditional teacher evaluation models failed to accurately measure teacher quality and effectively support teacher professional development (Marzano, 2012). Based on the research evidence (e.g. Stronge, 2006), VDOE has pointed out these problems, including failing to differentiate teachers' performance, low validity of evaluation instruments, lack of impact, and absence of meaningful and timely feedback to teachers (VDOE, 2015). RTTT and state legislation have called for "rigorous, transparent, and fair evaluation models that differentiate teacher effectiveness based on student achievement" (Marzano, 2011, p. 3).

In 2015, Virginia's Board of Education approved the revised Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers (Guidelines) (VDOE, 2015). Along with that, the Standards for the Professional Practice of Teachers (VDOE, 2011) also provides a conceptual model for effective teaching, and is used as a resource for teacher evaluation. The Code of Virginia requires that teacher evaluations be consistent with the performance standards included in the revised Guidelines (VDOE, 2015). The Guidelines document is designed to "support the continuous growth and development of each teacher by monitoring, analyzing, and applying pertinent data compiled within a system of meaningful feedback" (VDOE, 2015, p. 7).

The resulting Virginia teacher evaluation model was developed as a bifurcated tool. The first part of the

evaluation, standards one through six, "include[s] professional knowledge, instructional planning, instructional delivery, assessment of and for student learning, learning environment, and professionalism." (VDOE, 2015, p. 3-4). These six standards serve as a "trait instrument designed to rate an employee's overall performance against a pre-determined set of indicators" (Rebore, 2015, p. 211). The second part of the evaluation tool, standard seven, is required by Virginia law and seeks to connect teacher performance to measures of student academic progress. It serves as a "results-based form of evaluation, but instead of focusing on teachers comparatively rating themselves against their own growth over time in self-selected areas of need" (Rebore, 2015, p. 212), the standard uses student performance measures to quantify teacher effectiveness.

The Virginia evaluation model uses multiple data sources to document teacher performance, which include both formal and informal observations, student surveys, document logs, and teachers' self-evaluations (VDOE, 2015). The rating of the summative evaluation is determined by weighing the first six standards equally at 10% each, and giving the seventh standard (based on student performance) a weight of 40% (Marzano, 2011; VDOE, 2015). Within the 40%, about 20% of the student academic progress measure is comprised of student growth percentiles provided by VDOE, and the other 20% is recommended from one or more validated and quantitative measures (VDOE, 2015). A student growth percentile quantifies "how much progress a student has made relative to the progress of students whose achievement was similar on previous assessments" (VDOE, n.d., para. 1). The overall evaluation of a teacher relies on a scale that has four levels: exemplary, proficient, needs improvement, and unacceptable (VDOE, 2015). The first two levels are regarded as "effective performance" and the last two are "not meeting expectations" (VDOE, 2015, p. 58).

Functionalist Analysis of Virginia K-12 Teacher Evaluation Model

Abundant research evidence indicates that quite a proportion of student learning gains can be statistically explained by teacher effects (e.g., Muñoz & Chang, 2007; Rivkin et al., 2005; Stronge & Tucker, 2005), and that the comparative impact on student achievement for effective teachers versus ineffective teachers is significantly different (e.g., Stronge, 2010; Stronge, Ward, & Grant, 2013). For example, Stronge et al. (2013) found more than 30 percentile points of difference in fifth-grade students' achievement in reading and mathematics based on one year's teaching and learning experience can be attributed to the quality of teaching performed by top-quartile and bottom-quartile teachers. Therefore, including student academic progress in the teacher evaluation process is indispensable.

Evaluations of student progress, if used correctly, could go a long way toward improving educational outcomes for students by way of increased teacher effectiveness, which is the ultimate goal of the accountability movement. From the functionalist perspective, the primary responsibility of school is to produce students who can eventually contribute to the greater good of society. Given the understanding that teaching and learning

comprise the core work of schools, and that teachers "vary greatly in their ability to promote student achievement" (Taylor & Tyler, 2012, p. 3638), we come back to the question of whether the teacher evaluation model in Virginia serves its purpose when viewed through a functionalist lens.

By using student growth percentiles from standardized tests as a 40% measure of teacher effectiveness, the Virginia evaluation tool appears to serve the accountability movement well by focusing on the summative function of evaluation. While incorporating traditional measures of student achievement in Virginia's teacher evaluation tool appears to serve the summative function, it is questionable if these traditional measures, especially weighted as they are, complement the formative function of improving teacher quality/effectiveness. This is a critical consideration given that the formative function of teacher evaluation should be to improve teacher practice in the service of student achievement. In response to this perceived deficit, some are advocating in favor of incorporating more authentic measures of student progress into the teacher evaluation process.

The Virginia evaluation model shown in the revised Guidelines is certainly neither traditional nor all authentic. However, from a functionalist perspective there are some limitations to this model given its reliance on traditional/summative assessment, as outlined previously in this article. First, there is disproportionate weight given to student academic progress as a measure of teacher effectiveness. The problems with attributing student progress (or lack thereof) to particular teachers are well documented. There is also a troubling emphasis placed on student performance in math and reading; the only supporting evidence for linking student progress to teacher effectiveness in the Guidelines document reports teacher variance based only on scores from these two subjects (VDOE, 2015).

Secondly, VDOE teacher ratings highlight feedback problems, in both frequency and inattention to teacher development. The model allows for yearly summative evaluation of probationary teachers, and a three-year cycle of evaluation for continuing contract teachers. By the time a teacher receives feedback, it will be too late to implement any kind of meaningful change in that year. By the next year, teachers will have a new group of students, and the opportunity will have passed for teachers to instructionally impact those students for whom they now have feedback. From a functionalist perspective, this missed opportunity would not be a productive use of the evaluation tool, because the results of the evaluation would not inform teachers' instruction, which is the stated function of the evaluation tool. A final concern with the Virginia model if applying the functionalist lens is that the language used to rate teachers speaks to a disconnect in the emphasis on growth for all teachers. From the revised Guidelines: "the use of the scale enables evaluators to acknowledge effective performance (i.e., "exemplary" and "proficient") and provides two levels of feedback for teachers not meeting expectations (i.e., "developing/needs improvement" and "unacceptable")" (VDOE, 2015, p. 59). The current system

seems skewed towards identifying unsatisfactory performance, rather than functioning as a springboard for reflection. The current evaluation tool does not convey the expectation that all teachers should be working toward continuous improvement in their practice in fulfilling their functionalist role of preparing citizens who can contribute to society.

Discussion

There seems to be little question that traditional and authentic assessment methods both have their place in the evaluation of teachers. As Danielson and McGreal (2000) stated, "teaching is a highly complex process that defies traditional methodology for evaluating teachers" (p. 16). In order to support the function of education, any effective teacher evaluation system must recognize the difference between the two methods, and use each appropriately (Marzano, 2012).

Likewise, there is little argument that if teachers are more competent, students will learn more effectively, and grow throughout their academic careers. As Mielke and Frontier (2012) discussed, developing expertise in teachers should be the goal of a quality evaluation system, and that evaluation system should empower teachers through self-diagnosis, skill development, and peer interaction.

Tanner (2001) stated simply, "if we aim to improve evaluation, the greatest promise is to integrate both traditional and authentic assessment components as they complement one another" (p. 29). However, the two modes of assessment will not complement each other in the same way in every school. This is where the authors think a flat 40% assessment of improved test scores makes it difficult for teachers to function effectively. As mentioned before, teachers are often caught in a trap where they must teach to certain tests, while recognizing that the tests are narrow, and do not reflect the breadth of a student's learning or aptitude for learning.

The VDOE evaluation tool does include both traditional and authentic assessment methods, but is still ignoring key components of an evaluation system that in its most effective form, as described by Danielson and McGreal (2000), will be "designed to support teacher growth through a formative process [that produces] higher levels of satisfaction and more thoughtful, reflective practice while still satisfying accountability demands" (p. 15).

Conclusion

After analyzing the dual accountability functions of the VDOE teacher evaluation model, the authors found it fulfills the summative function by helping administrators make summative decisions, but it has limitations in the formative function of helping teachers increase the quality of their instruction in the service of student achievement. Also, a 40% assessment based on student achievement leads teachers to design test-driven curricula neglecting students' well-rounded development. Within the context of sociology, teacher evaluation is socially constructed and it is driven by social norms and expectations. Every society expects to develop competent citizens through the effective education system. However, the limitation in the functions of teacher evaluation would negatively affect teaching and learning in schools, which is the core of educational

systems. Analyzing the functions of teacher evaluation models through the lens of functionalism helped answer the question regarding how well the VDOE model serves the dual accountability functions. An analysis of the pros and cons of the teacher evaluation model's dual functions leads the authors to believe the VDOE should revise the model to focus more on the formative function, allowing the evaluation tool to serve its intended function of improving teachers' instructional capabilities in the service of students and society.

References

Bambrick-Santoyo, P. (2012). Beyond the scoreboard. *Educational Leadership, 70*(3), 26-30.

Bruce, S. (1999). *Sociology: A very short introduction.* Oxford, UK: Oxford University Press.

Corcoran, S. P. (2010). *Can teachers be evaluated by their students' test scores? Should they be? The use of value-added measures of teacher effectiveness in policy and practice.* Retrieved from: http://www .annenberginstitute.org/pdf/valueAddedReport.pdf.

Danielson, C., & McGreal, T. (2000). *Teacher evaluation to enhance professional practice.* Alexandria, VA: ASCD.

David, J. L. (2011). What students need to learn: Research says…high-stakes testing narrows the curriculum. *Educational Leadership, 68*(6), 78-80.

Martindale, D. (1965). *Functionalism in the social sciences: The strength and limits of functionalism in anthropology, economics, political science, and sociology.* Philadelphia, PA: American Academy of Political and Social Science.

Marzano, R. J. (2011). *Evaluation model alignment to Virginia teaching standards.* Retrieved from http://www.marzanoevaluation.com/files/Marzano_Alignment_to_Virginia_Teaching_Standards.pdf

Marzano, R. J. (2012). The two purposes of teacher evaluation. *Educational Leadership, 70*(3), 14-19.

Mielke, P., & Frontier, T. (2012). Keeping improvement in mind. *Educational Leadership, 70*(3), 10-13.

Muñoz, M. A., & Chang, F. C. (2007). The elusive relationship between teacher characteristics and student academic growth: A longitudinal multilevel model for change. J*ournal of Personnel Evaluation in Education, 20*(3-4), 147-164.

Popham, J. (1999). Why standardized tests don't measure educational quality. *Educational Leadership, 56*(6), 8-15.

ProCon.org. (2015, April 3). Is the use of standardized tests improving education in America? Retrieved from http://standardizedtests.procon.org/#background

Rebore, R. W. (2015). *Human resources administration in education: A management approach* (9th ed.). Boston, MA: Allyn & Bacon.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417-458.

Sawchuk, S. (2015, September). Teacher evaluation: An issue overview. *Education Week, 35*(3).

Stronge, J. H. (2010). *Effective teachers=Student achievement: What the research says.* Larchmont, NY: Eye on Education.

Stronge, J.H. & Tucker, P.D. (2005). *Linking teacher evaluation and student learning.* Alexandria, VA: ASCD.

Stronge, J. H., Ward, T. J., & Grant, L. W. (2011). What makes good teachers good? A cross-case analysis of the connection between teacher effectiveness and student achievement. *Journal of Teacher Education, 62*, 339-355.

Tanner, D. (2001). Authentic assessment: A solution, or part of the problem? *The High School Journal, 85*(1), 24-29.

Taylor, E. S., & Tyler, J. H. (2012). The effect of evaluation on teacher performance. *The American Economic Review, 102*, 3628-3651. doi:10.1257/aer.102.7.3628

Toch, T. (2008). Fixing teacher evaluation. *Educational Leadership, 66*(2), 32-37.

U.S. Department of Education. (2009, November). *Race to the Top program executive summary.* Retrieved from http://www.ed.gov/programs/racetothetop/executive-summary.pdf

Valli, L., Croninger, R. G., & Walters, K. (2007). Who (else) is the teacher? Cautionary notes on teacher accountability systems. *American Journal of Education, 113*(4), 635-662. doi: 10.1086/518492

Virginia Department of Education. (2011). *Standards for the Professional Practice of Teachers.* Retrieved from http://www.doe.virginia.gov/teaching/regulations/uniform_performance_stds_2011.pdf

Virginia Department of Education. (2015). *Guidelines for uniform performance standards and evaluation criteria for teachers.* Retrieved from http://www.doe.virginia.gov/teaching/performance_evaluation/guidelines_ups_eval_criteria_teachers.pdf

Virginia Department of Education. (n.d.). Student growth percentiles. Retrieved from http://www.doe.virginia.gov/testing/scoring/student_growth_percentiles/index.shtml

Whitaker, I. (1965). The nature and value of functionalism in sociology. In D. Martindale (Ed.), *Functionalism in the social sciences: the strength and limits of functionalism in anthropology, economics, political science, and sociology* (pp. 127-143). Philadelphia, PA: American Academy of Political and Social Science.

About the Authors

Brendan Bourdage is a doctoral student in the Educational Policy, Planning & Leadership Higher Education program at The College of William & Mary. His research interests include higher education access for underrepresented/disadvantaged students, and the balance of athletics and academics in higher education.

Kelly Erickson is a doctoral student in the Curriculum Leadership program at the College of William and Mary. She is also pursuing a leadership endorsement. Her goal is to be a practitioner able to apply research to resolve problems of practice in reading instruction.

Yi Hua is a doctoral candidate in the Educational Policy, Planning & Leadership K-12 program at The College of William & Mary. Her research interests include teacher evaluation, education assessment and parental involvement.