

LIES, DAMN LIES, AND TEST SCORES:  
NATIONWIDE MISTAKES AND FEASIBLE SOLUTIONS IN TEACHER EVALUATION

JIM CUMMINGS<sup>1</sup>

**Introduction**

As the battle to save education in America rages on, a new victim – or villain, depending on one’s opinion – has risen to the forefront. The gauntlet has been thrown upon America’s public school teachers. Teachers are being held accountable for their pensions, for their tenure, and for how some of them have left our nation’s children behind. Evaluation, however, is not an inquisition – for all the under-performing teachers in America, there are as many and more who are unrecognized and underdeveloped. Part of the problem of education in America is that the potential in many new teachers is not realized, and many good teachers are not given the guidance to be great. Like most any profession, there is nothing inherently wrong in evaluating teachers. Instead, America’s problem in evaluating its teachers is that, by and large, it does not know what it is doing.

This paper will discuss how the current educational and economic climate in America has allowed standardized testing to emerge as possibly the most significant factor in teacher evaluation, even though, if taken alone, test scores are an ineffective means of assessing teacher performance. Fortunately, feasible solutions appear on the horizon. The Bill and Melinda Gates Foundation has developed the Measures of Effective Teaching (MET) project to improve both the form and substance of teacher evaluation. Using combinations of varied and authentic assessments, MET has the potential to provide tangible data that will allow schools to measure teacher performance so America can close the achievement gap.

---

<sup>1</sup> Loyola University Chicago, J.D., May 2012. Prior to law school, the author spent two years teaching Middle School Language Arts, Religion, and Art at St. Mary’s Catholic Elementary School outside Birmingham, Alabama, and then six years teaching English at Marist High School in Chicago, Illinois. He admires and is indebted to all the teachers who have taught and worked alongside him, but most thanks go to his wife.

## I. Every Teacher Gets an A, Unless He Gets Fired

“Although there is growing consensus that effective teaching is the key to large-scale school reform, there is no agreement among education stakeholders about how to identify and measure effective teaching.”<sup>2</sup> The problem is not that a particular school or school district has the choice of using one of several informative assessments to evaluate its teachers, and schools are at a loss when these various methods conflict. Rather, no traditional, established method exists. As a whole, the effectiveness of America’s teachers – the most important factor for schools in improving student achievement – is not measured, recorded, or used to inform decision-making in any meaningful way.<sup>3</sup>

In 2009, the New Teacher Project explored this problem in a report called *The Widget Effect*.<sup>4</sup> This report surveyed teacher evaluation methods in twelve districts across four states (Arkansas, Colorado, Illinois, and Ohio).<sup>5</sup> Approximately 15,000 teachers, 1,300 administrators, and 80 local and state educational officials participated.<sup>6</sup> The districts ranged in size, geographic location, and evaluation policies, yet all employed some formal evaluation process for teachers.<sup>7</sup> The smallest district was Jonesboro Public Schools, serving approximately 4,450 students, and the largest was Chicago Public Schools, serving over 413,700 students.<sup>8</sup>

Despite the spectrum, the outcomes from the districts were similar, netting five conclusions: nearly all teachers are rated good or great, excellent teaching goes unrecognized, schools fail to provide useful professional development, new teachers are left to their own

---

<sup>2</sup> BILL AND MELINDA GATES FOUND., WORKING WITH TEACHERS TO DEVELOP FAIR AND RELIABLE MEASURES OF EFFECTIVE TEACHING 2 (2010), available at <http://www.metproject.org/downloads/met-framing-paper.pdf>.

<sup>3</sup> DANIEL WEISBERG ET AL., THE WIDGET EFFECT 3 (The New Teacher Project) (2009), available at <http://widgeteffect.org/downloads/TheWidgetEffect.pdf>.

<sup>4</sup> *Id.*

<sup>5</sup> *Id.* at 5.

<sup>6</sup> *Id.*

<sup>7</sup> *Id.*

<sup>8</sup> *Id.*

devices, and poor teaching goes unaddressed.<sup>9</sup> Exacerbating the problem is the practice and implementation of evaluations.<sup>10</sup> Most administrators do not have or give the time to conduct more than two class-long observations a year, and even then, the administrators usually lack the training to make the observation meaningful.<sup>11</sup> The results are that excellent teachers are not rewarded, chronically low-performing teachers languish or are unaware of their shortcomings until they are dismissed, and most teachers performing at moderate levels do not receive the individualized support and development they need to improve as professionals.<sup>12</sup>

This problem has recently come to a head in Wisconsin and other states when confronted by the issue of tenure. In Illinois, while a tenured teacher can be dismissed for incompetency, a decision to terminate must be supported by substantial evidence of failure to perform to the minimal standards articulated by the district.<sup>13</sup> When evaluation is infrequent and insubstantial, schools have difficulty removing low-performing teachers.

## **II. Standardized Tests: King of Education's New World Order**

Standardized testing has many purposes and flaws as an educational tool for students, none of which will be discussed here. Instead, this section will discuss how test scores have risen to prominence as a means to assess teachers and schools. This has occurred primarily because of two factors: developments in national education policy and the economy. These two factors have transformed the public school landscape so drastically that school districts and state legislatures are scrambling to change their education policies, allowing standardized testing to attain unprecedented sway over the educational experiences of students, teachers, administrators, schools, and states.

---

<sup>9</sup> *Id.* at 5-6.

<sup>10</sup> *Id.* at 6.

<sup>11</sup> *Id.*

<sup>12</sup> *Id.*

<sup>13</sup> Paul C. Burton, *Illinois Tenure: Intent, Standards, and Reality*, 19 ILL. ST. U. SCH. L.Q. 103, 108 (1999).

### **A. Race so Your School Is not Left Behind**

Without an accepted model to assess teachers, test scores have filled the void. By themselves a means to measure student achievement on a wide scale, standardized test scores are also used to assess schools themselves. The No Child Left Behind Act of 2001 (NCLB) requires all state-run schools receiving federal funding to administer standardized tests annually to all students, and the results of the test determine whether the school is doing well or poorly.<sup>14</sup>

In 2009, President Obama approved the Race to the Top Fund (RTTT), “a competitive grant program designed to encourage and reward States that are creating the conditions for education innovation and reform.”<sup>15</sup> Like NCLB, RTTT’s “education innovation and reform” is, at least in part, measured by standardized test scores, which are used to assess not only the schools, but the teachers and administrators as well.<sup>16</sup> To be eligible to receive federal funds through Phase 1 or 2 of this grant program, a state had to meet the following requirements:

- (a) The State’s application for funding under Phase 1 and Phase 2 of the State Fiscal Stabilization Fund Program must be approved by the Department [of Education] prior to the state being awarded a Race to the Top grant.
- (b) At the time the State submits its application, there must not be any legal, statutory, or regulatory barriers at the State level to linking data on student achievement (as defined in this notice) or student growth (as defined in this notice) to teachers and principals for the purpose of teacher and principal evaluation.<sup>17</sup>

To fall in line with requirement (b), school districts have employed value-added modeling (VAM) to create a meaningful link between student test scores and teacher evaluations.

---

<sup>14</sup> See U.S. DEP’T. OF EDUC., STATE STANDARDS AND ASSESSMENT UPDATE (2009), *available at* <http://www2.ed.gov/admins/lead/account/statesystems.html>.

<sup>15</sup> U.S. DEP’T. OF EDUC., RACE TO THE TOP PROGRAM: EXECUTIVE SUMMARY 2 (2009), *available at* <http://www2.ed.gov/programs/racetothetop/executive-summary.pdf>.

<sup>16</sup> *Id.* at 2-3.

<sup>17</sup> *Id.* at 4

VAM is a collection of complex statistical techniques that uses multiple years of students' test score data to estimate the effects of individual schools or teachers.<sup>18</sup> VAM's selling point is that it does not link test scores to teachers in a vacuum. Rather than simply judging teachers by the scores of their students in and of themselves, VAM considers the students' level of achievement as of the day they first sit in a particular teacher's class. VAM then "calculates the value teachers add to their students' achievement, based on the changes in test scores from year to year and how the students perform compared to others in their grade."<sup>19</sup>

VAM claims to be helpful in two ways. First, VAM purports that it can isolate the effects teachers and schools have upon student achievement from outside factors such as family background.<sup>20</sup> Second, VAM claims that it can identify measurable differences in the effectiveness of teachers.<sup>21</sup> Two separate possibilities emerge from this. One, these measurable differences can be compiled, allowing a school to rank its teachers, and thus objectively identify high-, moderate-, and low-performing teachers.<sup>22</sup> From this, by recognizing effective teachers, a school could then isolate specific practices of those teachers that were effective, thereby allowing those practices to be refined and proliferated.<sup>23</sup> If VAM's promises come to fruition, its potential for improvement of education could be great.<sup>24</sup>

## **B. It's the Economy, Sir**

RTTT was signed into law as part of the American Recovery and Reinvestment Act (ARRA) of 2009. At the time, while the recession had ended, America's unemployment rate was

---

<sup>18</sup> DANIEL F. MCCAFFREY ET AL., EVALUATING VALUE-ADDED MODELS FOR TEACHER ACCOUNTABILITY xi (RAND Education) (2003), available at [http://www.rand.org/pubs/monographs/2004/RAND\\_MG158.pdf](http://www.rand.org/pubs/monographs/2004/RAND_MG158.pdf).

<sup>19</sup> Sam Dillon, *Formula to Grade Teachers' Skill Gains Acceptance, and Critics*, N.Y. TIMES, Aug. 31, 2010, available at <http://www.nytimes.com/2010/09/01/education/01teacher.html?pagewanted=1>.

<sup>20</sup> MCCAFFREY ET AL., *supra* note 16, at xi.

<sup>21</sup> *Id.*

<sup>22</sup> Dillon, *supra* note 19.

<sup>23</sup> MCCAFFREY ET AL., *supra* note 18, at xi.

<sup>24</sup> *Id.*

9.9 percent and many states had deep budget shortcomings.<sup>25</sup> As part of its plan to stimulate the economy, ARRA provided \$4.35 billion for RTTT to hand out to the states which achieved the most in education innovation and reform – \$4 billion for statewide reform grants and \$350 to support states working together to improve the quality of their assessments.<sup>26</sup>

Phase 1 of RTTT was completed in March 2010.<sup>27</sup> Forty states and the District of Columbia submitted applications for grants. Sixteen finalists were selected to present proposals to panels of peer reviews, and Delaware and Tennessee won grants.<sup>28</sup> Delaware received approximately \$100 million and Tennessee received \$500 million to implement their comprehensive school reform plans over the next four years (the differences in funds accounting for population).<sup>29</sup> Applications for Phase 2 of RTTT were due on June 1, 2010.<sup>30</sup>

### **C. Call with Hundreds of Millions of Dollars, and the States Answer**

Witnessing the windfalls in Delaware and Tennessee, legislatures in several states passed bills that revamped their education polices to appear more in line with the requirements of RTTT.<sup>31</sup> Is it of no coincidence that these bills were passed in May 2010, weeks before RTTT's deadline for Phase 2. This section discusses changes in legislative policy in response to RTTT made by Colorado, New York, and New Jersey.

In May 2010, Colorado Governor Bill Ritter signed Senate Bill 191 that would tie teacher assessment closely with student achievement.<sup>32</sup> The bill focuses on teacher accountability as

---

<sup>25</sup> Bureau of Labor Statistics, "Economy at a Glance," <http://www.bls.gov/eag/eag.us.htm> (last visited May 5, 2011).

<sup>26</sup> U.S. DEP'T OF EDUC., *supra* note 13, at 2.

<sup>27</sup> Press Release, U.S. Dep't. of Educ., Delaware and Tennessee Win First Race to the Top Grants (Mar. 29, 2010), available at <http://www2.ed.gov/news/pressreleases/2010/03/03292010.html>.

<sup>28</sup> *Id.*

<sup>29</sup> *Id.*

<sup>30</sup> *Id.*

<sup>31</sup> Jessica Calefatti, *Should Teachers' Raises Depend on Kids' Test Scores?*, MOTHER JONES, Jun. 9, 2010, available at <http://motherjones.com/politics/2010/05/should-teachers-raises-depend-kids-test-scores>.

<sup>32</sup> Joseph Boven, *Ritter Signs Teacher Assessment Legislation into Law*, Colorado Independent, May 21, 2010, <http://coloradoindependent.com/53811/ritter-signs-teacher-assessment-legislation-into-law>.

“[t]eachers who fail to meet standards will be placed on probationary status after two years and could be dismissed after three years failing to achieve successful scores.”<sup>33</sup> The bill calls for using VAM to account for as much as 50 percent of a teacher’s evaluation.<sup>34</sup> In response to the Colorado Educators Association’s opposition to the bill, one of its authors stated that “[t]his bill is not about getting money for Race to the Top. It is about our children being in a race for the future of our country.”<sup>35</sup> Upon signing the bill, Governor Ritter echoed the bill’s focus upon children, however, he also pointed to an executive order he signed in January 2010 which created the Council on Educator Effectiveness, which was part of the effort to win RTTT funding in 2010.<sup>36</sup> In 2012, the Colorado legislature will vote on whether or not to continue the program set forth by the bill.<sup>37</sup>

Also in May 2010, New York’s State Education Department and teachers’ unions reached an agreement to link teacher evaluations to student test scores.<sup>38</sup> Under the previous system, teachers were simply rated satisfactory or unsatisfactory, and test scores could not play a role in teacher evaluations.<sup>39</sup> Under the new system, a 100-point scale is used to place teachers in one of four categories – highly effective, effective, developing, and ineffective.<sup>40</sup> Also, the scale devotes 40 percent of the points to student improvement on standardized exams.<sup>41</sup> Currently, half of the 40 percent is based on state exams, with the other half based on local tests, but in 2012, that will change, and 25 percent of a teacher’s evaluation will be a result of how her

---

<sup>33</sup> *Id.*

<sup>34</sup> Calefatti, *supra* note 31.

<sup>35</sup> Boven, *supra* note 32.

<sup>36</sup> *Id.*

<sup>37</sup> *Id.*

<sup>38</sup> Jennifer Medina, *Agreement Will Alter Teacher Evaluations*, N.Y. TIMES, May 10, 2010, available at <http://www.nytimes.com/2010/05/11/nyregion/11teacher.html>.

<sup>39</sup> *Id.*

<sup>40</sup> *Id.*

<sup>41</sup> *Id.*

students improve on state exams, and 15 percent resulting from the local tests.<sup>42</sup> The negotiations leading to the agreement were tense as the unions hesitated to accept the changes, while the Education Department felt pressure to meet the RTTT deadline of June 1, 2010.<sup>43</sup>

A similar May 2010 bill was passed in New Jersey, though it differed in some ways from the ones in Colorado and New York.<sup>44</sup> First, the state government was open to the fact that the plan to overhaul the state's education system was an appeal to RTTT.<sup>45</sup> Governor Chris Christie said adopting the proposed changes were "very necessary" to get the federal funding for New Jersey, which has a server budget deficit.<sup>46</sup> Second, the reforms called for a new statewide data system to track the achievement of each student in the state during every quarter of the school year.<sup>47</sup> Third, the bill required that student achievement accounted for at least 51 percent of their teachers' evaluation.<sup>48</sup> Lastly, the bill expanded the use of teacher evaluation – the assessments would not only be used to determine which teachers were retained or dismissed, but also to determine which teachers would receive bonuses, and how substantial those bonuses would be.<sup>49</sup>

On August 24, 2010, U.S. Secretary of Education Arne Duncan announced the winners of Phase 2.<sup>50</sup> Unlike Phase 1, nine states and the District of Columbia were awarded grants.<sup>51</sup> These grants ranged from \$75 million, won by Hawaii, Rhode Island, and the District of Columbia, to \$700 million, won by New York.<sup>52</sup> (Colorado and New Jersey were not among the

---

<sup>42</sup> *Id.*

<sup>43</sup> *Id.*

<sup>44</sup> Kelly Heyboer, *N.J. Education Chief Proposes Sweeping School Reform, Urges NJEA Cooperation*, N.J.com, May 7, 2010, [http://www.nj.com/news/index.ssf/2010/05/schundler\\_announces\\_sweeping\\_e.html](http://www.nj.com/news/index.ssf/2010/05/schundler_announces_sweeping_e.html).

<sup>45</sup> *Id.*

<sup>46</sup> *Id.*

<sup>47</sup> *Id.*

<sup>48</sup> *Id.*

<sup>49</sup> *Id.*

<sup>50</sup> Press Release, U.S. Dep't. of Educ., *Nine States and the District of Columbia Win Second Race to the Top Grants* (Aug. 24, 2010), <http://www.ed.gov/news/press-releases/nine-states-and-district-columbia-win-second-round-race-top-grants>.

<sup>51</sup> *Id.*

<sup>52</sup> *Id.*

winners.) Also, in January 2010, before the winners of Phases 1 and 2 were decided, President Obama announced his plans to continue the RTTT challenge, requesting \$1.35 billion for the program in his fiscal year 2011 budget.<sup>53</sup>

### **III. Problems with This Method**

“Educators know all too well that one-dimensional indicators such as test scores [cannot] begin to capture the complexities of effective teaching and learning.”<sup>54</sup> That said, this paper acknowledges that neither RTTT nor the laws passed by states hoping to win RTTT grants suggest that standardized tests provide the only means to evaluate teachers. Test scores can be useful when they are reliable and are used in conjunction with other measures. Therein reside the flaws of RTTT and the resulting state laws discussed above. For all the discussion of improving testing, little to no attention was paid to improving other forms of teacher evaluation, which continue to be below par. Further, standardized testing remains suspect. Studies have shown VAM results – which allegedly legitimizes test scores as an evaluation tool – to be imprecise. As well, the motivational effects for successful evaluations, such as performance pay, do not correspond to higher student test scores.

Because education is both a cumulative and a complex process, it is impossible to fully distinguish influences such as home life, school conditions, and other teachers from the effect an individual teacher has on his or her students.<sup>55</sup> A study by the Economic Policy Institute found that, while VAM claims to be able to accurately identify effective teachers, VAM results prove to be unstable.<sup>56</sup> In one study that applied VAM to test scores across five large urban districts,

---

<sup>53</sup> U.S. Dep’t. of Educ., Race to the Top Fund, <http://www2.ed.gov/programs/racetothetop/index.html>, (last visited May 5, 2011).

<sup>54</sup> GATES FOUNDATION, *supra* note 2 (quoting Michael Mulgrew, President of the United Federation of Teachers, New York City.)

<sup>55</sup> EVA L. BAKER ET AL., PROBLEMS WITH THE USE OF STUDENT TEST SCORES TO EVALUATE TEACHERS, 9 (Economic Policy Institute) (2010), *available at* [http://epi.3cdn.net/b9667271ee6c154195\\_t9m6ijj8k.pdf](http://epi.3cdn.net/b9667271ee6c154195_t9m6ijj8k.pdf).

<sup>56</sup> *Id.* at 2.

among the teachers who were ranked in the top fifth of effectiveness in the first year, less than one third of those remained in the top group the next year, and another third of that top 20 percent dropped all the way down to the bottom 40 percent.<sup>57</sup> The same dramatic fluctuations were found for teachers ranked at the bottom in the first year of analysis.<sup>58</sup> While one might assume that the quality of a teacher will change little over time, allowing for gradual improvements or dips due to experience, VAM results show that a teacher's effectiveness one year is not predictive of his effectiveness the next.<sup>59</sup> This is ironic as VAM intends to show that student achievement for one year creates a baseline upon which to reasonably assess what a teacher's effectiveness should be the next year. After completing similar research to that of the Economic Policy Institute, the Educational Testing Service Policy Information Center concluded that

VAM results should not serve as the sole or principal basis for making consequential decisions about teachers. There are many pitfalls to making causal attributions of teacher effectiveness on the basis of the kinds of data available from typical school districts. We still lack sufficient understanding of how seriously the different technical problems threaten the validity of such interpretations.<sup>60</sup>

On the opposite end of the spectrum, a study by the National Center on Performance Incentives at Vanderbilt University's Peabody College has reported that rewarding teachers with bonus pay does not raise student test scores.<sup>61</sup> The Project on Incentives in Teaching (POINT) was a three-year study conducted in the Metropolitan Nashville School System from 2006-07 through 2008-09.<sup>62</sup> POINT focused on the idea that part of the problem in American education is

---

<sup>57</sup> *Id.*

<sup>58</sup> *Id.*

<sup>59</sup> *See id.*

<sup>60</sup> *Id.* at 2-3.

<sup>61</sup> MATTHEW G. SINGER ET AL., *TEACHER PAY FOR PERFORMANCE: EXPERIMENTAL EVIDENCE FROM THE PROJECT ON INCENTIVES IN TEACHING*, xi (Vanderbilt Peabody College) (2010), *available at* [http://www.performanceincentives.org/data/files/gallery/ContentGallery/POINT\\_REPORT\\_9.21.10.pdf](http://www.performanceincentives.org/data/files/gallery/ContentGallery/POINT_REPORT_9.21.10.pdf).

<sup>62</sup> *Id.*

that teachers do not have incentives to be highly effective.<sup>63</sup> Hypothetically, if this incentive problem was corrected, student achievement would improve.<sup>64</sup> To test this hypothesis, POINT created two groups of teachers: one group of randomly selected teachers who would be eligible for bonuses if their students improved, and a control group which was not eligible for bonuses.<sup>65</sup> The results did not confirm the hypothesis as the students of the bonus-eligible teachers did not outperform the students learning from the control group teachers.<sup>66</sup>

In sum, the results of these studies leads to the conclusion that a focus on standardized tests does not lead to improved student achievement or better teachers. Even after adjusting to eliminate outside factors, test scores are an unreliable indicator of teacher performance. Also, even when teachers know that they will not only be evaluated, but rewarded, by improved test scores, this incentive does not improve those test scores. Thus, limiting education reform to improving test scores is flawed.

#### **IV. An Effective and Feasible Solution**

Fortunately, efforts are being made across America to remedy this situation. One of these is the MET project being developed by the Gates Foundation. The goal of MET is to improve how states and districts collect and evaluate information about teacher effectiveness.<sup>67</sup> The result is that a teacher's assessment is not dominated by any one aspect of the evaluation – while there is no one perfect tool to assess a teacher (or a student), implementing several varied and authentic methods of evaluation will negate the flaws of each method as the whole melds together to provide a more accurate portrait of a teacher's performance. What is more, the

---

<sup>63</sup> *Id.*

<sup>64</sup> *Id.*

<sup>65</sup> *Id.*

<sup>66</sup> *Id.*

<sup>67</sup> GATES FOUNDATION, *supra* note 2, at 1.

methods implemented by MET are feasible in an educational and economic climate where money and time are at a premium.

First, MET takes into consideration students' performance.<sup>68</sup> Rather than rejecting standardized testing, MET acknowledges its usefulness and supports it with additional assessments. To bolster the information that can be learned from standardized tests, MET incorporates supplemental assessments.<sup>69</sup> While a critique of focusing on standardized testing is that teachers are tempted or forced to teach to the test, the supplemental assessments used by MET assess higher-order conceptual understanding that might be unchecked by state tests but can show how the teacher is helping students use what they have learned in more complex ways.

Second, MET addresses the problems in teacher observation with video-based classroom observations accompanied by teacher reflections.<sup>70</sup> These observations would take place four times a year and would be scored by trained raters using established protocols.<sup>71</sup> As a result, several problems surrounding teacher observations are resolved: observations are expanded from what is often a once-a-year incident, the scoring is done by individuals who are trained and not otherwise occupied as administrators are, and the teachers are provided the opportunity to not only review the videos but also to reflect upon them. This may be the most important aspect of video observations. Unlike traditional observations in which the teacher either receives a written review and/or has a follow-up meeting, the reflection process forces the teacher to take an active role in her evaluation. This increases the likelihood of the teacher evaluating her own performance and taking steps to improve upon it.

---

<sup>68</sup> *Id.* at 6.

<sup>69</sup> *Id.*

<sup>70</sup> *Id.*

<sup>71</sup> *Id.* at 7.

This measure of MET evaluation is the least traditional and, at first glance, the least feasible. To lessen the need for manpower, MET incorporates panoramic digital video cameras that require minimal training to set up, are operated remotely by the teacher, and do not require a camera person.<sup>72</sup> After class, the teacher can upload the video lesson to a secure internet site.<sup>73</sup> To address the sheer amount of footage, MET contends that it is not necessary for raters to watch an entire lesson. Because MET is designed to assess wide range of teacher goals, and it is rare for a teacher to meet every type of instructional goal in a single lesson, raters only watch fifteen minutes of each lesson to both get a taste for what occurred in each class and also to provide an opportunity for the teacher to meet all the goals over several observations.<sup>74</sup> In all, MET ensures that this method of evaluation is not only effective, but feasible.

Third, MET assesses teachers on their general, specialized, and pedagogical content knowledge.<sup>75</sup> This goes beyond knowing the material. This measure ensures that teachers not only understand the content, but can identify errors in student reasoning and use this knowledge to develop a strategy to correct the errors and strengthen student understanding.<sup>76</sup>

Fourth, and perhaps even more intimidating than the prospect of being filmed, MET includes the students to report their experiences of the classroom instructional environment.<sup>77</sup> To ensure that the students provide useful feedback, the survey asks detailed questions to which the students respond if they agree or disagree.<sup>78</sup> The responses are not only sent to the evaluators but also to the teachers so they can improve their instructional methods.

---

<sup>72</sup> *Id.*

<sup>73</sup> *Id.*

<sup>74</sup> GATES FOUNDATION, THE PLATO PROTOCOL FOR CLASSROOM OBSERVATIONS, 3 (2010), available at [http://metproject.org/resources/PLATO\\_10\\_29\\_10.pdf](http://metproject.org/resources/PLATO_10_29_10.pdf).

<sup>75</sup> GATES FOUNDATION, *supra* note 2, at 7.

<sup>76</sup> *Id.*

<sup>77</sup> *Id.*

<sup>78</sup> *Id.*

Finally, MET requires that teachers also complete a survey about the working conditions, school environment, and instructional support in their schools.<sup>79</sup> This may be an antithesis to the VAM philosophy, as this measure incorporates numerous outside effects in hopes of filling in the blanks of a teacher's achievement. In addition, the results could be used to evaluate the effectiveness of principals in supporting effective instruction.<sup>80</sup>

MET is neither the only project of its kind, nor does it claim to be a perfect method of evaluation. However, by using varied measures of teacher effectiveness, MET can capture the complexities of effective teaching and learning that traditional teacher observations and standardized testing cannot accomplish. In addition, MET's evaluation tools are feasible, so that schools have a better chance of not having to rely on the bare minimums to evaluate teachers.

### **Conclusion**

Like our children, America's teachers need help. They need to be recognized for what they do well, supported to improve, and informed when they fail. Most of all, they need to be respected. Currently, America's methods for evaluation disrespect teachers. Classroom observations are largely vague and infrequent, and standardized tests alone cannot distill all a teacher does into a formula. While federal programs like NCLB and RTTT intend to foster legitimate education reform, they, in combination with the recent economic decline, force state governments to simplify their education policies by focusing on test scores. Fortunately, the Gates Foundation's MET program provides an alternative that is both effective and feasible for revolutionizing teacher evaluation. Teachers represent America's most essential profession, and if they are not viewed with the same earnest as America views its children, our children and nation will suffer.

---

<sup>79</sup> *Id.*

<sup>80</sup> *Id.*